Pro Gradu –tutkielma

Detecting effects of developmental mode variation on population genetic stability

Marina Mustonen

University of Jyväskylä

Department of Biological and Environmental Science

Ecology and Evolutionary Biology

14.11.2011

JYVÄSKYLÄN YLIOPISTO, Matemaattis-luonnontieteellinen tiedekunta

Bio- ja ympäristötieteiden laitos
Ekologia ja evoluutiobiologia

MUSTONEN, M. : Kehitysmuodon variaation vaikutus populaation geneettiseen rakenteeseen ja vakauteen.

Pro Gradu –tutkielma: 34s.
Työn ohjaajat: FT K. Emily Knott, FM Jenni Kesäniemi
Tarkastajat:
Marraskuu 2011

Hakusanat: alleelifrekvenssit, kehitysmuoto, levittäytymiskyky, *P. elegans*, temporaalinen vaihtelu

TIIVISTELMÄ

Kehitysmuodolla voidaan kuvata lajin elinkiertopiirteitä, esimerkiksi sitä onko meressä elävän selkärangattoman lajin toukalla pelagista vaihetta kehityksensä aikana. Tällä voi olla suuri vaikutus aikuisena huonosti liikkuvien tai paikallaan pysyvien lajien levittäytymispotentiaaliin, mikä vuorostaan vaikuttaa populaatioiden väliseen geenivirtaan ja geneettiseen rakenteeseen. Toukkakehitysmuotojen variaation taustalla ajatellaan olevan energeettinen vaihtokauppa jälkeläisten koon ja määrän välillä. Epävakaassa ympäristössä saattaa olla hyödyllisempää tuottaa paljon levittäytymiskykyisiä jälkeläisiä sen sijaan, että panostaisi paljon energiaa vain muutaman jälkeisen kasvattamiseen. Tutkimuksessani analysoin *Pygospio elegans* madon sekä ajan myötä tapahtuvaa, että alueiden välistä geneettistä vaihtelua mikrosatelliittimarkkerien avulla. *P. elegans* on kehitysmuodoltaan polymorfinen (poecilogonous) meressä elävä monisukasmato. Hypoteesini oli, että planktonisia eli pelagisen vaiheen omaavia toukkia tuottavissa populaatioissa on enemmän temporaalista vaihtelua kuin populaatioissa, joissa toukilta puuttuu pelaginen kehitysvaihe, ja tämä voi johtua planktonisten populaatioiden isommasta populaatiokoosta ja ns. ’sweepstakes’ lisääntymisestä (sattumanvaraiset erot yksilöiden lisääntymismenestyksessä paljon jälkeläisiä tuottavissa populaatioissa). Tulokset eivät tukeneet tätä hypoteesia: neljästä tutkimuspopulaatiostani temporaalista vaihtelua havaittiin suoraan kehittyviä toukkia tuottavassa Suomen populaatiosta (F_{ST}: 0.00757, p=0.04883) sekä tanskalaisesta populaatiosta jossa on havaittu useampia toukkamuotoja (F_{ST}: 0.01104, p=0.01855). Planktonisia toukkia tuottavasta Skotlannin populaatiosta ja useampia toukkamuotoja tuottavasta Alankomaiden populaatiosta ei havaittu temporaalista populaatiorakenteen muutosta. Nämä tulokset viittaavat siihen, että *P. elegans* –madolla ei ole ’sweepstakes’ lisääntymistä, vaan geenien ajautuminen on suurempi temporaaliseen populaatiorakenteeseen vaikuttava tekijä. Kiinnostavaa oli myös se, että samanlaisessa ympäristössä elävät populaatiot eivät olleet populaatioparametreiltaan samankaltaisia. Tutkimukseni oli yhdenmukainen aiempien tutkimusten kanssa siinä, että planktonisia toukkia tuottavilla populaatioilla oli enemmän geneettistä vaihtelua kuin populaatiossa, jossa toukilta puuttuu pelaginen vaihe (alleelien määrä Skotlannissa 13,197 ja Suomessa 10,933, geenidiversiteetti Skotlannissa 0,753 ja Suomessa 0,628).

UNIVERSITY OF JYVÄSKYLÄ, Faculty of Mathematics and Science

Department of Biological and Environmental Science

Ecology and Evolutionary Biology

MUSTONEN, M.: Detecting effects of developmental mode variation on population genetic stability

Master of Science Thesis: 34p.
Supervisors: PhD K. Emily Knott, MSc Jenni Kesäniemi
Inspectors:
November 2011

Key Words: allelic frequencies, developmental mode, dispersal, *P. elegans*, poecilogony, population structure, temporal variation, trade-off.

ABSTRACT

Developmental mode describes larval ecology, e.g. if larvae have a pelagic phase, and it is a key life-history characteristic in many marine invertebrates. Developmental mode has significant impact on the dispersal potential of sessile or sedentary invertebrates, which in turn affects gene flow, genetic variation and population structure of species. Variation in developmental mode is explained by energetic trade-offs between size and number of offspring, and in an unstable environment, energy allocated towards producing many larvae that have greater dispersal potential could be favored over producing and feeding few larger larvae. In my study, I use microsatellites to analyse both temporal and spatial variation in allele frequencies in the poecilogonous polychaete worm *Pygospio elegans.* My first hypothesis was that there is more temporal genetic structure in planktonic populations due to larger population size and sweepstakes reproduction (differences in reproductive success of an individual due to chance). This hypothesis was not supported by my results: out of my four study populations, temporal genetic structure was actually found in a population with brooded development from Finland (F_{ST}: 0.00757, p=0.04883) and in a population that has mixture of developmental modes from Denmark (F_{ST}: 0.01104, p=0.01855) and not in populations from Scotland and Netherlands, with planktonic and a mixture of developmental modes, respectively. These results suggest that there is no sweepstakes reproduction in *P. elegans* and that genetic drift might be a more powerful cause of temporal genetic structure. My second objective was to find if there was a connection between environmental stability and developmental mode. This was also not supported by my results. Population genetic parameters were not more similar in populations with similar habitats but differing in developmental mode, as I had expected, suggesting that developmental mode affects those parameters to a greater degree than the environment. My study did support the expectation that there is more genetic variation in planktonic populations (allelic richness in Scotland 13,197 vs. 10,933 in Finland, gene diversity in Scotland 0,753 vs. 0,628 in Finland). In this regard, my study is in agreement with previous studies on developmental mode.

Contents

1. INTRODUCTION **5**
1.1. Population genetic consequences of different developmental modes 5
1.2. Temporal genetic structure in populations 6
1.3. Poecilogony 8
1.4. Aims of the study 9
2. Materials and methods **9**
2.1. Study species *Pygospio elegans* 9
2.1.1. Reproduction in *P. elegans* 10
2.2. Population sites and sampling times 11
2.3. Laboratory methods 12
2.4. Data analysis 13
3. Results **14**
3.1. Genetic diversity 14
3.2. Population structure 19
3.3. Estimates of effective population size 21
4. Discussion **21**
4.1. Genetic diversity 22
4.2. Population structure 25
4.3. Estimates of effective population size 27
4.4. Developmental mode and stability of the environment 27
4.5. Possible problems in my study 28
4.6. Conclusions 29
ACKNOWLEDGEMENT **29**
REFERENCES **30**
APPENDICES **33**

1. INTRODUCTION

Developmental mode is a key life-history trait of marine invertebrates. Developmental mode describes the ecology of the larval stage, whether the larvae feeds or doesn't feed, whether it has a dispersive pelagic phase (a life stage that lives in water column) or not, and how long the pelagic phase lasts (McEdward 1995, Krug 2007). Because many marine invertebrates are sessile or sedentary as adults, their dispersal potential is dependent on the mobility of their larvae, which is determined by the developmental mode (McEdward 1995, Hoskin 1997, Kyle & Boulding 2000, Lee & Boulding 2009). Different modes of larval development are common in marine invertebrates (e.g. Blake & Arnofsky 1999) and even though developmental modes are usually characterized by extremes, (for example, pelagic larvae vs. benthic larvae), there also may be many intermediate modes.

Energetic trade-offs between size and number of offspring, are thought to explain variation in life-history traits, including developmental mode (McEdward 1995, Hart & Marko 2010, Kamel et.al. 2010). When a mother has limited resources, dividing them among offspring is very important for her reproductive success (Kamel et.al. 2010). When mothers produce many small eggs, they are typically poor in yolk. This trade-off leads to higher fecundity, but also higher larval mortality, since the larvae are smaller, hatch earlier and get less nutrition from the mother. However, when mothers produce fewer eggs they can be provided a lot of yolk, and although fewer offspring are produced, they might have a better chance of surviving (Krug 2009, Kamel et. al. 2010). Because of these trade-offs, developmental mode is expected to be a target of natural selection, but selection could favour either provisioning strategy.

Few studies have been done on how the stability of the environment affects developmental mode, but because developmental mode is tightly linked with larval dispersal and maternal provisioning, a relationship between environmental stability and developmental mode might also exist. For instance, in a stable environment less dispersal is favoured, whereas in a temporally variable environment it would be better for an organism to have greater dispersal potential (Cohen 1967, Gadgil 1971, Johannesson 1988). In a stable environment there is no advantage in searching for a better habitat since the organism is able to survive well in the current one, but if environmental conditions deteriorate, individuals could migrate or, in the case of sessile animals, their offspring could disperse to other, possibly better environments. Shull (1997) also suggests that in an environment where there is a lot of disturbance, it would be advantageous for an organism to be able to recolonize disturbed areas quickly. On the other hand, environmental stability may affect maternal provisioning and also indirectly affect dispersal. Krug (2009) suggests that in an unpredictable environment it might be better for a mother to allocate more energy into producing just a few larger eggs with a lot of nutrition for the young, as is done with benthic larvae, to hedge against juvenile mortality. In this study I survey temporal genetic stability in populations from different habitat types and discuss if environmental stability can affect population genetic parameters.

1.1. Population genetic consequences of different developmental modes

Developmental mode of the larvae, for example, whether they have a pelagic phase (planktonic larvae) or not (brooded or direct developing larvae), has a significant impact on the dispersal potential of the species. This, in turn, has an impact on migration and the population structure of the species, which further affects population size, genetic drift, within population genetic variation, and the species' geographic range (McEdward 1995, Hellberg 1996, Bohonak 1999). Brooded or directly developing larvae usually lack a pelagic phase, and therefore species with planktonic larvae are expected to have higher

dispersal potential than brooded developers (McEdward 1995, Hellberg 1996, Bohonak 1999). This difference in dispersal potential leads to more profound population structure in species with brooded development and less in species with planktonic development. When there is little gene flow between populations, the populations are separated from each other, and differentiation can happen more easily for brooded developers, whereas high gene flow in planktonic developers connects them into larger populations and prevents differentiation (McEdward 1995, Hellberg 1996, Hoskin 1997, Bohonak 1999). However, within populations, planktonic developers are expected to have more variation, due to the persistence of rare alleles in the large populations (genetic drift has less impact) and the input of different alleles into the population from other populations via migration (McEdward 1995, Hellberg 1996, Hoskin 1997, Bohonak 1999).

These population genetic consequences of developmental mode have been observed in many species. For example, Breton and colleagues (2003) compared the population genetic structure of planktonic *Neanthes virens* and brooding *Hediste diversicolor* and found that there was significant population structure in *H. diversicolor,* but no structure among *N. virens* populations. However, genetic diversity was very low in *N. virens*, which was surprising, since usually planktonic species have high within population genetic variation. Support for the connection between developmental mode and population structure comes also from a study by Hoskin (1997), where planktonic *Morula marginalba* and brooding *Cominella lineolata* and *Bedeva hanleyi* were compared. Results from that study showed that the brooding species had lower genotypic diversity within populations and more population structure between populations than did the planktonic species, which had more genetic diversity overall, and little or no structure between populations. Hoskin concluded that developmental mode is a reliable indicator for genetic variation within and between populations. Conflicting results have also been observed (e.g. Jokiel 1990, Vermeij et. al. 1990), but the general expectation of higher dispersal with planktonic development is well-supported (McEdward 1995, Hellberg 1996, Bohonak 1999).

One reason for results conflicting from this expectation is explained by sweepstakes reproductive success (Hedgecock 1994, Robainas et. al. 2005, Lee & Boulding 2007). Sweepstakes reproductive success means that high fecundity and high early mortality can result in very few parents producing the majority of the next generation by chance. Oceanographic conditions can be favorable to some individuals so that their offspring survive while others die, leading to large differences in the reproductive success of individuals in the population. This can affect the effective size of the population so that it is often many times lower than the census population size (Hellberg 1996, Robainas et. al. 2005, Lee & Boulding 2007). Brooded larvae are not exposed to predation in the pelagic environment and are not in danger of ending up in unfavourable environment like planktonic larvae, so brooded larvae usually have better survivorship, which means that there is not as much variation in the reproductive success of the brooded developers (Levin et. al. 1991, McEdward 1995).

1.2. Temporal genetic structure in populations

Studying the genetic structure of species with different kinds of developmental modes can give valuable information about ecological and evolutionary mechanisms affecting populations (e.g. dispersal possibilities and limitations, historical and physical barriers to gene flow), but often studies on population structure have used just one-time sampling and have not taken temporal (over time) genetic structure into account. Population genetic structure can change over time, and so, temporal studies can give a different perspective about the ecological and evolutionary mechanisms affecting genetic patterns of a species (Hedgecock 1994, Robainas et. al. 2005, Virgilio & Abbiati 2006, Lee & Boulding 2007).

