Dissertation

submitted to the

Combined Faculties for the Natural Sciences and for Mathematics of the Ruperto-Carola University of Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

Presented by

Diplom-Biologin Charlotte Luise Scheriau

born in Neustadt an der Weinstraße, Germany

Oral examination: 13.12.2011

Evolutionary history of *Hypericum perforatum* L.

Referees: Prof. Dr. Marcus Koch

Dr. Timothy Sharbel

Table of contents

Abstract

The genus *Hypericum* is a highly diverse group of about 500 species. Among the Eurasian taxa, *Hypericum perforatum* L. and *H. maculatum* Crantz are representative examples with a largely overlapping distribution area in Central Europe. As a medicinal plant and as a model plant for apomixis *H. perforatum* is a target of extended research. An important step to understand those complex processes is to illuminate the evolutionary history of *H. perforatum*. We were able to identify putative ancestral, fully sexual and diploid populations of *H. perforatum*. The *H. perforatum* populations were split into two major gene pools. An ancient contribution of *H. maculatum* or any other species to these gene pools and to the early evolution of *H .perforatum* could not be confirmed. However, we were able to show ongoing introgression and massive gene flow between *H. perforatum* and *H. maculatum*.

The geographical patterns observed in our data showed that (i) haplotype and nucleotide diversities are very low in Northern Europe, suggesting that single haplotypes rapidly colonized large areas; (ii) genetic diversity in South Europe is higher than in North Europe, suggesting that there were refugial zones during the Pleistocene.

Zusammenfassung

Die Gattung *Hypericum* ist eine hoch diverse Gruppe von ungefähr 500 Arten. Unter den eurasichen Taxa sind *Hypericum perforatum* L. und *H. maculatum* Crantz Representative mit einer weiten Überlappungszone in Zentraleuropa. *H. perforatum* wird als medizinische Pflanze und als Modelpflanze für Apomixis untersucht. Ein wichtiger Schritt um diese untersuchten Prozesse zu verstehen ist die evolutionäre Geschichte von *H. perforatum* aufzuklären. Wir konnten sexuelle und diploide Populationen von *H. perforatum* als nächste Verwandte von potentiellen Vorfahren identifizieren. Populationen von *H. perforatum* waren in zwei Hauptgenpools unterteilt werden. Einen historischen Beitrag von *H. maculatum* oder einer anderen Art zu den Genpools oder der frühen Evolution von *H. perforatum* konnte nicht bestätigt werden. Allerdings konnten wir aktuell anhaltende Introgression und massiven Genfluss zwischen *H. perforatum* und *H. maculatum* zeigen. Aus den beobachteten geographischen Mustern konnten folgende Schlussfolgerungen gemacht werden: (i) Die sehr niedrige Haplotyp- und Nukleotiddiversität in Nordeuropa deuted auf wenige Haplotypen hin, die in kurzer weite Areale kolonisiert haben. (ii) Die genetische Diversität in Südeuropa ist höher als im Rest von Europa was auf Refugialgebiete während des Pleistozäns in diesen Regionen hindeutet.

1 Introduction

1.1 Polyploidy

Common bananas, and normal salmon that we eat are examples of polyploid species. Polyploidy is found to occur in about 70 % of all angiosperms (Masterson J., 1994), but why actually rather all angiosperms are polyploids (Soltis D. *et al*, 2009) will be explained in the following chapters.

1.1.1 Presupposed glossary

Genome — As elucidated by Greilhuber *et al.*(2005), includes whole chromosome complement or monoploid genome. Therefore I want to use the proposed terminology.
Holoploid genome: refers to the whole chromosome complement with chromosome number n irrespective of the degree of generative polyploidy.
Monoploid genome: refers to one set of chromosomes.

C- value — Size of the holoploid genome; DNA content of the whole complement of chromosomes characteristic for the organism, irrespective of the degree of generative polyploidy, aneuploidies, etc (Greilhuber *et al.,* 2005)

Cx-value — Size of the monoploid genome; (averaged) DNA content of monoploid genome(s) in polyploids and non-polyploids. In generatively non-polyploid organisms C-value and Cx-value are congruent (Greilhuber *et al.,* 2005).

'n' — Haploid number, refering to nuclear phase status (Greilhuber *et al.,* 2005).
n: reduced, haplophasic; Indicates the meiotically reduced chromosome number
2n: unreduced, diplophasic; Indicates the non-reduced chromosome number
3n, 5n, etc: Indicates the endospermic chromosome number (in angiosperms)

'x' — Degree of generative polyploidy. A diploid embryophyte taxon has as a sporophyte 2n=2x and a tetraploid 2n=4x (Greilhuber *et al.,* 2005).

1.1.2 What is ploidy?

Ploidy is the number of sets of chromosomes in a cell. However there is more explanation necessary to define this.
The DNA copy number status of a cell can be differentiated in four ways (Greilhuber *et al.*, 2005):

1. Depending on the phase of the mitotic nuclear cycle (G1, S and G2 phase) the DNA content changes. This is expressed in 'C'
   e.g. 1C-value of a triploid sporophyte (2n=3x) is half of the 2C-value of a cell in telophase.
2. Alternation of nuclear phases associated with meiotic reduction and fertilization
   e.g. somatic cells have a chromosome number of 2n=16 and the gametes have a chromosome number of n=8.
3. Generative ploidy levels. The presence of one, two or more monoploid genomes (expressed in 'x') in a reduced genome.
   e.g. somatic cells have a chromosome number of 2n=4x=32 and the gametes have a chromosome number of n=2x=16.
4. somatic polyploidy, caused by endocycles (more rarely by mitotic disturbances) in somatic tissue.

Individuals declared as polyploids have more then one monoploid genome in a reduced holoploid genome.

1.1.3 Types of polyploidy

Polyploids can be distinguished by their mode of formation, based on genetic and cytogenetic criteria, time since polyploidization event and degree of diploidization and.

1.1.3.1 Modes of formation (as reviewed in Soltis D., 2003)

Three forms of formation can be distinguished, but which one is the most prominent one is debated. Somatic doubling seems to be less common, eventhough many examples have been reported (Ramsey *et al*., 1998). Unreduced gametes have been reported in numerous species (Ramsey *et al*., 1998). In *Dactylis glomerata* (Bretagnolle F. & Lumaret R., 1995) and *Tragopodon* species (Ownbey, 1950) the polyploidization via unreduced gametes is proposed. Some (Bretagnolle F. & Thompson J., 1995) suggested that the one-step process of union of unreduced gametes is more common than often considered. In Harlan *et al.* 1975 the triploid bridge is considered a significant pathway of polyploid formation, which could be true, because triploids are not complete sterile (Ramsey *et al*., 1998), in fact in studies by Husband (2004) they detected varying levels of fertility in triploids.

1. **Polyploidy via somatic doubling:** Doubling of the chromosome complement at the zygotic, embryonic, or meristematic stage of a plant's life cycle leads ultimately to the production of polyploid tissue and possibly to polyploid offspring.
2. **production of unreduced gametes:** polyploids can arise in one step after unreduced gamete formation by union of two unreduced gametes.
3. **triploid bridge:** triploids are formed wihin a diploid population. Tetraploids are then formed via backcrossing with reduced gametes from diploids or via selfing of triploids that produce reduced gametes.

1.1.3.2 Autopolyploidy and Allopolyploidy (as reviewed in Soltis D., 2003)

Grant V. (1981) divided polyploids in autopolyploids and amphipolyploids. Where autopolyploids (AAAA) can be strict or from interracial polyploidization events. Amphipolyploids are subdivided in Segmental allopolyploids ($A_sA_sA_tA_t$), genomic alloploy ploids (AABB) and autoallopolyploids (AAAABB). Other researchers classify polyploids using taxonomic rank as a criterion (Lewis W., 1980). In this study we consider allopolyploids individuals formed between different species, whereas autopolyploids form within species, typically involving crossing between individuals and not somatic polyploidisation events.

Allopolyploids are characterized by fixed (i.e., nonsegregating) heterozygosity, resulting from the combination of divergent parental genomes. Autopolyploids may exhibit multivalent formation at meiosis and are characterized by polysomic inheritance (Soltis P. 2000). In permanent polyploid hybrids, most single loci would be expected to exhibit a combination of two parental genomes. Except after complete homogenization or partial homogenization.

1.1.3.3 Paleopolyploidy and Neopolyploidy

Genome doubling is now known to be widespread in most groups of organisms (Soltis D. *et al.*, 2003), also for nonplant lineages ancient genome-wide duplication events are evident (Spring J. 2002). Additionally have many studies in plants and animals shown that recurrent polyploidy is the rule and not the exception (e.g. Soltis D. *et al.* 1999).

A process called diploidization was already noted by Grant V. (1981). He wrote 'Old polyploids tend to be more diploid-like than newly formed polyploids'. Diploidization is realized via genomic rearrangements, genomic downsizing and gene silencing (reviewed in Soltis D., 2003).

Now that those processes are studied problems in terminology occur e.g. is *Zea mays* a diploid or polyploid? Due to extensive genome rearrangements and gene silencing it was hard to discover an ancient polyploidization event (White S. *et al*., 1998).
Paleo-, meso and neopolyploidy are terms suggested to clearify that problem. However, often these terms are used in different ways (reviewed in Guerra M., 2008).
Schubert I. & Lysak M. (2011) refer to **neopolyploidy** when polyploidization has occurred recently in evolutionary time and the corresponding multiplication of chromosome number is still recognizable. Polyploidization followed by diploidization including chromosome number reduction, but still detectable by using comparative genomics and chromosome painting is termed **mesopolyploidy.** Ancient polyploidization events, blurred by many subsequent alterations of the genome, which can only be revealed in extant diploid taxa by bioinformatic searches for orthologous and paralogous sequence markers is termed **palaeopolyploidy**.

1.2 Modes of reproduction

Bell (1982) commented that "Sex is the queen of problems in evolutionary biology...Perhaps no other natural phenomenon has aroused so much interest; certainly none has sowed as much confusion". With those words in mind a little introduction to sexuality and asexuality:

Perennials use a combination of outcrossing sex, selfing and asexuality to fine tune their reproductivity to a changing ecological environment (Richards A. 2003).

1.2.1 Outcrossing sexuality

In plants and animals sexual reproduction and recombination among genes is the prevailing mode of reproduction. For many researchers it is still debatable why it is prevalent.
There are many beneficials to sex:

- Fisher -Muller hypothesis: Sexual reproduction enhances the fixation probability of beneficial mutations (Fisher R. 1930; Muller 1932)
- Purging of deleterious mutations. (Kondrashov A. 1988)
- Resistance to predators and parasites (the Red Queen hypothesis). Sexual organisms produce diverse progeny some will stay ahead in an evolutionary arms raise between host and parasite (Hamilton W. 1980).

To all these beneficials of sex come of course also costs of sex (Lewis W. 1987):

- Recombination scrambles genotypes, disrupting also favorably adapted gene combinations
- Meiosis and Syngamy take longer than mitosis (important to small organisms)
- Cost of producing males; mating may be risky, sexual competition, wastage of gametes, cost of maintenance of sexual dimorphisms
- At low population density difficult to coordinate
- Sexual females suffer from genome dilution

Reviews about evolution of sex see (Lewis W. 1987, Otto S. et al. 2002 and D'Souza T. et al. 2010).

1.2.2 Asexuality

Asexual reproduction is found all across the 'tree of life' and can be found in most plant and animal groups (excluding mammals) and takes a variety of forms. The definitions given here refer mostly to plants and not to animals.

1.2.2.1 Presupposed glossary

Asexual reproduction	Production of a new ramet (= plant module) by mitotic division without meiosis or sexual fusion
Apomixis	asexual formation of seed, avoiding meiosis and fertilization, usage today is synonymous with 'agamospermous' (Richards A. 1997)
Recurrent apomixis	Synonymous to gametophytic apomixis / gametophytic agamospermy
Vegetative apomixis	Synonymous to vegetative propagation / vegetative reproduction; new individuals arise without seed production or spores, propagation via propagules (e.g. bulbils).
Adventive embryony	Synonymous to sporophytic apomixis / sporophytic agamospermy / nucellar embryony
B_{III}-hybrid	Fertilization of unreduced egg cell
Pseudogamy	Only endosperm is fertilized
Autonomous apomixis	Endosperm develops autonomous, without fertilization
Autogamy	Self fertilization. Pollen from the same individual as the egg is fertilizing the egg and endosperm.
Apogamy	Embryo from other cell then egg, but of the gametophyte, e.g. synergids or antipodal cells
Parthenogensis	Embryo arises from an unfertilized egg cell (agamic)
Polyembryony	Two (or more) embryos in one seed. Often recognized in aposporous agamospermy where sexual and asexual embryos lie together in one seed.

