The origin of *Apatania zonella* in Iceland: A study based on molecular and morphological variation

Laurène Alicia Lecaudey

Faculty of Life and Environmental Sciences
University of Iceland
2013

The origin of *Apatania zonella* in Iceland: A study based on molecular and morphological variation

Laurène Alicia Lecaudey

90 ECTS thesis submitted in partial fulfillment of a *Magister Scientiarum* degree in biology

Advisor
Snæbjörn Pálsson

Examiner
Guðbjörg Ásta Ólafsdóttir

Faculty of Life and Environmental Sciences
School of Engineering and Natural Sciences
University of Iceland
Reykjavík, December 2013

The origin of *Apatania zonella* in Iceland:
A study based on molecular and morphological variation
90 ECTS thesis submitted in partial fulfillment of a *Magister Scientiarum* degree in biology

Faculty of Life and Environmental Sciences
School of Engineering and Natural Sciences
University of Iceland
Askja, Sturlugötu 7
101, Reykjavík
Iceland

Telephone: 525 4000

Bibliographic information:
Laurène Alicia Lecaudey, 2013, The origin of *Apatania zonella* in Iceland: A study based on molecular and morphological variation, Master's thesis, Faculty of Life and Environmental, University of Iceland, pp. 75.

Printing: Háskólafjölritun/Háskólaprent, Fálkagötu 2, 101 Reykjavík
Reykjavík, Iceland, December 2013

Abstract

This study focuses on the circumpolar species *Apatania zonella (Zetterstedt 1840),* a freshwater invertebrate from the order Trichoptera (caddisflies), living in cold streams, lakes and marshes. The species is widely distributed throughout the Arctic regions. As Iceland was fully covered by ice sheet during the last glacial maximum (LGM) 21 000 years ago, the current population of *A. zonella* in Iceland colonized the island after the last glaciation. Iceland is an interesting place for studying post glacial recolonization which started around 12.000 - 10.000 years ago, and species phylogeography due to its geographic position in the North Atlantic and its insular isolation. Several recent studies have shown evidences indicating that some Arctic species, including cold tolerant freshwater organisms, survived the LGM at high latitudes in cryptic refugia of the northern Palearctic within the permafrost areas. *A. zonella* is the only high Arctic Trichoptera species of Iceland indicating that this species had the potential ability to survive the LGM in northern refugia with harsher and colder conditions than the classically known refugia.

The aim of this work is to provide information regarding the origin of the Icelandic population of *A. zonella*, by studying their genetic and morphological relationships with conspecifics from neighboring countries. The phylogenetic structure and genetic diversity of *A. zonella*, were determined based on fragments of the mitochondrial gene: CO1, and three nuclear genes: Cadherin-like protein (CAD), RNA polymerase II (POL) and isocitrate dehydrogenase (IDH). Additionally, the morphological variation was investigated by measuring and statistically comparing several key traits of the genitalia of female *Apatania* with respect to their origin, including flies from Greenland and *A. hispida* from Norway.

The results from this study show that Iceland is an area of admixture or possibly a hybrid zone, colonized by at least two different lineages of *A. zonella* from Nearctic and Palearctic. Data analysis also revealed the lack of monophyly of *A. zonella*, indicating that the taxonomic status of this group should be re-evaluated. Finally, the morphological variations of the genital features allowed a good discrimination of the distinct populations into morphological and geographical groups.

Útdráttur

Þessi rannsókn fjallar um uppruna vorflugunnar lækjabyttu, *Apatania zonella (Zetterstedt 1840)* á Íslandi. Tegundin er dreifð umhverfis Norðurpól og lifir að miklu leyti sem lirfa í lækjum og vötnum. Þar sem Ísland var alþakið jökli við hámark ísþekju á síðasta kuldaskeiði fyrir um 21 þúsund árum, hefur lækjarbyttan numið land hér eftir að jöklar bráðnuðu og jafnvel eftir að síðasta kuldaskeiði lauk fyrir um 10 þúsund árum, rétt einsog nánast allar aðrar tegundir á Íslandi. Landnám Íslands og uppruni tegunda sem numu hér land er áhugavert rannsóknarefni vegna legu landsins, miðja vegu milli Ameríku og Evrópu. Ýmsar rannsóknir hafa sýnt merki um að arktískar tegundir, þ.á.m. kuldaþolnar ferskvatnstegundir, lifðu af á norðlægum breiddargráðum í kryptískum hælum (e. refugia) innan sífrera svæða ólíkt mörgum öðrum tegundum sem virðast hafa lifað af á suðlægari breiddargráðum. Lækjarbyttan er eina arktíska vorflugan sem lifir á Íslandi.

Markmið rannsóknarinnar er að varpa ljósi á uppruna lækjarbyttunar á Íslandi með athugun á erfðafræðilegri og útlitslegri aðgreiningu frá stofnum sömu tegundar í grannlöndum. Flokkunarfræðileg aðgreining og erfðafræðilegur breytileiki var greindur með raðgreiningu hvatberagensins COI og þriggja kjarna gena: Cadherin-like prótín (CAD), RNA polymerase II (POL) og isocitrat dehydrogenasa (IDH). Auk þess var útlitsbreytileiki í kynfærum kvendýra rannsakaður með tilliti til uppruna flugnanna, skoðaðar voru flugur frá Íslandi, Grænlandi og *Apatania hispida* frá Noregi.

Niðurstöður rannsóknarinnar sýna að Ísland er blendingssvæði sem hefur verið numið af tveimur ólíkum þróunarlínum innan *A. zonella*, annarri frá Ameríku og hinni frá Evrópu. Niðurstöðurnar sýna einnig að *A. zonella* er ekki einstofna, sem bendir til að rétt væri að endurskoða flokkunarfræði tegundarirnnar og jafnvel ættkvíslarinnar. Útlitsbreytileiki kynfæranna sýndi skýran mun milli þeirra þriggja hópa sem voru athugaðir.

Dedication

I dedicate this thesis to my loving parents who have always been supporting and assisting me throughout my life and education.

Table of Content

Acknowledgements

I wish to express my deepest gratitude to my advisor, Snæbjörn Pálsson, for providing me the opportunity to carry out this research project, and for his valuable advices, great assistance, encouragement, and guidance throughout the whole project. Also I am very thankful for his ideas and comments on my thesis drafts and his support during my stay in Iceland.

I am very grateful to Julien Amouret for all the help, assistance and advices he provided me in the lab and during the analysis process.

I am very thankful to Gísli Már Gíslason for his precious help for sampling and identification, and also to every person who contributed to the sampling.

I would also like to thank Etienne Kornobis and Dileepa De Croos for patiently teaching me all the methods when I started working in the lab. Thank you to all the friendly and kind coworkers in the lab. I want to thank as well my friends among the staff and students of Askja for their friendship and cheerfulness.

Thank you to all my friends in Iceland who have been a great help to survive the cold and dark winters and to enjoy the summers. I will never forget all the great times and experiences we shared. I also need to thank all my friends in France, and in the rest of the world, who supported and encouraged me when most needed.

I am deeply grateful to my parents and the members of my family for all their encouragements, patience and support in many levels during my studies.

Finally, my husband, Ehsan Pashay Ahi, gets my sincere and profound gratitude for believing in me, for helping me in every possible ways and for being a constant source of motivation throughout my studies.

General introduction

Biodiversity refers essentially to species diversity, yet intraspecies diversity is also an important component of biodiversity. Studying genotypic and phenotypic variations in natural populations of the same species, from different geographical areas, provides information about historical biogeography of the species (Avise 2000) and can lead to a better understanding of natural selection and environmental adaptation (Endler 1986). The intraspecific diversity in northern parts of the globe is one of the most critical elements on which biodiversity depends on in northern areas since the interspecies diversity tends to be significantly lower in the North area than southern areas due to harsher conditions and repetitive glaciations. The wide geographical distribution of northern species is therefore one of the key to maintain a relatively high intraspecies diversity (Pamilo & Savolainen 1999).

A species with a wide distribution includes individuals reproductively isolated which will eventually diverge genetically with time. Furthermore, species living in different environments may evolve local adaptations of some of its populations due to natural selection. Therefore genotypic and phenotypic diversity among populations is expected. Diversity at the genetical level plays a very important role in the survival and adaptability of a species, and thus in evolution of biological diversity (Frankham 2005).

Genetical data analysis is widely used nowadays for species identification, as well as to study species relationships, population structure, local adaptations, biodiversity, dispersion dynamics, colonization patterns, conservation priorities, etc. Estimations of intraspecific genetic diversity and genetical structure of populations constitute important sources of information to study the evolutionary history of a species.

The study of geographical variations in morphology is also essential to understand evolutionary processes. Morphological differences are the most visible sign of diversity, and may reflect important adaptive characters genetically determined; therefore intraspecies and interspecies morphological variations have been widely studied.

Differences, particularly in genital structures, between geographically isolated populations of the same species, are generally considered to be a precursor of allopatric speciation (Coyne & Orr 2004). The notion of species is an essential concept in biology, however various views and definitions exist for this notion, emphasizing different criteria (Agapow *et al.* 2004; Bock 2004), potentially leading to disagreement in interpretations of results, mainly regarding species identification and separation. One important assumption in species delimitation is the isolation in reproduction, forming the basis of the biological species concept (Coyne *et al.* 1988; Coyne & Orr 2004), yet currently in animal species the hybridization between species is estimated at 10% (Mallet 2005).

De Candolle (1820) was the first scientific author to hypothesize that ecological factors and historical parameters shaped the present day distribution of species. Indeed glaciation periods led to great consequences on most living species of the Arctic, subarctic and temperate areas, driving to complex vicariance and dispersal events. The intricate Pleistocene climatic variations, and its repeated glaciations and deglaciations, were essential driving forces in the restructuration of populations and the shaping of phylogeographic history, to an even greater extent for Arctic and subarctic species in northern Europe (Hewitt 2000; Tomasik & Cook 2005). The distribution, range shifts, changes in size of population, expansions, contractions are important demographic and geographic events, resulting from major climate changes that leave a great and specific imprint on the genetic diversity of the species.

Analyses of intraspecific genetic variations can help to elucidate how one specific event impacted a specific species and eventually form new hypothesis or support previous hypothesis on post-glacial events. This valuable information may provide a better understanding of the importance and the general or specific outcome of past events, and even help to predict consequences of future climatic events. Geographic distribution can be one of the main factors affecting the dynamic of the population, the gene flow, the genetic diversity and the local adaptations, which in long term might lead to reproductive isolation and subsequently to speciation (Schwenk *et al.* 2004).

The Quaternary climatic changes are the result of major variations in earth orbit around the sun (Hays *et al.* 1976). Boreal environments are harsh and unstable habitats; and due to the

last glacial maximum (LGM, 23 000 - 18 000 years ago), circumpolar species are relatively young on an evolutionary level, with a low intraspecies diversity and a wide range of distribution (Pamilo & Savolainen 1999); although it actually depends on the species and its generation rate.

Last glacial maximum turned large areas of northern Europe into uninhabitable environments for most species and therefore led to very drastic modifications regarding their habitats, distribution, genetic variability and populations size (Hewitt 2000, 2004; Schmitt 2007; Provan & Bennett 2008). This event had thus a profound influence on the later habitats colonization, gene flow level and pattern, demography, and also on present intraspecific and interspecific biodiversity (Schmitt & Seitz 2001; Wahlberg & Saccheri 2007; Previsić *et al.* 2009). Pleistocene glaciations had a deep impact on organisms of temperate areas of the northern hemisphere and even deeper on organisms of Arctic and subarctic areas. Many species became extinct and many others, as fossil records and molecular data suggest, survived solely by retreating to restricted southern refugia areas providing almost the only extensive suitable habitats during that period. Those refugia areas became the main sources of species and populations for the recolonization of the deglaciated areas after the LGM (Hewitt 2000; Petit *et al.* 2003). Northern areas and central or southern mountainous regions were covered by ice sheets and glaciers. The ice sheets covered virtually all the surface of Greenland (Gíslason 2005) and only few scattered nunataks coastal portions potentially remained ice-free during this period (Brochmann *et al.* 2003). The ice sheets from the late Weichselian covered all the high Arctic islands (Schäfer-Neth & Paul 2003) and also all the surface of Svalbard, Iceland, Scandinavia, Scotland, the Scottish islands, and reached even the South of Ireland and England (Brochmann *et al.* 2003). Most of central and eastern Europe were covered by extensive permafrost. However, some small ice-free areas existed even during the LGM presumably on the South-West coast of Norway, in the Faroe Islands and North-West of Scotland. Most of the South of England and all Ireland were ice-free. Population range of various species, both plants and animals, during the Pleistocene Ice Age, is considered to have been restricted to the main known refugia in southern Europe peninsulas (Iberian, Italian and Balkan). The species studied in this thesis, the Trichoptera *Apatania zonella (Zetterstedt 1840),* is a circumpolar freshwater species living at high latitudes. As the species is widely distributed in the Arctic, it might have survived in refugia in Beringia and North-America,

like many other plant and animal species, although their range may have been much wider and even along the ice edge, considering their current habitats.

The last Ice Age ended quite rapidly since the temperatures increased in just few decades transforming the frozen northern regions into areas with a climate warmer than present day (Dansgaard *et al.* 1993). One study of trichopteran fossils collected from Kråkenes Lake (western Norway, 62°N) has identify thermophilous Trichoptera species living around 11 400 years ago in this area which by all odds means that the summer temperatures, in this location, were at least 17°C in average (Solem & Birks 2000). The temperature decreased during the Younger Dryas leading to a new progress of the ice sheet (Atkinson *et al.* 1987) and allowing *Apatania spp.* to replace the other Trichoptera species. The post-glacial colonization of subarctic areas started around 12 000 - 10 000 years ago and was achieved by many species and populations with different ways and strategies, potentially creating allopatric populations of the same species. The recolonization of Arctic and subarctic areas by *Apatania* occurred during the Younger Dryas and early Holocene, likely with small populations and low diversity in the early period but then expanding rapidly throughout the areas (Solem & Birks 2000). For the last 8 000 years, the climate has been in an interglacial state and it can be considered rather stable (Wilcock *et al.* 2001).

Refugium has been a very important concept, especially in studies focusing on the consequences of climate changes of the late Quaternary on organisms. However, in the light of recent works on the subject, the definition of the term is somewhat widening and its usefulness is shrinking (Bennett & Provan 2008). Southern refugia areas usually present very high genetic diversity (high allelic variation and heterozygosity), while recently recolonized areas in northern Europe are characterized by lower diversity (Hewitt 1999). Therefore, the genetic variations tend to be lower in Arctic and subarctic species or populations mostly due to the almost complete coverage of these areas by glaciers or ice sheets during the last glacial maximum, and the relatively new recolonization (Hewitt 2004).

Iceland is an interesting place for studying post glacial recolonization and species phylogeography, in comparison to other countries, due to its geographic position in North Atlantic and insular isolation (Geirsdóttir *et al.* 2007). Since Iceland was almost or

completely covered by ice sheet during the LGM and recolonized after the glacier retreat, it is an important locality to study geology, paleoclimate, phylogeny and biogeography. Moreover, islands often present unique characteristics compared to continents as they show particular populations sizes, species diversity, geology, anthropization and colonization patterns. Subsequently to a severe glaciation and post-glaciation recolonization, the number of species, the population density and the distribution, in North-Atlantic islands like Iceland, are expected to follow the theory of biogeography (MacArthur & Wilson 1967). According to this theory, the number of species on an island is negatively correlated with the distance of the island from the continents and positively correlated with the size of suitable habitat provided by the island. Nonetheless, this theory does not take into account the human intervention, which greatly modified the type and number of species on most islands. The overall number of species of North-Atlantic isles appears to be higher among crustaceans than insects, which implies that crustaceans feature better adaptations for dispersal over large distances, due for instance to the dry stage of eggs allowing easier passive transport. On the contrary, aquatic insects do not possess any stage of life that allows them to be efficiently transported over long distance in feathers of waterfowls (Gíslason 2005). However, the floating ice detaching from the pack ice sheets during the deglaciation could possibly have helped the recolonization process by transporting a number of insect species from cryptic refugia in Scandinavia and the British isles to North Atlantic islands, like Iceland, which were starting to have suitable habitats for cold-adapted insects species (Buckland *et al.* 1986). More recently, freshwater insects could have been carried by humans across the Atlantic, along with domestic animals for example.

The studies of tertiary fossils reveal that the pre-Pleistocene fauna of Iceland was very similar to the current fauna of North-America (Friedrich & Simonarson 2002). However, the present day fauna of Iceland appears to be mainly originated from the Palaearctic. On the contrary, the Greenlandic fauna is predominantly originated from the Nearctic (Böcher 2001).

Bottleneck phenomenon, recolonizations, migrations, reproduction strategies, stochastic processes, genetic drift, dispersal abilities, populations isolation and local selections, among other events, can significantly affect the genetic and morphological diversity of a species between localities (Pamilo & Savolainen 1999). Changes in streams structure and

drainage patterns can also have a major effect on distribution and variations among freshwater organisms (McGlashan & Hughes 2000; Monaghan *et al.* 2001).

Phylogenetics is the most common method for studying evolutionary links between species, using mainly phylogenetic trees to display these relationships. Phylogeography is a branch of phylogenetics linking genetic variations and actual geographic distribution among and within species, using for instance information on population genetics, biogeography, geology, ecological environment, paleoclimatic events, etc. (Avise 2000). Both, phylogenetic trees and phylogeographic patterns should be considered carefully as they give only estimations and the most probable evolutionary scenario for a specific set of data, which can be different according to the method used to build the trees and the markers studied. Previously, phylogenetic trees were constructed mainly based on morphological data, first qualitative and then quantitative, but also based on physiological data or other phenotypic features. These methods were useful for taxonomic classification but for many taxa, species identification and segregation, based on phenotype, can only be done by some experts. Recently molecular data started to be used for phylogenetic purpose, reaching different conclusions than morphological methods in many cases. The traits previously used could be somehow influenced by environmental factors due to plasticity, or they could be the result of either evolutionary convergence or divergence due to selection acting on specific traits, leading therefore to erroneous conclusions about the historical relationships between individuals and species. Molecular phylogenetics uses DNA, RNA or protein sequences for the construction of phylogenetic trees. Considering the richness of information and the heritable nature, molecular phylogenetics is more reliable and also more widely accessible than phenotypic phylogenetics.

Arthropoda is by far the largest phylum of the animal kingdom with more than 1 million species already described and a global species richness estimation of 5 to 10 million species scattered in virtually all possible biotopes on Earth (Ødegaard 2000). *Apatania* belongs to the Trichoptera order (also known as caddisflies), which is the most species-rich group among aquatic insects (Solem & Birks 2000) with around 13.500 species known around the world (Holzenthal *et al.* 2007; Morse 2011), and includes 47 families (Morse 2011). Trichoptera are moth-like insects species, taxonomically and ecologically diverse, which play an important role in freshwater ecosystems (Wiggins 1996). Trichoptera is an

order of holometabolous insects which exhibit aquatic immature stages, and constitutes one of the four predominant orders of aquatic insects (Solem & Birks 2000). Arthropods are generally considered very good paleoclimate indicators because of the diversity of the phylum, their wide range of habitats, and their fast and efficient dispersal. Trichoptera are seldom used in palaeoecological studies despite the fact that their aquatic larvae are valuable primary bioindicators for the evaluation of the quality of freshwater habitats (Resh 1993; Wiggins 1996), and their aerial adults are great indicators of macroclimate conditions (Greenwood *et al.* 2003). Furthermore, such studies are facilitated by the strong sclerotized parts of Trichoptera, which are partially preserved in riverbeds or lake bottom sediments (Williams 1988, 1989) and allow species or family identification because of their characteristic features. *Apatania* genus includes about 100 species among which 30 are European (Malicky 2005; Holzenthal *et al.* 2007; Salokannel *et al.* 2010), they are the most northern caddisflies (Solem 1985; Wiggins 2004), the only high Arctic Trichoptera species of Iceland (Gíslason 1981) and the only Trichoptera species that can be found in North-West, North and North-East of Greenland (Stoltze 1981; Bennike *et al.* 2000). This particularity suggests that they had the potential ability to survive the LGM in northern refugia with harsher and colder conditions than the classically known refugia. *Apatania zonella* is one of the most northern Trichoptera as it can be found as far North as and 80° latitude on Svalbard (Norway) (Solem *et al.* 1977; Andersen & Wiberg-Larsen 1987; Gíslason 2005). It is also the only *Apatania* species occurring in high altitude lakes (above 1200m) (Solem 1985).

A. zonella is a circumpolar species, widely distributed in subarctic and mountainous areas of North America, and lowlands of Greenland, Iceland, North-Atlantic Islands, Scandinavia, Baltics, and Arctic Russia (Solem & Birks 2000; Gíslason 2005; Salokannel *et al.* 2010) while *A. hispida (Forsslund 1930)* is Scandinavia specific. *A. zonella* lives in cold and clear freshwaters in larval stage, in lowlands and mountains, mostly in cold running waters, springs, streams, rivers (lotic waters), but sometimes also in cold stony lake shores, ponds, marshes (lentic waters); feeding on periphyton, mainly by scraping diatoms and other organisms from stones in rivers and lakes (Nielsen 1943; Gíslason & Sigfusson 1987; Solem & Birks 2000; Holzenthal *et al.* 2007). It can be found at various altitudes, from mountains to valleys, occupying streams where the temperatures are generally between 0 and 5°C most of the year, and can reach about 10°C in summer. Its life

cycle may last over two years (Solem 1985) and is univoltine (Gíslason 1981) or semivoltine in some cases (Andersen & Wiig 1987). The larvae are highly cold tolerant and have the ability to endure rigorous winter conditions. Nonetheless, this aquatic larval stage is responsive to several abiotic factors such as water temperature (by extension, weather and climate conditions), pollution, flow and also biotic factors such as food availability, predators pressure and competition (Greenwood *et al.* 2003). Individuals become terrestrial in adult stage, occurring in lakeside or streamside habitats, reaching a length of 1cm and a wet weight of 15-20mg on average. This aerial adult stage, which is the crucial part of life for the dispersal and reproduction of the individual, is mostly responsive to the weather and climate conditions (Greenwood *et al.* 2003). The short-lived adults of *A. zonella* are greatly dependent on the summer temperature during and after the emergence period. *A. zonella*, and some other species of *Apatania* genus, are known to reproduce essentially by parthenogenesis. The influence of males on the population is currently unknown, on average less than 2% of males can be found among *A. zonella* populations, and even less in lowlands (Corbet 1966; Solem 1985; Andersen & Wiberg-Larsen 1987; Andersen & Wiig 1987; Raastad & Solem 1989; Salokannel *et al.* 2010). *Apatania hispida* shows no evidence of any existence of males, and is thus considered totally parthenogenetic (Malicky 2004; Salokannel *et al.* 2010).