One possible explanation for temporal genetic structure is also sweepstakes reproductive success (Hedgecock 1994, Robainas et. al. 2005, Lee & Boulding 2007). Another hypothesis that could explain temporal genetic structure is that natural selection is acting on the larvae during their pelagic phase causing changes in allelic frequencies (Lee & Boulding 2007). Additionally, some kind of natural event, such as a hurricane or El Nino, can reduce the population size and create an unfavorable environment for reproduction which will cause temporal variation in genetic structure of the population (Robainas et. al. 2005, Lee & Boulding 2007). Genetic drift in small populations can also result in temporal genetic structure (Tessier & Bernatchez 1999).

Temporal genetic studies on marine invertebrates can provide useful information about population parameters. Lee and Boulding (2007) found interesting results, for example when they studied the snail *Littorina keenae*, a planktonic developer. In their analysis, they used four different localities that were sampled twice. They found significant temporal genetic structure but no significant spatial variation and thought that the best explanation for it was a modification of the sweepstakes-theory. Rare alleles could spread widely if an aggregation of sisters would produce a lot of larvae and those larvae would encounter favourable oceanic currents that would spread them (larvae from other parents not being so lucky). Lee and Boulding (2007) concluded that more information was required to determine reliably the cause of the temporal genetic structure without spatial variation. In a different study, Wai-Chuen and colleagues (2010) studied the limpet *Cellana grata*, a planktonic developer. They sampled four locations along the shores of Hong Kong and found only weak temporal genetic structure and concluded that there was no consistent pattern in potential driving forces affecting genetic structure of limpets and that gene flow among shores counteracted the genetic differentiation.

Temporal studies are important because in addition to getting more information about population genetic structure and possible reasons affecting it, temporal genetic studies can also be used to get more information about potential genetic drift, the stability of populations, and determine if conservation efforts are needed. Robainas and colleagues (2005) studied *Farfantepenaeus notiali*, a shrimp, from four populations over an eight year period. They found significant changes in allelic richness and heterozygosities over time and concluded that that even though *F. notialis* has a large census population size it still might be vulnerable to unstable environmental and habitat conditions. They found that effective population size was very small in relation to the actual population size in this planktonic developer due to sweepstakes reproduction.

Temporal genetic studies have been used to get more information about population genetic structure of brooded developers also. Virgilio and Abbiati (2006) studied the brooding polychaete worm *Hediste diversicolor* with samples collected from four sites two times. They found both temporal genetic structure and variation between sampling locations. They were surprised that there was population structure within the estuary where they collected the samples since there is no apparent barrier to restrain gene flow, but concluded that the dispersal potential for this brooding species was sufficiently low to prevent the homogenising effect of gene-flow. They concluded that temporal genetic structure and spatial genetic fragmentation in *H. diversicolor* is best explained by genetic drift accompanied by periodical mortality and/or sweepstakes reproductive success.

As explained above, there have been studies that compared genetic structure in species with different developmental modes (Hellberg 1996, Hoskin 1997, Kyle & Boulding 2000, Breton et. al. 2003, Krug 2009) and there have been studies that assessed temporal variation in genetic structure (Robainas et. al. 2005, Virgilio & Abbiati 2006, Lee & Boulding 2007, Wai-Chuen et. al. 2010), but combining these two concepts has been rare. One study that does that is the study by Lee and Boulding (2009), in which they

compare spatially and temporally four littorinid gastropods of which two are brooded developers and two are planktonic developers. They collected the brooded developers *Littorina sitkana* and *Littorina subrotundata,* and the planktonic developers, *Littorina scutulata* and *Littorina plena,* from three different sites in two different years. Using DNA sequence data, they found that both of the planktonic species have more within-population variation and temporal variation in genetic structure. There was no significant temporal genetic structure in brooded developing species, as was expected, since brooded developers have lower fecundity and higher early survival, so there was not much differences in the reproductive potential of the individuals (no sweepstakes). Also, since the larvae do not have a pelagic phase they are probably safer from predation than planktonic developers. Even though Lee and Boulding did not find temporal genetic structure in brooded developers in their study, they still proposed a possible situation where there could be temporal genetic structure in brooded developers. Since brooded developers usually have smaller population sizes and less variation, they could be more vulnerable to extinction. Extinction and re-colonization of a population might cause temporal genetic structure.

1.3. Poecilogony

There are some species that have more than one developmental mode. The term poecilogony was used to describe this phenomenon by Giard (1905). Although Giard (1905) gives many examples of poecilogonous species, most of them have since been shown to be closely related cryptic species each with differing developmental modes (Hoagland & Robertson 1988). Although poecilogony is rare, there are some poecilogonous species, most of which are spionid polychaetes or opisthobranch molluscs (Rasmussen 1973, Blake & Arnofsky 1999, Morgan et.al. 1999, Bolam 2004, Krug 2007, 2009). There are many definitions of what kind of variation in development can be called poecilogonous, but Krug (2007) describes poecilogonous species as "...species where developmental mode differs between offspring due to variation in egg size or pre-hatching consumption of nurse eggs or extra-zygotic yolk". The developmental pathways that follow are feeding larvae that ingest planktonic food after hatching or larvae that metamorphose without feeding (Krug 2007).

Poecilogony could represent a transitional stage of divergence in developmental mode, meaning that when two populations differ in developmental mode, it could be a sign that they have started to differentiate from each other and will develop into different species. On the other hand, poecilogony could be a bet-hedging strategy to maintain populations in variable environments. The marine environment can be quite diverse and changing, so in an unpredictable environment, variation in developmental mode could be an adaptive response to selection (Krug 2009). In one environment one kind of developmental mode might be good, but when the conditions change, it may be an advantage for a species to be able to change to another developmental mode.

Studying developmental mode variation can give valuable information about ecological and evolutionary processes that influence genetic structure and stability of populations, but since many factors may influence population genetic parameters, isolating the effect of developmental mode from other possible influences is difficult. Even in comparisons of closely related species, such as the *Littorina* species discussed above, differences in ages of the species could confound interpretation of population size estimates based on temporal data. Studying poecilogonous species is a great opportunity to study developmental mode variation without the confusing effect of genetic variation due to speciation.

1.4. Aims of the study

My objective was to determine more about population genetic stability in populations that differ in developmental mode and habitat-type and investigate whether there is connection between developmental mode and the genetic stability of the population. Whether or not the stability of the environment affects the developmental mode can also be indirectly assessed. Since developmental mode, genetic stability and environmental stability are inter-related, teasing apart the relative roles of developmental mode and environmental variation on genetic stability is difficult. I evaluate temporal genetic structure in a poecilogonous species. Samples from four locations differing in developmental mode and habitat type were collected 2-3 times during a three year period. Seven microsatellite loci were amplified and sized, and population genetic analyses on a spatial and temporal scale were performed. My hypothesis is that temporal genetic structure in populations with planktonic developmental mode is high due to large population sizes, sweepstakes reproductive success, and the input of new alleles from larval recruits, whereas populations with brooded developmental mode are more stable genetically due to smaller population sizes, no sweepstakes and primarily local recruitment. In addition, I expected that populations polymorphic in development (showing both planktonic and brooded larvae) will also vary in temporal genetic structure depending on their environmental stability. In unstable habitats (open shores) these will show population parameters similar to those of planktonic developers, namely high temporal genetic structure; whereas those in stable habitats (sheltered shores) will show population parameters similar to that for brooded developers. I also aim to test different methods for estimating effective population sizes in the different populations.

2. MATERIALS AND METHODS

2.1. Study species *Pygospio elegans*

Pygospio elegans is a sedentary benthic polychaete worm. It is about 15mm long and builds sand-tubes with mucous interior that are about 1mm in diameter (Rasmussen 1973, Morgan et. al. 1999, Bolam 2004). *P. elegans* has a life span of 1-2 years on average. *P. elegans* has a wide habitat tolerance from shallow waters to down to 100m depths, from brackish waters with salinities as low as 2ppt to very saline marine waters, and from tidal open areas to estuarine environments with no tides (Rasmussen 1973, Anger 1984, Morgan et. al. 1999, Bolam 2004). *P. elegans* prefers sandy and mixed sediments, but there is a lot of variability in that also (Rasmussen 1973, Morgan et. al. 1999, Bolam 2004). Geographically *P. elegans* is wide-spread, it has been found in Arctic, Northern Atlantic, Northern Pacific Oceans, the Baltic, North, Othosk, and Mediterranean Seas and along the coast of South Africa (Anger 1984, Anger et. al. 1986, Morgan et. al. 1999, Bolam & Fernandes 2003, Bolam 2004). *P. elegans* is a cold-adapted pioneer species (meaning that it can re-colonize areas very fast after disturbance and recondition the sediment) and it can be the dominant species in many sandy and mudflat habitats in the Northern hemisphere (Anger 1984, Anger et. al. 1986, Morgan et. al. 1999). *P. elegans* can form dense populations that have a significant impact on the sediment (Anger 1984, Anger et. al. 1986, Morgan et. al. 1999) and can make it more favourable e.g to some bivalve species by stabilizing the sea bed with dense sand-tube aggregations, and the bivalves in turn can make the habitat unfavourable to *P. elegans* through competition, predation and sediment disturbance (Kube & Powilleit 1997, Bolam & Fernandes 2003). This can make some *P. elegans* populations that form sand-tube aggregations relatively short-lived.

2.1.1. Reproduction in *P. elegans*

P. elegans can reproduce throughout the year, but usually there are two reproductive peaks, e.g. in Drum Sands, Scotland the peaks are in May and December (Bolam 2004). There can also be just one reproductive peak in areas where circumstances are favourable to reproduction for only a short time of the year, such as on the coast of Finland. *P. elegans* can reproduce asexually by fragmentation of its body into segments from which new individuals regenerate (this form of asexual reproduction is called architomy) or it can reproduce sexually. After internal fertilization, female *P. elegans* produces egg strings inside their tubes. In every string there are capsules that are connected to each other and in those capsules there are genuine fertilised eggs that are clear and poor in yolk and there can also be unfertilised nurse eggs filled with yolk. The larvae in the capsules can be planktonic, e.g. have short brooding period in the egg capsule, after which they are released to the plankton to feed before settling in the sediment (planktonic larvae) or the larvae can be brooded and stay in the maternal tube and feed on nurse eggs provided by the mother (brooded larvae). After release, brooded larvae build their sand-tubes in the sediment without a pelagic phase. There are also intermediate type larvae, which spend some time developing and feeding in the egg capsule but also have a short pelagic phase. The planktonic, brooded and intermediate larvae differ morphologically but they metamorphose into the same kind of adults (Hannerz 1956, Rasmussen 1973).

There are fewer fertilised eggs in a capsule when the larvae are the brooded type and more when the larvae are planktonic (Rasmussen 1973, Morgan et. al. 1999). Likewise, the number of nurse-eggs provided by the mother to brooded and intermediate larvae also varies. Nurse-eggs are fragile and break easily into yolk granules that the developing larvae can feed upon (Blake & Arnofsky 1999). This feeding strategy is called adelphophagia (Rasmussen 1973, Blake & Arnofsky 1999, Morgan et. al. 1999). Planktonic larvae hatch from the egg capsules earlier and feed on the plankton and are typically not provided nurse-eggs. Egg capsules and larvae of different developmental modes can be seen in Figure 1.

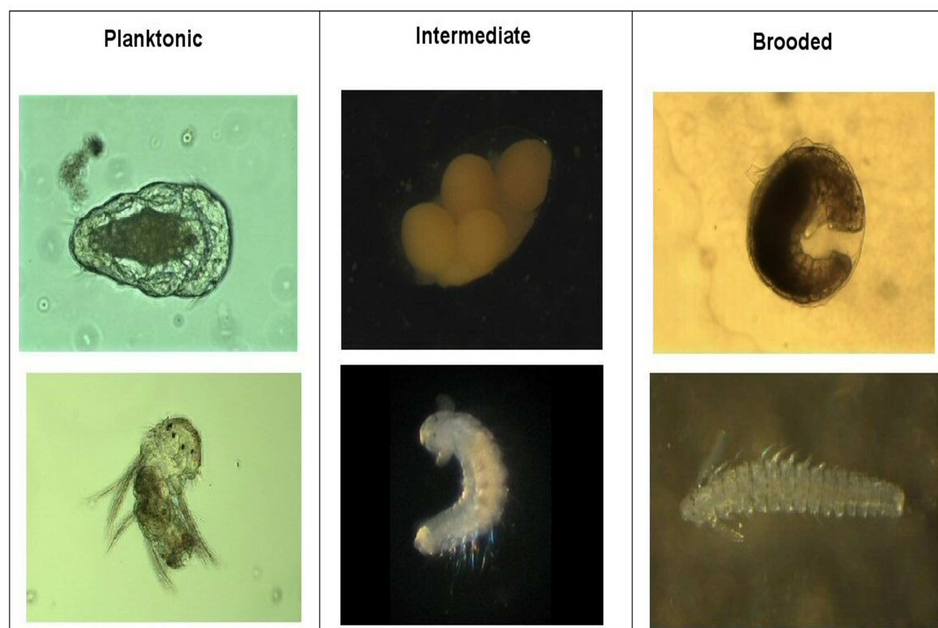

Figure 1. Egg capsules and larvae of different developmental modes, the planktonic type, intermediate type and brooded type. The scale is not same in these pictures, the planktonic larvae are much smaller than the other types (Photo credit:Jenni Kesäniemi).