1.2.2.2 Agamospermy (reviewed in Richards A. 2003):

Two types of agamospermy are generally distinguished. Gametophytic agamospermy is better studied, but sporophytic agamospermy seems to be the most widespread form of agamospermy.

1. ***Gametophytic agamospermy***:
   - Apospory: Embryo sac forms in nucellus, but not archesporial. Archesporium remains free, so that sex is possible.
   - Diplospory: Embryo sac is of archesporial origin. Only asexual embryos possible.
2. ***Sporophytic agamospermy***:
   - embryos are budded directly from integuments or sporophytic nucellus tissue of the ovule.

Agamospermy has evolved many times within certain families (Poaceae, Rosaceae, Asteraceae). It is not distributed at random, suggesting that certain taxa in high latitudes are preadapted to agamospermy.

1.2.2.3 Nonrecurrent apomixis (Maheshwari P. 1950)

In nonrecurrent apomixis the megaspore mother cell is meiotically reduced and the embryo sac develops without fertilization. In this study the seeds produced via this process are refered to as parthenogenetic pseudogamous or parthenogenetic autonomous.

1.2.2.4 Vegetative reproduction

Plants with vegetative reproduction remain mostly facultative, so that they can also set sexual seeds. In special cases (*Elodea canadensis, Hottonia palustris, Holcus mollis*) also obligate vegetative apomixis is present (reviewed in Richards A. 1997)

1.2.3 Self fertilization

Selfing is frequently associated with polyploidy, which can cause breakdown of self-incompatibility (Barringer B. 2007).
In selfers 50 % of allozyme diversity is found among populations, whereas in outcrossers only 12 % lays among populations (Hamrick J. et al. 1996)

1.3 Biogeography and Evolution

In the history of earth continental drift and several ice ages changed not only the landscape but also the conditions for living on earth. Most prominent for us, due to its accessibility with todays research methods is the last glacial epoch (started in the Quaternary) with its last glacial maximum (LGM) approximately 20.000 years ago. The environmental changes must have been most drastic around the glaciated regions, but climatic fluctuations during the Pleistocene also had a strong impact on the environment in other regions, causing drastic sea level variation to open or close migration routes (Randi E. 2007), altitudinal shifts of vegetation belts (Parolly *et al.* 2010) or aridification processes (Rebernik *et al.,* 2010).
As a consequence species expierienced range shifts through contractions into refugia and subsequent re-colonization (Taberlet et al. 1998; Hewitt 1999).
In this changing environment many species were extinct and many could arise particularly due to those changes (e.g. rapid radiation in *Dianthus* (Valente *et al*., 2010) and *Draba* (Jordon-Thaden *et al*. 2010)). To study these past migration or radiation processes it is important to note that phylogeographic inferences can be made by comparing gene diversity and nucleotide diversity according to Avise (2000).
Migration, surviving, and eventually re-colonized processes were often linked with local adaptation (e.g. Manel *et al*. 2010).
Populations adapt to an environment in that mutations arise in individuals. Natural selection favours some of these mutant individuals, who become more successful than those that do not have these mutations, and thus the composition of the gene pool of the population changes over time.
An important evolutionary force is Polyploidy. It is probably the most common mechanism of sympatric speciation (Otto S. et al. 2000). Polyploids form at relatively high frequency (one per 100000) in flowering plants (Ramsey 1998)
Hybridization resulting in allopolyploids has also been shown to be wide spread and a frequent mode of plant evolution and speciation (Soltis D. et al. 1993, Wendel J. et al. 1998, Soltis P. 2000). These events can often be detected from the incongruence of species relationships reconstructed from molecular markers with differing modes of inheritance

1.4 Introduction to studied species

1.4.1 The genus *Hypericum*

Hypericum belongs to the family Hypericaceae together with eight other genera (Stevens P., 2007; APG III 2009). It is a genus of about 500 species of shrubs, herbs and also trees. Members of *Hypericum* occur in all temperate parts of the world with a centre of species richness in the temperate regions of the Northern Hemisphere. *Hypericum* is absent from habitats that are extremely dry, hot or cold, and rarely found in water.

Robson provided in 1977 a revision of the genus, and proposed a new classification, defining 30 sections. This publication was the first in a series of monographs of subgroups of *Hypericum* in which detailed information on characters for species descriptions are given (Robson 1981), as well as the formal taxonomy of sections and species. 36 sections have been to date described and 472 species have been recognized. The section of interest in this study is sect. 9. *Hypericum* (core *Hypericum*), which includes 42 species and is distributed in Europe, NW Africa, Asia, NW America and introduced via *H. perforatum* into many other parts of the world. Section 9 is divided into two subsections: 1. *Hypericum* (19 species) and 2. *Erecta* (23 species). Subsection *Hypericum* contains two series 1. *Hypericum* (12 species) and 2. *Senanensia* (seven species; see Table 1).

sect.9 *Hypericum* subsect.1. *Hypericum* series 1. *Hypericum*	**Species**	**subspecies**			
	1. *H. maculatum* Crantz	1a *H. mac.* subsp. *immaculatum* (Murb.) A. Fröhl.	1b *H. mac.* subsp. *maculatum*	1c *H. mac.* subsp. *obtusiusculum* (Tourlet) Hayek	
	***2. H. undulatum* Schousb. ex Willd.**				
	***3. H. tetrapterum* Fr.**				
	4. H. triquetrifolium Turra				
	5. *H. perforatum* L.	5a *H. perf.* subsp. *perforatum*	5b *H. perf.* subsp. *Songaricum* (Ledeb. Ex Rchb.) N.Robson	5c *H. perf.* subsp. *Veronense* (Schrank) H. Lindb.	5d *H. perf.* subsp. *chinense* N. Robson
	6. *H. attenuatum* Fisch. ex Choisy				
	7. *H. elegans* Stephan ex Willd.				
	8. *H. tosaense* Makino				
	9. *H. iwatelittorale* H.Koidz.				
	10. *H. momoseanum* Makino				
	11. *H.yezoënse* Maxim.				
	12. *H. scouleri* Hook				
sect.9 *Hypericum* subsect. 1. *Hypericum* series 2. *Senanensia*	**13. *H. pibairense* (Miyabe & Y.Kimura) N. Robson**				
	14. *H. kamtschaticum* Ledeb.				
	15. *H. nakaii* H.Koidz.	15a *H. nak.* subsp. *nakaii*	15b *H. nak.* subsp. *miyabei* (Y.Kimura) N.Robson	15c *H. nak.* subsp. *tatewakii* (S.Watan.) N.Robson	
	16. *H. senanense* Maxim.	16a *H. sen.* subsp. *senanense*	16b *H. sen.* subsp. *mutiloides* (R.Keller) N.Robson		
	17. *H. faberi* R.Keller				
	18. *H. oliganthum* Franch. & Sav.				
	19. *H. sikokumontanum* Makino				
sect.9 *Hypericum* subsect. 2. *Erecta*	**31. *H. erectum* Thunb.**				
	35. *H. gracillimum* Koidz.				
	38. *H. hakonense* Franch. & Sav.				

Table 1 ***Species and subspecies included in Section 9 Hypericum***. In subsection 2 *Erecta* only two species shown, that were analysed in this study. Species analysed in this study in bold.

Species of the genus can be typically recognized by their leaves (opposite, simple and entire, lacking stipules), yellow flowers with petals free and several stamens in 3 or 5 fascicles, styles free, and the presence of pale and sometimes reddish to black glandular secretions (glands). The fruit is, in general, a dehiscent capsule, containing small cylindrical light brown to black seeds. For further and more comprehensive information about characters and descriptions of the genus, refer to Robson (1981) and Stevens (2007).
Within *Hypericum* karyology is quite diverse. Basic haplophasic chromosome numbers of n= 6 to 10 and 12 have been reported (Robson N. 1981; Kogi M. 1984). Robson (1981) suggested that n= 12 is the ancestral chromosome number within *Hypericum*. Aberrations of these numbers are frequent, as polyploidization is a common phenomenon in the genus (Robson N. 1981). Tetraploidy is reported on the base of n = 6 and n = 8–10. Higher degrees of polyploidy are confined to the sect. *Hypericum*, and are associated with the apomictic *H. perforatum* aggregate. Within *Hypericum* 16 apomictic species have been described, only one (*H. scabrum* from the Eastern Mediterranean to West Asia) is an obligate apomict (Matzk et al. 2003).
A complex mixture of bioactive secondary metabolites in several *Hypericum* species makes them valuable as herbal drugs (Crockett S. *et al*., 2005 and Mártonfi P. *et al*., 2006). *H. perforatum* L. (common St. John's wort) is certainly the best-known and most investigated species of the genus (Nahrstedt A. *et al.,* 2010).

1.4.2 *Hypericum perforatum* L.

In the following, an overview on the biology of *H. perforatum* is given, focusing on selected morphological characters and anatomical features, on floral development, cytology and on reproductive system.

1.4.2.1 Habit

H. perforatum is a perennial herb of 0.2 up to 1 m height. It is erect or rarely decumbent to procumbent from creeping and rooting base, which allows the plant to vegetatively propagate. Characters for each subspecies are summerized in Suppl.Mat. 1.
Characteristic for *H. perforatum* are 2–lines raised along each internode, often visible only in a young stadium of the plant. The stems are numerous to few and much branched with few black glands.
Leaves are sessile to petiolate with a lamina of 10-25 x 3-10 mm. Laminar glands are mainly pale and scatterd, but sometimes also black. Few glands on leaves are punctiform. Intramarginal glands are black and spaced, interspersed with small dense pale glands.

1.4.2.2 Habitat

H. perforatum has a wide ecological amplitude and can tolerate severe climatic conditions at high altitudes and latitudes (Robson N. 2002).
H. perforatum has practically no preferences in environment. It is found in open woodland, meadows, grassland and steppes, riverbanks, stony and grassy slopes, roadsides, in dry or well-drained habitats from 10 to 3150 m.s.l altitude.

1.4.2.3 Glands

The distribution and appearance of glands and canals (i. e. as dots, streaks or lines on leaves, stem or in flower organs) have been used in the subgeneric classification of the genus *Hypericum* (Robson N. 1977) and have been anatomically studied since the 19th century (Green J. 1884).
Pale glands are ubiquitous in the whole genus *Hypericum* at least on leaves, making this organ often look perforated. They are the site of hyperforin accumulation (Soelberg et al. 2007). It

has been shown, that the red pigment hypericin is mainly located in these dark glands, with the highest concentration in flower tissues, especially in stamens (Zobayed et al. 2006). They might be elongated as tubular nodules (Curtis J. et al. 1990) and may also be described as 'streaks' or 'lines' (Robson 1981).
Secretory canals of different types have been reported for *H. perforatum*. Ciccarelli and colleagues (2001) have counted three types of translucent canals. Type A canals are associated with veins. Type B canals resemble anatomically and ontogenetically elongated pale glands, and type C canals are located on the ovarya and enlarge in fruit, described as amber 'vitae' or 'vesicles' by Robson N. (1981).
Secretory structures in *H. perforatum* have been suggested to be involved in the plant's response to herbivore attack (Sirvent et al. 2003).

1.4.2.4 Infloresence

The bisexual flower of *H. perforatum* is stellate, 15-25 mm in diameter. The buds are narrowly ovoid.
Flowers of the genus *Hypericum* are generally nectarless. They are typical 'pollenflowers' visited by less-specialized insects, of which the Syrphidae (Diptera) are the most common to *Hypericum* (Robson 1981: 119 ff.). The yellow (flavonoid) colored petals of *H. perforatum* are characteristic for the whole genus *Hypericum* and are used to define borders of the genus (Robson 1977: 301).

1.4.2.4.1 Androecia

The specific arrangement of stamens in bundles, in the so-called stamen fascicles, is a characteristic feature for the whole genus *Hypericum* (Robson 1977, 1981). According to Leins P. et al (2008) the polyadelphus androecia of *Hypericum* are antepetalous and of centrifugal development. The arrangement in bundles results from a primordial burgeon, on which several primordial stamens form, until the burgeon is fully occupied. Connate filaments of a fascicle are the result from the primordial burgeon that is developing (growing) together with the stamens (Leins P. et al. 2008).
In several species of *Hypericum*, the number of fascicles appears to be smaller than that of the petals. This phenomenon results from the merging of adjacent pairs of fascicles. The stamens of *H. perforatum* are 3-fascicled and 6-8 mm long.

1.4.2.4.2 Gynoecium

The ovary in *H. perforatum* is superior, 3-locular and 3-5 x 1.3-1.8 mm. It is surmounted by 3 elongate and free styles, terminated by narrow stigmas.
The fruit is a septicidal capsule with a size of 0.7- 1.5 times the sepals or 3-10 x 3-6 mm. Glands, described as 'vittae' or 'vesicles', are present on the pericarp (Robson N. 2002) and used as an important character to discriminate species in *Hypericum*. In *H. perforatum* are valves with dorsal vitae and lateral vittae or yellowish, striiform to punctiform vesicles discribed (Robson N. 2002). Seeds are dark brown and ca. 1 mm long.