In *A. zonella*, DNA markers and genitalia morphology are two tools that could help to understand populations structure and colonization processes in Iceland, as variation patterns can be due to their distribution, ecological conditions, and also to the population past history like expansion events, vicariance, migration, bottleneck, etc. (Edwards & Beerli 2000). Therefore, it might be very informative to study the relationship between populations of the same species and closely related species in different countries; especially in the case of Iceland which is a geographically isolated and geologically young island. Repeated recolonization events and from different origins can though obscure such relationships of genetic and morphological distances with geographical distances.

Mitochondrial DNA (mtDNA) is in most species solely maternally inherited, enabling studies on maternal lineage; and as it does not include any recombination it facilitates the interpretations of the genealogical relationships. Mitochondrial DNA is a double-stranded DNA molecule containing 37 genes in animals, among which 13 are coding for proteins,

and other for tRNA and rRNA subunits. It has a high mutation rate, allowing a rather efficient tracking and understanding of evolutionary processes and genetic distances (Hartl & Clark 1997). The rate of sequence divergence in freshwater invertebrates is estimated to be between 1.4% to 2.6% per million years (Schubart 1998; Knowlton & Weigt 1998). Molecular markers with a high mutations rate can lead to unreliable phylogenetic tree of an ancient divergence due to saturation of the signal, while markers with low mutations rate may give insufficient data for the reconstruction of recent divergence. The mtDNA sequence comparisons play a major role in the construction of phylogenetic trees and the barcode of the life project (BOLD), but also for confirmation or invalidation of species and subspecies separation. Furthermore, the study of mtDNA enables estimations of genetic relationships between populations and groups of individuals within a species (Avise 2000; Watanabe *et al.* 2003), and appears to be effective with lower samples sizes than nuclear markers (Castro *et al.* 2010). Nonetheless, phylogenetic conclusions drawn from mtDNA data may be contentious as they can be inconsistent with those from nuclear markers (Watanabe *et al.* 2003; Rubinoff & Holland 2005), also it cannot provide information strictly related to males due to its maternal inheritance, thereby lacking that part of the genetic information. Assessment of variation of markers from the nuclear genome is thus warranted to get a better picture of the genealogical relationships among populations.

The aim of this study is to shed light on the origin of the current Icelandic population of *Apatania zonella*, by investigating the association of genetic or morphometric relationships among individuals in the Arctic and subarctic populations of this species, with respect to geographic origin. Studies focused on intraspecific genetic variation in this species could lead to a better understanding of populations distribution, colonization patterns and survival processes of Arctic freshwater invertebrates during glaciations, post-glacial recolonizations and expansions.

References

Agapow P-M, Bininda-Emonds ORP, Crandall KA *et al.* (2004) The impact of species concept on biodiversity studies. *The Quarterly Review of Biology*, **79**, 161–179.

Andersen T, Wiberg-Larsen P (1987) Revised check-list of NW European Trichoptera. *Insect Systematics & Evolution*, **18**, 165–184.

Andersen T, Wiig Ø (1987) Sex proportions, flight periods and morphometric variation in West Norwegian Apatania zonella (Zetterstedt, 1840) (Trichoptera: Limnephilidae). *Insect Systematics & Evolution*, **18**, 185–204.

Atkinson TC, Briffa KR, Coope GR (1987) Seasonal temperatures in Britain during the past 22,000 years, reconstructed using beetle remains. *Nature*, **325**, 587–592.

Avise JC (2000) *Phylogeography: The history and formation of species*. Harvard University Press.

Bennett K, Provan J (2008) What do we mean by "refugia"? *Quaternary Science Reviews*, **27**, 2449–2455.

Bennike OLE, Björck S, Böcher J, Walker IANR (2000) The Quaternary arthropod fauna of Greenland: a review with new data. *Bulletin of the Geological Society of Denmark*, **47**, 111–134.

Böcher J (2001) *Insekter og andre smådyr: i Grønlands fjeld og ferskvand*. Forlaget Atuagkat/Ilisimatusarfik.

Bock WJ (2004) Species: the concept, category and taxon. *Journal of Zoological Systematics and Evolutionary Research*, **42**, 178–190.

Brochmann C, Gabrielsen TM, Nordal I, Landvik J, Elven R (2003) Glacial survival or tabula rasa ? The history of North Atlantic biota revisited. *Taxon*, **52**, 417–450.

Buckland PC, Gerrard AJ, Larsen G *et al.* (1986) Late Holocene palaeoecology at Ketilsstadir in Mýrdalur, South Iceland. *Jökull*, **36**, 41–55.

De Candolle AP (1820) *Essai élémentaire de géographie botanique* (Levrault, Ed,).

Castro JA, Picornell A, Ramon M (2010) Mitochondrial DNA: a tool for populational genetics studies. *International Microbiology*, **1**, 327–332.

Corbet PS (1966) Parthenogenesis in caddisflies (Trichoptera). *Canadian Journal of Zoology*, **44**, 981–982.

Coyne JA, Orr HA (2004) *Speciation*. Sinauer Associates, Sunderland, MA, USA.

Coyne J, Orr H, Futuyma D (1988) Do we need a new species concept? *Systematic Zoology*, **37**, 190–200.

Dansgaard W, Johnsen SJ, Clausen HB *et al.* (1993) Evidence for general instability of past climate from a 250-kyr ice-core record. *Nature*, **364**, 218–220.

Edwards S, Beerli P (2000) Perspective: gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution*, **54**, 1839–1854.

Endler JA (1986) *Natural Selection in the Wild*. Princeton University Press.

Frankham R (2005) Genetics and extinction. *Biological Conservation*, **126**, 131–140.

Friedrich WL, Simonarson LA (2002) The Tertiary flora of Iceland - a witness of a migration route between Europe and North America. In: *6th European Palaeobotany - Palynology Conference* , pp. 183–184.

Geirsdóttir Á, Miller GH, Andrews JT (2007) Glaciation, erosion, and landscape evolution of Iceland. *Journal of Geodynamics*, **43**, 170–186.

Gíslason GM (1981) Distribution and habitat preferences of Icelandic Trichoptera. In: *Proceedings of the 3rd International Symposium on Trichoptera* (ed Moretti GP), pp. 99–109.

Gíslason GM (2005) Origin of freshwater fauna of the North-Atlantic islands: present distribution in relation to climate and possible migration routes. *Internationale Vereinigung für Theoretische und Angewandte Limnologie Verhandlungen*, **29**, 198–203.

Gíslason GM, Sigfusson AT (1987) The life cycle and food of Apatania zonella (Zett.) in a spring-fed stream in SW Iceland (Trichoptera: Limnephilidae). In: *Proceedings of the 5th International Symposium on Trichoptera* (eds Bournaud M, Tachet H), pp. 237–242.

Greenwood MT, Agnew MD, Wood PJ (2003) The use of caddisfly fauna (Insecta: Trichoptera) to characterise the Late-glacial River Trent, England. *Journal of Quaternary Science*, **18**, 645–661.

Hartl DL, Clark AG (1997) *Principles of Population Genetics*. Sinauer Associates Inc., Sunderland, Massachusetts.

Hays JD, Imbrie J, Shackleton NJ (1976) Variations in the Earth's orbit: pacemaker of the Ice Ages. *Science*, **194**, 1121–1132.

Hewitt G (1999) Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, **68**, 87–112.

Hewitt GM (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.

Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society. Series B, Biological Sciences*, **359**, 183–195.

Holzenthal RW, Blahnik RJ, Prather AL, Kjer KM (2007) Order Trichoptera Kirby, 1813 (Insecta), Caddisflies. *Zootaxa*, **1668**, 639–698.

Knowlton N, Weigt LA (1998) New dates and new rates for divergence across the Isthmus of Panama. *Proceedings of the Royal Society. Series B, Biological Sciences*, **265**, 2257–2263.

MacArthur RH, Wilson EO (1967) *The Theory of Island Biogeography*. Princeton University Press, Princeton.

Malicky H (2004) *Atlas of European Trichoptera., The Nederlands*. Springer.

Malicky H (2005) Ein kommentiertes Verzeichnis der Köcherfliegen (Trichoptera) Europas und des Mediterrangebietes. *Linzer Biologische Beiträge*, **37**, 533–596.

Mallet J (2005) Hybridization as an invasion of the genome. *Trends in Ecology & Evolution*, **20**, 229–237.

McGlashan DJ, Hughes JM (2000) Reconciling patterns of genetic variation with stream structure, earth history and biology in the Australian freshwater fish Craterocephalus stercusmuscarum (Atherinidae). *Molecular Ecology*, **9**, 1737–1751.

Monaghan MT, Thomas SA, Minshall GW *et al.* (2001) The influence of filter-feeding benthic macroinvertebrates on the transport and of particulate and diatoms in two streams deposition organic matter we thank. *Limnology and Oceanography*, **46**, 1091–1099.

Morse JC (2011) The Trichoptera World Checklist. *Zoosymposia*, **5**, 372–380.

Nielsen A (1943) Apatidea auricula Forsslund from a Norwegian mountain lake. Description of the imago and notes on the biology. *Entomologiske Meddelelser*, **23**, 18–30.

Ødegaard F (2000) How many species of arthropods? Erwin's estimate revised. *Biological Journal of the Linnean Society*, **71**, 583–597.

Pamilo P, Savolainen O (1999) Post-glacial colonization, drift, local selection and conservation value of populations: a northern perspective. *Hereditas*, **238**, 229–238.

Petit RJ, Aguinagalde I, de Beaulieu J-L *et al.* (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. *Science*, **300**, 1563–1565.

Previsić A, Walton C, Kucinić M, Mitrikeski PT, Kerovec M (2009) Pleistocene divergence of Dinaric Drusus endemics (Trichoptera, Limnephilidae) in multiple microrefugia within the Balkan Peninsula. *Molecular Ecology*, **18**, 634–647.

Provan J, Bennett KD (2008) Phylogeographic insights into cryptic glacial refugia. *Trends in Ecology & Evolution*, **23**, 564–571.

Raastad JE, Solem JO (1989) Autogeny as successful reproductive strategy in high altitude black-flies (Diptera, Simuliidae). *Annales de Limnologie*, **25**, 243–249.

Resh VH (1993) Recent trends in the use of Trichoptera in water quality monitoring. In: *Proceedings of the 7th International Symposium on Trichoptera* , pp. 285–291. Backhuys Publishers, Leiden, The Netherlands.

Rubinoff D, Holland B (2005) Between two extremes: mitochondrial DNA is neither the Panacea nor the Nemesis of phylogenetic and taxonomic inference. *Systematic Biology*, **54**, 952–961.

Salokannel J, Rantala MJ, Wahlberg N (2010) DNA-barcoding clarifies species definitions of Finnish Apatania (Trichoptera: Apataniidae). *Entomologica Fennica*, **21**, 1–11.

Schäfer-Neth C, Paul A (2003) The Atlantic Ocean at the last glacial maximum: 1. Objective mapping of the GLAMAP sea-surface conditions. In: *The South Atlantic in the Late Quaternary: Material Budget and Current Systems* (ed Wefer G., Mulitza S. and RV (eds)), pp. 531–548. Springer Berlin Heidelberg.

Schmitt T (2007) Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. *Frontiers in Zoology*, **4**, 11.

Schmitt T, Seitz A (2001) Intraspecific allozymatic differentiation reveals the glacial refugia and the postglacial expansions of European Erebia medusa (Lepidoptera: Nymphalidae). *Biological Journal of the Linnean Society*, **74**, 429–458.

Schubart CD (1998) Rapid evolution to terrestrial life in Jamaican crabs. *Nature*, **393**, 363–365.

Schwenk K, Junttila P, Rautio M *et al.* (2004) Ecological, morphological, and genetic differentiation of Daphnia (Hyalodaphnia) from the Finnish and Russian subarctic. *Limnology and Oceanography*, **49**, 532–539.

Solem JO (1985) Norwegian Apatania Kolenati (Trichoptera: Limnephiiidae): identification of larvae and aspects of their biology in a high-altitude zone. *Insect Systematics & Evolution*, **16**, 161–174.

Solem JO, Birks H (2000) Late-glacial and early-Holocene Trichoptera (Insecta) from Kråkenes Lake, western Norway. *Journal of Paleolimnology*, **23**, 49–56.

Solem JO, Senstad E, Bergvik T, Hegstad A (1977) New record of Apatania zonella Zett.(Trichoptera, Limnephilidae) from Svalbard. *Norwegian Journal of Entomology*, **24**, 85.

Stoltze M (1981) *Grønlandske vårfluers udbredelse og indvandringshistorie*.

Tomasik E, Cook JA (2005) Mitochondrial phylogeography and conservation genetics of wolverin (Gulo gulo) of Northwestern North America. *Journal of Mammalogy*, **86**, 386–396.

Wahlberg N, Saccheri I (2007) The effects of Pleistocene glaciations on the phylogeography of Melitaea cinxia (Lepidoptera : Nymphalidae). *European Journal of Entomology*, **104**, 675–684.

Watanabe K, Mori S, Nishida M (2003) Genetic relationships and origin of two geographic groups of the freshwater threespine stickleback, "Hariyo." *Zoological Science*, **20**, 265–274.

Wiggins GB (1996) *Larvae of the North American caddisfly genera (Trichoptera)*. University of Toronto Press, Scholarly Publishing Division.

Wiggins GB (2004) *Caddisflies: The Underwater Architects*. University of Toronto Press.

Wilcock HR, Hildrew AG, Nichols RA (2001) Genetic differentiation of a European caddisfly: past and present gene flow among fragmented larval habitats. *Molecular Ecology*, **10**, 1821–1834.

Williams NE (1988) The use of caddisflies (Trichoptera) in palaeoecology. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **62**, 493–500.

Williams N (1989) Factors affecting the interpretation of caddisfly assemblages from Quaternary sediments. *Journal of Paleolimnology*, **1**, 241–248.

Chapter 1: Genetic variation in *Apatania zonella*

Introduction

Phylogeography of Arctic species

The various analysis on the phylogeography of populations, from previously glaciated areas, have brought support and contradictions to previous theories on the postglacial recolonization of populations and species at high latitudes (Hewitt 2000). Previous studies have identified the main refugia, cryptic refugia and postglacial colonization routes for different organisms across the Holarctic (Hewitt 2001). Several recent studies have shown evidence of fossils or genetic patterns indicating that some Arctic species survived the Last glacial maximum (LGM) at high latitudes and have given support to the existence of previously unknown refugia, mainly in Nearctic areas and Asia (Abbott *et al.* 2000; Hewitt 2004; Kotlík *et al.* 2006; Brunhoff *et al.* 2006; Schmitt 2007; Bhagwat & Willis 2008; Provan & Bennett 2008). Cryptic refugia in northern Palearctic have also been described in former permafrost areas during the cold periods of Ice Age (Pinceel *et al.* 2005; Deffontaine *et al.* 2005; Gum *et al.* 2005; Pauls *et al.* 2006). However, phylogeographic studies of Arctic animals have mostly focused on terrestrial organisms (Conroy & Cook 2000; Milá *et al.* 2000; Latch *et al.* 2009; Bray *et al.* 2013). Yet, little is known about the differences and similarities between terrestrial vertebrates or invertebrates and freshwater invertebrates regarding evolutionary patterns and survival strategies during glaciations (Benke *et al.* 2009). It is commonly assumed that the responses to glaciation events might be very distinct between freshwater and terrestrial populations since their habitats are entirely different. Indeed, some studies show that the influence of Pleistocene glaciations had a different effect on aquatic organisms than terrestrial ones (Hänfling *et al.* 2002; Pauls *et al.* 2006; Brändle *et al.* 2007). Additionally, there are only few studies focusing on the effects of Pleistocene glaciations on aquatic invertebrates. However, recent studies have shown evidences of survival of cold tolerant freshwater organisms in small northern glacial refugia (Pauls *et al.* 2006; Benke *et al.* 2009).

Survival possibilities for freshwater invertebrate in northern areas

The survival of species at high latitudes during the glaciation period is depended on their type of habitat, tolerance to cold temperatures and dispersal abilities (Pinceel *et al.* 2005; Deffontaine *et al.* 2005). Dispersal of aquatic species such as the Trichoptera *Apatania zonella (Zettersted 1840)* is mostly limited by the hydrographic systems, speed flow and depth, at least at the larval stage. Nonetheless, permanent running waters usually do not freeze, potentially allowing the survival of several subarctic freshwater invertebrates in streams of central Europe, North America, North and central Asia and Siberia, without retreating to southern refugia such as the Mediterranean region (Pauls *et al.* 2006). Subarctic freshwater taxa are less restricted in terms of habitats and climate, therefore they could potentially have survived in small cryptic northern ice-free areas during the last glacial maximum (LGM, 23 000 - 18 000 years ago), unlike temperate terrestrial species which were more strictly confined to southern refugia (Verovnik *et al.* 2005; Pauls *et al.* 2006).

In Iceland, volcanic and geothermal activity could have helped to provide small patches of habitats, or even unglaciated perennial lakes, suitable enough to allow the survival of some multicellular organisms (Vogel 2008; Kornobis *et al.* 2010). However, it is generally agreed that during the coldest period of the LGM, no boreal species could have survived, in spite of possible existence of ice free areas in Iceland (Gíslason 2005). A recent counter-example was though found for two subterranean amphipods species (Kornobis *et al.* 2010) which survived in sub-glacial refugia in groundwater, along the tectonic plate boundary in Iceland. A cryptic survival of freshwater species within Iceland, and its adjacent areas such as Greenland and Scandinavia, during the LGM, is still considered unlikely. Although few strongly cold-resistant organisms may have survived cold periods in isolated nunataks, the fauna could hardly have survived the coldest glacial stage of the Quaternary (Bennike *et al.* 2000). The early colonizers are more likely to have originated from unglaciated patches within the permafrost, at the limit of the ice sheet, than from distant refugia (Kearney 2005). Although the new habitats opening during the deglaciation period were harsh, they may have offered low competition and low predation pressure. Such a scenario may have provided *A. zonella*, adapted to survive in cold environments, a great advantage for the recolonization of deglaciated areas (Geffen *et al.* 2007).

Dispersal

Migratory birds are considered as an occasional mean of transport of freshwater species (Bennike *et al.* 2000). However, aquatic insects do not present any life stages allowing transportation over long distance by waterfowl species (Gíslason 2005). Stones from freshwater streams, used as ballast by humans, could presumably have served as vectors for passive dispersal of *Apatania* between areas and countries. Active dispersal between streams for *A. zonella* is essentially limited to few kilometers during the adult stage (Sode & Wiberg-Larsen 1993; Petersen *et al.* 1999). The dispersal capacity is expected to have important consequences on gene flow, colonization and the evolutionary processes observed within the species.

Asexuality in Arctic areas

Apatania zonella reproduces essentially by parthenogenesis and the proportion of males is estimated to be 1-2% in average (Corbet 1966; Gíslason 1977; Raastad & Solem 1989; Salokannel *et al.* 2010). The short generation rate, the extent of parthenogenesis, the small body size and the ability to tolerate harsh environment and cold temperature, are major ecological features for a potential survival in northern glacial refugia (Bhagwat & Willis 2008). Fully and partial parthenogenetic populations, which are dominating in marginal habitats, are characterized by high genetic drift, high turn-over, weak biotic interactions and genetic isolation (Haag & Ebert 2004). Genetic diversity is commonly lower for insular populations of a species, as a result of geographical isolation, and generally smaller population sizes (Geffen *et al.* 2007). These island characteristics frequently lead to low diversity through genetic inbreeding (Frankham 1998), but parthenogenesis allow species to circumvent the disadvantages of inbreeding (Haag & Ebert 2004). The ability to colonize rapidly new areas, facilitated by parthenogenesis, and the instability of the marginal environments asexual populations live in, affects their geographical patterns of distribution (Law & Crespi 2002). Climatic changes have undoubtedly played a role in the success of asexual reproduction (Raastad & Solem 1989) but major changes in the geographical distribution of a species, following a climate change, have often lead to the transition from

sexual reproduction to parthenogenesis (Kearney 2005). The Quaternary glacial events have been claimed to be the main reason for the occurrence of parthenogenesis lineages in northern regions of Europe, which were repeatedly glaciated during the Pleistocene (Horne & Martens 1999; Kearney 2005).

Evolutionary analysis

In order to study the evolutionary history of species by using genetical markers, it is valuable to study several markers and compare the trees obtained from each marker, as different genes and genetic markers can have distinct evolutionary histories (Pamilo & Nei 1988). The different parts of the genome of each organism evolve at specific variable rates, and this can affect significantly the conclusions from different datasets (Avise 1994). The mitochondrial gene Cytochrome oxidase (CO1) is characterized by high mutation rate and relatively constant over time (Nei & Kumar 2000), it is a marker of choice in many studies aiming to describe the genetical structure at intra and interspecific level among invertebrates (Couceiro *et al.* 2007; de Croos & Pálsson 2010; Kornobis *et al.* 2011; Naro-Maciel *et al.* 2011). Mitochondrial and nuclear markers show different patterns of intraspecific variations, and mechanisms of evolution, partly due to major differences in the inheritance process and ploidy.

Genetic diversity indices are an efficient way to evaluate the diversity of individuals from different populations, and trace the historical relationships between populations. These indices can guide the interpretation of the phylogenetic trees from each marker and they can be supplemented by expansion estimation and neutrality tests. Molecular genetics is an important tool to identify different lineage of a species, and studying how it relates to the geographical distribution of the species might provide key clues in the past recolonization routes.