As mentioned previously, many species that are thought to be poecilogonous have after closer study turned out to be closely related species with different developmental modes. This has also been suspected of *P. elegans.* Anger (1984) made an experiment where she reared *P. elegans* worms in different salinities and temperatures for 14 months and found that *P. elegans* did not change its developmental mode regardless of the environmental fluctuations. She concluded that since temperature did not have an effect on the developmental mode, it might be genetic and that *P. elegans* populations with different developmental modes might be cryptic sister species. Morgan and colleagues (1999) studied morphological data and allozymes of *P. elegans* from four different localities and found that morphological characters did not give evidence for species divergence and that there was no correlation between genetic structuring and reproductive strategy, and they came to the conclusion that *P. elegans* is truly poecilogonous. Kesäniemi and colleagues (personal communication) have studied *P.elegans* from a broader geographic region and also concluded that it is indeed poecilogonous.

2.2. Population sites and sampling times

P. elegans worms were collected from four different localities: Finland (Ängsö), Denmark (Vellerup), Netherlands and Scotland, UK (See Figure 2). Worms were sampled 2-3 times in 2008-2010 (See Table 1) by Emily Knott's research group.

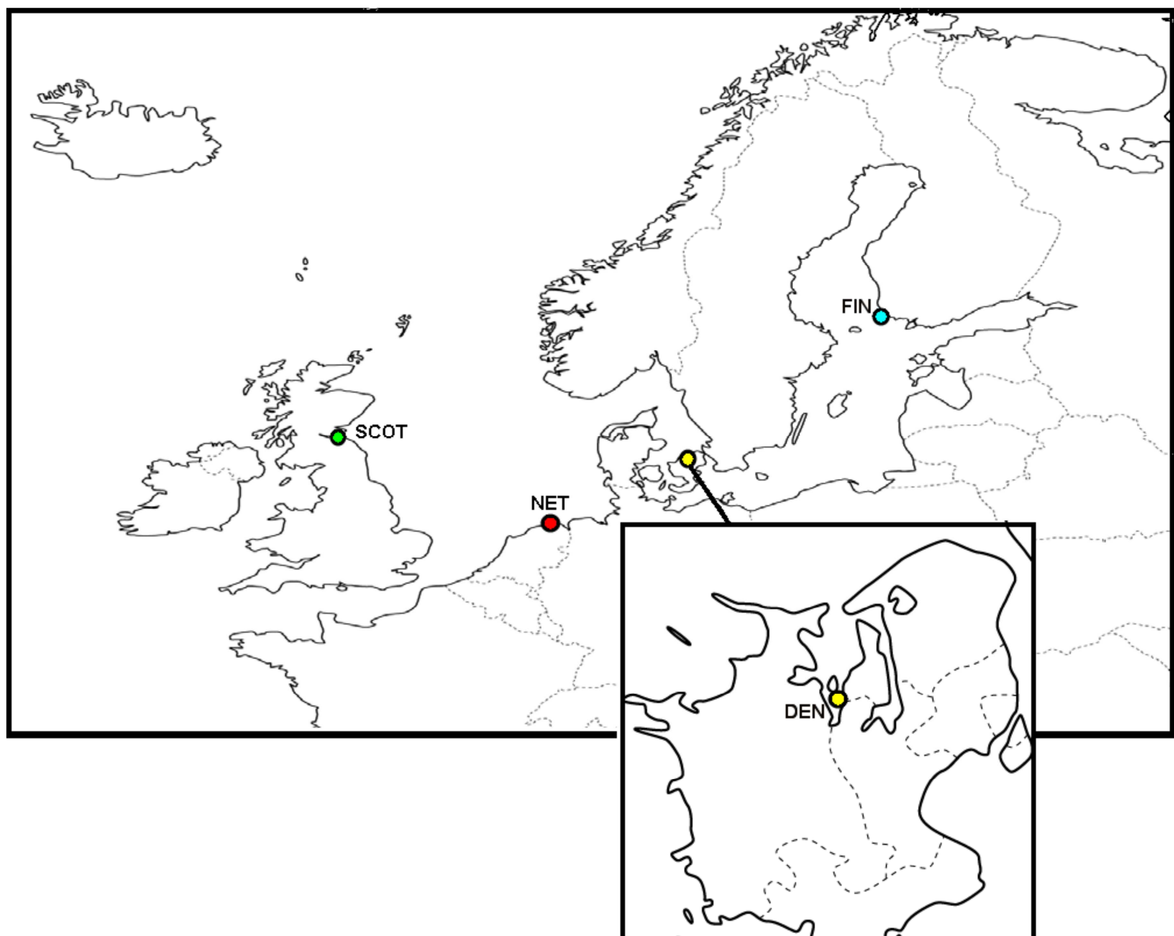

Figure 2. *P. elegans* sampling sites. FIN = Finland, DEN= Denmark, NET = Netherlands and SCOT = Scotland.

Table 1. Sampling sites and times, and the number of individuals sampled from each site in each year.

Population	Sampling time	Number of individuals
Finland	2008	58
	Spring 2009	30
	Fall 2009	48
	2010	42
Denmark	2008	43
	2009	47
	2010	44
Netherlands	2009	45
	2010	40
Scotland	2009	28
	2010	40

Sampling sites differ in the habitat-type they offer to *P. elegans* and in potential environmental variability. In Finland the habitat is sheltered, there are no tides, so the worms are under water all the time. There is a lot of vegetation (mostly *Zostera marina*) and *P. elegans* uses the vegetation as support when building sand tubes. Additionally, the *P. elegans* population in Finland is not dense (approx. 200-6000 individuals/m^2). Here, worms were sampled from about 5 meters depth. In Denmark the habitat is similar: sheltered, with weak tides and some vegetation. Worms were sampled from about 1 meter deep in Denmark. In Finland and Denmark, the sediment is sandy. Habitats in Scotland and Netherlands are different from the Baltic Sea sites, but are very similar to each other. These sites are open shores that experience strong tides and that have little vegetation, if any. At these sites, *P. elegans* live on intertidal mudflats, that are under water only during high tide and where the worms form large aggregations of sand tubes that give each other support (density is high, over 12000 individuals/m^2). In Scotland and Netherlands *P. elegans* samplings were done during low tide when the worm's tubes were exposed.

The study sites differ also in their developmental mode. In Finland the *P. elegans* population only produces benthic larvae. In Denmark and Netherlands there are brooded and planktonic larvae, and also intermediate type larvae. In Scotland only planktonic larvae have been observed.

2.3. Laboratory methods

P. elegans worms were preserved in ethanol after sampling. They were air dried on paper towels and crushed in separate 1,5 ml microcentrifuge tubes. DNA was extracted using Qiagen DNeasy Tissue Kit chemicals and protocols adjusted for use with a KingFisher magnetic particle processor (ThermoScientific).

A PCR reaction was used to amplify each of 7 microsatellite loci. All PCR reactions were 10 µl volume, but two different methods were used: a basic amplification using labelled primers and a M13 Tailing -method coupled with a labelled universal M13(-21) primer (Schuelke 2000). PCR for loci pe7, pe6, pe13 and pe19 were done with the basic reaction. The reaction mix contained 1µl DNA, 1mM buffer, 0,2mM dNTP's, 0,5 µM F-primer and 0,5 µM R-primer, 2-3mM $MgCl_2$ (See Table 2) and 0,5 units of Taq polymerase (Biotools). PCR for loci pe15, pe17 and pe18 were done with the M13 Tailing method. The reaction mix contained 1µl DNA, 1mM buffer, 0,2mM dNTP's, 0,04 µM F-primer, 0,16 µM R-primer, 0.16 µM M13(-21) primer, 1,5-2 mM $MgCl_2$ (See Table 2) and 0,5 units Taq-polymerase (Biotools). All reaction mixes for loci pe18 and pe13 and some reaction mixes for the other loci also contained 0,09 mM BSA. BSA is an additive that can help improve specificity of the amplification reaction (Kreader 1996).

The PCR –program for the basic method consisted of initial denaturation at 94 °C for 5 minutes, followed by 35 cycles of denaturation at 94°C, annealing at 54-57°C (see Table 2) and extension at 72°C, and a final extension at 72°C for 10 min. The PCR –program for the M13 Tailing method consisted of initial denaturation at 94 °C for 5 minutes, followed by 30 cycles of denaturation at 94°C, annealing at 54-57°C (see Table 2) and extension at 72°C, and 8 cycles of denaturation at 94°C, annealing at 53°C and extension at 72°C, and a final extension at 72°C for 10 min. Bio Rad Thermal Cyclers S1000 and C1000 were used. Success of PCR reactions was checked with agarose gel electrophoresis. PCR products were combined with 8 µl of a mix containing 1:129 GeneScan™ 500 LIZ™ Size Standard : formamide and then were separated with an ABI Prism 3130xl Genetic Analyzer. Alleles were sized using GeneMapper version 4.0.

Table 2. Names, repeats, primer sequences, T_a (PCR annealing temperature) and the amount of $MgCl_2$ used in PCR reactions for each locus. The M13(-21) tail sequence is TGTAAAACGACGGCAGT and it anneals at 53°C.

Locus	Repeat	Primer sequences 5'-3'	T_a (°C)	$MgCl_2$ (mM)
Pe6	$(CA)_{28}$	F: ACTACGGAAACTGCCTGCAC R: ATATGGCCACCGAAACCTCT	54	2
Pe7	$(CATA)_{13}$	F: CTCACCCTTTACACCCAAGG R: AGCGTCTGTTATGGGGTACAG	54	3
Pe13	$(GA)_{23}$	F: CCGGCGTCTCTACACAATAC R: CTGTGAACACTGCTGCGAAT	56	2
Pe15	$(GT)_{10}(GA)_{24}$	F: M13(-21)+ TAGTGATCACCCCACATCCA R: AACACACCTTCCCTCACACC	57	1.5
Pe17	$(TG)_4N_3(TG)_8$	F: M13(21)+CAAATGAGTTGTGGACTAGTAGGG R: CCCCCTGTGGGCTAGATAG	57	2
Pe18	$(CAA)_4N(CAA)_2$	F: M13(-21)+ TGGATACGGTCTCAACCTTTG R: AGCCATTGCCCAATGATAAC	57	2
Pe19	$(GC)_2(AA)_2$ $(GCAGCAA)_4$	F: TATCCAACGCACACCTACCA R: TTGAGTGATGGTGCGAGGTA	54	2

2.4. Data analysis

Populations were defined so that each temporal sampling from a sampling site is a population, so there are 11 temporal population samples (Finland 2008, spring 2009, fall 2009 and 2010, Denmark 2008, 2009 and 2010, Netherlands 2009 and 2010, and Scotland 2009 and 2010). Some analyses were done so that samplings from each site were grouped together for four population groups (Finland, Denmark, Netherlands and Scotland).

Microsatellite data were analysed using different population genetic programs. FSTAT version 2.9.3.2 (Goudet 1995) was used to see if there was linkage disequilibrium and to determine the gene diversity (GD), allelic richness (R) and inbreeding (F_{IS}) values for the populations. Arlequin version 3.0 (Excoffier & Schneider 2005) was used to analyse the expected and observed heterozygosities (He and Ho) of the loci. It was also used to test if the populations were in Hardy-Weinberg equilibrium (HWE) using a modified version of the Markov-chain random walk algorithm described by Guo and Thomson (1992). Since the data showed some Hardy-Weinberg disequilibrium, there was reason to suspect null alleles. This was tested using the Expectation Maximization (EM) algorithm (Dempster et.al. 1977) as it is implemented in FreeNA (Chapuis & Estoup 2007). The number of private alleles was determined using GenAlEx version 6.1 (Peakall & Smouse 2006). Differences between locations were analysed with FSTAT version 2.9.3.2 (Goudet 1995) by grouping temporal population samples for each locality together

so that there were four groups based on locality: Finland, Denmark, Netherlands and Scotland. Allelic richness, observed heterozygosity and gene diversity were calculated for each group and values were compared to each other to obtain p-values for variation between groups. A Kruskall-Wallis one-way analysis of variance was used in SPSS version 18 to determine if there was significant variation in gene diversity, allelic richness and expected and observed heterozygosities between the temporal population samples.

Arlequin version 3.0 (Excoffier & Schneider 2005) was also used to do AMOVA –analysis for the data (1000 permutations, otherwise default settings) and pairwise comparison of F_{ST} (distance matrix of the number of different alleles, 1000 permutations, otherwise default settings) to see both the differences between temporal population samples and differences between locations. STRUCTURE 2.3.3 (Pritchard et.al. 2000 & Falush et.al. 2003) was used to analyse population structure. In STRUCTURE, admixture model (individual has inherited some fraction of its genome from ancestor in population *k*) and independent allelic frequencies model were assumed, and default settings were used otherwise. Twelve K's (K=1-12) were tested with burnin=30000 and mcmc=1000000 and runs for each K were replicated 5 times. The consistency of the results was checked by testing K=4, K=5 and K=6 using the same settings otherwise, but with different burnin and mcmc lengths (burnin=300000 and mcmc=1000000, burnin=100000 and mcmc=900000, burnin=50000 and mcmc=500000). STRUCTURE analysis was also done to each of the four populations separately (Finland, Denmark, Netherlands and Scotland) to see if STRUCTURE could detect any substructuring within them. Admixture model was assumed, default settings were used with burnin=300000 and mcmc=1000000.

