Factors affecting the genotypic and genetic diversity of the dioecious clonal plant *Cirsium arvense* at the metapopulation level

Dissertation

zur Erlangung des akademischen Grades

doctor rerum naturalium (Dr. rer. nat.)

vorgelegt der

Mathematisch-Naturwissenschaftlich-Technischen Fakultät
(Fachbereich Biologie)
der Martin-Luther-Universität Halle-Wittenberg

von Frau Magali Solé

geb. am: 13.08.1975 in: Maubeuge (Frankreich)

Gutachterin bzw. Gutachter:

1. Prof. Dr. habil. Roland Brandl, Philipps-Universität Marburg

2. Prof. Dr. habil. Isabell Hensen, Universität Halle-Wittenberg

3. Prof. Dr. habil. Birgit Ziegenhagen, Philipps-Universität Marburg

Halle an der Saale, 16.12.2003 (Tag der öffentlichen Verteidigung)

Acknowledgements

Table of contents

General introduction

Genotypic and genetic diversity in natural populations

Natural populations of animals and plants are rarely composed of randomly distributed alleles[*] but reflect a complex mosaic of frequent and infrequent genotypes[**] (Murawski and Hamrick, 1990 ; Aspinwall and Christian, 1992; Lokker et al., 1994; Lynch and Milligan, 1994; Hänfling and Brandl, 1998; Schläpfer and Fischer, 1998; Ayres and Ryan, 1999). The relative abundance of these genotypes is not fixed, but may vary over the life span of populations. Hence, natural populations of plants and animals are genetically structured in space and time (Ewing, 1979; Chevillon et al., 1995; Hossaert-McKey et al., 1996). Such evidence for spatial and temporal genetic structures clearly emerged during the last three decades with the development of molecular markers and the access to the genotype of individuals. Then, with the increasing body of genetic data, research began to focus on the identification of the ecological and evolutionary processes responsible for such genetic structures. In plant species, Loveless and Hamrick (1984) reviewed the effect of several species traits on population genetic structure. They looked among others at the effect of breeding systems, floral morphology, pollination mechanisms, dispersal modes, life cycles, and size of populations. The main result of this review was that the breeding system is the principal factor designing the genetic structure of plant populations.

Genotypic and genetic diversity in clonal plants

In clonal plants, which have a mixed breeding system (sexual-asexual), clonality may affect the genetic variation and structure of natural populations. As clonal growth does not produce any genetic variation, the question whether clonal species harbour the same genetic diversity as non-clonal species has been often discussed (Williams, 1975; Harper, 1977; Abrahamson, 1980). An increasing amount of literature on this topic has demonstrated that a wide range of genetic diversity can also be found in populations of clonal plants (Ellstrand and Roose, 1987; Hamrick and Godt, 1990; Eckert and Barrett, 1993; Widen et al., 1994; Herlihy and Eckert, 2002).

---

[*] **Allele:** one of the different forms of a gene or DNA sequence that can exist at a single locus
[**] **Genotype**: the actual alleles present in an individual.

Nevertheless, although the genetic variation and structure of clonal plant populations have been intensively studied, the maintenance, evolution, and factors influencing clonal diversity remain poorly understood (Eckert, 1999). It is obvious that only descriptions of the genetic patterns are not enough to understand the structuring of genetic variation within and between populations (McLellan *et al.*, 1997). In contrast, an investigation of genetic variation combined with an investigation of factors responsible for its creation and maintenance (i.e. processes that have led to these patterns) might be more relevant to relate the genetic diversity to the ecology of the studied organism.

Factors influencing genotypic and genetic diversity

An important factor that must be taken into account when one studies the genetic diversity of natural populations is the temporal heterogeneity. It has already been demonstrated that temporal heterogeneity can influence the genetic variation within and between populations (Ewing, 1979; Tilman, 1984; Hossaert-McKey et al., 1996; Stanton et al., 1997). However, it remains difficult to follow natural populations through long periods. A bypass to this difficulty can be to investigate natural populations having contrasting biotic and abiotic selective pressures. For example, ecological succession is a case often studied in ecology where biotic and abiotic selective pressures change with the age of populations (Quinn et al., 1972; Beasleigh and Yarranton, 1974; Ross and Quinn, 1977; Escarré and Thompson, 1991; Piquot et al., 1998). Looking at natural populations of different successionnal stages may help to understand the evolution of genetic or life-history characters over time.

Another important source of variation that acts on genetic diversity of natural populations comes from the herbivory selective pressures. It has already been demonstrated that herbivores and pathogens can affect the genetic diversity of plants (Mopper et al., 1991; Krabel and Petercord, 2000; Koskela, 2002; Chen et al., 2001). For example, in Douglas-fir (*Pseudotsuga menziesii*), Chen et al. (2001) demonstrated that phenotypically resistant trees to defoliation by the western spruce budworm (*Choristoneura occidentalis*) were less heterozygote but had more rare alleles compared with susceptible trees. However, it is not clear whether differences in genetic variation of the host are a consequence or a cause induced by the biotic selection pressures, and the impact that herbivores and pathogens have on the neutral genetic diversity of populations remains unknown.

Although most of the variation expressed in genetic characters is neutral, variation expressed in life-history characters is usually non-neutral (i.e. induces changes in fitness). Life-history characters are assumed to undergo selection. Thus, the variation observed in life-history characters can be considered to result from heterogeneous selective pressures. For this reason, a mixed approach that combines the comparison of both characters (neutral genetic and selected) is important to localise selective pressures that act on natural populations (Bonnin et al., 1996).

Finally, most plant or animal species are organised as a set of local populations that interact via the dispersal of individuals. Such a spatial dynamic organisation of populations is called "metapopulation" (i.e. a population of populations which go extinct locally and recolonise; Levin, 1970). In a metapopulation organization, local populations do not obligatorily experience the same selective pressures. Therefore, certain characters of plants or animals might evolve differently in each local population (Hanski, 1998). For example, along a successional gradient, populations of early and late successional stage experience different local selection pressures. In clonal plants, which have a mixed reproduction system, these different local selection pressures might lead to different strategies of reproduction. In early stages of succession, the selection regime may favour individuals which display the best capacities for clonal reproduction (i.e. the best competitors, Ronce and Olivieri, 1997; Piquot et al., 1998). Sexual reproduction will be counter-selected at the population level. But because seeds are often the only mean of long distance dispersal for numerous clonal plants, sexual reproduction will be selected at the metapopulation level to assure the foundation of new populations in more favourable habitats.

In this case, the evolutionary equilibrium of characters related to the reproduction system (e.g. dispersal and reproductive effort) is not achieved at the population level, but at a higher level (the metapopulation level) while each local population is in disequilibrium (Olivieri et al., 1990; Olivieri and Gouyon, 1997).
Consequently, it is important to consider the metapopulation level while one studies traits that do not experience the same selective pressure during the different demographic stages of the population life span. Metapopulation approaches are also particularly relevant to study species in fragmented habitats (Hanski, 1998).

Cirsium arvense

This work is about the genetic diversity of a clonal plant, *Cirsium arvense*. *C. arvense* (Asteraceae), also called "Creeping thistle" in Europe or "Canada thistle" in North America, is one of the most frequent and most successful perennial weeds throughout Eurasia (Figure 1 from Meusel and Jäger, 1992).
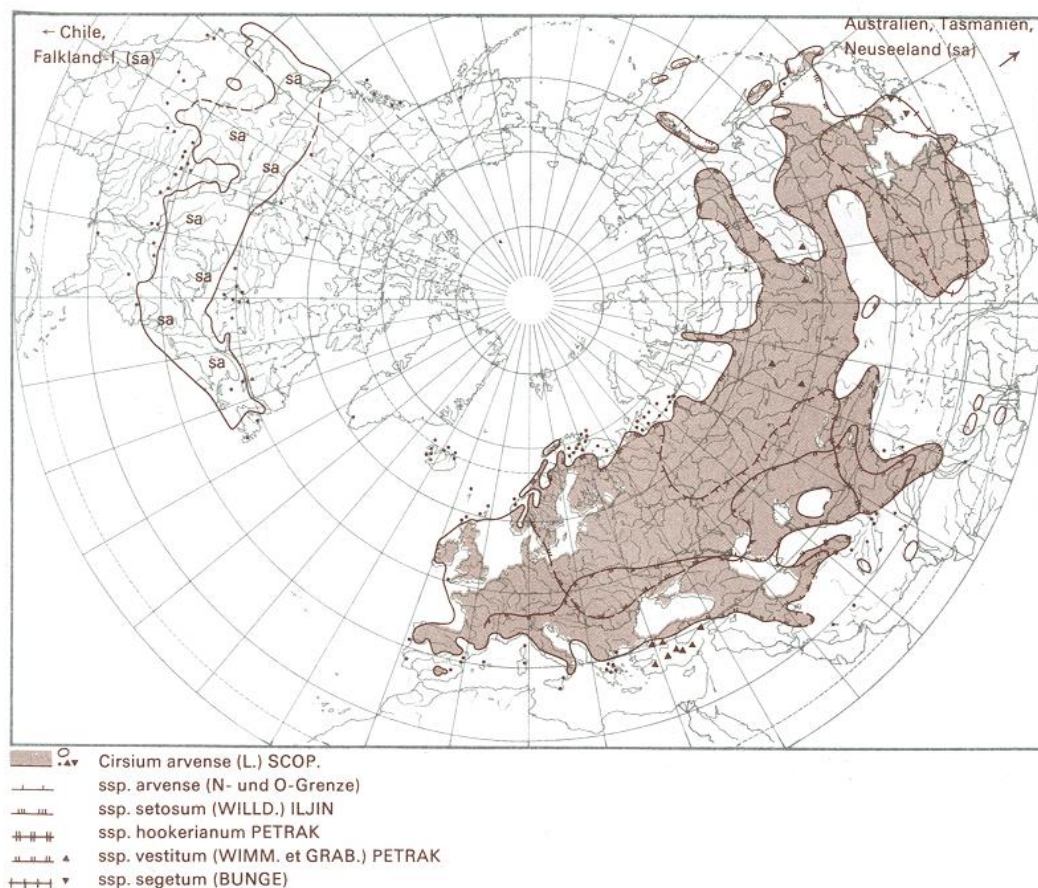
Figure 1: Geographical distribution of *Cirsium arvense* and subspecies within the northern hemisphere. Source: Meusel & Jäger (1992)

C. arvense occurs over a wide range of elevations from sea level (up to 2500m). The wide distribution of *C. arvense* suggests that the plant is adaptable to many soil types (Donald, 1994). *C. arvense* is an early successional species that emerges from seeds or root fragments shortly after disturbance. *C. arvense* grows best in open moist sunny sites. It usually grows in waste places, roadsides, fallows, or especially in arable fields and meadows or sites that have been impacted by disturbance (e.g. road building, fire, landslides…). Hence, the increasing intervention of humans in natural landscapes, which mostly opens the vegetation, multiplies suitable sites for *C. arvense*. The species is nowadays omnipresent in agricultural or semi-

natural landscapes. Furthermore, roads, streams and ditches offer new corridors for invasion. Thus, *C. arvense* invasion of native rangelands appears a problem especially for farmers who are worried about its invasion to arable fields.

Because of its economic importance (Moore, 1975; Donald, 1990) *C. arvense* has been intensively studied during the last century (Correns, 1916; Bakker, 1960; Lloyd and Myall, 1976; Hunter et al., 1985; Kay, 1985; Lalonde and Roitberg, 1989; Lalonde and Roitberg, 1994; Ang et al., 1995; Heimann and Cussans, 1996).

As most of perennial plants, *C. arvense* has a mixed reproduction system. Plants combine two modes of reproduction: sexual reproduction through seeds and clonal reproduction through vegetative propagation. *C. arvense* life cycle is presented in Figure 2.

Plants develop new roots and underground shoots in January and begin to elongate in February. New shoots of established *C. arvense* plants begin to emerge when average weekly temperature is at least equal to 8°C (Nuzzo, 2000). Plants remains short until a long day period activate stem elongation and flowering (Moore, 1975). The blooming period is bright, varies from place to place, but appears between May and October in Europe.
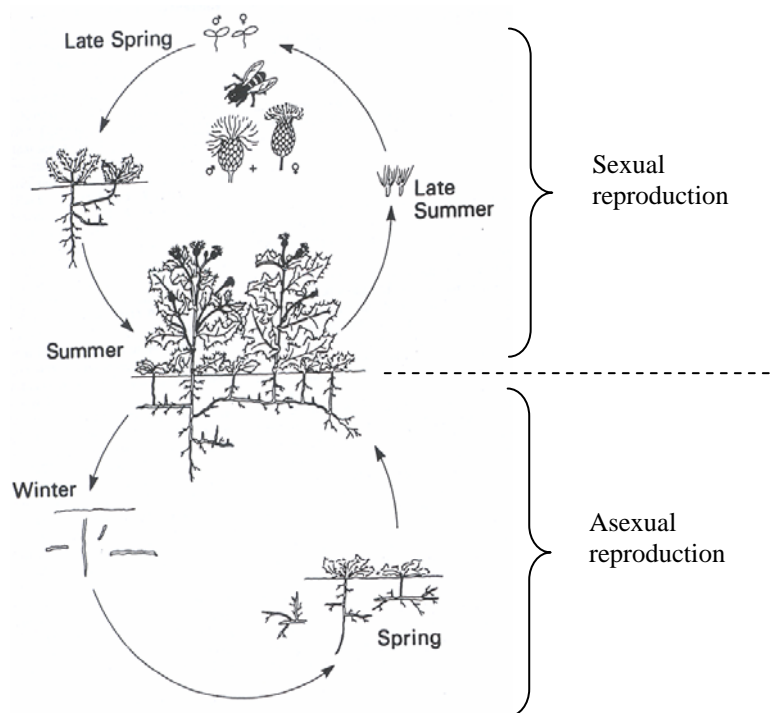
Figure 2: Life cycle of *Cirsium arvense* from Heimann & Cussans (1996)

In *C. arvense*, blooming capability as well as others life-history characters might be altered by the tephritid fly *Urophora cardui* (Peschken and Harris, 1975; Peschken *et al.*, 1982; Heeb *et al.*, 1999). *Urophora cardui* is a gall-forming insect which attacks *C. arvense* (Figure 3). In this plant-herbivore system, *U. cardui* is able to occur only within a narrow range of environmental conditions (Peschken *et al.*, 1997)

as the clonal *C. arvense* as well as *C. setosum* are the only host plants of this species (Frenzel et al., 2000). Adults of *U. cardui* emerge in early summer (June to July) and females lay eggs into suitable shoots of the host. With the development of the larvae the plant is forced to produce conspicuous multilocular stem galls (Peschken and Harris, 1975, Peschken *et al.*, 1982; Peschken and Derby, 1992).

Figure 3 :gall of *U. Cardui* on *C. arvense* (from Hegi, 1987)

C. arvense has an open breeding system. The plant is reported to be subdioecious-dioecious. Male plants are morphologically hermaphrodites, with a vestigial ovary (Delannay, 1979). *C. arvense* is insect pollinated, and pollinators mostly observed are bees (honey bees and wild bees), which can cover relatively important area (Walther-Hellwig and Frankl, 2000). Although *C. arvense* is mostly known for its efficient vegetative reproduction, the species also produces high quantity of achenes. Mayer (2000) found that thistle stands produced between 7600 and 21000 well-developed achenes per m² in fallow sites. *C. arvense* achenes are plumed, and can be easily dispersed by wind. Most achenes germinate in spring after the year in which they are produced (Bakker, 1960), but some achenes might germinate in the year, produce basal leaves before winter and then emerge to flower the next spring. The soil seed bank does not usually contain large numbers of *C. arvense* achenes (Heimann and Cussans, 1996). No seedling establishment is usually observed in dense natural populations (Bakker, 1960; Bostock and Benton, 1983).

C. arvense has a reputation to have a vigorous clonal propagation. *C. arvense* does not form rhizomes. Vegetative reproduction (or clonal growth) is realised through efficient laterals roots (Figure 4). Horizontal roots are very productive (six and 12m per year in Moore (1975) and Bostock and Benton (1979) respectively) and give rise to numerous aerial shoots. *C. arvense* allocates most of its resources to vegetative propagation. Total allocation of dry weight to sexual reproduction was only 7% for *C. arvense* grown in pots (Bostock and Benton, 1979). New *C. arvense* plants can also form from root fragments as short as 6 mm (Nadeau and Vanden Born, 1989). Therefore, *C. arvense* can easy survive disturbance by resprouting from buried root and stems fragments. Long distance dispersal by vegetative propagules is not common in *C. arvense*, as roots are deeply buried in the soil.

Figure 4: Two shoots of *C. arvense* connected by horizontal roots.

Although a lot of information has been accumulated about the ecology and biology of *C. arvense*, data about its genetic diversity are missing. The aim of this Ph-D was to conduct a mixed approach combining ecology and molecular biology to identify, on a landscape scale, the main factors acting and designing the genetic diversity of *C. arvense* in natural populations.

Outline

The primary step of an ecological and genetic study of a clonal plant is the identification of clones. In clonal plants, each individual shoot (ramet) can produce offspring through sexual and/or asexual reproduction. Physiological individuals do not fit with genotypic ones, and the identification of clones in the field is often impossible. Hence, the first objective of this work was the identification of clones. For this purpose, we required to molecular biology. The choice of the molecular marker we used, as well as essential information about molecular markers in general are developed in the appendix.

Chapter I concerns the genotypic and genetic diversity of the common weed *Cirsium arvense*. In many clonal species, seedling establishment is restricted to early successional stages when recruitment is still possible. Then, one expects that adapted genotypes become dominant and genotypic and genetic diversity should decrease with time. In order to check this hypothesis, I compared genotypic and genetic diversity of *C. arvense* populations in early and late successional stage. I used highly polymorphic AFLP markers and found that contrary to the general expectation, genotypic diversity, clonal evenness and genetic diversity (i.e. molecular variance) did not differ significantly between founder and established populations (respectively: U = 24, *P* > 0.3; U = 26.5, *P* > 0.3; U = 23; *P* > 0.44).
However, the most surprising result of this study was the extremely high genetic structure among populations (Φ_{sc} = 0.64) which occurred without loss of diversity. It is proposed that genetic differentiation among populations might result from founder effects and early selection in the seedling stage, while because of the particular reproductive system of the species (dioecious) seeds that are produced must be highly variable and therefore their recruitment (even sporadically) might contribute to maintain a high genotypic and genetic diversity. In *C. arvense* genotypic and genetic diversity mainly seem to reflect the status built up during the early stage of succession.

Chapter II looks at the influence of the phytophage insect *Urophora cardui* on *Cirsium arvense* neutral genotypic and genetic diversity. Several studies in controlled conditions have demonstrated that *U. cardui* has profound negative effects on life history traits of its host (*C. arvense*). The purpose of this chapter is to test rather these negative effects observed on an individual scale influence the neutral genotypic and genetic diversity of natural populations of the host plant. To investigate the effect of *U. cardui* on *C. arvense* neutral genotypic and genetic diversity, I used AFLP markers eight populations of *C. arvense*

having different history of infestation by *U. cardui*. Half of the populations were in the last 5 years infested by *U. cardui*, whereas the other half was not. The results were the following: average genotypic diversity and clonal evenness did not differ between infested and not-infested populations ($U = 6$; $P > 0.56$; $U = 5.5$; $P > 0.46$) for clonal evenness). Molecular variance due to infestation state of populations was also not significant (1.81 ± 1.05 versus 2.22 ± 0.82; $P > 0.11$) and explained less than five percent of the total variance. Hence, the results suggest that selection imposed by *U. cardui* on *C. arvense* was weaker on a population and/or metapopulation scale than it was supposed in experimental studies. This can be explained by the complex spatio-temporal population dynamics of the *C. arvense-U. cardui* system.

Chapter III further analyses the impact of ecological succession on natural populations of *C. arvense*. Whereas chapter I attempted to look at the impact of ecological succession on neutral molecular characters, chapter III is based on life-history characters, which are supposed to be selected. Certain life-history characters, especially those related to the breeding system, are generally considered to be able to evolve rapidly when affected by a change in the selection regime (Fisher, 1930; Reznick et al., 1990; Li and Margolies, 1994; Cody and Overton, 1996). Therefore, this chapter focuses on the reproductive effort and on the dispersal potential of *C. arvense*. Female and male of *C. arvense* plants were collected in three populations of early successional stage and in ten populations of late successional stage. Succession caused no change in the dispersal potential of *C. arvense* (length of the pappus), but important phenotypic changes in characters related to the reproductive outputs. These changes included a significant decrease in the number of flowering shoots per population (-48%; $P < 0.01$), the number of flower heads per shoot ($P < 0.01$ in females; not significant in males) and the number of flowers per flower head ($P < 0.1$ in females and not significant in males). Data on achene mass and germination rate show no relation to the successional stage of plants ($F = 0.07$; $P = NS$ and $F = 0.14$ $P > 0.71$). Based on recent theoretical investigations on the reproductive effort in metapopulation context (Ronce and Olivieri, 1997), it is suggested that the differences found in reproductive effort may result from evolutionary changes in the genotypic composition of populations because of increasing intraspecific selection pressures when the habitat is maturing. This idea of selection during early successional stage is supported by previous genotypic analysis of some populations with AFLP markers (Solé et al., 2004). However, we found no changes in the dispersal potential (length of the pappus) of *C. arvense*. Chapter III also discuss the role of the length pappus to estimate the dispersal potential in this species.

Own contribution

Since the scientific papers are co-authered by several people, the own contribution (%) of Magali Solé is listed subsequently.

Chapter I: Genotypic and genetic diversity of the common weed *Cirsium arvense* (Asteraceae). Together with Walter Durka, Sabine Eber, Roland Brandl.

1.	Sampling and data acquisition	80%
2.	Data analysis	90%
3.	Writing paper	70%

Chapter II: Does history of parasitism by *Urophora cardui* influence the genotypic and genetic diversity of *Cirsium arvense*? Together with Walter Durka, Sabine Eber, Roland Brandl.

1.	Sampling and data acquisition	70%
2.	Data analysis	90%
3.	Writing paper	80%

Chapter III: Environmental study of reproductive and dispersal efforts of the common weed *Cirsium arvense* on a metapopulation level. Together with Yves Piquot, Walter Durka, Roland Brandl

1.	Sampling and data acquisition	90%
2.	Data analysis	90%
3.	Writing paper	70%

Literature cited

Abrahamson WG. 1980 Demography and vegetative reproduction. Pages 89-106 *in* Solbrig OT ed. Demography and evolution of plant populations. Blackwell, Oxford.

Ang BN, LT Kok, GI Holtzman, DD Wolf 1995 Canada thistle (*Cirsium arvense* (L.) Scop.) response to density of *Cassida rubiginosa* Müller (Coleoptera: Chrysomelidae) and plant competition. Biological control 5:31-38.

Aspinwall N, T Christian 1992 Clonal structure, genotypic diversity, and seed production in populations of *Filipendula rubra* (Rosaceae) from the north central United States. American Journal of Botany 79:294-299.

Ayres DR, FJ Ryan 1999 Genetic diversity and structure of the narrow endemic *Wyethia reticulata* and its congener *W. bolanderi* (Asteraceae) using RAPD and allozyme techniques. American Journal of Botany 86:344-353.

Bakker D. 1960 A comparative life-history study of *Cirsium arvense* (l.) Scop., *Tussilago farfara* L, the most troublesome weeds in the newly reclaimed polders of the former Zuiderzee. Pages 205-222 *in* Harper JL ed. The biology of weeds. Blackwell Scientific Publications, Oxford.

Beasleigh WJ, GA Yarranton 1974 Ecological strategy and tactics of *Equisetum sylvaticum* during a postfire succession. Canadian Journal of Botany 52:2299-2318.

Bonnin I, JM Prosperi, I Olivieri 1996 Genetic markers and quantitative genetic variation in *Medicago truncatula* (Leguminosae) : A comparative analysis of population structure. Genetics 143:1795-1805.

Bostock SJ, RA Benton 1979 The reproductive strategies of five perennial Compositae. Journal of Ecology 67:91-107.

Bostock SJ, RA Benton 1983 Dry weight costs and establishment of seeds and vegetative propagules. Acta Oecologica 4:61-69.

Chen Z, TE Kolb, KM Clancy, VD Hipkins, LE DeWald 2001 Allozyme variation in interior Douglas-fir: association with growth and resistance to western spruce budworm herbivory. Canadian Journal of Forest Research 31:1691-1700.

Chevillon C, N Pasteur, M Marquine 1995 Population structure and dynamics of selected genes in the Mosquito *Culex pipiens*. Evolution 49:997-1007.

Cody ML, JM Overton 1996 Short-term evolution of reduced dispersal in island plant populations. Journal of Ecology 84:53-61.

Correns C 1916 Untersuchungen über Geschlechtsbestimmung bei Distelarten. Sitzungberichte der koeniglich preussischen Akademie der Wissenschaften 20:448-477.

Delannay X 1979 Evolution of male sterility mechanisms in gynodioecious and dioecious species of *Cirsium* (*Cynareae, Compositae*) Plant Systematic and Evolution 132: 327-332.

Donald WW 1990 Management and control of Canada thistle (*Cirsium arvense*). Reviews of Weed Science 5:193-250.

Donald WW 1994 The biology of Canada Thistle (*Cirsium arvense*). Reviews of Weed Science 6:77-101.

Eckert CG 1999 Clonal plant research: proliferation, integration, but not much evolution. American Journal of Botany 86:1649-1654.

Eckert CG, SCH Barrett 1993 Clonal reproduction and patterns of genotypic diversity in *Decodon verticillatus* (Lythraceae). American Journal of Botany 80:1175-1182.

Ellstrand NC, ML Roose 1987 Patterns of genotypic diversity in clonal plant species. American Journal of Botany 74:123-131.

Escarré J, JD Thompson 1991 The effects of successional habitat variation and time of flowering on seed production in *Rumex acetosella*. Journal of Ecology 79:1099-1112.

Ewing EP 1979 Genetic variation in a heterogeneous environment VII. Temporal and spatial heterogeneity in infinite populations. American Naturalist 114:197-212.

Fisher RA 1930 The genetical theory of natural selection. Clarendon Press, Oxford.

Frenzel M, S Eber, S Klotz, R Brandl 2000 Ecological comparisons across geographical distributions: the thisle gall fly *Urophora cardui* (Diptera:Tephritidae) on two different *Cirsium* hosts. European Journal of Entomology 97:183-189.

Hamrick JL, MJ Godt. 1990 Allozyme diversity in plant species. Pages 43-63 *in* Brown AHD, MT Clegg, AL Kahler, and BS Weir eds. Plant population genetics, breeding and genetic resources. Sinauer Associates Inc., Sunderland.

Hänfling B, R Brandl 1998 Genetic variability, population size and isolation of distinct populations in the feshwater fish *Cottus gobio* L. Molecular Ecology 7:1625-1632.

Hanski I 1998 Metapopulations dynamics. Nature 396:41-49.

Harper JL. 1977 Reproduction and growth. Pages 647-675 Population Biology of Plants. Academic Press, New York.

Heeb P, I Werner, AC Mateman, M Kölliker, MWG Brinkhof, CM Lessells, H Richner 1999 Ectoparasite infestation and sex-biased local recruitment of hosts. Nature 400:63-65.

Hegi G 1987 Illustrierte Flora von Mitteleuropa. Verlag Paul Parey ed., Berlin.

Heimann B, GW Cussans 1996 The importance of seeds and sexual reproduction in the population biology of *Cirsium arvense* - a literature review. Weed Research 36:493-503.

Herlihy CR, CG Eckert 2002 Genetic cost of reproductive assurance in a self-fertilizing plant. Nature 416:320-323.

Hossaert-McKey M, M Valero, D Magda, M Jarry, J Cuguen, P Vernet 1996 The evolving genetic history of a population of *Lathyrus sylvestris*: evidence from temporal and spatial genetic structure. Evolution 50:1808-1821.

Hunter JH, AI Hsiao, GI McIntyre 1985 Some effects of humidity on the growth and development of *Cirsium arvense*. Botanical gazette 146:483-188.

Kay QON 1985 Hermaphrodites and Subhermaphrodites in a reputedly dioeciuos plant, *Cirsium arvense* (L.) scop. New Phytologist 100:457-472.

Koskela T 2002 Variation in life-history traits among *Urtica dioica* populations with different history in parasitism by the holoparasitic plant *Cuscuta europaea*. Evolutionary Ecology 16:433-454.

Krabel D, R Petercord 2000 Genetic diversity and bark physiology of the European beech (*Fagus sylvatica*): a coevolutionary relationship with the beech scale (*Cryptococcus fagisuga*). Tree Physiology 20:485-491.

Lalonde RG, BD Roitberg 1989 Ressource limitation and offspring size and number trade-offs in *Cirsium arvense* (Asteraceae). American Journal of Botany 76:1107-1113.

Lalonde RG, BD Roitberg 1994 Mating system, life-history, and reproduction in Canada thistle (*Cirsium arvense*; Asteraceae). American Journal of Botany 81:21-28.

Levin R. 1970 Extinction. Some Mathematical Problems in Biology. American Mathematical Society, Providence, Rhode Island.

Li J, DC Margolies 1994 Responses to direct and indirect selection on aerial dispersal behaviour and life-history traits in T*etranychus urticae*. Heredity 72:10-22.

Lloyd DG, AJ Myall 1976 Sexual dimorphism in *Cirsium arvense*. Annals of Botany 40:115-123.

Lokker C, D Susko, L Lovett-Doust, J Lovett-Doust 1994 Population genetic structure of *Vallisneria america*, a dioecious clonal macrophyte. American Journal of Botany 81:1004-1012.

Loveless MD, JL Hamrick 1984 Ecological determinants of genetic structure in plant populations. Annual review of ecology and systematics 15:65-95.