Aims of the study

The objective of this work is to resolve the origin and past history of the Icelandic population of *Apatania zonella*, by studying genetic relationships among individuals of different populations and determining how they are linked to their geographical origin. This study aims to assess the influence of the Pleistocene glaciations on the current population distribution and species diversity of *A. zonella*, and to investigate if the species retreated in southern refugia or survived within the permafrost in Northern refugia during the LGM. Furthermore, this project intends to determine if the Icelandic population of *A. zonella* is originated from a single or multiple colonization events from North-America and/or mainland Europe. To achieve this goal, the phylogenetic structure and the genetic diversity were determined based on fragments of the mitochondrial gene: CO1, and three nuclear genes: Cadherin-like protein (CAD), RNA polymerase II (POL) and isocitrate dehydrogenase (IDH). In the case of a single colonization event in Iceland, we would expect few closely related haplotypes, while in the case of multiple colonization events, from the two continents, we would expect distinct lineages to be mixed within Iceland.

Material and methods

Sampling

Adults and larvae of *A. zonella* were collected between 2005 and 2012 from different locations in Iceland (37) by *Gísli Már Gíslason* and coworkers from the University of Iceland, and also by *Erling Ólafsson* from the Institute of Natural History (Figure 1, Appendix 1). Twelve individuals were collected in Greenland by *Gísli Már Gíslason* and *Snæbjörn Pálsson* in 2010 (Figure 1, Appendix 1). In addition, one individual from Alaska and 4 individuals of closely related *Apatania* species from Finland were provided by collaborators (Figure 1, Appendix 1). The Icelandic adults individuals of *A. zonella* were collected during flying period (between the beginning of May end the end of August) using nets and traps placed next to stony cold springs, mountains brooks and subarctic rivers in South, South-West, West and North-West (Westfjords) (Figure 1, Appendix 1). The larvae of *A. zonella* sampled in South-West of Iceland, were collected directly from the underside of stones in cold springs. A specimen of *Limnephilus griseus* was sampled and used as an outgroup in the analysis. The collected samples were preserved individually in 96% Ethanol. The species identification of *A. zonella* based on morphological characteristics was done by *Gísli Már Gíslason*, a Trichoptera specialist.

Several CO1 mitochondrial DNA sequences were provided by *Karl Kjer* working at the international barcode of life project *(unpublished):* United States n=6, Canada n=6, Iceland n=4, Greenland n=3, Norway n=2, Mongolia n=1. CO1 sequences from seven species of the genus *Apatania* from Finland were also obtained from Salokannel *et al.* (2010) *(A. stigmatella, A. wallengreni, A. muliebris, A. hispida, A. auricular, A.forsslundi* and *A. dalecarlica)*; and finally 11 Icelandic samples and one Greenlandic sample of *A. zonella* from previous unpublished students projects (Prieto-Arribas 2009; Sanz-Muñoz 2010).

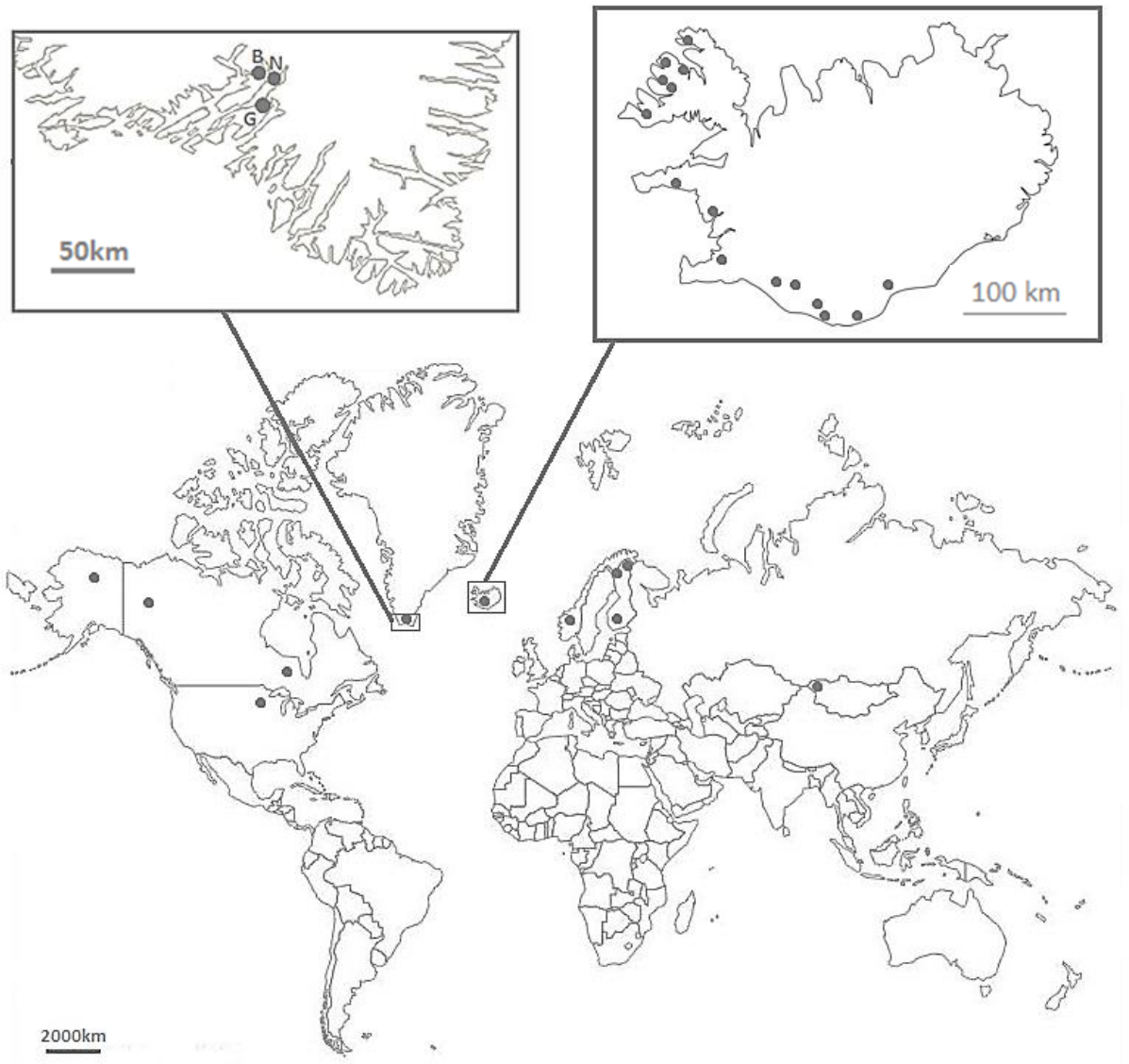

Fig 1. Schematic map of the world with all the sampling sites of *Apatania zonella* used in this study, with a zoom on the sampling areas in Greenland and Iceland (Grey dots: sampling sites, **B**: Brattahlid, **N**: Narsarsuaq, **G**: Gardar)

Genomic DNA Extraction

DNA was extracted from the head or one foreleg with part of the chest of the ethanol preserved individuals. The tissue was digested in 1.5ml Eppendorf containing 250µl of Chelex® 6% (Bio-Rad) and 2.5µl of 10% Proteinase K (final concentration=0.1mg/ml). The solution was incubated in a thermomixer at 65°C for at least 3 hours with 450 rpm shaking, and at 95°C for 10min to deactivate the proteinase K. Finally, the samples were

centrifuged at 12000 rpm for 2min to precipitate the Chelex® resin. The extracted DNA was then stored at 4°C.

Amplification

Polymerase chain reactions (PCR) were used to amplify each genomic region of interest for each sample in a 20µl reaction volume consisting of 9,86µl of deionized distilled H_2O, 1.65µl of dNTPs (0.2mM), 2.2µl Tween 20 (1%), 2.2µl Taq Buffer, 2.2µl BSA (0.5mg/ml) 0.7µl of each primer (10pM, forward and reverse), 0.2µl Taq polymerase (5U/µl) and 2µl of the template DNA from the sample (30 - 200ng). The DNA concentration was measured beforehand using a Thermo Scientific NanoDrop™ 1000 Spectrophotometer (V3.5, 2008 Thermo Fisher Scientific Inc.).

Fragments of 1190 base pairs of cytochrome oxidase subunit 1 (CO1) mitochondrial gene were amplified using the primers pair LCO1490/S20 (Folmer *et al.* 1994; Pauls *et al.* 2003) (Table 1). Fragments of 772 base-pair of POL, 850 base-pair of CAD and 720 base-pair of IDH nuclear genes were amplified using specific primers (Table 1).

Table 1. Detail description of the primers characteristics used for PCR and sequencing in this study.

Gene	Complete Name	Primer Name	Primer Direction	Primer sequence (5'-3')	Source
CO1	Cytochrome oxidase subunit 1	LCO1490	Forward	GGTCAACAAATCATAAAGATATTGG	Folmer *et al.* 1994
		S20	Reverse	GGGAAAAAGGTTAAATTTACTCC	Pauls *et al.* 2003
CAD	Cadherin-like protein	743nF-ino	Forward	GGIGTIACIACIGCITGYTTYGARCC	Johanson & Malm 2010
		1028r-ino	Reverse	TTRTTIGGIARYTGICCICCCAT	Johanson & Malm 2010
POL	RNA polymerase II	POLFOR2	Forward	TGGGAYGSYAAAATGCCKCAACC	Danforth *et al.* 2006
		POLREV2	Reverse	TYYACAGCAGTATCRATRAGACCTTC	Danforth *et al.* 2006
IDH	isocitrate dehydrogenase	deg27F-ino	Forward	GGWGAYGARATGACIAGRATHATHTGG	Malm & Johanson, 2011
		degR-ino	Reverse	TTYTTRCAIGCCCAIACRAAICCICC	Malm & Johanson, 2011

The optimized thermal cycling conditions for CO1 mtDNA fragment amplification were: initial denaturation at 94°C for 4min, followed by 40 cycles of denaturation at 94°C for 30sec; annealing at 45°C for 45sec, and extension at 72°C for 1min, with a final extension at 72°C for 6min. The optimized PCR conditions for CAD, POL and IDH fragment amplification were: initial denaturation at 94°C for 2min, followed by 10 cycles of

denaturation at 94°C for 30sec; annealing at 60°C for 30sec with a touchdown method of 1°C per cycle, and extension at 72°C for 40sec; followed again by 35 cycles identical except for the constant annealing temperature of 50°C, and a final extension at 72°C for 7 min.

The results of PCR amplification were visualized under ultraviolet light after performing electrophoresis with TAE buffer, on agarose gel 1.5% containing Ethidium Bromide.

Sequencing

The amplified DNA was purified in an ExoSAP reaction, including 0.75µl deionized distilled H_2O, 0.7µl Antarctic phosphatase, 0.5µl Antarctic phosphatase buffer and 0.05µl exonuclease (Exo1), for 30min at 38°C and 15min at 80°C.

After the purification, 3µl of each samples was sequenced in a mix including 5.1µl of deionized distilled H_2O, 1.75µl of buffer 5X, 0.75µl of ABI BigDye™ Terminator v3.1 (Applied Biosystems®) and 1.6µl of one primer diluted ten times (1pM). Due to the length of the CO1 gene fragment, it was sequenced from both directions (using both primers). POL and IDH genes were sequenced with the forward primer and the CAD gene using the reverse primer.

The thermal cycling conditions of the sequencing reaction were: 96°C for 10sec, followed by 25 cycles with 96°C for 10sec, 50°C for 5sec, and 60°C for 2min. After Ethanol precipitation and dissolution in Hi-Di™ Formamide, the products were sequenced on ABI PRISM™ 3100 Genetic Analyzer (Applied Biosystems®).

Data Analysis

Sequences were edited using BioEdit Sequence Alignment Editor program 7.0.9.0. (Hall 1999). Alignments were performed either manually or by using ClustalW in BioEdit with default parameters (Larkin *et al.* 2007). The four genes sequenced in this study are protein coding, therefore the nucleotide sequences were relatively well conserved which facilitated the alignment process.

For the phylogenetic analysis of the gene, members of the *Limnephilus* genus were used as an outgroup. *Limnephilus griseus* was used for CAD and CO1, and two samples of *Limnephilus centralis* from GenBank (accession number: FN600923, FN601233) were used for POL and IDH.

Relationships between haplotypes found among the samples were inferred and visualized using the medium-joining algorithm in Network Software 4.6.1.0 (Fluxus Technology Ltd 1999-2012) to construct networks. Phylogenetic trees of each gene were constructed using the R packages APE: Analyses of Phylogenetics and Evolution (Paradis *et al.* 2004) and phangorn (Schliep 2011) (R software 2.15.0, R Development Core Team, 2012) and using Mega5 (Tamura *et al.* 2011). The phylogenic tree that fits best the data for each gene, based on the maximum likelihood, were obtained with PhyML (Guindon & Gascuel 2003), using the Akaike Information Criterion (AIC) (Akaike 1974; Posada & Buckley 2004). Optimal trees were chosen by determining the evolutionary model with the lowest AIC score and were rooted with the outgroup individual, specific to each gene. Phylogeny trees were plotted using APE in R and FigTree (v1.3.1). The support for nodes in the trees was estimated by 1000 bootstrap replicates.

Mitochondrial DNA is normally solely maternally inherited, thus most individuals have only a single variant of the mitochondrial genes. However, individuals can be either homozygotes or heterozygote for nuclear genes, therefore the sequences from heterozygotes samples can show several double peaks in electropherograms. It can thus be challenging to determine the correct sequence of the paternal and maternal haplotypes when there are several heterozygous sites. In this study, the SeqPHASE program (Flot 2010) was used to determine the most likely haplotypes for heterozygotic individuals.

Genetic diversity was summarized with standard indices including number of segregating sites (S), number of different haplotypes (k), haplotype diversity (Hd), nucleotide diversity (π), and their corresponding standard deviations calculated according to Nei (1987), as implemented in the software Arlequin 3.5.1.2 (Excoffier & Lischer 2010) and DnaSP 5.10.01 (Librado & Rozas 2009). The haplotype diversity (Hd) is the probability that two haplotypes randomly selected from the same population or same group, will be different (Hartl & Clark 1997), which informs on the numbers and frequencies of different alleles at

a locus, regardless of their evolutionary relationships. The nucleotide diversity (π) represents the probability that a particular nucleotide site, randomly chosen, will be different between two individuals which correspond to a weighted sequence divergence between two individual sequences (Hartl & Clark 1997); and are, unlike Hd, independent of the length of the sequence. Overall, higher genetic diversity implies either an older or a larger population.

The proportion of synonymous and non-synonymous substitutions in the protein-coding genes was obtained with DnaSP (Librado & Rozas 2009). Nucleotide sequences were translated into amino acid sequences using the online tool for sequences conversion, with the appropriate settings for mtDNA of invertebrates for CO1 sequences, implemented by in-silico website (Joosse & Hannemann, in-silico 2006-2012).

Historical demographies were estimated with the Tajima's D test (Tajima 1989) and the Fu's Fs test (Fu 1997). In the neutral model, a population in demographic expansion produces a statistically significant large negative value of Fs, which tends to reflect an excess of rare alleles (recent mutations) and constitutes evidence against the neutrality of mutations. A Fs statistic value should be considered as significant at the 5% level, only if its p-value is below 0.02, and not below 0.05 (Fu 1997). Unlike the mtDNA, recombination may occur in the nuclear DNA, and in this case the actual value of Tajima's D takes more extreme values. Therefore, this statistical method might be conservative when nuclear DNA is analyzed. Significantly negative Tajima's D values can be due to directional selective effects, population expansion, bottleneck or heterogeneity of mutation rate, whereas a positive value can indicate balancing selection or result from admixture of divergent lineages.

The genetic differentiation among samples from different geographic origins was assessed by comparing corrected average pairwise differences ($D_A = Pixy - (Pix + Piy)/2$); Zvuloni *et al.* 2008, Nei & Li 1979), which gives the average percentage difference in nucleotide sites between two populations x and y, after taking the average differences within each population into account. To examine the partitioning of observed genetic variation and characterize the structure among and within populations, a standard analysis of molecular

variance (AMOVA) (Excoffier *et al.* 1992) based on pairwise differences (ϕ_{ST}) among haplotypes and on haplotype frequencies (F_{ST}), was computed using Arlequin.

To determine if geographic isolation of populations is partly responsible for the genetic structuration, Mantel tests with 1000 permutations (Mantel 1967) were performed on the matrices of genetic distances (D_A) and geographic distances between populations. The matrix of geographic distances was calculated from Euclidean distances between populations using coordinates of latitude and longitude. In addition, Neighbor-joining (NJ) trees were constructed using Mega5 for each gene, based on nucleotide divergence between populations (D_A), with negative divergence set to zero.

Results

In total, 237 sequences of *Apatania zonella* and 18 sequences of related *Apatania species* were obtained from four different protein-coding genes in this study. The main characteristics of the sequences for each marker are summarized in Table 2.

Table 2. Summary of the main characteristics of the sequences obtained from the samples of *A. zonella*, with a comparison between the different genes used in this study. Main haplotypes: percentage of haplotypes represented by more than two individuals. Samples in main haplotypes: percentage of the overall individuals included in main haplotypes.

Gene	CO1a	CO1	CAD	POL	IDH
Base pairs nb	636	1190	836	766	630
Sequences nb	69	39	62	64	42
Nb of haplotypes	21	27	32	24	16
Nb of polymorphic sites	19	38	14	23	22
Main haplotypes	33%	23%	19%	21%	25%
Samples in main haplotypes	78%	23%	40%	48%	57%

Considering nuclear markers, 40 to 57% of the sequences are shared by more than two individuals (Table 2). The corresponding ratio is lower for CO1 (23%) (Figure 3). CO1a, which length is slightly more than half of the CO1, has 78% of its sequences in main haplotypes, despite a higher number of sequences and a much more diverse sample origins. This indicates that the first half of the CO1 gene is more conserved.

CO1 mitochondrial gene overview

Twenty seven individuals were successfully sequenced for the CO1 gene in this study. By adding sequences from collaborators, a total of sixty-nine sequences of *A.zonella* and eighteen sequences of closely related *Apatania* species were obtained (Salokannel *et al.* 2010 and unpublished: Kjer *et al.* 2002, Prieto 2009, Sanz Muñoz 2010). The eighteen sequences obtained from Salokannel *et al.* (2010) and Kjer *et al. (unpublished)* had been sequenced with the same forward primer but a different reverse primer (HCO2198, Folmer *et al.* 1994) resulting in shorter mtDNA sequences. Therefore, 30 sequences were 636bp long (corresponding to base pairs number 1919 to 2555 in *Drosophila melanogaster* genome) and 39 were 1190bp long (corresponding to base pairs number 1919 to 3108). The gene was thus analyzed in two steps: The first half of the CO1 gene with all the 69 sequences: CO1a (636bp) and the entire sequence for 39 of the samples: CO1 (1190bp).

CO1a mitochondrial gene

The 69 sequences of CO1a (636bp) define 21 haplotypes and 19 polymorphic sites (3%). There are seven main haplotypes (found in more than 2 individuals) which represent 33% of the haplotypes and contain 78% of the sequences (Table 2). The global relative nucleotides composition is **C:** 18.2%, **T**: 36.3%, **A**: 31.2%, **G**: 14.3% and shows an A-T bias. All the mutations observed are synonymous.

Based on Akaike Information Criterion (AIC) (Akaike 1974, Posada & Buckley 2004), the best fit evolutionary model to build a tree from the CO1a sequences is GTR+I+G (Lanave *et al.* 1984). The phylogenetic tree including seven closely related *Apatania* species (Figure 2) shows that *A. stigmatella*, *A. wallengreni* and *A. muliebris* are well separated from *A. zonella*. The phylogenetic status of the more closely related species *A. hispida*, *A. zonella*, *A. auricular*, *A. dalecarlica* and *A. forsslundi* are less resolved and indicate that the group is not monophyletic as there are three *Apatania* species nested within *A. zonella*.

The phylogenetic tree of individuals from *A. zonella* species only, shows the genetical links between individuals in more details (Figure 3). The two phylogenic trees show mainly two clades within *A. zonella*, one Eurasian and one North-American (Figure 2, 3 and 4). The

first one includes Scandinavian individuals (Finnish and Norwegian), two individuals from Canada and one from Mongolia on one branch; and Icelandic individuals on the other branch. The second clade includes North-American individuals (United States, Alaska and Canada), all Greenlandic and part of the Icelandic ones. Interestingly, Icelandic individuals are found in both clades.

Different genetic diversity indices, within each location and for all samples, are summarized in Table 3. The CO1a nucleotide diversity at each location is uncorrelated with the sample size (r=0.30, p=0.47). Excluding the location represented by a single individual, most populations present low nucleotide diversity (Table 3). The relatively small sample from one area in the United States does not show any nucleotide or haplotype diversity as it is based on only a single haplotype. The highest pairwise differences between haplotypes are observed in Icelandic and Norwegian populations, which mean they include individuals that are very divergent genetically. The haplotype diversity is lower in Canada, Greenland and Finland, but samples are relatively small. Iceland shows a positive and very significant Tajima's D caused by the presence of very different lineages. The power of the Tajima's D neutrality test is though weak when the size of the sample is small. Other populations seem to follow the neutral mutation hypothesis (Table 3).

The corrected average pairwise differences between populations ranges from 0 (Mongolia-Norway) to 9.631 (Greenland-Finland) (Table 4). Alaska and Mongolia did not give any significant result with other populations due to the fact that these populations are represented by a single individual. Norway and Finland do not show significant genetic differentiation, mainly as the variation within Norway is large.

The correlation between genetic distances and geographical distance between populations is significant (r=0.47, p-value=0.021). The analysis of molecular variance based on pairwise differences shows that 43.8% of the variation is observed among population, and the other 56.2% is explained by genetic variation within populations (Table 5), meaning that both levels are important to explain the overall variation.

Neighbor-joining (NJ) tree for the CO1a, summarizes the relationship between localities, based on nucleotide divergence between populations (D_A) (Figure 4). According to the NJ tree, the populations in Europe and Mongolia originated from same branch, while North-

American and Greenlandic populations originated from another branch, and Icelandic individuals from South-West Iceland are genetically very close to individuals from Greenland and United-States, while Icelandic individuals from North-West and South Iceland are close to each other and somewhat genetically centered between European and North-American individuals. The position of the individuals from North-West Iceland and South Iceland could be explained by the admixture of haplotypes in regions with both North-American and European origin.