Effective population size (Ne) was estimated with TempoFs (Jorde & Ryman 2007), as well as using the moment based temporal method of Waples (1989), the linkage disequilibrium method of Hill (1981) and the TM3 -method (Berthier et. al. 2002) as they are implemented in NeEstimator (Overden et. al. 2007). The TempoFs method (Jorde & Ryman 2007) could only be used for Finland and Denmark since a minimum of three samplings was required for this method and I only had two samplings from Netherlands and Scotland. Sample plan 2 (individuals are sampled before they reproduce and are not returned to the population) was used and one generation/year was assumed for Finland and two generations/ year for Denmark in TempoFs (Jorde & Ryman 2007). In NeEstimator (Overden et. al. 2007) one generation/year for Finland and two generations/year for Denmark, Scotland and Netherlands was assumed.

3. RESULTS

3.1. Genetic diversity

Linkage equilibrium was tested permuting individuals within temporal populations in FSTAT version 2.9.3.2 (Goudet 1995). Loci pe7 and pe13 showed some association in Finland 2008. Other loci showed associations in the Danish populations: pe7 with pe15, pe7 with pe13 and pe13 with pe15 in both Denmark 2008 and 2009 populations. In addition, in the Denmark 2009 population there is also association between loci pe7 and 17 as well as pe13 and 17 (See Appendix 1 for details).

Guo & Thomson's Hardy-Weinberg equilibrium test, implemented in Arlequin version 3.0 (Excoffier & Schneider 2005), showed that some loci in the temporal population samples are not in Hardy-Weinberg equilibrium. Especially loci pe15 and pe13 deviated from Hardy-Weinberg equilibrium in many populations (See Table 3).

Table 3. Descriptive statistics for each temporal population and locus. GD=Gene diversity, R=allelic richness , He=expected heterozygosity , Ho=observed heterozygosity , F_{IS}=inbreeding. Significant deviations from Hardy-Weinberg equilibrium are marked with asterisks after He (p< 0.05 = *, p< 0.01 = **, p< 0.001 = ***).

Site	Year	N		Locus						
				Pe7	**Pe15**	**Pe17**	**Pe18**	**Pe6**	**Pe19**	**Pe13**
Finland	**2008**	53	**GD**	0.937	0.943	0.245	0.581	0.275	0.278	0.949
			R	21.033	19.160	4.777	4.537	2.955	4.596	20.100
			He	0.93555***	0.93932***	0.24510	0.58095	0.27409**	0.27835	0.94664***
			Ho	0.78431	0.59615	0.24528	0.54717	0.23077	0.30612	0.71429
			F_{IS}	0.163	0.368	-0.001	0.059	0.159	-0.101	0.248
	2009	30	**GD**	0.943	0.956	0.249	0.561	0.462	0.415	0.951
			R	21.272	17.036	4.990	2.998	3.000	3.749	21.000
			He	0.94124	0.95141***	0.24859	0.55876	0.46158**	0.41623	0.95122
			Ho	0.83333	0.66667	0.20000	0.43333	0.44828	0.50000	0.95238
			F_{IS}	0.116	0.303	0.198	0.227	0.029	-0.206	-0.001
	2009	48	**GD**	0.937	0.953	0.320	0.543	0.259	0.456	0.957
			R	19.871	19.775	5.983	3.430	3.123	5.168	20.415
			He	0.93640	0.94825***	0.31930*	0.54232	0.25943	0.45665	0.95458**
			Ho	0.87500	0.52083	0.29167	0.50000	0.29167	0.48936	0.79070
			F_{IS}	0.066	0.453	0.087	0.079	-0.126	-0.072	0.173
	2010	42	**GD**	0.945	0.955	0.242	0.587	0.354	0.444	0.966
			R	19.812	21.264	4.105	3.752	2.997	4.046	21.173
			He	0.94349	0.95181**	0.24211	0.58606	0.35441	0.44557	0.96245***
			Ho	0.78571	0.66667	0.26190	0.47619	0.41463	0.55000	0.74194
			F_{IS}	0.169	0.302	-0.083	0.189	-0.172	-0.238	0.232
Denmark	**2008**	43	**GD**	0.959	0.948	0.758	0.537	0.384	0.572	0.959
			R	21.290	21.445	9.965	3.000	5.314	5.142	21.006
			He	0.95841**	0.94550**	0.75294***	0.53706	0.38440	0.57212	0.95664***
			Ho	0.88095	0.75610	0.34884	0.57143	0.44186	0.60976	0.75610
			F_{IS}	0.082	0.202	0.540	-0.065	-0.152	-0.067	0.212

Table 3 continues

Site	Year	N		Locus						
				Pe7	Pe15	Pe17	Pe18	Pe6	Pe19	Pe13
Denmark	**2009**	47	**GD**	0.953	0.932	0.732	0.692	0.369	0.424	0.955
			R	19.364	19.693	10.008	4.718	3.452	4.546	19.779
			He	0.95181***	0.92833***	0.73005*	0.69056	0.37004	0.42424	0.95324*
			Ho	0.86667	0.60870	0.54348	0.55814	0.43478	0.47727	0.81818
			F_{IS}	0.090	0.347	0.258	0.194	-0.177	-0.127	0.143
	2010	44	**GD**	0.929	0.956	0.592	0.688	0.268	0.431	0.941
			R	16.279	21.077	7.758	4.591	3.229	2.500	17.670
			He	0.92857	0.95533	0.58950**	0.68835***	0.26840	0.43001	0.94073
			Ho	0.85714	0.86364	0.40476	0.72727	0.30233	0.35714	0.94872
			F_{IS}	0.078	0.097	0.316	-0.057	-0.128	0.171	-0.009
Netherlands	**2009**	45	**GD**	0.899	0.913	0.572	0.757	0.243	0.558	0.967
			R	17.755	15.683	7.197	6.719	2.729	5.545	22.122
			He	0.89557***	0.91026***	0.56881***	0.75473***	0.24347	0.55747	0.96421***
			Ho	0.62500	0.67442	0.32558	0.59524	0.27273	0.54545	0.78947
			F_{IS}	0.305	0.261	0.431	0.213	-0.122	0.022	0.183
	2010	40	**GD**	0.944	0.913	0.537	0.492	0.098	0.583	0.955
			R	21.169	17.881	5.428	2.999	2.828	5.841	19.714
			He	0.94306*	0.90918***	0.53608	0.49034	0.09778	0.58133*	0.95372**
			Ho	0.84615	0.57500	0.50000	0.41176	0.10000	0.42500	0.86486
			F_{IS}	0.104	0.371	0.068	0.162	-0.023	0.271	0.094
Scotland	**2009**	28	**GD**	0.961	0.969	0.874	0.525	0.374	0.623	0.968
			R	21.047	21.713	11.228	6.862	3.984	5.794	19.388
			He	0.96003	0.96494***	0.86723***	0.51586***	0.37338	0.62041	0.96011***
			Ho	0.92308	0.75000	0.51852	0.13636	0.35714	0.48000	0.58333
			F_{IS}	0.039	0.226	0.407	0.740	0.044	0.230	0.398
	2010	40	**GD**	0.941	0.965	0.848	0.579	0.402	0.560	0.965
			R	19.174	21.524	13.538	6.781	6.018	5.552	22.154
			He	0.94073	0.96076***	0.84351***	0.57433***	0.40260	0.55981	0.96187***
			Ho	0.89744	0.62500	0.55556	0.27778	0.43590	0.55000	0.70270
			F_{IS}	0.047	0.352	0.345	0.520	-0.084	0.018	0.272

Analysis with FreeNA (Chapuis & Estoup 2007) confirmed the presence of null alleles, which were suspected due to the deviation from Hardy-Weinberg equilibrium. Null alleles were shown to be present across loci and populations and pe15 and pe13 showed null alleles in all populations (See Table 4 for more details). ENA correction (Chapuis & Estoup 2007) was done to the null alleles and F_{ST} was estimated for the corrected data and the original data. F_{ST} –values for corrected and original data did not differ much, so the original data was used in other analyses (See Appendix 2).

Table 4. Estimate of null allele frequency for the temporal populations, per locus. Moderate null allele frequency ($0.05 < r < 0.20$) is marked with * and high null allele frequency ($r > 0.20$) is marked with **

	Locus						
Population	**Pe7**	**Pe15**	**Pe17**	**Pe18**	**Pe6**	**Pe19**	**Pe13**
Fin08	0.11468*	0.19287*	0.00001	0.00054	0.12675*	0.19344*	0.31268**
FinS09	0.03367	0.14375*	0.06856*	0.08713*	0.12760*	0.13774*	0.31581**
FinF09	0.02611	0.21457**	0.04280	0.00991	0.00000	0.06552*	0.18753*
Fin10	0.07742*	0.23126**	0.00001	0.12953*	0.06978*	0.09773*	0.35326**
Den08	0.06475*	0.14492*	0.22505**	0.31491**	0.00000	0.11595*	0.14730*
Den09	0.09242*	0.24010**	0.12293*	0.27582**	0.06115*	0.15080*	0.13370*
Den10	0.08834*	0.07922*	0.19906*	0.07563*	0.08370*	0.17263*	0.13124*
Net09	0.25248**	0.17005*	0.23159**	0.30391**	0.08729*	0.09086*	0.23831**
Net10	0.07827*	0.30023**	0.03181	0.39078**	0.00002	0.10761*	0.12567*
Scot09	0.09927*	0.10312*	0.21596**	0.51133**	0.03542	0.25182**	0.31204**
Scot10	0.05156*	0.19550*	0.25518**	0.61569**	0.07139*	0.02509	0.19863*

Calculations of gene diversity, allelic richness and expected and observed heterozygosities showed a general trend of greater gene diversity, allelic richness and expected heterozygosity in Scotland than in other populations and the least genetic variation in Finland. Observed heterozygosity is similar in all populations (See Table 3).

The number of alleles and private alleles was counted with GenAlEx version 6.1 (Peakall & Smouse 2006) and the results showed clearly that there are more private alleles in Scotland than in the other populations (See Figure 3).

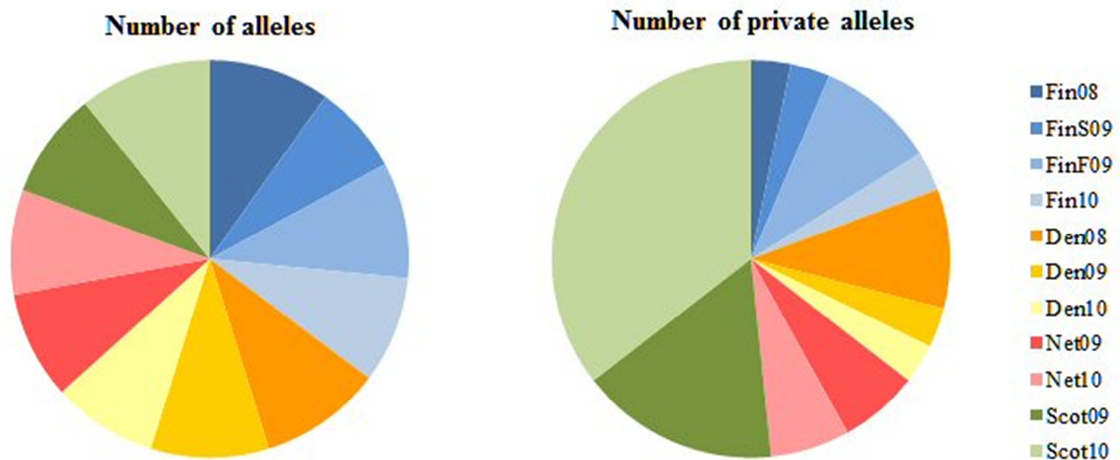

Figure 3. Pie charts of number of alleles and number of private alleles in all populations

Gene diversity, allelic richness and expected and observed heterozygosity were not significantly different among temporal population samples within any of the four locations (analyzing each sampling site separately with Kruskall-Wallis one way analysis of variance) (see Table 5). Nevertheless, there was a general trend of more variation between years in Denmark, Netherlands and Scotland populations than in Finland. This general trend with most variation in Scotland overall can be seen in Figure 4 plotting gene diversity and allelic richness (see medians).

Table 5. Kruskall-Wallis one-way analysis of variance for gene diversity (GD), allelic richness (R) and expected and observed heterozygosity (He and Ho). Each sampling site was analysed separately. No significant variation among temporal populations was detected.