Lynch M, BG Milligan 1994 Analysis of population genetic structure with RAPD markers. Molecular Ecology 3:91-99.

Mayer F 2000 Long distance dispersal of weed diaspores in agricultural landscapes - The Scheyern approach. Aachen.

McLellan AJ, D Prati, O Kaltz, B Schmid. 1997 Structure and analysis of phenotypic and genetic variation in clonal plants. Pages 185-210 *in* De Kroon H and J Van Groenendael eds. The ecology and evolution of clonal plants. Backhuys Publishers, Leiden.

Meusel H, EJ Jäger 1992 Vergleichende Corologie der zentraleuropäischen Flora. Jena.

Moore RJ 1975 The biology of Canadian weeds. 13. *Cirsium arvense* (L.) Scop. Canadian Journal of Plant Sciences 55:1033-1048.

Mopper S, JB Mitton, TG Whitman 1991 Genetic differentiation and heterozygocity in Pinyon pine associated with resistance to herbivory and environmental stress. Evolution 45:989-999.

Murawski DA, JL Hamrick 1990 Local genetic and clonal structure in the tropical terrestrial bromeliad, *Aechmea magdalenae*. American Journal of Botany 77:1201-1208.

Nadeau LB, WH Vanden Born 1989 The root system of Canada thistle. Canadian Journal of Plant Science 69:1199-1206.

Nuzzo V 2000 Element stewardship abstract: *Cirsium arvense*. In: Weeds on the web: The Nature Conservancy wildland invasive species program http://tncweeds.ucdavis.edu/esadocs/cirsarve.html

Olivieri I, D Couvet, PH Gouyon 1990 The genetics of transient populations: research at the metapopulation level. TREE 5:207-210.

Olivieri I, P-H Gouyon. 1997 Evolution of migration rate anf other traits: the metapopulation effect. Pages 293-323 *in* Hanski I and ME Gilpin eds. Metapopulation biology. Academic Press, San Diego.

Peschken DP, JL Derby 1992 Effect of *Urophora cardui* (L.) (Diptera: Tephritidae) and *Ceutorhynchus litura* (F.) (Coleptera: Curculionidae) on the weed Canada thistle, *Cirsium arvense* (L.) Scop. The canadian entomologist 124:145-150.

Peschken DP, JL Derby, DT Spann. 1997 Establishment of *Urophora cardui* (Diptera: Tephritidae) on canada thistle, *Cirsium arvense* (Asteraceae), and colony development in relation to habitat and parasitoids in Canada. Pages 53-66 *in* Dettner K, G Bauer, and W Völkl eds. Vertical food web interactions. Springer, Berlin.

Peschken DP, DB Finnamore, AK Watson 1982 Biocontrol of the weed Canada thistle: releases and development of the gall fly *Urophora cardui* (Diptera: Tephritidae) in Canada. The canadian entomologist 114:349-357.

Peschken DP, P Harris 1975 Host specificity and biology of *Urophora cardui* (Diptera: Tephritidae). A biocontrol agent for canada thistle (*Cirsium arvense*). The canadian entomologist 107:1101-1110.

Piquot Y, D Petit, M Valero, J Cuguen, P de Laguerie, P Vernet 1998 Variation in sexual and asexual reproduction among young and old populations of the perennial macrophyte *Sparganium erectum*. Oikos 82:139-148.

Quinn JA, J Rotsettis, DE Fairbrothers 1972 Inflorescence characters and reproductive proficiency in *Danthonia sericea* populations. American Journal of Botany 59:627-631.

Reznick DN, H Bryga, JA Endler 1990 Experimentally induced life-history evolution in natural population. Nature 346:357-359.

Ronce O, I Olivieri 1997 Evolution of reproductive effort in a metapopulation with local extinctions and ecological succession. American Naturalist 150:220-221.

Ross FH, JA Quinn 1977 Phenology and reproductive allocation in *Andropogon scoparius* (Graminaceae) populations in communities of different successional stages. American Journal of Botany 64:535-540.

Schläpfer F, M Fischer 1998 An isozyme study of clone diversity and relative importance of sexual and vegetative recruitment in the grass *Brachypodium pinnatum*. Ecography 21:351-360.

Solé M, W Durka, S Eber, R Brandl 2004 Genotypic and genetic diversity of the common weed *Cirsium arvense* (Asteraceae). International Journal of Plant Sciences *(in press)*.

Stanton ML, C Galen, J Shore 1997 Population structure along a steep environmental gradient: Consequences of flowering time and habitat variation in the snow buttercup, *Ranunculus adoneus*. Evolution 51:79-94.

Tilman GD 1984 Plant dominance along an experimental nutrient gradient. Ecology 65:1445-1453.

Walther-Hellwig K, R Frankl 2000 Foraging habitats and foraging distances of bumblebees, *Bombus* spp. (Hym., Apidae), in an agricultural landscape. Journal of Applied Entomology 124:299-306.

Widen B, N Cronberg, M Widen 1994 Genotypic diversity, molecular markers and spatial distribution of genets in clonal plants, a literature survey. Folia Geobotanica Phytotaxonomica 29:245-263.

Williams GC 1975 Sex and Evolution. Princeton University Press, Princeton.

Chapter I

Genotypic and genetic diversity of the common weed *Cirsium arvense* (Asteraceae)

With Walter Durka, Sabine Eber & Roland Brandl

Abstract: In many clonal species, seedling establishment is restricted to early successional stages when recruitment is still possible. Then, one expects that adapted genotypes become dominant and genotypic and genetic diversity should decrease with time. We investigated genotypic and genetic diversity within recently founded and established populations of the common weed *Cirsium arvense.* We used highly polymorphic AFLP markers. All populations were multiclonal and highly diverse (proportion of distinguishable genotypes was 0.73 ± 0.25 (mean $\pm$ SD)). Clonal evenness was variable and ranged from 0.2 to 1. Independent of successional stage, we found on the small geographic scale of our study (< 5 km) a considerable differentiation between populations ($\Phi_{sc} = 0.63$). This amount of differentiation was similar between founder and established populations and could result from selection in the early stage of selection, as well as founder effects. Contrary to the general expectation, genotypic and genetic diversity were maintained through time, and molecular variance did not differ between successional stages (1.9 ± 0.89 versus 2.5 ± 1.41). We suggest that this pattern is a consequence of the particular reproductive system of *C. arvense* that combines clonality with dioecy. The combination of clonal reproduction with recruitment of sexually outcrossed seedlings in the first years allows the species to perform efficient colonisation even with founder effects, to undergo selection without loss of diversity, and to persist locally. This strategy appears to be very efficient in *C. arvense* and may have contributed to the worldwide success of this species.

Key words: Amplified Fragments Length Polymorphism (AFLP), *Cirsium arvense*, clonal plant, genotypic diversity, molecular variance, population differentiation, succession.

Introduction

Clonal plants can reproduce by sexual and asexual reproduction. Whereas sexual reproduction accounts for recombination and dispersal, clonality propagates the same genotype locally. Therefore it is often suggested that spatial distribution of genotypic and genetic diversity within and across populations reflects the balance between clonal growth and successful sexual reproduction (Sackville Hamilton et al., 1987; Schmid, 1990; McLellan et al., 1997). In correlation with these two modes of reproduction, it is important to distinguish between genotypic and genetic diversity. Genotypic diversity is usually called clonal diversity in clonal plants and stands for the number of genotypes within populations, whereas genetic diversity represents the variability between genotypes. Clonal reproduction together with selection and mutation will principally affect the genotypic diversity and spatial distribution of clones within populations, whereas sexuality together with gene flow and recombination will act more on the genetic variability within and among populations.

C. arvense is a long-lived perennial weed, abundant in agricultural land. It has a wide distribution in Europe from the Atlantic ocean to the Ural mountains. *C. arvense* is also widely distributed in North America where it has been introduced in the 17th century. Since the introduction, *C. arvense* has spread throughout Canada and the United States where it is one of the most tenacious and economically important weeds (Donald, 1990). *C. arvense* has a mixed sexual-asexual reproduction system. It reproduces vegetatively by means of very efficient lateral roots which assure a rapid expansion of ramets. Moore (1975) reported on two independent studies which found a spread via lateral roots of 6 m per season, whereas in Bostock and Benton (1979) another study even note a spread of 12 m a year. *C. arvense* is often described as a subdioecious-dioecious species. Female plants are strictly female. Male plants possess vestigial ovaries and are morphological hermaphrodites, but are considered to be functional males (Moore, 1975). *C. arvense* is insect-pollinated and produces plumed achenes with a long pappus. Natural long distance dispersal by vegetative propagules is limited (except if roots are transport by man e.g. in soil), hence new populations must be founded by seeds.

Because of its dioecious breeding system, genotypic diversity of seedling founding a new population is expected to be high compared to other none strictly out-crossed species. However, after the establishment of seedlings, genotypes with vigorous clonal growth may become favoured, as clonality is a better strategy to spread and colonize a habitat for plants

having a mixed sexual asexual reproduction system. Hence, after some time only a few genotypes should dominate the habitat due to competitive exclusion (Eriksson, 1993; Gray, 1987). This process should then lead to a general decrease of genotypic diversity when populations are aging i.e. in later successional stages.

Studies that followed genotypic diversity through time (e.g. Hartnett and Bazzaz, 1985 in *Solidago canadensis*) or compared genotypic diversity between sites which differed in successional stage (e.g. Maddox et al., 1989 in *Solidago altissima*; McNeilly and Roose, 1984 in *Lolium perene*) showed a decline in the number of genets over time. Studies, which compared populations of clonal plants in habitats with different disturbance regime reached similar conclusions. Populations in unstable habitats tended to have a higher genotypic diversity compared to populations in stable habitats (Piquot et al., 1996, Xie et al., 2001). However, in some cases no decrease of genotypic diversity was found (Verburg et al., 2000). Some studies which followed genetic diversity through time with quantitative markers supported the theoretical predictions (Aarssen and Turkington, 1985). But in some others no decrease of genetic diversity over time was observed (Taylor and Aarssen, 1988; Hartnett et al., 1987).

The present study is based on molecular data. We looked at genotypic and genetic diversity within recently founded as well as established populations of *C. arvense*. Because of their high power of resolution in distinguishing genotypes, we used Amplified Fragment Length Polymorphism (Vos et al. 1995). Our paper aims to address the following questions: (1) Is there a decrease in genotypic diversity and evenness of *C. arvense* populations with time? (2) How is the genetic variation partitioned within and among populations?

Material and methods

Sampling

All sampling populations belong to a 15 km² rural area located in Southern Germany (11°50'E 49°35'N). This area is a complex mosaic of roadsides, meadows, cultivated and abandoned agricultural fields, wastelands. In September 1999, samples were collected from 16 natural populations of *C. arvense* from two contrasting successional stages: founder and established. We identified each successional stage according to information available from previous studies (Eber and Brandl, 1994; Eber and Brandl, 1996) and ecological criteria. Founder populations of *C. arvense* occurred in habitats where the vegetation was in a typical early successional stage. The vegetation cover was below 75 % and the plant community was exclusively composed of herbaceous species. *C. arvense* was the dominant plant species. Shoots of *C. arvense* were easily recognizable as freshly germinated as they significantly differ from shoots from mature successional stages (Solé in prep). Moreover, during our previous attempts to map all populations of *C .arvense* in the same area these populations were not existent (Eber and Brandl, 2003). Thus, we are sure that these populations date from the year of our study. Because of the difficulty to find founder populations, we were able to sample only six populations. Established populations of *C. arvense* were sampled in old fallows. Within theses communities, *C. arvense* populations were in an advanced or regressive successional phase. There, the vegetation cover was totally closed and woody species were present. *C. arvense* shoots were particularly high and ligneous. All these established populations were already mapped in this advanced stage in 1994. Although the date of foundation of these populations is unknown, we can guess they have been found for 10 years.

In 1999, the size (number of shoots) of all 16 populations was recorded. For populations having less than 100 shoots, the number of shoots was counted and rounded to the nearest ten. For bigger populations, average number of plants was counted in one square meter and extrapolated to the population area covered by *C. arvense*. In these cases, the number of shoots was rounded to the nearest hundreds or thousands.

Populations differed markedly in their size, density and patchiness. Therefore, we could not apply a regular spatial sampling design to all populations. Moreover, as we were interested in estimating the genotypic diversity of *C. arvense* rather than the spatial extent of certain genotypes we adopted a random sampling strategy. We randomly sampled fresh leaves of

C. arvense shoots according to population extent and patchiness in order to have a sampling effort reflecting the ramet density (Table 1).

Table 1. Demographic, genotypic and genetic characteristics of 16 German populations of the clonal dioecious *Cirsium arvense*.

Population name	Successional stage	Population size (number of shoots)	Number of plants analysed (N)	Number of genotypes detected (G)	Genotypic diversity (i)	Evenness index ($E_{1/D}$)	Molecular variance
F1	Founder	30	7	6	0.86	0.78	2.12
F2	Founder	40	15	7	0.47	0.28	0.48
F3	Founder	200	17	16	0.94	0.89	4.59
F4	Founder	10	7	7	1.00	1	3.45
F5	Founder	100	19	9	0.47	0.2	1.78
F6	Founder	120	10	10	1.00	1	2.63
E1	Established	500	53	42	0.79	0.58	2.48
E2	Established	100	20	12	0.60	0.29	1.93
E3	Established	30	20	20	1.00	1	3.01
E4	Established	3000	54	52	0.96	0.93	2.92
E5	Established	400	28	13	0.46	0.29	1.07
E6	Established	40	12	6	0.50	0.33	1.06
E7	Established	10	8	2	0.25	0.2	0.75
E8	Established	800	11	11	1.00	1	2.42
E9	Established	8000	18	13	0.72	0.5	0.76
E10	Established	100	8	5	0.63	0.5	2.39

DNA extraction and AFLP

We extracted DNA from young leaves with CTAB (Doyle and Doyle, 1988). DNA quality and concentration were estimated from 5.5 µl of the extract on a 0.8 % agarose gel. For AFLP (Vos et al., 1995) we used 0.5 µg of DNA per sample. We followed the *Ligation and Preselective Amplification Modules for Small Plant Genomes* procedure from Applied Biosystems except that the digestion and the ligation were performed in a MWG-Biotech Primus 96 thermocycler at 37°C during 2 hours. An initial screening using 64 selective primer combinations was performed on a random sample of 10 individuals across all sampled populations. From that analysis the two primer combinations EcoRI-ACC / MseI-CTG and EcoRI-ACG / MseI-CTT appeared to be sufficiently polymorphic to discriminate clones within populations. Fragments were separated on an ABI PRISM 310 genetic analyser with 100 units as a minimum height threshold for peak detection. Data were imported to the

analysing software *Genotyper*, but we were not able to genotype our data automatically because of the strong sensitivity of the *Genotyper* program and the heterogeneity among runs. Such a lane-to-lane variation has been already observed in previous studies (De Riek et al., 1999). Following De Riek *et al.*, we only used the Genotyper program to produce a preliminary presence / absence matrix, which was subsequently checked manually. For the present study we used 93 polymorphic loci, 42 for EcoRI-ACC / MseI-CTG and 51 EcoRI-ACG / MseI-CTT. Samples showing the same multilocus AFLP phenotype were considered to be the same genotype.

Analyses

AFLP data were analyzed at two hierarchical levels: within and among populations. As an estimator of the intrapopulation genotypic diversity we used the proportion of distinguishable genotypes (i = G/N, Ellstrand and Roose, 1987), where G is the number of genotypes and N the number of sampled shoots. Clonal evenness was calculated as the relative abundance of each genotype within a population. A number of evenness indices are available and there is no consensus on which one is the best (Smith and Wilson, 1996). We choose the evenness index

$E_{1/D} = \dfrac{1/D}{G}$ (Williams, 1964), which is based on Simpson's index ($D = \sum_{i=1}^{G} p_i^2$ where p_i is the

relative abundance of the ith genotype). $E_{1/D}$ ranges from 0 to 1. High values characterize populations with an even distribution of clones; low values characterize populations dominated by only few genotypes. We compared the genotypic diversity and the evenness index between the two successional stages with a Mann-Whitney U-test.

For each population we calculated the molecular variance as the average number of mismatches of bands within a population (sum of mismatches divided by 2N (N-1), where N equals the number of samples), and we tested for correlations between the molecular variance and population size. We tested for correlations between the population size and the three factors molecular variance, genotypic diversity and clonal evenness, using Spearman rank correlations.

To study the partitioning of genetic variance among populations we conducted an analysis of molecular variance (AMOVA, Excoffier et al., 1992) with the program Arlequin (Schneider et al., 2000). We constructed a hierarchical model with genotypes nested within populations and populations nested within successional stage (Table 1). Variance components (σ_a^2, σ_b^2, σ_c^2; see Table 3 for explanations), the sum of all squared differences and analogs of F-

statistics (Φ) were calculated. The significance of estimated parameters was tested by a permutation procedure (Excoffier et al., 1992). Φ_{ct} and σ_a^2 were tested by random permutation of genotypes of whole populations across successional stages. In this study Φ_{ct} estimated the successional stage effect. Φ_{sc} and σ_b^2 were tested by random permutation of individuals across populations but within the same successional stage. Φ_{sc} estimated the population differentiation and is the equivalent of the Wright's F_{st} index (Wright, 1965). Φ_{st} and σ_c^2 were tested by permuting randomly AFLP phenotypes among populations and between successional stages. We used a matrix correlation test with 1000 permutations to test whether the matrix of genetic distances ($\Phi_{sc}/(1-\Phi_{sc})$; Rousset, 1997) was correlated with the matrix of geographic distances (Table 2). The genetic distance matrix was calculated by the program Arlequin as a pairwise population Φ_{sc} matrix (Schneider et al., 2000).

Table 2: Matrices of pairwise genetic and geographic distances among 16 German populations of *Cirsium arvense*. Genetic distances (Φ_{sc}) are presented in the lower part of the table; the geographic distances (km) in the upper part.

	F1	F2	F3	F4	F5	F6	E1	E2	E3	E4	E5	E6	E7	E8	E9
F1		0.13	1.83	0.73	0.85	0.88	0.7	0.72	0.6	0.59	0.82	0.8	0.83	2.95	0.8
F2	0.87		1.8	0.75	0.75	0.83	0.62	0.76	0.7	0.58	0.85	0.87	0.94	3.05	0.88
F3	0.34	0.55		2.28	2.48	2.03	0.53	0.46	0.44	0.43	0.58	0.53	0.65	1.68	0.57
F4	0.53	0.77	0.22		1.25	1.53	0.56	0.52	0.5	0.47	0.74	0.66	0.79	3	0.73
F5	0.71	0.80	0.44	0.43		0.63	0.63	0.56	0.64	0.6	0.78	0.77	0.85	3.75	0.77
F6	0.61	0.73	0.35	0.51	0.64		0.55	0.66	0.57	0.5	0.77	0.74	0.81	3.48	0.77
E1	0.70	0.62	0.53	0.56	0.63	0.55		0.56	0.62	0.49	0.64	0.6	0.77	4.13	0.65
E2	0.72	0.76	0.46	0.52	0.56	0.66	0.56		2.15	1.8	0.66	0.72	0.82	0.65	0.75
E3	0.60	0.70	0.44	0.50	0.64	0.57	0.62	0.58		0.38	0.73	0.65	0.75	2.15	0.67
E4	0.59	0.58	0.43	0.47	0.60	0.50	0.51	0.51	0.50		0.63	0.6	0.72	1.88	0.64
E5	0.82	0.85	0.58	0.74	0.78	0.77	0.64	0.66	0.73	0.63		0.28	2.18	2.83	2.2
E6	0.80	0.87	0.53	0.66	0.77	0.74	0.60	0.72	0.65	0.60	0.78		2.13	2.75	1.75
E7	0.83	0.94	0.65	0.79	0.85	0.81	0.77	0.82	0.75	0.72	0.90	0.89		0.63	3.43
E8	0.63	0.77	0.45	0.53	0.71	0.60	0.61	0.62	0.57	0.56	0.70	0.71	0.82		0.72
E9	0.80	0.88	0.57	0.73	0.77	0.77	0.65	0.75	0.67	0.64	0.82	0.83	0.92	0.72	
E10	0.69	0.82	0.43	0.43	0.68	0.62	0.60	0.60	0.49	0.54	0.79	0.71	0.79	0.55	0.79

Results

From our AFLPs we identified 93 polymorphic loci. We distinguished 231 haplotypes (clones) across the 307 sampled *C. arvense* shoots; 86 % of these clones were found only once. Genotypes shared by several individuals (14%) always occurred in the same population. The number of clones per population ranged from 2 to 52. The mean genotypic diversity over all populations was $i = 0.73 \pm 0.25$ (mean $\pm$ SD). However genotypic diversity differed considerably between populations (range: 0.25 – 1.0; Table 1). The clonal evenness ($E_{1/D}$) was also very variable and ranged from 0.2 to 1 (mean $E_{1/D} = 0.6 \pm 0.32$). Both indices (i and E) were highly correlated (Spearman rank-correlation coefficient $r_s = 0.98$, $P < 0.0001$). Genotypic diversity and the clonal evenness did not differ significantly between founder and established populations (genotypic diversity: $U = 24$, $P > 0.3$; evenness: $U = 26.5$, $P > 0.3$) and were not correlated with the population size (genotypic diversity: $r = 0.1$; $P > 0.70$; evenness $r = 0.03$; $P > 0.92$).

The average molecular variance was 2.1 ($\pm$ 1.1). Established populations had a lower mean molecular variance than the founder ones (1.9 ± 0.89 versus 2.5 ± 1.41) but the difference was not significant ($U = 23$; $P > 0.44$). Molecular variance was not correlated with the population size (Figure 1).
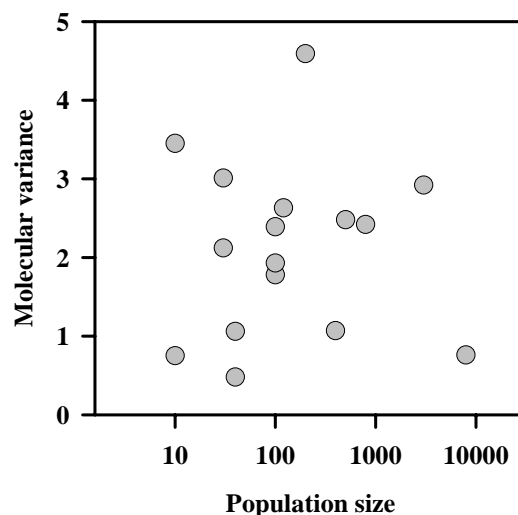

Figure 1: Population size (number of shoots counted per population) in 16 German populations of *Cirsium arvense* was not correlated with the molecular variance ($r_s = 0.09$; $P > 0.72$).

Most of the haplotype diversity was found among populations within successional stages ($\sigma_b^2 = 60$ %; Table 3), which results in a high Φ_{sc} value (0.63). Nevertheless, a considerable amount of diversity was found within populations ($\sigma_c^2 = 35$ %; Table 3). The variance due to differentiation between founder and established populations was significant but explained only a small part of the total variance ($\sigma_a^2 = 5$ %; Table 3).

Table 3. Hierarchical analysis of molecular variance testing for differentiation between successional stages, among populations within successional stages and within populations. The significance tests are based on 1000 permutations.

Variation Component		Variance	% total	Significance	Φ-statistics
Between successional stages	σ_a^2	0.58	5	$P < 0.01$	$\Phi_{ct} = 0.05$
Among populations within successional stages	σ_b^2	7.46	60	$P < 0.001$	$\Phi_{sc} = 0.63$
Within populations	σ_c^2	4.44	35	$P < 0.001$	$\Phi_{st} = 0.64$

Pairwise genotypic distances were not correlated to geographic distances (matrix correlation $= 0.25$, $P > 0.2$; Figure 2).
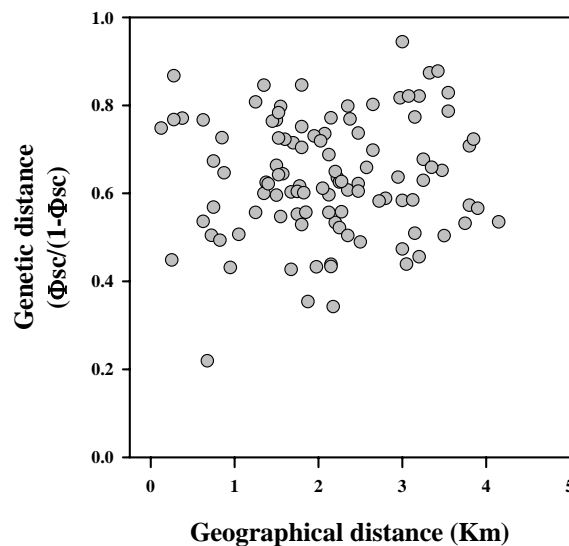

Figure 2: The relation between the genetic and the geographical distances in 16 German populations of *Cirsium arvense* show no correlation between the two matrices (matrix correlation $= 0.12$; $P > 0.18$). High population differentiation observed in *C. arvense* is not due to geographical distances.

Discussion

Genotypic diversity and clonal evenness of *C. arvense* varied greatly among the sixteen sampled populations. The average genotypic diversity was equal to 0.73. Direct comparisons with data from the literature is difficult to interpret as reviews are based on different sample sizes, spatial scales, molecular markers and compare plant species with different life and phylogenetic histories (Gitzendanner and Soltis, 2000). However, compared to the most widely cited reviews of genetic diversity patterns in clonal plants (Ellstrand and Roose 1987 and Widen et al., 1994) average genotypic diversity in *C. arvense* is much higher (0.73 against 0.17). This big difference might mostly result from the highest power of resolution of the DNA-based molecular markers (we had 93 polymorphic loci), whereas the above-mentioned reviews are mostly based on allozymes studies, which offer fewer loci. Nevertheless, the high genotypic diversity found in *C. arvense* is confirmed by the fact that 100 % of the sampled populations were multiclonal, against an average of 62 % in Ellstrand and Roose (1987) and Widen et al. (1994).

Like genotypic diversity, the clonal evenness may vary greatly among clonal species. Populations can be mono- or multiclonal (Piquot et al., 1996; Eckert and Barrett, 1993; Aspinwall and Christian, 1992; McClintock and Waterway, 1993). For multiclonal populations, almost all patterns of genotypic diversity combined with clonal evenness can be found. (Ayres and Ryan, 1997, Ivey and Richards, 2001b; Gabrielsen and Brochmann, 1998). In *C. arvense*, the genotypic diversity is high and populations are not dominated by one clone, a pattern which was also found in other species (Chung and Epperson, 1999; Xie et al., 2001; Auge et al., 2001).

Population differentiation has frequently been analyzed with allozymes. Three main results emerged: (1) Differentiation between populations is rather similar in plant species with mixed and purely sexual reproduction (G_{st} values 0.21 and 0.23, respectively, Hamrick and Godt, 1990) (2) most variation occurs within populations (Baur and Schmid, 1996) (3) population differentiation is lower in outbreeding than in inbreeding species (G_{st} values < 20% and > 50%, respectively). Studies based on dominant markers (RAPD) lead to the same conclusions (Bussel, 1999).

In *C. arvense* most of the genetic variation was among populations (60 % of the total variation). All haplotypes were strictly local. The Φ_{sc}-statistic (equivalent in our study to Wrights' F_{st} value) was very high ($\Phi_{sc} = 0.64$). Furthermore, differentiation among

populations was not correlated to geographic distance. A high differentiation among populations independent of geographical distance, like in *C. arvense*, is more common in rare, partially selfing, locally dispersed, or gene flow limited species (Fischer and Matthies, 1998; Travis et al., 1996; Schmidt and Jensen, 2000). Thus, the high population differentiation independent of geographical distance in *C. arvense* is surprising, as *C. arvense* is an abundant outbreeding species.

In plants with mixed reproduction systems, clonal propagation can strongly bias the estimation of F-statistics (McLellan et al., 1997). For example, in the clonal *Cladium jamaicense* the F_{st} based on the ramet (i.e. all aerial shoots coming from the clonal propagation of a single root) was 0.68, whereas the Fst based on the genets (genetic individuals) was 0.035 (Ivey and Richards, 2001a). In *C. arvense*, the analysis on the genet level also resulted in a high population differentiation ($\Phi_{sc} = 0.53$; $P < 0.001$). Therefore, clonality is not the reason of population differentiation. A result, which was already evident from the fact that populations were not dominated by one or few clones.

Besides clonality, high population differentiation independent of geographical distance can also result from strong selection (Endler, 1986), as well as from founder effects acting together with drift, low gene flow among populations or low seedling recruitment within populations (Slatkin, 1977, Whitlock, 1992).

Similarity among genotypes within populations, and thus dissimilarity among populations, is expected to increase in small and isolated populations as a consequence of genetic drift, bottlenecks and inbreeding (Hartl and Clark, 1989; Barrett and Kohn, 1991). A decline in molecular variance in small populations was frequently found, e.g. in the non-clonal *Gentianella germanica* (Fischer and Matthies, 1997) as well as in the clonal herb *Ranunculus reptans* (Fischer et al., 2000). In our study, population size was neither correlated with the molecular variance, nor with the genotypic diversity. The same correlation performed on established populations only, which are more likely to have experienced genetic drift, was not significant. Hence, *C. arvense* populations need not have been strongly affected by non-selective processes through time.

