# Population genetic structure and phylogeography of invasive aquatic weed, *Elodea canadensis* (Hydrocharitaceae) and comparative analyses with *E. nuttallii*

**Tea Huotari**

Department of Agricultural Sciences
Faculty of Agriculture and Forestry

University of Helsinki
Finland

ACADEMIC DISSERTATION

*To be presented, with the permission of the Faculty of Agriculture and Forestry of the University of Helsinki, for public criticism in Auditorium 1041, Biocenter 2 (Viikinkaari 5, Helsinki), on October 5th, 2012, at 12 noon.*

HELSINKI 2012

Supervised by: Dr Helena Korpelainen
Department of Agricultural Sciences
University of Helsinki, Finland

Dr Elina Leskinen
Department of Environmental Sciences
University of Helsinki, Finland

Reviewed by: Dr Jouni Aspi
Department of Biology
University of Oulu, Finland

Dr Alain Vanderpoorten
Department of Life Sciences
University of Liége, Belgium

Examined by: Prof. Katri Kärkkäinen
The Finnish Forest Research Institute Oulu, Finland

Custos: Prof. Teemu Teeri
Department of Agricultural Sciences
University of Helsinki, Finland



ISBN 978-952-10-8258-0 (PAPERBACK)
ISBN 978-952-10-8259-7 (PDF)

Yliopistopaino
Helsinki, Finland 2012

Äidille

# List of original publications

**THIS THESIS IS BASED ON** the following publications and a manuscript, which are referred to in the text by their Roman numerals:

**I** Huotari, T., Korpelainen, H. and Kostamo, K. 2010. Development of microsatellite markers for the clonal water weed *Elodea canadensis* (Hydrocharitaceae) using inter-simple sequence repeat (ISSR) primers. – Molecular Ecology Resources 10: 576–579.

**II** Huotari, T., Korpelainen, H., Leskinen, E. and Kostamo, K. 2011. Population genetics of invasive water weed *Elodea canadensis* in Finnish waterways. – Plant Systematics and Evolution 294: 27–37.

**III** Huotari, T. and Korpelainen, H. 2012. Complete chloroplast genome sequence of *Elodea canadensis* and comparative analyses with other monocot genomes. – Gene 508: 96–105.

**IV** Huotari, T. and Korpelainen, H. 2012. Comparative analyses of plastid sequences between native and introduced populations of aquatic weeds *Elodea canadensis* and *E. nuttallii*. – Submitted.

# Contributions

**THE FOLLOWING TABLE** shows the major contributions of authors to the original articles or manuscripts.

| | **I** | **II** | **III** | **IV** |
|---|---|---|---|---|
| Original idea | HK, KK | HK, KK | TH, HK | TH, HK |
| Field work | TH, KK | TH,KK | TH | TH |
| Molecular data | TH | TH | TH | TH |
| Data analyses | TH | TH | TH | TH |
| Manuscript preparation | TH, HK, KK | TH, HK, KK, EL | TH, HK | TH, HK |

TH = Tea Huotari, HK = Helena Korpelainen, KK = Kirsi Kostamo, EL = Elina Leskinen

# Contents

# Abstract

**THE INTRODUCTIONS OF** invasive species are one of the most important threats to global biodiversity and ecosystem function. In addition to severe environmental impacts, invasive species often cause large economical and social consequences. Genetic characteristics of introduced populations have an impact on their capacity of range expansion in the non-native areas. Therefore, understanding the evolutionary consequences of invasions will provide knowledge for the design of appropriate methods for managing introduced populations. In addition, detailed genetic data enables the design of molecular genetic markers useful in monitoring risky species and in early detection of new invasions.

In this thesis, I developed novel genetic markers to investigate population genetic structure and phylogeography of two invasive aquatic weeds, *Elodea canadensis* Michx and *E. nuttallii* (Planch.) St. John (Hydrocharitaceae). Furthermore, these markers offer a valuable tool in species identification, as both *E. canadensis* and *E. nuttallii* show a wide range of morphological variation and, therefore, are difficult to discriminate.

I investigated the genetic population structure of introduced Finnish *E. canadensis* populations using microsatellite markers. The results revealed a moderate level of variation within and among Finnish populations analysed. This result indicates that *E. canadensis* could have been introduced to Finland more than once, considering the asexual reproduction mode and recent introduction of this species. However, the possibility of only one introduction followed by post-establishment evolution cannot be rejected based on the results in this thesis. Furthermore, I surveyed the geographical distribution of the chloroplast (cp) DNA haplotypes within the native and introduced ranges of *E. canadensis* and *E. nuttallii* in order to reconstruct the spreading histories of these species. Only a single haplotype was found in the introduced range in both species and these haplotypes were widespread also in the native range. Therefore, I was not able to identify either the geographic origin of the introduced populations or test the hypothesis of single versus multiple introductions. Moreover, the low level of cpDNA variation detected in the introduced range further supports one introduction event to Finland.

I sequenced the complete cp genome sequence of *E. canadensis*, characterized the cp genome organization of this basal monocot, and evaluated the level of similarity among monocot cp genomes. The results showed that the cp genome of *E. canadensis* has gone through less rearrangements or gene losses when compared to assumed ancestral species than have the other monocots studied. The inverted repeat region (IR) of *E. canadensis* has a unique structure among the monocot species studied so far. Only few cp genomes representing early lineages of monocots have been sequenced and, therefore, this thesis provides valuable information about the course of evolution in the divergence of monocot lineages.

This thesis addresses key issues in the biology of invasive species and gives novel information on the population genetics of *Elodea* species, and on the genetic patterns at native and introduced ranges of *E. canadensis* and *E. nuttallii*. It also provides a source of genetic markers for future investigations of the population genetics of invasive species. The results highlight the need for further investigation of the *Elodea* species, as well as studies on other invasive plant species. Future research should focus on predictive analyses of potential future invaders and other preventive methods to minimize new introductions.

# 1 Introduction

**NON-INDIGENOUS SPECIES** (NIS) are species distributed outside their historic and native range. Dispersal of NIS may occur either intentionally or accidentally, both being promoted by many human activities, such as agriculture, aquaculture, recreation and transportation (Kolar & Lodge 2001). Currently, the rate of invasions by plants and other organisms is accelerating (Mack et al. 2000; Levine & D´Antonio 2003; Lockwood et al. 2008), and warming climate will further increase the probability of invasions, especially in boreal regions, where the northern range limit is typically determined by minimum winter temperatures (Heino et al. 2009). Only a small fraction of NIS is able to spread widely and become invasive (Williamson & Fitter 1996). Nevertheless, together with habitat loss and fragmentation, the introductions of invasive species are one of the most important threats to global biodiversity and ecosystem function (Walker & Steffen 1997; Wilcove et al. 1998; Sala et al. 2000). In addition to severe environmental impacts (Vitousek 1990; Vitousek et al. 1997), invasive species often cause large economical and social consequences (Sakai et al. 2001; Cox 2004). Invasive species are, no doubt, a threat, but they also offer significant opportunities to study basic processes in population biology. The consequences of biological invasions allow studies of basic evolutionary processes, as invaders often evolve rapidly in response to novel abiotic and biotic conditions, and native species evolve in response to the invasion (Sakai et al. 2001).

In order to become invasive, a species must go through three phases (Fig 1.) (Heger & Trepl 2003; Williamson 2006; Lockwood et al. 2008). 1) Individuals of the species must have dispersed naturally or human-assisted from their native range to a new area. 2) After the introduction, they must survive and reproduce within the non-native range in order to establish. 3) Once established, the species must increase in number and expand its geographic range. In the following chapter, I will demonstrate these phases more thoroughly.

## 1.1 From non-indigenous to invasive

One of the most important life history processes for invasive species is dispersal. All species possess a life history stage adapted to dispersal, such as spore, seed, egg, larva or mature organism. Furthermore, the most successful invaders combine effective mechanisms of natural and human-assisted dispersal (Cox 2004). Successful invaders often share characteristics, which help them to overcome several barriers they face in order to disperse into new areas, such as geographical barriers, severe abiotic conditions, biotic interactions and landscape factors (Prentis et al. 2008; Heikkinen et al. 2009). These characteristics include, *e.g.*, long lasting life cycle stages, wide tolerance capability, and active and passive transportation ability (Heikkinen et al. 2009).

In addition to effective dispersal ability, successful invaders must be adapted to rapid establishment in new environments, as only 10% of NIS is able to establish, survive and reproduce in the invaded ecosystem (Kolar & Lodge 2001). For this purpose, they need to possess mechanisms for quickly capturing the resources required for growth and reproduction (Cox 2004). One advantage for an exotic species in a novel environment is clonality, as one sex can start to reproduce and spread immediately after the introduction (Pyšek 1997). In fact, there is evidence of a positive relationship between plant invasiveness and the occurrence of vegetative reproduction (Kolar & Lodge 2001). There are several examples of

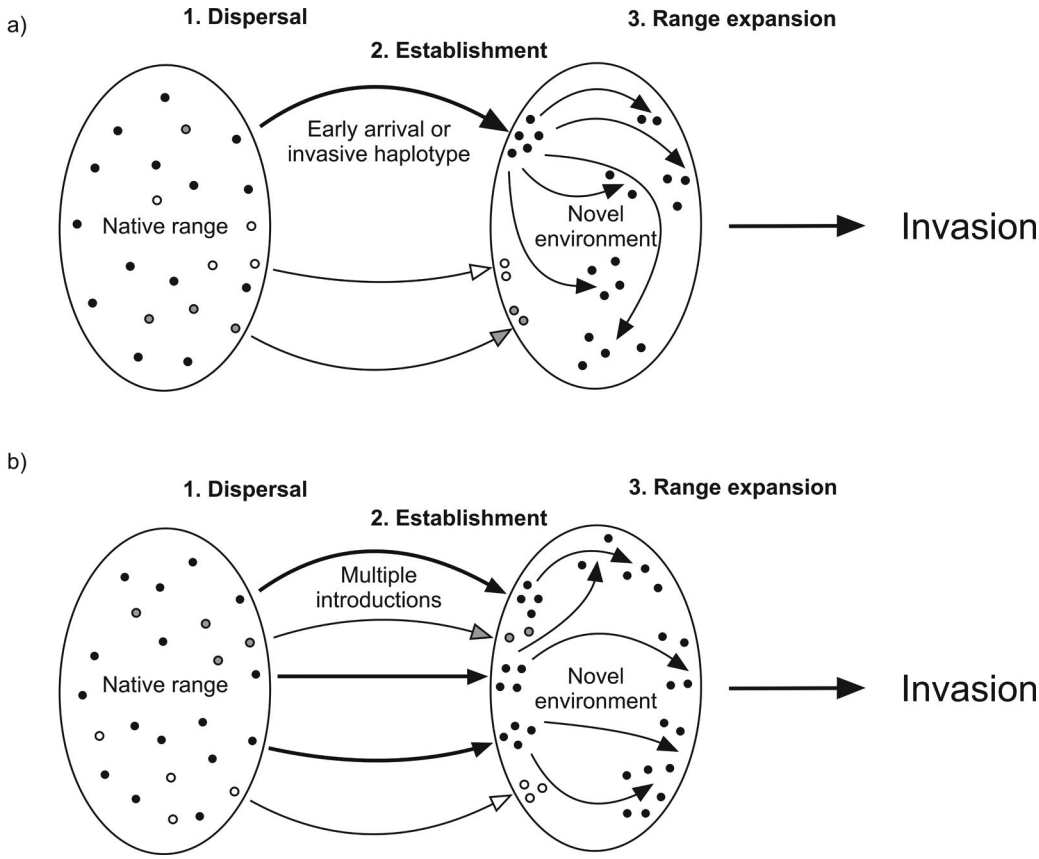


**Figure 1** Steps and stages a non-indigenous species passes in order to become invasive.
a) One haplotype may be more invasive than others or arrive earlier into a novel environment, thus having an advantage in the competition at the introduced area.
b) One haplotype may have larger propagule pressure due to multiple introductions or higher number of individuals released and, therefore, become more successful at the introduced area. Each colour represents one haplotype of the same species.

mainly asexually reproducing aquatic plant species, in which only one sex has been introduced, yet the species has managed to spread successfully (Philbrick & Les 1996; Sakai et al. 2001). In general, the probability of establishment increases with increasing propagule pressure. Propagule pressure is a combination of the number of individuals released at a release event (propagule size) and the number of discrete release events (propagule number) in a non-native region (Lockwood et al. 2005; Simberloff 2009).

Once established, a species must increase in number and spread in order to become invasive. A species can be considered invasive in one location but not in another (Lockwood et al. 2008). Therefore, in order to minimise new introductions of invasive species, the most effective management efforts are the predictive and preventive actions (Ricciardi & Rasmussen 1998). It has been suggested that effects of many invasive species have an acute phase, right after the introduction, followed by a chronic phase after ecological and

evolutionary processes have passed, which represents the eventual outcome of the invasion. However, even the acute phase might last long enough to have serious ecological and economic consequences (Strayer et al. 2006).

## 1.2 Aquatic plant invasions to fresh water environment

Among all habitat types, freshwater systems are particularly sensitive to invasions (Shea & Chesson 2002). Reasons for this are that aquatic environments are homogeneous on a large spatial scale, which makes it possible for aquatic plants to survive and establish more easily outside their native geographic range. Aquatic environments are also difficult to monitor, and an early detection of introduction of a submerged species is seldom possible. In addition, water is a very effective vector of propagules, and seeds and vegetative fragments are easily dispersed over long distances (Barrett et al. 1993; Santamaria 2002; Larson 2007). Over all, aquatic environment is complex, several factors influence at different levels and times and, therefore, predicting possible invaders and their impacts is difficult.