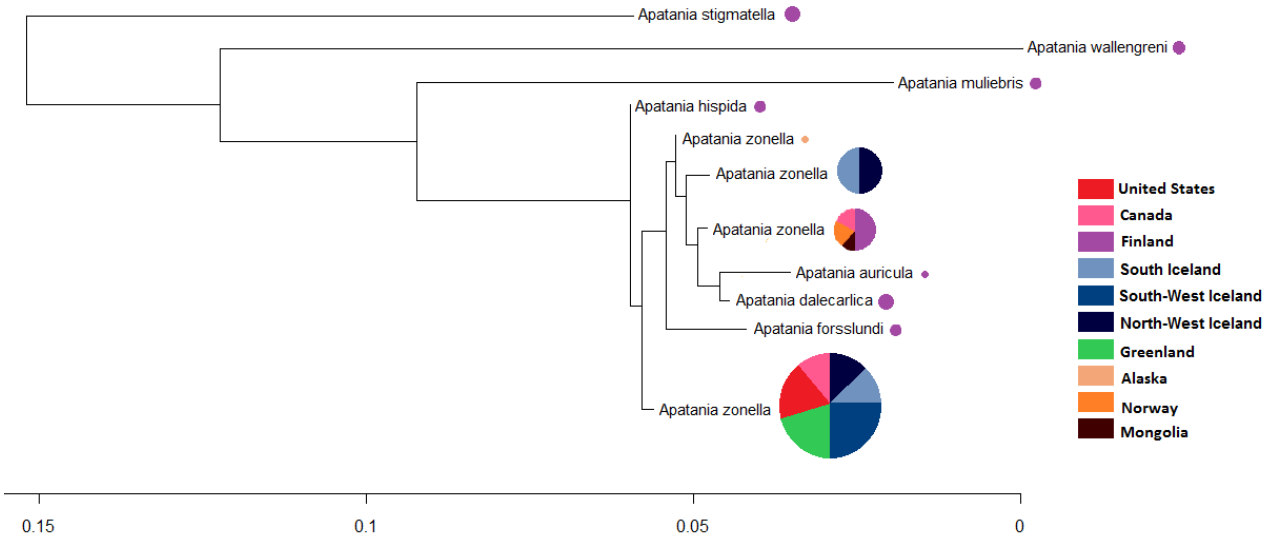

Fig 2. Phylogenetic tree of several closely related *Apatania* species from Finland and *A. zonella*, based on CO1a sequences, performed using the GTR+I+G model (best fit model based on AIC) (Lanave *et al.* 1984) using the Phangorn R package (*Limnephilus griseus* outgroup is not shown). Each circle represents the individuals included in the corresponding branch of the tree, and their surface is proportional to the number of individuals. Scale unit of branches length is estimated nucleotide substitutions number per hundred sites. The different colors represent the geographic origins.

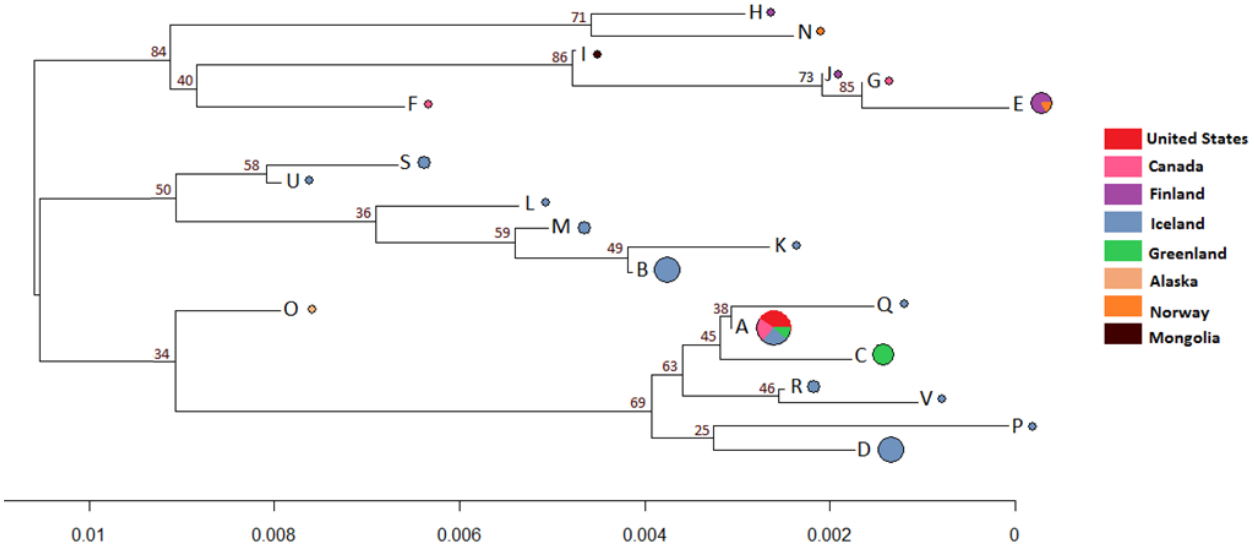

Fig 3. Phylogenetic tree of *A. zonella* based on CO1a sequences, performed using the GTR+I+G model (best fit model based on AIC) (Lanave *et al.* 1984) with Phangorn on R (*Limnephilus griseus* outgroup is not shown). Numbers on nodes show bootstrap values. Letters correspond to the different haplotypes, each circle represents one haplotype and their surface is proportional to the relative frequencies. Scale unit of branches length is estimated nucleotide substitutions number per hundred sites. The different colors represent the geographic origin.

Table 3. Genetic diversity of *A. zonella* within CO1a, by sampling area and for all samples combined. The following information is presented: **N** number of sequences, **S** number of segregating sites, **k** number of different haplotypes, **π** nucleotide diversity, **SD** standard deviation, **Hd** haplotype diversity, **p** mean pairwise differences, and corresponding values to Tajima's D and Fu's *Fs* Tests and significance level.

Population	N	S	k	π (x100) ± SD	Hd ± SD	Tajima's D	Fu's *Fs*
United States	7	0	1	---	---	---	---
Canada	6	11	3	0.77 ± 0.50	0.60 ± 0.22	0.062	2.895
Greenland	8	2	3	0.07 ± 0.08	0.43 ± 0.17	-0.448	0.536
Alaska	1	1	1	---	---	---	---
Iceland	37	12	13	0.90 ± 0.49	0.84 ± 0.04	2.753 **	0.538
Finland	7	9	3	0.43 ± 0.30	0.52 ± 0.21	-1.319	2.167
Norway	2	9	2	1.42 ± 1.49	1.00 ± 0.50	---	---
Mongolia	1	1	1	---	---	---	---
All samples	69	19	22	0.97 ± 0.51	0.90 ± 0.02	1.808	-2.487

Signif. codes: <0.001 '***' 0.01 '**' 0.05 '*' >0.05 ' '

Table 4. Matrix of D_A distances, corrected average pairwise differences between populations, and p-values with statistical significance level corrected for multiple tests (1023 permutations), for CO1a in *A. zonella*.

	United-States	Canada	Greenland	Iceland	Finland	Norway	Alaska
Canada	0.400	---					
Greenland	0.536 *	0.936 **	---				
Iceland	2.388 **	0.985	2.924 **	---			
Finland	9.095 ***	4.400 *	9.631 ***	6.185 ***	---		
Norway	6.000 *	1.900	6.536 *	2.934 *	-1.333	---	
Alaska	4.000	2.067	4.536	2.691	6.810	3.000	---
Mongolia	8.000	4.067	8.536	5.267	1.952	0.000	6.000

Signif. codes: <0.001 '***' 0.01 '**' 0.05 '*' >0.05 ' '

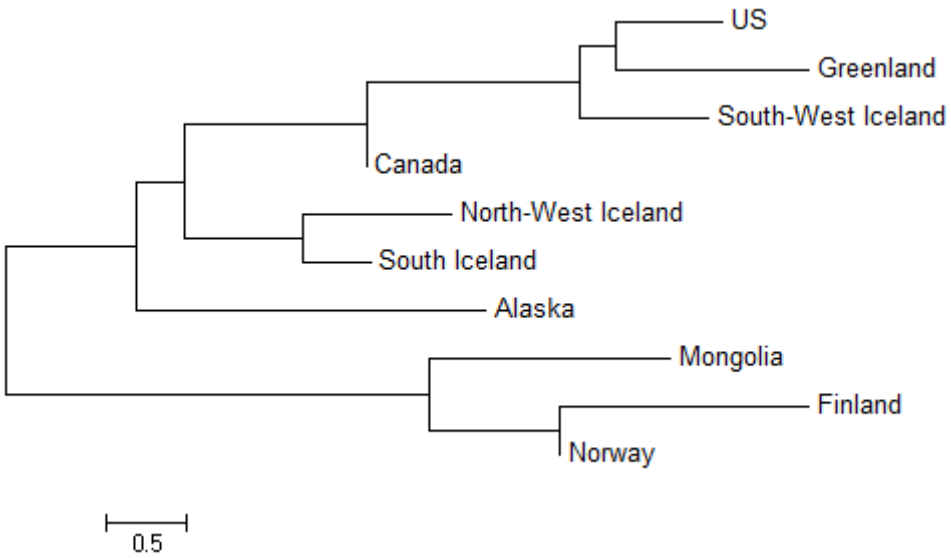

Fig 4. Neighbor-joining (NJ) tree of localities, constructed on Mega5, for CO1a, based on nucleotide divergence between populations (D_A), with negative divergence set to zero.

Table 5. Analysis of molecular variance (AMOVA), of *A. zonella* populations studied, based on CO1a mtDNA variation, and Fixation Index with its significance level.

Source of Variance	d.f.	Sum of Squares	Variance Components	Percentage of Variation
Among Populations	7	87.315	1.64818 Va	43.81 ***
Within Populations	56	118.357	2.11352 Vb	56.19
Total	63	205.672	3.76170	
Fixation Index FST : 0.43815 *				

Signif. codes: <0.001 '***' 0.01 '**' 0.05 '*' >0.05 ' '

CO1 mitochondrial gene

The 39 sequences of CO1 (1190bp) revealed 27 haplotypes and 38 polymorphic sites (3.2%). There are three main haplotypes (more than 2 individuals) and all together they represent 23% of haplotypes and contain 23% of the samples (Table 2). The global relative nucleotide composition is **C**: 16.3%, **T**: 37.8%, **A**: 31.4%**, G**: 14.5% and shows an A-T bias. In a total of 38 mutations, 34 are synonymous and 4 are non-synonymous, which represents 1% of the amino acids.

This sequence analysis focuses on Iceland, Greenland and Alaska but it allows studying in details the links between different populations within Iceland.

Based on Akaike Information Criterion (AIC) (Akaike 1974, Posada & Buckley 2004), the best fit model to build a tree from the CO1 sequences is GTR+I+G (Lanave *et al.* 1984).

The phylogenetic tree shows distinctly two main clades (Figure 5), clade A with only individuals from South Iceland and North-West Iceland, and clade B from all locations. Therefore, Icelandic individuals in each location come from the two different clades and they can be more closely related to the individuals in other locations than the individuals within the same population. In clade A, the individuals from North-West Iceland are genetically similar to individuals from South of Iceland, and thus they seem to share a more recent common ancestor than South-West of Iceland.

The shared individuals between CO1a and CO1 show a neat correspondence. The second clade of CO1a phylogenetic tree (Figure 3) corresponds to the second clade of the CO1 tree (Figure 5), and the first clade of CO1a tree matches the first clade of CO1 tree. However, one Icelandic individual from North-West Iceland however is an exception to this trend.

Different genetic diversity indices, within each location and for all samples, are summarized in Table 6. The CO1 nucleotide diversity at each location is uncorrelated with the sample size ($r=0.75$, $p=0.14$). Low nucleotide diversity is observed in South-West Iceland and Greenland, which are only located in one of the clades. High pairwise difference between haplotypes is observed in all populations (except Alaska only represented by one individual) which means samples are highly divergent within populations (Table 6). South Iceland shows a positive and very significant Tajima's D,

which is here again due to the presence of very different haplotypes in the population. Other populations seem to follow the neutral mutation hypothesis.

The corrected average pairwise differences between populations ranges from -0.075 (NWI-SI) to 12.302 (Greenland-NWI). Greenland and SWI show significant results with all other populations except with Alaska (Table 7).

The Mantel test, analyzing the correlation between the matrix of genetic distances and matrix of geographical distance between populations, is insignificant for D_A matrix (r=0.11, p-value=0.308). The analysis of molecular variance based on pairwise differences shows that 37.3% of the variation is observed among population, and the other 62.7% is explained by genetic variation within populations (Table 8). Therefore, almost two-thirds of the overall variation is due to genetic variability inside each population, nonetheless the variation among population is still substantial. This result indicates a rather weak geographical structure for this gene in *Apatania zonella*.

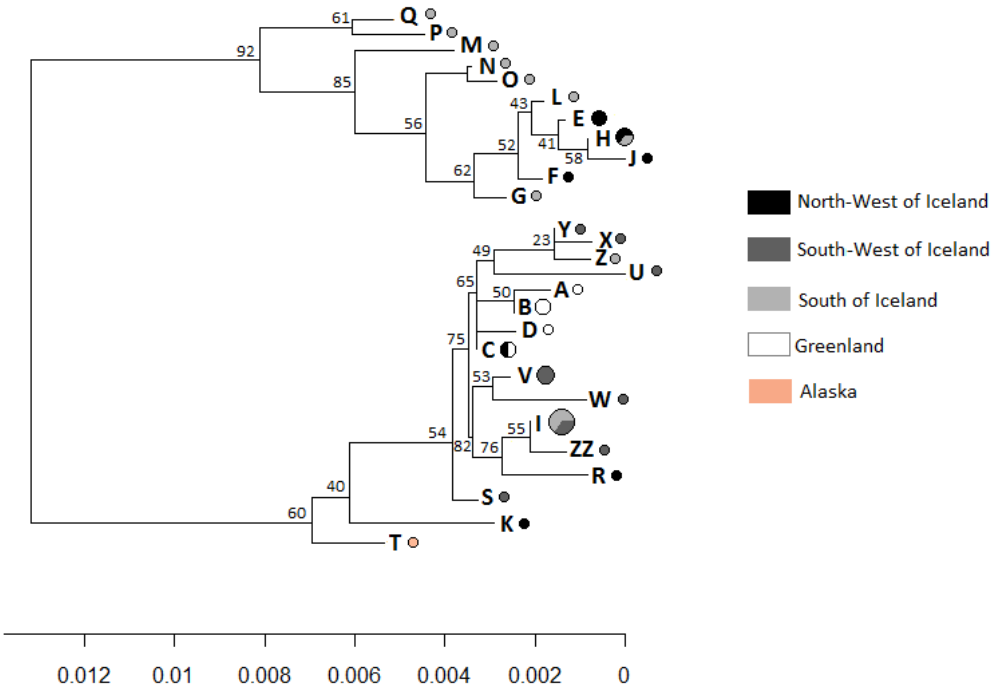

Fig 4. Phylogenetic tree of *A. zonella* based on CO1 sequences, performed using the GTR+I+G model (best fit model based on AIC) (Lanave *et al.* 1984) using the Phangorn R package (outgroup is not shown). Numbers on nodes show bootstrap values. Numbers on nodes show bootstrap values. Letters correspond to the different haplotypes, each circle represents one haplotype and their surface is proportional to the relative frequencies. Scale unit of branches length is estimated nucleotide substitutions number per hundred sites. The different colors represent the geographic origin.

Table 6. Genetic diversity of *A. zonella* within CO1 gene, by sampling area and for all samples combined. The following information is presented: **N** number of sequences, **S** number of segregating sites, **k** number of different haplotypes, **π** nucleotide diversity, **SD** standard deviation, **Hd** haplotype diversity, **p** mean pairwise differences, and corresponding values to Tajima's D and Fu's *Fs* Tests and significance level.

Population	N	S	k	π (x100) ± SD	Hd ± SD	Tajima's D	Fu's *Fs*
Greenland	**5**	**3**	**4**	**0.11 ± 0.09**	**0.87 ± 0.13**	**-0.185**	**-1.350**
Alaska	1	0	1	---	---	---	---
North-West of Iceland	9	30	7	1.17 ± 0.66	0.94 ± 0.07	1.321	0.974
South-West of Iceland	11	13	8	0.31 ± 0.19	0.93 ± 0.07	-0.719	-2.318
South of Iceland	13	32	10	1.32 ± 0.71	0.92 ± 0.07	2.296 *	0.281
Iceland	**33**	**38**	**23**	**1.29 ± 0.66**	**0.96 ± 0.02**	**2.289 ***	**-3.142**
All samples	39	38	27	1.22 ± 0.54	0.97 ± 0.03	1.998	-4.880

Signif. codes: <0.001 '***' 0.01 '**' 0.05 '*' >0.05 ' '

Table 7. Above the diagonal: matrix of D_A distances, corrected average pairwise differences between populations, and p-values with statistical significance level corrected for multiple tests (1023 permutations), for CO1 in *A. zonella*.

	Greenland	Alaska	N-W Iceland	South Iceland
Greenland	---			
Alaska	5.867	---		
North-West Iceland	12.302 **	9.917	---	
South Iceland	7.918 **	7.615	-0.075	---
South-West Iceland	0.527 **	6.327	11.719 ***	6.992 **

Signif. codes: <0.001 '***' 0.01 '**' 0.05 '*' >0.05 ' '

Table 8. Analysis of molecular variance (AMOVA), of *A. zonella* populations studied, based on CO1 mtDNA variation, and Fixation Index with its significance level.

Source of Variance	d.f.	Sum of Squares	Variance Components	Percentage of Variation
Among Populations	4	106.606	2.91912 Va	37.31 **
Within Populations	35	171.644	4.90411 Vb	62.69
Total	39	278.250	7.82323	
**Fixation Index FST : 0.37314 **				

Signif. codes: <0.001 '***' 0.01 '**' 0.05 '*' >0.05 ' '

CAD nuclear gene

Sequences of 836bp were obtained for the nuclear gene CAD from 31 individuals, among which 8 were heterozygotes. All the Icelandic and Finnish individuals, and also one Greenlandic individual, were homozygotes. Seven Greenlandic individuals and the single one from Alaska were heterozygotes. SeqPHASE could resolve all the haplotypes for the heterozygote individuals.

These sequences revealed 14 polymorphic sites (1.7%) and 32 haplotypes. There are six main haplotypes (found in more than 2 individuals) and all together they represent 19% of haplotypes and contain 40% of the samples (Table 2). The global relative nucleotide composition is **C**: 24.6%, **T**: 28.8%, **A**: 27.9%, **G**: 18.7% and shows an A-T bias. For a total of 14 mutations, 12 are synonymous and 2 are non-synonymous, which represents 0.7% of the amino acid.

The best fit model to build a tree from the Cad sequences is GTR+I+G (Lanave *et al.* 1984), with Mega5 (Tamura *et al.* 2011).

The phylogenetic tree does not show a very clear genetic structuration for this marker (Figure 6). Nevertheless, all the individuals from South-West Iceland (which correspond to the area near Reykjavík) are gathered in one clade of the tree, except for one individual genetically much closer to the population in South Iceland. Populations from North-West and South Iceland present closely related haplotypes which are divided in two main sub-clades of the trees. However, these two sub-clades are closely linked together. The Finnish individuals have all the same haplotype and the heterozygote individual from Alaska has very closely related sequences. On the contrary, the Greenlandic sequences are scattered in the genetical tree, most likely meaning several colonization events at different times.

All the Icelandic individuals are homozygous and in South-West Iceland each individual has a unique haplotype which supports asexual reproduction in this area. However, the Greenlandic individuals are mostly heterozygote and four of them have genetically distant haplotypes.

Different genetic diversity indices, within each location and for all samples, are summarized in Table 9. The Cad nucleotide diversity at each location is positively

correlated to the sample size (r=0.92, p=0.009). All populations show a very low nucleotide diversity except Iceland. The haplotype diversity is high in all populations except in North-West Iceland which has a medium Hd and Finland which presents only one haplotype (Table 9). Greenland presents a very significant Fu's *Fs* negative value indicating that the population is in expansion. Other populations seem to follow the neutral mutation hypothesis.

The corrected average pairwise differences between populations ranges from 0.596 (Greenland-NWI) to 3.662 (Greenland-Finland). All results are significant except for Alaska-Greenland and Alaska-Finland (Table 10).

The Mantel test is insignificant for the D_A matrix (r=-0.05, p-value=0.50). The analysis of molecular variance based on pairwise differences shows that 34.4% of the variation is observed among population, and the other 65.6% is explained by genetic variation within populations (Table 11). Despite a significant contribution of the variation among population, two-third of the overall variation is explained by the genetic variability inside each population. This result indicates a rather weak geographical structure for this gene in *A. zonella*.

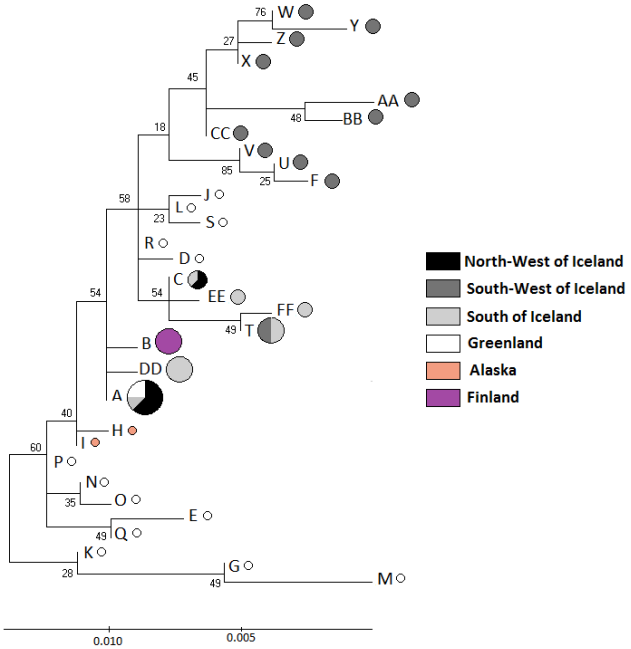

Fig 5. Phylogenetic tree of *A. zonella* based on CAD nuclear sequences, performed using the GTR+I+G model (best fit model based on AIC) (Lanave *et al.* 1984) on Mega5 (outgroup is not shown). Numbers on nodes show bootstrap values. Letters correspond to the different haplotypes, each circle represents one haplotype and their surface is proportional to the relative frequencies. Scale unit of branches length is estimated nucleotide substitutions number per hundred sites. The different colors represent the geographic origin.