Gene flow among populations can be distinguished into pollen and seeds. The pollinators of *C. arvense*, mainly bumble bees, are very mobile (Walther-Hellwig and Frankl, 2000). The observation that seeds were produced in all populations even with 100 % of females (Solé unpublished data) and the fact that apomixis does not occur (Solé personal observation)

confirm an efficient transport of pollen among populations. However, low seed dispersal and absence of seedling recruitment in *C. arvense* are potential causes that generate and maintain differentiation. Despite producing plumed seeds, the effectiveness of seed dispersal in *C. arvense* is arguable. The pappus often breaks off the seed, so that at a distance of 1000 m only 0.2% of trapped pappi carried a seed (Bakker, 1960; Bostock and Benton, 1979). Like in other clonal plants (Wolf et al., 2000, Eriksson, 1992) the recruitment of *C. arvense* seedlings is possible only during the early phases of the colonization. *C. arvense* seedlings are very susceptible to shading and competition (Bakker, 1960). Our observations as well as reports in the literature indicate that the recruitment of *C. arvense* seedlings is not possible in natural or artificial plant communities with a dense plant cover (Bostock and Benton, 1983; Solé unpublished data). Thus, once populations have reached closed canopy seedling recruitment will stop. This phenomenon could then reinforce the population differentiation (see Gibson and Wheelwright 1995).

Two potential driving forces remain to explain the surprisingly high genetic differentiation among *C. arvense* populations: founder effects and selection. We tested for founder effects (i.e. restricted-source origin of founder populations) by comparing the level of genetic differentiation among founder and established populations (Whitlock and McCauley, 1990). The Φ_{st} values were 0.55 ($P < 0.001$) for founder populations versus 0.64 ($P < 0.001$) for established populations. As differentiation among founder populations was already high, this result may indicate foundation of new populations by a non random sample of the propagule pool of *C. arvense*.

Selection in *C. arvense* can also be an important factor, which designs the genetic differentiation between populations. Strong selection (e.g. for genotypes that have the highest capacity for clonal growth) can occur in the early stages of succession. During this phase, many genotypes may become established but seedling mortality is high (Brandl et al. unpublished data).

In *C. arvense*, founder effects and selection are not mutually exclusive. However, founder effects and strong selection usually tend to decrease the within-population genetic variability (Pannell and Charlesworth, 1999; Endler, 1986). In our study, genotypic and genetic diversity did not decrease from founder to established populations. We offer the following explanation. We sampled plants in September and thus the founder populations already passed the seedling stage. If mortality and strong selection occur during the seedling stage, our sampling scheme was not able to retrieve the decrease of genetic diversity as the founder populations had

already passed the filter of a first selection phase. As long as the selection regime is not spatially autocorrelated this will lead to high differentiation between populations.

However, the maintenance of genetic diversity through time could also be fostered by the obligatory outcrossed breeding system of *C. arvense*. Comparable results (high genetic diversity coupled with a high differentiation of populations) were already found in other clonal dioecious species (Sherman-Broyles et al., 1992 in *Rhus* species). Because of dioecy, *C. arvense* seeds are strictly outcrossed and thus must be highly variable. Their recruitment during the early stages of succession could compensate the loss of diversity within population, and therefore maintain a high level of neutral genetic variability. Recently theoretical studies about the genetic diversity in clonal plants lead to similar conclusions (Bengtsson, 2003). In his model, Bengtsson looked at the "Genotypic Identity" of population (i.e. the probability that two randomly sampled adult individuals from a population have the same genotype) depending on population growth parameters, rate of sexuality and recruitment of sexually derived offspring. From his simulations, Bengtsson proposed that clonal populations possess an effective "memory" of their earlier genetic history, in the way that "a population which started by a number of sexually derived propagules may thus retain its initial genotypic variation for a very long period of time, even if its later reproduces almost exclusively asexually".

Conclusion

In contrast to the expected decline of genotypic diversity over time, genotypic diversity and evenness did not vary between founder and established populations of *C. arvense*. We found on a small geographic scale at the same time an extremely high population differentiation together with a high within population variability. We interpret these patterns of genotypic and genetic diversity as the result of the particular reproductive system of the species. Founder effects and early selection in the seedling stage may contribute to genetic differentiation among populations. Then, combination of recruitment of sexually outcrossed seedlings in the early stages of succession and clonal reproduction allows the species to perform efficient colonization and to persist locally. Diversity of genotypes, as well as genetic diversity, are maintained through time and mainly seem to reflect the status built up during the early stage of succession. This strategy appears to be very efficient in *C. arvense* and may have contributed to the worldwide success of this species.

Literature cited

Aarssen LW, R Turkington 1985 Within-species diversity in natural populations of *Holcus lanatus*, *Lolium perenne* and *Trifolium repens* from four different aged pastures. Journal of Ecology 73:869-886.

Aspinwall N, T Christian 1992 Clonal structure, genotypic diversity, and seed production in populations of *Filipendula rubra* (Rosaceae) from the north central United States. American Journal of Botany 79:294-299.

Auge H, B Neuffer, F Erlinghagen, R Grupe, R Brandl 2001 Demographic and random amplified polymorphic DNA analyses reveal high level of genetic diversity in clonal violet. Molecular Ecology 10:1811-1819.

Ayres DR, FJ Ryan 1997 The clonal and population structure of a rare endemic plant, *Wyethia reticulata* (Asteraceae): allozyme and RAPD analysis. Molecular Ecology 6:761-772.

Bakker D. 1960 A comparative life-history study of *Cirsium arvense* (l.) Scop., *Tussilago farfara* L, the most troublesome weeds in the newly reclaimed polders of the former Zuiderzee. Pages 205-222 *in* Harper JL ed. The biology of weeds. Blackwell Scientific Publications, Oxford.

Barrett SCH, JR Kohn. 1991 Genetic and evolutionary consequences of small population size in plants: Implication for conservation. Pages 3-10 *in* Falk D and K Holsinger eds. Genetics and conservation of rare plants. Oxford University Press, Oxford, UK.

Baur B, B Schmid. 1996 Spatial and temporal patterns of genetic diversity within species. Pages 169-201 *in* Gaston KJ ed. Biodiversity - a biology of numbers and difference. Blackwell, Oxford.

Bengtsson BO 2003 Genetic variation in organisms with sexual and asexual reproduction. Journal of Evolutionary Biology 16:189-199.

Bostock SJ, RA Benton 1979 The reproductive strategies of five perennial Compositae. Journal of Ecology 67:91-107.

Bostock SJ, RA Benton 1983 Dry weight costs and establishment of seeds and vegetative propagules. Acta Oecologica 4:61-69.

Bussel JD 1999 The distribution of random amplified polymorphic DNA (RAPD) diversity amongst populations of *Isotoma petraea* (Lobeliaceae). Molecular Ecology 8:775-789.

Chung MG, BK Epperson 1999 Spatial genetic structure of clonal and sexual reproduction in populations of *Adenophora grandiflora* (Campanulaceae). Evolution 53:1068-1078.

De Riek J, J Dendauw, M Mertens, M De Loose, J Heursel, E Van Bockstaele 1999 Validation of criteria for the selection of AFLP markers to assess the genetic variation of breeder's collection of evergreen azalas. Theoretical and Applied Genetics 99:1155-1165.

Donald WW 1990 Management and control of Canada thistle (*Cirsium arvense*). Reviews of Weed Science 5:193-250.

Doyle JJ, JL Doyle 1988 Isolation of plant DNA from fresh tissue. Focus 12:13-15.

Eber S, R Brandl 1994 Ecological and genetic spatial patterns of *Urophora cardui* (Diptera: Tephritidae) as evidence for population structure and biogeographical processes. Journal of Animal Ecology 63:187-199.

Eber S, R Brandl 1996 Metapopulation dynamics of the tephritid fly *Urophora cardui*: An evaluation of incidence-function model assumptions with field data. Journal of Animal Ecology 65:621-630.

Eber S, R Brandl 2003 Regional patch dynamics of the weed *Cirsium arvense*, and possible implications for plant-animal interactions. Journal of Vegetation Science 14:259-266.

Eckert CG, SCH Barrett 1993 Clonal reproduction and patterns of genotypic diversity in *Decodon verticillatus* (Lythraceae). American Journal of Botany 80:1175-1182.

Ellstrand NC, ML Roose 1987 Patterns of genotypic diversity in clonal plant species. American Journal of Botany 74:123-131.

Endler JA 1986 Natural selection in the wild. Princeton University Press, Princeton, New Jersey.

Eriksson O 1992 Evolution of seed dispersal in clonal plants. Oikos 63:439-448.

Eriksson O 1993 Dynamics of genets in clonal plants. TREE 8:313-316.

Excoffier L, PE Smouse, JM Quattro 1992 Analysis of Molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. Genetics 131:479-491.

Fischer M, R Husi, D Prati, M Peintinger, M van Kleunen, B Schmid 2000 RAPD variation among and within small and large populations of the rare clonal plant *Ranunculus reptans* (Ranunculaceae). American Journal of Botany 87:1128-1137.

Fischer M, D Matthies 1997 Mating structure and inbreeding and outbreeding depression in the rare plant *Gentianella germanica* (Gentianaceae). American Journal of Botany 84:1685-1692.

Fischer M, D Matthies 1998 RAPD variation in relation to population size and plant fitness in the rare *Gentianella germanica* (Gentianaceae). American Journal of Botany 85:811-819.

Gabrielsen TM, C Brochmann 1998 Sex after all: high levels of diversity detected in the arctic clonal plant *Saxifraga cernua* using RAPD markers. Molecular Ecology 7:1701-1708.

Gibson JP, N Wheelwright 1995 Genetic structure in a population of a tropical tree *Ocotea tenera* (Lauraceae): influence of avian seed dispersal Oecologia 103:49-54.

Gitzendanner MA, PS Soltis 2000 Patterns of genetic variation in rare and widespread plant congeners. American Journal of Botany 87:783-792.

Gray AJ. 1987 Genetic change during succession in plants. Pages 273-293 *in* Gray AJ, DE McCauley, and PJ Edwards eds. Colonization, succession and stability. Blackwell Scientific Publications, Oxford.

Hamrick JL, MJ Godt. 1990 Allozyme diversity in plant species. Pages 43-63 *in* Brown AHD, MT Clegg, AL Kahler, and BS Weir eds. Plant population genetics, breeding and genetic resources. Sinauer Associates Inc., Sunderland.

Hartl DL, AG Clark 1989 Principles of population genetics. Sinauer, Sunderland, Massachusetts, USA.

Hartnett DC, FA Bazzaz 1985 The genet and ramet population dynamics of *Solidago canadensis* in an abandoned field. Journal of Ecology 73:407-413.

Hartnett DC, BB Hartnett, FA Bazzaz 1987 Persistence of *Ambrosia trifida* populations in old fields and responses to successional changes. American Journal of Botany 74:1239-1248.

Ivey CT, JH Richards 2001a Genetic diversity of everglades sawgrass, *Cladium jamaicense* (Cyperaceae). International Journal of Plant Sciences 162:817-825.

Ivey CT, JH Richards 2001b Genotypic diversity and clonal structure of everglades sawgrass, *Cladium jamaicense* (Cyperaceae). International Journal of Plant Sciences 162:1327-1335.

Maddox DG, RE Cook, PH Winberger, S Gardescu 1989 Clone structure in four *Solidago altissima* (Asteraceae) populations: rhizome connections within genotypes. American Journal of Botany 76:318-326.

McClintock KA, MJ Waterway 1993 Patterns of allozyme variation and clonal diversity in *Carex lasiocarpa* and *C. pellita* (Cyperaceae). American Journal of Botany 80:1251-1263.

McLellan AJ, D Prati, O Kaltz, B Schmid. 1997 Structure and analysis of phenotypic and genetic variation in clonal plants. Pages 185-210 *in* De Kroon H and J Van Groenendael eds. The ecology and evolution of clonal plants. Backhuys Publishers, Leiden.

McNeilly T, ML Roose 1984 The distribution of perennial ryegrass genotypes in swards. New Phytologist 98:503-513.

Moore RJ 1975 The biology of Canadian weeds. 13. *Cirsium arvense* (L.) Scop. Canadian Journal of Plant Sciences 55:1033-1048.

Pannell JR, B Charlesworth 1999 Neutral genetic diversity in a metapopulation with recurrent local extinction and recolonization. Evolution 53:664-676.

Piquot Y, P Soumitou-Laprade, D Petit, P Vernet, JT Epplen 1996 Genotypic diversity revealed by allozymes and oligonucleotide DNA fingerprinting in French populations of the aquatic macrophyte *Sparganium erectum*. Molecular Ecology 5:251-258.

Rousset F 1997 Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. Genetics 145:1219-1228.

Sackville Hamilton NR, B Schmid, JL Harper 1987 Life-history concepts and the population biology of clonal organisms. Proceedings of the Royal Society of London B 232:35-57.

Schmid B 1990 Some ecological and evolutionary consequences of modular organization and clonal growth in plants. TREE 4:25-34.

Schmidt K, K Jensen 2000 Genetic structure and AFLP variation of remnant populations in the rare plant *Pedicularis palustris* (Scrophulariaceae) and its relation to population size and reproductive components. American Journal of Botany 87:678-689.

Schneider S, D Roessli, L Excoffier 2000 Arlequin ver. 2.000: a software for population genetic data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Sherman-Broyles SL, JP Gibson, JL Hamrick, MA Bucher, MJ Gibson 1992 Comparisons of allozyme diversity among rare and widespread *Rhus* species. Systematic botany 17:551-559.

Slatkin M 1977 Gene flow and genetic drift in a species subject to frequent local extinctions. Theoretical Population Biology 12:253-262.

Smith B, JB Wilson 1996 A consumer's guide to evenness indices. Oikos 76:70-82.

Taylor DR, LW Aarssen 1988 An interpretation of phenotypic plasticity in *Agropyron repens* (Gramineae). American Journal of Botany 75:401-413.

Travis SE, J Maschinski, P Keim 1996 An analysis of genetic variation in *Astragalus cremnophylax* var. *cremnophylax*, a critically endangered plant, using AFLP markers. Molecular Ecology 5:735-745.

Verburg R, J Maas, HJ During 2000 Clonal diversity in differently-aged populations of the pseudo-annual clonal plant *Circaea lutetiana* L. Plant Biology 2:646-652.

Vos P, R Hogers, M Bleeker, M Reijans, T van de Lee, M Hornes, A Frijters, J Pot, J Peleman, M Kuiper, M Zabeau 1995 AFLP: a new technique for DNA fingerprinting. Nucleic Acid Research 23:4407-4414.

Walther-Hellwig K, R Frankl 2000 Foraging habitats and foraging distances of bumblebees, Bombus spp. (Hym., Apidae), in an agricultural landscape. Journal of Applied Entomology 124:299-306.

Whitlock MC 1992 Temporal fluctuations in the demographic parameters and the genetic variance among populations. Evolution 46:608-615.

Whitlock MC, DE McCauley 1990 Some population genetic consequences of colony formation and extinction: genetic correlations within founding groups. Evolution 44:1717-1724.

Widen B, N Cronberg, M Widen 1994 Genotypic diversity, molecular markers and spatial distribution of genets in clonal plants, a literature survey. Folia Geobotanica Phytotaxonomica 29:245-263.

Williams CB 1964 Patterns in the balance of nature. Academic Press, London.

Wolf AT, RW Howe, JL Hamrick 2000 Genetic diversity and population structure of the serpentine endemic *Calystegia collina* (Convolvulaceae) in northern California. American Journal of Botany 87:1138-1146.

Wright S 1965 The interpretation of population structure by F-statistics with special regards to systems of mating. Evolution 19:395-420.

Xie Z, Y Lu, S Ge, D Hong, F Li 2001 Clonality in wild rice (*Oryza rufipogon*, Poaceae) and its implications for conservation management. American Journal of Botany 88:1058-1064.

Chapter II

Does history of parasitism by *Urophora cardui* influence the genotypic and genetic diversity of *Cirsium arvense*?

With Walter Durka, Sabine Eber & Roland Brandl

Abstract: Several studies in controlled conditions have demonstrated that the parasite (*U. cardui*) has profound negative effects on life history traits of its host (*C. arvense*). Such negative effects may affect the neutral genotypic and genetic diversity of natural populations of the host plant. To test this hypothesis, we investigated eight populations of *C. arvense* with different history of infestation by *U. cardui*. Half of the populations were in the last 5 years infested by *U. cardui*, whereas the other half was not. We investigated genotypic and genetic diversity with AFLP markers. Contrary to what we expected, average genotypic diversity, clonal evenness and molecular variance did not differ between infested and not-infested populations (0.73 ± 0.21 *versus* 0.78 ± 0.26 (U = 6; $P > 0.56$) for clonal diversity; 0.58 ± 0.27 *versus* 0.71 ± 0.34 (U = 5.5; $P > 0.46$) for clonal evenness). Molecular variance due to infestation state of populations was not significant (1.81 ± 1.05 versus 2.22 ± 0.82; $P > 0.11$) and explained less than five percent of the total variance. Hence, our results suggest that selection imposed by *U. cardui* on *C. arvense* was weaker on a population and/or metapopulation scale than it was supposed in experimental studies. This can be explained by the complex spatio-temporal population dynamics of the *C. arvense-U. cardui* system. Within each population, we did not find any significant correlation between the genetic dissimilarity matrix of *C. arvense* shoots and the corresponding "infestation state" matrix. In our study, *C. arvense* shoots appear to be randomly infested by *U. cardui.*

Introduction

The tephritid fly *Urophora cardui* is a gall-forming insect which attacks the common weed *Cirsium arvense*. In this plant-herbivore system, *U. cardui* is able to occur only within a narrow range of environmental conditions (Peschken *et al.*, 1997) as the clonal *C. arvense* as well as *C. setosum* are the only host plants of this species (Frenzel et al., 2000). Adults of *U. cardui* emerge in early summer (June to July) and females lay eggs into suitable shoots of the host. With the development of the larvae the plant is forced to produce conspicuous multilocular stem galls (Peschken and Harris, 1975, Peschken *et al.*, 1982; Peschken and Derby, 1992).

C. arvense is a perennial clonal weed of considerable economic importance. (Moore, 1975; Donald, 1990). Furthermore, it is an aggressive invader throughout North American continent since the 17th century. Hence, the potential of *U. cardui* as a biocontrol agent for *C. arvense* was intensively investigated during the last decades (Peschken *et al.*, 1997; Peschken *et al*., 1982). Experimental studies demonstrated that *U. cardui* can influence the fecundity and survival of infested plants. For example *U. cardui* may reduce the production of mature seed-heads (Laing, 1977) as well as below and above ground biomass (65% and 47% respectively). In extreme *U. cardui* may cause the death of infested ramet (Peschken and Harris, 1975). Furthermore, Gange and Nice (1997) showed that the larvae manipulate the nitrogen metabolism of infested ramets in order to maintain an optimal nutrient level within galls. Hence, galls act as sinks (Shorthouse and Watson, 1976). Thus, ramets with galls are at a disadvantage, especially in nitrogen-limited habitats.

Although the impact of phytophagous insects on host fitness is well-known, their effect on other life history traits is less obvious. (see Thomas *et al.*, 2000 for review). For example, phytophages may influence the dispersal of their host (Heeb *et al.*, 1999). In the *C. arvense-U. cardui* system, the growth of galls leads to a reduction of plant height (Peschken and Harris, 1975; Peschken *et al*., 1982). In turn this may reduce dispersal distances of seeds (Sheldon and Burrows, 1973) which may affect genotypic diversity and the genetic structure of the host populations. Furthermore, reduced fecundity and survival of infested plants may reduce effective population size and thus may also influence genetic diversity. On the other hand, in *C. arvense*, a reduced growth of infested ramets can lead to an increased growth of axillary shoots (Peschken and Harris, 1975). This suggests a reallocation of resources between infested and non-infested shoots within clones. This evidence of changes in life history of

host plants after infestation calls for an analysis of the repercussions of the attack by *U. cardui* in the population genetic of *C. arvense*.

The aim of our work was twofold. Firstly, using AFLP markers we characterised the genotypes of infested and non-infested *C. arvense* shoots in local populations. Although the biology of the *U. cardui-C. arvense* system is well-known, it still remains elusive how *U. cardui* chooses suitable shoots for oviposition. We wanted to investigate whether *U. cardui* selects particular genotypes for oviposition. Secondly, we explored at a landscape scale the effects of the phytophage on the genetic diversity of its host. For that purpose, we compared the genotypic and genetic diversity of populations with different history of parasitism.

Material and methods

Sampling

Our study area in north-eastern Bavaria (11°50'E 49°35'N) was a 15 km² rural area composed of diverse array of habitats (meadows, agricultural fields, wastelands, old fallows, roadsides). Within this area, the biology and local dynamics of both *C. arvense* and *U. cardui* are well-known (Eber and Brandl, 1994; Eber and Brandl, 1996; Eber and Brandl, 1997; Eber and Brandl, 2003). For the present study, we focused on eight populations of *C. arvense*. All populations were in a similar successional phase, judged from other co-occurring plant species. Half of these eight populations were infested by *U. cardui* for at least five years, whereas there were no previous records of infestation for the other half. In September 1999 we recorded the number of *C. arvense* shoots. Fresh thistle leaves were sampled at all sites. In thistle populations with *U. cardui* galls we collected leaves from infested and non-infested thistle shoots separately. Details of the samples are presented in Table 1.

Table 1: Demographic, ecological, genotypic and genetic characteristics of eight populations of the clonal weed *C. arvense*. In brackets we give the number of shoots with galls of the tephritid fly *U.cardui*, which were sampled additionally in infested populations.

Population labels	Infestation group	Population size (number of shoots)	Number of plants analysed (N)	Number of genotypes detected (G)	Genotypic diversity (i)	Evenness index ($E_{1/D}$)	Molecular variance
8	infested	500	46 (7)	42	0.79	0.58	2.48
61	infested	3000	44 (10)	52	0.96	0.93	2.92
119	infested	400	18 (10)	13	0.46	0.29	1.07
155	infested	8000	10 (8)	13	0.72	0.5	0.76
60	non-infested	30	20	20	1.00	1	3.01
120	non-infested	40	12	6	0.50	0.33	1.06
148	non-infested	800	11	11	1.00	1	2.42
156	non-infested	100	8	5	0.63	0.5	2.39

Genotypic and genetic variability among infested and non-infested populations

AFLP were processed according to Solé et al., 2004. Genotypic diversity (i) was estimated according to Ellstrand and Roose (1987; i = G/N where G is the number of genotypes and N the number of sampled shoots). Clonal evenness, which estimates the relative abundance of each genotype within a population, was calculated according to Williams (1964; $E_{1/D} = \dfrac{1/D}{G}$,

where D is Simpson's index (D = $\sum\limits_{i=1}^{G} p_i^2$; p_i = the relative abundance of the ith genotype). We compared the average genotypic diversity (i) and the clonal evenness ($E_{1/D}$) of infested and non-infested populations with a Mann-Whitney U-test. To study the partitioning of genetic variance among populations we performed an analysis of molecular variance (AMOVA, Excoffier *et al.*, 1992) using the program Arlequin (Schneider *et al.*, 2000). As we were interested in comparing genetic variability of infested and non-infested populations, we constructed a hierarchical model where genotypes were nested within populations and populations were nested within either an infested or non-infested group. The program calculated the associated variance components (σ_a^2, σ_b^2, σ_c^2; see Table 3 for explanations), the significance of estimated parameters was then tested by a permutation procedure (Excoffier *et al.*, 1992). We used the number of pairwise differences as genetic distance. For each population we also calculated the molecular variance (sum of number of pairwise differences within a population divided by 2N (N-1), where N equals the number of samples). The molecular variance is a measure of genetic diversity within populations. We tested for a difference in the average molecular variance of infested and non-infested populations using a Mann-Whitney U-test.

Pattern of infestation within populations

For the four populations with *U. cardui* galls, we calculated a matrix of genetic dissimilarities based on the AFLP data for all possible pairwise comparisons of sampled shoots (1-Jaccard coefficient, Jaccard, 1908), and a matrix which codes for the infestation state. In the second matrix, a shared state (compared shoots either infested or not infested) was coded by 0 and individuals having a contrasting state were coded by 1. Finally, we used matrix correlation to test whether the genetic similarity of shoots with contrasting infestation states was lower than compared to pairs of shoots with identical infestations states (negative matrix correlation) using the NTSYS-pc-p package (Rohlf, 1993).

Results

Our AFLP-fingerprints allowed to analyse 93 polymorphic loci. With these 93 loci we distinguished 162 haplotypes among the 194 sampled *C. arvense* shoots (i.e. 83% distinguishable genotypes). Genotypes shared by several individuals (32 out of 194) always occurred in the same population. Within the 169 shoots not infested by *U cardui* we found 135 individual genotypes (85%), whereas within the 35 shoots infested by *U cardui* we found 27 individual genotypes (77%).

Average genotypic diversity, clonal evenness and molecular variance did not differ between infested and not-infested populations (Table 2).

Table 2: Average genotypic diversity, clonal evenness and molecular variance (± standard deviation) in total eight populations of *C. arvense* according to their infestation by the tephritid fly *U. cardui* (see Table 1 for raw data).

	Genotypic Diversity (i)	Clonal Evenness (E1/D)	Molecular Variance
Infested populations	0.73 (± 0.21)	0.58 (± 0.27)	1.81 (± 1.05)
Non-infested populations	0.78 (± 0.26)	0.71 (± 0.34)	2.22 (± 0.82)
Mann-Whitney U-test	U = 6; P > 0.56	U = 5.5; P > 0.46	U = 7; P > 0.77

Most of the genetic diversity was found among populations within groups (58 %; Table 3). Nevertheless, a large amount of diversity was still present within populations (37%). The variance due to grouping into infested and non-infested populations was not significant ($P > 0.11$) and explained less than five percent of the total variance.

Table 3: Results of a hierarchical analysis of molecular variance between infested versus non-infested populations, among populations within each infestation group and within populations. σ_a^2 was tested by random permutation of genotypes of whole populations across infested and non-infested populations. σ_b^2 was tested by random permutation of individuals across populations but within the same group. σ_c^2 was tested by permuting AFLP phenotypes randomly among populations and between groups. The significance tests are based on 1000 permutations

Variance component		Variance	% total	Significance
Among groups ("*infestation effect*")	σ_a^2	0.56	4.66	$P > 0.11$
Among populations within groups	σ_b^2	7.02	58.38	$P < 0.001$
Within populations	σ_c^2	4.44	36.97	$P < 0.001$

Within each population, we did not find any significant correlation between the genetic dissimilarity matrix of *C. arvense* shoots and the corresponding matrix coding for the similarity in the infestation state (Table 4). In the studied populations, shoots of *C. arvense* appear to be infested randomly by *U. cardui*.

Table 4: Results of matrix correlation for the four infested populations. For each population we correlated the matrix of genetic dissimilarity (1-Jaccard coefficient) with matrix coding for the similarity in the infestation state of individual shoots. Error probabilities are based on 1000 randomisations.

Populations	Matrix correlation	Probability
8	$r = -0.02$	$P > 0.39$
61	$r = -0.003$	$P > 0.48$
119	$r = 0.06$	$P > 0.19$
155	$r = -0.02$	$P > 0.38$

Discussion

U. cardui may affect population genetics of *C. arvense* along two pathways: (1) If infestation imposes selection pressure on morphological, phenological or physiological traits with a genetic background, some non-neutral genetic differentiation between infested and non-infested host plants should emerge within populations. However, the chance to detect such differences depends on the number of loci affected and on whether selection influences reproductive isolation. (2) If infestation is random with respect to the traits, neutral genetic effects are expected if infestation changes the demography of the population. An increased mortality or decreased fecundity decreases effective population size and hence influences the neutral genetic diversity.

If selection occurs, resistance or tolerance in host plants will evolve in order to reduce the costs of phytophages or pathogens (Crawley, 1983; Marquis, 1992). The system *C. arvense-U. cardui* should be no exception. Resistance or tolerance of certain genotypes are expected to create non-random infestation patterns in natural populations as well as genetic variation in host plant resistance. Significant genetic differentiation and variation associated with resistance have been demonstrated for several plant species (Mopper *et al.*, 1991; Vargas *et al.*, 2002; Strong *et al.*, 1993; Krabel and Petercord, 2000). However, in most of these cases it is not clear whether genetic differentiation of the host is a consequence of selection induced by infestation or whether differential infestation is a consequence of genetic differentiation due to other causes.

In our analyses, populations infested by *U. cardui* were not dominated by particular genotypes. In an UPGMA cluster analysis (results not shown), genotypes clustered according to the population, and we found no clusters of either infested or non-infested genotypes within populations. This supports the fact that we found no correlation between the similarity of the infestation state and the genetic similarity of shoots. Hence, in our study area, infestation of *C. arvense* by *U. cardui* appears to be random among and within populations referring to the genetic markers used.

Like mentioned above, if infestation is random, effects on genetic diversity of the host plant can only be expected if infestation changes the demography of the population. In species with mixed reproduction system (like in clonal plants), the clonal propagation of genotypes may decrease the genetic evenness of the population and thus promote the evolution of parasite specialization to particular host genotypes (Barrett, 1981). Hence, sexual reproduction

appears to be an efficient strategy to escape from parasitism. In the system *Urtica dioica-Cuscuta europaea* experiments which compared resource allocation into sexual and asexual reproduction showed that plants from infested populations had been selected for sexual reproduction (Koskela, 2002). Consequently, if one expects that parasitism selects for sexual reproduction, this should lead to an increase of genetic diversity within the infested populations. Although a decrease in the fitness of *C. arvense* by the attack of *U. cardui* has been demonstrated (Peschken and Harris, 1975), we found no effect of the phytophagous insect on the genotypic and genetic diversity of its host.