Among aquatic plant species, the main transport vectors are shipping, including ballast water transport, transport of ornamental plants, fishing gear and aquaculture (aquarium trade), and unintentional release or escapes (Lockwood et al. 2008). Most aquatic plant species introduced from outside their native range do not become invasive. Those that are able to establish and spread, have a potential to cause large changes in aquatic ecosystems at both ecosystem and population level. Mass occurrences of aquatic weeds may affect native plant species by competition or habitat alteration as well as block waterways, alter hydrological cycling, and change such basic fresh water ecosystem functions like nutrient cycling (Gordon 1998; Cox 2004; Strayer et al. 2006).

## 1.3 Genetic consequences of invasions

Plant invasions may involve a number of evolutionary phenomena, such as founder effect, genetic drift, hybridization and adaptation to new environments (Lee 2002; Frankham et al. 2003). These factors may result in rapid genetic changes causing genetic differentiation between native and introduced populations (Lee 2002; Müller-Schärer et al. 2004; Bossdorf et al. 2005).

Genetic characteristics of introduced populations have an impact on their capacity of range expansion in the non-native areas (Tsutsui et al. 2000; Lee 2002). Genetic diversity of introduced populations may be reduced due to founder effect, which is the loss of genetic variation that occurs when a new population is established by a small number of individuals from a larger population (Nei et al. 1975; Lee 2002; Frankham et al. 2003). Lower genetic diversity may further reduce the capacity to adapt to novel conditions (Sakai 2001; Lee 2002). Reduced genetic diversity can have additional consequences, such as inbreeding depression, which may limit the production of propagules and population growth (Ellstrand & Elam 1993). In small populations, also stochastic processes, such as genetic drift, take place more easily. On the other hand, genetic diversity may increase by multiple introductions providing new genotypes (Frankham 2005), which may create additional variation by recombination in sexually reproducing populations (Ellstrand & Schierenbeck 2006).

After introduction, differentiation among introduced populations may be caused by somatic mutations. The level of molecular divergence due to novel mutations may differ among populations, partly because novel mutations may occur in response to stressful environmental conditions (McClintock 1983; Gill et al. 1995). Invasion may also lead to spatial sorting, where alleles that facilitate the highest rate of dispersal accumulate at the expanding range edge even in the absence of conventional natural selection (Shine et al. 2011). Genetic variability in non-indigenous populations is also affected by the breeding system of the species (Sakai et al. 2001; Silvertown & Charlesworth 2001). Furthermore, asexual reproduction

can complicate interpretations of genetic diversity estimates because multiple plants derived from asexual reproduction have identical genotypes (Barret et al. 1993).

In order to control the dispersal and effects of invasive species, it is important to know the origin of invasive taxa, the size of the introduction, the level of genetic variation compared to the native range, and whether multiple introductions have occurred. These questions can be answered by using molecular genetic techniques (Schaal et al. 2003; Bossdorf et al. 2005; Hufbauer & Sforza 2008; Ward et al. 2008). Molecular genetic markers provide a simple way to estimate allele frequencies, genetic diversity and genetic distances between populations (Avise 2004). Molecular methods can provide important information for predicting population responses to management and preventing new invasions, for tracking introduction routes and solving the mechanisms of local dispersal and adaptation (Ward et al. 2008). In recent years, the number of genetic markers available for evolutionary studies has increased due to improvements in the efficiency, cost and accuracy of high-throughput sequencing technologies. In the following, I introduce the properties of the genetic markers used in this thesis.

Microsatellites or simple sequence repeats (SSRs) are DNA sequences consisting of short (1–6 bp) tandemly repeated motifs assumed to be randomly distributed throughout the nuclear, chloroplast and mitochondrial DNA (Goldstein & Schlötterer 1999; Zane et al. 2002). They usually represent selectively neutral markers, which often show high levels of length polymorphism, thus enabling fine-scale population genetic studies (Goldstein & Schlötterer 1999; Hedrick 1999; Ellegren 2004). SSRs have been successfully used to analyse the genetic diversity of native and introduced populations, to detect recent bottlenecks (Meimberg et al. 2006) and to reveal the introduction history of many invasive plant species, (*e.g.* Walker et al. 2003; Durka et al. 2005; Williams et al. 2005; Besnard et al. 2007). As a downside, technical problems, such as non-amplifying alleles, problems in allele scoring and DNA sample quality, may introduce bias to the inferences based on the SSR markers (Bonin et al. 2004; Pompanon et al. 2005).

Several studies have successfully utilized chloroplast (cp) DNA diversity when revealing spreading routes and possible sources of introductions on invasive species (*e.g.* McIvor et al. 2001; Saltonstall 2002; Trewick et al. 2004; Gaskin et al. 2005; Williams et al. 2005; Hufbauer & Sforza 2008; Pock Tsy et al. 2009). The gene order and content of cp genomes are generally highly conserved and the substitution rate in cpDNA is much lower than that in nuclear DNA in plants (Wolfe et al. 1987; Korpelainen 2004). Chloroplasts are usually inherited maternally and transmitted only through seeds, and hence, cpDNA has less potential for gene flow than nuclear genes, which are spread also by pollen dispersal. Also, the effective population size $N_e$ of plastids is only one quarter of the value for nuclear genes (Charlesworth 2009). Consequently, the genetic variation of the cp genome is often more geographically structured than that of the nuclear genome, making chloroplasts valuable sources of genetic markers for intraspecific phylogeographic analyses (Soltis et al. 1997; Provan et al. 2001; Raubeson et al. 2007; Ravi et al. 2008).

On the other hand, the lack of informative loci has been limiting the use of cpDNA in phylogenetic and phylogoegraphic analyses (Schaal et al. 1998; Thiebaut & Di Nino 2009). Although several informative cpDNA loci have been discovered and exploited since the pioneering work in 1991 (Taberlet et al. 1991), many variable loci remain undetected. Shaw et al. (2005; 2007) suggested that the widely used non-coding areas of cp genome might actually be among the least variable, while the most variable regions are rarely exploited. The number of completely sequenced cp genomes is growing rapidly (Ravi et al. 2008), and lower sequencing costs enable genome-wide intraspecific comparative studies, which aid in detecting the most informative genomic regions for further analyses. There are also risks in rely-

ing exclusively on organelle DNA, considering the frequency of chloroplast capture, where the cytoplasm of one species is replaced by that of another species through hybridization or introgression, and which has been reported to occur in natural ecosystems (Riesenberg & Soltis 1991; Fehrer et al. 2007). On the other hand, cpDNA markers are required in order to understand the evolutionary history of the uniparental organelles and to distinguish maternal and paternal lineages in hybridizations (Yuan & Olmstead 2008).

### 1.4 Genus *Elodea*

*Elodea* (Hydrocharitaceae) is a New World genus with at least five submerged aquatic angiosperm species living in fresh water environments. *Elodea canadensis* Michx, *E. nuttallii* (Planch.) St. John and *E. bifoliata* are native to temperate North America, while *E. potamogeton* and *E. callitrichoides* are native to South America (Cook & Urmi-König 1985). Of the species native to North America, *E. canadensis* and *E. nuttallii* have been introduced to several other continents (Cook & Urmi-König 1985; Bowmer et al. 1995). *Elodea canadensis* was brought to Europe in 1836, first to the United Kingdom (Sculthorpe 1967), and at present it is widespread in the whole Europe. It was introduced to New Zealand in 1868 (Chapman 1970), to Australia in 1931 (Aston 1973), and it has been considered a noxious weed also in many regions of Asia and Africa (Bowmer et al. 1995). *Elodea nuttallii* was first reported in Europe in 1939, in Belgium (Simpson 1984). It has not yet been found in the northernmost Europe, but it is likely to spread to new areas with a high risk of being invasive (Simpson 1984; Larson 2007). *Elodea nuttallii* was introduced to Japan in 1961 (Kadono 2004) and to China in 1980 (Xu et al. 2007), while it has not yet been found in Australia or New Zealand. Both *Elodea* species were reputedly brought to Europe as aquarium plants or with timber (Cook & Urmi-König 1985). Once in Europe, water currents and birds have spread these plants locally, while botanists and botanic gardens were responsible for their long distance dispersal (Cook & Urmi-König 1985; Galera & Sudnik-Wójcikowska 2010).

*Elodea canadensis* was introduced to Finland in 1884 to be displayed at the Botanical Garden of the University of Helsinki (Hintikka 1917). After the introduction some individuals were intentionally planted and some escaped into the environment, after which they rapidly spread across the country. At present, *E. canadensis* is common in the Southern and Central Finland and has lately invaded many lakes in the Kuusamo district in the Northeastern Finland. Since the first introduction, it has been argued whether the whole Finnish population originates from this first introduction event.

Both *Elodea* species have attracted great attention due to their invasive nature in the non-native range (Cook & Urmi-König 1985; Bowmer et al. 1995). They form dense stands, which may change the balance of lake and river ecosystems by outcompeting native species and by changing the pH and nutrient levels, as well as by reducing the oxygen concentration of the water column. In addition, mass occurrences of *Elodea* can make the recreational use of lakes difficult (Simpson 1984; Bowmer et al. 1995; Sarvala 2005). After rapid dispersal in Europe, there was a subsequent decline of *E. canadensis* in several locations, but new invasions are occurring at least in its northern distribution limit (Heikkinen et al. 2009; S. Hellsten 2012, pers. comm., 31 Sep.). In several European locations, *E. nuttallii* has been reported to displace already well-established *E. canadensis* populations by its faster growth rate (Simpson 1990; Barrat-Segretain & Arnaud 2004).

In both native and introduced *Elodea* populations, males and females are rarely found in the same population (Cook & Urmi-König 1985), and the dominating mode of reproduction is vegetative by way of fragmentation or specialized buds (Bowmer et al. 1984; Les 1988). Both *E. canadensis* and *E. nuttallii* have a branching stem up to three meters long. Both species have oblong stalk-less leaves in whorls of three. The leaves of *E. nuttallii* are often strongly recurved and typi-

cally narrow with sharp tips (Fig. 2), whereas the leaves of *E. canadensis* are typically broader with blunt tips (Fig. 3) (Cook & Urmi-König 1985). Both species show a wide range of morphological variation, which makes the species identification difficult (Simpson 1984; 1988; Vanderpoorten et al. 2000; Thiébaut & Di Nino 2009). No specific molecular markers have been formerly available for *Elodea* which has impeded population level analyses. Previously, introduced populations of *E. canadensis* have been surveyed using AFLP (amplified fragment length polymorphism) markers (Vanderpoorten 2000; Lambertini et al. 2010).

### 1.5 Aims of the study

The general objective of this study is to increase knowledge of the genetic diversity, population structure and spreading history of two invasive water weeds, *Elodea canadensis* and *E. nuttallii*, as well as to contribute to the general understanding of evolutionary consequences of invasions. Introductions of invasive species have occurred recurrently throughout the history, but still the genetic consequences of range expansions have been little investigated (Excoffier et al. 2009). This knowledge is needed for the design of appropriate methods for managing introduced populations as well as for preventing new invasions and tracking introduction routes. A further aim is to develop novel molecular markers for population level, phylogeographic and phylogenetic research, and reliable tools for species identification between *Elodea* species. The research is composed of four interlinked substudies, hereafter referred to by their Roman numerals **I–IV**.

**Figure 2** Fragments of *Elodea nuttallii*. The leaves are in whorls of three, often strongly recurved and typically narrow with sharp tips.

**Figure 3** Fragments of *Elodea canadensis*. The leaves are in whorls of three, typically slightly boarder than in *E. nuttallii* with blunt tips.

In the substudy **I**, I developed and tested ten polymorphic and presumably neutral microsatellite markers (SSR) for *Elodea canadensis*. The objectives of the substudy **II** were to clarify the dispersal history of *E. canadensis* to Finland and to analyze the genetic population structure of Finnish *E. canadensis* populations using the SSR markers developed in the substudy **I**. In the substudy **III**, I sequenced the complete chloroplast genome sequence of *E. canadensis*, characterized the chloroplast genome organization of this basal monocot, and evaluated the level of similarity among monocot chloroplast genomes. The aim of the substudy **IV** was to examine the level of variation in the plastid sequence in native and introduced populations of *E. canadensis* and *E. nuttallii*, and to survey the geographical distribution of cpDNA haplotypes in order to reconstruct the spreading histories of *E. canadensis* and *E. nuttallii*.

# 2 Material and methods

## 2.1. Sampling and populations studied

The plant material collected covers widely the native and introduced distributions of *E. canadensis* and *E. nuttallii.* Samples for population genetic analyses (**II**) were collected from seven introduced *E. canadensis* populations in Finland (n = 183) (Fig. 4). Plant materials for the cp genome sequencing (**III**) were collected from native population in Sucker Creek in Minnesota United States of America, and from introduced population in Ditch Nurmijärvenoja in Southern Finland (Fig. 5). Samples for the phylogeographic analyses (**IV**) were collected from seven native and 17 introduced *E. canadensis* populations (n = 61), and 10 native and three introduced *E. nuttallii* populations (n = 13) (Fig. 5). In addition, herbarium specimens representing nine *E. canadensis* (n = 9) and 12 *E. nuttallii* (n = 12) populations from North America were obtained from herbarium collections (Fig. 5). Sampling of all populations is described in more detail in papers **II**, **III** and **IV**.