1.4.2.5 Chromosome numbers

Tetraploidy with a haplophasic chromosome number n=2x=16 was recognized by Robson N. et al. 1968 and also by Brutovská et al. 2000b.
H. perforatum is according to Campbell *et al.* (1984) and Robson N. (2002) of allopolyploid origin. Based on morphological studies and geographical distribution *H. maculatum* subsp. *immaculatum* and *H. attenuatum* are refered to as potential parents.
Cytological results (Brutovská et al. 2000a) already point instead to an autopolyploid origin for *H. perforatum*. In wild populations the tetraploid cytotype (2n = 4x = 32) occurs most

frequently. Although both diploid (2n = 2x = 16) and hexaploid (2n = 6x = 48) individuals can be found (Robson 1981; Matzk et al. 2001). Moreover, individuals of all three ploidy forms hybridize with the diplo- and tetraploid subspecies of *H. maculatum*, which results refering to Robson N. (1981, 2002) in the morphologically extremely plastic *Hypericum* x *desetangii* complex.
Our study here will shed some light in the distribution of cytotypes, origin of polyploidy and the morphological plasticity in *H. perforatum*.

1.4.2.6 Reproduction biology

H. perforatum may be self-incompatible, due to the fact that self-incompatiblity is wide spread in the genus *Hypericum* (Robson N. 1977; 1981), but little information is available for *H. perforatum.*
Apomixis was first described for *H. perforatum* by Noack (1939). Since Matzk and colleagues (2001) determined eleven different routes of reproduction in *H. perforatum*, the species has become a model plant for apomixis research (Matzk et al. 2003; Barcaccia et al. 2006; Qu et al. 2010; Schallau et al. 2010).
It has been shown that tetraploid *H. perforatum* is a facultative apomict with the gametophytic mode of agamopermy. It includes apospory and the pseudogamous mode of endosperm development (for review see Matzk *et al.* 2001, Barcaccia et al 2007 and Schallau et al. 2010). The evolution of apomixis is not yet in detail understood.
Matzk *et al.* (2001) suggested a dosage effect for the trait apomixis, due to his findings of diploid sexual progeny from tetraploid apomictic parents.
Pseudogamy selects for self-fertility, so that self-fertility may also have been a preadaptive for evolution of agamospermy (Noirot *et al.* 1997), but others find that apomixis is associated with expression of mechanisms that limit self-fertilization (Bicknell et al. 2004).
10 % of animals and 25 % of plants are known to hybridize with at least one species (Mallet J., 2005). In the genus *Hypericum* natural hybridization is relatively rare (Robson N., 1981), except in series 1. *Hypericum*, were it is quite common due to the biology of *H. perforatum*. A number of morphological hybrids are described in Robson N. (2002). A summary of morphological characters is given in Suppl.Mat. 1.
Noak (1939) performed varified that crosses between *H. perforatum* (2n=4x=32) and *H. maculatum* (2n=2x=16) are possible. Triploid hybrids (2n=3x= 24) were described when *H. perforatum* was the father. They were intermediate habit o their parents, but rather hard to distinguish. Pentaploid hybrids (2n=5x=40) were detected, when *H. perforatum* was the mother. These individuals could not be distinguished morphologically from *H. perforatum*.

2 Aims of this doctoral thesis

We aim to unravel the evolutionary history of *H. perforatum* and *H. maculatum* during the pleistocene and to characterize genetical contact zones reflecting actual and past gene flow between both. Does the evolutionary scenario of those wide-spread lowland species confirm patterns of refuge areas and principle migration routes revealed from the analysis of woody plants, annuals, arctic-alpine representatives and others?
First we test the main alternative hypotheses on the origin of *H. perforatum*: Auto- versus allopolyploidization.
Second, we want to answer the question if the morphologically highly variable *H. perforatum* had a polytopic/polyphyletic origin.
Third, we want to give a data evaluation of the morphology and cytology of *H. perforatum* and *H. maculatum* (incl. *H. attenuatum*).

3 Material and Methods

3.1 Plant material

For *H. perforatum* and *H. maculatum* a total of 112 populations plus additional 155 accessions with single individuals (669 individuals in total) were analysed.
For plastid and nuclear DNA analysis (outgroups) another seven accessions were investigated (*H. tetrapterum, H. attenuatum, H. pibairense, H. erectum, H. kamtchaticum, H. gracillimum*, and *H. oliganthum*).
Herbarium material was acquired from Herbarium München (MSB), Staatsherbarium München (M) and Herbarium Heidelberg (HEID). Additionally herbarium material was provided by the Hungarian Natural History Museum and the Natural History Museum London. Further leave material was collected in the wild by several botanists (M. Koch, C. Bräuchler, C. Dobes, J. Paule, Y. Zhao, C. Scheriau) and from plants grown from seeds which derived from several botanical gardens in Europe. Furthermore we made use of *Hypericum* DNA samples from T. Sharbel and F. Blattner (IPK, Gatersleben).

3.2 DNA isolation, amplification and sequencing

Total genomic DNA was obtained from 50-75 mg dried leaf tissue from single individuals by the cetyltrimethyl ammonium bromide (CTAB) method as applied in Doyle J. J. and Doyle J. L. (1987), with minor changes. Leaf tissue was grinded in a Precellys 24 homogenizator (Bertin technologies). The DNA pellet was washed twice with 70 % ethanol and dissolved in 50 µl TE buffer. For each reaction 2U of RNase A were added. Samples were then incubated at 37 °C for 1 h. The concentration of the samples was measured using the NanoDrop ND-1000 spectrophotometer.
In general Polymerase chain reactions (PCRs) included 5x GoTaq Reaction buffer (containing 1.5 mM MgCl), 5 pmol of each primer, 5 nmol dNTPs (1.25 nmol each), 0.5 U GoTaq DNA polymerase and 1 µl undiluted template DNA (1-150 ng) in 25 µl reaction. Specific DNA was amplified using a PTC 200 Peltier Thermal Cycler (MJ Research) under the following conditions: 2 minutes initial DNA denaturation at 95 °C, 30 cycles of DNA amplification with 20 seconds at 95 °C, 45 seconds at 61 °C and 1 minute at 72 °C, and 10 seconds of final elongation at 72 °C.
PCR products were checked for length and quality by electrophoresis on 1 % agarose gels in TAE-buffer and stained with GelRed (Biotium Inc., Hayward, CA). Before sequencing, the DNA was purified by NucleoFast Kit (Macherey-Nagel, Germany).
Sequencing was carried out by the commercial sequencing service GATC (Konstanz, Germany) using M13 sequencing primer pair (forward -5´-GCATGTTTTCCCAGTCACGAC-3´, reverse -5´-ACTTCAGGAAACAGCTATGAC-3´). Cycle sequencing was performed on both strands.
Sequences were checked and trimmed using DNASTAR Lasergene (GATC Biotech, Konstanz, Germany).

3.3 DNA sequence variation

3.3.1 Plastid DNA sequence variation

After an initial screening of various plastid regions (trnL intron, trnL-F intergenic spacer, trnF intron, trnH intron, trnH-psbA intergenic spacer, psbA intron, trnC intron, trnC-ycf6 intergenic spacer, ycf6 intron, psbA intron, psbA-matK intergenic spacer, matK intron, rps16 intron, trnT intron, trnT-trnL intergenic spacer, trnS intron, trnS-psbZ intergenic spacer, psbZ intron, psbZ-trnG intergenic spacer, trnG intron, rpoC1 intron) three DNA regions were chosen providing sufficient DNA sequence variation: psbB-psbH (psbB, psbT, psbT-psbN

intergenic spacer, psbN, psbN-psbH intergenic spacer), trnL intron, trnS-ycf9 (trnS, trnS-psbZ intergenic spacer, psbZ, ycf9).
For DNA amplification the following primer pairs were used (the extending M13 sequencing primer sequence is indicated in small letters): psbB-psbH: 5´-gcatgttttcccagtcacgacTCCAAAAANKKGGAGATCCAAC-3´ (forward psbB) - 5´-acttcaggaaacagctatgacTCAAYRGTYTGTGTAGCCAT-3´ (reverse psbH), trnL intron: 5´-gcatgttttcccagtcacgacCGAAATCGGTAGACGCTACG-3´ (forward trnL-L) - 5´-acttcaggaaacagctatgacGGGGATAGAGGGACTTGAAC-3´ (reverse trnL-L), 5´-gcatgttttcccagtcacgacGAGAGAGAGGGATTCGAACC-3´, trnS: (forward trnS uga)- 5´-acttcaggaaacagctatgacCAAAMACAGCCAATTGGAAAGC-3´ (reverse ycf9P) (Taberlet *et al.* 1991, Shaw et al. 2005, Heinze 2007, with minor modifications).
Due to bad quality of some sequences the three markers were separately cloned. Representative PCR products were purified with the Wizard SV Gel and PCR Clean-up System (Promega) and ligated into a pGEM-T Easy plasmid vector (Promega). The plasmid was than transformed into α-select competent cells of E.coli. Approximately five colonies of each cloning reaction were selected and further grown in LB medium containing 100ug/ml ampicillin. Plasmids were purified with the NucleoSpin Plasmid Kit (Macherey-Nagel) and digested with SacII and NdeI. Plasmids were then sequenced from both directions.
Clones of identical or similar type were grouped and their consensus sequence was analysed for polymorphic sites.

3.3.2 Nuclear DNA sequence variation

For DNA amplification the following primer pairs were used (extending M13 sequencing primer sequence is indicated in small letters): ITS (Internal Transcribed Spacer 1 and 2): 5´-gcatgttttcccagtcacgacGGAAGGAGAAGTCGTAACAAG-3´ (forward ITS18F) 5´-acttcaggaaacagctatgacGGGTAATCCCGCCTGACCTGG-3´ (reverse ITS25R) (originally designed by White *et al.* (1990) with minor modifications by Mummenhoff *et al.* (1997)).
Due to polymorphic nucleotide positions in direct sequences of purified PCR products further analysis was necessary.
To identify the two differing genotypes two values were calculated to measure the percentage to which they explain the profile of the mixed sequence in each individual. The first value ‘A’ gives the percentage of ambiguous sites in the sequence that can be explained by a mixture of two genotypes. Value ‘B’ gives a percentage of how many of the sites that are polymorphic between *H. perforatum* and *H. maculatum* in the alignment (29 in total) can be explained by either one of the genotypes or both. The combination of the two geotypes which gained the highest percentage number was then assigned to the individual. 30 individuals could not be assigned to any genotype. As a consequence they were excluded from the analyses.
One genotype was only found in hybrid sequences and was therefore excluded from the alignment, but the genotype could be assigned to one of the major groups (red gene pool) in the net work. The individuals carrying it could, therefore be included in further analyses as group members.
In cooparation with N. Nürk (IPK) we could also analyse several cloned individuals to varify the genotypes found in hybridogen sequences.

3.4 Plastid-nuclear data and network analysis

For DNA sequence based analysis several outgroup taxa have been included (*Hypericum tetrapterum, H. attenuatum, H. pibairense, H. erectum, H. kamtchaticum, H. gracillimum*, and *H. oliganthum*). These outgroup taxa are a representative set of the closest relatives of *H. perforatum* and *H. maculatum* based on DNA sequence data (Nürk and Blattner, in prep.).
Sequences were edited and assembled with DNASTAR Lasergene (GATC Biotech, Konstanz, Germany). Sequences were further assembled and aligned by ClustalW (Labarga A. *et al.*

2007) with later manual adjustments using PhyDE 0.9971 (Müller *et al.* 2005). Network analyses were performed with TCS v.1.21 (Clement *et al.*, 2000), which is using parsimony (Templeton *et al.*, 1992) as underlying method. Gaps were coded separately and added to the final data matrix. Length variation in poly-A or poly-T regions was excluded from the analyses. The acceptance limit was set to 95 % (Clement *et al.*, 2000).
To estimate the genetic differentiation in *H. perforatum*, individuals were grouped according to their geographical distribution. Europe was divided in six regions as follows: ES: Iberian Peninsula, IT: Apennine Peninsula, ATHU: unglaciated part of Austria and Hungary, Balk: Balkan Peninsula, Nglac: glaciated regions of Northern Europe and Nunglac: unglaciated regions of Northern Europe (glaciation was checked according to map material from Haak Weltatlas for google earth 6.0.2).
Several calculations were made with DNASP ver. 5.10.01 (Rozas, J. *et al.*, Universitat de Barcelona, 2010) and Arlequin ver. 3.5.1.2 (Excoffier L., CMPG, University of Berne, 2009).
For georeferencing Google Earth 6.0.2 was used and mapping was done with ESRI ArcView 9.