The failure to detect neutral genetic patterns in relation to infestation may be explained by at least two arguments. Firstly, the capacity of *C. arvense* to reallocate resources after infestation provides an "ecological buffer" for the host which reduces the direct effects of the phytophage on its host plant. Secondly, it may be dangerous to generalize the effects of *U. cardui* on *C. arvense* found in small-scale laboratory experiments (e.g. Peschken and Harris, 1975; Peschken *et al*., 1982) to the complex spatial and temporal dynamic of the *C. arvense-U.cardui* system on the landscape scale. The importance of spatial and temporal dynamics in host-parasite has already been intensively investigated for non-neutral genetic variability (Thrall and Burdon, 1997; Frank, 1997; Frank, 1996) Our previous studies of the temporal and spatial dynamics of local thistle populations suggested that *U. cardui* forms a very dynamic metapopulation system (Eber and Brandl, 1994; Eber and Brandl, 1996). Thereby, the spatial and temporal dynamics of *C. arvense* population is crucial for the spatial and temporal abundance of *U. cardui* (Eber and Brandl, 2003). Consequently, the dynamics of the system may preclude any consistent selection within populations. These results also comply with the fact that the introduction of *U. cardui* as a biocontrol agent for *C. arvense* failed in Canada (Peschken and Derby, 1992; Peschken *et al.*, 1997). *U. cardui* was not able to reduce the density of thistles at a landscape scale. Hence, selection pressure imposed by *U. cardui* on *C. arvense* appeared to be weaker on a population and/or metapopulation scale than it was supposed to be in experimental studies.

Moreover, other factors like competition may also be important for the impact of a phytophageous insect on the fitness of its host. In *C. arvense*, Peschken *et al*. (1982) showed that infested and non-infested thistle shoots cultivated without competitors showed no significant difference in vigour. Plants grown with competitors, however, showed significant differences. Hence, simple experiments with isolated plants may lead to a biased view about the impact of the phytophageous insect at the landscape level.

In conclusion, we suggest that in the *C. arvense-U. cardui* system the spatial scale of the host plant-phytophage interaction may be crucial to estimate the effects of the phytophage on host plants at a landscape scale. In contrast to experiments with isolated plants, infestation in natural populations of *C. arvense* occurs at the ramet level, whereas natural selection occurs at a genet level. Thus, the metapopulation dynamics may preclude any consistent selection within local populations. In part this may explain the persistence of the *C. arvense-U. cardui* system.

Literature cited

Barrett JA. 1981 The evolutionary consequences of monoculture. Pages 209-248 *in* Bishop JA and LM Cook eds. Genetic consequences of man-made change. Academic Press, London.

Crawley MJ 1983 Herbivory – the dynamics of animal-plant interactions. University of California Press, Berkely, USA.

Donald WW 1990 Management and control of Canada thistle (*Cirsium arvense*). Reviews of Weed Science 5:193-250.

Eber S, R Brandl 1994 Ecological and genetic spatial patterns of *Urophora cardui* (Diptera: Tephritidae) as evidence for population structure and biogeographical processes. Journal of Animal Ecology 63:187-199.

Eber S, R Brandl 1996 Metapopulation dynamics of the tephritid fly *Urophora cardui*: An evaluation of incidence-function model assumptions with field data. Journal of Animal Ecology 65:621-630.

Eber S, R Brandl 1997 Genetic differentiation of the tephritid fly *Urophora cardui* in Europe as evidence for its biogeographical history. Molecular Ecology 6:651-660.

Eber S, R Brandl 2003 Regional patch dynamics of the weed *Cirsium arvense*, and possible implications for plant-animal interactions. Journal of Vegetation Science 14:259-266.

Ellstrand NC, ML Roose 1987 Patterns of genotypic diversity in clonal plant species. American Journal of Botany 74:123-131.

Excoffier L, PE Smouse, JM Quattro 1992 Analysis of Molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. Genetics 131:479-491.

Frank SA 1996 Statistical properties of polymorphism in host-parasite genetics. Evolutionary Ecology 10:307-317.

Frank SA. 1997 Spatial Processes in Host-Parasite Genetics. Pages 325-352 *in* Hanski I and ME Gilpin eds. Metapopulation Biology Ecology,Genetics, and Evolution. Academic Press, New-York.

Frenzel M, S Eber, S Klotz, R Brandl 2000 Ecological comparisons across geographical distributions: the thisle gall fly *Urophora cardui* (Diptera:Tephritidae) on two different *Cirsium* hosts. European Journal of Entomology 97:183-189.

Gange AC, HE Nice 1997 Performance of the thistle gall fly, *Urophora cardui*, in relation to host plant nitrogen and mycorrhizal colonization. New Phytologist 137:335-343.

Heeb P, I Werner, AC Mateman, M Kölliker, MWG Brinkhof, CM Lessells, H Richner 1999 Ectoparasite infestation and sex-biased local recruitment of hosts. Nature 400:63-65.

Jaccard P 1908 Nouvelles recherches sur la distribution florale. Bulletin de la Société Vaudoise des Sciences Naturelles 44:223-270.

Koskela T 2002 Variation in life-history traits among *Urtica dioica* populations with different history in parasitism by the holoparasitic plant *Cuscuta europaea*. Evolutionary Ecology 16:433-454.

Krabel D, R Petercord 2000 Genetic diversity and bark physiology of the European beech (*Fagus sylvatica*): a coevolutionary relationship with the beech scale (*Cryptococcus fagisuga*). Tree Physiology 20:485-491.

Laing JE 1977 Establishment of *Urophora cardui* L. (Diptera: Tephritidae) on Canada Thistle in sourthern Ontario. Proceedings of the Entomological Society of Ontario 108:

Marquis RJ. 1992 The selective impact of herbivores. Pages 301-325 *in* Fritz RS and EL Simms eds. Plant resistance to herbivores and pathogens: ecology, evolution, and genetics. University of Chicago Press, Chicago, Illinois, USA.

Moore RJ 1975 The biology of Canadian weeds. 13. *Cirsium arvense* (L.) Scop. Canadian Journal of Plant Sciences 55:1033-1048.

Mopper S, JB Mitton, TG Whitman 1991 Genetic differentiation and heterozygocity in Pinyon pine associated with resistance to herbivory and environmental stress. Evolution 45:989-999.

Peschken DP, JL Derby 1992 Effect of *Urophora cardui* (L.) (Diptera: Tephritidae) and *Ceutorhynchus litura* (F.) (Coleptera: Curculionidae) on the weed Canada thistle, *Cirsium arvense* (L.) Scop. The canadian entomologist 124:145-150.

Peschken DP, JL Derby, DT Spann. 1997 Establishment of *Urophora cardui* (Diptera: Tephritidae) on canada thistle, *Cirsium arvense* (Asteraceae), and colony development in relation to habitat and parasitoids in Canada. Pages 53-66 *in* Dettner K, G Bauer, and W Völkl eds. Vertical food web interactions. Springer, Berlin.

Peschken DP, DB Finnamore, AK Watson 1982 Biocontrol of the weed Canada thistle: releases and development of the gall fly *Urophora cardui* (Diptera: Tephritidae) in Canada. The canadian entomologist 114:349-357.

Peschken DP, P Harris 1975 Host specificity and biology of *Urophora cardui* (Diptera: Tephritidae). A biocontrol agent for canada thistle (*Cirsium arvense*). The canadian entomologist 107:1101-1110.

Rohlf FJ 1993 NTSYS: numerical taxonomy and mutivariable analysis system. Exeter Ltd., Setauket, New York, NY.

Schneider S, D Roessli, L Excoffier 2000 Arlequin ver. 2.000: a software for population genetic data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Sheldon JC, FM Burrows 1973 The dispersal effectiveness of the ahene-pappus units of selected compositae in steady winds with convection. New Phytologist 72:665-675.

Shorthouse JD, AK Watson 1976 Plant galls and the biological control of weeds. Insect World Digest 3:8-11.

Solé M, W Durka, S Eber, R Brandl 2004 Genotypic and genetic diversity of the common weed *Cirsium arvense* (Asteraceae). International Journal of Plant Sciences *(in press)*

Strong DR, S Larsson, U Gullberg 1993 Heritability of host plant resistance to herbivory changes with gallmidge density during an outbreak on willow. Evolution 47:291-300.

Thomas F, JF Guegan, Y Michalakis, F Renaud 2000 Parasites and host life-history traits: implications for community ecology and species co-existence. International Journal of Parasitology 30:669-674.

Thrall PH, JJ Burdon 1997 Host-Pathogen Dynamics in a Metapopulation Context: The Ecological and Evolutionary Consequences of Being Spatial. Journal of Ecology 85:743-753.

Vargas CF, F López, H Sánchez, B Rodríguez 2002 Allozyme analysis of host selection by bark beetles in central Mexico. Canadian Journal of Forest Research 32:24-30.

Williams CB 1964 Patterns in the balance of nature. Academic Press, London.

Chapter III

Environmental study of reproductive and dispersal efforts of the common weed *Cirsium arvense* on a metapopulation level

With Yves Piquot, Walter Durka, Roland Brandl

Abstract: We examined the effect of succession on *Cirsium arvense* reproductive and dispersal outputs. *Cirsium arvense* is a dioecious, insect pollinated clonal plant. Female and male plants were collected in three populations of early successional stage and in ten populations of late successional stage. Among others, succession caused a number of important phenotypic changes in *C. arvense* including a significant decrease in the number of flowering shoots per population (-48%; $P < 0.01$), the number of flower heads per shoot ($P < 0.01$ in females; not significant in males) and the number of flowers per flower head ($P < 0.1$ in females and not significant in males). However, there was no relationship between successional stage and average mass of an achene or germination rate (F = 0.07; $P = NS$ and F = 0.14 $P > 0.71$ respectively). Based on recent theoretical models about the reproductive effort in metapopulation context (Ronce and Olivieri, 1997; Ronce et al., 2000b), we suggest that the differences found may result from evolutionary changes in the genotypic composition of populations because of increasing intraspecific selection pressures when the habitat is maturing. This idea is supported by previous AFLP analysis of *C. arvense* populations (Solé, et al., 2004). However, we found no changes in the dispersal potential (pappus length) of *C. arvense* and we discuss the suitability of this trait for estimating the dispersal potential in this species.

Key words: Reproductive effort, dispersal capacity, succession, mixed reproductive system, *Cirsium arvense*, metapopulation.

Introduction

Succession, which corresponds to all biotic and abiotic changes associated with the development of a juvenile community into a mature one, is often studied in ecology and more recently in population biology (Quinn et al., 1972; Beasleigh and Yarranton, 1974; Escarré et al., 1985; Escarré and Thompson, 1991; Escarré et al., 1994; Welham and Setter, 1998; Ross and Quinn, 1977; Peroni, 1994; Piquot et al., 1998). During succession, populations encounter different selection pressures: from density-independent selection pressures following colonisation of a new habitat to density-dependent selection when the habitat reaches its carrying capacity. Therefore, succession can lead to temporal and spatial variability in the expression of some plants life history characters, especially reproduction and dispersal, which can evolve rapidly when affected by changes in the selection regime (Fisher, 1930; Reznick et al., 1990; Li and Margolies, 1994; Cody and Overton, 1996).

Most perennial plants have a mixed reproductive system with sexual reproduction through seeds and clonal reproduction through different modes of vegetative propagation. These plants are usually called clonal plants. Limited availability of resources or a limited capacity for resource uptake may generate, at every age in clonal plants a trade-off between different life history components, if these compete for the same resources (Stearns, 1992). Therefore, if reproductive effort represents the proportion of resources that an individual allocates at a specific age into reproduction compared with the proportion allocated into maintenance (survival) or growth (Williams, 1966), a trade-off is expected between reproductive effort and clonal reproduction (van Noordwijk and de Jong, 1986; Hautekeete et al., 2001; Stearns 1992). If selection pressures change through time during succession, then we can expect population age-specific patterns of reproductive effort and clonal reproduction in each successionnal stage (Ronce and Olivieri, 1997; Piquot et al., 1998).

Likewise, different selection pressures may act on dispersal rate over time (Quinn et al., 1972; Olivieri et al., 1983; Peroni, 1994; Welham and Setter, 1998). For many clonal plants, dispersal by vegetative means is limited and seed dispersal is the only way to colonize a new habitat. Thus, genotypes with a high dispersal and reproductive output will more likely occur in early successional habitats. In late successional habitats however, density dependant selection should favour genotypes with high competitive ability i.e. plants that allocate more resources into vegetative reproduction and survival. Density dependant selection may also act against seed dispersal because of the immediate cost in local competitivity once the carrying

capacity is reached (Ronce et al., 2000a). However, the evolutionary relation between reproductive effort and dispersal is not trivial as demonstrated by recent theoretical models (Ronce et al., 2000b).

Studies looking at the evolution of these two traits in natural conditions are rare (Quinn et al., 1972). The aim of this study was to investigate reproductive and dispersal outputs in natural populations of the clonal dioecious plant *Cirsium arvense* in different successional states. We asked the following questions: 1) is there a relation between the reproductive output and succession? 2) is there a relation between dispersal and successionnal state? 3) do dispersal and reproduction coevolve?

In order to determine which possible factors might be responsible for changes in *C. arvense* reproductive effort and dispersal, we looked among others in the field at the plant height, sex ratio, fruit-set frequency distribution, seed weight and seed germination.

Materials and Methods

Plant biology

Canada thistle (*Cirsium arvense*, Asteraceae) is a major perennial weeds of Eurasia and Northern America. As most perennial plants, *C. arvense* has a mixed sexual-asexual reproduction system. Clonal reproduction occurs by means of very efficient lateral roots, which can spread up to 12m a year (Donald, 1994; Moore, 1975; Bostock and Benton, 1979) and lead to the production of large ramets (i.e. all aerial shoots coming from the clonal propagation of a single root). With regard to its flowering biology, *C. arvense* is often described as a subdioecious-dioecious species. Female plants are strictly female with only pistilate flowers (Moore, 1975). Males are morphologically hermaphrodites with vestigial ovaries, but functionally males. *C. arvense* is insect pollinated and produces a large number of plumed achenes with a long pappus easily dispersed by wind. New populations are mainly founded by achenes, as natural long distance dispersal of vegetative propagules is strongly limited or quasi-impossible (roots are deeply buried in the soil).

Early and late successional stages

Our study area is an experimental zone of 15 km² located in Southern Germany (11°50'E 49°35'N) that we regularly investigated over the last ten years (Eber and Brandl, 1994; Eber and Brandl, 1996; Eber and Brandl, 2003). It is a rural area composed of several types of habitats (meadows, agricultural fields, wastelands, old fallows, roadsides). In September 2000, we sampled three natural populations of *C. arvense* in early successional stage and 10 populations in late successional stage. For easiness, populations in early successional stage are called young populations whereas populations in late successional stages are called old populations. Young populations represent the first step after population establishment (i.e. colonisation phase). To locate young populations, we prospected and recorded *C. arvense* stands that had never been observed in previous studies. Young populations of *C. arvense* typically occurred on bare patches in disturbed grazing area, or in disturbed areas recently ploughed or mowed. Within these habitats, the vegetation cover was below 75 %. The plant communities were exclusively composed of herbaceous species in which *C. arvense* was the dominant species.

Old populations of *C. arvense* were sampled in old fallows. The vegetation cover was totally closed and woody species were present. In this habitat *C. arvense* shoots were particularly tall and ligneous. Old populations had already been mapped in this advanced successional stage in

1994. Although, the date of foundation of old populations is unknown, they must exist for at least ten years.

Populations characteristics

Field observations and plant collection were done in early September 2000. In this paper, we define the population size as the number of shoots within populations. In 2000, the population size, percentage of flowering shoots and the sex ratio of all 13 populations was estimated. For populations having less than 100 shoots, the number of flowering and non-flowering shoots, and the number of males and females were counted as exact as possible. Then, counts were rounded to the nearest ten to estimate the population size and percentages were calculated to express sex ratio (proportion of females) and proportion of flowering shoots (Table 1). For bigger populations, average number of plants was counted within randomly placed square meters. Counts were extrapolated to the population area covered by *C. arvense*. In these cases, the number of shoots was rounded to the nearest hundreds or thousands.

Table 1: Ecological characteristics of 13 natural populations of the perennial dioecious *C. arvense*

Population name	Successional stage	Population size (number of shoots)	Sex ratio (% of females)	Percentage of flowering shoots	Number of plants sampled (N)
E1	Early	300	100	86	30
E2	Early	250	100	94	30
E3	Early	900	75	95	50
L1	Late	8000	100	60	72
L2	Late	100	0	31	9
L3	Late	500	100	49	37
L4	Late	100	73.3	27	9
L5	Late	30	0	25	5
L6	Late	3000	98	60	51
L7	Late	400	100	50	29
L8	Late	40	53.3	47	7
L9	Late	800	69	88	36
L10	Late	10	*No data*	0	0

To avoid possible "time sampling artefacts", we recollected the same information in 2001 for three populations of each successionnal stage at the beginning, middle and end of the flowering time (mi-July; mi-August, mi-September). Because of the high disturbance in pioneer habitats, we lost two young populations due to mowing after the first survey. Thus, full data set 2000/2001 was available only for one of the young populations (E3), but for the

three old populations L1, L2, and L3 (Table 2). Additionally, in 2001 we measured directly in field the height of plants in these four populations (E3, L1, L2 and L3).

Reproductive and dispersal outputs

Populations differed markedly in their area, plant distribution and density, therefore sampling was adjusted in each population. Plants were collected on a ramet level. To avoid sampling the same genotype twice, we used adequate minimal distances between samples. Sampling effort was correlated to population size (r = 0.96; *P < 0.001*).

Plants were cut in the field and brought to the lab for measurements of reproductive and dispersal outputs. As we conducted a field survey, we could not estimate the reproductive effort according to Williams (1966). The reproductive effort was estimated independently from the investment into growth. Therefore in this paper, the reproductive effort represents the reproductive output we measured in the field for each plant.

For each plant collected (n = 365), we determinate its sex and counted its number of flower heads. Then we randomly selected four flower heads and counted the number of florets and fruits. In *C. arvense*, there are three categories of achenes: non-pollinated, aborted (i.e. ovaries that displayed pericard development, but did not contain a healthy embryo), or successfully pollinated and developed (Lalonde and Roitberg, 1989). We checked the quality of each achene and we used the third category to define the fruit set. In addition, we calculated the average mass of an achene by dividing the mass of all filled achenes of a flower head by the fruit set of the same flower head.

Additionally, we tested for germination capacity in one population of each successional stage (21 and 24 families for early and of late successional stage respectively). Achenes were sowed in Petri dishes with 3 mm Whatman filter paper and 2.5 ml of distilled water. All Petri dishes were kept in a culture chamber at: 26°C with a 12 hours photoperiod (Moore, 1975; Heimann and Cussans, 1996). We recorded germination percentages after four weeks.

Pappus length was used as an indicator of the dispersal capacity. In *C. arvense* the pappus derives from the corolla, and both males and females can develop a pappus after flowering. We limited our measures to females only. We measured the length of two random pappy per flower head.

Statistical analysis

All variables representing a proportion or a percentage (sex ratio, flowering, fruit set and germination) were analysed with generalised linear models with the program GLIM (Crawley, 1993). In this model, we fitted our data to a binomial distribution. The variables sex ratio (counts of females) and flowering (counts of flowering shoots) were fitted to the total number of plants observed in the populations. For the fruit set, we took into account only the flower heads that have produced at least one achene. The number of achenes produced by flower head was fitted to the number of florets. For the germination test, the number of seedling obtained per achene family was fitted to the number of achenes sown for this family.

We conducted a variance analyse for variables that measured the reproductive and dispersal outputs (number of flower heads per plant, number of florets per flower head, pappus length and average mass of an achene). Analyse was done with the General Linear Models module of the program SAS (SAS Institute, Cary, NC). In clonal plants, it has already been proved that resources acquired by a genet can be but are not necessary equitably allocated between all ramets (clonal integration, Marshall, 1990; Marshall and Price, 1997). Therefore, as we sampled plants on a ramet level, we strongly increased the variation inter-individuals. Hence, we suppressed the individual level from our statistical analysis. We analysed our data on a two-level nested ANOVA with population nested within successional stage. Because of strongly unbalanced design between male and female sample sizes, males and females were analysed separately. Finally, differences in plant height between young and old populations were tested with a t-test

Results

The sex ratio was biased in favour of females (Figure 1A) but did not significantly differ between the two successional stages (F = 0.00041, *P* > 0.98). Temporal surveys of sex ratio did not show any differences in the proportion of females during the flowering time (Figure 2).
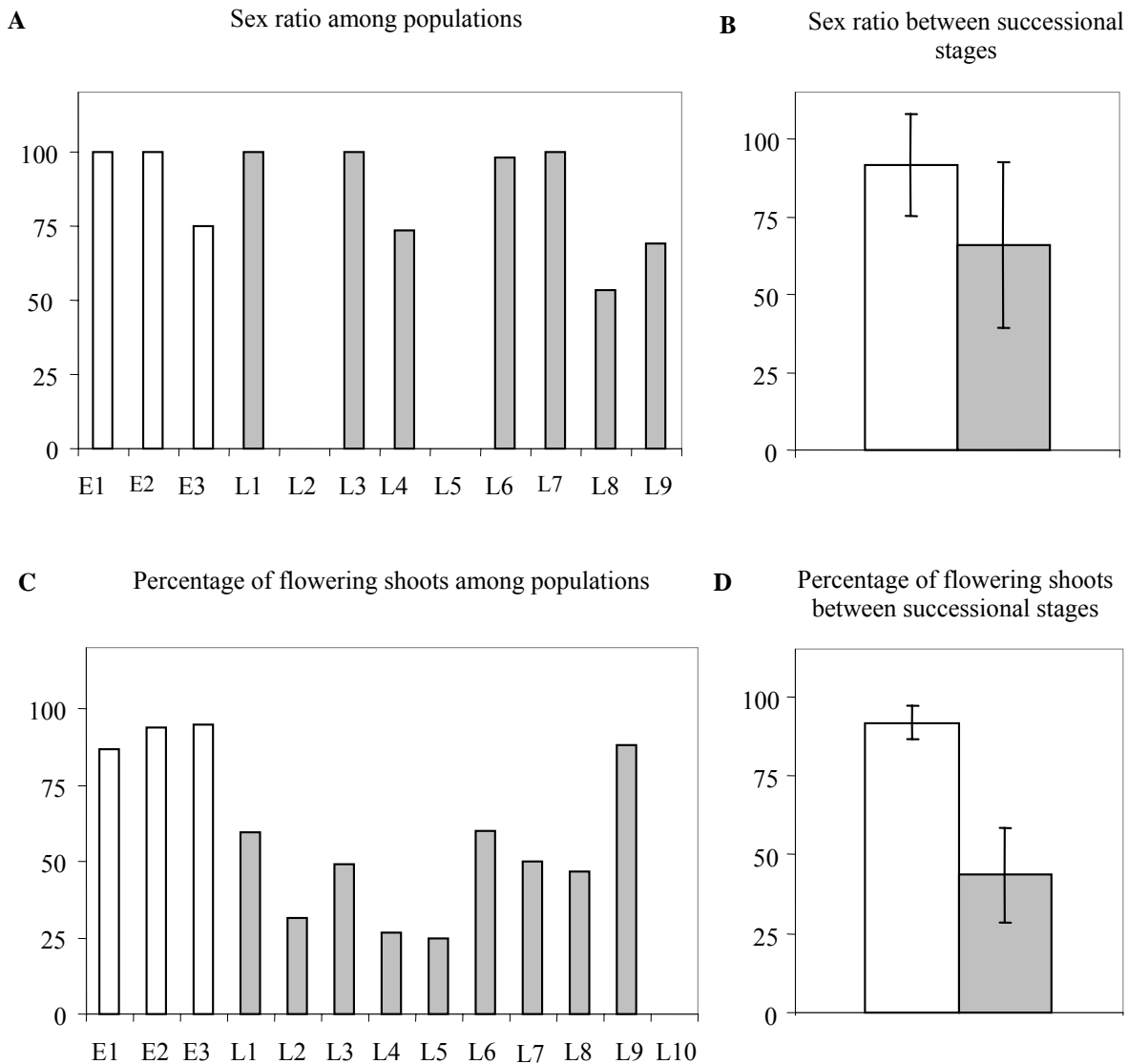

Figure 1: (**A**) Sex ratio (proportion of females) and percentage of flowering shoots (**C**) observed in the field among respectively 12 and 13 populations of *C. arvense*. Sex ratio (proportion of females) (**B**) and percentage of flowering (**D**) expressed in relation to the successional stage of populations. Open and shaded bars represent respectively young and old populations. Error bars represent the 95% confidence interval. Average sex ratio between young and old populations did not differ significantly (F = 0.00041; *P* > 0.98), whereas the two successional stages significantly differ in their importance of flowering (F = 10.60; *P* < 0.008).

The proportion of flowering shoots was significantly different between the two successional stages (F = 10.60, $P < 0.01$). Old populations had a lower (43.7%) proportion of flowering shoots than young populations (91.67 %, Figure 1B).
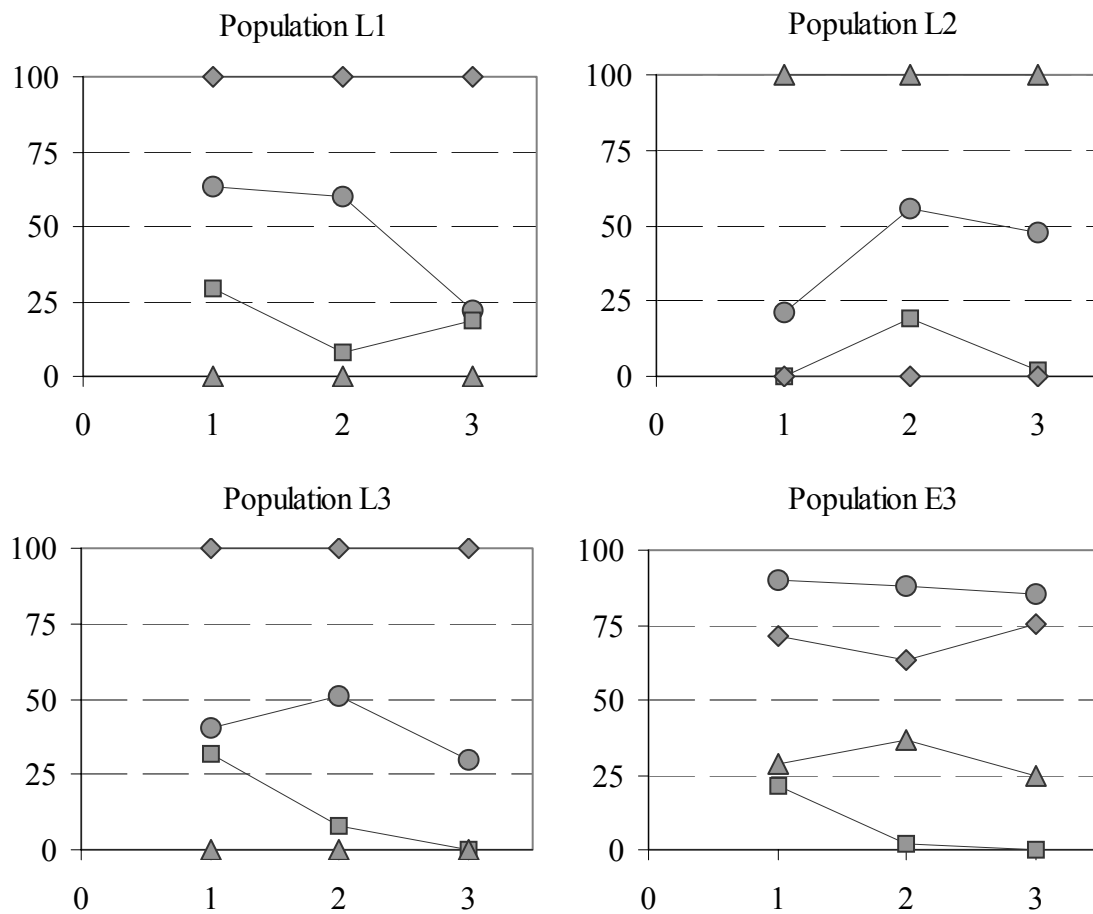

Figure 2: Proportion of males (▲), females (♦), buds (■) and flowering shoots (●) observed in four natural populations of *C. arvense* during the flowering time. On the abscise axis 1, 2, 3 represent respectively the beginning, middle and the end of flowering. The proportion of males and females were calculated as the percentage of males and females observed within the flowering shoots. The percentage of buds and flowering shoots were calculated over the total number of plants observed within populations.

Comparison of reproductive outputs between males and females did not reveal significant differences for the number of flower heads (Table 2), but showed significant differences for number of florets per flower head in old populations ($t_{(838)} = 3.8$ *P<0.01*).

Table 2: Descriptive statistics of variables measured in *C. arvense* females (♀) and males (♂) in populations of early and late successional stages (respectively three and nine populations). In the table (N) represents the sample size and (Ci) the 95% confidence interval.

		Early successional stage			Late successional stage		
		Mean	N	Ci	Mean	N	Ci
Plant size (cm)		93.5	32	7.4	123	49	5.6
Reproductive output	Flower head ♀	23.5	102	2.8	15.8	230	1.1
	Flower head ♂	31.8	18	13.0	12.6	25	6.2
	Floret ♀	77.4	293	3.3	65.1	755	1.2
	Floret ♂	80.4	65	14.4	70.9	84	12.5
	Fruit set	34.2	231	8.8	24.5	407	3.8
	Achene mass (mg)	1.2	214	0.06	1.2	362	0.03
	Germination (%)	32	24	7.6	34.2	21	8.2
Dispersal	Pappus ♀	22.9	734	0.7	21.9	1433	0.3

Males in old populations had more florets per flower head than females. The analyse of variance did not reveal successional stage effect for characters measured in male plants (Table 3). However, in females the same analyse showed significant differences between the two successional stages for the number of flower heads per plant and number of florets per flower head (Table 3).