## 2.1. Microsatellite marker development and analyses of population genetic structure

I developed ten polymorphic microsatellite markers for *E. canadensis*. The microsatellite loci were discovered using genomic screening with inter-simple sequence repeat (ISSR) primers (Korpelainen et al. 2007) to locate genomic areas with a high microsatellite frequency. Specific primers were designed for both sides of the microsatellites detected within the sequenced ISSR amplification products (**I**). These markers were used to analyze samples collected from the Finnish *E. canadensis* populations (Fig. 4), in order to investigate the genetic characteristics of introduced populations and to clarify the dispersal history of *E. canadensis* in Finland (**II**). Primer design, PCR reactions and genotyping are described in more detail in papers **I** and **II**.

The genetic characteristics estimated of the introduced Finnish *E. canadensis* populations include the observed ($H_O$) and expected ($H_E$) heterozygosities, the mean number of alleles ($A$) and the level of inbreeding ($F_{IS}$, Weir & Cockerham 1984) (**II**). Genetic differentiation was calculated using an allele frequency estimator ($F_{ST}$) and an estimator taking into account the contribution of stepwise mutations ($R_{ST}$). In addition, the mean genetic diversity ($H_S$, Nei et al. 1975) over loci was calculated in the Finnish populations (**II**). I identified all specimens sharing an identical multilocus genotype (MLG), and calculated the $P_{SEX}$ values (probability that two individuals would share the same MLG by chance) for all MLGs found more than once in each population (**II**). One copy of each MLG with a significant $P_{SEX}$ value was retained for further analyses. If duplicated MLGs are simply removed from the analyses of clonal populations, rare alleles may be overrepresented and common alleles underrepresented in the data, as common alleles may occur by chance and form identical genotypes (Ivey & Richards 2001).

The distribution of genetic variation in introduced populations was examined using the analysis of molecular variance (AMOVA) (**II**). In order to determine the number and distribution of genetic clusters in the Finnish data set, and to reveal the number of distinct introductions of *E. canadensis* to Finland, a Bayesian analysis of population structure was performed using STRUCTURE 2.2 (Pritchard et al. 2000) (**II**). Levels of differentiation among introduced *E. canadensis* populations were studied with an assignment test with a Bayesian-based approach (Rannala & Mountain 1997) using GENECLASS2

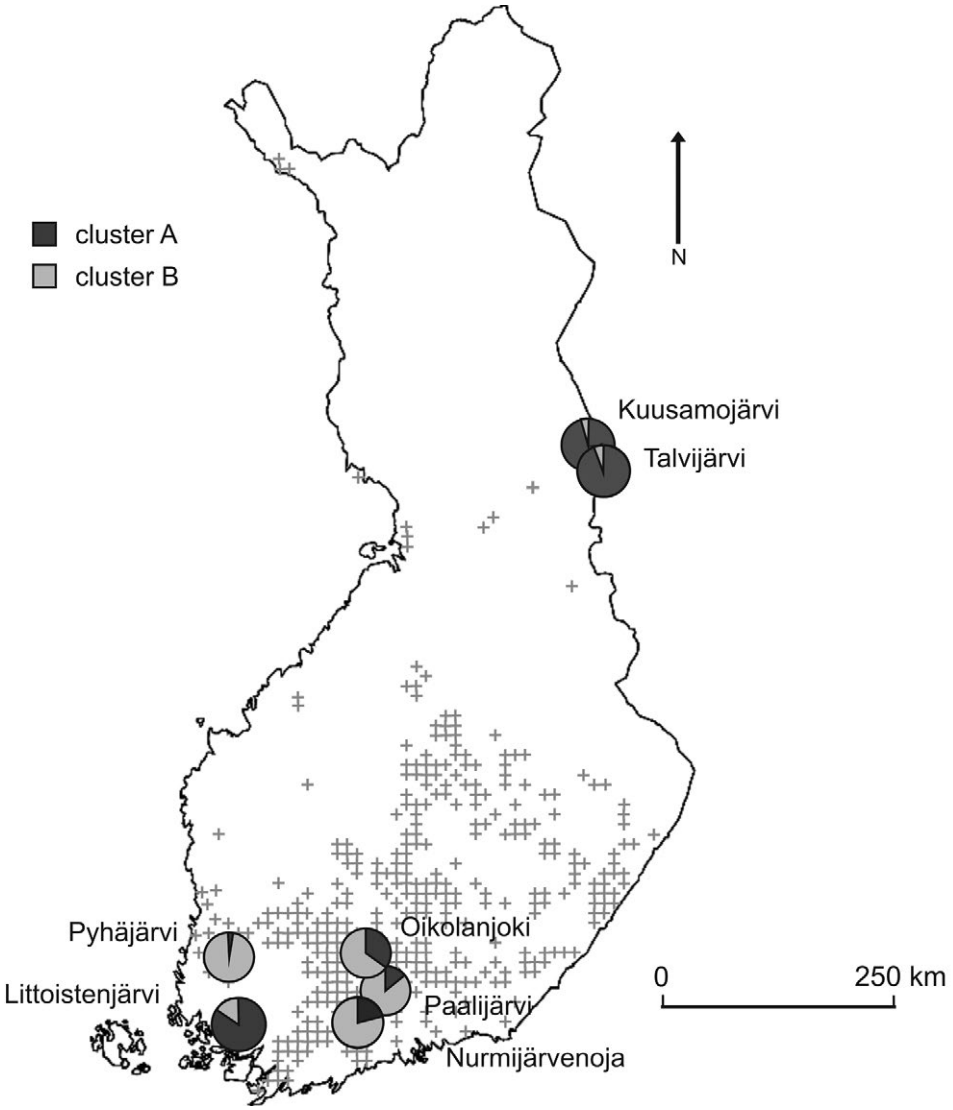


**Figure 4** The main distribution area of *Elodea canadensis* in Finland, and the sampling locations for the population genetic analyses. Pies represent the Bayesian estimates of population structure in these introduced populations based on ten microsatellite markers. Two colours represents the two clusters (A and B) formed in the analysis (K = 2).

(Piry et al. 2004) (**II**).

To reveal if the amount of genetic differences increases with geographic distance, we tested the isolation by distance by regressing pairwise $F_{ST}$ / (1-$F_{ST}$) (Rousset 1997) against the log-transformed geographical distances (km) between all introduced populations and separately between the five populations located Southern Finland. Significances were tested by a Mantel test with 10000 randomizations conducted using Isolation

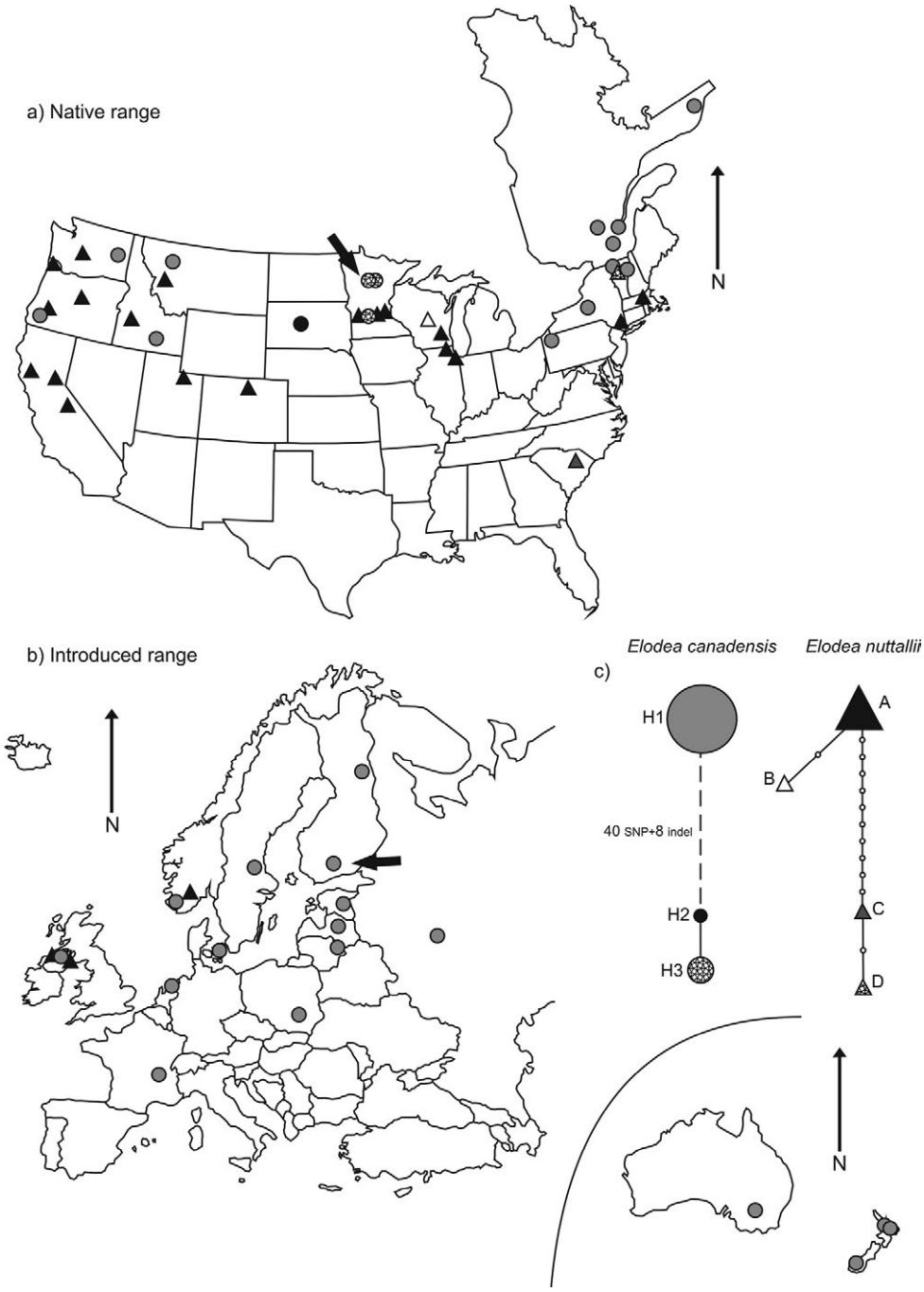


**Figure 5** Sampling locations for the phylogeographic analysis and geographical distribution of cpDNA haplotypes of *Elodea canadensis* (circles) and *E. nuttallii* (triangles). Arrows indicate populations where the plant materials for the chloroplast DNA extraction were collected. a) Sampled populations in the native range in the North America. b) Sampled populations in the introduced range in Europe, New Zealand and Australia. c) Haplotype network showing the relationships among chloroplast DNA haplotypes in *E. canadensis* and *E. nuttallii*. (For details, see Fig. 4 in **IV**)

by Distance web service (Jensen et al. 2005).

## 2.2. Chloroplast genome organization and phylogeographic analyses

I sequenced the whole cp genome of *E. canadensis* from plant material collected from the introduced range (hereafter FIN) (Fig. 6). I was also able to sequence the majority (86.1%) of the cp genome of plant material collected from the native range (hereafter USA). Both cp genomes were sequenced with 454 FLX pyrosequencer (Roche Applied Science). Extraction of cpDNA and the sequencing of cp genomes are described in more detail in paper **III**. Chloroplast-related contigs obtained from both of the pyrosequencings were identified by performing a database search using the BLAST algorithm at the National Center for Biotechnology Information (NCBI), and by aligning with reference cp genomes of other species. Both cp genomes were annotated using DOGMA (Dual Organellar GenoMe Annotator, Wyman et al. 2004) (**III**, **IV**).

The organization of *E. canadensis* FIN cp genome was compared with other monocot chloroplasts using Advanced PipMaker (Schwartz et al. 2000). Monocot cp genomes were also screened for the SSR content using Msatcommader 0.8.2 (Faircloth 2008) and for the number of repeats using REPuter (Kurtz et al. 2001) (**III**). In order to estimate the phylogenetic relationships of early-diverging monocots, two phylogenetic analyses were performed by Bayesian inference (BI) using MrBayes 3.1.2 (Ronquist & Huelsenbeck 2003). These analyses were based on the DNA sequences of 81 chloroplast genes from 17 monocot taxa with *Amborella trichopoda* as an outgroup. Separate analyses were performed for the unpartitioned data set of 81 genes and for the data set divided into 81 partitions (**III**).

In order to detect variable sites between native and introduced *E. canadensis* individuals, the contigs of USA and FIN cp genome sequencings were compared using Gap4 (Staden et al. 2003). Altogether, 29 specific primer pairs were developed for the flanking sequence outside the most variable regions between these plastid sequences (Fig. 6). These cpDNA markers comprised 11,464 bp and contained 78 SNPs, 24 indels and two inversions (**IV**). All these primers were found to work well for amplifying the cpDNA of *E. nuttallii* as well. These markers were used to compare the level of variation in plastid sequence within *E. canadensis* and between *E. canadensis* and *E. nuttallii*. PCR reactions are described in more detail in paper **IV**.