Table 9. Genetic diversity of *A. zonella* within CAD gene, by sampling area and for all samples combined. The following information is presented: **N** number of sequences, **S** number of segregating sites, **k** number of different haplotypes, **π** nucleotide diversity, **SD** standard deviation, **Hd** haplotype diversity, **p** mean pairwise differences, and corresponding values to Tajima's D and Fu's *Fs* Tests and significance level.

Population	N	S	k	π (x100) ± SD	Hd ± SD	Tajima's D	Fu's *Fs*
Greenland	**16**	**7**	**14**	**0.37 ± 0.23**	**0.98 ± 0.03**	**0.736**	**-5.054 ***
Alaska	**2**	**1**	**2**	**0.12 ± 0.17**	---	---	---
North-West of Iceland	6	1	2	0.06 ± 0.07	0.53 ± 0.17	0.851	0.625
South-West of Iceland	22	8	8	0.41 ± 0.25	0.95 ± 0.02	1.498	-2.236
South of Iceland	12	3	4	0.29 ± 0.19	0.85 ± 0.07	1.914	0.242
Iceland	**40**	**8**	**11**	**0.51 ± 0.29**	**0.96 ± 0.01**	**1.903**	**-4.677**
Finland	**4**	**1**	**1**	---	---	---	---
All samples	62	14	32	0.36 ± 0.22	0.92 ± 0.02	1.357	**-14.080 ***

Signif. codes: <0.001 '***' 0.01 '**' 0.05 '*' >0.05 ' '

Table 10. Matrix of D_A distances, corrected average pairwise differences between populations, and p-values with statistical significance level corrected for multiple tests (1023 permutations), for CAD in *A. zonella*.

	Greenland	Alaska	N-W Iceland	S-W Iceland	South Iceland
Greenland	---				
Alaska	0.977	---			
North-West Iceland	0.596 *	0.900 *			
South-West Iceland	1.347 ***	2.481 *	2.365 ***	---	
South Iceland	1.005 ***	2.288 *	1.077 **	2.299 ***	---
Finland	1.571 ***	1.000	0.733 **	3.662 ***	2.121 ***

Signif. codes: <0.001 '***' 0.01 '**' 0.05 '*' >0.05 ' '

Table 11. Analysis of molecular variance (AMOVA), of *Apatania zonella* populations studied, based on CAD DNA variation, and Fixation Index with its significance level.

Source of Variance	d.f.	Sum of Squares	Variance Components	Percentage of Variation
Among Populations	5	44.082	0.78878 Va	35.59 ***
Within Populations	56	79.934	1.42739 Vb	64.41
Total	61	124.016	2.21617	
Fixation Index FST : 0.35592 *				

Signif. codes: <0.001 '***' 0.01 '**' 0.05 '*' >0.05 ' '

POL nuclear gene

Sequences of 766bp were obtained from 32 individuals for the POL gene, four were heterozygotes: 2 Icelandic and 2 Greenlandic. The sequences gave a total of 24 haplotypes and 23 polymorphic sites (3%). There are five main haplotypes (with more than 2 individuals) and all together they represent 21% of the haplotypes and contain 48% of the samples (Table 2). The global relative nucleotide composition is **C**: 22%, **T**: 25.2%, **A**: 28.8%**, G**: 24%. All mutations are synonymous.

Based on the Akaike Information Criterion (AIC) (Akaike 1974, Posada & Buckley 2004), the best fit model to build a tree from the Pol sequences is HKY+I+G (Hasegawa *et al.* 1985).

The Phylogenetic tree presents three main clades (Figure 7). The first clade includes only Icelandic individuals, with all the individuals from North-West Iceland and South-Iceland. The second clade encompasses mainly sequences from Greenlandic individuals, one individual from Finland and one from South Iceland. The last clade includes the rest of the sequences originated from Greenland and South-West Iceland, as well as one individual from Finland.

The two Greenlandic heterozygotes present genetically closely related sequences on same clade and the two Icelandic heterozygotes are on the same clade as well but with much higher genetical distances. However, the two haplotypes of these individuals are similar between these two individuals (Figure 7).

Different genetic diversity indices, within each location and for all samples, are summarized in Table 12. The Pol nucleotide diversity at each location is uncorrelated with the sample size ($r=0.15$, $p=0.80$). The nucleotide diversity is relatively high in all population except in North-West Iceland. High pairwise difference between haplotypes is also observed in all locations, albeit slightly lower for South-West Iceland, which means individuals are highly genetically divergent within each location (Table 12). Greenland and the overall population of Iceland show a positive and very significant Tajima's D indicating a recent diminution of these populations size. Other populations seem to follow the neutral mutation hypothesis.

The corrected average pairwise differences between populations ranges from -0.114 (Greenland-Finland) to 6.820 (NWI-SWI). All results are significant except for Greenland-Finland (Table 13).

The Mantel test is insignificant for the D_A matrix (r=-0.24, p-value=0.749). The analysis of molecular variance based on pairwise differences shows that 44.2% of the variation is observed among population, and the other 55.8% is explained by genetic variation within populations (Table 14), meaning that both levels are important to explain the overall variation.

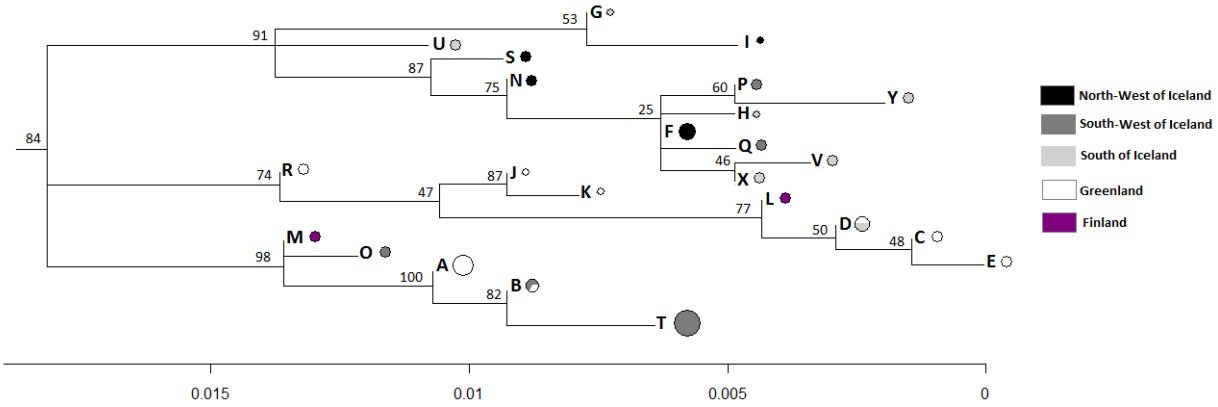

Fig 7. Phylogenetic tree of *A. zonella* based on POL nuclear sequences, performed using the HKY+I+G model (best fit model based on AIC) (Hasegawa *et al.* 1985) using the Phangorn R package (*Limnephilus griseus* outgroup is not shown). Numbers on nodes show bootstrap values. Letters correspond to the different haplotypes, each circle represents one haplotype and their surface is proportional to the relative frequencies. Scale unit of branches length is estimated nucleotide substitutions number per hundred sites. The different colors represent the geographic origin.

Table 12. Genetic diversity of *A. zonella* within POL gene, by sampling area and for all samples combined. The following information is presented: **N** number of sequences, **S** number of segregating sites, **k** number of different haplotypes, **π** nucleotide diversity, **SD** standard deviation, **Hd** haplotype diversity, **p** mean pairwise differences, and corresponding values to Tajima's D and Fu's *Fs* Tests and significance level.

Population	N	S	k	π (x100) ± SD	Hd ± SD	Tajima's D	Fu's *Fs*
Greenland	**16**	**15**	**8**	**0.92 ± 0.51**	**0.91 ± 0.05**	**2.090 ***	**0.805**
North-West of Iceland	10	8	4	0.34 ± 0.22	0.73 ± 0.12	-0.348	1.197
South-West of Iceland	22	16	5	0.78 ± 0.44	0.63 ± 0.11	1.228	4.904
South of Iceland	12	16	7	0.70 ± 0.41	0.92 ± 0.05	0.078	0.176
Iceland	**44**	**23**	**17**	**1.10 ± 0.58**	**0.91 ± 0.034**	**2.174 ***	**-0.004**
Finland	**4**	**10**	**2**	**0.87 ± 0.62**	**0.67 ± 0.20**	**2.223**	**4.664**
All samples	64	23	24	1.17 ± 0.16	0.94 ± 0.02	2.666 *	-2.377

Signif. codes: <0.001 '***' 0.01 '**' 0.05 '*' >0.05 ' '

Table 13. Above the diagonal: matrix of D_A distances, corrected average pairwise differences between populations, and p-values with statistical significance level corrected for multiple tests (1023 permutations), for POL in *A. zonella*.

	Greenland	N-W Iceland	S-W Iceland	South Iceland
Greenland	---			
North-West Iceland	6.512 ***	---		
South-West Iceland	3.519 ***	6.820 ***	---	
South Iceland	4.266 ***	1.103 ***	6.136 ***	---
Finland	-0.114	4.867 **	2.867 *	2.886 **

Signif. codes: <0.001 '***' 0.01 '**' 0.05 '*' >0.05 ' '

Table 14. Analysis of molecular variance (AMOVA), of *A. zonella* populations studied, based on POL DNA variation, and Fixation Index with its significance level.

Source of Variance	d.f.	Sum of Squares	Variance Components	Percentage of Variation
Among Populations	4	114.728	2.22563 Va	44.17 ***
Within Populations	56	157.567	2.81369 Vb	55.83
Total	60	272.295	5.03932	
Fixation Index FST : 0.44165 * **				

Signif. codes: <0.001 '***' 0.01 '**' 0.05 '*' >0.05 ' '

IDH nuclear gene

The sequencing of IDH gene was in overall less successful than for the other markers. Sequences of 630bp were obtained for the nuclear gene IDH from 21 individuals, among which 3 were heterozygotes: 2 Icelandic samples and the single sample from Alaska. The sequences revealed 16 different haplotypes and 22 polymorphic sites (3.5%), with 3 sites that show deletion in the sample from Alaska and all the samples from Iceland. There are four main haplotypes (more than 2 individuals) and all together they represent 25% of haplotypes and contain 57% of the samples (Table 2). The global relative nucleotide composition is **C**: 22.8%, **T**: 24.4%, **A**: 28.4%, **G**: 24.4%. For a total of 22 mutations, all are synonymous except 4 segregation sites, at the end of the coding region, which correspond to three insertions and one deletion compared to the sequence of *Limnephilus centralis* for this gene. These four segregations sites are all at the end of the coding region and represent 0.6% of the amino acid.

Based on Akaike Information Criterion (AIC) (Akaike 1974, Posada & Buckley 2004), the best fit model to build a tree from the Pol sequences is SYM+I+G (Zharkikh 1994).

The Phylogenetic tree shows two main clades and a separate branch for the individual from Alaska (Figure 8). The first clade includes one individual from Greenland and all Icelandic individuals from the South population except one which has one sequence in each clade. The second clade contains the other Greenlandic individual, all the Icelandic individuals from South-West area, one sequence from South-Iceland and one sequence from North-West Iceland. The other sequence of the individual from North-West Iceland has a more ancestral haplotype. The two mix Icelandic individuals and the individual from Alaska indicate occasional occurrence of sexual reproduction *in A. zonella* as those present admixtures of haplotypes.

Different genetic diversity indices, within each location and for all samples, are summarized in Table 15. The CO1 nucleotide diversity at each location is negatively correlated to the sample size ($r=-0.90$, $p=0.03$). The nucleotide diversity is average in all samples except South-West Iceland where π is low. The haplotype diversity is high in all populations reflecting high genetical differences within populations (Table 15). According to neutrality tests, all populations seem to follow the neutral mutation hypothesis.

The corrected average pairwise differences between populations ranges from -0.011 (Greenland-SI) to 13.210 (Alaska-SWI) (Table 16).

The mantel test is significant for the D_A matrix ($r=0.95$, p-value=0.019). The analysis of molecular variance based on pairwise differences shows that 39.5% of the variation is observed among population, and the other 60.5% is explained by genetic variation within populations (Table 17).

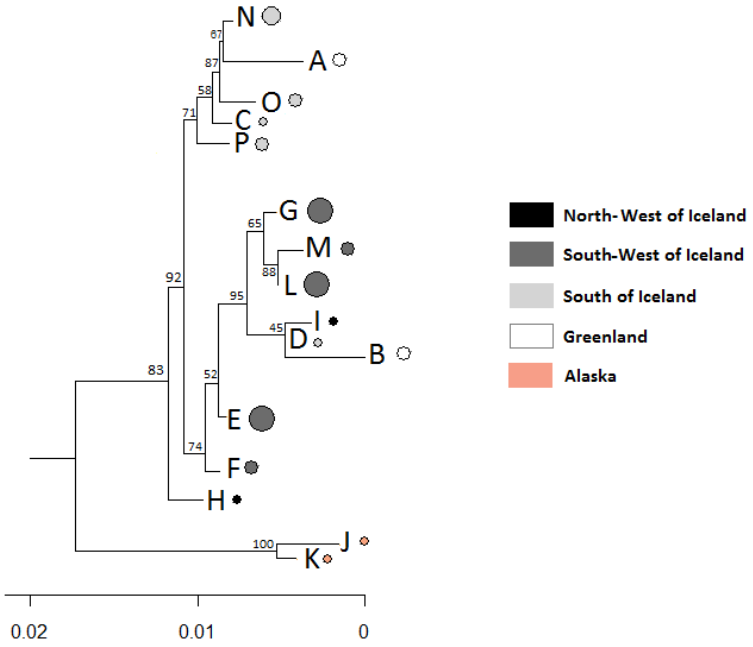

Fig 8. Phylogenetic tree of *A. zonella* based on IDH nuclear sequences, performed using the SYM+I+G model (best fit model based on AIC) (Zharkikh 1994) using the Phangorn R package (outgroup is not shown). Numbers on nodes show bootstrap values. Letters correspond to the different haplotypes, each circle represents one haplotype and their surface is proportional to the relative frequencies. Scale unit of branch length is estimated nucleotide substitutions number per hundred sites. The different colors represent the geographic origin of the samples.

Table 15. Genetic diversity of *A .zonella* within IDH gene, by sampling area and for all samples combined. The following information is presented: **N** number of sequences, **S** number of segregating sites, **k** number of different haplotypes, **π** nucleotide diversity, **SD** standard deviation, **Hd** haplotype diversity, **p** mean pairwise differences, and corresponding values to Tajima's D and Fu's *Fs* Tests and significance level.

Population	N	S	k	π (x100) ± SD	Hd ± SD	Tajima's D	Fu's *Fs*
Greenland	**4**	**6**	**2**	**0.64 ± 0.48**	**0.67 ± 0.20**	**2.156**	**3.526**
Alaska	**2**	**3**	**2**	**0.48 ± 0.55**	**1.00 ± 0.50**	---	---
North-West of Iceland	2	4	2	0.64 ± 0.71	1.00 ± 0.50	---	---
South-West of Iceland	24	4	5	0.20 ± 0.15	0.79 ± 0.04	0.447	-0.316
South of Iceland	10	7	5	0.39 ± 0.26	0.82 ± 0.10	-0.089	-0.167
Iceland	**36**	**10**	**12**	**0.52 ± 0.25**	**0.90 ± 0.03**	**1.480**	**-2.442**
All samples	42	22	16	0.61 ± 0.58	0.90 ± 0.02	-0.453	-1.764

Signif. codes: <0.001 '***' 0.01 '**' 0.05 '*' >0.05 ' '

Table 16. Above the diagonal: matrix of D_A distances, corrected average pairwise differences between populations, and p-values with statistical significance level corrected for multiple tests (1023 permutations), for IDH in *A.zonella*.

	Greenland	Alaska	N-W Iceland	S-W Iceland
Greenland	---			
Alaska	11.500	---		
North-West Iceland	-1.000	12.000	---	
South-West Iceland	0.377 *	13.210 **	0.043	---
South Iceland	-0.011	10.789 ***	0.489	2.166 ***

Signif. codes: <0.001 '***' 0.01 '**' 0.05 '*' >0.05 ' '

Table 17. Analysis of molecular variance (AMOVA), of *A. zonella* populations studied, based on IDH DNA variation, and Fixation Index with its significance level.

Source of Variance	d.f.	Sum of Squares	Variance Components	Percentage of Variation
Among Populations	4	43.219	1.55779 Va	62.40
Within Populations	37	34.733	0.93874 Vb	37.60
Total	41	77.952	2.49653	
Fixation Index FST : 0.62398 *				

Signif. codes: <0.001 '***' 0.01 '**' 0.05 '*' >0.05 ' '

Discussion

The main finding of this study is that *Apatania zonella* in Iceland originates from more than one colonization event, with at least one from North-America through Greenland and one from Europe, most likely from Scandinavia or British islands and Ireland. The multidirectional colonization is reflected by high genetic diversity in all markers used in this study, mainly echoed by the high nucleotide diversity in Iceland which is comparable to the overall nucleotide diversity of all the sequences. Studies on high latitude species living in areas previously affected by Quaternary glaciations often reveal moderate genetic diversity, ambiguous clades division and superficial phylogenic trees (Hewitt 2004; Kornobis *et al.* 2010). These characteristic elements support a scenario of recolonization from several directions or in several steps.

Another interesting result from this study is the lack of monophyly of *A. zonella*. The distinct lineages in North-America and Europe are genetically more different than different *Apatania* species within Europe. Indeed three *Apatania* species: *A. auricular*, *A. dalecarlica* and *A. forsslundi*, are nested within *A. zonella* and appear, from the phylogenetic tree, to be genetically more closely related to its European branch (Figure 2). This result questions the actual species classification within *Apatania* genus and justifies a re-evaluation of the taxonomic status of this group. Either the two main branches of *A. zonella* in the CO1a phylogenetic tree (Figure 2) should be regarded as two different species; or *A. auricular*, *A. dalecarlica* and *A. forsslundi* should be considered as subspecies of *A. zonella*. The comparison of the clades, between the different phylogenetic trees, reveals that these clades are not consistent across markers and indicate recombinations. Therefore sexual reproduction appears to have occurred between individuals of *A. zonella* species, which suggests that redefining *A. auricular*, *A. dalecarlica* and *A. forsslundi* as subspecies of *A. zonella* would be more appropriate. Moreover, the same phylogenetic relationship between the different *Apatania* species was also observed in a study by Salokannel *et al.* (2010), in which they consequently defined the group of *Apatania zonella sensu lato* including *A. zonella*, *A. hispida*, *A. auricular*, *A. dalecarlica* and *A. forsslundi*.

The mitochondrial DNA (CO1a), supports two major lineages within *A. zonella*, one with individuals originated from North-America and Greenland, and the other one with individuals from mainland Europe. However, Icelandic samples are shared between these two lineages. The NJ tree based on $D_{A,}$ from the CO1a marker (Figure 4), shows that Icelandic individuals from South-West Iceland are genetically very close to individuals from Greenland and United-States, while Icelandic individuals from North-West and South Iceland are more close to each other and somewhat genetically between European and North-American individuals. As samples from Alaska, Western-most Canada (Yukon Territory) and Mongolia have the most similar genotype to the most recent common ancestor, present *A. zonella* populations may have dispersed from the unglaciated area in Beringia, and colonized Asia and Europe from one side and North-America and Greenland on the other, to finally intermix again within Iceland. The split between those two branches of *A. zonella* is though much older than the LGM. This recolonization hypothesis is also supported by the Neighbor-joining tree of CO1a based on D_A. Large areas in North-West of

American continent and North-East of Asia remained unglaciated, even at these high latitudes, during the LGM, and therefore probably constituted a refugial area for a part of *Apatania* population which likely contracted in this areas, in particular in Beringia (Abbott & Brochmann 2003; Hewitt 2004; Pielou 2008). Due to the substantial tolerance to cold and harsh environmental conditions of the species, it is conceivable that at least part of the population survived in unglaciated areas closer to the ice sheet, in the periglacial areas, which would constitute cryptic northern refugia. This in situ survival would have facilitated a rapid and multidirectional recolonization leading to a complex evolutionary history in the species. The surviving individuals, after the glaciation, recolonized Arctic and subarctic areas by different routes, blending in central locations like Iceland, leading to common haplotype for individuals from different locations, and reducing apparent geographic-linked genetic patterns. Malicky (1983) was the first to hypothesize an in situ survival of the Pleistocene glaciation, for freshwater insects, within the permafrost. One recent study has confirmed this hypothesis by revealing evidences that the montane caddisfly *Drusus discolor*, which is currently considered as a temperate species, was able to survive during the LGM in periglacial mountain streams inside the permafrost areas of central Europe (Pauls et al. 2006). This indicates that food was available for caddisflies species and leads to assume that arctic species like *Apatania* would have the ability and the possibility to survive in periglacial streams. *Apatania* feeds on periphyton which shows a very high adaptability, suggesting that food would not be a limiting factor for its survival and recolonization process. Evidences provided by several studies suggest that biogeography of aquatic fauna can be substantially different from terrestrial fauna biogeography (Englbrecht et al. 2000; Stewart & Lister 2001; Gum et al. 2005; Pauls et al. 2006).

Overall, the phylogenetic trees for *A. zonella* show mixed haplotypes at different locations indicating a relatively weak geographic structure. Small proportion of the genetic diversity between populations is explained by the distance separating them. The mantel tests show that the genetic variation between populations does not fit with the model of isolation by distance, except for CO1a and IDH. For all of the markers used in this study, the phylogenetic tree branches do not correspond well with geographic origin and haplotypes from each population do not cluster in specific clades (Figure 3, 4, 5, 6, 7 and 8). The phylogeographic structures of *Apatania* in different markers are quite shallow. For all markers, the analysis of molecular variance based on pairwise differences reveals that most

of the genetic variation is actually due to the variance within populations (between 55.8% and 65.6%) (Table 5, 8, 11, 14 and 17). Therefore, the geographical structure explains only a minority of the molecular variance in *A. zonella*.