		χ^2	df	p-value
Finland	GD	0,624	3	0,891
	R	0,070	3	0,995
	He	0,728	3	0,867
	Ho	0,386	3	0,943
Denmark	GD	0,423	2	0,809
	R	0,987	2	0,610
	He	0,386	2	0,825
	Ho	0,030	2	0,985
Netherlands	GD	0,065	1	0,798
	R	0,037	1	0,848
	He	0,102	1	0,749
	Ho	0,037	1	0,848
Scotland	GD	0,102	1	0,749
	R	0,037	1	0,848
	He	0,004	1	0,949
	Ho	0,102	1	0,749

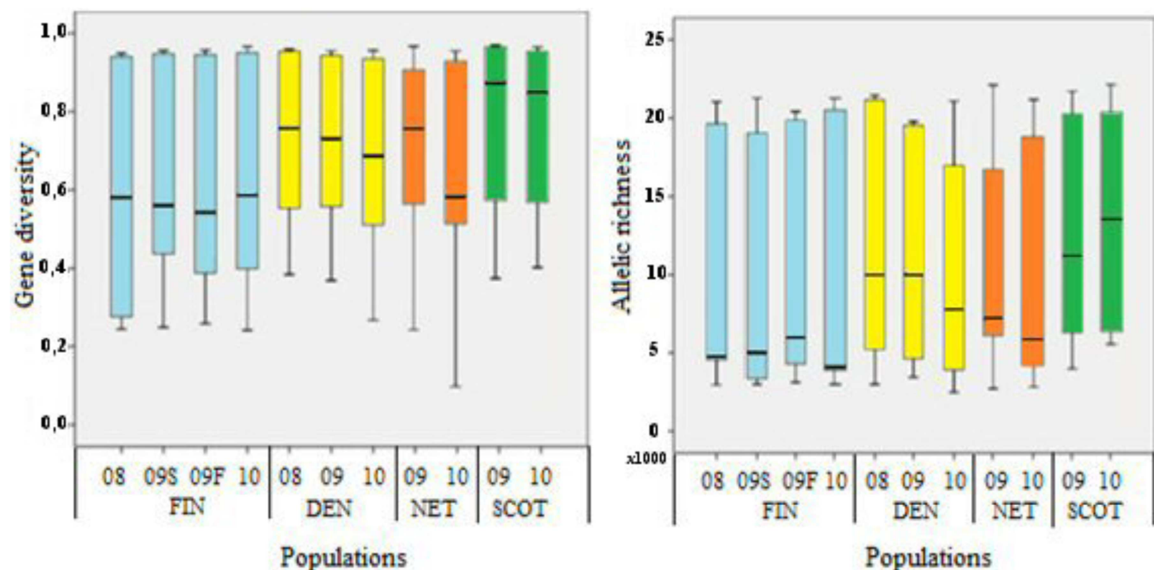

Figure 4. Boxplot figures of gene diversity and allele richness in all the populations. The lower line of the box represents the 25 percentile, middle line 50 percentile (median) and upper line 75 percentile. Ends of the whiskers represent the lowest and highest values.

FSTAT version 2.9.3.2 was used to calculate allelic richness, gene diversity and observed heterozygosity for grouped populations corresponding to our sampling sites. These showed that Scotland has the highest allelic richness and gene diversity, while Finland has the least. Diversity values similar to the Finnish population were also seen in Netherlands (see Table 6). Allelic richness and gene diversity are significantly different between groups (Allelic richness p= 0.01550, gene diversity p= 0.01100). Observed heterozygosity does not vary significantly between groups (p= 0.08150).

Table 6. Allelic richness (R), observed heterozygosity (Ho) and within population gene diversity (Hs) when temporal population samples are grouped based on sampling site.

Site	R	Ho	Hs
Finland	10.933	0.534	0.628
Denmark	11.515	0.626	0.714
Netherlands	10.972	0.541	0.676
Scotland	13.197	0.561	0.753

3.2. Population structure

AMOVA analysis showed that most of the variation is among individuals within populations (93,76 %). There was some variation between groups or sampling locations (5,79 %), meaning that Finland, Denmark, Scotland and Netherlands differed from each other. However, there was not much variation among the temporal population samples within a group (0,46 %), meaning there is not much variation between sampling times within each locality. Variation in all levels is significant (Table 7 for details).

Table 7. Results of AMOVA -analysis.

Source of variation	d.f.	Sum of squares	Variance components	% of variation	p-value
Among locations	3	79.969	0.10977 Va	5.79	0,00010 (+-0,00010)
Among temporal samples within locations	7	17.520	0.00867 Vb	0.46	0,0205 (+-0,00139)
Within locations	909	1616.589	1.77843 Vc	93.76	0,00000 (+-0,00000)
Total	919	1714.078	1.89686		

Pairwise F_{ST} comparisons showed that there is significant spatial genetic structure among the locations (all p-values for comparisons between different sampling sites were highly significant, p< 0,001, see Table 7 for details). Between temporal population samples within a locality there is a significant difference seen in Denmark between years 2009 and 2010 (F_{ST}= 0.01104, p=0.01855) and in Finland between years 2008 and 2010 populations (F_{ST}= 0.00757, p=0.04883) (See Table 8).

Population genetic structure analysis done with STRUCTURE 2.3.3 estimated that there are five genetic clusters in the data (K=5, Estimated Ln Prob of Data = -12369.5). These clusters likely correspond to the four sampling sites plus an additional group (See Figure 5). Consistency of the results was checked with different burnin and mcmc lengths, and they were consistent (results are not shown). Also, STRUCTURE was used to check if it could detect any substructuring in any of the four localities when analysed separately, but K=1 got most support in each population, so substructuring was not detected with STRUCTURE.

Table 8. Pairwise comparisons of F_{ST}. Significant Fst-values are marked with asterisks ($p < 0.05$ = *, $p < 0.01$ = **, $p < 0.001$ = ***). Within population comparisons are bold.

Population	Fin08	FinS09	FinF09	Fin10	Den08	Den09	Den10	Net09	Net10	Scot09	Scot10
Fin08	0.00000										
FinS09	0.00273	0.00000									
FinF09	0.00429	0.00441	0.00000								
Fin10	0.00757*	-0.00036	0.00141	0.00000							
Den08	0.06467***	0.05268***	0.04758***	0.04284***	0.00000						
Den09	0.06854***	0.06145***	0.05998***	0.05396***	0.00393	0.00000					
Den10	0.04341***	0.03820***	0.03448***	0.02898***	0.00460	0.01104*	0.00000				
Net09	0.15274***	0.13915***	0.13582***	0.13438***	0.06983***	0.07015***	0.08606***	0.00000			
Net10	0.12007***	0.11082***	0.10843***	0.10888***	0.06007***	0.06302***	0.07373***	0.00902	0.00000		
Scot09	0.04722***	0.03938***	0.03749***	0.03606***	0.02457***	0.02575***	0.01962***	0.09233***	0.11425***	0.00000	
Scot10	0.03765***	0.04099***	0.02628***	0.02789***	0.02465***	0.03454***	0.02106***	0.07703***	0.09093***	-0.00008	0.00000

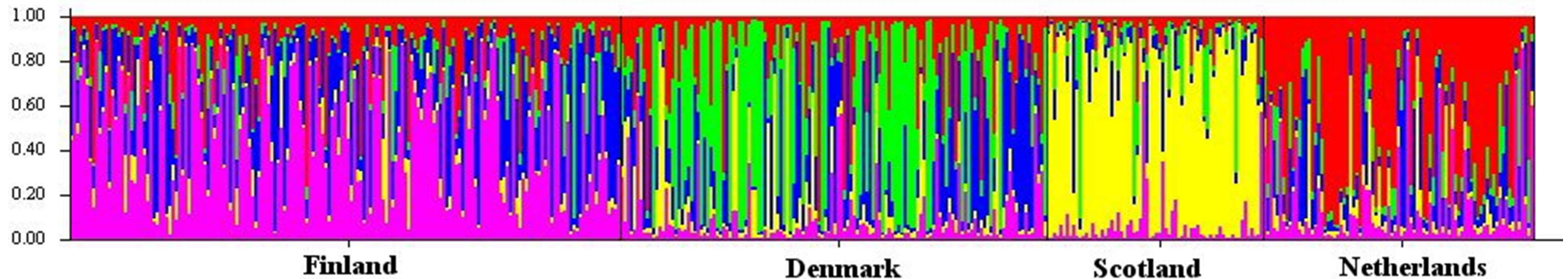

Figure 5. STRUCTURE grouping picture for K=5, which received the most support in the analysis. Each color represents one K and each individual is represented by a single vertical line broken into K=5 colored segments, meaning that the vertical line tells how much an individual is part of each cluster.

3.3. Estimates of effective population size

Estimates of effective population size varied depending on the estimation method used. TempoFs (Jorde & Ryman 2007) showed that in Finland the population was growing from 2008 to 2010, but the population sampled in Denmark was getting smaller from 2008 to 2010. In contrast, the linkage disequilibrium method implemented in NeEstimator, suggested that the Finnish population was growing until 2010, when Ne dropped, and that Denmark was growing from 2008 to 2010. This method also indicated that the Scotland population was growing from 2009 to 2010 while the Netherlands population got smaller from 2009 to 2010. Moreover, the moment based temporal -method implemented in NeEstimator showed both Finnish and Danish populations declining in Ne. Regardless of these inconsistencies, almost every method gave the highest Ne estimate to Scotland. See Table 9 for details.

Table 9. Effective population size (Ne) estimated with TempoFs, and NeEstimator

Population	**Year**	**Effective population size (Ne) estimated with different methods**				
		TempoFs	**NeEstimator**			
			Linkage disequilibrium	Moment based temporal		TM3
Finland	2008	65	102,9	140,6	Total 112,5	Total 197,2
	2009 Spring		400,9			
	2009 Fall	118	3780,3	61,0		
	2010		142,5			
Denmark	2008		44,4		Total 184,4	
	2009	69	89,4	182,9		Total 132,0
	2010	44	170,4	112,2		
	2010		137,6			
Netherlands	2009	-	749,1	74,3		Total 132,4
	2010		137,6			
Scotland	2009	-	103,2	195,7		Total 2157
	2010		441,5			

4. DISCUSSION

Even though *Pygospio elegans* is poecilogonous, having multiple developmental modes, some populations appear to be fixed for one developmental mode while others are variable (all larval types are found). I expected to find more genetic variation in populations that have planktonic larvae than in those that have brooded larvae, based on previous studies done on species with different developmental modes (Hellberg 1996, Hoskin 1997, Bohonak 1999). Results from different analyses show that my study populations are in accordance with those expectations. Gene diversity, allelic richness and expected heterozygosity were higher in Scotland which has exclusively planktonic larvae and lowest in Finland, which has exclusively brooded larvae (allelic richness: Scotland 13,197, Finland 10,933, gene diversity: Scotland 0,753, Finland 0,628, see also Table 3).

One of the goals of my study was to investigate if populations with planktonic development are temporally more unstable than populations with brooded development. AMOVA analysis did give a significant value for temporal genetic structure (p=0,0205), though the majority of the variation was found within populations (93,76 %). Pairwise F_{ST} comparisons revealed two significant values for temporal genetic structure: between Denmark 2009 and 2010 samplings and between Finland 2008 and 2010 samplings.

Denmark has all larval types, but Finland has only brooded larvae, so significant temporal genetic structure was not expected in this population. A Kruskall-Wallis analysis did not find significant values for temporal genetic structure in the population genetic parameters within any of the populations (see Table 5). These results did not support my hypothesis of more temporal genetic structure in planktonic populations.

My other goal was to show that not only developmental mode but also stability of the environment affects the population genetic parameters and temporal genetic structure. I expected to find that populations with similar habitats would give similar population genetic parameters (Finland and Denmark have stable habitats, Scotland and Netherlands have unstable habitats). This was not supported by the results. FSTAT analysis for grouped populations actually gave similar values of gene diversity and allelic richness for Finland and Netherlands (see Table 6). I also expected to find more temporal genetic structure in populations from the unstable habitats, but pairwise F_{ST} comparisons found temporal genetic structure in Finland and in Denmark, and not in Scotland and Netherlands. So, based on my study, it would seem that developmental mode has a stronger impact on population genetic parameters than the environment, but more study is needed for more conclusive results.

4.1. Genetic diversity

There was some linkage disequilibrium found, mostly in Denmark (see Appendix 1), but it did not pose a problem for the other analyses since linkage was not found in all populations. If linkage disequilibrium for the same loci had been found in all populations, independence of the loci would have been unlikely. The inconsistent linkage disequilibrium found here was more likely a spurious result. One possible reason to explain linkage disequilibrium in some populations but not in others is inbreeding (Li & Merilä 2010). This is probably not a good explanation for the linkage disequilibrium seen in my study, because there is more inbreeding in Finland than in Denmark where most of the linkage disequilibrium was detected (See F_{IS} values in Table 3). Another possible reason for linkage disequilibrium could be that population is experiencing strong natural selection, though that can not be safely assumed with small populations (Hill & Robertson 1968). Recent admixture of populations is also one possible reason for linkage disequilibrium (Peltonen et. al. 2000, Pritchart et. al. 2000, Goldstein & Weale 2001, Barilani et. al. 2007). Banks and colleagues (2000) concluded that the linkage disequilibrium that they detected in their study of chinook salmon was due to individuals from the spring-run in their winter-run sampling. Similar situation could explain the linkage disequilibrium in my study: there are two reproductive peaks in Denmark, so admixture in population samples, meaning that a sampling includes some individuals that are from the previous generation, could cause the linkage disequilibrium that was detected.