Nine potentially highly variable cpDNA markers comprising 4072 bp (Fig. 6) were used for phylogeographic analyses. These regions were chosen to represent both coding and non-coding cp genome regions, and to contain SNP-, SSR- and indel-variation based on the initial comparison between the two *E. canadensis* plastid sequences USA and FIN. The geographical distributions of the cpDNA haplotypes were analyzed in native and introduced ranges of *E. canadensis* and *E.nuttallii*. Haplotype networks from the sequence data were constructed using the network building software TCS 1.2.1 (Clement et al. 2000).

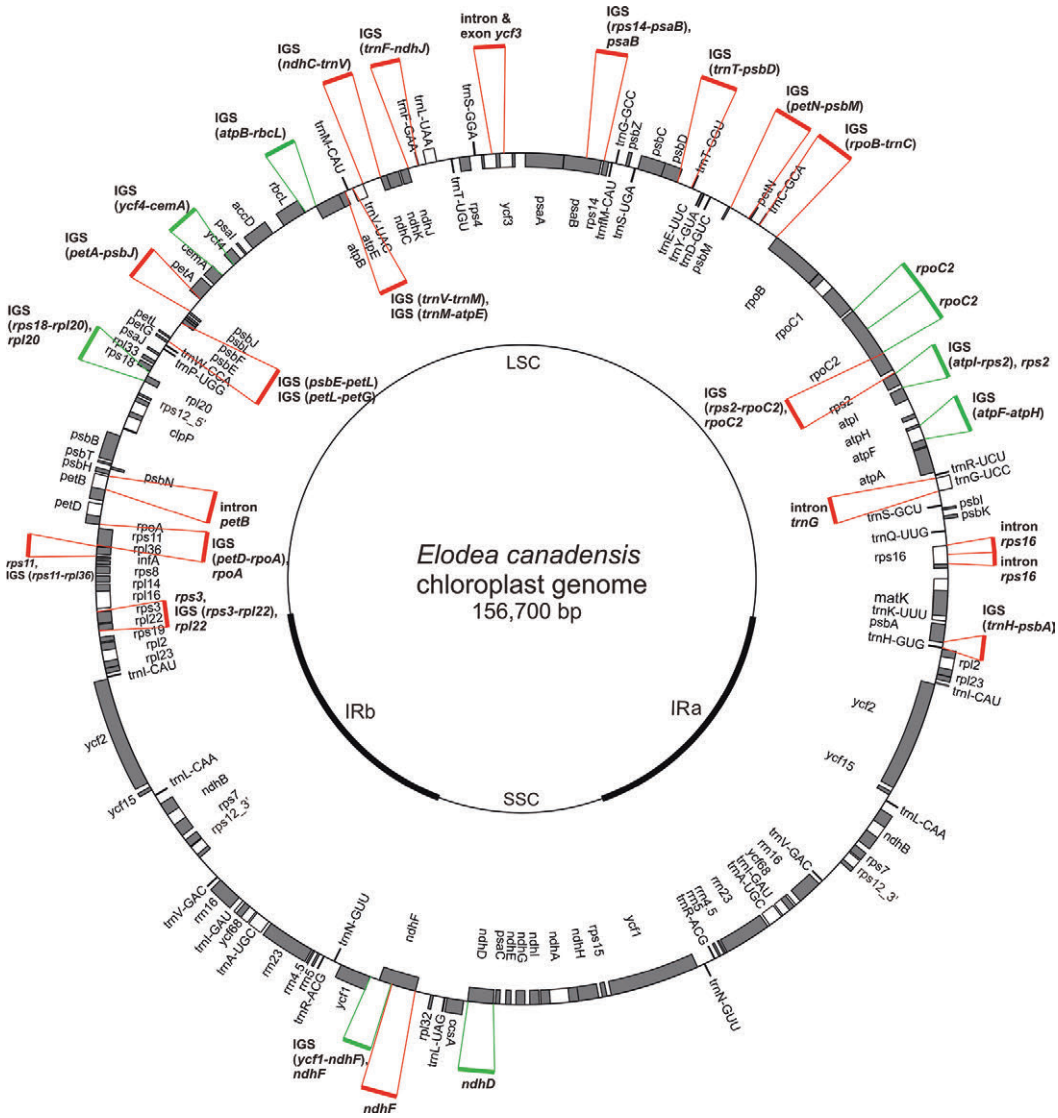


**Figure 6** Gene organization of the *Elodea canadensis* chloroplast genome. Genes shown inside the circle are transcribed clockwise, while those located outside are transcribed counterclockwise. The 29 cpDNA markers developed for variable regions are indicated with red and green colours. The nine markers used in the phylogeographic analysis are indicated with green colour.

# 3 Main results and discussion

## 3.1 Genetic characteristics of the introduced *Elodea canadensis* populations in Finland

Ten polymorphic microsatellite markers developed (**I**) in order to explore the genetic characteristics of native and introduced populations of *E. canadensis* revealed adequate levels of polymorphism to study the genetic structure of introduced *E. canadensis* populations. However, only four out of ten markers amplified reliably in samples collected from the native range. This result could indicate that these microsatellite sequences contain highly variable regions showing sequence variation also within the primer-binding regions. Other option for the misamplification could be the poor quality of the native plant material. The same DNA extracted from the native specimens was, however, used also in the phylogeographic analyses using cpDNA markers without any problems in amplification. In order to reveal the actual reason for the failure in amplification, a new set of primers should be developed located outside the current primer sequences. Alternatively, new microsatellite markers could be developed using native *E. canadensis* plant material.

Ten microsatellite markers were used to analyse samples from seven Finnish *E. canadensis* populations (Fig. 4) in order to reveal the introduction history of *E. canadensis* in Finland (**II**). If there was only one introduction of *E. canadensis* to Finland, a very low level of genetic variation could be expected. On the other hand, multiple introductions could raise the level of genetic variation to the level of the source population or even higher.

All studied populations were found to be multiclonal. Eleven multilocus genotypes (MLG) were shared by more than one sample. Only one of these MLGs possessed a significant $P_{SEX}$ value ($p < 0.05$), indicating a common clonal ancestry. The high genotypic diversity detected indicates surprisingly high levels of genetic variation within introduced populations of *E. canadensis*, albeit it is known that the fast mutation rate in microsatellite markers may even compromise their application in evolutionary studies (Paetkau et al. 1997).

Aquatic vascular plants, in general, show low levels of genetic variation, especially variation within populations has been found to be rather low (Lesica et al. 1988; Barrett et al. 1993; Waycott et al. 1995). On the other hand, microsatellite markers have revealed considerable levels of genetic variation even in aquatic plants reproducing mainly vegetatively. For example in *Zostera marina*, ten alleles were found at one microsatellite locus in a population (Reusch et al. 2000). In my study, the microsatellite markers amplified from two to five alleles per locus, and the mean number of alleles per population varied between 1.6 and 2.4 in seven Finnish populations surveyed. The allele numbers as well as the levels of genetic diversity were rather low ($H_E$ 0.19–0.37) but in agreement with results from other highly clonal aquatic plant populations, such as *Posidonia oceanica* (Procaccini & Mazzella 1998), *Phragmites australis* (Saltonstall 2003; Engloner et al. 2010) and the mainly selfing *Typha* sp. (Tsyusko et al. 2005).

Wrights *F*-statistics is a common tool for describing deviations from random mating (Wright 1951; Balloux et al. 2003). Negative $F_{IS}$ values indicating an excess of heterozygosity relative to random mating were found in three *E. canadensis* populations. Negative $F_{IS}$ values are also suggested to indicate clonality in diploid populations (Balloux et al. 2003; Halkett et al. 2005). In this study, ca. 24% (based on $F_{ST}$) of the genetic varia-

tion present in introduced populations was found among populations. This result is in agreement with those in the clonal *Zostera marina* (30%) (Reusch et al. 2000) and the mainly selfing *Typha angustifolia* and *T. latifolia* (ca. 25%) (Tsyusko et al. 2005). Clonal populations are able to differentiate from each other if gene flow between populations is restricted, for example due to restricted water connection between aquatic plant populations (Schoen & Brown 1991; Barrett et al. 1993). In addition, novel mutations remain in the genome of clonal plants with long lifespans living in isolated populations (deWitte & Stöcklin 2010). It is therefore likely that the asexual reproduction of isolated freshwater populations of *E. canadensis* has resulted in the substantial divergence in microsatellite allele frequencies. The moderate level of population differentiation detected also correlates with the moderate to high pairwise $F_{ST}$ values (0.007–0.516, average 0.222) discovered between the Finnish populations, as well as with significant overall $F_{ST}$ (0.241) and $R_{ST}$ (0.217) values over all Finnish populations ($p < 0.001$).

A Bayesian analysis of population structure was performed to determine the number and distribution of genetic clusters in the Finnish data set, and to demonstrate the number of distinct introductions of *E. canadensis* to Finland. In the Bayesian analysis, most of the existing regional population structure in the Finnish populations were captured with two clusters (K = 2). This division clustered the two northern populations with one of the southern populations, Lake Littoistenjärvi (cluster A), and combined the rest of the southern populations to another cluster (B) (Fig. 4). Cluster A could be further divided into two clusters, suggesting that the two northern populations are at some level distinct from all southern populations, and the Lake Littoistenjärvi population is distinct from the rest of the southern populations. The $F_{ST}$ values further confirm this regional structuring, since populations within cluster B were more similar to each other than to other populations. Also the results from the assignment test support the formation of cluster A and its further division into two clusters (Table 6 in **II**).

However, it has been suggested that isolation by distance (IBD) in a data set could cause the Bayesian methods to overestimate genetic structure, especially in irregularly collected samples (Frantz et al. 2009; Schwartz & McKelvey 2009). In our data set, genetic differences between populations increased significantly with geographic distance ($r = 0.446$, $p < 0.05$) when all the Finnish populations were included in the analysis. However, when only the five populations in Southern Finland were included, no isolation by distance was detected. Therefore, part of the genetic structure detected in Finnish *E. canadensis* populations could be caused by the IBD. On the other hand, *E. canadensis* was introduced to Finland only recently, and, therefore, there has not been much time for evolutionary changes. However, some of the microsatellite loci developed are evidently located in genomic areas with very high levels of variation. Moreover, absence of sexual reproduction after a single introduction event does not necessarily prevent long-term persistence or the development of significant microsatellite genetic diversity (Johnson et al. 2011; Karlin et al. 2011a; 2011b).

The results of this study revealed that the level, structure and distribution of genetic variation both within and among Finnish populations were higher than expected, considering the species' asexual mode of reproduction. This pattern of genetic variation may implicate that *E. canadensis* has been introduced to Finland more than once. On the other hand, unique alleles were found only from three populations, and the differences between the allele lengths detected were rather small. These results support the possibility that at least part of the variation detected may have occurred after the introduction. Also, the genetic structure detected in Finnish *E. canadensis* populations may be overestimated due to the IBD detected. Therefore, the possibility of only one introduction followed by post-establishment evolution cannot be rejected based on

the results of this thesis.

## 3.2 Chloroplast genome organization of *Elodea canadensis* and other monocots

Despite of the high number of cp genome sequences available, only few monocot cp genomes representing early lineages of monocots have been sequenced so far. I report the complete cp genome sequence of *E. canadensis* (**III**), which represents a taxonomically basal monocot. This is the first cp genome belonging to the family Hydrocharitaceae sequenced to date. The cp genome of *E. canadensis* is a circular double-stranded DNA molecule, 156,700 bp in length, and has a typical structure with large (LSC 86,194 bp) and small (SSC 17,810 bp) single-copy regions separated by a pair of inverted repeat regions (IRs 26,348 bp each). The *E. canadensis* cp genome contains 113 unique genes and 16 duplicated genes in the IR regions (Fig. 6).

Tobacco (*Nicotiana tabacum*) was the first angiosperm cp genome sequenced completely (Shinozaki et al. 1986) and, therefore, it has often been used as a reference genome in terms of the gene content and organization. A comparative analysis showed that the gene order and organization of the *E. canadensis* cp genome is almost identical to those of tobacco (Shinozaki et al. 1986) and a basal angiosperm *Amborella trichopoda* (Goremykin et al. 2003). The ancestral gene content and organization of the cp genome have been modified by structural rearrangements and gene loss and gain events in several monocot lineages (Chang et al. 2006; Guisinger et al. 2010). However, in comparison with other monocots, *E. canadensis* cp genome has gone through less rearrangements or gene losses when compared to assumed ancestral species. The percent identity plot generated by Advanced PipMaker (Schwartz et al. 2000) shows that intergenic spacers are the most divergent regions (shown in gray or white in Fig. 7), while the rest of the divergent regions represent intron and gene losses among monocots (see arrows in Fig. 7). *Phoenix dactylifera* (Yang et al. 2010) and *Typha latifolia* (Guisinger et al. 2010) have the only published monocot cp genomes with the same gene content with *E. canadensis*.

I identified the number of SSRs in 14 monocot cp genomes. The total number of SSRs varied from 106 in *Oryza sativa* to 209 in *Phoenix dactylifera*, being 127 in the *E. canadensis* cp genome. The ratio of genome size and the total number of SSRs was 1234 in *E. canadensis*. This was the second highest value after *Oryza*, indicating that *E. canadensis* cp genome contains a low number of SSRs relative to the mean genome size among monocots. *E. canadensis* also had the lowest number of mononucleotide SSRs among monocot cp genomes studied.