Based on D_A of CO1 and POL, Greenland and South-West Iceland populations are closely related to each other, and are jointly more related genetically to Alaska and Finland than they are to North-West Iceland and South Iceland. This indicates that North-West Iceland and South Iceland populations of *A. zonella* have a different origin than the South-West Iceland, which was likely colonized by individuals from Greenland. The comparisons including samples from Iceland are sensitive to the proportions of the haplotypes characteristic for mainland Europe and North-America found in the different regions in Iceland. Despite the large overall distance within Iceland, similar haplotypes can be found in different regions. The high nucleotide diversity within Greenlandic samples and their phylogenetic relationships for POL and IDH markers indicate that Greenland could also have been colonized by individuals from North-America and from Europe through Iceland. However, the mitochondrial data does not support this assumption. This hypothesis could be verified by adding more data from Greenland, Canada and the United-States. Iceland might have served as a stepping stone for *A. zonella* to colonize Greenland. The Icelandic settlers could have carried *A. zonella* in its larval form when establishing in South of Greenland. In the same way, individuals from Greenland with North-American genotype could have been carried to Iceland.

The different patterns observed for the different markers in the Greenland samples, as for CAD and POL, indicates admixture and signs of recombination. Although the percentage of males is very low in *A.zonella* and sexual reproduction is assumed to be relatively rare, it might still be sufficient to rearrange the alleles to different genetic backgrounds. The Icelandic individuals do not show evident sign of recombination and sexual reproduction.

The nucleotide diversity and the haplotype diversity are relatively high in all the markers for *A. zonella* in Iceland. For all markers, a large majority of the haplotypes are unique to one location, but the genealogy reveals their historical relationships and supports multiple colonization routes for Iceland, from North-America through Greenland, and from mainland Europe.

In general, the data provided by this study suggest that the population structure of *Apatania* in Iceland has been deeply influenced by the late Quaternary Ice Age and post glaciation recolonization processes. The results indicate that Iceland has undergone intricate processes of recolonization by *A. zonella* after the LGM. Nonetheless, the genetic patterns should therefore be interpreted cautiously as they might be affected by different factors, leading to somewhat biased conclusions. Moreover, adults of *Apatania* emerge and reproduce, asynchronously; as a result, samples of adults collected at a certain time might represent only a small fraction of the offspring. Freshwater insects can also present significant population size fluctuations between generations which can affect the genetic diversity (Enders & Wagner 1996). Further researches on this topic and this species would greatly benefit from a larger amount of samples in each location and also more locations to compare with the Icelandic populations. Additionally, for a greater and more complete understanding of the influence of past events, and glaciations periods in particular, on current phylogeography of species, future studies would be greatly enhanced by expanding the parallel use of complementary domains, mainly genetics, ecology, paleontology; palynology and climatology.

References

Abbott RJ, Brochmann C (2003) History and evolution of the arctic flora: in the footsteps of Eric Hulten. *Molecular Ecology*, **12**, 299–313.

Abbott RJ, Smith LC, Milne RI *et al.* (2000) Molecular analysis of plant migration and refugia in the Arctic. *Science*, **289**, 1343–1346.

Akaike H (1974) A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, **19**, 716–723.

Avise JC (1994) *Molecular Markers: Natural History and Evolution*. Springer.

Benke M, Brändle M, Albrecht C, Wilke T (2009) Pleistocene phylogeography and phylogenetic concordance in cold-adapted spring snails (Bythinella spp.). *Molecular Ecology*, **18**, 890–903.

Bennike OLE, Björck S, Böcher J, Walker IANR (2000) The Quaternary arthropod fauna of Greenland: a review with new data. *Bulletin of the Geological Society of Denmark*, **47**, 111–134.

Bhagwat SA, Willis KJ (2008) Species persistence in northerly glacial refugia of Europe: a matter of chance or biogeographical traits ? *Journal of Biogeography*, **35**, 464–482.

Brändle M, Heuser R, Marten A, Brandl R (2007) Population structure of the freshwater flatworm Crenobia alpina (Dana): old lineages and low gene flow. *Journal of Biogeography*, **34**, 1183–1192.

Bray SCE, Austin JJ, Metcalf JL *et al.* (2013) Ancient DNA identifies post-glacial recolonisation, not recent bottlenecks, as the primary driver of contemporary mtDNA phylogeography and diversity in Scandinavian brown bears. *Diversity and Distributions*, **19**, 245–256.

Brunhoff C, Yoccoz NG, Ims RA, Jaarola M (2006) Glacial survival or late glacial colonization? Phylogeography of the root vole (Microtus oeconomus) in north-west Norway. *Journal of Biogeography*, **33**, 2136–2144.

grConroy CJ, Cook JA (2000) Phylogeography of a post-glacial colonizer: Microtus longicaudus (Rodentia: Muridae). *Molecular Ecology*, **9**, 165–175.

Corbet PS (1966) Parthenogenesis in caddisflies (Trichoptera). *Canadian Journal of Zoology*, **44**, 981–982.

Couceiro L, Barreiro R, Ruiz JM, Sotka EE (2007) Genetic isolation by distance among populations of the netted dog whelk Nassarius reticulatus (L.) along the European Atlantic coastline. *The Journal of Heredity*, **98**, 603–610.

De Croos MDST, Pálsson S (2010) Mitochondrial DNA variation and population genetic structure of white shrimp Fenneropenaeus indicus along the coastal belt of Sri Lanka. *Aquatic Living Resources*, **23**, 315–323.

Danforth BN, Fang J, Sipes S (2006) Analysis of family-level relationships in bees (Hymenoptera: Apiformes) using 28S and two previously unexplored nuclear genes: CAD and RNA polymerase II. *Molecular Phylogenetics and Evolution*, **39**, 358–372.

Deffontaine V, Libois R, Kotlík P *et al.* (2005) Beyond the Mediterranean peninsulas: evidence of central European glacial refugia for a temperate forest mammal species, the bank vole (Clethrionomys glareolus). *Molecular Ecology*, **14**, 1727–1739.

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of Molecular Variance Inferred From Metric Distances Among DNA Haplotypes: Application to Human Mitochondrial DNA Restriction Data. *Genetics*, **131**, 479–491.

Flot J-F (2010) SeqPHASE: a web tool for interconverting phase input/output files and fasta sequence alignments. *Molecular Ecology Resources*, **10**, 162–166.

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, **3**, 294–299.

Frankham R (1998) Inbreeding and Extinction: Island Populations. *Conservation Biology*, **12**, 665–675.

Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**, 915–925.

Geffen E, Waidyaratne S, Dalén L *et al.* (2007) Sea ice occurrence predicts genetic isolation in the Arctic fox. *Molecular ecology*, **16**, 4241–55.

Gíslason GM (1977) Aspects of the biology of Icelandic Trichoptera, with comparative studies on selected species from Northumberland. University of Newcastle upon Tyne, England.

Gíslason GM (2005) Origin of freshwater fauna of the North-Atlantic islands: present distribution in relation to climate and possible migration routes. *Internationale Vereinigung für Theoretische und Angewandte Limnologie Verhandlungen*, **29**, 198–203.

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, **52**, 696–704.

Gum B, Gross R, Kuehn R (2005) Mitochondrial and nuclear DNA phylogeography of European grayling (Thymallus thymallus): evidence for secondary contact zones in central Europe. *Molecular ecology*, **14**, 1707–1725.

Haag CR, Ebert D (2004) A new hypothesis to explain geographic parthenogenesis. *Annales Zoologici Fennici*, **41**, 539–544.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, **41**, 95–98.

Hänfling B, Hellemans B, Volckaert FAM, Carvalho GR (2002) Late glacial history of the cold-adapted freshwater fish Cottus gobio, revealed by microsatellites. *Molecular Ecology*, **11**, 1717–1729.

Hartl DL, Clark AG (1997) *Principles of Population Genetics*. Sinauer Associates Inc., Sunderland, Massachusetts.

Hewitt GM (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.

Hewitt GM (2001) Speciation, hybrid zones and phylogeography - or seeing genes in space and time. *Molecular Ecology*, **10**, 537–549.

Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society. Series B, Biological Sciences*, **359**, 183–195.

Horne DJ, Martens K (1999) Geographical parthenogenesis in European non-marine ostracods: post-glacial invasion or Holocene stability? *Hydrobiologia*, **391**, 1–7.

Johanson KA, Malm T (2010) Testing the monophyly of Calocidae (Insecta: Trichoptera) based on multiple molecular data. *Molecular Phylogenetics and Evolution*, **54**, 535–541.

Kearney M (2005) Hybridization, glaciation and geographical parthenogenesis. *Trends in Ecology & Evolution*, **20**, 495–502.

Kornobis E, Pálsson S, Kristjánsson BK, Svavarsson J (2010) Molecular evidence of the survival of subterranean amphipods (Arthropoda) during Ice Age underneath glaciers in Iceland. *Molecular Ecology*, **19**, 2516–2530.

Kornobis E, Pálsson S, Sidorov DA, Holsinger JR, Kristjánsson BK (2011) Molecular taxonomy and phylogenetic affinities of two groundwater amphipods, Crangonyx islandicus and Crymostygius thingvallensis, endemic to Iceland. *Molecular Phylogenetics and Evolution*, **58**, 527–539.

Kotlík P, Deffontaine V, Mascheretti S *et al.* (2006) A northern glacial refugium for bank voles (Clethrionomys glareolus). In: *Proceedings of the National Academy of Sciences of the United States of America* , pp. 14860–14864.

Larkin MA, Blackshields G, Brown NP *et al.* (2007) Clustal W and Clustal X version 2.0. *Bioinformatics*, **23**, 2947–2948.

Latch EK, Heffelfinger JR, Fike JA, Rhodes OE (2009) Species-wide phylogeography of North American mule deer (Odocoileus hemionus): cryptic glacial refugia and postglacial recolonization. *Molecular Ecology*, **18**, 1730–1745.

Law JH, Crespi BJ (2002) The evolution of geographic parthenogenesis in Timema walking-sticks. *Molecular Ecology*, **11**, 1471–1489.

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–1452.

Malicky, H. (1983). Chorological patterns and biome types of European Trichoptera and other freshwater insects. Archiv für Hydrobiologie, 96, 223–244.

Malm T, Johanson KA (2011) A new classification of the long-horned caddisflies (Trichoptera: Leptoceridae) based on molecular data. *BMC Evolutionary Biology*, **11**, 10–26.

Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research*, **27**, 209–220.

Milá B, Girman DJ, Kimura M, Smith TB (2000) Genetic evidence for the effect of a postglacial population expansion on the phylogeography of a North American songbird. *Proceedings of the Royal Society. Series B, Biological Sciences*, **267**, 1033–1040.

Naro-Maciel E, Reid B, Holmes KE *et al.* (2011) Mitochondrial DNA sequence variation in spiny lobsters: population expansion, panmixia, and divergence. *Marine Biology*, **158**, 2027–2041.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Nei M, Kumar S (2000) *Molecular Evolution and Phylogenetics*. Oxford University Press.

Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences*, **76**, 5269–5273.

Pamilo P, Nei M (1988) Relationships between gene trees and species trees. *Molecular Biology and Evolution*, **5**, 568–583.

Paradis E, Claude J, Strimmer K (2004) APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics*, **20**, 289–290.

Pauls S, Lumbsch HT, Haase P (2003) Genetische isolation von Drusus discolor rambur 1842 (Trichoptera: Limnephilidae) in montanen mittelgebirgen mitteleuropas - Erste ergebnisse. In: *Deutsche Gesellschaft für Limnologie* , pp. 378–379. DGL-Tagungsbericht.

Pauls SU, Lumbsch HT, Haase P (2006) Phylogeography of the montane caddisfly Drusus discolor: evidence for multiple refugia and periglacial survival. *Molecular Ecology*, **15**, 2153–2169.

Petersen I, Winterbottom JH, Orton S *et al.* (1999) Emergence and lateral dispersal of adult Plecoptera and Trichoptera from Broadstone Stream, U.K. *Freshwater Biology*, **42**, 401–416.

Pielou EC (2008) *After the Ice Age: the return of life to glaciated North America*. University of Chicago, Chicago.

Pinceel J, Jordaens K, Pfenninger M, Backeljau T (2005) Rangewide phylogeography of a terrestrial slug in Europe: evidence for Alpine refugia and rapid colonization after the Pleistocene glaciations. *Molecular Ecology*, **14**, 1133–1150.

Posada D, Buckley TR (2004) Model selection and model averaging in phylogenetics: advantages of akaike information criterion and bayesian approaches over likelihood ratio tests. *Systematic Biology*, **53**, 793–808.

Prieto-Arribas D (2009) Phylogeography of Apatania zonella (Zett.) and Potamophylax cingulatus (Stph.). University of Iceland.

Provan J, Bennett KD (2008) Phylogeographic insights into cryptic glacial refugia. *Trends in Ecology & Evolution*, **23**, 564–571.

Raastad JE, Solem JO (1989) Autogeny as successful reproductive strategy in high altitude black-flies (Diptera, Simuliidae). *Annales de Limnologie*, **25**, 243–249.

Salokannel J, Rantala MJ, Wahlberg N (2010) DNA-barcoding clarifies species definitions of Finnish Apatania (Trichoptera: Apataniidae). *Entomologica Fennica*, **21**, 1–11.

Sanz-Muñoz S (2010) Mitochondrial variation of the caddisflies Apatania zonella and Potamophylax cingulatus. University of Iceland.

Schliep KP (2011) phangorn: phylogenetic analysis in R. *Bioinformatics*, **27**, 592–593.

Schmitt T (2007) Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. *Frontiers in Zoology*, **4**, 11.

Sode A, Wiberg-Larsen P (1993) Dispersal of adult Trichoptera at a Danish forest brook. *Freshwater Biology*, **30**, 439–446.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.

Tamura K, Peterson D, Peterson N *et al.* (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, **28**, 2731–2739.

Verovnik R, Sket B, Trontelj P (2005) The colonization of Europe by the freshwater crustacean Asellus aquaticus (Crustacea: Isopoda) proceeded from ancient refugia and was directed by habitat connectivity. *Molecular Ecology*, **14**, 4355–4369.

Vogel SW (2008) Cryosphere: Fire and ice. *Nature Geoscience*, **1**, 91–92.

Zvuloni A, Mokady O, Al-Zibdah M *et al.* (2008) Local scale genetic structure in coral populations: a signature of selection. *Marine Pollution Bulletin*, **56**, 430–438.

Chapter 2: Morphological variation of genitalia in *Apatania zonella* and *Apatania hispida*

Introduction

Morphological variation between species and populations of the same species from different geographical areas is a key element in the study of evolutionary history of species and early stages of speciation. Traits which are variable between species are expected to be variable between geographically distinct populations if these traits evolve rapidly. Genital structures have served as one of the main traits in studies of morphological variation and species identification in insects.

Insect genitalia reveal useful patterns of variation as they are considered to have been shaped by rapid divergent evolution and be more species-specific than non-genital features (Eberhard 1985), thus they have been commonly used by taxonomists for identification of genus, species and subspecies among insects. Based on the lock-and-key hypothesis, internal and external genital differences, in size and shape, lead to mechanical isolation between species in terms of sexual reproduction (Shapiro & Porter 1989), and such variation may thus be informative when defining species based on reproductive barriers.

Qualitative description for species delimitation is subjective and thus substantially unreliable, however quantitative morphometric analysis are assumed to be statistically more powerful and important tool for discrimination of species (Mutanen & Pretorius 2007; Holwell 2008). Despite high similarities among closely related individuals that enable clustering of different groups of individuals, the morphology of genitalia is unique to each individual. Therefore, the individual variation should be carefully considered, in studies focusing on variation between species, subspecies or geographical populations.

Although taxa or species delimitation between populations showing consistent variations looks practical and well-founded, it may need to be confirmed or at least supported by genetical, behavioral or ecological data (Silva-Brandão *et al.* 2009). This morphological classification method could have problematic bias in some studies of genitally polymorphic species (Huber & Pérez-González 2001), as it could lead to the classification of different morphotypes of one species into different species (Isaac *et al.* 2004), while the variation

could actually reflect morphological plasticity due to geographical and environmental effects. Furthermore, this method could also hide genetical divergences in case of interpopulation uniformity (Zeh & Zeh 1994). Observations of variations in genitalia between populations geographically isolated are scarce, potentially resulting from this classification method bias (Mutanen & Kaitala 2006; Mutanen & Pretorius 2007). Most studies on geographical variation in genital morphology are focusing on closely related species rather than on a single polymorphic species (Sullivan *et al.* 1990; Mutanen & Kaitala 2006; Holwell 2008) or focusing on males genitalia (Kelly *et al.* 2000; Mutanen *et al.* 2007; Andrade *et al.* 2009), as they are usually assumed to be more divergent and easier to measure than female genitalia. Therefore, there are fewer descriptions and comparisons of female genitalia in the literature, at both interspecific and intraspecific level (Arnqvist 1997).

The most common way to explore population divergences in morphology is to study geographic variation in different populations of one species, and compare closely related species. In this study, the main focus is the patterns of morphological variation in female genitalia among populations of *Apatania zonella*, from South Greenland and Iceland, and including a comparison with a Norwegian population of *Apatania hispida*, a closely related species. These two species are comparable as the morphological characters of their body and genitalia are quite similar, and it can be challenging to differentiate them, especially when considering the individual variations within each species (Solem 1985; Salokannel *et al.* 2010). The molecular genetic analysis of the populations, in chapter 1, shows that the phylogenetic classification of the genus *Apatania* needs reevaluation as there is larger genetic divergence within *A. zonella* than between *A. hispida* and some *A. zonella* populations.

A. zonella presents a taxonomic problem as the female genitalia morphology is highly variable (Schmid 1953; Solem 1985; Andersen & Wiberg-Larsen 1987). However, there are no clear or sufficient morphological or ecological evidences, leading to a classification of the individuals in more than one species or in subspecies (Solem 1985).

A. zonella reproduces mainly by parthenogenesis, but a small percentage of males persists in most populations, about 1-2% (Corbet 1966; Gíslason 1977; Raastad & Solem 1989;

Salokannel *et al.* 2010). Conversely, *A. hispida* is assumed to be a completely parthenogenetic species. Asexual reproduction is mainly observed at high altitudes or latitudes, in insular environments, in ecologically disturbed areas, in dry and arid habitats, and also in various marginal and sparsely inhabited environments (Glesener & Tilman 1978; Lynch 1984; Peck *et al.* 1998; Kearney *et al.* 2006). In harsh environments, asexual reproduction may be more advantageous than sexual reproduction (Law & Crespi 2002; Błoszyk *et al.* 2004; Kearney 2005; Ben-Ami & Heller 2007) since finding mates and keeping sexual reproduction consume a lot of energy (Bell 1982; Peck *et al.* 1998). Partial parthenogenesis could also be a result of biased sex-ratio due to higher mortality in males than females (Ben-Ami & Heller 2007), e.g. by predation or parasites infections. Likewise juveniles males might have lower survival chances than juveniles females, leading to imbalanced sex-ratio in adults (Ben-Ami & Heller 2007).

Two of the main advantages of parthenogenetic reproduction are the possibility of fast and massive colonization of new habitats and the ability of reproduction in rough environments. This reproduction type enables an easier colonization of habitats that are hardly accessible for populations with sexual reproduction (Raastad & Solem 1989; Błoszyk *et al.* 2004). On the other side, parthenogenesis may result in decreased genetic variability among populations and is expected to lead to an accumulation of deleterious mutations accumulation which could be a significant disadvantage in long term for adaptation and survival of the population (Bell 1982; Butlin *et al.* 1998).

During the last glacial maximum of Ice Age (LGM, 23 000 - 18 000 years ago), most species in Europe were confined to southern Mediterranean refugial areas (Iberian Peninsula, Italian Peninsula, and Balkan Peninsula), due to repeated glaciations, and recolonized the northern territories afterwards (Hewitt 2004). Due to its wide circumpolar distribution through Arctic and subarctic areas, *A. zonella* could have survive in other refugia of the Holarctic: Beringia and southern North-America. The LGM had an important impact on the current distribution and genetic diversity of species, especially in northern areas (Hewitt 2004). Iceland, Greenland and Norway were almost completely covered by ice sheet during this glaciation period (Ehlers & Gibbard 2007). Nonetheless, it is possible that *Apatania* survived along the glacial edge, within the permafrost in northern Europe. This phenomenon was observed for some other Arctic species (Ægisdóttir & Þórhallsdóttir

2004; Denk *et al.* 2011). Populations of Greenland, Iceland and Norway could have survived in different refugia during the glacial period. Furthermore, the present Icelandic populations of *Apatania* could potentially have originated from both North-America and Europe. This hypothesis is well supported by the analysis genetical data in chapter 1, especially from the mitochondrial marker CO1.

Parthenogenesis provided a significant advantage in restricted areas and newly available habitats after the glaciation period, as it enhances adaptability and dispersal ability (Suomalainen 1962; Downes 1965). Some studies have revealed the possible occurrence of small and constricted ice-free areas in Iceland along coastal mountains in North-West and North-East, that could be defined as nunataks, possibly hosting highly resilient plant species (Rundgren & Ingólfsson 1999; Andrews *et al.* 2000; Geirsdóttir *et al.* 2007). However, it is commonly accepted that almost no boreal species of animal could have survived the coldest period in the LGM even if Iceland had ice free areas (Gíslason 2005). Therefore, it is generally considered that most of the species currently living in Iceland have colonized the territory during the last deglaciation period (Buckland *et al.* 1986; Coope *et al.* 1986). Nonetheless, some small nunataks probably existed on the South-West of Norway, Faeroes Islands, some North-Atlantic islands, South Greenland and North-West Scotland (Brochmann *et al.* 2003; Gíslason 2005), that could have potentially enabled *A. zonella* to survive in northern areas.

The aim of this work is to shed light on the origin of the Icelandic populations of *A. zonella*, by studying morphologic relationships between individuals of different populations and determining how they are linked to their geographical origin. To achieve this goal, several key traits of the genitalia were measured and statistically compared with respect to their origin, including flies from Greenland and *A. hispida* from Norway.

Material and methods

Sample collection

The variation in the morphology of the genitalia was studied among 54 females *Apatania*, collected between 1982 and 2012, in Iceland, Greenland and Norway. Ten individuals of *A. hispida* (closely related species of *A. zonella*) were obtained from the Norwegian University of Science and Technology, collected from Vitenskapsmuseet Trondheim (Norway), between 1988 and 2007. Fourteen Greenlandic individuals of *A. zonella* were obtained from the Natural History Museum of Denmark, collected from Godhavn (also known as Qeqertarsuaq), Kap Farvel and Upernaviarssuk, in 1982 and 1984. Finally 30 individuals were sampled from South and South-West Iceland (Reykjavík area), between 2005 and 2012 (Figure 1). The species identification of *A. zonella* based on morphological characteristics was done by Gísli Már Gíslason, a Trichoptera specialist.