Many of my temporal population samples did not conform to Hardy-Weinberg equilibrium expectations due to heterozygosity deficiency. Heterozygosity deficiency has been detected in many marine invertebrates and possible reasons for it are inbreeding, mixing of populations (Wahlund effects), bottlenecks and founder effects, selection, and null alleles (Borsa et. al 1991, Creasey et. al. 1996, Banks et. al. 2007). Inbreeding is not a likely reason for heterozygosity deficiency in *P. elegans* since we would have expected to see more heterozygosity deficiency in Finland, where inbreeding is more likely due to brooded development of the larvae that have higher probability of settling near their parents, than it is in the other locations where planktonic development of the larvae probably reduces inbreeding. In Table 3 you can see that even though there is more inbreeding in Finland (according to the F_{IS} values), there is Hardy-Weinberg disequilibrium in all populations.

Morgan and colleagues (1999) also found heterozygosity deficiency in *P. elegans* when they used allozymes to examine genetic variability in four different populations. They considered the possibility that it was caused by bottlenecks and founder effects, since *P. elegans* is opportunistic species and there can be substantial temporal and spatial variation in its population densities. Bottlenecks and founder effects can lead to rare alleles being lost more easily and so also lead to heterozygosity deficiency. Morgan and colleagues (1999) rejected this as an explanation for the heterozygosity deficiency in *P. elegans* on the grounds that all loci should be influenced equally when there are bottlenecks and founder effects. In their study, heterozygosity deficiency was detected only in two of their five polymorphic loci. Likewise in my study not all loci are affected equally, there is very little heterozygosity deficiency in loci pe6 and pe19 (see Table 3), so this explanation probably does not apply in my study either.

Morgan and colleagues (1999) concluded that the heterozygosity deficiency they observed was probably caused by selection at specific loci. Even though neutrality of the markers used in population genetic studies is often assumed, sometimes it turns out that allozymes can be under direct selection, especially in marine invertebrates (Lemaire 2000). I used microsatellites in my study and they can be assumed to be neutral much more reliably than allozymes, so selection is not a likely cause for heterozygosity deficiency in my study. Even if the microsatellite loci that I used were close to genes that are under selection, which would mean that they would behave similarly to being under selection themselves, we would have to assume that there is similar selection in all of the four very different kinds of locations (which is unlikely), since there was Hardy-Weinberg disequilibrium equally in all of them. Also, there would probably not be heterozygosity deficiency in so many loci (five out of seven) if selection was the cause.

A Wahlund effect can come from samples collected from subpopulations that differ in allele frequencies (subpopulations can be in Hardy-Weinberg equilibrium but overall heterozygosity is reduced), which is a ‘spatial’ Wahlund effect, or it can come from samples that are from different age classes that differ in allele frequencies (‘temporal’ Wahlund effect)(Borsa et. al. 1991). A ‘Spatial’ Wahlund effect seems unlikely in *P. elegans* since my samples were collected from relatively small areas in each location, where there would probably not be subpopulation structure. Dispersal of the larvae could be expected to homogenize the populations at least in Scotland, Netherlands and Denmark. In Finland, some subpopulation structure would be more probable due to the limited dispersal of brooded larvae, but there was not more Hardy-Weinberg disequibrium here than in the other locations (See Table 3). Differences in allele frequencies between age classes could come from differences in selection that each age class experiences, or from genetic drift caused by differences in the reproductive success of individuals (sweepstakes) (Borsa et. al. 1991, Hedgecock 1994). Differences between age classes caused by differences in selection in *P. elegans* seems unlikely since the sampling locations differ so much and it would not be reasonable to think that selection differs between age classes in all of them. At the very least, differences between age classes caused by differences in selection would be very unlikely in the stable conditions of Finland where brooding predominates. Differences between age classes caused by sweepstakes could be one of the reasons for heterozygosity deficiency in *P. elegans*, although in that case there would probably be more heterozygosity deficiency in the populations where sweepstakes is more likely (populations where there is planktonic development), so Scotland should have more Hardy-Weinberg disequilibrium and Finland less, and, as mentioned before, that is not the case. Nevertheless, a ‘temporal’ Wahlund effect could be one of the reasons for heterozygosity deficiency in *P. elegans.*

The presence of null alleles means that there is a potential mutation in the primer sequence used in amplification of the loci and that all alleles are not amplified. When this occurs, an individual heterozygous for the null allele will appear to be a homozygote with only the allele that can be detected (Reece et. al. 2004). Individuals homozygous for the null allele appear as a failed reaction and might be treated as missing data. The presence of null alleles leads to detection of an excess of homozygotes and deficiency in heterozygotes. Null alleles were detected in my analysis of *P. elegans* populations, so null alleles explain at least part of the heterozygosity deficiency detected in *P. elegans*. In fact, every locus showed either high or moderate null allele frequency in some populations, and two loci, pe13 and pe15, showed null alleles in every population. FreeNa (Chapuis & Estoup 2007) is a null allele estimation and correction program that I used for my data. FreeNa first estimates the null and visible allele frequencies and then estimates the amount of real homozygotes and false ones. When using ENA correction, null alleles are excluded from the data set and F_{ST} is estimated for both the original data and to the data where null alleles are excluded. Since the results did not show that much difference between F_{ST} estimates done with the original data and the corrected data (Appendix 2), I concluded that the null alleles do not have that much effect and the original data set could be used in other analyses. Maier et. al. (2005) faced a similar situation in their study of the coral *Seriatopora hystrix*. They also detected null alleles in their microsatellite data, but used the original data in their analyses after confirming that correcting the data did not change F_{ST} results. I acknowledge that the presence of null alleles could still have some effect on my results, but the other option of removing the two loci with most of the null alleles from the analyses was also problematic. High frequencies of null alleles could be caused by higher mutation rate of the loci (null alleles are caused by mutations in the primer sequence), which could also mean that the loci are more variable (Holm et. al 2001, Hedgecock et. al. 2004). In my data set the loci (pe13 and pe15) that have the most null alleles are also the most variable loci (along with pe7, see Table 3), so removing them from the analysis would have eliminated a lot of data, reducing the power of the statistical tests. Keeping them even with the risk of some effect on the results seemed a better option.

Many studies done on marine invertebrates have found out that there is more genetic variation in species with planktonic development than in species with brooded development (Hellberg 1996, Hoskin 1997, Bohonak 1999). In light of that, I expected to find most genetic diversity in Scotland, where larvae go through planktonic development, and the least genetic diversity in the brooded developing Finland population. Results for gene diversity, allelic richness and expected heterozygosity were as expected: higher values for Scotland, lowest values for Finland (see Table 3). These results indicate that the Scotland population really is the most genetically diverse out of the four localities studied.

Since planktonic developers are expected to have larger populations and receive more alleles due to larval input from other populations (Hellberg 1996, Bohonak 1999), it could be that Scotland's large number of private alleles (alleles found only in one population) is due to the population's large size, where rare alleles are more easily preserved, and there is a flow of new alleles from other populations. Another thing that affects the number of private alleles is gene flow between populations. A barrier preventing gene flow between populations could lead to higher number of private alleles (Baus et. al. 2005). In this regard, it is surprising that the highest number of private alleles was found in Scotland, where dispersal potential of the larvae is highest due to planktonic development. It could be that in Scotland there is gene flow between other populations, but not between Scotland and the other populations that were in my study. This could be either because the distance is too long, even for planktonic larvae, or there could be some oceanographic barrier preventing gene flow.

No significant differences among temporal population samples was detected when employing a Kruskall-Wallis one way analysis of variance test on values of gene diversity, allelic richness and expected and observed heterozygosity for each location (Finland, Denmark, Scotland and Netherlands; see Table 5). These results are in conflict with the results from pairwise F_{ST} comparisons that did detect temporal genetic structure in Finland and Denmark (discussed later in Discussion section 4.2). The conflict is likely due to the differences in power between the tests. The Kruskall-Wallis test analysed summary statistics rather than the raw data and are expected to have less power (McDonald 2009). Despite these conflicts, my hypothesis of high temporal genetic structure in populations with planktonic development was not supported. There is a trend of more temporal genetic structure in the populations with planktonic development (Denmark, Netherlands and Scotland) than in the population where there is only brooded development (Finland) that can be seen in the boxplot figure 4, but the since the differences between the values that the boxplots are based are not significant, this trend can not be considered as support for my hypothesis of more temporal genetic structure in planktonic populations. The trend of overall more genetic variation in planktonic populations (Denmark, Netherlands and Scotland) than in brooded population (Finland) is nicely illustrated in this boxplot figure (see medians in Figures 4).

Analysis of gene diversity, allelic richness and observed heterozygosity after grouping the temporal population samples by location allowed statistical comparison. Results confirmed that Scotland has more gene diversity and allelic richness than the other populations, and that Finland has least, and that differences between populations are significant (see Table 6). Finland and Netherlands have similar allelic richness and gene diversity values (see Table 6). Observed heterozygosity did not vary between the locations. These results are, again, as expected and give more confirmation that genetic variation in my study populations are consistent with most of the earlier studies done with planktonic developing species and brooded developing species even though here I focus on populations that differ in developmental mode within a single species. Other studies done on the effect of developmental mode on genetic variation and population structure have considered other factors also contributing to the variation. For example, Lee and Boulding (2009) concluded in their study on four littorinid gastropods (two of which were brooded developing and two planktonic developing), that the differences in genetic variation and population structure that they detected between the brooded developers and planktonic developers were probably best explained by differences in developmental mode, but they could not completely exclude shared common ancestry of the brooded developing species and planktonic developing species as possible reason for the differences. In my study there is certainly shared ancestry between the planktonic and brooded populations since they are the same species. In light of this my results give additional support for the general expectation that population genetic parameters are heavily influenced by the developmental mode of a species, e.g. higher genetic variation in planktonic developers.

4.2. Population structure

AMOVA was used to analyse molecular variance at different levels. I found that most of the variance is among individuals within a location (93,76%). Even though the effect of temporal genetic structure within locations is small, it is significant (0,46%). AMOVA analysis also confirms that the four locations, Finland, Denmark, Netherland and Scotland, are significantly different from each other (5,79 % of the variation). Based on these results it can be said that there is temporal instability in some of my study populations, but with this analysis it can not be determined in what populations or between what years the temporal variation is. Lee & Boulding (2009) used AMOVA analysis also

in their study on four littorinid gastropods. They had larger data set than I so they could use AMOVA separately for each of their study species to determine if there was temporal genetic structure in their populations. They found only marginally significant temporal genetic structure in their planktonic developers and no temporal genetic structure in their brooded developers. This sort of analysis could be informative for *P. elegans* also (but it requires more populations/group than I have in my data set) to see in what populations there is temporal genetic structure. For my study AMOVA is good for revealing that there is temporal genetic structure, and I use other analyses to see in what populations it is concentrated (discussed below).

One way to characterize the temporal genetic structure is to do pairwise F_{ST} comparisons of the samples. Robainas and colleagues (2005) used pairwise comparisons of F_{ST}, in addition to AMOVA analysis, to find out if there was temporal genetic structure in their shrimp populations. They found that there was not significant variation between their study populations (spatial variation), but there was significant variation between sampling times (temporal genetic structure). In contrast to that, my pairwise F_{ST} comparisons showed that there was significant spatial variation (the locations differ from each other, see Table 8), but there were only two significant p-values for temporal genetic structure: between Finland 2008 and 2010 samples and between Denmark 2009 and 2010 samples. This is a very surprising result, as I had expected to find most temporal genetic structure in planktonic populations (Scotland, Netherlands and Denmark) and least in the brooded developing Finland population. Temporal genetic structure in Denmark is not as surprising since the Denmark population is larger and there are planktonic larvae also (in addition to brooded and intermediate larvae), which could increase migration between adjacent populations and make sweepstakes possible. It is harder to explain the temporal genetic structure detected in brooded developing Finland population. Genetic drift can cause temporal genetic structure in small populations (Tessier & Bernatchez 1999), and that could be one explanation. Temporal genetic structure was only detected between 2008 and 2010, and this would fit with genetic drift causing gradual change that is not detected when comparing samplings one year apart, but would show when comparing samplings two years apart. Even though temporal genetic structure in brooded developers is unexpected it does happen. Virgilio and Abbiati (2006) found temporal genetic structure in their brooded developing *Hediste diversicolor,* and they also concluded that it was probably due to genetic drift. The meaning of these results from my pairwise comparisons of F_{ST} is that my hypothesis of more temporal genetic structure in planktonic developing population is not supported. These results imply that actually planktonic developing populations are temporally stable and brooding could lead to unstable population structure on a temporal scale.

As an alternative approach, I used STRUCTURE to estimate the number of genetic clusters in the data without *a priori* definition of populations. STRUCTURE assumes a model where there are K populations and then assigns individuals to these populations based on their allele frequencies at each locus. Probabilities for different numbers of populations (K) are calculated and the most probable K represents the likely number of populations in a data set. Benefits of STRUCTURE are that it can correctly assign individuals to groups even with a small number of loci, no prior knowledge of population structure is needed, data is used more efficiently since raw data is used instead of summarized data and it can be used to study admixed populations (Pritchard et. al. 2000, Rosenberg et. al. 2001). Applying STRUCTURE to identify genetic clusters from temporal samples is a novel application of the program. For my data the most probable K is 5 (Estimated Ln Prob of Data = -12369,5), meaning that there are five different genetic groups in my data. These can be visualized in Figure 5 where different colors represent the

different genetic groups and vertical lines represent individuals. Four of those groups correspond well to the four populations based on location (Finland, Denmark, Netherlands and Scotland, mapped onto Figure 5). The fifth could come from temporal substructuring, and since Finland and Denmark are colored similarly in Figure 5, it could mean that the temporal substructuring is in these populations. If so, this STRUCTURE result supports the results from pairwise F_{ST} comparisons and there is temporal genetic structure in Finland and Denmark.