3.5 AFLP analysis

The AFLP data were acquired and analysed in the context of a diploma thesis by Anja Landau (2009).

3.5.1 Preparation of AFLPs

Total genomic DNA was extracted from 50-75 mg dried leaf tissue following the procedure of Doyle & Doyle (1987, 1991) with minor modifications (Gong & al., 2008). Leaf tissue was ground in a Precellys 24 (Bertin technologies) homogenizator. The DNA pellet was washed twice with 70 % ethanol and dissolved in 50 µl TE-buffer. Per reaction 2U of RNase A were added and incubated at 37 °C for 1h. The concentrations of the samples were measured using the NanoDrop ND-1000 spectrophotometer, and each sample was diluted with ddH2O to a final DNA concentration of 100 ng/µl.
AFLP analysis was performed according to Vos *et al*. (1995) and modified by Gong & al. (2008). The analysis was carried out using EcoRI-A and MseI-C as pre-selective primers and three selective primer combinations: EcoRI-AAC (FAM)/MseI-CAC, EcoRI-ACG (TET)/MseI-CAC, EcoRI-AGC (HEX)/MseI-CAA.
The three differently fluorescence-labeled primer pairs were multiplexed (2 µl TET, 2 µl FAM, 6 µl HEX) and diluted 30 times with ddH2O, of which 1 µl was mixed with 0.2 µl ET-ROX 550 size standard and 5.8µl ddH2O. After 2 min denaturation at 95 °C samples were run on a MegaBase 500 automated sequencer (Amersham Biosiences).
Raw data was scored and exported as a presence/absence matrix using Genemarker v1.6 (SoftGenetics LLC).
In order to maximize the reliability of the data set the scored fragment size was restricted from 60 to 350 nucleotides.
Each experiment was running with 48 samples on the 48-capillary DNA sequencer, one general standard probe was placed twice (to check for consistency within a single run and among different runs) and an additional negative control.

3.5.2 AFLP-Data analysis

The error rate of AFLP genotyping was calculated following the procedure described by Bonin *et al.* (2004). Fragments occurring with a frequency lower than 1.5 % were excluded from the dataset.
The AFLP data was analysed and visualized with SPSS 19.
As an alternative approach, the genetic structure of the data set was examined by genetic admixture analysis using the program STRUCTURE vers. 2.3.1 (Pritchard & al., 2000; Falush

et al. 2007). The data was analysed with K ranging from 2 to 10, with 10 replicate runs for each K, and a burn-in period of 1×10^5 and 1×10^6 iterations. The "admixture model" and uncorrelated allele frequencies were chosen for the analysis. The likelihood of Ks ranging from 1 to 10 was calculated using the R-script Structure-sum (Ehrich, 2006). R-script Structure-sum allows the comparison of the likelihood of the runs (Rosenberg *et al.* 2002), the similarity coefficient between the runs and the delta K as defined by Evanno *et al.* (2005). K was assumed as optimal chosing the model with the highest probability of the data. In the visualization of Evanno´s delta K a peak has to match the optimal fitting model and the results over multiple runs have to be consistent.
For a detailed description of the used parameters, refer to Ehrich et al. (2007) and Evanno & al. (2005).
AMOVA (Analysis of molecular variance) was performed with Arlequin ver. 3.5.1.2 (Excoffier L., CMPG, University of Berne, 2009) to deduce the level of molecular variation among and within different groups.
The input file for Arlequin was created with AFLPdat (Ehrich, 2006, available under http://www.nhm.uio.no/forskning-samlinger/forskning/forskningsgrupper/ncb/Online_publications), embedded in the software R 1.9.0 (R Development Core Team, 2006), available under http://www.rproject.org). Several statistical statistical calculations concerning the genetic diversity at population level were performed with the script AFLPdat in R:

- The function 'Diversity' calculated Nei's gene diversity: D=n/(n-1)*(1-(*freq*(1)^2+*freq*(0)^2)) whereas n is the number of individuals in the population and freq1 and freq0 indicate the presence or absence of a marker in the population. D was first calculated for each marker in a population afterwards the average was calculated; the function of AFLPdat also counted the proportion of polymorphic markers for each population.
- The frequency down-weighted marker index DW according to Schönswetter and Tribsch (2005) was estimated using the function 'DW' whereas for each population the occurrence of adequate AFLP marker in that population was divided by the number of occurrences of that marker in the complete data set. DW is expected to be high for isolated populations with many rare markers and low for newly established populations, where rare markers have not yet accumulated through mutations. DW was only calculated for populations with three or more individuals.

3.6 Cytogenetic analysis

Part of the cytogenetic data used in this thesis was acquired in the context of a bachlor thesis by Nora Hohmann (2010).

3.6.1 Karyotyping

Ploidy level estimates are either based on flow cytometric genome size estimates or direct counts of chromosome numbers from root tips.
For karyotyping root tips were collected from plants cultivated in the Heidelberg Botanical Garden and immediately put in 2mM 8-hydroxychinolin at approximately 15 °C. After 2-3h of incubation the root tips were transferred to ethanol–glacial acetic acid (3:1) fixative (Carnoy fixative). For long term storage they were transferred to 70 % ethanol after at least 1h incubation in Carnoy fixative. In preparation for cell spread assays the root tips were washed in water and then enzymatically digested (0.1 % cellulase(in citrate buffer), 0.1 % cytohelicase and 0.1 % pectohelicase) for 90min at 37 °C. The cell suspension was than mixed on a slide with one drop of 45 % acetic acid for 1-2min at 45 °C on a hot plate. For fixation a ring of Carnoy fixative was dropped on the slide and left to dry on the hot plate. Objects were finally stained with Vectashield containing DAPI solution (Axxora, Germany)

and sealed with clear nail polish for permanent use. Chromosome numbers were determined and documented as digital images.

3.6.2 Flow cytometry

3.6.2.1 Ploidy level estimation

For ploidy level estimates, based on flow cytometry, young and living leaves were used. The majority of samples was measured via a high-through-put method established by the group of T. Sharbel at the IPK. A small piece of leave tissue was added to a stainless steel ball (3.22mm) in each well of a 96-well sample plate. After adding 400µl OttoI buffer the plates were shaken for 15sec with 600 strokes per second. After 10 min incubation on ice and in the dark the suspensions were filtered using 25µm Partec filters. 40µl of the filtered suspension was then mixed with 200µl of OttoII buffer and incubated on ice and in the dark for 5 min. In each plate two samples with known ploidy level were measured as external standard.
Analyses were performed with a CyFlow Space (Partec, Münster, Germany) flow cytometer and data evaluation was carried out using the FloMax analysis program (Partec, Münster, Germany).
To this previous data ploidy information of several populations was added in collaboration with Marta Molins (2011).

3.6.2.2 Reproductive mode analysis

The reproductive mode was analysed with an extensive seed screen using flow cytometry and following the concept presented first by Matzk *et al.* (2000, 2001, 2003) as Flow Cytometric Seed Screen (FCSS). Following this analysis individual-based seed material was analysed for ploidy level variation of the endosperm compared to the embryo and referring to the known ploidy level status of the original mother plant. For any of the various accessions 24 randomly selected seeds were chosen, cut and incubated individually with 0.5 ml extraction buffer (Partec CyStain PI kit, Partec, Münster) and filtered. The samples were then stained by adding 2 ml of the Partec CyStain PI solution containing DNase-free RNase following the instructions of the manufacturer (Partec, Münster, Germany) and incubated for 4 hours at 4 °C in darkness. The diploid material of accession Hyp1410 was used as external standard and reference to calculate relative C-values. Analyses were performed with a CyFlow Space (Partec, Münster, Germany) flow cytometer (green laser with 532 nm), and data evaluation was carried out using the FloMax analysis program (Partec, Münster, Germany).
In a normally reproducing sexual plant the ploidy level ratio (genome size ratio) of embryo versus endosperm can be expected as 2:3 (diploid versus triploid in a diploid individual, or tetraploid versus hexaploid in a tetraploid plant). Any deviation from these expectations can be interpreted with some contribution as apomictic or non-sexual reproduction (for details refer to Matzk et al, 2001).

3.6.2.3 Genome size estimation

For all cultivated accessions genome size estimates were performed using young and fresh leaf material to determine the ploidy level independently. Small amounts of leaf tissue of a sample and internal standards were co-chopped in 0.5 ml of extraction buffer (Partec CyStain PI kit, Partec, Münster) and prepared as described in chapter 3.6.2.2..
Solanum lycopersicum cv. Stupicke (1.48 pg/1C; Doležel *et al.* 1992) and Zea mays CE-777 (2.71 pg/1C; Lysak and Doležel 1998) were used as primary standards, and the following *Hypericum* accessions were chosen as external standands: diploid Hyp1410 and tetraploid Hyp1517. From the databases we expected a value for *H. perforatum* of 0.325 pg/1Cx since Bennet and Leitch (2010) provided a value of 0.65 pg/1C (most likely they made use of a tetraploid individual).

3.7 Morphological analyses

In the context of a diploma thesis by Xiaorui Sun (2010) a dataset was established to define *H. perforatum* and *H. maculatum* by morphological characters.
Furthermore, the use of independent morphological characters precluded the inclusion of various taxonomical concepts describing not only different subspecies for both taxa, but also providing names for different putative hybrids between *H. maculatum* and *H. perforatum* (e.g broadly summarized in past literature as *H.* X *desetangsii*).
For *H. perforatum* and *H. maculatum* a total of 125 populations plus 25 additional accessions with single individuals (450 individuals in total) were analysed.
Furthermore, for all individuals collected in the wild for which seed material was available from the original mother plants, offspring was grown in a common garden experiment at the Botanical Garden in Heidelberg (257 individuals in total) were also analysed. This material was used to test the principal discriminatory power of the selected morphological characters, in addition to character variability (or lack of) under different environmental conditions (original habitat versus common garden experiment) to thus confirm the general reliability of the original morphomatrix.
All morphological analyses were performed with dried herbarium vouchers deposited at the herbarium at Heidelberg Botanical Garden (HEID) and at Munich Staatssammlung (M).
The various measurements were scored by hand using a high-power binocular microscope to measure dimensions on digital live-images in high resolution. Coding of discrete character states does not imply any evolutionary progression (Möller *et al.*, 2007).
The removal of some plant organs (e.g. flowers and flower buds) from the original voucher was necessary for appropriate analysis. Here 100 µl of a detergent was added to the removed plant organ, and then covered with boiling water (total volume 10 ml) for 2 minutes. The organs were then dissected onto a glass slide, and after removing the water the samples were fixed and mounted on blotting paper and finally dried for two days at 40 °C.
In total 28 characters (continuous, discrete, multistate) were scored from leaves, stem, flowers and capsules (see extra material 5 on CD). From these characters a final matrix of 15 suitable variables available for all individuals was extracted (8 quantitative characters of which five are based on ratios, 4 multistate characters, and 3 two-state characters, see extra material 6 on CD).

3.7.1 principal component analysis

A principal component analysis (PCA) was run to analyse the multivariate data using single individuals as OTU (operational taxonomic unit) using SPSS (version 19 for Windows). SPSS was also used to correlate other variables such as ploidy level, growing conditions (wild origin versus common garden experiment), genetic differentiation and taxon designation, with the morphometric data. PCA requires that continuous character states are normally distributed. Our tests demonstrated that no character showed significant skewness. Therefore data transformation was not necessary.

3.7.2 Canonical discriminant analysis

In order to derive a morphology-based index of intermediacy between *H. maculatum* and *H. perforatum* for single individuals, we conducted multiple discriminate analysis (MDA, also termed canonical discriminant analysis, CDA). For this analysis eight individuals had to be excluded, due to their genetic connection to H. tetrapterum. Diploid and genetically not introgressed (see AFLP and cpDNA data) individuals from *H. perforatum* and *H. maculatum* served as reference groups indicative for the most divergent and putatively parental morphotypes.
The intermediacy index for each individual genetically analysed was calculated from the two dimensional MDA bar blot, placing the value of the individual to be tested on a re-scaled

distance (0 to 1) between the mean values of diploid *H. perforatum* (left border) and *H. maculatum* (right border) reference individuals. Consequently, the intermediacy index varies between 0 (pure *H. perforatum*) and 1 (pure *H. maculatum*).
In order to derive a morphology-based differentiation between *H. maculatum* and *H. perforatum*, we conducted a CDA for the whole data set of 442 individuals.
In the course of the CDA the program calculates a probability suggesting the placement of each individual to a certain group. For this analysis the same eight individuals had to be excluded, due to their genetic connection to *H. tetrapterum*. Diploid and genetically not introgressed (see AFLP and cpDNA data) individuals from *H. perforatum* and *H. maculatum* served as reference groups (group1 and group2) indicative for the most divergent and putatively parental morphotypes. Genetical hybrids moreover served as an intermediate group (group3; See AFLP and cpDNA data). To verify the results of the CDA the individuals in a PCA of the morphological characters were coloured due to their calculated groups and checked for their position.