Only mature individuals were used in this study, and all the individuals collected were females. The samples were mainly obtained from previous studies on *Apatania* and some new samples have been collected in Iceland in 2010 and 2012. The sampling took place during the flight period, which starts at the beginning of May and lasts until the end of August and was carried out using nets and traps placed next to stony cold springs, mountains brooks and subarctic rivers. Specimens were stored in 96% ethanol.

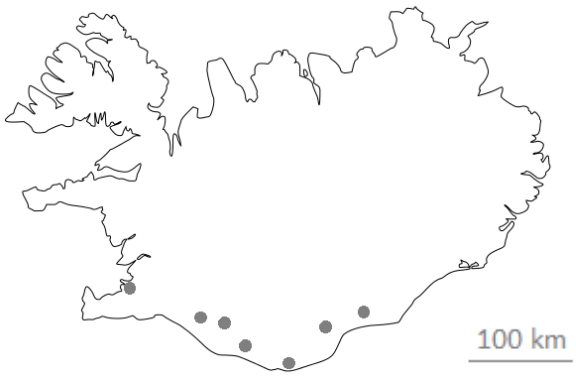

Fig 1. Sampling sites of *A. zonella* within Iceland (dots).

The genitalia of each individual was photographed, using an Evolution™ LC camera mounted on a stereomicroscope and using the software PixelLINK®. All pictures were

taken at the exact same magnification. Individuals with damaged genitalia were discarded from the study.

Morphological traits and measurements

A total of 7 measurements were taken from the genitalia area (figure 2). Straight distances between easily identifiable points were chosen with the aim to cover the main morphological characteristics of the female *Apatania* genitalia.

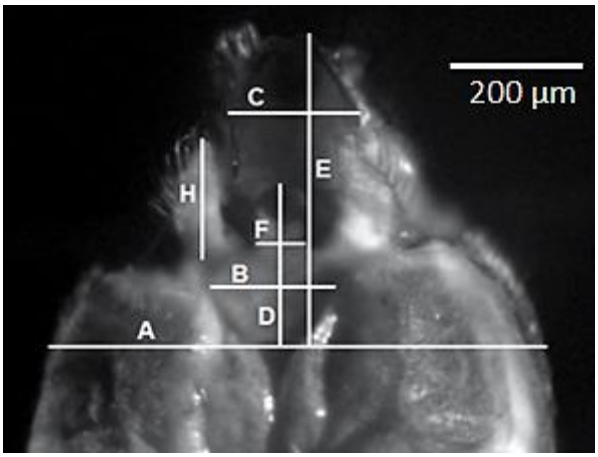

Fig 2. Picture of the genitalia of a female *A. zonella* with the 7 measurements used for the morphometric analysis.

To test the size effect on the measurements studied, the total width was measured exactly under the central part of the genitalia (**A**), and also the total length of the genitalia area from the line for the total width to the extremity (**E**). Other measurements represent the width of the central part (**B**), the length of the central part (**D**), the width of the protuberance of the central part (**F**), the width of the extremity (**C**) and the length of the side part (**H**).

As the genitalia structure is in three dimensions, several pictures were taken per individual, focused on different parts of the genitalia area. These pictures were then assembled using

the *GNU Image Manipulation Program (GIMP 2.6.11,* © 2001-2012 The GIMP Team*)* into one picture showing clearly all the different parts of the genitalia. The measurements were taken from each combined picture for every individual using GIMP.

Morphometric statistical analysis

Shapiro-Wilk's test and Bartlett test were performed on each set of measurements, globally and also among each group, to test respectively the normal distribution of the data and the homogeneity of the variances. Pearson correlation coefficient (r) was calculated between all pairs of variables to evaluate their dependencies. To avoid redundancy and dependency of the variables on size, two methods were used to find the best model to fit the data: the linear regression method (Hastie et al. 2009) and the allometric correction method (Lleonart *et al.* 2000). As the R-squared were either lower or similar in the allometric correction method, the linear regression was selected as the best fit to the measurements. Each variable was then regressed on the total width (**A**) and the residuals from the predicted linear regressions were calculated for each measurement. The following analysis was based on these residuals as size-adjusted values for these measurements.

One way analysis of variance (ANOVA) was conducted on each variable to test whether the *Apatania* genitalia differs between countries. Tukey's post hoc tests were applied to explore further which countries differed from each other. The boxplots of the residuals for each variable were plotted and interpreted along with the results from the ANOVA and Tukey's post hoc tests.

As the differences between countries can also result from the combinations of the morphological traits, the dataset was analyzed with a multivariate analysis: Linear Discriminant Analysis (LDA). The method finds a linear combination of the variables which maximizes the ratio of the variance among groups to the total variance. Before performing the LDA, stepwise variable selection was conducted, using the Wilk's lambda criterion, to keep only the independent variables that have a relatively important explanatory power. The LDA was then done on the morphological traits selected by this method, using the country of origin as the dependent variable. The patterns of genital

morphology differences and similarities given by the result of the LDA were presented on a scatterplot. All statistical tests were performed using the R software 2.15.0 (R Development Core Team, 2012).

Results

Shapiro-Wilk and Bartlett tests show that the measurements are normally distributed, both globally and among groups, and the variances are homogeneous. Correlation coefficients shows 11 significant correlations ($p<0.05$) for a total of 21 possible pairs of variable (Table 1). Most of the variables were significantly correlated with the size of the individual. To eliminate the size effect in the measurements, in order to avoid redundancy, the variables were adjusted by size.

Table 1. Pearson correlation coefficients (r) between morphometric variables and significance levels.

	A	B	C	D	E	F	H
A	---	---	---	---	---	---	---
B	0.27 *	---	---	---	---	---	---
C	0.42 **	0.27	---	---	---	---	---
D	0.37 **	0.28 *	0.38 **	---	---	---	---
E	0.63 ***	0.09	0.31 *	0.51 ***	---	---	---
F	0.25	0.18	0.18	0.11	0.22	---	---
H	0.42 **	0.30 *	0.05	0.08	0.23	0.28 *	---

Signif. codes: <0.001 '***' 0.01 '**' 0.05 '*' >0.05 ' '

Shapiro-Wilk and Bartlett tests were run again on adjusted values, which revealed a normal distribution, both globally and among groups, and homogeneous variances. The new variables (residuals from size-adjusted variables) were independent from each other as the pairwise correlations were all non-significant ($p>0.05$), except for variables **DA r** and **EA r** ($p<0.05$), for a total of 21 possible pairs of variables (Table 2).

Table 2. Pearson correlation coefficients between A (global width) and size-adjusted variables (residuals), and significance levels. Names of the variables refer to variables in figure 2.

	A	EA r	BA r	CA r	DA r	FA r	HA r
A	---	---	---	---	---	---	---
EA r	0	---	---	---	---	---	---
BA r	0	-0.11	---	---	---	---	---
CA r	0	0.05	0.18	---	---	---	---
DA r	0	0.39 **	0.18	0.25	---	---	---
FA r	0	0.08	0.13	0.08	0.01	---	---
HA r	0	-0.05	0.21	-0.15	-0.10	0.20	---

Signif. codes: <0.001 '***' 0.01 '**' 0.05 '*' >0.05 ' '

Table 3. ANOVA statistical tests for each variable, Tukey's post hoc tests between groups and significance levels

	ANOVA		Tukey's post hoc tests		
	F	**P**	Iceland versus Greenland	Iceland versus Norway	Greenland versus Norway
A	4.5531	0.01516 *	0.3964	0.0121 *	0.2539
EA r	5.2777	0.008256 **	0.4561	0.0060 **	0.1459
BA r	3.4931	0.03787 *	0.5225	0.0310 *	0.3218
CA r	6.3683	0.003397 **	0.0023 **	0.7617	0.0885
DA r	7.9861	9.61e-04 ***	9.9e-04 ***	0.9999	0.0117 *
FA r	0.3623	0.6978	---	---	---
HA r	1.157	0.3225	---	---	---

Signif. codes: <0.001 '***' 0.01 '**' 0.05 '*' >0.05 ' '

The ANOVA results reveal that fives measurements (new variables), out of seven, show significantly differences among locations: **A**, **EA r**, **BA r**, **CA r**, and **DA r**. (Table 3). The results of the Tukey's post hoc tests reveals a significant difference between *A. zonella* from Iceland and *A. hispida* from Norway in the total width of the genitalia area (**A**), in the total length (**EA r**) and in the width of the central part, both corrected for the size effect (**BA r**). However, the width of the extremity size-adjusted (**CAr**) shows only a significant difference among *A. zonella*, between Iceland and Greenland. Despite a high significant correlation coefficient between the residuals **EAr** and **DAr**, they show completely different results in the Tukey's post hoc test. The length of the central part corrected for size effect (**DAr**) revealed a significant difference between *A. zonella* from Greenland and *A. hispida* from Norway, and a highly significant difference among *A. zonella* between Iceland and Greenland.

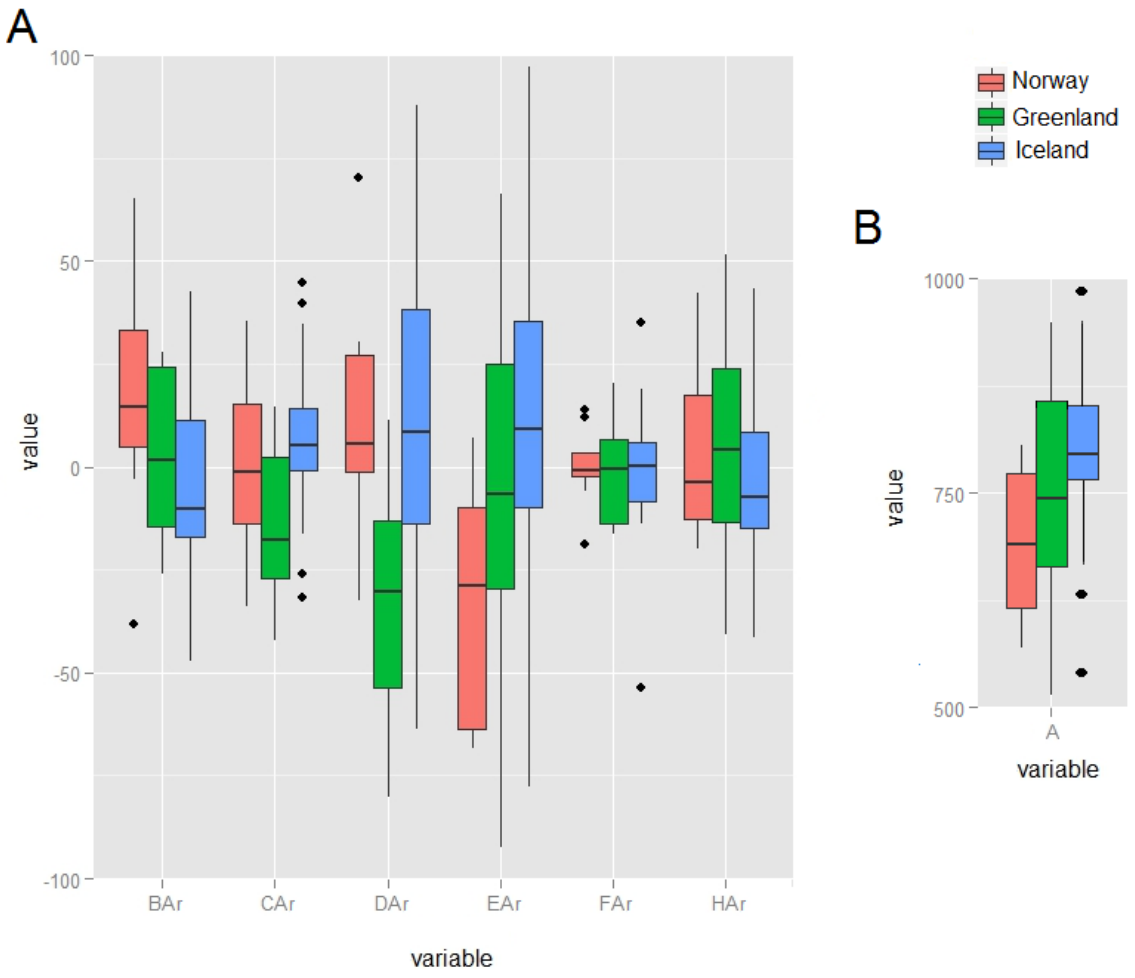

Fig 3. A: Boxplots of the size-adjusted variables (residuals) for each country. **B**: Boxplot of the measurements for the variable A for each country.

The boxplots, obtained from the adjusted variables in R, specify the results from Tukey's post hoc tests (Figure 3A-B). It reveals that the values of the variables **A** and **EAr** tend to be much higher for the Icelandic individuals than for the Norwegians ones, which represent the global size of the genitalia area (width and length). The boxplot shows the opposite concerning the variable **BAr** (width of the central part). For these three variables, Greenland individuals stand in-between. Nonetheless, for the variable **CAr** (width of the extremity), the boxplot indicates a greater length for Icelandic individuals than for Greenlandic ones, with the group from Norway clearly in the middle. Concerning the variable **DAr** (length of the central part), it reveals that flies from Iceland and Norway have larger values for this measurement but flies from Greenland have much smaller values in average. The boxplots of **FAr** (adjusted width of the protuberance of the central part) and

HA r (adjusted length of the side part) show that these measurements are very stable between locations and also between the two species independently from the size. Despite significant differences in size between the three groups, these measurements are perfectly constant, meaning they are not proportional with the global size of the genitalia area.

The stepwise variable selection using the Wilk's lambda criterion reveals that only five of the seven variables have a significant explanatory power and should therefore be used in the Linear Discriminant Analysis, with the location as predicator (Table 4). These selected variables match exactly the variables showing significant results in the ANOVA.

Table 4. Stepwise variable selection using Wilk's lambda criterion.

Step	Variable added	Wilk's λ	F	p value
1	DAr	0.76	7.986	9.61e-04 ***
2	EAr	0.58	7.672	1.22e-03 **
3	A	0.47	5.577	6.51e-03 **
4	CAr	0.40	4.299	1.90e-02 *
5	BAr	0.32	5.790	5.59e-03 **

Signif. codes: <0.001 '***' 0.01 '**' 0.05 '*' >0.05 ' '

The Linear Discriminant Analysis resulted in two linear combinations of the variables. The first axis explains 60% (eigenvalue 4.852) of the variance in genitalia morphology between countries and separate mostly Icelandic flies from the rest (Figure 4A); the second axis covers 40% (eigenvalue 3.944) of the variation and separate mostly Greenlandic and Norwegian flies (Figure 4A), with Icelandic ones in between. In the reclassification performed by the LDA, 11 individuals, out 54, were classified in the incorrect country. Therefore, the LDA could correctly classify 80% of the individuals, and this percentage of accuracy is perfectly constant between groups. The result of Pearson's Chi-Square test shows a very high and significant similarities between the real classes and the classes predicted by the LDA (p-value=2.15e-10).

Standardized Linear Discriminant Functions:

LD1 = $1.12e^{-2}$***DAr** + $9.50e^{-3}$***EAr** + $6.33e^{-3}$***A** + $3.42e^{-2}$***CAr** - 3.10***BAr** – 4.86

LD2 = - $2.66e^{-2}$***DAr** + $2.15e^{-2}$***EAr** + $3.77e^{-3}$***A** - $1.18e^{-2}$***CAr** - $7.47e^{-4}$***BAr** – 2.90

The individuals from different countries are relatively well distinguished by the Linear Discriminant Analysis (Figure 4A). The means of countries show that samples from Norway tend to be morphologically more different from Greenland and Iceland regarding the genitalia area, than Greenlandic and Icelandic samples are between each other (Figure 6).

The LDA, in particular the mismatches, shows that the overall characteristics of Icelandic *Apatania* overlap with Norwegians and Greenlandic while this last two do not overlap (figure 4B). The result of the LDA (Figure 4 and 5) allows us to deduce average genitalia morphotype for each country. Individuals from Iceland have longer and wider genital area, wider extremity, longer and thinner central part. Individuals from Greenland have a medium global size genitalia, thin extremity, short length and medium width of the central part. Norwegian individuals have smaller and thinner genital area, the width of the extremity is very similar to Icelandic ones, just like for the length of the central part. However, *A. hispida* from Norway shows much wider central part than *A. zonella*.

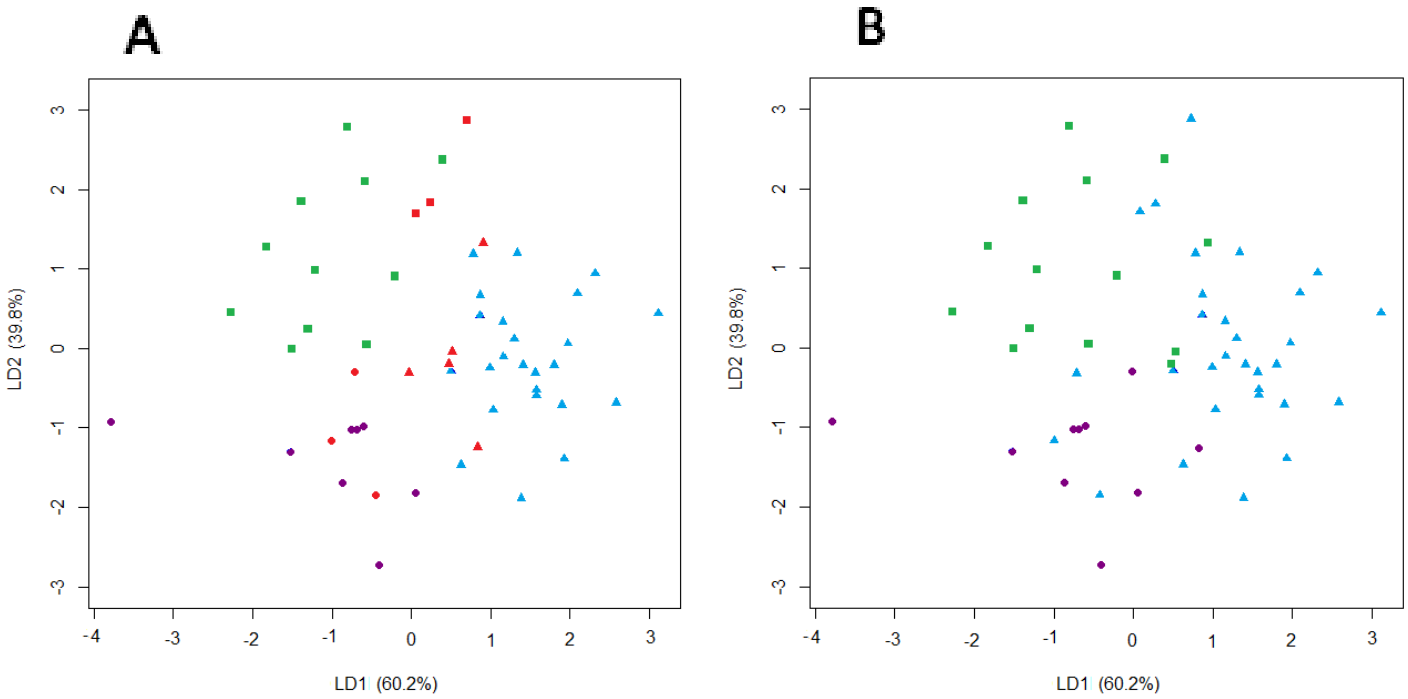

Fig 5. A: Graph of the classification of *Apatania zonella* individuals, from different countries, performed by the Linear Discriminant Analysis, with the percentage of variance explained by each axis. **B**: Graph of the real country of origin of each individual. ■ = Greenland, ● = Norway, ▲ = Iceland, red symbols (■ ● ▲) represents the incorrect country classification by the LDA.

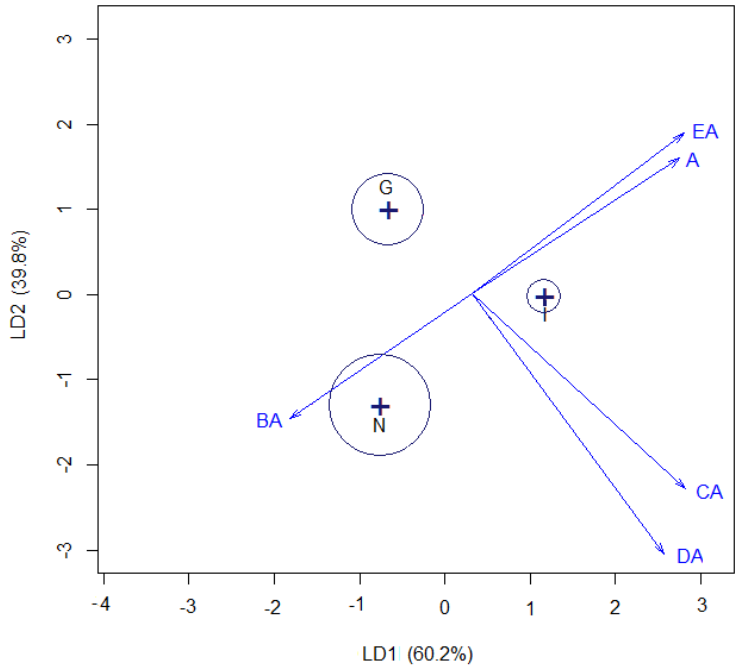

Fig 6. Biplot showing the means (+) of the average linear discriminant for each country Greenland (G), Iceland (I) and Norway (N). The circles correspond to a 95% confidence limit for the mean. The contribution and direction of the variables to the axes are presented with arrows.

Discussion

The results from this study show clear patterns of morphometric variation in female genitalia of *Apatania zonella* from Greenland, Iceland and *Apatania hispida* from Norway. Overall, the genital features allow to discriminate the different populations into morphological and geographical clusters quite well; especially considering the samples size.

The result of the LDA shows greater morphological differences between *A. hispida* and *A. zonella*, than between Greenlandic and Icelandic populations of *A. zonella* (Figure 5 and 6). This result was expected considering that *A. hispida* and *A. zonella* are considered as two different species. Nonetheless, the higher interspecies variation is relatively small, and there is a clear overlap in genitalia features between some Icelandic individuals and some Norwegians ones (Figure 4B). This result is somehow surprising, especially given the fact that most of the species in insect taxa have been differentiated and classified based on their genital features. A more distinct interspecies variation in genital morphology would be expected.