Table 3: Effect of successional stage on the reproductive and dispersal out put in male and female of *C. arvense*. The table summarises the results of the nested analysis of variance showing the degree of freedom, the F-values and the level of significance (NS $P > 0.05$; *$P < 0.05$; **$P < 0.01$; ***$P < 0.0001$). The number of flower heads per plant (Flower head) and the number of flowers per flower head (Floret) were measured in both males and females. The pappus length (Pappus) and the average mass of a seed (mass of a seed) were measured in females only. The analysis of variance was performed separately for male and female plants. Populations were nested into successional stage.

		Females			Males		
		df	F	*P*	df	F	*P*
Flower head	Successional stage	1; 8	9.95	**	1; 2	3.43	NS
	Population	8; 322	3.12	***	2; 39	12	***
Floret	Successional stage	1; 8	5.26	*	1; 2	0.38	NS
	Population	8; 1038	13.78	***	2; 145	22.02	***
Pappus	Successional stage	1; 8	1.97	NS	-	-	-
	Population	8; 2157	38.82	***	-	-	-
mass of a seed	Successional stage	1; 8	0.07	NS	-	-	-
	Population	8; 566	12.42	***	-	-	-

Plants in young populations produced more flower heads per plant and more florets per flower head than plants in old populations (Figure 3).
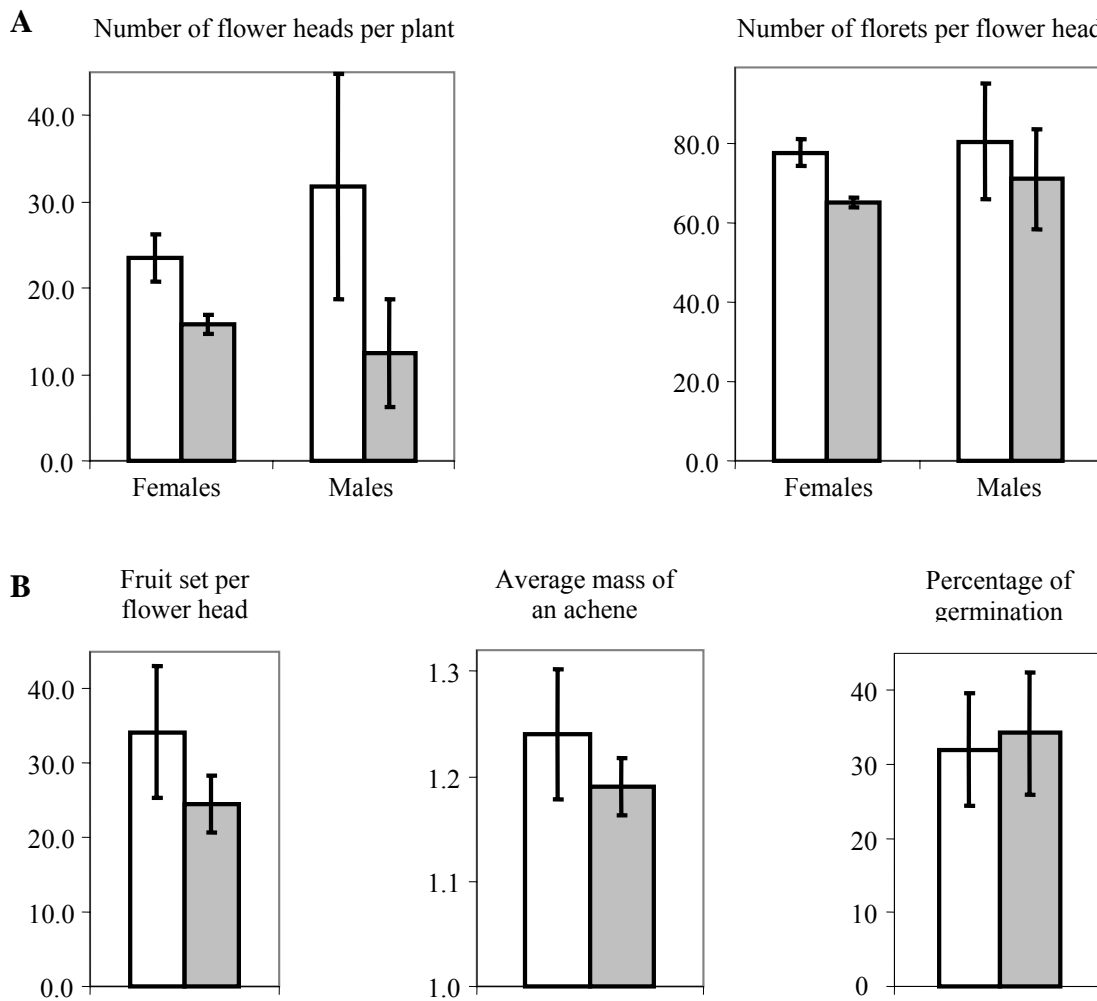

Figure 3: Reproductive output measured in the field in 12 populations of *C. arvense*. Open and shaded bars represent respectively young and old populations. Error bars represent the 95% confidence interval. (**A**) Females in young populations have a significantly greater number of flower heads per plant and number of florets per flower head than females in old populations. The same characters do not present significant differences between the two successional stages in males. (**B**) All characters measured in achenes do not show any significant differences between young and old populations.

The fruit set and the average mass of an achene did not significantly differ between young and old populations (respectively $F = 1.30$; $P > 0.28$ and $F = 0.07$; $P > 0.79$; Figure 3). The percentage of germination did not significantly differ among achenes families coming from young and old populations ($F = 0.14$; $P > 0.71$; Figure 3). However, we tested a single population for each successional stage; therefore, we could not directly test for a successional effect.

Fruit set class distribution revealed similar patterns of relative frequency between the two successional stages (Figure 4).
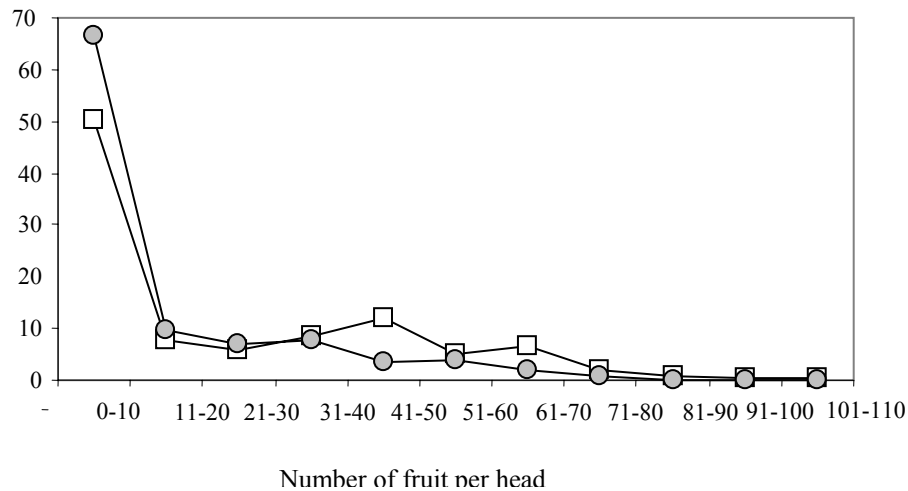

Figure 4: Frequency distribution of the average fruit set per flower head measured in young and old populations of *C. arvense*. The plot with squared symbols represents young populations, whereas the plot with round symbols represents old populations.

For the dispersal capacities, the pappus length did not differ significantly between the two successional stages (F = 1.97; *P* > 0.05; Figure 5). Finally, the average height of plants was significantly different between the two successional stages ($t_{(79)}$ = -6.5; *P* < 0.001). Plants in old populations were about 25% taller than plants in young populations (Figure 5).
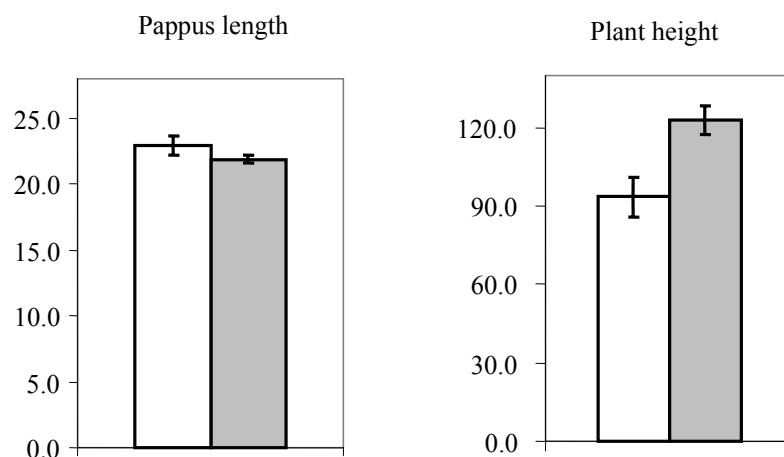

Figure 5: Dispersal capacity (pappus length) and plant height measured in the field in young and old populations of *C. arvense*. Open and shaded bars represent young and old populations respectively. Error bars correspond to the 95% confidence interval. Pappus length and plant height are expressed in mm and cm respectively.

Discussion

The overall sex ratio in natural populations of *C. arvense* was heavily biased towards females (72 %) irrespectively of the population successionnal status. The sex ratio did not change over time during the flowering season in the four populations surveyed. Similar female to male ratios of 3 to 1 have already been reported for *C. arvense* in Lake Yssel in the Netherlands by Bakker (1960) and in England by Kay (1985). However, two old populations consisted almost entirely of males (L2 & L5) and a single small old population (L8) had an equilibrate 1:1 ratio.

The inheritance of sex in *C. arvense* is poorly known. The only investigation of this issue was done by Correns (1916), who proposed that males were heterogametic and the female homogametic. However, as the crosses were interspecific (*Cirsium arvense* ♀ + *Cirsium oleraceum* ♂), the results are hardly interpretable. Moreover, other studies that looked at the sex ratio of progeny clearly showed strong females biases (around 1 male for 80 females, Lalonde and Roitberg, 1994). These results strongly suggest a complicated inheritance of sex in *C arvense*, with possible nuclear-cytoplasmic factors involved in the turn of offspring into females (like in *Thymus vulgaris*; Couvet et al., 1986).

The principal effect of a female biased sex ratio is the poor availability of pollen donors, which might lead to important changes in seeds quantity and quality. For example, Lalonde and Roitberg (1994) demonstrated that a reduction of pollination success increases the average achene mass. In our study, there were no significant differences in the average mass of an achene between young and old populations. Moreover, populations of both stages present a similar pattern in their frequency distribution of the number of fruits produced per flower head. Although seed abortion was high (more than 50% of flower heads did not produce a seed), it appears invariant with respect to the successional stage. These results, adding to the fact that seeds were produced even in populations composed by 100 percent of female plants, suggest that young and old populations experience similar pollination conditions. This can be explained by the high mobility of bumble bees (Walther-Hellwig and Frankl, 2000), which are the main pollinators of *C. arvense*.

Young and old populations of *C. arvense* differ in their reproductive output. Young populations have a higher percentage of flowering shoots. Plants in young populations produce more flower heads per shoot and more florets per flower head. These results were found only for female plants; however, we observed similar trends in males. The non-

significance of the variance analysis in males might result from the low size of the samples. Because of the highly female-biased sex ratio, male plants were difficult to find and to sample in natural populations.

Plant size, average mass of an achene and germination did not vary significantly with population age. This discards the hypothesis that the decrease of reproductive output is due to environmental factors (like a reduction of resources in old populations) or maternal effects. Moreover, in clonal species, a decrease of reproductive output occurring with an increase age of populations might be due to the accumulation of somatic mutations, which affect sexual reproduction (Barrett, 1980; Carson et al., 1982; Klekowski, 1988; Klekowski and Godfrey, 1989; Eckert et al., 1999). As the average mass of an achene and the germination rate are not affected by population age, reduction of reproductive output should result from a selective process.

In a metapopulation context, local populations of different successional stage may experience opposing selection pressures. In early successional habitats, where seedling recruitment is frequent relative to late successional habitats, plants that maximise their resource allocation to sexual reproduction and dispersal will be selected. In late successional habitats, plants that maximise allocation to vegetative growth and survival will be selected by density dependant selective pressures. Because achenes are the primary means of dispersal in *C. arvense*, sexual selection will be selected at the metapopulation level for the foundation of new populations. Consequently, new populations are generally composed of genotypes with a high output for sexual reproduction. However, when the population reaches the habitat's carrying capacity, density dependant selection pressures will be maximal and genotypes that the best at clonality will be selected (Harper, 1977; Abrahamson, 1980; Gray, 1987; Eriksson, 1993). As clonal extension is favoured within populations, while sexual reproduction is counter-selected, a decrease of reproductive output over time is expected and should result from evolutionary changes in the genetic composition of populations. This is supported by several theoretical models (Ronce and Olivieri, 1997). Ronce and Olivieri assume that differences between populations are genetic (i.e. distinctive genotypes with different selective values for reproductive output have been selected in different local populations) and predict a decrease of the average reproductive effort when populations get older (Ronce and Olivieri, 1997). The model developed by Ronce and Olivieri was developed for species that have overlapping generations but no senescence and could be applied to *C. arvense*.

Studies that followed the reproductive effort through time (Beasleigh and Yarranton, 1974; Antos and Allen, 1999) or compared populations in sites which differed in successional stage or disturbance (Gadgil and Solbrig, 1972; Quinn et al., 1972; Hickman, 1975; Holler and Abrahamson, 1977; Ross and Quinn, 1977; Grace and Wetzel, 1981; Escarré et al., 1985; Escarré and Thompson, 1991; Piquot et al., 1998) generally found a decrease of reproductive effort with an increasing maturity of the habitat.

However, variation in reproductive output maybe genotypic or may result from phenotypic plasticity. Not all studies cited above attempted to separate environmental and genetic components. In some cases, differences between habitats were found to be genetic (e.g. Gadgil and Solbrig, 1972; Escarré et al., 1985; Escarré et al., 1994; Grace and Wetzel, 1981). In others, they were attributed to an environmental plastic response of plants (Beasleigh and Yarranton, 1974; Hickman, 1975; Holler and Abrahamson, 1977; Ross and Quinn, 1977; Antos and Allen, 1999), or both (Reinartz, 1984).

Here, we recorded characters in natural populations of *C. arvense*. The hypothesis that changes in reproductive output were due to phenotypic plasticity can not be excluded. However, in a molecular study of the same area we found an extremely high population differentiation (Solé et al., 2004) independent of geographical distance (even within each successional stage). This high population differentiation was partly due to founder effects, but the occurrence of a strong selection during the early successional stages can not be rejected.

Similarly to reproductive output, two opposing selection pressures act on dispersal when a habitat is maturing, selection for dispersal during recolonisation, and selection against dispersal once a population is founded. Therefore, changes in dispersal capacity over time might also be explained by evolutionary changes in the composition of populations in a metapopulation context (Olivieri et al., 1995). Assuming local extinctions to be the sole source of environmental variation, and dispersal capacities to have a genetic basis, the model of Olivieri et al. (1995) predicts a decrease of dispersal with an increase of habitat maturity. Environmental studies that looked at dispersal capacities during succession are rare. Some studies found genetic changes in populations composition (Olivieri et al., 1983; Peroni, 1994; Cody and Overton, 1996), whereas in others the variation in dispersal capacities was clearly due to phenotypic plasticity (Imbert and Ronce, 2001) or maternal effects (Donohue and Schmitt, 1998; Acosta et al., 1997).

Although the pappus is one of the most obvious characteristics of *C. arvense* achenes, we did not detect differences in the pappus length of *C. arvense* between young and old populations. This result calls for the suitability of this trait for estimating the dispersal potential in plants. Many methods (including falling velocity experiments, diaspore flight angle, releasing height…) have been developed to determine dispersal potential in anemochorous species (Sheldon and Burrows, 1973; Greene and Johnson, 1990; Hensen and Müller, 1997), but there is no consensus about which estimator is the best. Peroni (1994) noted that differences in dispersal potential appeared smaller when dispersal capacity was measured by a continuous variable (wing loading ratio in her case) than when dispersal capacity was measured by discontinuous variable (presence / absence of a pappus).

Furthermore, the pertinence of the length pappus for estimating the dispersal potential in *C. arvense* may be arguable, since the pappus often breaks off the achene (Bakker, 1960; Bostock and Benton, 1979). Thus, an estimation of the percentage of achenes released might have been a better parameter than the length of the pappus to assess dispersal potential in *C. arvense*. Recently, new models about wind dispersal and assessment of dispersal potential, which include topography, weather conditions, turbulence and thermal up-drafts, have been developed (Nathan et al., 2002; Tackenberg et al., 2003; Tackenberg, 2003). Tackenberg (2003) showed that wind dispersal potential was more sensitive to weather conditions than falling velocity. These new results should be taken into account in the future when choosing a variable to estimate dispersal potential of plant species

Literature cited

Abrahamson WG. 1980 Demography and vegetative reproduction. Pages 89-106 *in* Solbrig OT ed. Demography and evolution of plant populations. Blackwell, Oxford.

Acosta FJ, JA Delgado, F López, JM Serrano 1997 Functional features and ontogenetic changes in reproductive allocation and partitionong strategies of plant modules. Plant Ecology 132:71-76.

Antos JA, GA Allen 1999 Patterns of reproductive effort in male and female shrubs of *Oemleria cerasiformis*: a 6-year study. Journal of Ecology 87:77-84.

Bakker D. 1960 A comparative life-history study of *Cirsium arvense* (l.) Scop., *Tussilago farfara* L, the most troublesome weeds in the newly reclaimed polders of the former Zuiderzee. Pages 205-222 *in* Harper JL ed. The biology of weeds. Blackwell Scientific Publications, Oxford.

Barrett SCH 1980 Sexual reproduction in *Eichhornia crassipes* (water yacinth). I. Fertility of clones from diverse regions. Journal of Applied Ecology 17:101-112.

Beasleigh WJ, GA Yarranton 1974 Ecological strategy and tactics of *Equisetum sylvaticum* during a postfire succession. Canadian Journal of Botany 52:2299-2318.

Bostock SJ, RA Benton 1979 The reproductive strategies of five perennial Compositae. Journal of Ecology 67:91-107.

Carson HL, LS Chang, TW Lyttle 1982 Decay of female sexual behavior under parthenogenesis. Science 218:68-70.

Cody ML, JM Overton 1996 Short-term evolution of reduced dispersal in island plant populations. Journal of Ecology 84:53-61.

Correns C 1916 Untersuchungen über Geschlechtsbestimmung bei Distelarten. Sitzungberichte der koeniglich preussischen Akademie der Wissenschaften 20:448-477.

Couvet D, F Bonnemaison, PH Gouyon 1986 The maintenance of females among hermaphrodites: the importance of nuclear-cytoplasmic interactions. Heredity 57:325-330.

Crawley MJ 1993 GLIM for ecologists. Blackwell Scientific Publications, Oxford, UK.

Donald WW 1994 The biology of Canada Thistle (*Cirsium arvense*). Reviews of Weed Science 6:77-101.

Donohue K, J Schmitt. 1998 Maternal environmental effects in plants : Adaptative plasticity? Pages 137-158 *in* Mousseau TA and JA Fox eds. Oxford University Press, New York.

Eber S, R Brandl 1994 Ecological and genetic spatial patterns of *Urophora cardui* (Diptera: Tephritidae) as evidence for population structure and biogeographical process. Journal of Animal Ecology 63:187-199.

Eber S, R Brandl 1996 Metapopulation dynamics of the tephritid fly *Urophora cardui*: An evaluation of incidence-function model assumptions with field data. Journal of Animal Ecology 65:621-630.

Eber S, R Brandl 2003 Regional patch dynamics of the weed *Cirsium arvense*, and possible implications for plant-animal interactions. Journal of Vegetation Science 14:259-266.

Eckert CG, ME Dorken, SA Mitchell 1999 Loss of sex in clonal populations of a flowering plant, *Decodon verticillatus* (*Lythraceae*). Evolution 53:1079-1092.

Eriksson O 1993 Dynamics of genets in clonal plants. TREE 8:313-316.

Escarré J, C Houssard, JP Briane. 1985 Genotypic and phenotypic variation within *Rumex acetosella* L. populations along a secondary succession. Pages 223-239 *in* Jacquard P, G Heim, and J Antonovics eds. Genetic differentiation and dispersal in plants. Springer, Berlin.

Escarré J, C Houssard, JD Thompson 1994 An experimental study of the role of seedling density and neighbor relatedness in the persistence of *Rumex acetosella* in an old-field succession. Canadian Journal of Botany 72:1273-1281.

Escarré J, JD Thompson 1991 The effects of successional habitat variation and time of flowering on seed production in *Rumex acetosella*. Journal of Ecology 79:1099-1112.

Fisher RA 1930 The genetical theory of natural selection. Clarendon Press, Oxford.

Gadgil M, O Solbrig 1972 The concept of r- and k-selection: evidence from wild flowers and some theoretical considerations. American Naturalist 106:14-31.

Grace JB, RG Wetzel 1981 Phenotypic and genotypic components of growth and reproduction in Typha latifolia: experimental studies in marshes of differeing successional maturity. Ecology 62:789-801.

Gray AJ. 1987 Genetic change during succession in plants. Pages 273-293 *in* Gray AJ, DE McCauley, and PJ Edwards eds. Colonization, succession and stability. Blackwell Scientific Publications, Oxford.

Greene DF, EA Johnson 1990 The aerodynamics of plumed seeds. Functional Ecology 4:117-125.

Harper JL 1977 Population biology of plants. Academic Press, London.

Hautekeete NC, Y Piquot, H Van Dijk 2001 Investment in survival and reproduction along a semelparity-iteroparity gradient in the Beta species complex. Journal of Evolutionary Biology 14:795-804.

Heimann B, GW Cussans 1996 The importance of seeds and sexual reproduction of in the population biology of Cirsium arvense - a literature review. Weed Research 36:493-503.

Hensen I, C Müller 1997 Experimental and structural investigations of anemochorous dispersal. Plant Ecology 133:169-180.

Hickman JC 1975 Environmental unpredictability and plastic energy allocation strategies in the annual *Polygonum cascadense* (Polygonaceae). Journal of Ecology 63:689-701.

Holler LC, WG Abrahamson 1977 Seed and vegetative reproduction in relation to density in *Fragaria virginia* (Rosaceae). American Journal of Botany 64:1003-1007.

Imbert E, O Ronce 2001 Phenotypic plasticity for dispersal ability in the seed heteromorphic *Crepis sancta* (Asteraceae). Oikos 93:126-134.

Kay QON 1985 Hermaphrodites and Subhermaphrodites in a reputedly dioeciuos plant, *Cirsium arvense* (L.) scop. New Phytologist 100:457-472.

Klekowski EJ 1988 Progressive cross- and self-sterility associated with aging in fern clones and perhaps others plants. Heredity 61:247-253.

Klekowski EJ, PJ Godfrey 1989 Aging and mutation in plants. Nature 340:389-391.

Lalonde RG, BD Roitberg 1989 Ressource limitation and offspring size and number trade-offs in *Cirsium arvense* (Asteraceae). American Journal of Botany 76:1107-1113.

Lalonde RG, BD Roitberg 1994 Mating system, life-history, and reproduction in Canada thistle (*Cirsium arvense*; Asteraceae). American Journal of Botany 81:21-28.

Li J, DC Margolies 1994 Responses to direct and indirect selection on aerial dispersal behaviour and life-history traits in T*etranychus urticae*. Heredity 72:10-22.

Marshall C. 1990 Source- sink relations of interconnected ramets. *in* Van Groenendael J and H De Kroon eds. Clonal growth in plants: regulation and function. SPB Academic Publishing, The Hague.

Marshall C, EAC Price. 1997 Sectoriality and its implications for physiological integration. Pages 79-107 *in* De Kroon H and J Van Groenendael eds. The ecology and evolution of clonal plants. Blachuys Plublishers, Leiden, The Netherlands.

Moore RJ 1975 The biology of Canadian weeds. 13. *Cirsium arvense* (L.) Scop. Canadian Journal of Plant Sciences 55:1033-1048.

Nathan R, GG Katul, HS Horn, SM Thomas, R Oren, R Avissar, SW Pacala, SA Levin 2002 Mechanisms of long-distance dispersal of seeds by wind. Nature 418:409-413.

Olivieri I, Y Michalakis, PH Gouyon 1995 Metapopulation genetics and evolution of dispersal. The American Naturalist 146:202-228.

Olivieri I, M Swan, P-H Gouyon 1983 Reproductive system and colonizing strategy of two species of *Carduus* (compositae). Oecologia 60:114-117.

Peroni PA 1994 Seed size and dispersal potential of *Acer rubrum* (Aceraceae) samaras produced by populations in early and late successional environments. American Journal of Botany 81:1428-1434.

Piquot Y, D Petit, M Valero, J Cuguen, P de Laguerie, P Vernet 1998 Variation in sexual and asexual reproduction among young and old populations of the perenial macrophyte *Sparganium erectum*. Oikos 82:139-148.

Quinn JA, J Rotsettis, DE Fairbrothers 1972 Inflorescence characters and reproductive proficiency in *Danthonia sericea* populations. American Journal of Botany 59:627-631.

Reinartz JA 1984 Life history variation of common mullein (*Verbascum thapsus*): III. differences among sequential cohorts. Journal of Ecology 72:927-936.

Reznick DN, H Bryga, JA Endler 1990 Experimentally induced life-history evolution in natural population. Nature 346:357-359.

Ronce O, I Olivieri 1997 Evolution of reproductive effort in a metapopulation with local extinctions and ecological succession. American Naturalist 150:220-221.

Ronce O, F Perret, I Olivieri 2000a Evolutionarily stable dispersal rates do not always increase with local extinction rates. American Naturalist 155:485-496.

Ronce O, F Perret, I Olivieri 2000b Landscape dynamics and evolution of colonizer syndromes: interactions between reproductive effort and dispersal in a metapopulation. Evolutionary Ecology 14:233-260.

Ross FH, JA Quinn 1977 Phenology and reproductive allocation in *Andropogon scoparius* (Graminaceae) populations in communities of different successional stages. American Journal of Botany 64:535-540.

Sheldon JC, FM Burrows 1973 The dispersal effectiveness of the achene-pappus units of selected compositae in steady winds with convection. New Phytologist 72:665-675.

Solé M, W Durka, S Eber, R Brandl 2004 Genotypic and genetic diversity of the common weed *Cirsium arvense* (Asteraceae). International Journal of Plant Sciences *(in press)*

Stearns SC 1992 The evolution of life histories. Oxford University Press, Oxford.

Tackenberg O 2003 Modeling long distance dispersal of plant diaspores by wind. Ecological Monographs 73:

Tackenberg O, P Poschlod, S Bonn 2003 Assessment of wind dispersal potential in plant species. Ecological Monographs 73:

van Noordwijk AJ, G de Jong 1986 Acquisition and allocation of resources: their influence on variation in life history tactics. American Naturalist 128:137-142.

Walther-Hellwig K, R Frankl 2000 Foraging habitats and foraging distances of bumblebees, *Bombus* spp. (Hym., Apidae), in an agricultural landscape. Journal of Applied Entomology 124:299-306.

Welham CVJ, RA Setter 1998 Comparison of size-dependent reproductive effort in two dandelion (*Taraxacum officinale*) populations. Canadian Journal of Botany 76:166-173.

Williams CC 1966 Natural selection, the cost of reproduction and a refinement of Lack's principle. American Naturalist 100:687-690.

Discussion and outline

This chapter discusses the results of this work on two levels - generally together with the ecology of clonal plants, and more specifically with the biology of *Cirsium arvense*.

Part I

Genotypic diversity has long been controversial issue in clonal plants. Whereas sexual reproduction allows recombination and production of new allele combinations, clonality propagates the same genotype (mutations excepted). This difference between the two means of reproduction has led people to hypothesise that clonality should lower the genetic variation in clonal plants in comparison to non-clonal plants. (Abrahamson and Gadgil, 1973; Williams, 1975; Harper, 1977; Abrahamson, 1980).

In *C. arvense*, the genotypic diversity was high (proportion of distinguishable genotypes equalled 0.73 ± 0.25 (mean $\pm$ SD)). This result is in accord with the literature accumulated on this topic during the last three decades, which has led to reviews that demonstrate that clonal plants are as variable to non-clonal plants (Ellstrand and Roose, 1987; Hamrick and Godt, 1990; Widen et al., 1994). With regard to this surprising result, genotypic diversity in clonal plants became controversial and research strongly focused on factors responsible for this considerable level of variation (Silander, 1985; Klekowski and Godfrey, 1989; De Kroon and Van Groenendael, 1997; Klekowski, 1997).

In clonal plants, which have mostly a mixed reproduction system, occasional establishment of seedlings (if not inbred) is often proposed as a powerful mechanism generating genotypic variation (Soane and Watkinson, 1979; Marshall and Weir, 1979; Watkinson and Powell, 1993; Bengtsson, 2003).

In *C. arvense* seedling recruitment is not excluded in the first years following the establishment of populations, but must be improbable once populations have reached the carrying capacity (Bakker, 1960). Moreover, Eriksson (1989) reported no evidence for seedling recruitment in established populations in 60% of the 68 clonal species that he reviewed; a statement that moderates the power of seedling recruitment to explain the high genotypic variation in clonal plants. According to Bengtsson (2003), the fact that populations with limited or missing seedling recruitment nevertheless display a high genotypic variation, can be explained by the population effective "memory" of their earlier genetic history. In his

theoretical work, Bengtsson looked at the "Genotypic Identity" of populations (i.e. the probability that two randomly sampled adult individuals from a population have the same genotype) depending on population growth parameters, rate of sexuality and recruitment of sexually derived offspring. From his simulations, he concluded that: "a population which was started by a number of sexually derived propagules may thus retain its initial genotypic variation for a very long period of time, even if it later reproduces almost exclusively asexually".