The two IRs in cp genome contain four junctions, $J_{LB}$, $J_{LA}$ (between the two IRs and the LSC region), $J_{SB}$ and $J_{SA}$ (between the two IRs and the SSC region) (Shinozaki et al. 1986). Size variation by extensions or contractions within IR sequences is the main reason for size variation between cp genomes of different taxa (Goulding et al. 1996; Chang et al. 2006; Raubeson et al. 2007; Wang et al. 2008; Yang et al. 2010), together with loss of genes and differences, such as indels in the intergenic regions (Ravi et al. 2008). At the same time, IRs are the most conserved regions in the cp genome, as the rate of neutral nucleotide substitutions is lower within IRs than in single-copy regions (Wolfe et al. 1987).

I discovered that the structure of IRs in *E. canadensis* cp genome is unique among monocot species with the whole cp genome sequenced. In most monocots, $J_{LA}$ is located downstream of the *psbA* gene and IRs have expanded to include *trnH-GUG-rps19* gene cluster (Wang et al. 2008). However, in *E. canadensis* and in other monocot *Lemna minor*, both of these genes are included in LSC and are not duplicated (Fig. 8). This type of organization is similar to that of most eudicots and basal angiosperms, such as *Amborella*. However, there is variation in the extent of duplications among other monocots as well (Fig. 8). Variation in the extent of duplications has also been detected at IR/SSC boundaries of monocot cp genomes. The structure of $J_{SA}$ in *E. canaden-*

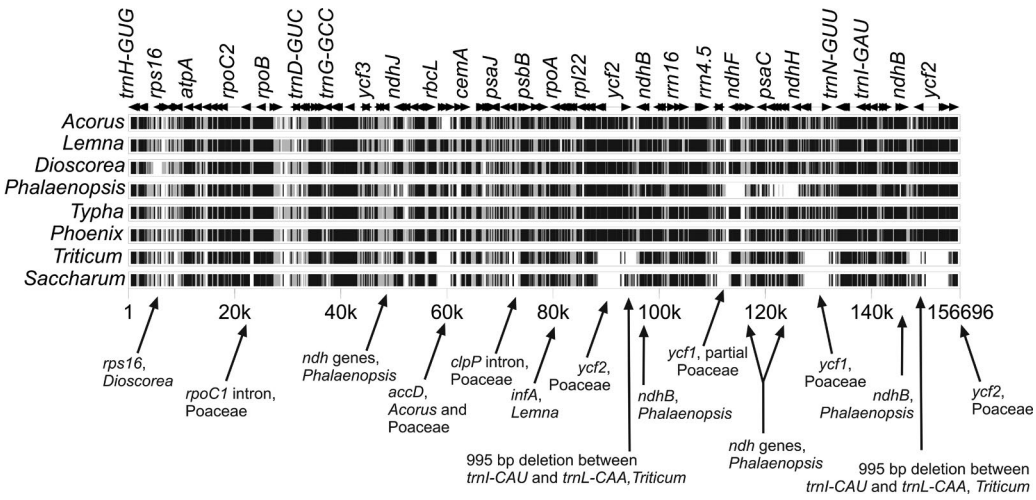


**Figure 7** A percent identity plot showing the overall sequence similarity of chloroplast genomes of nine monocots using *Elodea canadensis* as the reference genome. Levels of sequence similarity are indicated by black (75–100%), gray (50–75%) and white (<50%).

*sis*, *Typha latifolia* and *Acorus* is similar to that of *Amborella*, where only a part of the *ycf1* gene is duplicated in IR (Fig. 8).

The fluxes in the IR boundaries have been suggested to implicate taxonomic relationships among angiosperms, and the existence of *trnH-rps19* gene cluster in IR of most monocots has been proposed to be an evidence of a duplication event prior to the divergence of monocot lineages (Chang et al. 2006). There are, however, contradictory opinions on whether the IR expansions correlate with the divergence pattern of monocot phylogeny (Wang et al. 2008; Yang et al. 2010). Genus *Acorus* has been suggested as the earliest splitting lineage in monocots (Chase 2004), and, therefore, the IR/LSC junction of Acorales has been suggested to represent a primitive state compared to other monocot lineages (Wang et al. 2008). In *Acorus*, the *rps19* sequence of $IR_A$ is incomplete, and, therefore, it has been speculated that the expansion of IR to include *trnH-rps19* cluster has happened early in monocot lineage, while the contraction of IR in Acorales is secondary and happened only after the expansion. It has also been proposed that IR contractions at IR/LSC junctions of most Alismatales (*Elodea* and *Lemna*) and Dioscoreales (*Dioscorea*) might have happened separately and independently (Wang et al. 2008). The flanking sequences at the IR/LSC boundaries in Alismatales have been found to be more similar to other monocots than to non-monocot angiosperms. This result indicates that the IR contraction is likely due to an early termination of the repair-extension reaction or a contraction after the expansion of IR in monocot lineage. The sequence flanking IR/LSC junctions are also found to be typically A-rich and there is much more variation in this sequence in monocots than in eudicots (Wang et al. 2008). This kind of polyA tract sequences might be recombination hotspots (Goulding et al. 1996), which supports the hypothesis suggesting several independent contraction events in the IR evolution of monocots (Wang et al. 2008).

*Elodea canadensis* is the only monocot species studied so far with all IR junctions similar to those of basal angiosperm *Amborella* (Fig. 8). This result supports the position of *E. canadensis*

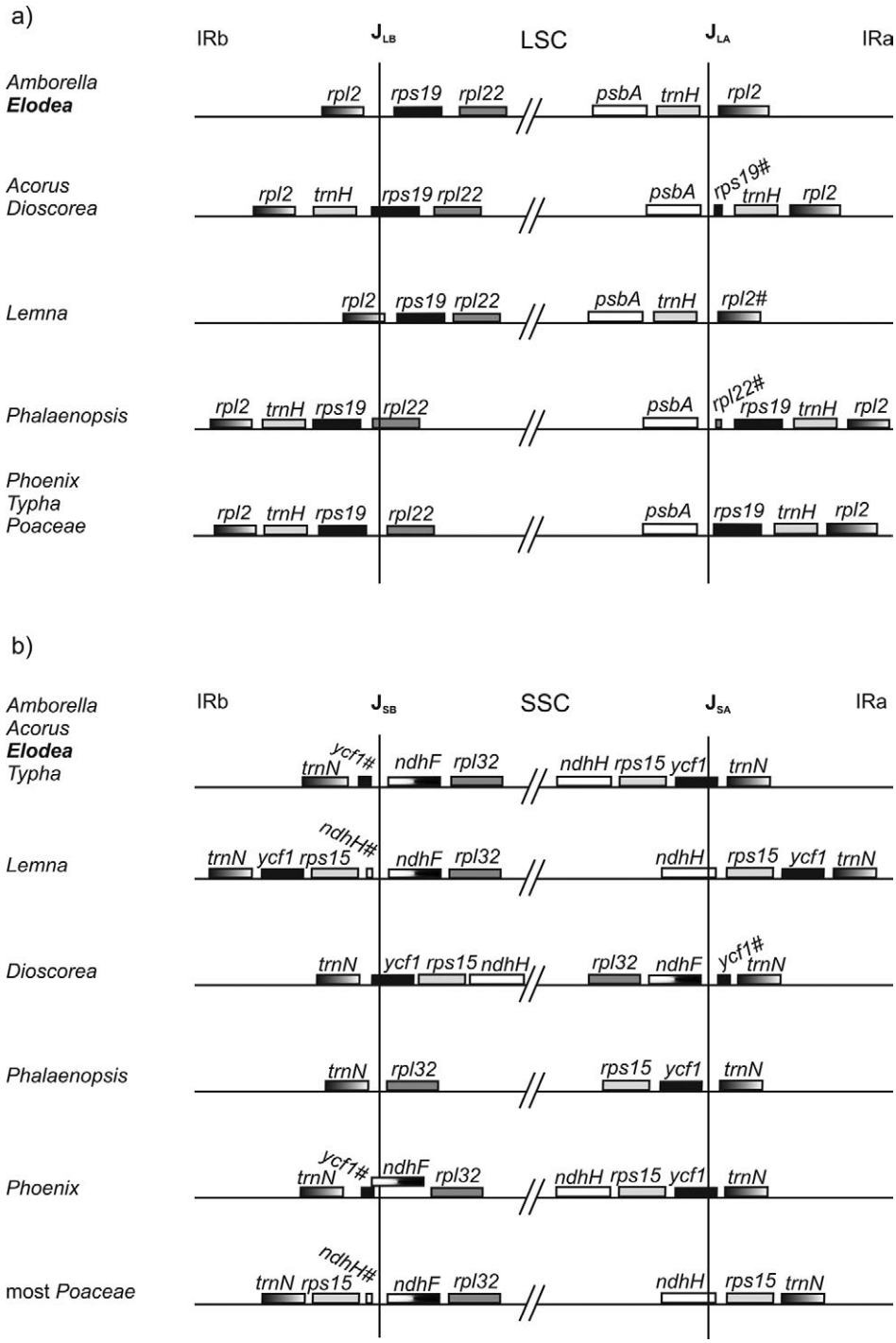


**Figure 8** Comparison of the junction positions ($J_{LB}$, $J_{LA}$, $J_{SB}$ and $J_{SA}$) between a) inverted region (IR) and large single-copy (LSC) region and b) IR and small single-copy (SSC) region in eight representative monocot taxa and *Amborella trichopoda*, a basal angiosper.

as a basal monocot and gives valuable information about the course of evolution and divergence in monocot lineages. Phylogenetic analyses performed using 81 chloroplast genes from 17 sequenced monocot plastid genomes further support the placement of *E. canadensis* as a basal monocot. In Bayesian tree, *E. canadensis* was located together with *Lemna*, as the next diverging lineage of monocots after Acorales (Fig. 4 in **III**). It may be speculated that contractions of IR at $J_{LB}$ in basal monocots and also the expansions of IR to include *rps15* (*Lemna* and *Poaceae*) and *ycf1* (*Lemna*) at $J_{SA}$ have happened separately and independently, and that the signals of the course of evolution given by both ends of IR interpreted simultaneously could be misleading.

## 3.3 Chloroplast DNA phylogeography of *Elodea canadensis* and *E. nuttallii*

### *3.3.1. Level of chloroplast DNA variation between native and introduced Elodea canadensis populations*

Most intraspecific studies of cpDNA variation have utilized a limited number of markers from a few selected gene regions. Only few studies have examined intraspecific variation in the whole cp genome sequence (Tang et al. 2004; Diekmann et al. 2009; Young et al. 2011). I was able to investigate variation in an 112,193 bp (86.1%, only one IR included) region of two *E. canadensis* plastid sequences (USA and FIN) and found a total of 235 variable sites (186 SNPs, 47 indels and two inversions) (**IV**).

The mutation rate of SNPs (often about $10^{-8}$–$10^{-9}$) is low compared to SSRs (about $10^{-4}$). On average, one SNP can be expected every 500–1000 bp in coding regions and every 200–500 bp in non-coding regions (Brumfield et al. 2003). The evolution rate of cpDNA is only half of that in nuclear DNA in plants (Wolfe et al. 1987) and, therefore, also the incidence of SNPs in cp genome is expected to be lower. I found on average one SNP every 863 bp in coding and every 420 bp in non-coding regions. The total frequency was one SNP every 603 bp. Based on these figures, the frequency of SNPs between *E. canadensis* plastid sequences is slightly higher than in average. In the *E. canadensis* plastid sequences, SNPs occurred at a rate of 16.6 in 10,000 bp, giving the intraspecific polymorphism rate for SNPs 0.16%, which is high in comparison to rates previously reported for angiosperm cp genomes (0.05% for rice (Tang et al. 2004) and 0.07% for *Panicum* (Young et al. 2011)). Over all, the level of variation between the two *E. canadensis* plastid sequences was constantly higher than that discovered in earlier intraspecific comparisons in angiosperms (Tang et al. 2004; Diekmann et al. 2009; Young et al. 2011). Possible reasons for this phenomenon are discussed in chapter 3.3.3.

### *3.3.2. Level of chloroplast DNA variation between Elodea canadensis and E. nuttallii*

The 29 cpDNA primer pairs flanking the most variable genomic areas between the USA and FIN plastid sequences, comprising 11,464 bp (Fig. 6), were used to compare the level and pattern of intraspecific cpDNA variation within *E. canadensis* and interspecific variation between *E. canadensis* and *E. nuttallii* (**IV**). I discovered on average 61% more SNPs between the two *Elodea* species than between the native and introduced *E. canadensis* cp genomes, while the number of indels was 10% larger in the within-species comparison of *E. canadensis*. On the whole, the level of variation between the two *Elodea* species was 53% higher than that within *E. canadensis*.

### *3.3.3. Geographical distribution of cpDNA haplotypes*

The number of haplotypes in an invasive range is a function of many factors, including the number of introductions, the size of each introduction, the population structuring in the native range, and subsequent drift and selection pressures that occur after the introduction (Gaskin et al. 2005). In the phylogeographic analysis, only a single haplotype was found in the introduced range in both species. These haplotypes H1 (*E. canadensis*)

and A (*E. nuttallii*) were also widespread in the native range, covering the majority of native populations analysed (Fig. 5). Therefore, I was not able to identify either the geographic origin of the introduced populations or test the hypothesis of single versus multiple introductions. The result could indicate one introduction event, but multiple introductions of the same haplotype are just as possible.