In general, most genital features are positively correlated with the body size. The results show indeed high correlation between global size and five of the morphological traits studied in the genitalia area. As the size is significantly different between countries, all the morphological traits have been adjusted for the global size. Without size-adjustment, differences observed in genitalia traits may only reflect differences in size of the individuals from different populations. Aside from the size effect on genitalia morphology, four of the traits show significant differences between populations. Therefore these differences have another origin than the differences observed in overall body size. The difference in overall size is significant between Iceland and Norway and might reflect a difference more specific to the species or populations. The dissimilarity or morphological distances of the populations can reflect the time since they diverged, or it may result from environmental factors such as climate, competition, availability, quality and type of food (Iversen 1974; Svensson 1975), developmental conditions, etc. (Garland & Losos 1994). Morphological plasticity within a species in response to different environments can further add to the divergence and has been documented in other insect species (Blanckenhorn

1991; Taylor & Merriam 1995). Variations in temperature, linked to latitude, altitude, origin and distance from the source, during the development of insects larvae has previously shown to lead to significant variations in adults size (Van den Heuvel 1963; McCafferty & Pereira 1984; Lyimo *et al.* 1992; Panov & McQueen 1998). Nonetheless, these differences could also be simply linked with shifts in the reproduction period and consequently in the developmental period, due to photoperiod or weather conditions differences for instance. It is known that populations of *A. zonella* present unsynchronized life cycle and new adults emerge during the whole summer (Gíslason & Sigfusson 1987; Gíslason 2005). A common garden experiment, where larvae from different origins would grow at the same time and in the same environmental conditions, would help to understand if environment and timing could have an effect on adult size. *A. zonella* has a wide distribution from lowlands to mountainous areas in Arctic and subarctic environment, but undergoes a strong competition with other freshwater invertebrate species in lowlands areas and this might also have an effect on adult size of individuals experiencing higher competition in larval stage (Andersen & Wiig 1987). The variation in genitalia morphology could also be the result of genetic differences between populations from different locations. An ancient separation between these populations and the absence of contact between them would lead individuals from different locations to evolve differently, especially in morphological aspects. Genetic data results suggest that *A. zonella* colonized Iceland by two or more colonization events (Chapter 1), and this could be reflected in *A. zonella* morphology. However, little is known yet about the genetic basis of genital morphology. Andersen and Wiig (1987) found a significant correlation between shape differences and geographical distances in *A. zonella*, but also between wings/legs dimensions and altitude.

The required morphological divergence leading to mechanical isolation in terms of sexual reproduction, according to lock-and-key hypothesis, is unknown, but it is improbable that the variations observed in this study could result in mechanical isolation. One of the most obvious reasons is that individuals with extreme morphologies from the same population are clearly more different from each other, than individuals with the most similar morphologies from different populations. Although the sexual mechanical isolation might not be really relevant in this case since *A. hispida* is assumed to be completely parthenogenetic and *A. zonella* reproduces essentially through parthenogenesis. However,

even occasional sexual reproduction could be enough for significant gene flow between populations.

Males genitalia are among the most divergent and the fastest evolving morphological traits in animals, especially the intermittent organ which can even delimitate cryptic species in some cases (Tidon-Sklorz & De Melo-Sene 1995). However, analyze males genitalia in this study was not possible as no males could be found during the sampling due to the parthenogenetic reproduction strategy of the two species studied (Corbet 1966; Raastad & Solem 1989; Salokannel *et al.* 2010).

Sexual selection hypothesis assumes that all sexual related traits are constantly going under sexual selection leading to variations in genitalia according to geographical area and ecological living conditions (Andersson 1994; King Sirot 2003; Rundle *et al.* 2005). Nonetheless, the influence of sexual selection in this study is most likely low considering the high preponderance of parthenogenesis as reproduction strategy in these two species. Therefore, the biogeographic pattern observed in this study could be better explained by vicariance, natural selection and geographical isolation creating geo-morphotype (Endler 1986), and therefore this clear variations between countries would be most likely maintained by the lack of gene flow between Iceland, Greenland and Norway. Furthermore, in partial or almost complete parthenogenesis cases like in this study, the gene flow in the population would be much lower as most eggs produced are female clones of their biological mother.

Long term allopatry, bottleneck events and glaciation events followed by rapid expansion with small founder populations as a starting point, may result in morphological divergences between geographically isolated populations. These phenomena, among other, are at the origin of speciation processes. Morphological variation in *Apatania zonella* is most likely a combination of the result of different responses to different ecological conditions, and also post-glacial recolonization with a founder effect, and isolation between populations, both in a geographic and reproduction level.

References

Ægisdóttir HH, Þórhallsdóttir ÞE (2004) Theories on migration and history of the North-Atlantic flora: a review. *Jökull*, **54**, 1–16.

Andersen T, Wiberg-Larsen P (1987) Revised check-list of NW European Trichoptera. *Insect Systematics & Evolution*, **18**, 165–184.

Andersen T, Wiig Ø (1987) Sex proportions, flight periods and morphometric variation in West Norwegian Apatania zonella (Zetterstedt, 1840) (Trichoptera: Limnephilidae). *Insect Systematics & Evolution*, **18**, 185–204.

Andersson MB (1994) *Sexual Selection* (PU Press, Ed,). Princeton University Press.

Andrade CAC, Vieira RD, Ananina G, Klaczko LB (2009) Evolution of the male genitalia: morphological variation of the aedeagi in a natural population of Drosophila mediopunctata. *Genetica*, **135**, 13–23.

Andrews JT, Hardardóttir J, Helgadóttir G *et al.* (2000) The N and W Iceland Shelf: insights into Last Glacial Maximum ice extent and deglaciation based on acoustic stratigraphy and basal radiocarbon AMS dates. *Quaternary Science Reviews*, **19**, 619–631.

Arnqvist G (1997) The evolution of animal genitalia: distinguishing between hypotheses by single species studies. *Biological Journal of the Linnean Society*, **60**, 365–379.

Bell G (1982) *The Masterpiece of Nature: The Evolution and Genetics of Sexuality*. University of California Press, Berkeley.

Ben-Ami F, Heller J (2007) Temporal patterns of geographic parthenogenesis in a freshwater snail. *Biological Journal of the Linnean Society*, **91**, 711–718.

Blanckenhorn WU (1991) Fitness consequences of food-based territoriality in water striders, Gerris remigis. *Animal behaviour*, **42**, 147–149.

Błoszyk J, Adamski Z, Napierala A, Dylewska M (2004) Parthenogenesis as a life strategy among mites of the suborder Uropodina (Acari: Mesostigmata). *Canadian Journal of Zoology*, **82**, 1503–1511.

Brochmann C, Gabrielsen TM, Nordal I, Landvik J, Elven R (2003) Glacial survival or tabula rasa ? The history of North Atlantic biota revisited. *Taxon*, **52**, 417–450.

Buckland PC, Gerrard AJ, Larsen G *et al.* (1986) Late Holocene palaeoecology at Ketilsstadir in Mýrdalur, South Iceland. *Jökull*, **36**, 41–55.

Butlin RK, Schön I, Griffiths HI (1998) *Introduction to reproductive modes. Sex and Parthenogenesis: Evolutionary Ecology of Reproductive Modes in Non-Marine Ostracods* (K Martens, Ed,). Backhuys Publishers, Leiden.

Coope GR, Moore PD, Gibbs A (1986) The Invasion and colonization of the North Atlantic islands: a palaeoecological solution to a biogeographic problem [and

Discussion]. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, **314**, 619–635.

Corbet PS (1966) Parthenogenesis in caddisflies (Trichoptera). *Canadian Journal of Zoology*, **44**, 981–982.

Denk T, Grimsson F, Zetter R, Símonarson LA (2011) Introduction to the nature and geology of Iceland. *Late Cainozoic Floras of Iceland*, **35**, 1–29.

Downes JA (1965) Adaptations of Insects in the Arctic. *Annual Review of Entomology*, **10**, 257–274.

Eberhard WG (1985) *Sexual selection and animal genitalia*. Harvard University Press, Cambridge, Mass. (USA).

Ehlers J, Gibbard PL (2007) The extent and chronology of Cenozoic global glaciation. *Quaternary International*, **164-165**, 6–20.

Endler JA (1986) *Natural Selection in the Wild*. Princeton University Press.

Garland TJ, Losos JB (1994) Ecological morphology of locomotor performance in squamate reptiles. In: *Ecological Morphology: Integrative Organismal Biology* , pp. 240–302.

Geirsdóttir Á, Miller GH, Andrews JT (2007) Glaciation, erosion, and landscape evolution of Iceland. *Journal of Geodynamics*, **43**, 170–186.

Gíslason GM (1977) Aspects of the biology of Icelandic Trichoptera, with comparative studies on selected species from Northumberland. University of Newcastle upon Tyne, England.

Gíslason GM (2005) Origin of freshwater fauna of the North-Atlantic islands: present distribution in relation to climate and possible migration routes. *Internationale Vereinigung für Theoretische und Angewandte Limnologie Verhandlungen*, **29**, 198–203.

Gíslason GM, Sigfusson AT (1987) The life cycle and food of Apatania zonella (Zett.) in a spring-fed stream in SW Iceland (Trichoptera: Limnephilidae). In: *Proceedings of the 5th International Symposium on Trichoptera* (eds Bournaud M, Tachet H), pp. 237–242.

Glesener RR, Tilman D (1978) Sexuality and the components of environmental uncertainty: clues from geographic parthenogenesis in terrestrial animals. *The American Naturalist*, **112**, 659–673.

Van den Heuvel MJ (1963) The effect of rearing temperature on the wing length, thorax length, leg length and ovariole number of the adult mosquito, Aedes aegypti (l.). *Transactions of the Royal Entomological Society of London*, **115**, 197–216.

Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society. Series B, Biological Sciences*, **359**, 183–195.

Holwell GI (2008) Geographic variation in genital morphology of Ciulfina praying mantids. *Journal of Zoology*, **276**, 108–114.

Huber BA, Pérez-González A (2001) Female genital dimorphism in a spider (Araneae: Pholcidae). *The Zoological Society of London*, **255**, 301–304.

Isaac NJB, Mallet J, Mace GM (2004) Taxonomic inflation: its influence on macroecology and conservation. *Trends in Ecology & Evolution*, **19**, 464–469.

Iversen TM (1974) Ingestion and growth in Sericostoma personatum (Trichoptera) in relation to the nitrogen content of ingested leaves. *Oikos*, **25**, 278–282.

Kearney M (2005) Hybridization, glaciation and geographical parthenogenesis. *Trends in Ecology & Evolution*, **20**, 495–502.

Kearney M, Blacket MJ, Strasburg JL, Moritz C (2006) Waves of parthenogenesis in the desert: evidence for the parallel loss of sex in a grasshopper and a gecko from Australia. *Molecular Ecology*, **15**, 1743–1748.

Kelly CD, Godin JJ, Abdallah G (2000) Geographical variation in the male intromittent organ of the Trinidadian guppy (Poecilia reticulata). *Canadian Journal of Zoology*, **78**, 1674–1680.

King Sirot L (2003) The evolution of insect mating structures through sexual selection. *Florida Entomologist*, **86**, 124–133.

Law JH, Crespi BJ (2002) The evolution of geographic parthenogenesis in Timema walking-sticks. *Molecular Ecology*, **11**, 1471–1489.

Lleonart J, Salat J, Torres GJ (2000) Removing allometric effects of body size in morphological analysis. *Journal of Theoretical Biology*, **205**, 85–93.

Lyimo EO, Takken W, Koella JC (1992) Effect of rearing temperature and larval density on larval survival, age at pupation and adult size of Anopheles gambiae. *Entomologia Experimentalis et Applicata*, **63**, 265–271.

Lynch M (1984) Destabilizing hybridization, general-purpose genotypes and geographic parthenogenesis. *Quarterly Review of Biology*, **59**, 257–290.

McCafferty WP, Pereira C (1984) Effects of developmental thermal regimes on two mayfly species and their taxonomic interpretation. *Annals of the Entomological Society of America*, **77**, 69–87.

Mutanen M, Kaitala A (2006) Genital variation in a dimorphic moth Selenia tetralunaria (Lepidoptera, Geometridae). *Biological Journal of the Linnean Society*, **87**, 297–307.

Mutanen M, Pretorius E (2007) Subjective visual evaluation vs. traditional and geometric morphometrics in species delimitation: a comparison of moth genitalia. *Systematic Entomology*, **32**, 371–386.

Mutanen M, Rytkönen S, Lindén J, Sinkkonen J (2007) Male genital variation in a moth Pammene luedersiana (Lepidoptera: Tortricidae). *European Journal of Entomology*, **104**, 259–265.

Panov VE, McQueen DJ (1998) Effects of temperature on individual growth rate and body size of a freshwater amphipod. *Canadian Journal of Zoology*, **76**, 1107–1116.

Peck JR, Yearsley JM, Waxman D (1998) Explaining the geographic distributions of sexual and asexual populations. *Nature*, **391**, 889–892.

Raastad JE, Solem JO (1989) Autogeny as successful reproductive strategy in high altitude black-flies (Diptera, Simuliidae). *Annales de Limnologie*, **25**, 243–249.

Rundgren M, Ingólfsson O (1999) Plant survival in Iceland during periods of glaciation ? *Journal of Biogeography*, **26**, 387–396.

Rundle HD, Chenoweth SF, Doughty P, Blows MW (2005) Divergent selection and the evolution of signal traits and mating preferences (J Kingsolver, Ed,). *PLoS Biology*, **3**, 368.

Salokannel J, Rantala MJ, Wahlberg N (2010) DNA-barcoding clarifies species definitions of Finnish Apatania (Trichoptera: Apataniidae). *Entomologica Fennica*, **21**, 1–11.

Schmid F (1953) Contribution à l'étude de la sous-famille des Apataniinae (Trichoptera, Limnophilidae). *Tijdschrift Voor Entomologie*, **96**, 109–167.

Shapiro AM, Porter AH (1989) The lock-and-key hypothesis: evolutionary and biosystematic interpretation of insect genitalia. Annual Review of Entomology, 34, 231–245.

Silva-Brandão KL, Lyra ML, Freitas AVL (2009) Barcoding lepidoptera: current situation and perspectives on the usefulness of a contentious technique. *Neotropical Entomology*, **38**, 441–451.

Solem JO (1985) Norwegian Apatania Kolenati (Trichoptera: Limnephiiidae): identification of larvae and aspects of their biology in a high-altitude zone. *Insect Systematics & Evolution*, **16**, 161–174.

Sullivan RM, Calhoun SW, Greenbaum IF (1990) Geographic variation in genital morphology among insular and mainland populations of Peromyscus maniculatus and Peromyscus oreas. *Journal of Mammalogy*, **71**, 48–58.

Suomalainen E (1962) Significance of parthenogenesis in the evolution of insects. *Annual Review of Entomology*, **7**, 349–366.

Svensson BW (1975) Morphometric variation of adult Potamophylax cingulatus (Trichoptera) reflecting environmental heterogeneity in a South Swedish stream. *Oikos*, **26**, 365–377.

Taylor PD, Merriam G (1995) Wing Morphology of a forest damselfly is relted to landscape structure. *Oikos*, **73**, 43–48.

Tidon-Sklorz R, De Melo-Sene F (1995) Drosophila seriema n. sp.: new member of the Drosophila serido (Diptera: Drosophilidae) superspecies taxon. , 139–142.

Zeh DW., Zeh JA. (1994) When morphology misleads: interpopulation uniformity in sexual selection masks genetic divergence in harlequin beetle-riding pseudoscorpion. *Evolution*, **48**, 1168–1182.

Conclusion

The main finding of this study is that *Apatania zonella* in Iceland originates from more than one colonization event, with at least one from North-America through Greenland and one from Europe. *A. zonella* populations may have survived in unglaciated areas close to the ice sheet and in the unglaciated area in Beringia. At the time of the deglaciation, individuals from these areas contributed to the recolonization of Asia and Europe from one side and North-America and Greenland on the other, and eventually intermixed within Iceland. Populations in North-West Iceland and South Iceland appear to have a Palearctic origin while the population in South-West Iceland was likely colonized by individuals from Greenland. Overall, the data provided by this study suggest that the population structure of *A. zonella* in Iceland was shaped by the late Quaternary Ice Age and the post glaciation recolonization process.

Another interesting result from this study is the lack of monophyly of *A. zonella*. The distinct lineages in North-America and Europe are genetically more different than different *Apatania* species within Europe. Indeed three *Apatania* species are nested within *A. zonella*. This result questions the actual species classification within *Apatania* genus, and implies that the taxonomic status of this group should be reevaluated.

Finally, the morphological variations of the genital features allow a good discrimination of the different populations into morphological and geographical groups. The results show greater morphological differences, especially in overall size, between *A. hispida* and *A. zonella* than among *A. zonella* between the Greenlandic and Icelandic populations. The morphological variation in *A. zonella* may reflect different responses to variable ecological conditions, but also the post-glacial recolonization with founder effect, following reproductive isolation of geographically separated populations during the glacial periods of Ice age.

Appendix 1.Summary Table of the list of samples: collection sites, geographic coordinates (given in decimal degrees) and number of individuals per collection site (n).

Country	Area	Location	Geographic coordinates	Date or Reference	n	Species
United-States	Minnesota	Lake Superior, Beaver Bay	47° 16' 13.0872"N; 91° 16' 32.1492"W	Kjer et al. (unpublished)	6	*Apatania zonella*
Canada	Ontario	Pukaskwa Natural Park	48° 35' 24"N; 86° 17' 24.0036"W	Kjer et al. (unpublished)	2	*Apatania zonella*
	Yukon Territory	Haines Highway, Dezadeash Lake	60° 22' 1.2036"N; 137° 3' 0.0108"W		4	*Apatania zonella*
Finland	Lapland	Enontekiö, Eanodat	68° 23' 0.5778"N; 23° 38' 20.9214"E	Salokannel et al. (2010)	4	*Apatania zonella*
		Utsjoki	69° 49' 55.146"N; 27° 1' 15.1458"E		3	*Apatania zonella*
					1	*Apatania hispida*
		Kittilä, Pallastunturi	68° 8' 45.5712"N; 24° 13' 49.2486"E		1	*Apatania forsslundi*
	Pirkanmaa	Valkeakoski, Hiuspalvelu Apia	61° 16' 6.8952"N; 24° 2' 30.552"E		1	*Apatania auricula*
	Satakunda	Jämijärvi, Uhrilähde	61° 46' 52.0242"N; 22° 46' 8.6844"E		1	*Apatania dalecarlica*
Iceland	North-West Iceland	Westfjords, Ketilseyrará	65° 51' 11.1744"N; 23° 20' 32.64"W	Kjer et al. (unpublished)	2	*Apatania zonella*
		Westfjords, River Bjarnardalsa	65° 47' 57.483"N; 21° 20' 9.9594"W	Kjer et al. (unpublished)	1	*Apatania zonella*
		Westfjords, Hvalskurdara, Skoetufjoerdur	65° 53' 56.0508"N; 22° 49' 49.44"W	Kjer et al. (unpublished)	1	*Apatania zonella*
		Westfjords, Aðalvik staðará	66° 21' 35.9994"N; 23° 1' 8.7738"W	01/10/2008	3	*Apatania zonella*
		Snæfellsnes, Staðará	64° 49' 8.4"N; 22° 59' 9.6"W	28/06/2006	1	*Apatania zonella*
		Westfjords, Bjarnadalsá	65° 58' 22.7994"N; 23° 26' 9.5994"W	28/06/2006	1	*Apatania zonella*
		Westfjords, Ketilseyrará	65° 51' 10.7994"N; 23° 20' 31.1994"W	28/06/2006	1	*Apatania zonella*
		Westfjords, Á hjá Fremri-Breidadal	65° 52' 27.933"N; 23° 25' 46.2108"W	28/06/2006	1	*Apatania zonella*
		Westfjords, Patreksfjörður	65° 35' 24"N; 23° 58' 26.4"W	28/06/2006	1	*Apatania zonella*
		Westfjords, Arnardalsá	66° 5' 45.6"N; 23° 3' 7.1994"W	28/06/2006	1	*Apatania zonella*
	South-West Iceland	Reykjavík, Vífilsstaðavatn	64° 4' 44.2302"N; 21° 52' 41.826"W	26/04/2012	12	*Apatania zonella*
	South Iceland	Tumastaðir	63° 43' 52.2222"N; 20° 3' 18.5616"W	09/07/2010 - 10/09/2010	5	*Apatania zonella*
		Rauðafell	63° 32' 58.8984"N; 19° 36' 2.9592"W	09/07/2010 - 09/09/2010	4	*Apatania zonella*
		Ytri Rangá	63° 50' 9.5994"N; 20° 24' 14.3994"W	14/07/2005	2	*Apatania zonella*
		Hofsá	63° 31' 19.221"N; 19° 30' 11.9046"W	14/07/2005	1	*Apatania zonella*
		Blautakvísl	63° 28' 15.5994"N; 18° 35' 31.1994"W	15/07/2005	1	*Apatania zonella*
		Núpsstadir	63° 59' 42.6078"N; 17° 28' 33.6288"W	15/07/2005	1	*Apatania zonella*
Greenland	Tunulliarfik Fjord	Brattahlíð	61° 10' 29.481"N; 45° 47' 27.657"W	Kjer et al. (unpublished)	11	*Apatania zonella*
		Narsarsuaq	61° 6' 45.759"N; 45° 24' 23.3784"W	Sanz Muñoz 2010 (unpublished)	3	*Apatania zonella*
		Garðar	60° 57' 18.3168"N; 45° 24' 53.0424"W	Sanz Muñoz 2010 (unpublished)	1	*Apatania zonella*
Alaska	North Slope Borough	Galbraith Lake	68° 27' 57.8154"N; 149° 24' 53.0964"W	Sanz Muñoz 2010 (unpublished)	1	*Apatania zonella*
Norway	Oppland	Dovre, Rondane National Park	61° 58' 54.8358"N; 9° 48' 16.3434"E	Kjer et al. (unpublished)	2	*Apatania zonella*
Mongolia	Bayan-Ölgiy	Tsengel	49° 0' 10.7922"N; 88° 31' 58.7886"E	Kjer et al. (unpublished)	1	*Apatania zonella*

Appendix