Coalescence theory (Kingman, 1982) can be used to illustrate the influence of clonality on population genetic variation. The coalescence theory aims to reconstruct the genealogy of a sample of observed gene copies, according to a specific mutation model of the molecular marker used to estimate the genotypic diversity, and according to the demographical history of populations. The coalescence time of a sample of observed gene copies is the time to the most recent common ancestor of these observed gene copies (expressed in number of generations). Whereas the normal coalescence time is 2N (for a large and constant population size N), Bengtsson (2003) obtained greater coalescence time in his model developed for partial asexuality (i.e. 4N for gene copies taken from the same individual, and 3N for gene copies taken from different individuals). This result means that clonality enlarges the coalescence time and thus allows genes copies to persist longer within populations.

From all these theoretical and experimental data, the general hypothesis that clonality should lower the genetic variation in clonal plants in comparison to non-clonal plants appears outdated. Predominantly asexual populations can display any pattern of genotypic and genetic variation, and can be as variable as non-clonal populations.

As already emphasized by McLellan et al. (1997), more relevant than the amount of variation is the distribution of the genotypic and genetic pattern within and among populations and the evolutionary processes responsible for it. Several reviews already attempted to determine which ecological variables are good predictors of genotypic and genetic structures (Hamrick et al., 1979; Loveless and Hamrick, 1984; Baur and Schmid, 1996; Hamrick and Godt, 1996; Bachmann, 2001). Nevertheless, conclusions drawn from these reviews are quite general and the distribution of the genotypic and genetic pattern within and among populations must still be analysed case by case. In *C. arvense* genotypic diversity and population differentiation were extremely high (proportion of distinguishable genotypes = 0.73 ± 0.25 and F_{ST} = 0.63) and most similar genotypic and genetic patterns were found in the literature in species that

have similar ecology, e.g. *Rhus* species that are widespread, clonal, long-lived and dioecious (Sherman-Broyles et al., 1992).

Part II

Natural populations of *C. arvense* were mostly multiclonal and almost all patterns of clonal diversity were combined with clonal evenness. Thus, genotypic diversity must be studied case by case in *C. arvense*.

However, the absence of isolation by distance was surprising and might bring some new information about seed dispersal in this species. In *C. arvense* the fact that the pappus often breaks off the achene has lead people to hypothesize that *C. arvense* possesses a low or limited seed dispersal (Bakker, 1960; Bostock and Benton, 1979). Under this condition, in a metapopulation context where new populations are mostly founded by seeds coming from older ones, one could have expected to observe an increasing genetic divergence of populations with an increasing geographical distance. However, this pattern was not observed with nucleic molecular markers and suggests that genetic divergence among populations might occur, with these markers, at a bigger scale than the 4 km² of the study area. This, this absence of isolation by distance could defend the hypothesis of a long distance dispersal occurring in *C. arvense*, which is not in accord with the few available data about field experiment seed traps (maximal dispersal distance measured = 139m in Mayer, 2000; less than 0.2% of achenes bore a seed 1 Km from the mother plant, references in Heimann and Cussans, 1996). However, recent models about wind dispersal based on landscape topography, weather conditions, turbulence and thermal up-drafts, have been developed (Nathan et al., 2002 ; Tackenberg et al., 2003; Tackenberg, 2003). Tackenberg (2003) showed that distance dispersal is significantly increased relative to traditional models using terminal velocity. Still, these models need calibrations (e.g. landscapes parameters, falling velocity, seed release rate) and the measurement of real maximum dispersal distance for comparison is impossible. A solution to study seed dispersal in natural populations would be to investigate genetic differentiation and isolation by distance with maternally inherited molecular markers e.g. mitochondrial DNA or chloroplastic DNA more specifically in plants (McCauley, 1995; Ouborg et al., 1999). Indeed, the comparison of patterns of differentiation obtained with nuclear and cytoplasmic markers can inform about the ratio of pollen versus seed gene flow (Oddou-Muratorio et al., 2001; Lian et al., 2003) and might help to determine the seed

dispersal distance. For example, it has already been demonstrated that the ratio of pollen to seed ratio changes with the spatial scale (McCauley 1997 in Ouborg et al., 1999). In *Silene alba*, the author found an increase of pollen dispersal, in comparison to seed dispersal, with an increase of the study area (pollen to seed ratio equalled 3.4, 9.2, and 124.0 at large, intermediate and small scales respectively). In *C. arvense*, the absence of isolation by distance observed with nuclear inherited molecular markers might be explained by the long distance dispersal of honey bees (up to kilometres; Walther-Hellwig and Frankl, 2000), which leads to an extensive dispersal of pollen. However, it is not excluded that isolation by distance occurs at the same scale according to maternal inherited molecular markers (i.e. in case of limited seed dispersal).

The absence of isolation by distance also questions the application of the metapopulation concept to *C. arvense*. It is not as straightforward to decide whether populations are organised as a metapopulation. Levins (1970) originally defines the metapopulation as "a population of populations that go extinct and recolonise". Since this first definition the concept of metapopulation rapidly evolved and become a paradigm in ecology (Hanski, 1998). Currently, the concept of metapopulation can be applied to situations of heterogeneous habitats (islands or patches). Due to isolation from other patches, individual patches have their own dynamics. On the other hand, patches are not completely isolated since they are still connected via gene flow (pollen or seed flow). Under this hypothesis, a genotypic isolation is often found within the metapopulation because of distance-dependant dispersal of propagules (Planes et al., 1996; Colas et al., 1997).

However, recent studies of the demography of *C. arvense* conducted over five years in the same area have demonstrated a dynamic pattern in time and space with important changes in the number, distribution and size of plant stands (Eber and Brandl, 2003). Whereas large patches were more or less persistent, Eber and Brandl found important turnover rates in small patches with an increasing probability for a patch to become extinct as patch size decreases. This result suggests important extinction–recolonisation events, which fit to the metapopulation concept. In a metapopulation, both local extinction and recolonization events occur continuously. When the recolonization rate of patches is lower than the rate of local extinction, the metapopulation itself becomes extinct and the species disappears from the landscape, which is not the case for *C. arvense*. Besides, this metapopulation dynamic of populations of *C. arvense* is confirmed by our study looking at the neutral genetic diversity of populations of *C. arvense* infested by the phytophageous insect *Urophora cardui*. No specific

influence of *U. cardui* on genotypic and genetic diversity of *C. arvense* infested populations was found. The metapopulation dynamics of the host plant might play a role to escape the selection pressures imposed by *U. cardui*. Thus, selection imposed by *U. cardui* on natural populations of *C. arvense* appeared weaker in than it was supposed in experimental studies.

Additionally, the results of the chapter III concerning the evolution of reproductive effort strongly suggest that local populations of *C. arvense* of different successional stages undergo different selection pressures. Once a population is founded, the selection regime will favour clonality (because of its better strategy for colonisation) leading to a decrease of reproductive effort in female plants of *C. arvense*. Consequently, the maintenance of sexual reproduction can not be explained at the population level, but at the metapopulation. Because of seed dispersal, which allows the colonisation of new favourable sites, sexual reproduction is selected and maintained at the metapopulation level.

Nevertheless, the genetic basis of the characters linked to the reproductive effort still need to be demonstrated as similar patterns could also result from phenotypic plasticity (Beasleigh and Yarranton, 1974; Hickman, 1975; Holler and Abrahamson, 1977; Ross and Quinn, 1977; Antos and Allen, 1999). The genetic basis of the trade-off between sexual and asexual reproduction in relation to the successional stage of populations has already been demonstrated in *Sparganium erectum* (Piquot et al., 1998). However, as the authors started from ramets instead of seeds, the demonstration was not complete as the decrease of reproductive effort could still result from senescence of plants.

To summarise, patterns of reproductive effort as well as those found in chapter II suggest a spatio-temporal dynamic of *C. arvense* populations, as it is classically found in metapopulation context, with numerous extinction-recolonisation events. However, this still needs to be confirmed by cytoplasmic data (i.e. seed dispersal) to find out at which scale recolonisation events occur.

Literature cited

Abrahamson WG. 1980 Demography and vegetative reproduction. Pages 89-106 *in* Solbrig OT ed. Demography and evolution of plant populations. Blackwell, Oxford.

Abrahamson WG, M Gadgil 1973 Growth form and reproductive effort in goldenrods (*Solidago*, compositae). American Journal of Botany 107:651-661.

Antos JA, GA Allen 1999 Patterns of reproductive effort in male and female shrubs of *Oemleria cerasiformis*: a 6-year study. Journal of Ecology 87:77-84.

Bachmann K 2001 Evolution and the genetic analysis of populations: 1950-2000. Taxon 50:7-45.

Bakker D. 1960 A comparative life-history study of *Cirsium arvense* (l.) Scop., *Tussilago farfara* L, the most troublesome weeds in the newly reclaimed polders of the former Zuiderzee. Pages 205-222 *in* Harper JL ed. The biology of weeds. Blackwell Scientific Publications, Oxford.

Baur B, B Schmid. 1996 Spatial and temporal patterns of genetic diversity within species. Pages 169-201 *in* Gaston KJ ed. Biodiversity - a biology of numbers and difference. Blackwell, Oxford.

Beasleigh WJ, GA Yarranton 1974 Ecological strategy and tactics of *Equisetum sylvaticum* during a postfire succession. Canadian Journal of Botany 52:2299-2318.

Bengtsson BO 2003 Genetic variation in organisms with sexual and asexual reproduction. Journal of Evolutionary Biology 16:189-199.

Bostock SJ, RA Benton 1979 The reproductive strategies of five perennial Compositae. Journal of Ecology 67:91-107.

Colas B, I Olivieri, M Riba 1997 *Centaurea corymbosa*, a cliff-dwelling species tottering on the brink of extinction: A demographic and genetic study. Proceedings of the National Academy of Sciences of the USA 94:3471-3476.

De Kroon H, J Van Groenendael 1997 The ecology and evolution of clonal plants. Backhuys Publishers, De Kroon H and Van Groenendael J eds., Leiden.

Eber S, R Brandl 2003 Regional patch dynamics of the weed *Cirsium arvense*, and possible implications for plant-animal interactions. Journal of Vegetation Science 14:259-266.

Ellstrand NC, ML Roose 1987 Patterns of genotypic diversity in clonal plant species. American Journal of Botany 74:123-131.

Eriksson O 1989 Seedling dynamics and life histories in clonal plants. Oikos 55:231-238.

Hamrick JL, MJ Godt. 1990 Allozyme diversity in plant species. Pages 43-63 *in* Brown AHD, MT Clegg, AL Kahler, and BS Weir eds. Plant population genetics, breeding and genetic resources. Sinauer Associates Inc., Sunderland.

Hamrick JL, MJ Godt 1996 Effects of life history traits on genetic diversity in plant species. Phil.Trans.of the Royal Society of London Ser B 351:1291-1298.

Hamrick JL, YB Linhart, JB Mitton 1979 Relationships between life-history characteristics and electrophoretically detectable genetic variation in plants. Annual review of ecology and systematics 10:173-200.

Hanski I 1998 Metapopulation dynamics. Nature 396:41-49.

Harper JL. 1977 Reproduction and growth. Pages 647-675 Population Biology of Plants. Academic Press, New York.

Heimann B, GW Cussans 1996 The importance of seeds and sexual reproduction in the population biology of *Cirsium arvense* - a literature review. Weed Research 36:493-503.

Hickman JC 1975 Environmental unpredictability and plastic energy allocation strategies in the annual *Polygonum cascadense* (Polygonaceae). Journal of Ecology 63:689-701.

Holler LC, WG Abrahamson 1977 Seed and vegetative reproduction in relation to density in *Fragaria virginia* (Rosaceae). American Journal of Botany 64:1003-1007.

Kingman JFC 1982 The coalescent. Stochastic Processes and their Applications 13:235-248.

Klekowski EJ. 1997 Somatic mutation theory of clonality. Pages 227-241 *in* De Kroon H and J Van Groenendael eds. The ecology and evolution of clonal plants. Backhuys Publishers, Leiden, Netherlands.

Klekowski EJ, PJ Godfrey 1989 Aging and mutation in plants. Nature 340:389-391.

Levins R. 1970 Extinction. Pages 75-107 *in* Gerstenhaber M ed. Some mathematical problems in biology. American Mathematical Society, Providence, Rhode Island.

Lian CL, R Oishi, N Miyashita, K Nara, H Nakaya, B Wu, ZH Zhou, T Hogetsu 2003 Genetic structure and reproduction dynamics of *Salix reinii* during primary succession on Mount Fuji, as revealed by nuclear and chloroplast microsatellite analysis. Molecular Ecology 12:609-618.

Loveless MD, JL Hamrick 1984 Ecological determinants of genetic structure in plant populations. Annual review of ecology and systematics 15:65-95.

Marshall DR, BS Weir 1979 Maintenance of genetic variation in apomictic plant populations. Heridity 42:159-172.

Mayer F 2000 Long distance dispersal of weed diaspores in agricultural landscapes - The Scheyern approach. Aachen.

McCauley DE 1995 The use of chloroplast DNA polymorphism in studies of gene flow in plants. TREE 10:198-202.

McLellan AJ, D Prati, O Kaltz, B Schmid. 1997 Structure and analysis of phenotypic and genetic variation in clonal plants. Pages 185-210 *in* De Kroon H and J Van Groenendael eds. The ecology and evolution of clonal plants. Backhuys Publishers, Leiden.

Nathan R, GG Katul, HS Horn, SM Thomas, R Oren, R Avissar, SW Pacala, SA Levin 2002 Mechanisms of long-distance dispersal of seeds by wind. Nature 418:409-413.

Oddou-Muratorio S, RJ Petit, B Le Guerroue, D Guesnet, B Demesure 2001 Pollen- versus seed-mediated gene flow in a scattered forest tree species. Evolution 55:1123-1135.

Ouborg NJ, Y Piquot, J Van Groenendael 1999 Population genetics, molecular markers and the study of dispresal in plants. Journal of Ecology 87:551-568.

Piquot Y, D Petit, M Valero, J Cuguen, P de Laguerie, P Vernet 1998 Variation in sexual and asexual reproduction among young and old populations of the perennial macrophyte *Sparganium erectum*. Oikos 82:139-148.

Planes S, R Galzin, F Bonhomme 1996 A genetic metapopulation model for reef fishes in oceanic islands: The case of the surgeonfish, *Acanthurus triostegus*. Journal of Evolutionary Biology 9:103-117.

Ross FH, JA Quinn 1977 Phenology and reproductive allocation in *Andropogon scoparius* (Graminaceae) populations in communities of different successional stages. American Journal of Botany 64:535-540.

Sherman-Broyles SL, JP Gibson, JL Hamrick, MA Bucher, MJ Gibson 1992 Comparisons of allozyme diversity among rare and widespread *Rhus* species. Systematic botany 17:551-559.

Silander JA. 1985 Microevolution in clonal plants. Pages 107-152 *in* Jackson JBC, LW Buss, and RE Cook eds. Population biology and evolution of clonal organisms. Yale University Press, New Haven and London.

Soane ID, AR Watkinson 1979 Clonal variation in populations of Ranunculus repens. New Phytologist 82:557-573.

Tackenberg O 2003 Modeling long distance dispersal of plant diaspores by wind. Ecological Monographs 73:

Tackenberg O, P Poschlod, S Bonn 2003 Assessment of wind dispersal potential in plant species. Ecological Monographs 73:

Walther-Hellwig K, R Frankl 2000 Foraging habitats and foraging distances of bumblebees, *Bombus* spp. (Hym., Apidae), in an agricultural landscape. Journal of Applied Entomology 124:299-306.

Watkinson AR, JC Powell 1993 Seedling recruitment and the maintenance of clonal diversity in plant populations - a computer simulation of *Ranunculus repens*. Oikos 81:707-717.

Widen B, N Cronberg, M Widen 1994 Genotypic diversity, molecular markers and spatial distribution of genets in clonal plants, a literature survey. Folia Geobotanica Phytotaxonomica 29:245-263.

Williams GC 1975 Sex and Evolution. Princeton University Press, Princeton.

Summary

The purpose of this thesis is to study the genetic diversity of the clonal plant *Cirsium arvense* (Asteraceae). *C. arvense* is one of the most frequent and most successful perennial weeds throughout the temperate zone. The increasing intervention of humans in natural landscapes, which mostly opens the vegetation, multiplies suitable sites for *C. arvense*. The species is nowadays omnipresent in agricultural or semi-natural landscapes. Because of its economic importance, *C. arvense* has been intensively studied during the last century. Although a lot of information has been accumulated about the ecology and biology of *C. arvense*, data about its genetic diversity are missing. Thus, the aim of this thesis was to conduct a mixed approach combining ecology and molecular biology to identify, on a landscape scale, the main factors acting and designing the genetic diversity of *C. arvense* in natural populations.

As a first step, Amplified Fragment Length Polymorphisms (AFLP) were developed to investigate the neutral genotypic and genetic diversity of natural populations of *C. arvense*. The study focuses on one hand on the impact of the ecological succession, and on the other hand on the occurrence of the phytophagen insect *Urophora cardui*. Another part of this work concerns the evolution of the reproductive effort of *C. arvense* and its dispersal capacity along the ecological succession.

To summarise, a high genotypic and genetic diversity was found in natural populations of *C. arvense*. This diversity seems to be maintained through time as populations from early and late successional stages present similar patterns of diversity. The AFLP molecular markers, which are mostly based on nucleic DNA, did not show an isolation by distance. This result suggests that an important nuclear gene flow might occur at our study scale (< 4 km). However, the same markers showed an important genetic differentiation between populations. The absence of influence of the phytophageous insect *U. cardui* on the genetic diversity of *C. arvense*, as well the patterns of evolution of the reproductive effort of *C. arvense* along a successional gradient, can both be understood in a metapopulation context. Therefore, these results suggest a spatio-temporal dynamic of populations of *C. arvense*, with extinction-recolonisation events.

Zusammenfassung

Das Ziel dieser Doktorarbeit ist die Untersuchung der genetischen Diversität der klonalen Pflanze *Cirsium arvense* (Asteraceae). *C. arvense* ist eine der häufigsten und erfolgreichsten mehrjährigen Unkräuter der gemässigten Klimazone. Der zunehmende Eingriff der Menschen in natürliche Landschaften, welcher meistens eine Öffnung der Vegetation zur Folge hat, vervielfacht geeignete Habitate für *C. arvense*. Die Spezies ist heutzutage allgegenwärtig in Agrar- und semi-natürlichen Landschaften. Wegen des durch sie verursachten ökonomischen Schadens wurde *C. arvense* während des letzten Jahrhunderts intensiv untersucht. Obwohl viele Informationen über die Ökologie und Biologie der Pflanze *C. arvense* vorhanden sind, fehlen Daten über ihre genetische Diversität. Das Ziel dieser Doktorarbeit war deshalb, ökologische und molekularbiologische Ansätze zu kombinieren, um herauszufinden, welche Hauptfaktoren die genetische Diversität von *C. arvense* in natürlichen Populationen im Landschaftsmassstab beeinflussen.

Es wurden molekulare Marker (Amplified Fragment Length Polymorphisms-AFLP) angewandt, um die neutrale genotypische und genetische Diversität natürlicher Populationen von *C. arvense* zu untersuchen. Dabei stand einerseits die Bedeutung der Sukzzession und andererseits das Auftreten des phytophagen Insekts *Urophora cardui* im Mittelpunkt. Ein weiterer Teil der Arbeit beschäftigt sich mit der Evolution des Reproduktionsaufwandes und der Ausbreitungsfähigkeit von *C. arvense* im Verlaufe der Sukzzession.

Die Ergebnisse zeigen, dass eine hohe genotypische und genetische Diversität in natürlichen Populationen von *C. arvense* vorhanden ist. Die Diversität scheint über die Zeit aufrechterhalten zu werden, da Populationen von frühen und späten Sukzzessionsstadien ähnliche Diversitätsmuster zeigen. Mit den molekularen Markern, die vor allem auf Kern-DNA beruhen, wurde keine Isolierung durch die Entfernung gefunden. Dies lässt vermuten, dass in unserem Untersuchungsmassstab (<4km) ein bedeutender nuklearer Genfluss auftritt. Trotzdem wurde eine starke Differenzierung zwischen den Populationen festgestellt. Der nichtvorhandene Einfluss des phytophagen Insekts *U. cardui* auf die genetische Diversität von *C. arvense* sowie das Evolutionsmuster des Reproduktionsaufwandes können in einem Metapopulations-Kontext verstanden werden. Diesbezüglich lassen die Ergebnisse eine räumlich-zeitliche Dynamik von *C. arvense* Populationen mit Auslöschungs- und Rekolonisierungsereignissen vermuten.

Magali Solé   *Curriculum vitae*

Personal details
    Date of birth :   August 13, 1975 in Maubeuge (France)
    Marital status :   Single
    Citizenship   :   French

Education

1993-1997

Undergraduate studies: University of Lille 1, France
Bachelor's Dregree (DEUG-Licence-Maîtrise) in Biology of Populations and Ecosystems, University of Lille 1
Title of dissertation: **"**Study of the polymorphism in two worker caste ants of the genus *Atopomyrmex*: *A. Calpocalycolus* an "ant-plant" and *A. Mocquerysii* a generalist ant"
Supervisor: Pr. Doyle McKey
Laboratory: Centre d'Ecologie Fonctionnelle et Evolutive CEFE/CNRS, 1919, route de Mende, 34293 Montpellier Cedex, France

09/1997-09/1998

Master's Degree in "Biodiversity: Genetics, History and Evolutionary Mechanisms", Universities of Orsay (Paris XI), Pierre et Marie Curie (Paris VI), Museum National d'Histoire Naturelle of Paris, Institut National d'Agronomie of Paris-Grignon
Title of dissertation: "Population biology and adaptation to environments polluted by heavy metals in *Arabidopsis halleri*"
Supervisor: Dr. Patrick de Laguérie
Laboratory: Génétique et Evolution des Populations Végétales, Université de Lille 1, France

07/1999-11/2002

Ph-D student at the Umweltforschungszentrum Leipzig-Halle GmbH
Title of the dissertation: "Factors affecting the genotypic and genetic diversity of the dioecious clonal plant *Cirsium arvense* at the metapopulation level"
Supervisor: Pr. Roland Brandl
Laboratory: Department of Community Ecology, UFZ Leipzig-Halle GmbH, Theodor-Lieser-Strasse 4, 06120 Halle an der Saale (Germany)

Actual position

12/2002-03/2004

working for the Groundwater Microbiology Research Group at the Umweltforschungszentrum Leipzig-Halle GmbH
Fellowship from the DFG-Graduate College "Adaptive Physiological and biochemical reactions to ecological important substances" at the Martin-Luther-University Halle-Wittenberg, Germany

AFLP: Principle and Application

AFLP: Principle and Application

1. Molecular Markers: General Introduction

In recent years, molecular markers and especially DNA-based markers, have been extensively used in many areas such as gene mapping and tagging (kliebenstein et al., 2001; Karp and Edwards, 1997), characterisation of sex (Flachowsky et al., 2001; Martinez et al., 1999), analysis of genetic diversity (Erschadi et al., 2000; Palacios et al., 1999; Lerceteau and Szmidt, 1999; Godt and Hamrick, 1999) or genetic relatedness (Mace et al., 1999; Roa et al., 1997; Brookfield, 1992). In population genetics, protein-based markers (allozymes) were the first markers developed and widely used (see Hamrick and Godt, 1990 for review). DNA-based methodologies are now the method of choice to differentiate closely related organisms (Widen et al., 1994; Ouborg et al., 1999; Avise, 1994). Moreover, the use of DNA-based markers allows efficient comparisons because genetic differences are detectable at all stages of development of the organism unlike allozymes which may show age dependent changes.
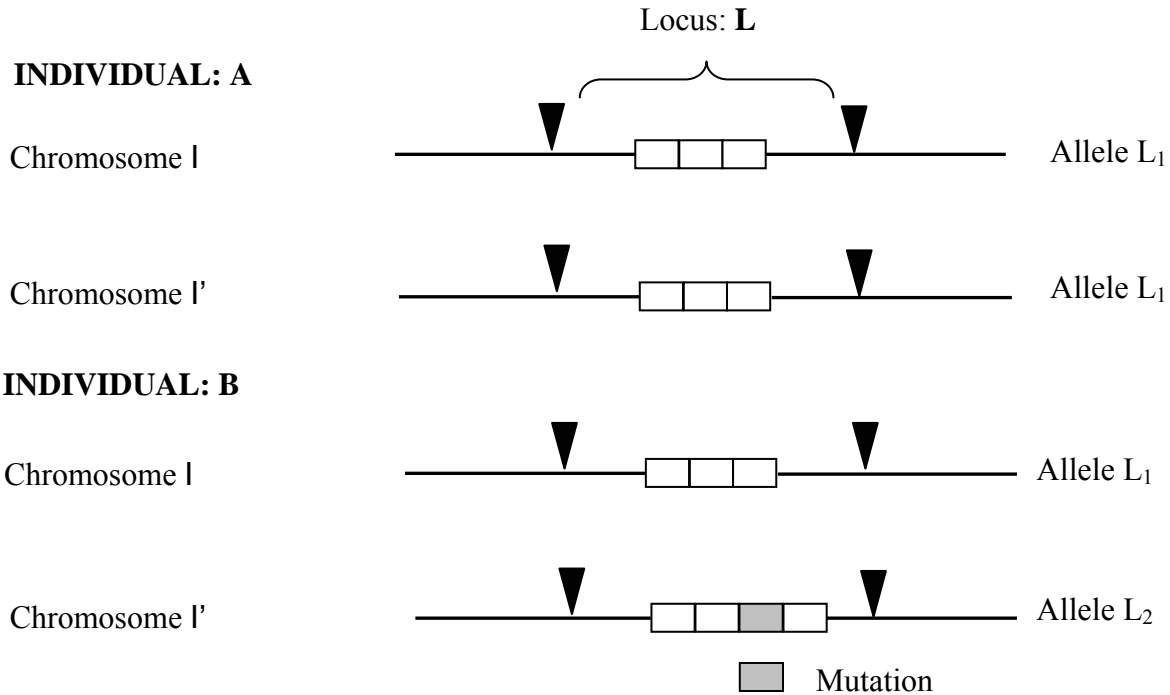
1.1. Definitions

According to Stansfield (1986), the term MARKER is usually used for "LOCUS MARKER". Each gene has a particular place along the chromosome called LOCUS (Box 1). Due to mutations, genes can be modified in several forms mutually exclusives called ALLELES (or allelic forms). All allelic forms of a gene occur at the same locus on homologous chromosomes. When allelic forms of one locus are identical, the genotype is called HOMOZYGOTE (at this locus), whereas different allelic forms constitute a HETEROZYGOTE. In diploid organisms, the GENOTYPE is constituted by the two allelic forms of the homologous chromosomes. All these definitions are illustrated in the Box 1 with a codominant marker as example.

Thus, MOLECULAR MARKERS are all loci markers related to DNA (markers can also be biochemical, or morphological).

Box 1: Molecular Markers: an example of a codominant marker

LOCUS: DNA fragment comprised between the two arrows (▼).

Locus: **L**

INDIVIDUAL: A

Chromosome l — Allele L_1

Chromosome l' — Allele L_1

INDIVIDUAL: B

Chromosome l — Allele L_1

Chromosome l' — Allele L_2

▨ Mutation

In this example, the locus **L** has two different **ALLELES**: L_1 and L_2.
GENOTYPE of individual A is "$L_1 L_1$", and genotype of Individual B "$L_1 L_2$"

In a case of a codominant marker, the pattern observed after a gel electrophoresis will be:

Ind A   Ind B
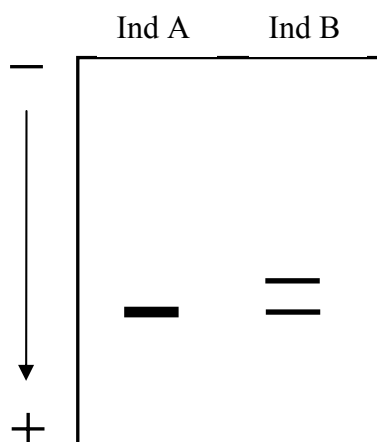
Individual A is **HOMOZYGOTE**
for the locus L

Individual B is **HETEROZYGOTE**
for the locus L

1.2. What is a good molecular marker for a population geneticists?

A good molecular marker should be/have:

> 1/ Mendelian inheritance: transmit from one generation to another.
>
> 2/ Polymorphic: present several alleles at the locus investigated (multiallelic).
>
> 3/ Codominant: allow the discrimination between homo and heterozygotes.
>
> 4/ Neutral: all alleles have the same fitness
>
> 5/ Not epistatic: one can determine the genotype of a phenotype irrespective of the genotype of the other loci.
>
> 6/ Independent of environment: no phenotypic plasticity.
>
> 7/ Frequent occurrence in the genome
>
> 8/ Even distribution throughout the genome
>
> 9/ Highly reproducible

The most frequently used markers in population genetics are allozymes (biochimical), RAPD (Random Amplified Polymorphic DNA; Williams et al., 1990), RFLP (Restriction Fragment Length Polymorphism, Botstein et al., 1980), AFLP (Amplified Fragment Length Polymorphism, Zabeau and Vos, 1993), minisatellite fingerprints, microsatellites and SSR (Single Sequence Repeats, Tautz and Renz, 1989).

The choice of a specific molecular marker depends of its suitability to answer a particular ecological question. For this purpose, the main difference among molecular markers is their degree of dominancy. Co-dominant markers enable for an easy estimation of allele frequencies. Therefore, they are suitable to estimate gene flow between populations or will be preferred, for example, to study dispersal. On the other hand, dominant markers can estimate genotypes but not the allele frequencies. Dominant markers are preferably used as fingerprints (Mueller and Wolfenbarger, 1999; Hongtrakul et al., 1997; Weising et al., 1995) and can be helpful in the identification of clones.