Despite the limited variation, the data presented here provides interesting information of the genetic patterns in the native and introduced ranges of *E. canadensis* and *E. nuttallii*. The cpDNA haplotype homogeneities at introduced range are most likely attributed to bottleneck followed by a single introduction event or few introductions of similar haplotypes. This pattern is further supported by the vegetative regeneration, combined to the relatively recent and fast expansions of both species in their introduced ranges. Several studies have explored whether certain genotypes within an invading population are more succesfull and explain or predict invasive behaviour across species (reviewed by Bossdorf et al. 2005). Saltonstall (2002) found one haplotype of *Phragmites australis* dominating the Atlantic Coast tidal marshes and replacing native haplotypes. Certain haplotype may also be introduced to a novel environment earlier than others, or have larger propagule pressure due to multiple introductions or higher number of individuals released and, therefore, become more successful at the introduced area. These factors may result in the situation where one haplotype dominates over the others (Fig. 1).

The haplotype network constructed indicates that *E. canadensis* haplotype H1 is very divergent from haplotypes H2 and H3, as it could not be connected within the limits of parsimony (95%) (Fig. 5). The diverged haplotypes were also geographically clustered in both species, although, a more detailed sampling would be necessary to get more support for the results presented here. The large divergence between *E. canadensis* haplotypes lets us hypothesize the evolutionary event behind this phenomenon.

The divergent haplotypes could be explained by chloroplast capture, where the cytoplasm of one species is replaced by that of another species through hybridization or introgression, as reported to occur in natural ecosystems (Riesenberg & Soltis 1991; Fehrer et al. 2007). Chloroplast capture can occur between species with sympatric distribution and reproductive compatibility, and the process is facilitated by weak reproductive barriers between species (Acosta & Premoli 2010). In *Elodea*, males and females are rarely found in the same population (Cook & Urmi-König 1985), indicating strong reproductive barrier between populations within species. However, this kind of a situation might lead to a weak reproduction barrier between species, if the appearance of another sex, even representing another *Elodea* species, would allow sexual reproduction between species.

There is some support for naturally occurring hybrids between *E. canadensis* and *E. nuttallii* (Cook & Urmi-König 1985). However, most of the variable sites found between *E. canadensis* cp genomes FIN and USA were not detected in *E. nuttallii*, indicating that hybridization between these species does not explain the large divergence between *E. canadensis* haplotypes. Interestingly, the distributions of *E. canadensis* and *E. bifoliata* would enable a hybrid zone allowing the cp capture event involving the parapatric *E. bifoliata* (Fig. 5 in **IV**). However, an extensive survey of genetic characteristics of *E. bifoliata* would be needed to firmly retrace the possibility of cp capture.

I found several variable sites between *E. nuttallii* haplotypes as well, and the level of variation between *E. canadensis* and *E. nuttallii* cp genomes was still 53% higher than that between the two *E. canadensis* plastid sequences. In addition, while the ITS sequences in *E. nuttallii* have been found highly homologous, those of *E. canadensis* include several polymorphic sites (Gross et al. 2003). Therefore, based on our results, the possibility of high level of naturally occurring variation in both nuclear and cp genomes of *E. canadensis* cannot be ruled out.

# 4 Conclusions and future perspectives

**IN THIS THESIS**, I examined essential aspects of the biology of invasive species using *Elodea* as a model. *Elodea* is one of the key genuses of invasion biology and provides a perfect model for studying the evolutionary consequences of invasion, since the timing of the first introduction events are well known. The studies of this thesis gave new information on the genetic diversity, population structure and distribution history in two invasive aquatic weeds, *Elodea canadensis* and *E. nuttallii*. The results also revealed the unique structure of chloroplast genome of *E. canadensis* and gave new information on the course of evolution and divergence in monocot lineages. This thesis provide great source of novel molecular markers that will be useful in future studies on *Elodea* species. The results highlight the need for further investigations in the *Elodea* species, as well as in other invasive plant species.

The results of the substudy **II** revealed a moderate level of microsatellite variation within and among seven introduced *E. canadensis* populations. The results indicate that *E. canadensis* could have been introduced to Finland more than once, considering the asexual reproduction mode and recent introduction. On the other hand, the genetic structure detected in Finnish *E. canadensis* populations may be overestimated due to the isolation by distance detected. Moreover, some of the microsatellite loci developed are evidently located in genomic areas with very high levels of variation, and the low level of cpDNA variation detected in substudy **IV** further suggest one introduction event followed by effective range expansion. Therefore, the possibility of only one introduction followed by post-establishment evolution cannot be rejected based on the results of this thesis. Furthermore, Finnish populations of *E. canadensis* are assumedly derived from the primary areas of introduction in Europe, which makes it more difficult to interpret the results according to one or multiple introductions. The results of this study indicate that Finnish populations are more similar to each other than to native populations, and the native populations studied have probably not been the direct source of the Finnish populations.

In order to further estimate the number of introductions to Finland, genetic variation should be analysed more widely within both native and introduced range of *E. canadensis*. The microsatellite markers were originally developed using introduced *E. canadensis* material collected from Finland (**I**), and most of them proved not to be suitable for analysing native populations of *E. canadensis*. Those loci that did not amplify in native samples may be located in genome areas with a high evolutionary rate, while the amplifying loci may be located in more conservative areas. Therefore, a new set of microsatellite markers developed using native *E. canadensis* material could give valuable information on the level of variation in native *E. canadensis* populations. Also, a detailed study of *Elodea*'s mating system using sex-linked genetic markers would be valuable in order to reveal the actual level of clonality in both native and introduced ranges.

Substudy **III** showed that the cp genome of *E. canadensis* has gone through less rearrangements or gene losses when compared to assumed ancestral species than other monocots. The IR of *E. canadensis* has a unique structure among the monocot species studied so far, as its structure is similar to that of *Amborella*, a basal angiosperm. The position of *Elodea* as a basal monocot was further supported by the phylogenetic analyses. Only few cp genomes representing early lineages of monocots have been sequenced and, therefore, this study provides valuable information on the

course of evolution and divergence in monocot lineages.

The results of the substudy **IV** revealed interesting results in the phylogeography of cpDNA haplotypes in *E. canadensis* and *E. nuttallii*. Only three *E. canadensis* haplotypes and four *E. nuttallii* haplotypes were found. In both species, only a single haplotype was found in the introduced range. These dominant haplotypes covered most populations analysed in the native range as well. Therefore, we were not able to identify either the geographic origin of the introduced populations or test the hypothesis of single versus multiple introductions. More samples from the native and introduced ranges of both species would be valuable for making inferences about the area of introduction.

Over all, the levels of variation between the native and introduced *E. canadensis* plastid sequences and between the most diverged *E. canadensis* cpDNA haplotypes were constantly higher than those discovered in earlier intraspecific comparisons in angiosperms (Tang et al. 2004; Diekmann et al. 2009; Young et al. 2011). The diverged haplotypes were also geographically clustered. The divergent haplotypes could be explained by chloroplast capture event involving the parapatric *E. bifoliata*. However, the possibility of a higher level of naturally occurring variation in the nuclear and cp genome sequence of *E. canadensis* cannot be ruled out based on this thesis and, therefore, an extensive survey of genetic characteristics of *E. bifoliata* would be needed to firmly retrace the possibility of chloroplast capture. Other issues that require further study are the possible differences in morphology, habitat requirements and invasiveness between the diverged haplotypes of *E. canadensis* and *E. nuttallii*.

This thesis helps to fill in some gaps in the knowledge of population genetics of *Elodea* species. Unfortunately, biological research can only solve a small fraction of the problems required in managing invasive species and practical efforts are needed. Repeated cutting is one of the most common methods used to control *Elodea* populations. This method may, however, give only short-time effects, and eventually even increase the rate of spread. After mechanical clipping, the cut-off fragments need to be removed carefully. If the fragments are not collected, they can either be transported by water currents to new areas, where they may establish new colonies, or they may start to grow vigorously in the original water system and magnify their population density in short time. Mechanically pulled rope seines, facilitating the removal of submerged aquatic vegetation without leaving fragments, have been the most effective method to remove *Elodea* from water systems so far (Laita et al. 2007). However, a more comprehensive control or eradication programme should be adopted. With the warming climate, new species will have the potential to establish if introduced, and possibly become invasive. Therefore, future research should focus on predictive analyses of potential future invaders and other preventive methods to minimize new introductions. In several countries, national strategies on invasive alien species have already been prepared and put into practice.

In Finland, relatively few non-indigenous freshwater plant species have succeeded in all invasion phases (Pienimäki & Leppäkoski 2004). At present, *E. canadensis* is the most harmful invasive aquatic weed in Finland and its mass occurrences have proved to be difficult to manage (Laita et al. 2007), whereas *E. nuttallii* has not yet established within Finnish borders but has a high probability to arrive to Finland. Both, *E. canadensis* and *E. nuttallii* are still expanding their range and new invasions are occurring at least in their northern distribution limit (Heikkinen et al. 2009). Both species are, therefore, included in the Finnish national strategy on invasive alien species (www.mmm.fi/vieraslajit). Due to the difficulties in the discrimination between *E. canadensis* and *E. nuttallii*, the molecular markers developed in this thesis will offer a valuable tool in monitoring purposes and in early detection of *E. nuttallii* in Finland and other locations with high risk of invasion.

# 5 Acknowledgements

**FIRST OF ALL,** I would like to thank my supervisors Helena Korpelainen and Elina Leskinen. Helena's positive attitude helped me accomplish what ever was needed, even if I otherwise might have given up. I'm very grateful for your support and amazingly fast comments to my work. Elina, I truly enjoyed our discussions. Thank you for a great inspiration and support. I also owe my gratitude to Kirsi Kostamo, who basically formed my unofficial advisory group. I am especially grateful for your patience when teaching me my first steps in the laboratory and your help in the field when collecting *Elodeas*. I am most grateful for Jouni Aspi and Alain Vanderpoorten for the valuable reviewing work of this thesis.

A number of people have provided *Elodea* samples for my work. Special thanks to Ilkka Sammalkorpi, Ari Mäkelä, Seppo Hellsten, Anna Väisänen, Riku Paavola and the Oulanka Research Station for co-operation and help in the sample collection. I owe my sincere thanks to David Biesboer, Ray Newman and Mark Gernes for your invaluable help during the collection of native *Elodea* samples. I am grateful for Jouko Sarvala for important advices in the beginning of my work. I want to acknowledge The Finnish Museum of Natural History, Oregon State University Herbarium and New York Botanical Garden Herbarium for providing the herbarium specimens of *Elodea* for my research.

I thank all the researchers and technicians at the Department of Agricultural Sciences who gave me methodological advice. Thank you Marjo for answering all the questions and for making me laugh countless times in the lab during discussions often far away from the science.

I thank members of our special research group for the good times spent during these years. Johanna, Kirsi, Hanna, Sakina, Sanna, and both Marias, I loved to come to work every day because of you! I thank Inka for the friendship and many interesting conversations. I also thank Tuula for sharing the office and piece of your interesting life with me. I am forever grateful for Maria for sharing the office as well as all the joys and sorrows in life. I learned from you that life does not need to be that serious.

I wish to acknowledge the Academy of Finland, the Finnish Cultural Foundation, Emil Aaltonen Foundation, Oskar Öflund Foundation and University of Helsinki for financial support.

I am grateful for all my friends for company and good times during these years. Jenni, you have been one of my most important inspirations during this work. I feel I can always count on your support and your endless optimism is just amazing!

Thank you Iskä for always supporting my decisions and letting me do my own mistakes in life. Hanne, my dear sister, you have made my day so many times by just being there and listening. You are the person with whom I can simply be myself, always. Jussi, thank you for your support and patience during this project. You have seen the good days and the bad, and always reminded me of the things that really matter. Mosku, I am fortunate in having you in my life. I thank you for taking my thoughts away from the work. Finally, this thesis is dedicated to my Mom, who had a big influence on me to become brave enough to start this work and to believe in myself enough to complete this project. Mom, I wish I made you proud.

# 6 References


Aston HI (1973) Aquatic plants of Australia. Melbourne University Press, Melbourne.

Avise JC (2004) Molecular markers, natural history and evolution. Sinauer Associates inc., Sunderland, MA.

Balloux F, Lehmann L, de Meeûs T (2003) The population genetics of clonal and partially clonal diploids. *Genetics*, **164**, 1635–1644.

Barrat-Segretain M-H, Arnaud E (2004) Experiments on growth interactions between two invasive macrophyte species. *Journal of Vegetation Science*, **15**, 109–114.

Barrett SCH, Eckert CG, Husband BC (1993) Evolutionary processes in aquatic plant populations. *Aquatic Botany*, **44**, 105–145.

Besnard G, Henry P, Wille L, Cooke D, Chapuis E (2007) On the origin of the invasive olives (*Olea europaea* L., Oleaceae). *Heredity*, **99**, 608–619.

Bonin A, Bellemain E, Bronken Eidesen P, Pompanon F, Brochmann C, Taberlet P (2004) How to track and assess genotyping errors in population genetics studies. *Molecular Ecology*, **13**, 3261–3273.

Bossdorf O, Auge H, Lafuma L, Rogers WE, Siemann E, Prati D (2005) Phenotypic and genetic differentiation between native and introduced plant populations. *Oecologia*, **144**, 1–11.