4 Results

4.1 AFLP analysis detects no private fragments for interspecific hybrids

Three AFLP markers were analysed in 195 individuals (*H. perforatum*, N:153 , *H. maculatum*, N:42). Altogether, a total of 408 polymorphic fragments were scored. The number of fragments for each taxon differed significantly (*H. perforatum*: 123.1 ±23.2 and *H. maculatum*:97.6 ±22.2). Polyploid individuals showed in average more fragments than diploid individuals (see Table 2). Polyploid *H. perforatum* had in average 121.7 fragments where polyploid *H. maculatum* had in average 104.6 fragments. Diploid *H. perforatum* had in average 112.5 fragments whereas diploid *H. maculatum* had 79.6 fragments.
H. perforatum owned 99 private polymorphic fragments. Whereas there were only 15 in *H. maculatum* 15 (see Table 2).

Species	N_p (total:102)	N_i (total:195)	N_f (total 408)	Average gene diversity over loci	f_{pp}
H. perforatum (red genotype/diploids with red genotype/polyploid with red genotype)	52 (21/2/15)	153 (49/4/35)	123.1 ± 23.2 (134.1 ± 17.0/ 123.3 ± 3.8/ 135.1 ± 18.2)	0.239 ± 0.114 (0.200 ± 0.097/ 0.109 ± 0.072/ 0.192 ± 0.094)	99 (1/0/1)
H. perforatum (green genotype/diploids with green genotype/polyploids with green genotype)	52 (12/5/2)	153 (29/11/3)	123.1 ± 23.2 (108.1 ± 16.4/ 107.9 ± 15.5/ 111.2 ± 14.7)	0.239 ± 0.114 (0.208 ± 0.103 / 0.176 ± 0.093/ 0.234 ± 0.175)	99 (4/2/1)
H. perforatum (mixed genotype red-green/ diploids with mixed genotype red-green/polyploids with mixed genotype red-green)	52 (26/2/15)	153 (53/4/27)	123.1 ± 23.2 (126.4 ± 22.3/ 111.2 ± 9.3/ 128.0 ± 21.9)	0.239 ± 0.114 (0.228 ± 0.110/ 0.167 ± 0.110/ 0.225 ± 0.111)	99 (8/0/3)
H. perforatum (interspecific hybrid genotype/ diploids with interspecific hybrid genotype/polyploids with interspecific hybrid genotype)	52 (19/3/16)	153 (29/3/25)	123.1 ± 23.2 (112.7 ± 18.6/ 107.5 ± 12.0/ 112.5 ± 18.4)	0.239 ± 0.114 (0.231 ± 0.114/ 0.235 ± 0.176/ 0.229 ± 0.114)	99 (0/0/0)
H. maculatum (blue genotype/diploids with blue genotype/polyploids with blue genotype)	27 (24/14/9)	42 (35/14/20)	97.6 ± 22.2 (93.5 ± 18.5/ 79.6 ± 5.2/ 104.6 ± 16.9)	0.185 ± 0.090 (0.170 ± 0.083/ 0.124 ± 0.064/ 0.172 ± 0.086)	15 (6/0/3)

Table 2 ***Gene diversity by AFLP analysis***. In this table are shown: number of populations **(N_p)**; number of individuals (N_i); number of AFLP fragments scored (N_f); **average gene diversity over loci** and number of polymorphic private fragments (**f_{pp}**) for species, major genetic groups and cytotypes.

4.1.1 AFLP data shows two genetic groups in *H. perforatum*

16% of variance could be visualized due to a PCA in Fig. 1. Three groups are visible in this graph, but not clear cut. One group could be assigned to *H. maculatum* and the remaining two groups contain only *H. perforatum* individuals. In each of the three groups a cluster of diploid individuals could be identified. No geographic pattern was visible.

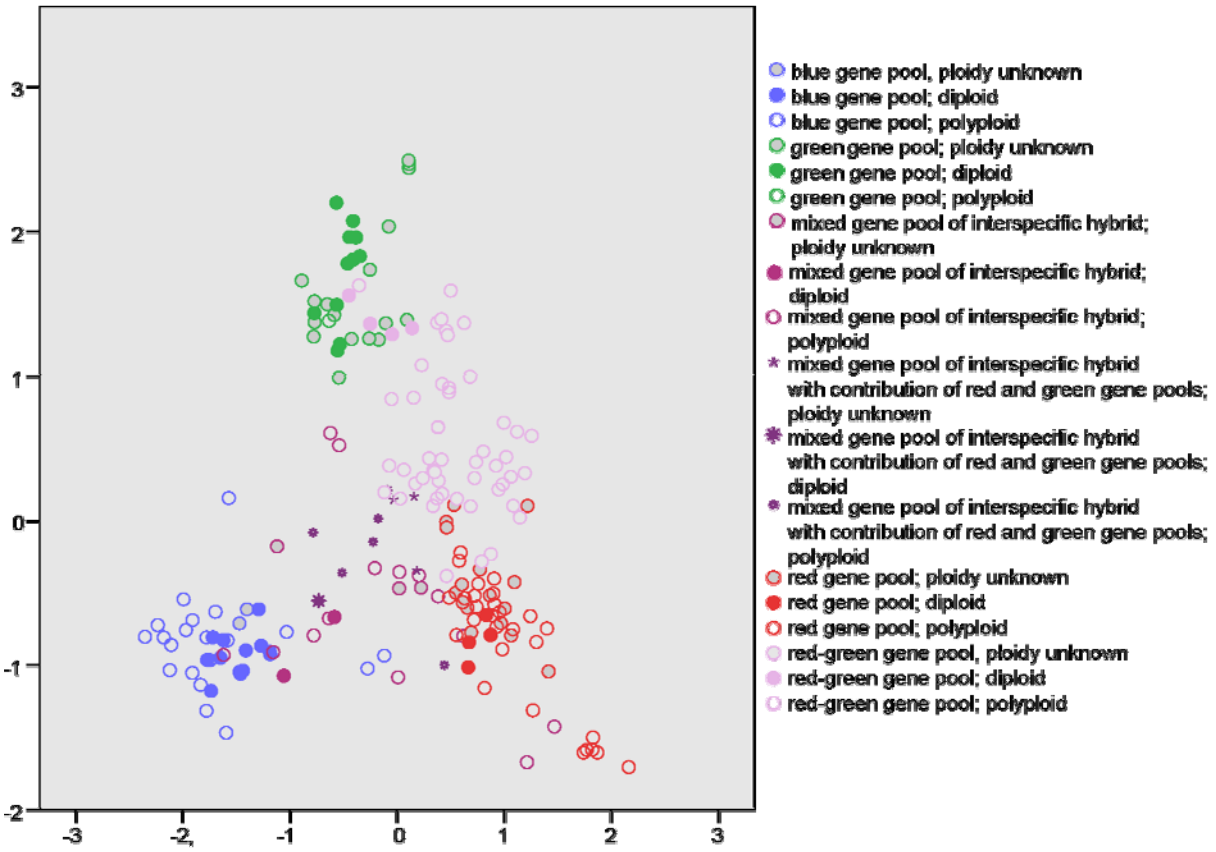


Fig. 1 ***Graph shows the results of the PCA of AFLP data***. The x-axis explains 10.7 % and the y-axis explains 5.3 % of the total variation. Each circle represents one individual. Symbols are explained in figure.

4.1.2 Admixture analysis shows gene flow between all three gene pools detected by AFLP

An admixture analysis in structure showed the cross breeding of three gene pools (from now on refered to as green, red and blue). 30 individuals (15 % of all individuals analysed with AFLP) were to 90 % of the green gene pool, 44 individuals (23 %) were to 90 % of the red gene pool, 32 individuals (16 %) were to 90 % of the blue gene pool, 47 individuals (24 %) had a mixed genotype of red and green, 19 individuals (10 %) had a mixed genotype of red or green and blue and 8 individuals (4 %) had a mixed genotype of red and green and blue (see details in extra material 1 on CD). Fig. 2 shows the geopgraphical distribution of the gene pools. The pie charts represent the mean value of the different clusters.

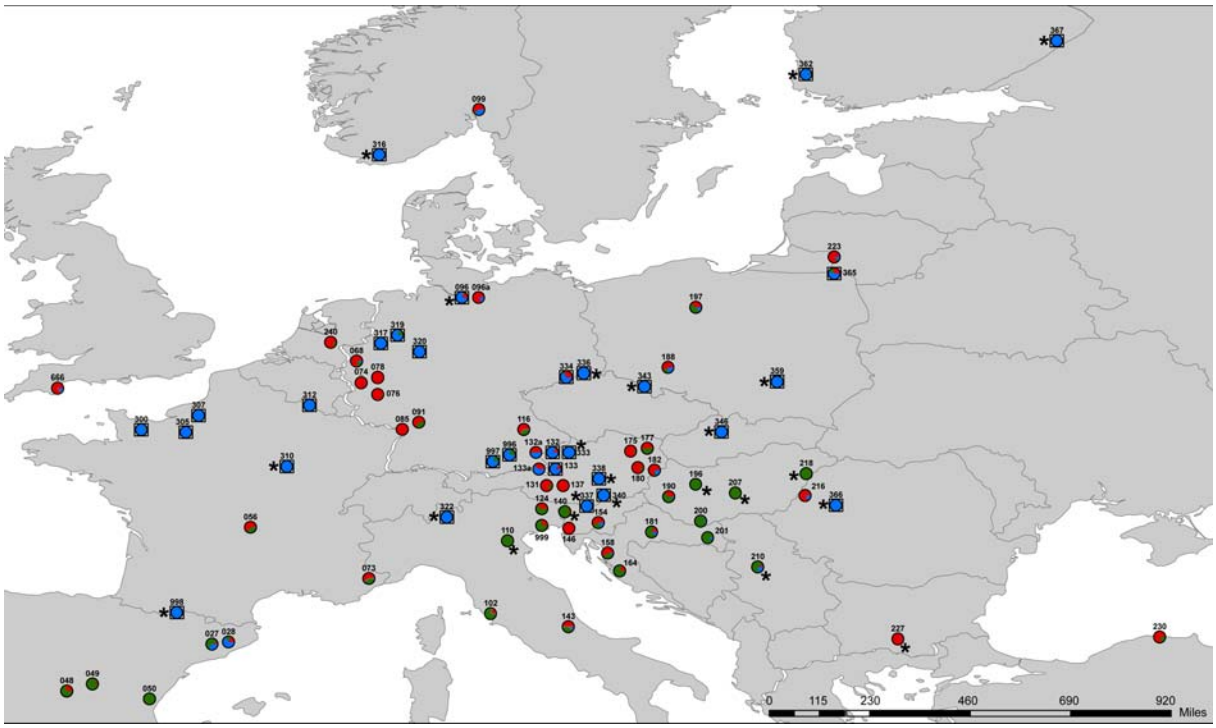


Fig. 2 ***Geographical distribution of 125 populations (195 individuals) and their shared genotypes illustrated as pie charts.*** Populations identified as *H. perforatum* are shown as circles; *H. maculatum* as squares; asterisks indicates diploid individuals. The STRUCTURE analysis has its highest significance with K=3 corresponding to the blue (*H. maculatum*), green and red gene pool (*H. perforatum*).

4.1.3 High variation within populations in AFLP data

When an AMOVA was applied to the data separated in two groups according to their species the majority of variation was found within populations (39.79 %) and among populations within each species (34.16 %). Less variation was found among groups (26.05 %). The same is valid when an AMOVA was applied to five groups according to their genetic background (populations with only green, red, blue, red and green or red/green and blue genotypes). Here again the highest variation was calculated within population (45.02 %; for details see Table 3).

Groups	Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Species	among groups	1	759.474	16.81972 Va	26.05
	among populations within groups	48	4771.304	22.05808 Vb	34.16
	within populations	117	3006.3	25.69487 Vc	39.79
Genetic groups	among groups	4	1233.251	9.44046 Va	17.35
	among populations within groups	30	2810.174	20.47116 Vb	37.63
	within populations	84	2057.533	24.49444 Vc	45.02

Table 3 ***Molecular variance of AFLP phenotypes.*** Comparing species (*H. perforatum* and *H. maculatum*) and five genetically defined groups (green, red, blue, red-green and interspecific mixed gene pool).