To study *Cirsium arvense* genetic diversity, we choose to develop a recent DNA marker called "AFLP" (Amplification Fragment Length Polymorphism). They full fill all the previous characteristics of "good molecular markers", except for the codominance. AFLP markers are dominant markers. Nevertheless, because of the high amount of polymorphism they can detect (Mueller and Wolfenbarger, 1999), AFLP markers were a priori the most

efficient markers for our study; e.g. to identify individual genotypes at the landscape scale, in a species described to be highly clonal (Moore, 1975; Bostock and Benton, 1979; Donald, 1994) and where genotypes were thus expected to be closely related. Moreover, from a technical point of view, no prior DNA sequence information is needed, many markers can be analysed in a short time, and only small amount of DNA is needed.
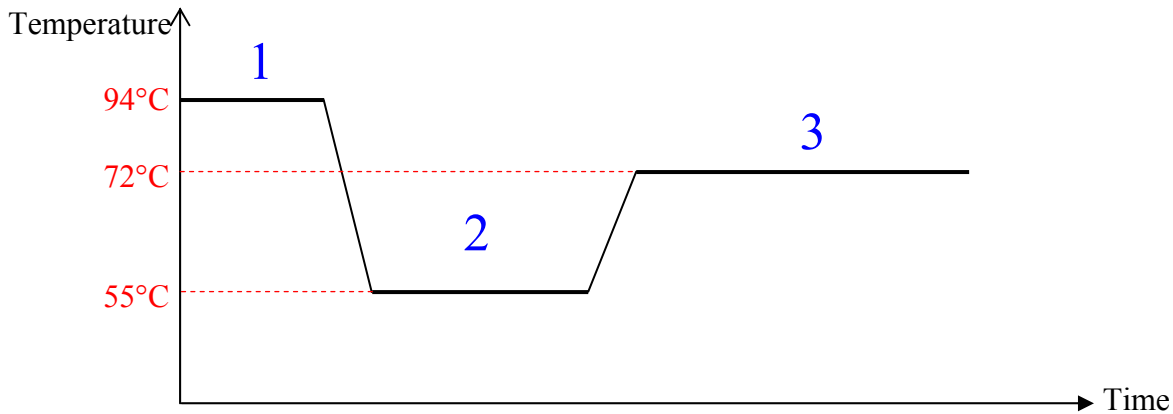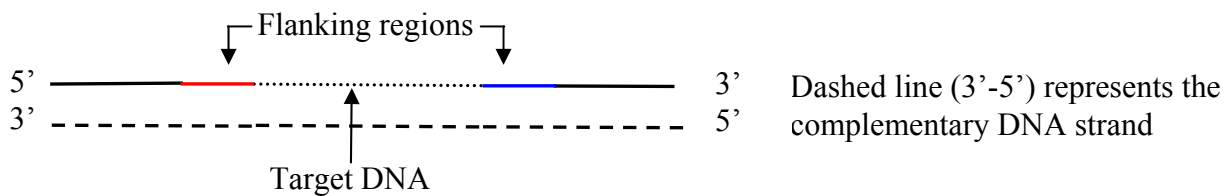
2. AFLP

2.1. Summary of the method

Amplified Length Fragments Polymorphism is a recent DNA fingerprinting technique developed by Zabeau and Vos (1993); but see also Vos et al. (1995) and Vos and Kuiper, 1997). This method is based on PCR amplification of selected restriction fragments of a total digested genomic DNA. Once labelled, amplified products are separated by electrophoresis. DNA fragments obtained range from 60 to 500 base pairs.

To be visualised, DNA polymorphism, which is usually made of small DNA fragments of few base pairs (up to 500), must be amplified. This amplification is commonly done by Polymerase Chain Reaction (Mullis et al., 1986; Mullis and Faloona, 1987). The PCR method can amplify specific DNA fragments through a precise priming of the polymerisation reaction occurring at each end of the target DNA. This precise priming is done by short oligonucleotidic sequences (Primers) able to anneal to the template DNA in the target zone. Primers are 18-24 base pairs long, synthesised in laboratory and correspond to a complementary DNA sequence designed in the flanking regions of the heavy strand of the target DNA (Box 2). The Polymerase Chain Reaction starts first with a high temperature phase (denaturation) that produces single-stranded DNA. Then, once temperature has reached the TM (Box 2), primers will bind to the template DNA. The
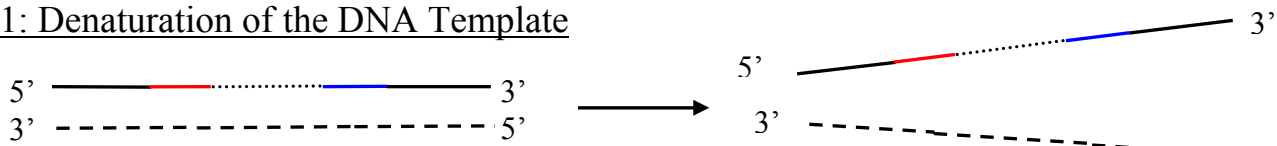
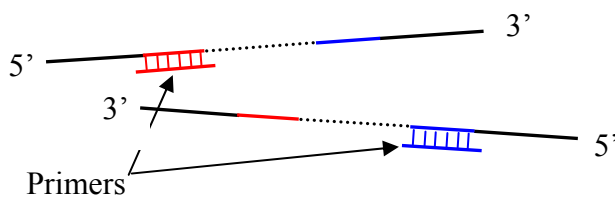# Box 2: Illustration of the Polymerase Chain Reaction

Temperature

94°C — 1

72°C — 3

55°C — 2

Time

Primers are designed one DNA strand

Flanking regions

5'
3'

Target DNA

Dashed line (3'-5') represents the complementary DNA strand

1: Denaturation of the DNA Template

5' ———— 3'
3' - - - - 5'

5' ... 3'
3' - - - 5'

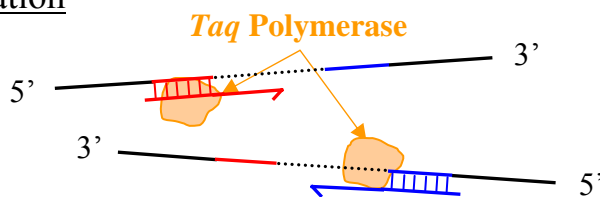Single stranded DNA is produced by high temperature

2: Annealing of the primers on the DNA template

5'
3'
3'
5'

Primers

The annealing temperature is defined by the oligonucleotides composition of the primers.

TM = 2(N# of AT) + 4(N# of GC)

3: Elongation

Taq Polymerase

5'
3'
3'
5'

The *Taq* polymerase optimally works at 72°C.

During the elongation, each DNA strand is polymerised in the direction 5'->3'

Taq polymerase recognises each double-stranded DNA as a start of synthesis and will continue the polymerisation reaction in the direction 5' → 3' as soon as the temperature has reached 72°C (optimal elongation temperature).

Therefore, in order to design specific primers, the sequences of the flanking regions of the target DNA must be known. This supposes detailed knowledge about the genome or further elaborated investigations to get it. This step usually requires high laboratory equipments and are most of the time, time consuming.

The originality of the AFLP method was to design and synthesise arbitrary primers first, and then to ligate them to target DNA fragments (Box 3). The AFLP arbitrary primers are called "adapters" and consist of a known sequence of 20 nucleotides. The target DNA sequences are DNA fragments generated by restriction enzymes. Fragments are produced from total genomic DNA by the combined action of two restriction enzymes. Then, adapters are ligated at each end of a restriction fragment by a protein ligase (New England Biolab®). Finally, adapters are used in a PCR as priming sites to amplify the restriction fragments. AFLP markers reveal a "restriction site" polymorphism and must be treated as dominant markers, since homozygotes and heterozygotes cannot be established unless breeding/pedigree studies are carried out to determine inheritance patterns of each fragment. However, the large number of fragments gives an estimate of variation across the entire genome, which thus gives a good general picture of the level of genetic variation of the studied organism.
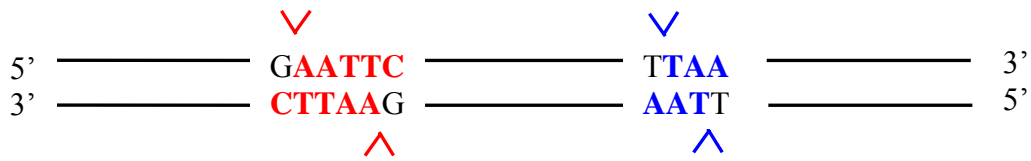
2.2. Basic steps of AFLP fingerprinting
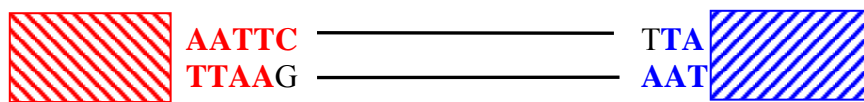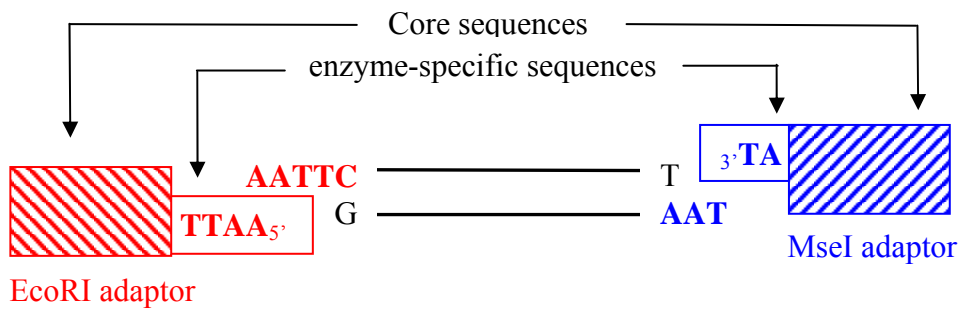
2.2.1. DNA extraction

Clean and high molecular weight DNA is a prerequisite for AFLP. In our study, we extracted DNA according to Doyle & Doyle method (Doyle and Doyle, 1988). This method is based on the CTAB procedure. For more details, refer to the protocol (3.3) and troubleshooting (3.4) parts.
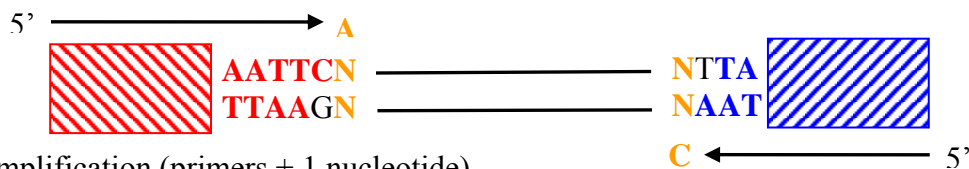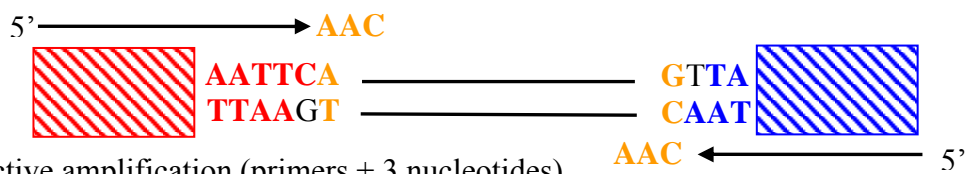
Box3: AFLP: Principle

Genomic DNA

```
            ∨                        ∨
5' ———————— GAATTC ———————— TTAA ———————— 3'
3' ———————— CTTAAG ———————— AATT ———————— 5'
            ∧                        ∧
```

Action of the Restriction Enzymes **EcoRI (>)** & **MseI (>)**

Core sequences
enzyme-specific sequences

```
  AATTC ——————— T          3'TA
G ——————— AAT
TTAA5'
```

EcoRI adaptor   MseI adaptor

```
AATTC ——————————— TTA
TTAAG ——————————— AAT
```

Ligation of the adaptors (known DNA sequence = futur primers)

```
5' ——————————→ A
AATTCN ——————————— NTTA
TTAAGN ——————————— NAAT
                   C ←——————————— 5'
```

Preamplification (primers + 1 nucleotide)

```
5' ——————————→ AAC
AATTCA ——————————— GTTA
TTAAGT ——————————— CAAT
                   AAC ←——————————— 5'
```

Selective amplification (primers + 3 nucleotides)

```
AATTCAAC ——————————— TTCTTA
TTAAGTTG ——————————— AACAAT
```
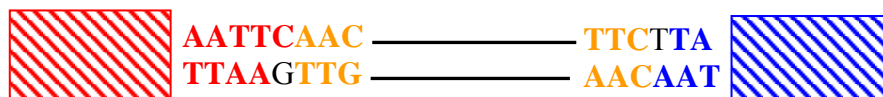
Amplified fragments will be separated by electrophoresis

2.2.2.   Restriction

Restriction fragments of the genomic DNA are produced by using two different restriction enzymes: a frequent cutter (the four-base restriction enzyme *Mse*I) and a rare cutter (the six-base restriction enzyme *Eco*RI) (Box 3). The frequent cutter serves to generate small fragments, which amplify well and which have the optimal size range for separation on a sequence gel, whereas the rare cutter limits the number of fragments to be amplified.

2.2.3. Ligation of oligonucleotide adapters

Double-stranded adapters consist of a core sequence and an enzyme-specific sequence (Box 3). Therefore, adapters are specific for either the *Eco*RI site or the *Mse*I site. Usually restriction and ligation take place in a single reaction. Ligation of the adapter to the restricted DNA alters the restriction site in order to prevent a second restriction from taking place after ligation has occurred. The core sequence of the adapters consists of a known DNA sequence of 20 nucleotides, which will be used later as primer in the PCR.

2.2.4. Pre-amplification

This step is a normal PCR where the adapters are used as primers. This first PCR, called preamplification, allows a first selection of fragments by only amplifying the DNA restriction fragments that have ligated an adapter to both extremities. Additionally to the adapter sequences, the primers used for the pre-selective amplification have a supplementary base. This extra base enables another first selection by amplifying ¼ of the fragments that have ligated an adapter to both extremities. These first three steps (DNA extraction, restriction/ligation and preamplification) can be run and visualised on a 1.6 % agarose gel (Figure 1).
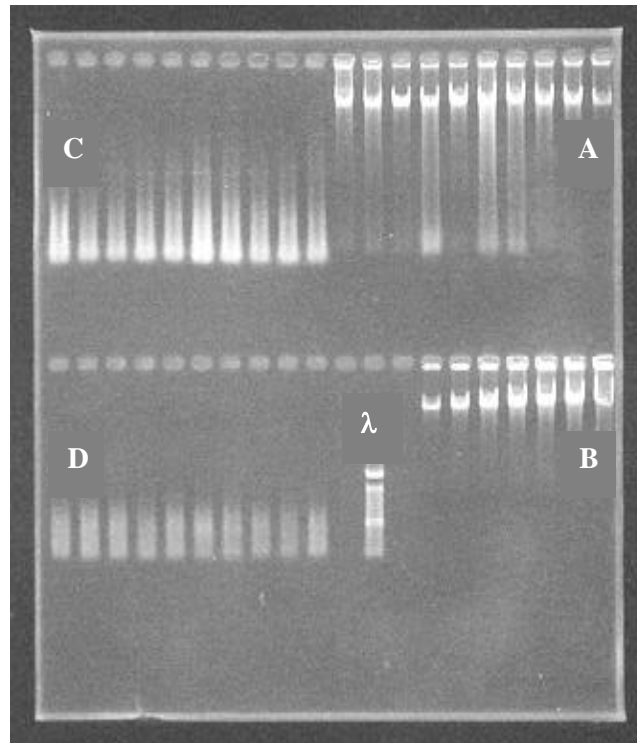
Figure 1: The three first steps of the AFLP process (DNA extraction, restriction/ligation and preamplification) run on a 1,6% agarose gel. The 10 first lanes (**A**) represent genomic DNA of 10 *Cirsium arvense* samples. According to the Lambda DNA concentration standards (**B** from left to right: 0.125; 0.25; 0.5; 0.75; 1; 1.5; 2 μg) *C. arvense* DNA concentration can be estimated to be in around 5ng per lane (5μl DNA load). (**C**) restriction/ligation of the same 10 samples. Genomic DNA was restricted with Mse I and EcoR I enzymes. The restriction produces a large quantity of small size DNA fragments. During the preamplification (**D**), only a part of the restricted fragments are amplified. The preamplification leads to a homogeneous DNA smear ranging from 100 to 800 bp. (λ) represents a DNA size marker.

2.2.5. Amplification

The aim of this step is to restrict the level of polymorphism and to label the DNA. For this second amplification, we added three more nucleotides at the 3' end of the primer sequence used for the preamplification (= adapters sequence + 3 nucleotides; Box 3). These two additional nucleotides make the amplification more selective and will decrease the number of restriction fragments amplified (polymorphism). Moreover, one of the primers (usually the EcoRI primer) is labelled with a fluorescent dye, and will allow the visualisation of DNA during the migration.

2.2.6. Electrophoresis

The PCR products are denaturated and run on acrylamide gel (DNA sequencer). In our study, samples were run on an ABI Prism 310 (Figure 2).

Figure 2: ABI Prism sequencer

A thin capillary containing a polymer replaces the usual acrylamide gel. The electrophoresis conditions we used for fragments analysis can resolve DNA fragments differing just by one base pair. Samples are loaded in a track, and run one after the other through the capillary.

All fragments are separated with regard to length, smaller fragments running first. Once passing the laser, a dye attached to the primer is excited and emits a fluorescent signal that is then collected by a computer. The results of fluorescence are visualised on the computer as peaks, called Electropherogramms (Figure 3). Each peak corresponds to a band on a normal acrylamide gel. Amplified fragments range from 30 to 400 base pairs.
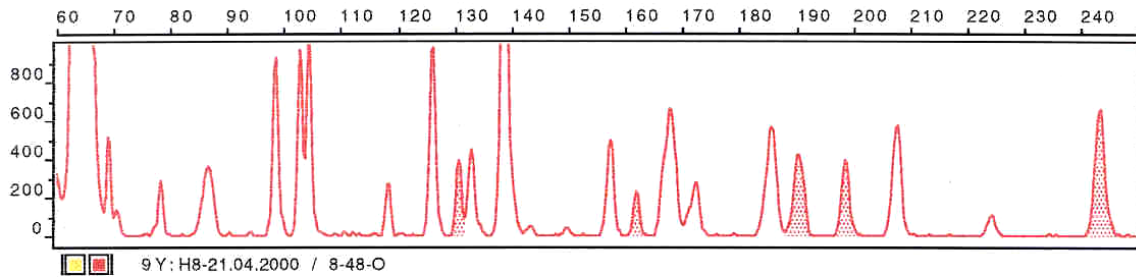
Figure 3: AFLP electropherogramm. On the figure, each peak corresponds to a DNA restriction fragment of a precise length. From the left to the right, the 5 shaded peaks correspond to DNA fragments having respectively 131; 161; 190; 198 and 242 base pairs.

3. How to develop AFLP markers

3.1. Primer choice

Choice of number and sequence of primers that will be used for the selective amplification is an important step of the AFLP process because they will later determine the level of polymorphism accessible in the studied species.

To test for good primer combination we used the *Selective Amplification Modules for Small Plant Genomes* from Applied Biosystems. The *Selective Amplification Modules for Small Plant Genomes* from Applied Biosystems provided eight EcoRI-labelled primers and eight MseI primers. Thus, 64 selective primer combinations were available. An initial screening using all these 64 primer combinations was performed on a random sample of 10 individuals across all sampled populations.

For our study, we choose primers according to their level of inter- and intrapopulation polymorphism (figures 4 and 5). A good primer must discriminate individuals coming from different populations, but should still be polymorphic enough at the population level to precisely identify clones.

From the screening of the 64 primers, two primer combinations EcoRI-ACC / MseI-CTG and EcoRI-ACG / MseI-CTT appeared efficient at the inter-population level, and sufficiently polymorphic to discriminate clones within populations.

Figures 4 and 5 represent respectively inter and intrapopulation polymorphism obtained with the primer combination EcoRI-ACC / MseI-CTG.

In Figure 4, about interpopulation polymorphism, the electropherograms of the same colour represent two individuals coming from the same population. From up to down: individuals 8-48-O, 8-44-O, 47-5-O, 47-6-O, N2-1-O, N2-2-O were respectively 2 individuals of populations 8, 47 and N2. In each case, both individuals of the same population present an identical electropherogramm. Therefore, according to the AFLP patterns they have identical genotypes and thus belong to the same clone. In figure 4, shaded peaks indicate interpopulation polymorphism. Shaded peaks are just present in the population they belong, and offer a good discrimination between populations.

Figure 5 illustrates an example of intrapopulational polymorphism. In figure 5, the electropherograms of the same colour represent two individuals coming from the same population and having different genotypes. From up to down: 8-48-O, 8-45-O, 47-6-O, 47-10-O, N2-2-O, N2-3-O were respectively 2 individuals of populations 8, 47 and N2. Shaded peaks indicate intrapopulation polymorphism, e.g. peaks that are specific to one individual within a population. These results show that the primer combination EcoRI-ACC / MseI-CTG was also efficient to discriminate between individuals of a same population.

Remarks: We found similar results with the primer combination EcoRI-ACG / MseI-CTT (results not shown).
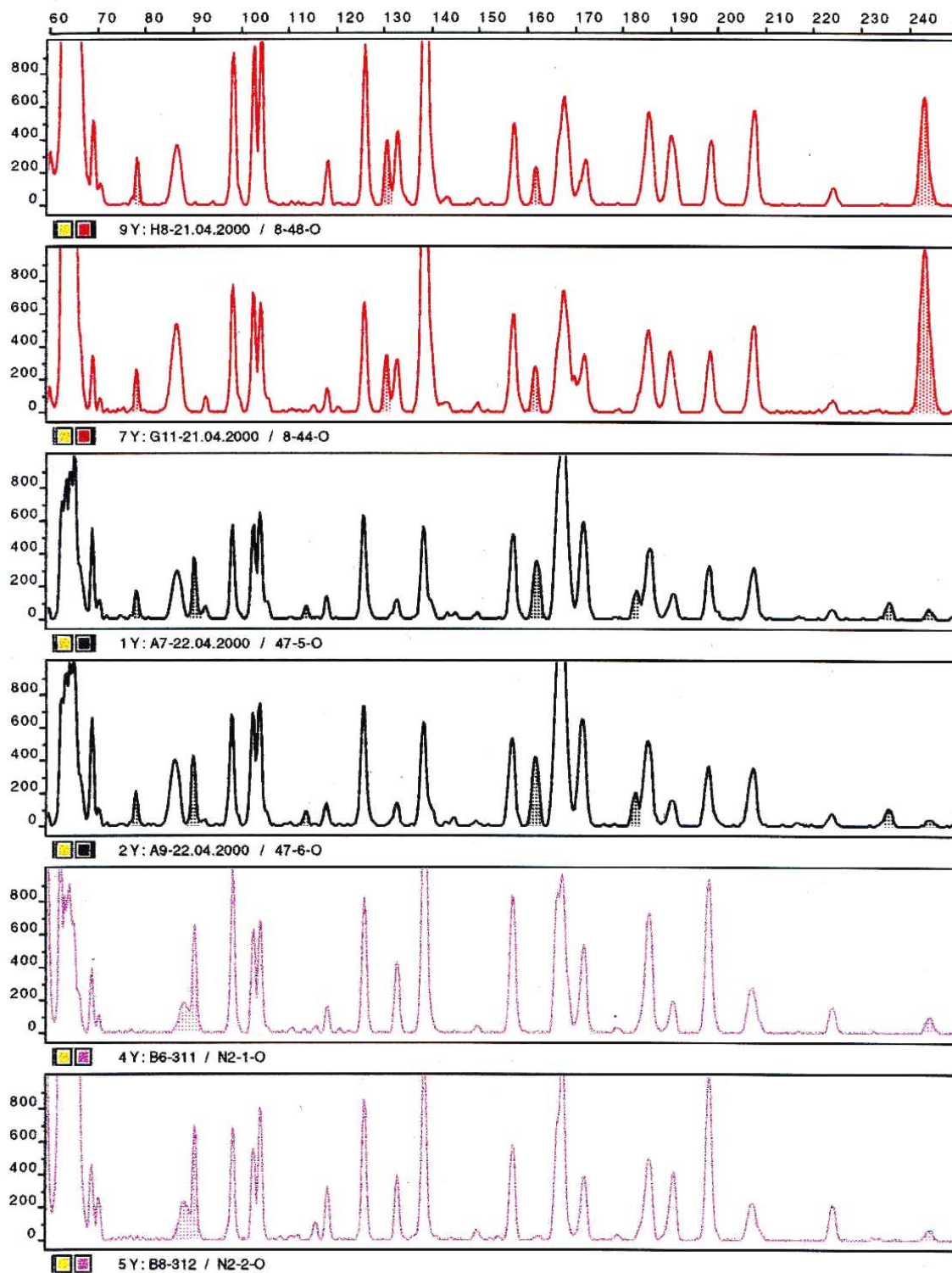
Figure 4: Example of polymorphism interpopulational obtained with the primer combination EcoRI-ACC / MseI-CTG. Electropherograms of the same color represent two individuals coming from the same population and having the same genotypes (clones). Shaded peaks indicate interpopulation polymorphism.
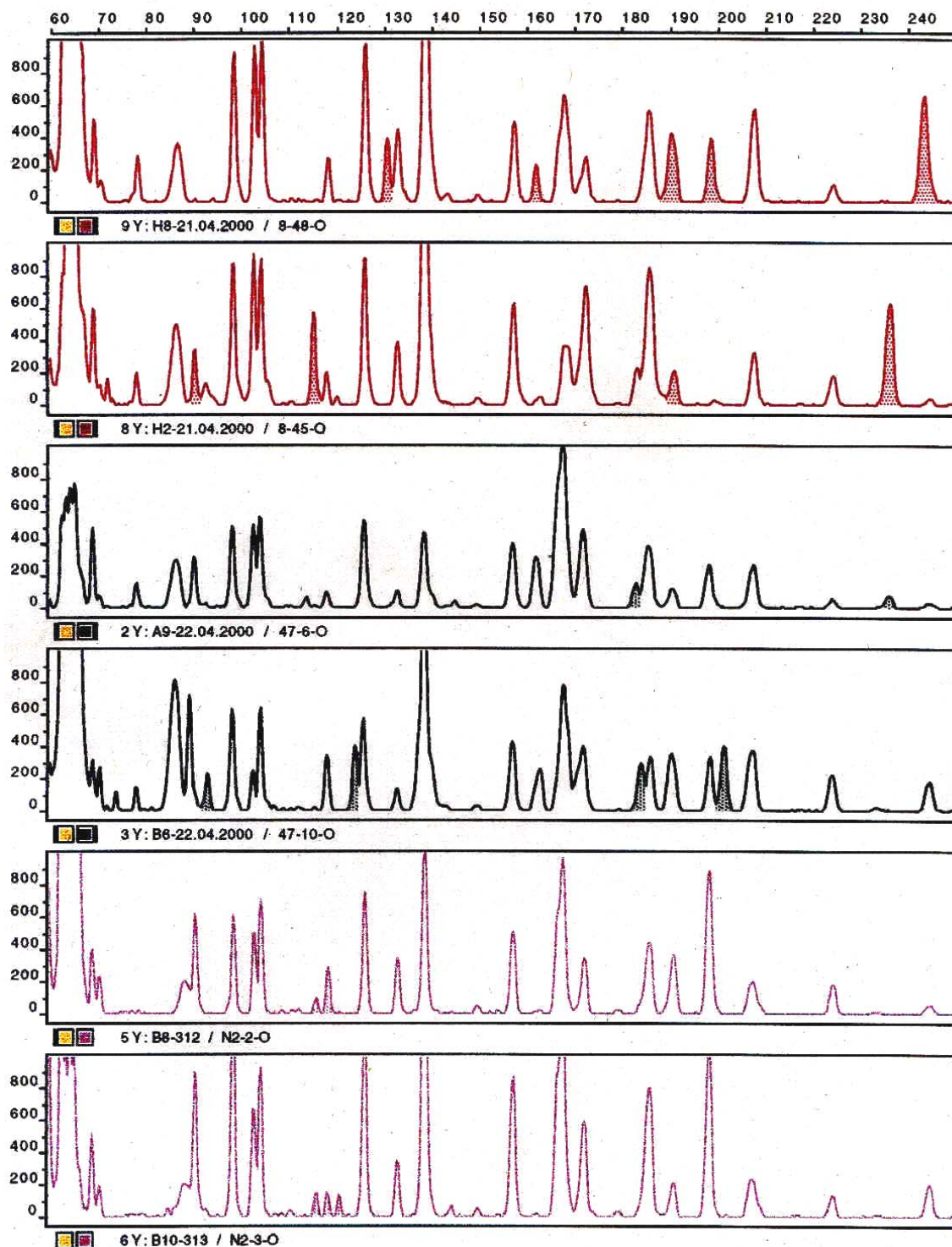
Figure 5: Example of polymorphism intrapopulational obtained with the primers EcoRI-ACC / MseI-CTG. Electropherograms of the same colour represent two individuals coming from the same population and having different genotypes. Shaded peaks indicate intrapopulation polymorphism.