Bowmer KH, Mitchell DS, Short DL (1984) Biology of *Elodea canadensis* Mich. and its management in Australian irrigation systems. *Aquatic Botany*, **18**, 231–238.

Bowmer KH, Jacobs SWL, Sainty GR (1995) Identification, biology and management of *Elodea canadensis*, Hydrocharitaceae. *Journal of Aquatic Plant Management*, **33**, 13–19.

Brumfield RT, Beerli P, Nickerson DA, Edwards SV (2003) The utility of single nucleotide polymorphisms in inferences of population history. *Trends in Ecology and Evolution*, **18**, 249–256.

Catalano SA, Saidman BO, Vilardi JC (2009) Evolution of small inversions in chloroplast genome: a case study from a recurrent inversion in angiosperms. *Cladistics*, **25**, 93–104.

Chang C-C, Lin H-C, Lin I-P, Chow T-Y, Chen H-H, Chen W-H, Cheng C-H, Lin C-Y, Liu S-M, Chang C-C, Chaw S-M (2006) The chloroplast genome of *Phalaenopsis Aphrodite* (Orchidaceae): comparative analysis of evolutionary rate with that of grasses and its phylogenetic implications. *Molecular Biology and Evolution*, **23**, 279–291.

Chapman VJ (1970) A history of the lake-weed infestation of the Rotorua lakes and the lakes of the Waikato hydro-electric system. *New Zealand Department of Scientific and Industrial Research Information Series*, **78**, 1–52.

Charlesworth B (2009) Effective population size and patterns of molecular evolution and variation. *Genetics*, **10**, 195–205.

Chase MV (2004) Monocot relationships: an overview. *American Journal of Botany*, **91**, 1645–1655.

Clement M, Posada D, Crandall K (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1660.

Cook CDK, Urmi-König K (1985) A revision of the genus *Elodea* (Hydrocharitaceae). *Aquatic Botany*, **21**, 111–156.

Cox GW (2004) Alien species and evolution: The evolutionary ecology of exotic plants, animals, microbes and interacting native species. Island Press, Washington, DC.

de Witte LC, Stöcklin J (2010) Longevity of clonal plants: why it matters and how to measure it. *Annals of Botany*, **106**, 859–870.

Diekmann K, Hodkinson TR, Wolfe, KH, van den Bekerom R, Dix PJ, Barth S. (2009) Complete chloroplast genome sequence of a major allogamous forage species, perennial ryegrass (Lolium perenne L.). *DNA Research*, **16**, 165–176.

Durka W, Bossdorf O, Prati D, Auge H (2005) Molecular evidence for multiple introductions of garlic mustard (*Alliaria petiolata*, Brassicaceae) to North America. *Molecular Ecology*, **14**, 1697–1706.

Ellegren H (2004) Microsatellites: simple sequences with complex evolution. *Nature Reviews Genetics*, **5**, 435–445.

Ellstrand NC, Elam DR (1993) Population genetic consequences of small population size: implications for plant conservation. *Annual Review of Ecology and Systematics*, **24**, 217–242.

Ellstrand NC, Schierenbeck KA (2006) Hybridization as a stimulus for the evolution of invasiveness in plants? *Euphytica*, **148**, 35–46.

Engloner AI, Major A, Podani J (2010) Clonal diversity along a water depth gradient in a declining reed stand as detected by three different genetic methods. *Aquatic Botany*, **92**, 1–8.

Excoffier L, Foll M, Petit RJ (2009) Genetic consequences of range expansions. *Annual Review of Ecology, Evolution and Systematics*, **40**, 481–501.

Faircloth BC (2008) MSATCOMMANDER: detection of microsatellite repeat arrays and automated, locus-specific primer design. *Molecular Ecology Resources*, **8**, 92–94.

Fehrer J, Gemeinholzer B, Chrtek Jr J, Bräutigam S (2007) Incongruent plastid and nuclear DNA phylogenies reveal ancient intergeneric hybridization in *Pilosella* hawkweeds (*Hieracium*, Cichorieae, Asteraceae). *Molecular Phylogenetics and Evolution*, **42**, 347–361.

Frankham R (2005) Resolving the genetic paradox in invasive species. *Heredity*, **94**, 385.
Frankham R, Ballou JD, Briscoe DA (2003) Introduction to Conservation Genetics. Cambridge university press, Cambridge.
Frantz AC, Cellina S, Krier A, Schley L, Burke T (2009) Using spatial methods to determine the genetic structure of a continuously distributed population: clusters or isolation by distance? *Journal of Applied Ecology*, **46**, 493–505.
Galera H, Sudnik-Wójcikowska B (2010) Central European botanic gardens as centres of dispersal of alien plants. *Acta Societatis Botanicorum Poloniae*, **79**, 147–156.
Gaskin JF, Zhang D-Y, Bon M-C (2005) Invasion of *Lepidium draba* (Brassicaceae) in the western United States: distributions and origins of chloroplast DNA haplotypes. *Molecular Ecology*, **14**, 2331–2341.
Gill DE, Chao L, Perkins SL, Wolf JB (1995) Genetic mosaicism in plants and clonal animals. *Annual Review of Ecology and Systematics*, **26**, 423–444
Goldstein DB, Schlötterer C (1999) Microsatellites, Evolution and Applications. Oxford university press, Oxford.
Gordon DR (1998) Effects of invasive, non-indigenous plant species on ecosystem processes: lessons from Florida. *Ecological Applications*, **8**, 975–989.
Goremykin VV, Hirsch-Ernst KI, Wölfl S, Hellwig FH (2003) Analysis of the *Amborella trichopoda* chloroplast genome sequence suggests that *Amborella* is not a basal angiosperm. *Molecular Biology and Evolution*, **20**, 1499–1505.
Goremykin VV, Holland B, Hirsch-Ernst KI, Hellwig FH (2005) Analysis of *Acorus calamus* chloroplast genome and its phylogenetic implications. *Molecular Biology and Evolution*, **22**, 1813–1822.
Goulding SE, Olmstead RG, Morden CW, Wolfe KH (1996) Ebb and flow of the chloroplast inverted repeat. *Molecular and General Genetics*, **252**, 195–206.
Guisinger MM, Chumley TW, Kuehl JV, Boore JL, Jansen RK (2010) Implications of the plastid genome sequence of *Typha* (Typhaceae, Poales) for understanding genome evolution in Poaceae. *Journal of Molecular Evolution*, **70**, 149–166.
Halkett F, Simon J-C, Balloux F (2005) Tackling the population genetics of clonal and partially clonal organisms. *Trends in Ecology and Evolution*, **20**, 194–201.
Hansen DR, Dastidar SG, Cai Z, Penaflor C, Kuehl JV, Boore JL, Jansen RK (2007) Phylogenetic and evolutionary implications of complete chloroplast genome sequences of four early-diverging angiosperms: *Buxus* (Buxaceae), *Chloranthus* (Chloranthaceae), *Dioscorea* (Dioscoreaceae), and *Illicium* (Schisandraceae). *Molecular Phylogenetics and Evolution*, **45**, 547–563.
Hedrick PW (1999) Perspective: Highly variable loci and their interpretation in evolution and conservation. *Evolution*, **53**, 313–318.
Heger T, Trepl L (2003) Predicting biological invasions. *Biological Invasions*, **5**, 313–321.
Heikkinen RK, Leikola N, Fronzek S, Lampinen R, Toivonen H (2009) Predicting distribution patterns and recent northward range shift of an invasive aquatic plant: *Elodea canadensis* in Europe. *BioRisk*, **2**, 1–32.
Heino J, Virkkala R, Toivonen H (2009) Climate change and freshwater biodiversity: detected patterns, future trends and adaptations in northern regions. *Biological Reviews*, **84**, 39–54.
Hintikka TJ (1917) Kanadalaisesta vesirutosta ja sen levenemisestä Euroopassa, eritoten Suomessa. *Luonnon Ystävä*, **21**, 79–90.
Hufbauer R, Sforza R (2008) Multiple introductions of two invasive *Centaurea* taxa inferred from cpDNA haplotypes. *Diversity and Distributions*, **14**, 252–261.
Ivey CT, Richards JH (2001) Genetic diversity of everglades sawgrass, *Cladium jamaicense* (Cyperaceae). *International Journal of Plant Sciences*, **162**, 817–825.
Jensen, J.L., Bohonak, A.J., and Kelley, S.T. 2005. Isolation by distance, web service. BMC Genetics 6: 13. v.3.16 http://ibdws.sdsu.edu/
Johnson MTJ, FitzJohn RG, Smith SD, Rausher MD, Otto SP (2011) Loss of sexual recombination and segregation is associated with increased diversification in evening primeroses. *Evolution*, **65**, 3230–3240.
Kadono Y (2004) Alien aquatic plants naturalized in Japan: history and present status. *Global Environmental Research*, **8**, 163–169.
Karlin EF, Andrus RE, Boles SB, Shaw J (2011a) One haploid parent comtributes 100% of the gene pool for a widespread species in northwest North America. *Molecular Ecology*, **20**, 753–767.
Karlin EF, Hotchkiss SC, Boles SB, Stenøien HK, Hassel K, Flatberg KI, Show AJ (2011b) High genetic diversity in a remote island population system: sans sex. *New Phytologist*, **193**, 1088–1097.
Kelchner SA, Wendel JF (1996) Hairpins create minute inversions in non-coding regions of chloroplast DNA. *Current Genetics*, **30**, 259–262.
Kolar CS, Lodge DM (2001) Progress in invasion biology: predicting invaders. *Trends in Ecology and Evolution*, **16**, 199–204.
Korpelainen H (2004) The evolutionary processes of mitochondrial and chloroplast genomes differ from those of nuclear genomes. *Naturwissenschaften*, **91**, 505–518.
Korpelainen H, Kostamo K, Virtanen V (2007) Microsatellite marker identification using genome screening and restriction-ligation. *BioTechniques*. **42**, 479–486.
Kurtz S, Choudhuri JV, Ohlebusch E, Scheleiermacher C, Stoye J, Giegerich R (2001) REPuter: the manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Research*, **29**, 4366–4642.
Laita M, Tarvainen A, Mäkelä A, Sammalkorpi I, Kemppainen E, Laitinen L (2007) Uposkasvien runsastumisesta 2000-luvun alussa. Suomen ympäristökeskuksen raportteja 20, Suomen ympäristökeskus, Helsinki.
Lambertini C, Riis T, Olesen B, Clayton JS, Sorrell BK, Brix H (2010) Genetic diversity in three invasive clonal aquatic species in New Zealand. *BMC Genetics*, **11**, 52.