4.3. Estimates of effective population size

Estimating the effective size of a population (Ne) is very useful, but also very difficult (Waples 2006, Demant 2010). This can also be seen in my results: the different methods gave very different results and determining which method is most reliable is not straight-forward. The method based on linkage disequilibrium can give overestimation of the Ne when sample sizes are small and loci are not tightly linked (Hill 1981). My sample sizes are fairly small and linkage disequilibrium was found in just a few populations and loci (see Appendix 1), so this method is probably not very reliable for my data. TempoFs and the moment based temporal method are both based on F-statistics, they measure change in allele frequencies over time to estimate Ne (Waples 1989, Jorde & Ryman 2007). Both methods assume discrete generations, no selection and no migration. That is rarely true for actual populations, so results can be skewed if a population has overlapping generations, selection or migration. Jorde and Ryman (2007) showed that for TempoFs the bias from overlapping generations was very small if the time between samplings was long enough. Demant (2010) compared these methods and concluded that, at least for his data, TempoFs was more reliable since it handles fluctuations in population size better than the moment based temporal method. Berthier et. al (2002) compared their TM3 –method, which is likelihood-based, to methods based on F-statistics and concluded that their method gave more reliable results. Precision of all of the methods compared increased when number of loci or individuals was increased, but the TM3 method required fewest loci and individuals for reliable results. Based on these previous studies, it would seem that the TM3 method might be the most reliable method for my data, since only seven loci were used and ~30-50 individuals were sampled from each location/year. A very high estimate of Ne for Scotland (Ne = 2157) may raise questions about the suitability of this method for my data, especially since estimates from the other methods are not even close (see Table 9), but it is not unreasonable. Density in Scotland ranges from about 6000 to over 12000 individuals/m^2, much higher than the estimated Ne (Bolam 2004). It is difficult to draw conclusions from the Ne estimates that the different methods gave, since there is so much variation in them and reliability of them is hard to determine. One notable thing in the results was that two of the methods (moment based temporal and TM3) gave the highest Ne estimate for Scotland, which was expected for this population which has only planktonic development and high densities.

4.4. Developmental mode and stability of the environment

At least one study has tried to relate developmental mode in *P. elegans* and some environmental factors, such as temperature and salinity (Anger 1984), but these factors do not say anything about the stability of the environment. So, the connection between stability of the environment and developmental mode has not been studied in *P. elegans* before. There are disadvantages to planktonic dispersal of the larvae: pelagic conditions can be unpredictable, there is more predation in the pelagic environment than in the benthos, larvae can end up in an unsuitable environment after dispersal and there can be difficulty in finding a mate when larvae disperse over a large area (Johannesson 1988,

Pechenik 1999). So in an environment where benthic conditions are stable, like in Finland and Denmark, brooded development could be a better strategy for the mother: larvae get more nutrition from the mother and avoid pelagic predation (have low juvenile mortality), larvae stay in a good environment, and mates are easy to find. In Scotland and Netherlands the environment is not stable. These open habitats experience strong tides and there is not a lot of vegetation to stabilize the sediment. *P. elegans* can defend against this instability by forming dense aggregations of sand tubes which stabilize the sediment. The downside for this is that it does not stabilize the sediment just for *P. elegans*, but other marine invertebrates which are able to colonize *P. elegans* patches, such as bivalves, which in turn can lead to a decrease of *P. elegans* in the patch (Bolam & Fernandes 2003). This leads to cycle where *P. elegans* colonizes an area, stabilizes it, bivalves come to the patch and *P. elegans* has to colonize a new area and it all happens again. In this kind of situation the ability to colonize new areas fast is an advantage, and that is better achieved through planktonic development than brooded development.

In my study I look at how the stability of the environment and developmental mode affect genetic parameters and temporal genetic structure in populations differing in developmental mode and habitat type. It is difficult to distinguish how much of the variation in the parameters is due to the developmental mode and how much due to the stability of the environment. In this regard, the two interesting populations in my study are Denmark and Netherlands. Scotland has unstable environment and only planktonic larvae and Finland has stable environment and only brooded larvae. They serve as the extremes to which the other two can be compared. Denmark has both brooded and planktonic larvae (and intermediate larvae) and a habitat similar to Finland, and Netherlands has both brooded and planktonic larvae and a habitat similar to Scotland. If environment has an effect on the population genetic parameters and temporal genetic structure, not just developmental mode, then we could expect Denmark and Finland to have similar parameters and Netherlands and Scotland to have similar parameters. In my data this was not the case, and I found that Finland and Netherlands populations have very similar values of allelic richness and gene diversity (see Table 6). I expected to find more temporal genetic structure in the unstable Scotland and Netherlands populations, but instead significant temporal change in genetic structure was observed in Finland and Denmark,. It seems, based on my study, that developmental mode has a stronger effect on population genetic parameters than the environment. Understanding the connection between developmental mode and stability of the environment would require further studies with more populations and larger sample sizes. For example, it would be ideal to have multiple populations with fixed developmental mode as well as multiple populations with variable developmental mode, so a pattern could be revealed.

4.5. Possible problems in my study

As mentioned earlier, the presence of null alleles in my data could be a problem. Chapuis and Estoup (2007) studied the effects of null alleles on estimation of population differentiation and found that null alleles led to overestimation of F_{ST} and genetic distance when there was population structure (low level of gene flow) in a population. This could explain why the F_{ST} method indicated significant difference between Denmark 2009 and 2010 populations, since Denmark was also the locality where a high probability of null alleles was detected. Perhaps a similar argument could be made for Finland. Many of the analyses used in my study assume Hardy-Weinberg equilibrium and the null alleles cause deviation from Hardy-Weinberg equilibrium. So null alleles and possibly other things that cause Hardy-Weinberg disequilibrium (e.g. a ’temporal’ Wahlund effect) could have some

effect on my results. I still felt confident about using my data set despite the null alleles since ENA correction for my data did not chance the F_{ST} results (see appendix 2).

Another possible problem in my study, which could affect the estimation of temporal genetic structure is the difference in the average number of reproductive peaks between the different localities. In Finland there is only one reproductive peak/year, while in Denmark, Netherlands and Scotland there are two peaks/year. There have been fewer generations in Finland during the study period than in the other populations, and still temporal genetic structure was detected in Finland. This does not really change the conclusions I made, but it could mean that the temporal genetic structure is stronger in Finland than was detected compared to the other populations, where temporal genetic structure was not detected even though more generations had passed.

A suggestion for future study would include larger sample sizes and longer time span between samplings, which might give more statistical power and definitive answers to the questions I asked in my study. Differences in the temporal genetic structure between the populations that differ in developmental mode might be clearer if time span was longer. More individuals and more loci often improve the accuracy of the results, and if there were more populations to compare that might make the results more generalizable.

4.6. Conclusions

My study supports the expectation that there is more variation in planktonic population than in brooded populations. But when it comes to my hypothesis of more temporal genetic structure in planktonic populations, my results showed almost the opposite: temporal genetic structure was detected in a brooded population and in a population where there is both (and intermediate) developmental modes. These results could suggest that there is no sweepstakes reproductive success in *P. elegans*, which has been shown to cause temporal genetic structure in planktonic species (Hedgecock 1994, Robainas et. al. 2005, Lee & Boulding 2007). Genetic drift can cause temporal genetic structure in small populations (Tessier & Bernatchez 1999) and it probably caused the temporal genetic structure detected in Finland, so my results could mean that in *P. elegans* genetic drift has more powerful influence on temporal fluctuations in allele frequencies. My other hypothesis, that there is a connection between stability of the environment and developmental mode, was not supported either, but it could be a good candidate for further study.

ACKNOWLEDGEMENT

Huge thanks goes to my supervisors Emily Knott and Jenni Kesäniemi for all their guidance, advice and encouragement. Their help was invaluable. They also provided this very interesting research topic for my master’s thesis, which was a big part in helping me figure out what I want to do “when I grow up”, and for that I am very grateful. I also want to thank Centre of Excellence in Evolutionary Research in Jyväskylä for the 1100 € grant and *Societas pro Fauna et Flora Fennica* foundation for the 900 € grant. I also want to thank University of Jyväskylä for the Konnevesi-grant, which enabled me to go and write my thesis in the lovely and calm environment of Konnevesi research station. Thank you goes also to Sari Viinikainen and Elina Virtanen for their advice in the lab.

I also really want to give a big thank you to my husband Paavo for supporting, loving and encouraging me the whole time. My friends deserve thanks also for their support, especially Sanna Hallman for reading my text and giving me feedback. And I also want to thank my mom and little brother for their encouragement and for providing me a place where I could unwind.

REFERENCES

Anger V. 1984: Reproduction in *Pygospio elegans* in relation to its geographical origin and to environmental conditions: a preliminary report. *Fortscritte der Zoologie* 29: 45-51.

Anger K., Anger V. & Hagmeier E. 1986: Laboratory studies on larval growth of *Polydora ligni, Polydora ciliata,* and *Pygospio elegans* (Polychaeta: Spionidae). *Helgol. Meeresunters.* 40: 377–395.

Banks M.A., Rashbrook V.K., Calavetta M.J., Dean C.A. & Hedgecock D. 2000: Analysis of microsatellite DNA resolves genetic structure and diversity of Chinook salmon (*Oncorhynchus tshawytscha*) in California's Central Valley. *Can. J. Fish. Aquat. Sci.* 57: 915-927.

Banks S.C., Piggott M.P., Williamson J.E. & Beheregaray L.B. 2007: Microsatellite DNA markers for analysis of population structure in the sea urchin *Centrostephanus rodgersii. Mol. Ecol. Notes* 7: 321-323.

Barilani M., Sfougaris A., Giannakopoulos A., Mucci N., Tabarroni C. & Randi E. 2007: Detecting introgressive hybridization in rock partridge populations (*Alectoris graeca*) in Greese through Bayesian admixture analyses of multilocus genotypes. *Conserv. Genet.* 8: 343-354.

Baus E., Darrock D.J. & Bruford M.W. 2005: Gene-flow patterns in Atlantic and Mediterranean populations of the Lusitanian sea star *Asterina gibbosa. Mol. Ecol.* 14: 3373-3382.

Berthier P., Beaumont M.A., Cornuet J. & Luikart G. 2002: Likelihood-based estimation of the effective population size using temporal changes in allele frequencies: a genealogical approach. *Genetics* 160: 741-751.

Blake J.A. & Arnofsky P.L. 1999: Reproduction and larval development of the spioniform Polychaeta with application to systematics and phylogeny. *Hydrobiologia* 402: 57-106.

Bohonak A.J. 1999: Dispersal, gene flow, and population structure. *Q. Rev. Biol.* 74: 21-45.

Bolam S.G. & Fernandes T.F. 2003: Dense aggregations of *Pygospio elegans* (Clapare'de): effect on macrofaunal community structure and sediments. *J. Sea Res.* 49: 171-185.

Bolam S.G. 2004: Population structure and reproductive biology of *Pygospio elegans* (Polychaeta: Spionidae) on an intertidal sandflat, Firth of Forth, Scotland. *Invertebr. Biol.* 123(3): 260-268.

Borsa P., Zainuri M. & Delay B. 1991: Heterozygote deficiency and population structure in the bivalve *Ruditapes decussates. Heredity* 66: 1-8.

Breton S., Dufresne F., Desrosiers G. & Blier P.U. 2003: Population structure of two northern hemisphere polychaetes, *Neanthes virens* and *Hediste diversicolor* (Nereididae), with different life-history traits. *Mar. Biol.* 142: 707-715.

Chapuis M. & Estoup A. 2007: *Mol. Biol. Evol.* 24 (3): 621-631.

Cohen D. 1967: Optimization of seasonal migratory behaviour. *Amer. Nat.* 101: 5-17.

Creasey S., Rogers A.D. & Tyler P.A. 1996: Genetic comparison of two populations of deep-sea vent shrimp *Rimicaris exoculata* (Decapoda: Bresiliidae) from Mid-Atlantic Ridge. *Mar. Biol.* 125: 473-482.

Demandt M.H. 2010: Temporal changes in genetic diversity of isolated populations of perch and roach. *Corserv. Genet.* 11: 249-255.

Dempster A.P., Laird N.M. & Rubin D.B. 1977. Maximum likelihood from incomplete data via the *EM* algorithm. *J. R. Stat. Soc., Series B*, 34:1-38.

Excoffier, Laval L.G. & Schneider S. 2005: Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evol. Bioinform. Online* 1:47-50.

Falush D., Stephens M. & Pritchard J.K. 2003: Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics* 164: 1567-1587.

Gadgil M. 1971: Dispersal: Population consequences and evolution. *Ecology* 52: 253-261.

Giard A.C. 1905: La Poecilogonie. *Comp. Rend. Six. Congr. International de Zool. Berne* 1904: 617-646.