The genetic diversity D (Nei's diversity), calculated with AFLPdat for populations containing more than three individuals ranged from 0.016 to 0.241 (mean values of pairwise differences between the individuals of a population). The proportion of variable markers within the populations varied from 0.025 to 0.483 (also indicated in Suppl.Mat. 2). The down-weighted marker index is lowest in population no. 188 (5.59) and highest in population no. 164 (23.61). A comparison of the DW indices of the putative refugial areas and the formerly glaciated area

resulted in a mean DW index of 12.43 for the refugial and 9.23 for the colonization populations. Population no. 96 was excluded from these calculations (the reasons for the composition of individuals in this poulation seems to be the closeness to a railway station and not due to natural migration/gene flow patterns).

4.2 Chloroplast DNA sequences detect 21 haplotypes

577 individuals were screened with three primer pairs (91 individuals could only be assigned to a group not to a single haplotype) and outgroup individuals representing seven additional taxa. All were gathered into an alignment of 1620 bp in length representing 21 haplotypes in total (alignment on CD extra material 2), corresponding GenBank accession numbers are shown in Suppl.Mat. 3. Parsimony analysis running PAUP revealed 54 shortest trees each 85 steps in length and corresponding consistency (CI) and retention (RI) indices of 0.650 and 0.864, respectively. In summary 63 variable characters were identified with 38 parsimony informative character (10 based on gaps, 28 based on nucleotide polymorphisms), and 25 parsimony uninformative characters (8 based on gaps, 17 based on nucleotide polymorphisms). Within *H. perforatum* 15 polymorphic characters were found, which separate the species in ten haplotypes, whereas in *H. maculatum* 20 polymorphic characters were screened, which separate the species in four haplotypes (alignment on CD extra material 2).

The parsimony network analysis running TCS involved 89 mutational steps and clearly separated *H. maculatum* from *H. perforatum* (see Fig. 3).

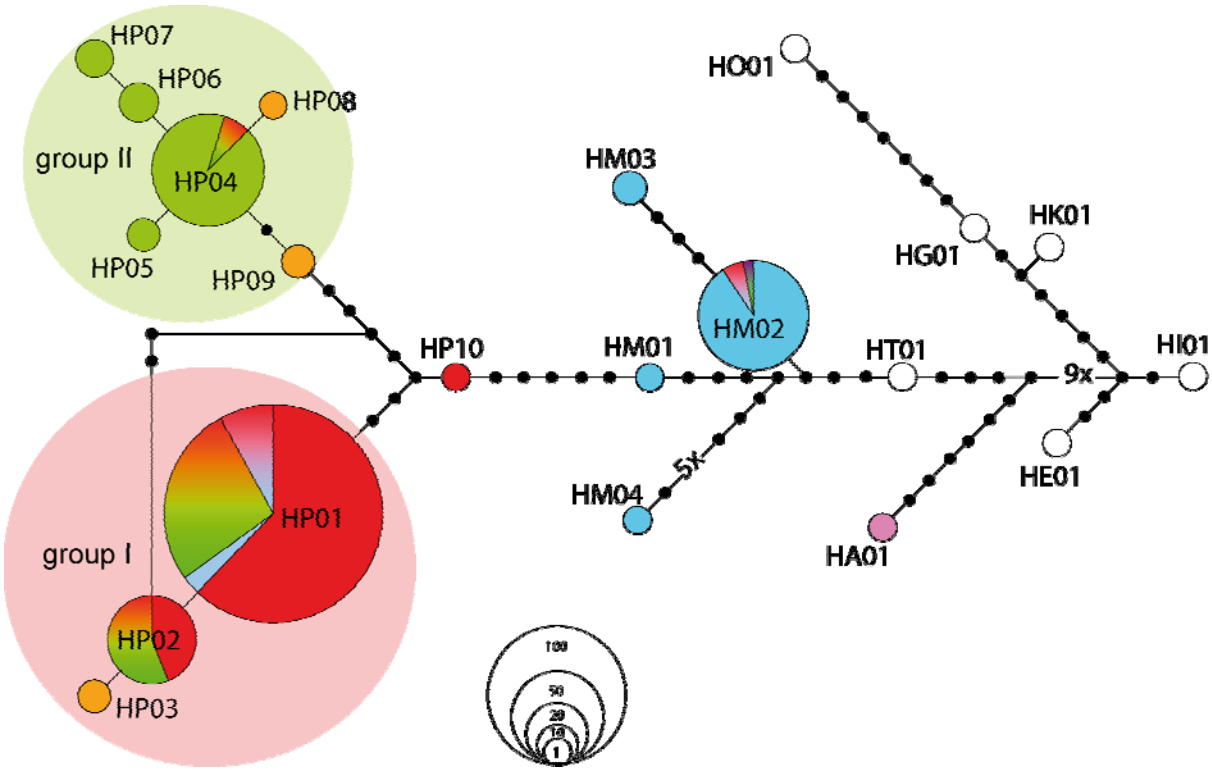


Fig. 3 ***TCS network of 21 cpDNA haplotypes***: Each circle represents one haplotype. Size represents number of individuals. **HP01-10**: *H. perforatum* and **HM01-HM04**: *H. maculatum*. Red, green and blue colours indicate AFLP genetic clusters (gene pools); for orange haplotypes no AFLP data was available; pink haplotype **HA01** represents *H. attenuatum.* Circles with no colour indicate outgroup species that are closely related to *H .perforatum*: **HT01**: *H. tetrapterum*, **HE01**: *H. erectum*, **HI01**: *H. pibariense*, **HK01**: *H. kamtschaticum*, **HG01**: *H. gracillimum* and **HO01**: *H. oliganthum*.

4.2.1 Chloroplast sequence data reveals two strongly differentiated lineages in *H. perforatum*

H. perforatum forms two genetic lineages which are separated by six mutational steps. Group I contains four haplotypes (HP01, HP02, HP03, HP10) separated by a maximum of six mutational steps and group II contains six haplotypes (HP04 to HP09) separated by a maximum of five mutational steps (se also Fig. 3). The two lineages display a large overlapping distribution area in the South of Europe. Furthermore group I is also distributed in the North of Europe (see Fig. 8).
In both groups diploid and polyploid individuals were found (see Fig. 4), whereas individuals with pure red AFLP genotype were found only in group I contrasting the pure green ones which were only found in group II. Green and red mixed genotypes could be recognized in individuals of both groups both groups (see Fig. 3).

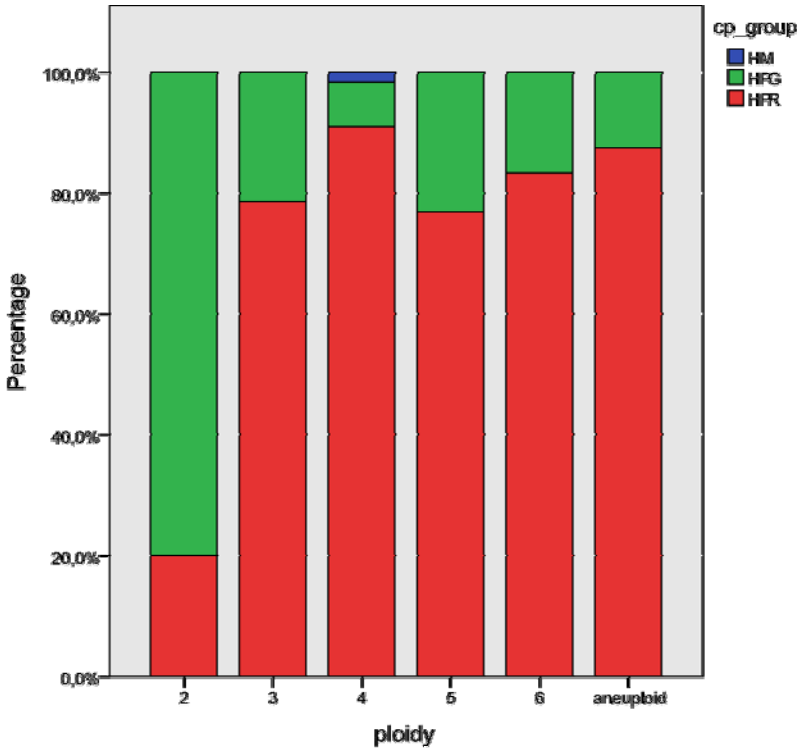


Fig. 4 ***Distribution of haplotype groups over ploidy level from individuals determined as H. perforatum***. **Green** stands for group II, **red** represents haplotypes of groupI and **blue** stands for haplotypes of *H. maculatum* group.

In *H. maculatum* four haplotypes were detected (HM01 to HM04) and separated by a maximum of 17 mutational steps. The closest relative to *H. maculatum* is *H. tetrapterum* with only four mutational steps separating it from *H. maculatum*. The other species, as well as *H. attenuatum* were well separated from *H. perforatum* and *H. maculatum* (see Fig. 3).

4.2.2 Chloroplast sequence data reveals high haplotype diversity in South Europe

Haplotype diversity was highest in the Iberian Peninsula (0.753), but similar to the other regions from South Europe (0.639 to 0.648), lowest haplotype diversity was calculated in the glaciated North (0.354). Nucleotide diversity was in all six regions rather low (0.00035 to 0.00233), but the lowest values were calculated for the unglaciated and glaciated North, 0.00127 and 0.00035 respectively (Table 4).

Region	N_i	N_h	Haplotype diversity	Sd_h	Mean number of pairwise differences	Sd_p	Nucleotide diversity	Sd_n
AT/HU	46	6	**0.6386**	0.053	3.516908	1.823756	**0.002332**	0.00134
Balk	72	4	**0.6405**	0.033	2.788732	1.49165	**0.001849**	0.0011
ES	31	6	**0.7527**	0.052	2.494624	1.382742	**0.001654**	0.00102
IT	68	5	**0.6475**	0.033	3.516681	1.812912	**0.002332**	0.00133
Nglac	118	4	**0.3536**	0.046	0.531508	0.446163	**0.000352**	0.00033
Nunglac	62	3	**0.4426**	0.067	1.918033	1.106399	**0.001272**	0.00081

Table 4 Number of individuals (N_i); number of haplotypes (N_h); haplotype diversity; standard deviation of haplotype diversity (sd_h); mean number of pairwise differences; standard deviation of mean number of pairwise differences (sd_p); nucleotide diversity and standard deviation of nucleotide diversity (sd_n) of *H. perforatum* from six regions: **AT/HU**: unglaciated parts of Austria and Hungary; **Balk**: Balkan region; **ES**: Iberian Peninsula; **IT** Apennine Peninsula; **Nglac**: glaciated regions in Northern Europe and **Nunglac**: unglaciated regions in Northern Europe.

4.3 Nuclear DNA

4.3.1 *H. perforatum* and *H. maculatum* form two genetically distinct groups with interspecies hybrids frequently found in contact zones

The internal transcribed spacer (ITS) was sequenced in 334 individuals. 79 individuals could be assigned to a single genotype, whereas 169 individuals could be assigned to two genotypes and 86 individuals could only be assigned to a group not to a single genotype. Outgroup individuals representing eight additional taxa were joined into an alignment of 666 bp in length representing 35 genotypes in total (alignment in extra material 3 on CD), corresponding GenBank accession numbers are shown in Suppl.Mat. 4. A Parsimony analysis running PAUP revealed one shortest tree with 67 steps in length and corresponding consistency (CI) and retention (RI) indices of 0.914 and 0.980, respectively. In summary 63 variable characters were identified with 31 parsimony informative characters (three based on gaps, 28 based on nucleotide polymorphisms), and 32 parsimony uninformative characters (four based on gaps, 28 based on nucleotide polymorphisms). Within *H. perforatum* 18 polymorphic characters were found, which separate the species in 15 genotypes, whereas in *H. maculatum* seven polymorphic characters were observed, which separate the species in six genotypes (see alignment in extra material 3 on CD).

The parsimony network analysis running TCS involved 69 mutational steps and clearly separated *H. maculatum* from *H. perforatum* (Fig. 5). The 15 genotypes in *H. perforatum* (HPI to HPXV) were separated by a maximum of 11 mutational steps. Five genotypes (HPXI to HPXV) could be separated from the others due to their chloroplast and AFLP assignement, both part of the green group. Four genotypes could be assigned to the red group (HPIII, HPV, HPIX and HPX), two genotypes carried an AFLP type of mixed red and green cluster (HPI and HPII) in addition to a chloroplast haplotype from group I, and for four genotypes no chloroplast or AFLP data was available. *H. perforatum* is well separated from *H. maculatum* by a minimum of five mutational steps. The six genotypes in *H. maculatum* (HMI to HMVI) were separated by a maximum of six mutational steps. Closely related species were all well separated from both, *H. perforatum* and *H. maculatum* (Fig. 5).