3.2. Repeatability

Compared with RAPD, the AFLP technique is highly reproducible (Mueller and Wolfenbarger, 1999). In order to check for the reproducibility of our AFLP method, we repeated the whole process with different DNA extractions. Results are presented in figure 6.
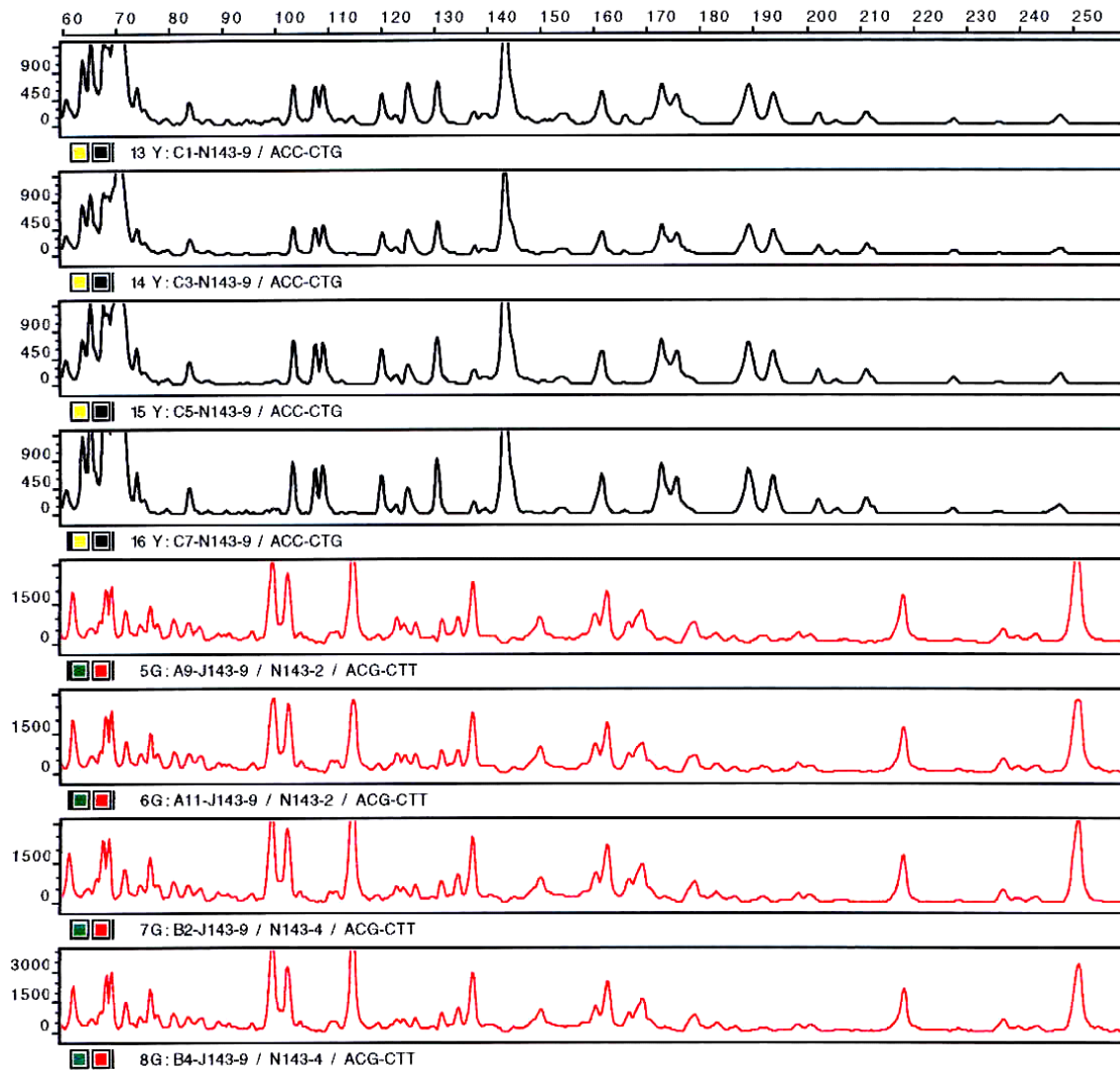

Figure 6: Repeatability test. The figure shows four independent AFLP processes realised from four DNA extractions of the same sample (N1-143-9) for the primers ACC-CTG (in black) and for the primers ACG-CTT (in red).

Box 4: DNA extraction: components and their effects

Components	Name	Nature	Effects
pH			- Inhibition of degradative enzymes. (ex: DNases act at pH=7)
Tris		Buffer	- Maintains the pH
EDTA	Ethylenediamine-tetraacetate		- Chelation of divalent cations (Ca^{2+}, Mg^{2+}...) - Inhibition of metal-dependant enzymes (ex: nucleases, DNases)
Na or K		Salt	- Stabilisation of nucleic
Proteinase K		Enzyme	- Digestion of proteins
SDS	Sodium Dodecyl Sulfate	Anionic detergent	- Solubilisation of cellular membranes - Denaturation of proteins
CTAB	Hexadecyltrimethyl ammoniumbromide	Cationic detergent	- Solubilisation of cellular membranes - Denaturation of proteins - Formation of a complex with the DNA
CIA	Chloroform-isoamyl-alcohol		- Extraction proteins
Isopropanol			- Precipitation of the CTAB-DNA complex
ß-mercaptoethanol		Reducing agents (antioxidants)	- Inhibition of the oxidation processes - Protection of DNA from quinones, disulfites, peroxidases, polyphenoloxydases
Glutathione			
Cysteine			
DDT	Dithiothreitol		
Ascorbic acid			
$Na_2S_2O_5$	Sodium bisulfite		
PVP	Polyvinylpyrro-lidone	Polyphenols absorbents	- Decrease the effect of polyphenols, quinones, tanins - BSA: denaturation of degradative enzymes
BSA	Bovine Serum Albumin		
Diethylpyrocarbonate		Polyamines	- Protection from RNases
Bentonite			
Spermine			
Spermidine			
DIECA	Diethydithiocar-bamic acid	Phenoloxidase inhibitors	- Cu^{2+} chelation
Cyanide			- Protection from heavy metal oxidases
NH			- Protection from H^+
Aurintriarbolylic acid			- Inhibition of nucleases

3.3. <u>Protocols</u>

3.3.1. <u>DNA extraction</u>

We extracted DNA according to the Doyle & Doyle method (Doyle and Doyle, 1988). This method is based on the CTAB procedure (Webb and Knapp 1990). The CTAB (Hexadecyltrimethylammoniumbromid) is a cationic detergent (Box 4) that forms a complex with the DNA. The CTAB-DNA complex is then separated from the cellular debris by chloroform. In this step, we can observe two phases: a superior clear aquatic phase containing the DNA, and a denser inferior one containing the chloroform and all the secondary components (polysaccharides, proteins etc…). After centrifugation, cellular debris can usually be observed at the interface. The purification by chloroform can be repeated several times (2, 3 times).

DNA is soluble in water and can be precipitate with salts (NH_4^+, Na^+, NaOAc, NaCl, NH_4OAC...) and ethanol 100% or isopropanol. To precipate DNA, one can equally use one of the following salts: Na^+(3M) or NH_4^+(7.5M), with one of the following alcohols: ethanol 100% or isopropanol. The precipitation is done in these proportions: 10% and 50% of the final volume for Na^+ and NH_4^+ respectively, and 250% and 70% of the final volume for ethanol and isopropanol respectively.

After the precipitation, the DNA molecule must be washed with ethanol to remove the salts used for the precipitation. DNA is then dissolved and stored in a Tris/EDTA buffer (TE buffer: Tris (10mM) / EDTA (1mM)) pH 8.

The protocol we used to extract DNA was the following. About 150 mg of plant tissue were first frozen in liquid nitrogen and ground in a mixer mill. Then, we added 700 μl of extraction buffer (Box 5) to the ground material and incubated the mixture 45 minutes at 60°C. At this stage, it is important to vortex to get a homogenous solution free of clumps to prevent DNA degradation (3.4.1). At the beginning of the incubation, DNA is still in the cells and can not be broken by vortexing. Afterwards, vortexing should be avoided.

Box 5: DNA extraction: buffers and stock solutions

Extraction buffer

	Final concentration	Stock solution	For 500 ml of buffer
NaCl	1.4M	5 M	140 ml
Tris-HCl pH 8.0	100 mM	1 M	50 ml
EDTA pH 8.0	20 mM	0.5 M	20 ml
CTAB	2 %		10.0 g
$Na_2S_2O_5$	1 %		5.0 g

Autoclave the solution and then add ß-Mercaptoethanol (0.2 %).
Check the pH again.

CIA-solution (Chloroform:Isoamylalcohol 24:1)

	Final concentration	For 250 ml of CIA
Chloroform	96 %	240 ml
Isoamylalcohol	4 %	10 ml

Store the solution at 4°C.

TE-Buffer (TE 10/1)

	Final concentration	Stock solution	For 250 ml of TE
Tris-HCl pH 8.0	10 mM	1 M	1.0 ml
EDTA pH 8.0	1 mM	0.5 M	0.2 ml

After incubation, we added 700 µl of CIA 4°C (Box 5), mixed carefully and let the solution 5 min at room temperature. We centrifuged 10 min at 10.000 rpm, and recuperated the floating phase and repeated this purification step several times (2 times).

DNA was then precipitated with a 3M Na$^+$ salt (10%) and 0.7 volume of cold isopropanol (-20°C). We recuperated the DNA pellet after centrifugation (5 min 10.000 rpm), washed for 15 min in 500 µl of 75% and 100 % ethanol, dried at room temperature and then redissolved in 100 µl of TE buffer.

For each sample, we checked DNA quality and quantity on a 1.6% agarose gel.

3.3.2. Restriction/Ligation

The Mix prepared for the restriction / ligation is presented in the Box 6. Because of their high concentrations, we first diluted the enzymes in order to have reasonable volumina for pipetting. As enzymes are conserved in glycerol, in order to limit the amount of glycerol and to prevent the reaction from inhibition, the final volume of enzymes should not exceed 5% of the total volume (0.55 µl). Therefore, we divided this volume in three, and got the optimal volume of 0.18 µl for each enzyme. Then, we diluted enzymes within their respective buffers and prepared relevant stock solutions to have just to pipette 0.18 µl to get the right final concentration (Box 6, Restriction / Ligation Mix). Then, we added 11µl of the Mix to 5.5 µl of DNA containing 500 ng of genomic DNA. The reaction mixture was incubated 2 hours at 37°C in a MWG-Biotech Primus 96 thermocycler.

3.3.3. Preamplification

Products of the restriction ligation were diluted 1:2. We used 4 µl of the restriction / ligation as DNA template for the preamplification PCR (Box 6, preamplification). Amplification reaction was performed in a MWG-Biotech Primus 96 thermocycler with the following program: 2' 72°C, (20'' 94°C, 30'' 56°C, 2' 72°C)X30, 30' 60°C, 4°C for ever.

Restriction / Ligation

	Stock solution	Final concentration	For 5.5 µl (one reaction)
H_2O			0.21 µl
T4 DNA ligase buffer	X 10	X 1	1.1 µl
NaCl	0.5 M	0.05 M	1.1 µl
BSA	1 mg/ml	0.05 mg/ml	0.55 µl
MseI adaptor	50 pmol/µl	5 pmol/µl	1 µl
EcoRI adaptor	5 pmol/µ	0.5 pmol/µl	1 µl
MseI	5.5 U/µl	1 U	0.18 µl
EcoRI	27.77 U/µl	5 U	0.18 µl
T4 DNA Ligase	372.22 U/µl	67 U	0.18 µl
final Volume			*5.5 µl*

Preamplifiction

	Stock solution	Final concentration	For 20 µl (one reaction)
PCR buffer	X 10	X 1	2 µl
dNTPs	2mM (X 10)	200 µM	2 µl
Primers MIX	2mM (X 10)	200 µM (X 1)	1 µl
Taq polymerase	5U / µl	2U / 50 µl	0.16 µl
H_2O			10.84 µl
ADN			4 µl
Volume final			*20 µl*

Selective Amplification

	Stock solution	Final concentration	For 20 µl (one reaction)
PCR buffer	X 10	X 1	2 µl
dNTPs	2mM (X 10)	200 µM (X 1)	2 µl
MseI primer	2mM (X 10)	200 µM (X 1)	1 µl
EcoRI primer	2mM (X 10)	200 µM (X 1)	1 µl
Taq polymerase	5U / µl	2U / 50 µl	0.16 µl
H_2O			10.84 µl
ADN			3 µl
Volume final			*20 µl*

3.3.4. Amplification

Products of the preamplification were diluted 1:10 and we used 3 μl as DNA template for the selective amplification PCR (Box 6, selective amplification). Amplification reaction was performed in a MWG-Biotech Primus 96 thermocycler with a touchdown PCR.

In a touchdown PCR the annealing temperature decreases by 1°C every cycle to a 'touchdown' annealing temperature which is then used for further cycles. Touchdown PCR is usually used to reduce non-specific amplification by optimising the optimal annealing temperatures. The idea of a touchdown PCR is that any differences in TM (Box 2) between correct and incorrect annealing gives a 2-fold difference in product amount per cycle. Therefore, touchdown PCR enriches for the correct product over any incorrect products. The touchdown PCR program we used was the following: 2' 72°C, (20'' 94°C, 30'' (66°C decreased by 1°C during 10 cycles), 2' 72°C)X10, (20'' 94°C, 30'' 56°C, 2' 72°C)X20, 4°C for ever.

3.3.5. Electrophoresis

We used 1μl of each selective amplifications (two primer combinations) that we mixed to 12 μl deionised formamide and 0.3 μl of Genescan-500 (ROX) size standard. ROX 500 is made of labelled DNA fragments of known size and is run with each sample. After a run, the internal standard will allow the calibration of all AFLP signals produced by primer dies recorded during the electrophoresis. The use of an internal standard makes comparisons between independent samples feasible.

To get a good signal it is important to mix properly the PCR products with the formamide and the size standard (quickly vortex). DNA was then denatured by heating each tube to 95°C for 5 minutes and quick chill on ice. Samples were finally loaded in the ABI Prism 310. The electrophoresis was performed at 60°C, DNA was injected into the capillary during 20'' and we used the GS STR POP 4 module to detect peaks.

3.4. Critical phases of the process & troubleshooting

The AFLP process comprises five consecutive steps (DNA extraction, restriction/ligation, preamplification, amplification, electrophoresis) and only a success of all these steps will lead to good electropherograms (e.g. good signal intensity and quality of peaks). Compared with

other molecular markers, like microsatellites or RAPD that just consist in a single PCR, AFLP markers are more sophisticated. The implication of enzymes (restriction and ligation enzymes) makes the method particularly sensitive and sometimes problematic. In particularly, presence of secondary compounds in the DNA (like polyphenols or polysaccharides) can damage DNA, inhibit restriction enzymes or the *Taq* polymerase. Therefore, in order to avoid further technical problems, the first step (DNA extraction) appears to be of the most importance. This section is about the most common problems encountered during DNA extraction, and advices to resolve them.
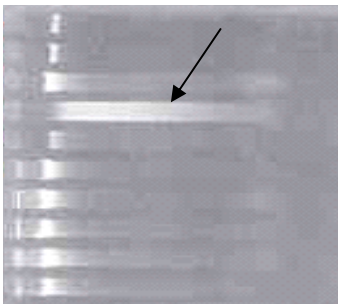
3.4.1. DNA degradation

Figure 7: Detail of a 1.6% agarose gel of *C. arvense* genomic DNA. Whereas all lanes show a clear single band attesting for a high molecular weight DNA, the lane number 4 (cf arrow on the picture) presents a degradated DNA. No clear band of DNA is visible; DNA appears in a smear

DNA degradation is characterised by the generation of DNA fragments through the action of endogenous nucleases. Endogenous nucleases are active at room temperature, but inhibited are low and high temperature. Therefore, it is important to pay attention of temperatures all over the DNA extraction. For example, the presence of clumps in the solution containing ground tissues and the extraction buffer during the 60°C incubation phase (cf. 3.3.1.) can lead to a decrease of the temperature within a clump and therefore activate the endogenous nucleases. However, DNA degradation can also be due to a break of the cold chain before the extraction.

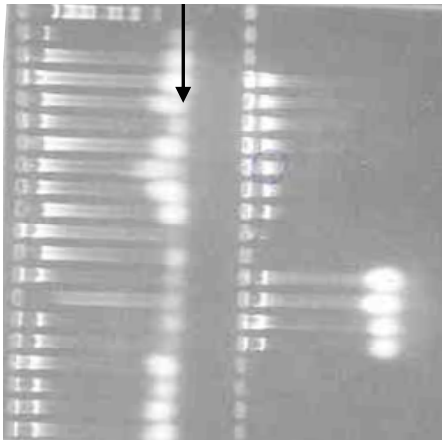### 3.4.2. Simultaneous isolation of RNA

Figure 8: Detail of a 1.6% agarose gel of *C. arvense* genomic DNA showing a simultaneous extraction of RNA. On the picture, genomic DNA is represented by a clear single band closed to the lanes, whereas spots at the end of the lanes (cf arrows on the picture) represent small size RNA.

In some cases, RNA can be extracted simultaneously with DNA and then disturb the AFLP process. To avoid extracting RNA it is important to check the pH of the extraction buffer. A too acidic pH could lead to RNA extraction. RNA can also be removed by an additional step including a RNase A. The RNase A should be prepared as follows:

> Dilute 10 mg/ml of RNase (use a DNase-free RNAses to avoid a digestion of the DNA) in a RNase buffer (RNase buffer: 10 mM Tris HCl; 15 mM NaCl; pH: 7.5). Boil the solution for 15 min (95°C); cool it at room temperature, stored at -20°C. RNase should be used with a final concentration of 100ug/ml.

The RNase step can be done "as preventive" during the incubation of tissues with extraction buffer at 60°C, or just after the protein extraction by CIA and before the precipitation of the CTAB-DNA complex. In both cases, adequate volume of RNase leading to a final concentration of 100ug/ml must be added.

In case RNA has already been extracted (present on the agarose gel), the RNase step can still be done after the DNA extraction. For that, add the adequate volume of RNase to the extracted DNA (TE buffer and DNA), mix and incubate 30min room temperature (DNA can eventually be precipitated again).

3.4.3. Isolation of polysaccharides

Polysaccharides can easily be recognised because of their high viscosity. To remove polysaccharides one can either increase the CTAB concentration, or increase the concentration of NaCl (2M) prior to the first alcohol precipitation (polysaccharides remain soluble in high salt concentration). If DNA is already extracted, one can remove polysaccharides by redissolving them by increasing the NaCl concentration of the TE buffer up to 2M. Then, DNA must be precipitated again with 2 volumes of ethanol.

3.4.4. Isolation of polyphenols

Polyphenols lead to brown colouration of the DNA. They can damage DNA, inhibit restriction enzymes or *Taq* polymerase. Polyphenols can be removed by increasing the concentration of polyphenol absorbents up to 2%. One can also add phenoloxidase inhibitors, or elevate the concentration of antioxidants.

4. Genotyping

Genotyping of *C. arvense* electropherograms was one of the most time consuming phase of the study. First, we used the *Genotyper* software of ABI PRISM *PE biosystems*. In *Genotyper*, it is possible to standardise electropherogram peaks into categories. Each peak is defined in a category according to its width and its intensity (e.g. to its high). Categories are recorded and then, runs are screemed and automatically genotyped. Such an automatic labelling of peaks appears relevant especially when numerous samples must be treated (307 in our study). Although we spent a lot of time to define all possible categories, we did not manage to genotype automatically our data. This inability of using *Genotyper* can be explained by a too high sensitivity of the program coupled with a heterogeneity among runs. Because of the size reconstruction method of fragments (3.3.5) each AFLP peak gets a very precise position: for example: 55.6; 55.9; 56.1 base pairs and 56.2; 56.4; 56.7 base pairs. In the case that the first three peaks belong to the length "55 bp" and the next 3 peaks belong to the length 56 bp, it is difficult to define 2 independent categories which do not overlap. This example, repeated over 307 samples and 93 AFLP markers, made an automatic genotyping impossible.

Therefore, in a second approach, we only used the program *Genotyper* to produce a preliminary presence / absence matrix, which we subsequently checked manually.

5. Interpretation of results

This section will be about how to estimate the "consistency" of the markers.

Each same-sized fragment from different electropherograms, are supposed to be homologous (e.g. have the same origin) but this condition is not fully filled in AFLP. An absence of fragment in one run can either be due to a mutation in the restriction site of the enzymes, a mutation in the primers sequence, or a mutation in the restriction fragment. All these processes leading to the same phenotype: an absence of peak. Nevertheless, the numerous markers produced by AFLP can compensate for eventual homoplasy.

According fits Kimuras infinite allele model (Kimura and Crow, 1964) mutations happen randomly throughout the genome and tend to form a new allele from an ancestral allele, rather than an allele already existing. This random mutation process can lead to an asymmetric distribution of alleles (relatively high occurrences at high and low frequencies) when one looks at the allele class frequency. Checking the allele class frequency of AFLP could thus be a way to estimate the quality (randomness) of the makers (Figure 9).
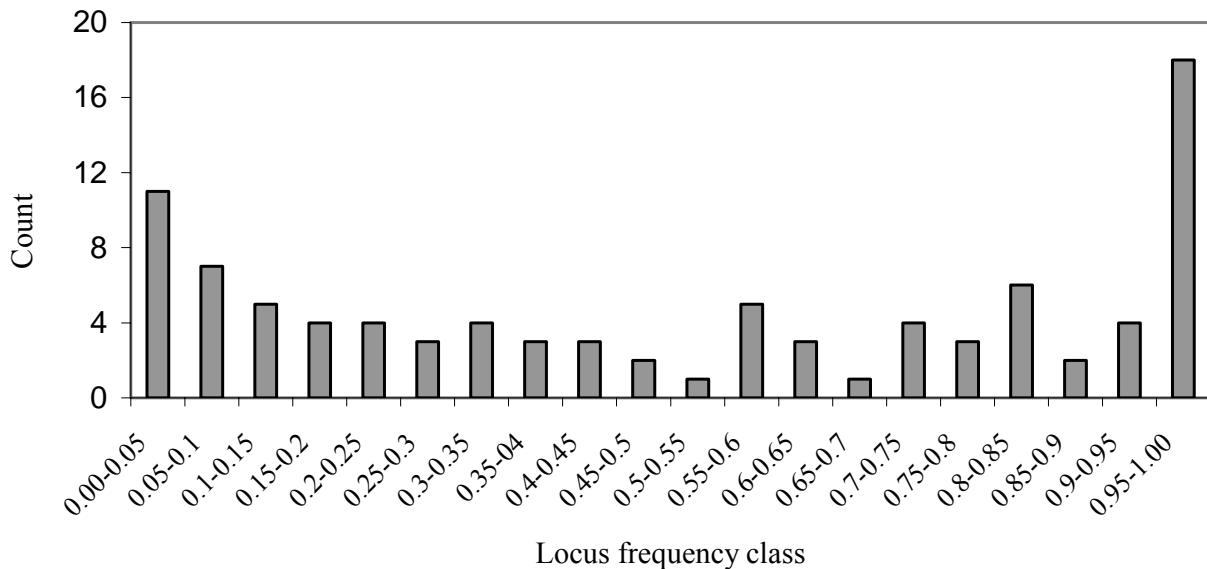

Figure 9: Locus frequency class distribution of 93 AFLP loci over 307 *C. arvense* individuals.

Therefore, quality of AFLP data was assessed by calculating the allele frequency class distribution for the 93 loci (42 for EcoRI-ACC / MseI-CTG and 51 for EcoRI-ACG / MseI-CTT) (Figure 9). The graph shows relatively high occurrences at high and low frequencies (common and rare bands). This asymmetric distribution fits Kimuras infinite allele model (Kimura and Crow, 1964) and leads to a bimodal distribution with common and rare alleles. This pattern of class frequency distribution of AFLP alleles has already been found in *Arabidopsis thaliana* (Miyashita et al., 1999; Sharbel et al., 2000) and attests the quality of our data.

Literature cited

Avise JC 1994 Molecular markers, natural history and evolution. Chapman & Hall, New York.

Bostock SJ, RA Benton 1979 The reproductive strategies of five perennial Compositae. Journal of Ecology 67:91-107.

Botstein D, RL White, M Skolnick, RW Davis 1980 Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32:314-331.

Brookfield JFY 1992 DNA Fingerprint in clonal organisms. Molecular Ecology 1:21-26.

Donald WW 1994 The biology of Canada Thistle (*Cirsium arvense*). Reviews of Weed Science 6:77-101.

Doyle JJ, JL Doyle 1988 Isolation of plant DNA from fresh tissue. Focus 12:13-15.

Erschadi S, G Haberer, M Schoniger, RA TorresRuiz 2000 Estimating genetic diversity of *Arabidopsis thaliana* ecotypes with amplified fragment length polymorphisms (AFLP). Theoretical and Applied Genetics 100:633-640.

Flachowsky H, E Schumann, WE Weber, A Peil 2001 Application of AFLP for the detection of sex-specific markers in hemp. Plant Breeding 120:305-309.

Godt MJ, JL Hamrick 1999 Population genetic analysis of *Elliottia racemosa* (Ericaceae), a rare Georgia shrub. Molecular Ecology 8:75-82.

Hamrick JL, MJ Godt. 1990 Allozyme diversity in plant species. Pages 43-63 *in* Brown AHD, MT Clegg, AL Kahler, and BS Weir eds. Plant population genetics, breeding and genetic resources. Sinauer Associates Inc., Sunderland.

Hongtrakul V, GM Huestis, SJ Knapp 1997 Amplified fragment lengh polymorphisms as a tool for DNA fingerprinting sunflower germplasm: genetic diversity among oilseed inbred lines. Theoretical and Applied Genetics 95:400-407.

Karp A, KJ Edwards. 1997 DNA markers: a global overview. Pages 1-13 *in* Caetano-Anollés G and PM Gresshoff eds. DNA markers: protocols, applications and overviews. Wiler-VCH, New York.

Kimura M, JF Crow 1964 The number of alleles that can be maintained in a finite population. Genetics 49:725-738.

kliebenstein DJ, J Gershenzon, T Mitchell-Olds 2001 Comparative quantitative trait loci mapping of aliphatic, indolic and benzylic glucosinolate production in *Arabidopsis thaliana* leaves and seeds. Genetics 159:359-370.

Lerceteau E, AE Szmidt 1999 Properties of AFLP markers in inheritance and genetic diversity studies of Pinus sylvestris L. Heredity 82:252-260.

Mace ES, CG Gebhardt, RN Lester 1999 AFLP analysis of genetic relationships in the tribe Datureae (Solanaceae). Theoretical and Applied Genetics 99:634-641.

Martinez EA, C Destombe, MC Quillet, M Valero 1999 Identification of random amplified polymorphic DNA (RAPD) markers highly linked to sex determination in the red alga *Gracilaria gracilis*. Molecular Ecology 8:1533-1538.

Miyashita NT, A Kawabe, H Innan 1999 DNA variation in the wild plant *Arabidopsis thaliana* revealed by amplified fragment length polymorphism analysis. Genetics 152:1723-1731.

Moore RJ 1975 The biology of Canadian weeds. 13. *Cirsium arvense* (L.) Scop. Canadian Journal of Plant Sciences 55:1033-1048.

Mueller UG, LL Wolfenbarger 1999 AFLP genotyping and fingerprinting. TREE 14:389-394.

Mullis KB, FA Faloona 1987 Specific sythesis of DNA *in vitro* via a polymerase-catalized chain reaction. Methods in Enzymology 155:335-339.

Mullis KF, FA Faloona, R Scharf, G Saiki, G Horn, H Erlich 1986 Specific enzymatic amplification of DNA in vitro: THe polymerase chain reaction. Cold Spring Harbor Symp.Quant.Biol. 51:263-

Ouborg NJ, Y Piquot, J Van Groenendael 1999 Population genetics, molecular markers and the study of dispresal in plants. Journal of Ecology 87:551-568.

Palacios C, S Kresovich, F Gonzáles-Candelas 1999 A population genetic study of the endangered plant species *Limonium dufourii* (Plumbaginaceae) based on amplified fragment length polymorphism (AFLP). Molecular Ecology 8:645-657.

Roa AC, MM Maya, MC Duque, J Tohme, AC Allem, MW Bonierbale 1997 AFLPs analysis of relationships among cassava other *Manihot* species. Theoretical and Applied Genetics 95:741-750.

Sharbel TF, B Haubold, T Mitchell-Olds 2000 Genetic isolation by distance in *Arabidopsis thaliana*: biogeography and postglacial colonization in Europe. Molecular Ecology 9:2109-2118.

Stansfield WD 1986 Theory and Problems of Genetics. McGraw-Hill Book Company, New York.

Tautz D, M Renz 1989 Simple sequences are ubiquitous components of eukaryotic genomes. Nucleic Acids Res 18:6531-6535.

Vos P, R Hogers, M Bleeker, M Reijans, T van de Lee, M Hornes, A Frijters, J Pot, J Peleman, M Kuiper, M Zabeau 1995 AFLP: a new technique for DNA fingerprinting. Nucleic Acid Research 23:4407-4414.

Vos P, M Kuiper. 1997 AFLP analysis. Pages 115-131 *in* Caetano-Anollés G and PM Gresshoff eds. DNA markers: protocols, applications and overviews. Wiley-VCH, New York.

Weising K, H Nybon, K Wolff, W Meyer 1995 DNA Fingerprinting in Plants and Fungi. Boca Raton.

Widen B, N Cronberg, M Widen 1994 Genotypic diversity, molecular markers and spatial distribution of genets in clonal plants, a literature survey. Folia Geobotanica Phytotaxonomica 29:245-263.

Williams JGK, AR Kubelik, JA Livak, JA Rafalski, SV Tingey 1990 DNA polymorphisms amplified by arbitraty primers are useful as genetic markers. Nucleic Acids Research 18:6531-6535.

Zabeau M, P Vos1993Selective Restriction Fragment Amplification: A general method for DNA fingerprints