Goldstein D.B. & Weale M.E. 2001: Population genomics: Linkage disequilibrium holds the key. *Curr. Biol.* 11: R576-R579.

Goudet J. 1995: FSTAT (version 1.2): a computer program to calculate *F*-statistics. *J. Hered* . 86: 485–486.

Guo S.W. & Thompson E.A. 1992: Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics*: 48(2):361-72.

Hannerz L. 1956: Larval development of the polychaete families Spionidae SARS, Disomidae, MESNIL, and Poecilocheatidae *n. fam.* in the Gullmar Fjord (Sweden). *Zool. Bidr. Uppsala* 31: 1-204.

Hart M.W. & Marko P.B. 2010: It's about time: Divergence, demography, and the evolution of developmental modes in marine invertebrates. *Integr.Comp. Biol.* 50: 643-661.

Hedgecock D. 1994: Does variance in reproductive success limit effective population sizes of marine organisms? In: *Genetic and Evolution of Aquatic Organisms* (ed. Beaumont A.R.) Chapman & Hall, London. 122–134.

Hedgecock D., Li G., Hubert S., Bucklin K. & Ribes V. 2004: Widespread null alleles and poor cross-species amplification of microsatellite DNA loci cloned from Pacific oyster, Crassostrea gigas. *J. Shellfish Res.* 23: 379-385.

Hellberg M.E. 1996: Dependence of gene flow on geographic distance in two solitary corals with different larval dispersal capabilities. *Evolution* 50: 1167–1175.

Hill W.G. & Robertson A. 1968: Linkage disequilibrium in finite populations. *Theor. Appl. Genet.* 38: 226-231.

Hill W.G. 1981: Estimation of effective population-size from data on linkage disequilibrium. *Genet. Res* 38: 209-216.

Hoagland K.E. & Robertson R. 1988: An assessment of poecilogony in marine invertebrates: phenomenon or fantasy? *BioL. Bull.* 174: 109-125.

Holm L.E., Loeschcke V. & Bendixen C. 2001: Elucidation of the molecular basis of null allele in a rainbow trout microsatellite. *Mar. Biotechnol*. 3: 555-560.

Hoskin M.G. 1997: Effects of contrasting modes of larval development on the genetic structures of populations of three species of prosobranch gastropods. *Mar. Biol.* 127: 647–656.

Johannesson K. 1988: The paradox of Roekall: why is a brooding gastropod *(Littorina saxatilis)* more widespread than one having a planktonic larval dispersal stage *(L. littorea)? Mar. Biol.* 99: 507-513.

Jokiel P.L. 1990: Transport of reef corals into the Great Barrier Reef. *Nature* 347: 665-667.

Jorde P.E. & Ryman N. 2007: Unbiased estimator for genetic drift and effective population size. *Genetics* 177: 927-935.

Kamel S.J., Oyarzun F.X. & Grosberg R.K. 2010: Reproductive biology, family conflict and size of offspring in marine invertebrates. *Integr.Comp. Biol.* 50: 619-629.

Kreader C.A. 1996: Relief of amplification inhibition in PCR with Bovine Serum Albumin or T4 gene 32 protein. *Appl. Environ. Microbiol.* 62: 1102-1106.

Krug P.J. 2007: Poecilogony and larval ecology in the gastropod genus Alderia. *Am. Malacol. Bull.* 23: 99-111.

Krug P.J. 2009: Not my "type": Larval dispersal dimorphisms and bet-hedging in opisthobranch life histories. *Biol. Bull.* 216: 355-372.

Kube J. & Powilleit M. 1997: Factors controlling the distribution of *Marenzelleria* cf. *viridis*, *Pygospio elegans* and *Streblospio shrubsoli* (Polychaeta: Spionidae) in the southern Baltic Sea, with special attention for the response to an event of hypoxia. *Aquat. Ecol.* 31: 187-198.

Kyle C.J. & Boulding E.G. 2000: Comparative population genetic structure of marine gastropods (*Littorina* spp.) with and without pelagic larval dispersal. *Mar. Biol* 137: 835–845.

Lee H.J. & Boulding E.G. 2007: Mitochondrial DNA variation in space and time in the northeastern Pacific gastropod, *Littorina keenae*. *Mol. Ecol.* 16: 3084–3103.

Lee H.J. & Boulding E.G. 2009: Spatial and temporal population genetic structure or four northeastern Pacific littorinid gastropods: the effect of mode of larval development on variation at one mitochondrial and two nuclear DNA markers. *Mol. Ecol.* 18: 2165-2184.

Lemaire C., Allegrucci G., Naciri M., Bahri-Sfar L., Kara H. & Bonhomme F. 2000: Do discrepancies between microsatellite and allozyme variation reveal differential selection between sea and lagoon in the sea bass (*Dicentrarchus labrax*)? *Mol. Ecol.* 9: 457-467.

Levin L.A., Zhu J. & Creed E. 1991: The genetic basis of life-history characters in a poluchaete exhibiting planktotrophy and lecithotrophy. *Evolution* 45: 380-397.

Li M.H. & Merilä J. 2010: Sex-specific population structure, natural selection and linkage disequilibrium in a wild bird population as revealed by genome-wide microsatellite analyses. *BMC Evol. Biol* 10: 66.

Maier E., Tollrian R., Rinkevich B. & Nurrnberger B. 2005: Isolation by distance in the scleractinian coral *Seriatopora hystrix* from the Red Sea. *Mar. Biol* 147: 1109-1120.

McDonald J.H. 2009: Handbook of Biological Statistics (2nd ed.) Sparky House Publishing. 165-172.

McEdward L. 1995: Ecology of marine invertebrate larvae. CRC Press.

Morgan T.S., Rogers A.D., Paterson G.L.J., Hawkins L.E. & Sheader M. 1999: Evidence for poecilogony in *Pygospio elegans* (Polychaeta: Spionidae). *Mar. Ecol. Prog. Ser.* 178: 121-132.

Overden J.R., Peel D., Street R., Courtney A.J., Hoyle S.D., Peel S.L. & Podlich H. 2007: The genetic effective and adult census size of an Australian population of tiger prawns (*Penaeus esculentus*). *Mol. Ecol.* 16: 127-138.

Peakall R. & Smouse P.E., 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol. Ecol. Notes* 6, 288-295.

Pechenik J.A. 1999: On the advantages and disadvantages of larval stages in benthic marine invertebrate life cycles. *Mar. Ecol. Prog. Ser:* 177: 269-297

Peltonen L., Palotie A. & Lange K. 2000: Use of population isolates for mapping complex traits. *Nat. Rev. Genet.* 1: 182-190.

Pritchard J.K., Stephens M. & Donnelly P. 2000: Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959.

Rasmussen E. 1973: Systematics and ecology of the Isefjord marine fauna (Denmark). *Ophelia* 11: 1-495.

Reece K.S., Ribeiro W.L., Gaffney P.M., Carnegie R.B. & Allen S.K. 2004: Microsatellite marker development and analysis in the eastern oyster (*Crassolea virginica*): Confirmation of null alleles and non-Mendelian segregation ratios. *J.Hered* 95: 346-352.

Robainas B.A., Espinosa López G., Hernández D. & García-Machado E. 2005: Temporal variation of the population structure and genetic diversity of *Farfantepenaeus notialis* assessed by allozyme loci. *Mol. Ecol.* 14: 2933–2942.

Rosenberg N.A., Burke T., Elo K., Feldman M.W., Freidlin P.J., Groenen M.A.M., Hillel J., Mäki-Tanila A., Tixier-Boichard M.,Vignal A., Wimmers K. & Weigend S. 2001: Empirical evaluation of genetic clustering methods using multilocus genotypes from 20 chicken breeds. *Genetics* 159:699-713.

Schuelke M. 2000: An economic method for fluorescent labeling of PCR fragments. *Nat. Biotechnol.* 18: 233-234.

Shull D.H. 1997: Mechanisms of infaunal polychaete dispersal and colonization in an interdital sandflat. *J. Mar. Res.* 55: 153-179.

Tessier N. & Bernatchez L. 1999: Stability of population structure and genetic diversity across generations assessed by microsatellites among sympatric populations of landlocked Atlantic salmon (*Salmo salar* L.). *Mol. Ecol.* 8: 169-179.

Vermeij G.J., Palmer A.R. & Lindberg D.R. 1990: Range limits and dispersal of mollusks in the Aleutian Islands, Alaska. *The Veliger* 33: 346-354.

Virgilio M. & Abbiati M. 2006: Temporal changes in the genetic stucture of intertidal populations of Hediste diversicolor (Polychaeta: Nereididae). *J. Sea Res.* 56: 53-58.

Wai-Chuen N.G., Leung F.C.C., Solomon T.C.C., Slingby G. & Williams G.A. 2010: Temporal genetic variation in populations of the limpet *Cellana grata* from Hong Kong shores. *Mar. Biol.* 157: 325-337.

Waples R.S. 1989: A generalized approach for estimating effective population size from temporal changes in allele frequency. *Genetics* 121:379–391.

Waples R.S. 2006: A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. *Conserv. Gen.* 7:167–184.

APPENDICES

Appendix 1. Loci deviating from linkage equilibrium in each temporal population. The Bonferroni adjusted p-value for the 5 % nominal level is 0,000216 (*) and for the 1% nominal level, it is 0,000043 (**)

	FIN08	FIN09S	FIN09F	FIN10	DEN08	DEN09	DEN10	SCOT09	SCOT10	NET09	NET10
Pe7 X Pe15	0.15247	1.00000	0.04935	1.00000	0.00004**	0.00009*	0.02121	1.00000	1.00000	0.20701	0.09571
Pe7 X Pe17	0.69680	0.58251	0.97506	1.00000	0.00515	0.00004**	0.06797	1.00000	1.00000	0.73442	0.01026
Pe7 X Pe18	0.02628	1.00000	0.39195	0.21571	0.38667	0.00372	0.66693	0.40320	0.20061	0.70091	0.03212
Pe7 X Pe6	0.81818	1.00000	0.23528	1.00000	0.01506	0.14883	0.10580	1.00000	1.00000	0.43450	0.77394
Pe7 X Pe19	0.98463	1.00000	0.83468	0.32662	0.00342	0.22978	0.71095	1.00000	1.00000	0.31879	0.30009
Pe7 X Pe13	0.00013*	1.00000	0.01312	1.00000	0.00004**	0.00004**	0.02861	1.00000	1.00000	0.02489	0.03113
Pe15 X Pe17	0.91610	0.32693	0.55082	0.21857	0.00853	0.00537	0.38610	1.00000	1.00000	0.29004	0.35169
Pe15 X Pe18	0.71550	0.44537	0.30320	0.48078	0.16355	0.24593	1.00000	1.00000	0.78286	0.66680	0.33498
Pe15 X Pe6	0.62212	1.00000	0.05762	0.34918	0.02818	0.47831	0.24987	1.00000	0.87152	0.49009	0.58182
Pe15 X Pe19	0.04247	1.00000	0.23931	0.69017	0.18896	0.06052	0.58403	1.00000	0.30870	0.76805	0.57004
Pe15 X Pe13	0.05957	1.00000	0.03325	1.00000	0.00004**	0.00004**	0.00732	1.00000	1.00000	0.02355	0.07645
Pe17 X Pe18	0.55732	0.13095	0.34468	0.34567	0.37952	0.02853	0.45377	0.11883	0.85675	0.30299	0.76316
Pe17 X Pe6	0.70208	1.00000	0.41299	0.73861	0.51944	0.27961	0.06563	0.10991	0.55078	0.15437	0.32693
Pe17 X Pe19	0.78026	0.94571	0.85485	0.22619	0.04004	0.23238	0.31801	0.56325	0.19329	0.55403	0.26753
Pe17 X Pe13	0.64571	0.58216	0.47918	1.00000	0.05069	0.00004**	0.51450	1.00000	0.24619	0.44987	0.25043
Pe18 X Pe6	0.11905	0.82420	0.90104	0.14680	0.00602	0.69035	0.74918	0.03364	0.90052	0.96723	0.44693
Pe18 X Pe19	0.38610	0.43290	0.21242	0.27866	0.07260	0.20091	0.02061	0.98126	0.87177	0.00325	0.81052
Pe18 X Pe13	0.73459	0.25182	0.16541	1.00000	0.34567	0.15961	0.57160	1.00000	1.00000	0.16528	0.64095
Pe6 X Pe19	0.88792	0.04775	0.97848	0.61216	0.00398	0.53043	0.67247	0.35541	0.56338	0.07589	0.10247
Pe6 X Pe13	0.89494	1.00000	0.56831	0.38502	0.01632	0.02182	0.88879	1.00000	0.80667	0.82372	0.39584
Pe19 X Pe13	0.94727	1.00000	0.65515	1.00000	0.02416	0.10952	0.73017	1.00000	0.21433	1.00000	0.20593

Appendix 2. Estimates of global F_{ST} with and without the ENA correction for null alleles.

Locus	**F_{ST} not using ENA**	**F_{ST} using ENA**
Pe7	0.022619	0.019011
Pe15	0.017439	0.013054
Pe17	0.102196	0.131332
Pe18	0.106061	0.088899
Pe6	0.016726	0.022833
Pe19	0.024455	0.016272
Pe13	0.006584	0.005052
All loci	0.039090 (95% CI 0.015773-0.075081)	0.038869 (95%CI 0.013151-0.076623)