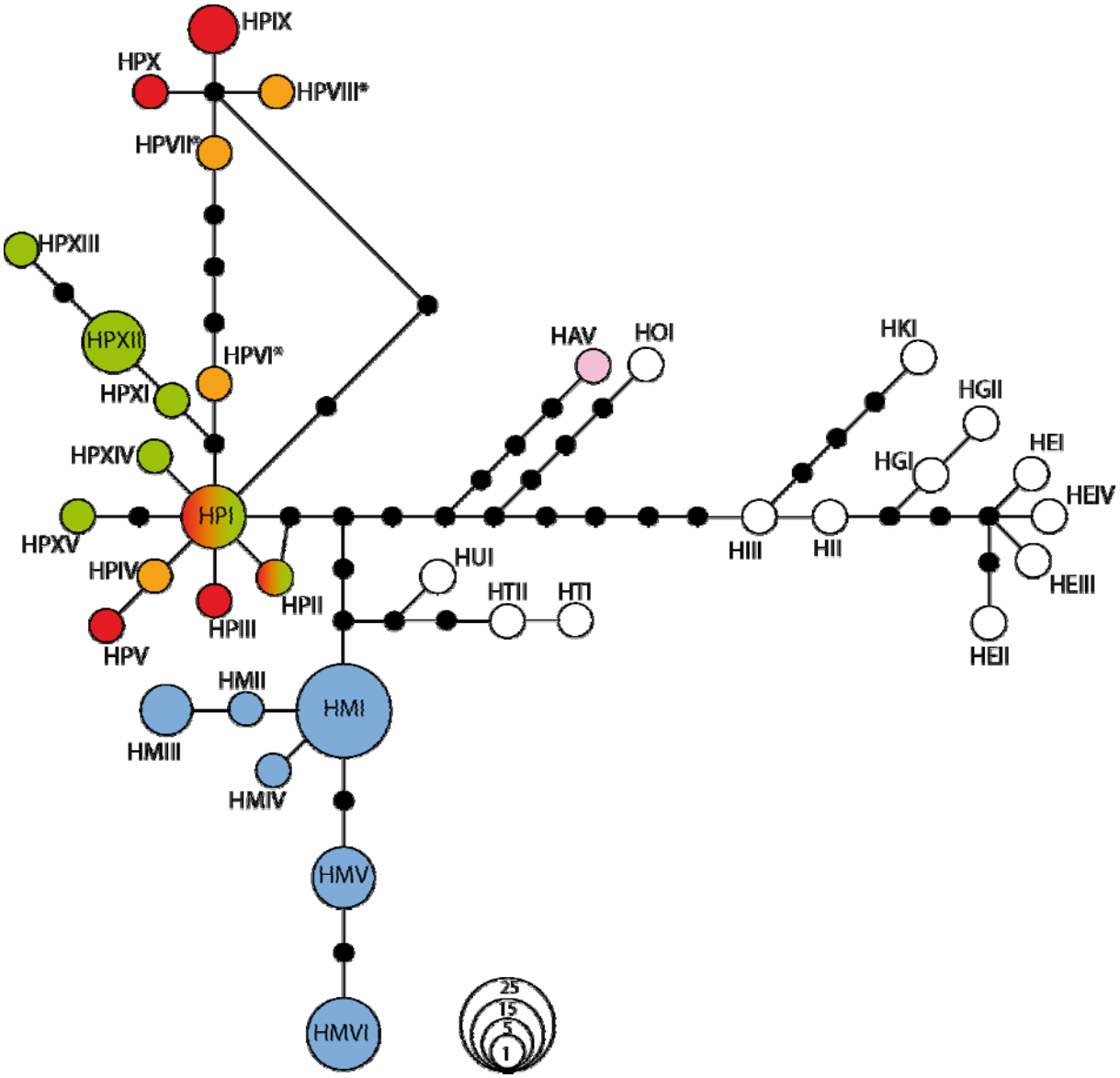


Fig. 5 ***TCS network of 35 nrDNA sequence genotypes***: Each circle represents one genotype; **HPI-HPXV**: *H. perforatum* and **HMI-HMVI**: *H. maculatum*. Red, green and blue colours indicate AFLP genetic clusters and/or cpDNA groups (gene pools); for orange genotypes no AFLP or cp sequence data was available; pink genotype **HAV** represents *H. attenuatum.* Circles with no colour indicate outgroup species that are closely related to *H .perforatum*; **HU**: *H. undulatum*, **HT**: *H. tetrapterum*, **HOI**: *H. oliganthum*, **HII-HIII**: *H. pibariense,* **HKI**: *H. kamtschaticum*, **HGI-II**: *H. gracillimum* and **HEI-HEIV**: *H. erectum*.

In 208 individuals two differing copies of sequences could be identified (for details see extra material 1 on CD). 45 individuals carried genotypes from *H. perforatum* and *H. maculatum*. All interspecific hybrid sequences had a chloroplast haplotype of *H. perforatum* (group I: Red), except one individual that carried a chloroplast haplotype of *H . maculatum*. 68 individuals carried genotypes from the green and the red gene pool and of those carried all a chloroplast haplotype of *H. perforatum*, with 60 % of group II (green) and 40 % of group I (red) (see Suppl.Mat. 6). 95 individuals carried two genotypes from the same gene pool (72 in *H. perforatum* and 23 in *H. maculatum*).

4.3.2 ITS sequence data shows weak geographical pattern in *H. perforatum*

Genotypes HPXI to HPXV are only found in the South of Europe, were also most of the unique genotypes are detected. Other genotypes showed no geographic pattern.

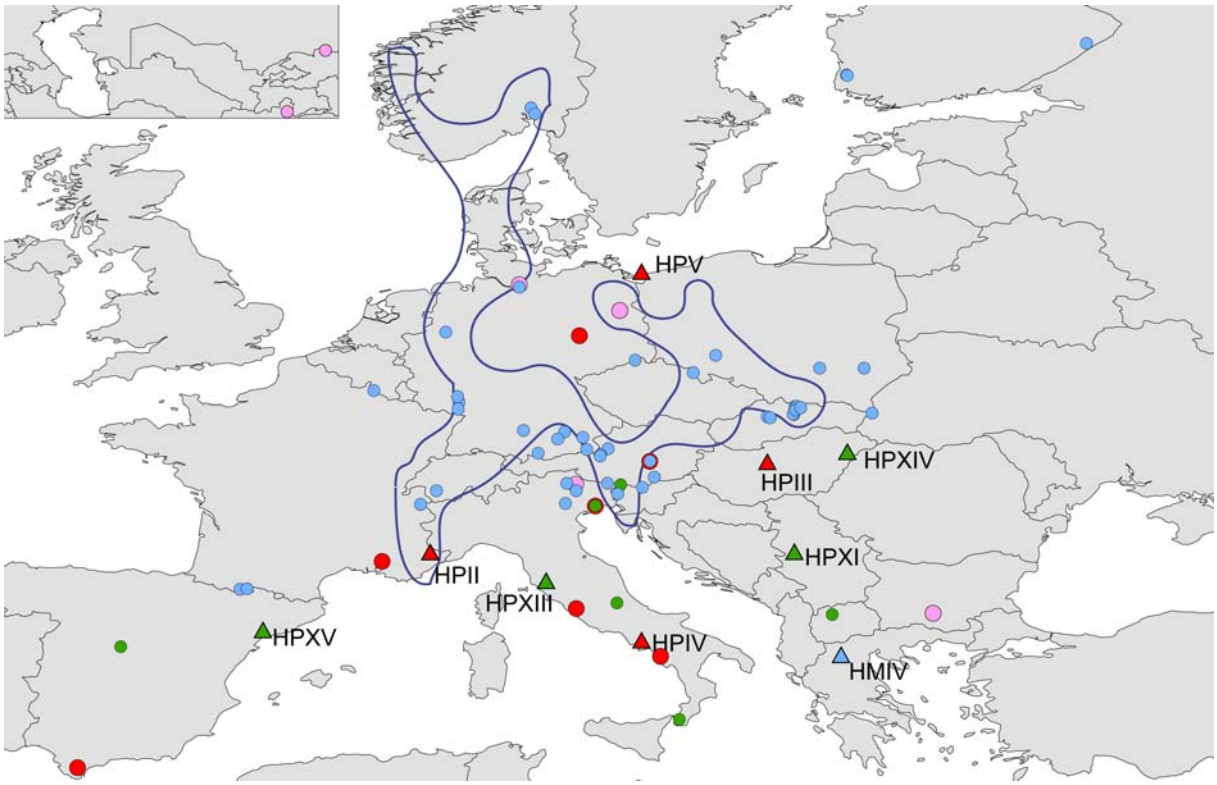


Fig. 6 ***Geographic distribution of ITS genotypes of H. perforatum and H. maculatum***. In the left upper corner are two individuals from Afghanistan and Kyrgistan respectively. Red circles represent genotype HPI, pink circles indicate individuals with HPIX or HPX, green circles stand for genotype HPXII and blue circles represent genotype HMI and HMIII to HMVI. Triangles represent unique genotypes. Blue line indicates zone where interspecific hybrids were detected via ITS.

4.4 Gene flow between and within H. perforatum and H. maculatum

92 Populations with mixed genetic background and 82 populations with pure genetic background were detected. The majority of populations with pure background was of the red gene pool and the majority of populations with a mixed genetic background was from red and green gene pool.

4.4.1 AFLP and ITS data detect gene flow between *H. perforatum* and *H. maculatum*

37 populations with a mixed genetic background of *H. perforatum* and *H. maculatum* were detected by AFLP analysis and by analysing their ITS sequences. From those individuals 18 % carried a chloroplast haplotype of *H. maculatum*, 77 % carried a chloroplast haplotype of *H. perforatum* group I and 5 % carried a chloroplast haplotype of *H. perforatum* group II. 35 populations were distributed within the contact zone of the two species. In Fig. 7 the contact zone is visualized as the overlap of chloroplast haplotypes from *H. perforatum* and *H. maculatum* over whole Europe. It is noteworthy that *H. maculatum* is rather restricted to arboreous habitats, whereas *H. perforatum* has no strict preferences. It is also visible that hybrids have not succeeded to spread further than the contact zone of their parental species.

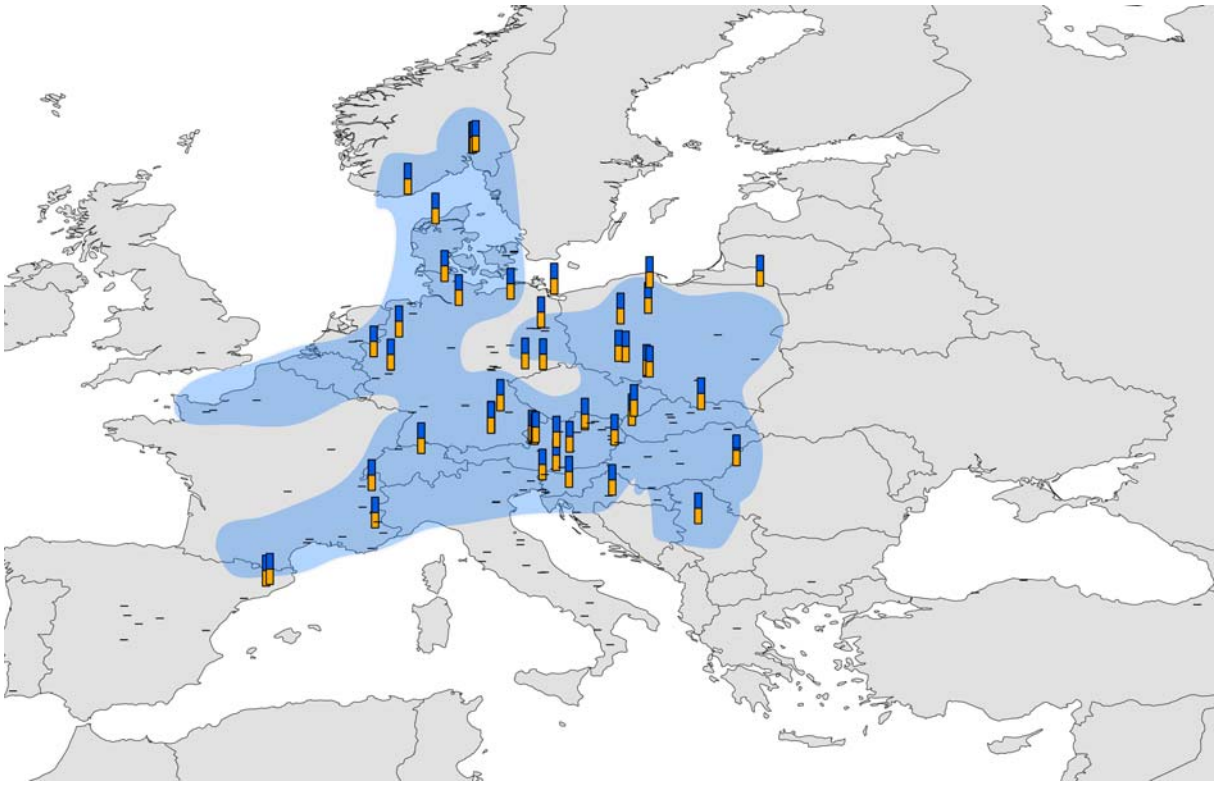


Fig. 7 ***Geographical distribution of hybrids between H. maculatum and H. perforatum***. Hybrids are represented by stacked bars. The blue cloud represents the contact zone of *H. perforatum* and *H. maculatum* detected by chloroplast sequence haplotypes.