Larson D (2007) Reproduction strategies in introduced *Nymphoides peltata* populations revealed by genetic markers. *Aquatic Botany*, **86**, 402–406.
Lee CE (2002) Evolutionary genetics of invasive species. *Trends in Ecology and Evolution*, **17**, 386–391.
Les DH (1988) Breeding systems, population structure, and evolution in hydrophilous angiosperms. *Annals of the Missouri Botanical Garden*, **75**, 819–835.
Lesica P, Leary RF, Allendorf FW, Bilderback DE (1988) Lack of genic diversity within and among populations of an endangered plant, *Howellia aquatilis. Conservation Biology*, **2**, 275–282.
Levine JM, D´Antonio CM (2003) Forecasting biological invasions with increasing international trade. *Conservation Biology*, **17**, 322–326.
Lockwood JL, Cassey P, Blackburn T (2005) The role of propagule pressure in explaining species invasions. *Trends in Ecology and Evolution*, **20**, 223–228.
Lockwood JL, Hoopes MF, Marchetti MP (2008) Invasion Ecology. Blackwell publishing, Malden.
Mack RN, Simberloff D, Lonsdale WM, Evans H, Clout M and Bazzaz FA (2000) Biotic invasions: causes, epidemiology, global consequences, and control. *Ecological Applications*, **10**, 689–710.
Mardanov AV, Ravin NV, Kuznetsov BB, Samigullin TH, Antonov AS, Kolganova TV, Skyabin KG (2008) Complete sequence of the duckweed (*Lemna minor*) chloroplast genome: structural organization and phylogenetic relationships to other angiosperms. *Journal of Molecular Evolution*, **66**, 555–564.
McClintock B (1983) The significance of responses of the genome to challenge. *Science*, **26**, 792–801.
McIvor L, Maggs CA, Provan J, Stanhope MJ (2001) *rbcL* sequences reveal multiple cryptic introductions of the Japanese red alga *Polysiphonia harveyi*. *Molecular Ecology*, **10**, 911–919.
Meimberg H, Hammond JI, Jorgensen CM, Park TW, Gerlach KJ, Rice KJ, McKay JK (2006) Molecular evidence for an extreme genetic bottleneck during introduction of an invading grass to California. *Biological Invasions*, **8**, 1355–1366.
Müller-Schärer H, Schaffner U, Steinger T (2004) Evolution in invasive plants: implications for biological control. *Trends in Ecology and Evolution*, **19**, 417–422.
Nei M, Maruyama T, Chakraborty R (1975) The bottleneck effect and genetic variability in populations. *Evolution*, **29**, 1–10.
Paetkau D, Waits, LP, Clarkson PL, Craighead L, Strobeck C (1997) An empirical evaluation of genetic distance statistics using microsatellite data from bear (Ursidae) populations. *Genetics*, **147**, 1943–1957.
Philbrick CT, Les DH (1996) Evolution of aquatic angiosperm reproductive systems; What is the balance between sexual and asexual reproduction in aquatic angiosperms? *BioScience*, **46**, 813–826.
Pienimäki M, Leppäkoski E (2004) Invasion pressure on the Finnish lake district: invasion corridors and barriers. *Biological Invasions*, **6**, 331–346.
Piry S, Alapetite A, Cornuet J-M, Paetkau D, Baudouin L, Estoup A (2004) GENECLASS2: a software for genetic assignment and first-generation migrant detection. *Journal of Heredity*, **95**, 536–539.
Pock Tsy J-ML, Lumaret R, Mayne D, Vall AOM, Abutaba YIM, Sagna M, Raoseta SOR, Danthu P (2009) Chloroplast DNA phylogeography suggests a West African centre of origin for the baobab, *Adansonia digitata* L. (Bombacoideae, Malvaceae). *Molecular Ecology*, **18**, 1707–1715.
Pompanon F, Bonin A, Bellemain E, Taberlet P (2005) Genotyping errors: causes consequences and solutions. *Nature Reviews Genetics*, **6**, 847–859.
Prentis PJ, Wilson JRU, Dormontt EE, Richardson DM, Lowe AJ (2008) Adaptive evolution in invasive species. *Trends in Plant Science*, **13**, 288–294.
Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.
Procaccini G, Mazzella L (1998) Population genetic structure and gene flow in the seagrass *Posidonia oceanica* assessed using microsatellite analysis. *Marine Ecology Progress Series*, **169**, 133–141.
Provan J, Powell W, Hollingsworth PM (2001) Chloroplast microsatellites: new tools for studies in plant ecology and evolution. *Trends in Ecology and Evolution*, **16**, 142–147.
Pyšek P (1997) Clonality and plant invasions: can a trait make a difference? In: *The ecology and evolution of clonal plants* (eds de Kroon H, van Groenendael J), pp 405–427. Blackhuys Publishers, Leiden.
Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. Proceedings of the *National Academy of Sciences, USA*, **94**, 9197–9201.
Raubeson LA, Peery R, Chumley TW, Dziubek C, Fourcade HM, Boore JL, Jansen KJ (2007) Comparative chloroplast genomics: analyses including new sequences from the angiosperms *Nuphar advena* and *Ranunculus macranthus*. *BMC Genomics*, **8**, 174.
Ravi V, Khurana JP, Tyagi AK, Khurana P (2008) An update on chloroplast genome. *Plant Systematics and Evolution*, **271**, 101–122.
Reusch TBH, Stam WT, Olsen JL (2000) A microsatellite-based estimation of clonal diversity and population subdivision in *Zostera marina*, a marine flowering plant. *Molecular Ecology*, **9**, 127–140.
Ricciardi A, Rasmussen JB (1998) Predicting the identity and impact of future biological invaders: a priority for aquatic resource management. *Canadian Journal of Fisheries and Aquatic Sciences*, **55**, 1759–1765.
Riesenberg LH, Soltis DE (1991) Phylogenetic consequences of cytoplasmic gene flow in plants. *Evolutionary Trends in Plants*, **5**, 65–84.
Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics Applications Note*, **19**, 1572–1574.
Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics*, **145**, 1219–1228.
Sakai AK, Allendorf FW, Holt JS, Lodge DM, Molofsky J,

With KA, Baughman S, Cabin RJ, Cohen JE, Ellstrand NC, McCauley DE, O'neil P, Parker IM, Thompson JN, Weller SG (2001) The population biology of invasive species. *Annual Review of Ecology and Systematics*, **32**, 305–332.

Sala OE, Chapin FS, Armesto JJ, Berlow E, Bloomfield J, Dirzo R, Huber-Sandwald E, Huenneke LF, Jackson RB, Kinzig A, Leemans R, Lodge DM, Mooney HA, Oesterheld M, Poff NL, Sykes MT, Walker BH, Walker M, Wall DH (2000) Global biodiversity scenarios for the year 2100. *Science*, **287**, 1770–1774.

Saltonstall K (2002) Cryptic invasion by a non-native genotype of the common reed, *Phragmites australis*, into North America. *Proceedings of the National Academy of Sciences, USA*, **99**, 2445–2449.

Saltonstall K (2003) Microsatellite variation within and among North American lineages of *Phragmites australis*. *Molecular Ecology*, **12**, 1689–1702.

Santamaria L (2002) Why are most aquatic plants widely distributed? Dispersal, clonal growth and small-scale heterogeneity in a stressful environment. *Acta Oecologica*, **23**, 137–154.

Sarvala J (2005) Littoistenjärven ekologisen tilan kehitys ja hoitovaihtoehdot. Turun yliopiston Biologian laitoksen julkaisuja 24, Turun yliopiston biologian laitos, Turku.

Schaal BA, Gaskin JF, Caicedo AL (2003) Phylogeography, haplotype trees, and invasive plant species. *Journal of Heredity*, **94**, 197–204.

Schoen DJ, Brown AHD (1991) Intraspecific variation in population gene diversity and effective population size correlates with the mating system in plants. *Proceedings of the National Academy of Sciences, USA*, **88**, 4494–4497.

Schwartz S, Zhang Z, Frazer KA, Smit A, Riemer C, Bouck J, Gibbs R, Hardison R, Miller W (2000) PipMaker – a web server for aligning two genomic DNA sequences. *Genome Research*, **10**, 577–586.

Schwartz MK, McKelvey KS (2009) Why sampling scheme matters: the effect of sampling scheme on landscape genetic results. *Conservation Genetics*, **10**, 441–452.

Sculthorpe CD (1967) The biology of aquatic vascular plants. Edward Arnold, London.

Shaw J, Lickey EB, Beck JT, Farmer SB, Liu W, Miller J, Siripun KC, Winder CT, Schilling EE, Small RL (2005) The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany*, **92**, 142–166.

Shaw J, Lickey EB, Schilling EE, Small RL (2007) Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *American Journal of Botany*, **94**, 275–288.

Shea K, Chesson P (2002) Community ecology theory as a framework for biological invasions. *Trends in Ecology and Evolution*, **17**, 170–176.

Shine R, Brown GP, Phillips BL (2011). An evolutionary process that assembles phenotypes through space rather than through time. *Proceedings of the National Academy of Sciences, USA*, **108**, 5708–5711.

Shinozaki K, Ohme M, Tanaka M, Wakasugi T, Hayashida N, Matsubayashi T, Zaita N, Chunwongse J, Obokata J, Yamaguchi-Shinozaki K, Ohto C, Torazawa K, Meng BY, Sugita M, Deno H, Kamogashira T, Yamada K, Jusuda J, Takaiwa F, Kato A, Tohdoh N, Shimada H, Sugiura M (1986) The complete nucleotide sequence of the tobacco chloroplast genome: its gene organizations and expression. *The EMBO Journal*, **5**, 2043–2049.

Silvertown J, Charlesworth D (2001) Introduction to plant population biology. Blackwell Science Ltd, Oxford.

Simberloff D (2009) The role of propagule pressure in biological invasions. *Annual review of Ecology, Evolution and Systematics*, **40**, 81–102.

Simpson DA (1984) A short history of the introduction and spread of *Elodea* Michx in the British Isles. *Watsonia*, **15**, 1–9.

Simpson DA (1988) Phenotypic plasticity of *Elodea nuttallii* (Planch.) H. St John and *Elodea canadensis* Michx in the British Isles. *Watsonia*, **17**, 121–132.

Simpson DA (1990) Displacement of *Elodea canadensis* Michx by *Elodea nuttallii* (Planch.) H. St John in the British Isles. *Watsonia*, **18**, 173–177.

Soltis DE, Gitzendanner MA, Strenge DD, Soltis PS (1997) Chloroplast DNA intraspecific phylogeography of plants from the Pacific Northwest of North America. *Plant Systematics and Evolution*, **206**, 353–373.

Staden R, Judge DP, Bonfield JK (2003) Managing sequencing projects in the GAP4 environment. In: *Introduction to bioinformatics. A theoretical and practical approach* (eds Krawetz SA, Womble DD), pp. 393–410. Human Press Inc, Totawa.

Strayer DL, Eviner VT, Jeschke JM, Pace ML (2006) Understanding the long-term effects of species invasions. *Trends in Ecology and Evolution*, **21**, 645–651.

Taberlet P, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology*, **17**, 1105–1109.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, **28**, 2731–2739.

Tang J, Xia H, Cao M, Zhang X, Zeng W, et al. (2004) A comparison of rice chloroplast genomes. *Plant Physiology*, **135**, 412–420.

Thiébaut G, Di Nino F (2009) Morphological variations of natural populations of an aquatic macrophyte *Elodea nuttallii* in their native and in their introduced ranges. *Aquatic Invasions*, **4**, 311–320.

Trewick SA, Morgan-Richards M, Chapman HM (2004) Chloroplast DNA diversity of *Hieracium pilosella* (Asteraceae) introduced to new Zealend: reticulation, hybridization, and invasion. *American Journal of Botany*, **91**, 73–85.

Tsumura Y, Suyama Y, Yoshimura K (2000) Chloroplast DNA inversion polymorphism in populations of *Abies* and *Tsuga*. *Molecular Biology and Evolution*, **17**, 1302–1312.

Tsutsui ND, Suarez AV, holway DA, Case TJ (2000) Reduced genetic variation and the success of an invasive species.

*Proceedings of the National Academy of Sciences, USA*, **97**, 5948–5953.
Tsyusko OV, Smith MH, Sharitz RR, Glenn TC (2005) Genetic and clonal diversity of two cattail species, *Typha latifolia* and *T. angustifolia* (Typhaceae), from Ukraine. *American Journal of Botany*, **92**, 1161–1169.
Vanderpoorten A, Lambinon J, Tignon M (2000) Morphological and molecular evidence of the confusion between *Elodea callitrichoides* and *E. nuttallii* in Belgium and Northern France. *Belgian Journal of Botany*, **133**, 41–52.
Vitousek PM (1990) Biological invasions and ecosystem processes: towards an integration of population biology and ecosystem studies, *Oikos*, **57**, 7–13.
Vitousek PM, D'antonio CM, Loope LL, Rejmánek M and Westbrooks R (1997) Introduced species: a significant component of human-caused global change. *New Zealand Journal of Ecology*, **21**, 1–16.
Walker B, W Steffen (1997) An overview of the implications of global change for natural and managed terrestrial ecosystems. *Conservation Ecology* 1: http://www.consecol.org/vol1/iss2/art2.
Walker NF, Hulme PE, Hoelzel AR (2003) Population genetics of an invasive species, *Heracleum mantegazzianum*: implications for the role of life history, demographics and independent introductions. *Molecular Ecology*, **12**, 1747–1756.
Wang R-J, Cheng C-L, Chang C-C, Wu C-L, Su T-M, Chaw S-M (2008) Dynamics and evolution of the inverted repeat-large single copy junctions in the chloroplast genomes of monocots. *BMC Evolutionary Biology*, **8**, 36.
Ward SM, Gaskin JF, Wilson LM (2008) Ecological genetics of plant invasion: What do we know? *Invasive Plant Science & Management*, **1**, 98–109.
Waycott M, Walker DI, James SH (1995) Genetic uniformity in *Amphibolis antarctica*, a dioecious seagrass. *Heredity*, **76**, 578–585.
Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.
Wilcove DS, Rothstein D, Dubow J, Phillips A, Losos E (1998) Quantifying the relative importance of habitat destruction, alien species, pollution, overexploitation, and disease. *BioScience*, **48**, 607–615.
Williams DA, Overholt WA, Cuda JP, Hughes CR (2005) Chloroplast and microsatellite DNA diversities reveal the introduction history of Brazilian peppertree (*Shcinus terebinthifolius*) in Florida. *Molecular Ecology*, **14**, 3643–3656.
Williamson M (2006) Explaining and predicting the success of invading species at different stages of invasion. *Biological Invasions*, **8**, 1561–1568.
Williamson M, Fitter A (1996) The varying success of invaders. *Ecology*, **77**, 1661–1666.
Wolfe KH, Li W-H, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proceedings of the National Academy of Sciences of the USA*, **84**, 9054–9058.
Wright S (1951) The genetical structure of populations. *Annals of Eugenics*, **15**, 323–354.
Wyman SK, Jansen RK, Boore JL (2004) Automatic annotation of organellar genomes with DOGMA, *Bioinformatics*, **20**, 3252–3255.
Xu J, Li W, Liu G, Zhang L, Liu W (2007) Inter-specific competition between two submerged macrophytes, *Elodea nuttallii* and *Hydrilla verticillata*. *Journal of Plant Ecology*, **31**, 83–92
Yang M, Zhang X, Liu G, Yin Y, Chen K, Yun Q, Zhao D, Al-Mssallem IS, Yu J (2010) The complete chloroplast genome sequence of date palm (*Phoenix dactylifera* L.). *Plos One*, **5**, e12762.
Young HA, Lanzatella CL, Sarath G, Tobias CM (2011) Chloroplast genome variation in upland and lowland switchgrass. *Plos One*, **6**, e23980.
Yuan Y-W, Olmstead RG (2008) A species-level phylogenetic study of the *Verbena* complex (Verbenaceae) indicates two independent intergeneric chloroplast transfers. *Molecular Phylogenetics and Evolution*, **48**, 23–33.
Zane L, Bargelloni L, Patarnello T (2002) Strategies for microsatellite isolation: a review. *Molecular Ecology*, **11**, 1–16