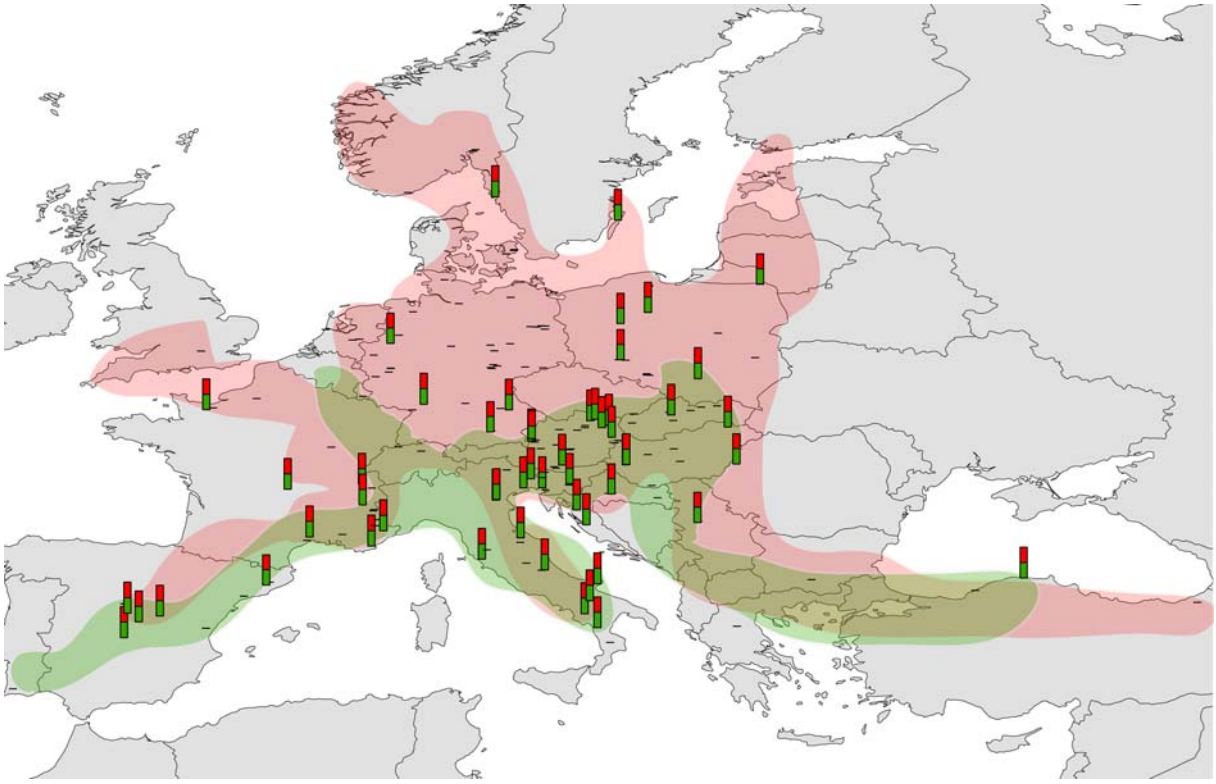


Fig. 8 ***Geographical distribution of intraspecific hybrids between red and green gene pools of H. perforatum***. Stacked bars represent individuals with mixed background. The red and green cloud represent the distribution of groupI and groupII haplotypes, respectively.

4.4.2 AFLP and ITS data detect gene flow between genetic groups in *H. perforatum*

45 populations with a mixed genetic background of the two *H. perforatum* gene pools (referred to as red and green gene pool) were detected by AFLP analysis and in ITS sequences. 33 populations were distributed within the contact zone of the two gene pools that were detected by chloroplast sequences. The wide contact zone is restricted to the South of Europe (Fig. 8), but individuals with mixed genetic background are also found in the North of Europe.

4.4.3 *H. perforatum* and *H. maculatum* form two genetically distinct groups.

Four individuals determined *H. perforatum* carried chloroplast haplotype HM02, which is characteristic for *H. maculatum* (see also Suppl.Mat. 5). Two of those individuals were shown to be of hybrid origin by AFLP analysis and are tetraploids. One individual determined *H. perforatum* carried a chloroplast haplotype HP01 (characteristic for *H. perforatum*) and carried an ITS genotype HMI (characteristic for *H. maculatum*) (see also Suppl.Mat. 6). Two individuals determined *H. maculatum* carried haplotype HP01, which is characteristic for *H. perforatum*. One of those individuals showed in the AFLP analysis hybrid origin and is diploid, whereas the other belonged to 100 % to the red cluster and is tetraploid. One individual determined *H. maculatum* carried a chloroplast haplotype HM03 (characteristic for *H. maculatum*) and carried an ITS genotype HPI (characteristic for *H. perforatum*), the AFLP anaylsis showed 99 % blue genotype (characteristic for *H. maculatum*).

4.5 Ploidy level estimations detected diploids and polyploids in H. perforatum and H. maculatum

Ploidy level estimations for 969 individuals are primarily based on flow cytometry, but in several independent cases confirmed by chromosome counts. The results are summarized in extra material 1 on CD and visualized in Fig. 9 to Fig. 12.

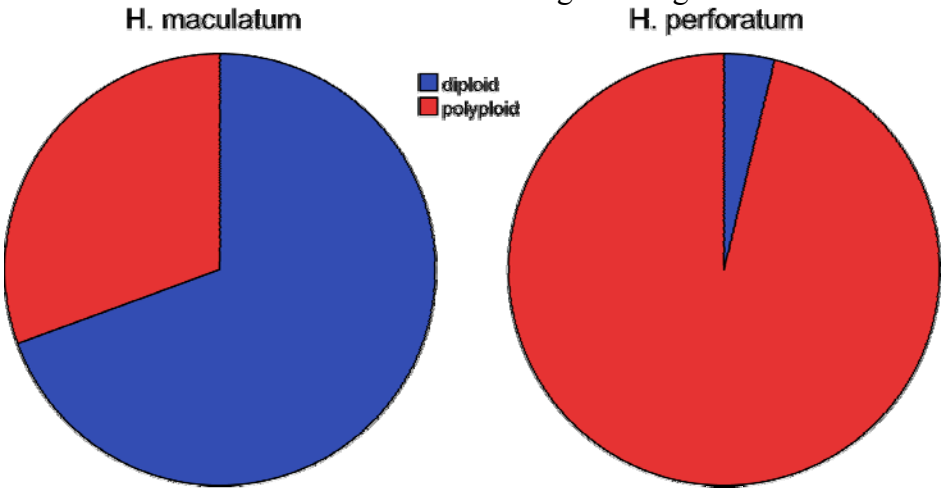


Fig. 9 Percentage of diploid (blue) and polyploid (red) individuals in *H. maculatum* and *H. perforatum.*

4.5.1 Ploidy measurements identify diploid individuals in both gene pools of *H. perforatum*

Most important was the recognition of many diploid individuals of *H. perforatum* (polyploids: 685; diploids: 42, see also Fig. 9). All diploid individuals were found in southern populations carrying either the “green AFLP genotype” and cpDNA haplotype HP04 from the “green haplotype cluster” or they carried the “red AFLP genotype” and cpDNA haplotye HP02 from

the "red haplotype cluster". Three individuals do not fit into these results, they occurred further North and two of them carried the haplotype HP01 (see also Fig. 10).
In *H. maculatum* numerous polyploids were detected, but we mainly found diploids (83 polyploids and 159 diploids, see also Fig. 9).

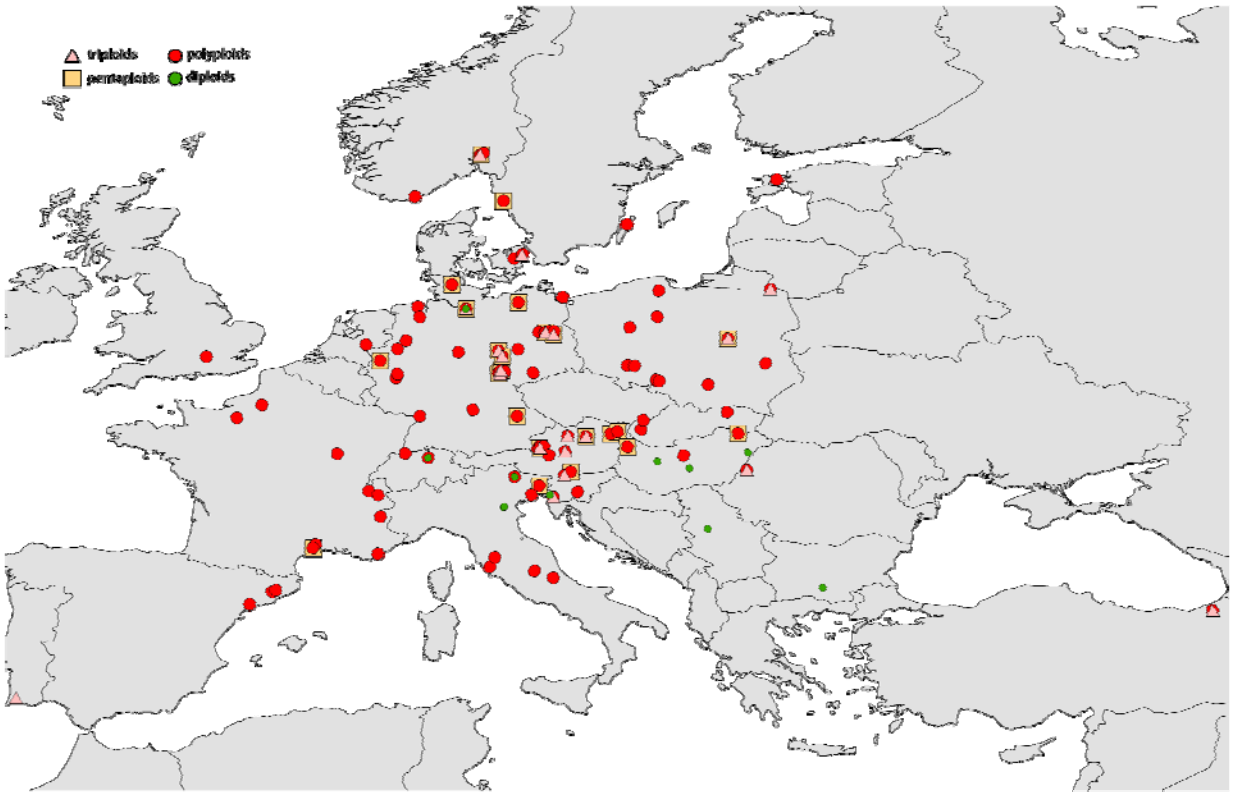


Fig. 10 ***Geographical distribution of ploidy levels in H. perforatum.*** Green circles represent diploids; red circles represent polyploids; pink triangles represent triploids and the orange squares stand for pentaploids.

4.5.2 Populations with mixed genetic background are not restricted to populations with mixed ploidy levels

Hybrids of any combination were always identified as polyploid with four exceptions of diploid *H. maculatum* X *H. perforatum* individuals in Switzerland, Germany and Serbia. Three diploid populations were found in Italy, which had a green genotype in the AFLP analysis and green haplotypes from chloroplast sequence analysis, but a mixed ITS profile (see also Suppl.Mat. 7).
The majority of populations comprised tetraploids only or tetraploids and few hexaploids (see Suppl.Mat. 8). 34 populations included mixed ploidy levels with tetra-, tri-, penta-, hexa- and higher polyploids, those populations did not show higher percentages of a genetic mixed background (from either ITS, AFLP or mixed with chloroplast data; see also Fig. 11). We found pure diploid populations with pure red or green genetic background, but also three populations with a red-green mixed background in Italy and one population in the Balkan region with a red-green-blue genetic background. Diploids were further detected in two populations which mainly included tetraploids and also a few hexaploids (in formerly glaciated Switzerland and formerly glaciated Italy) and in one population with mixed ploidy levels (population number 096 from Germany).

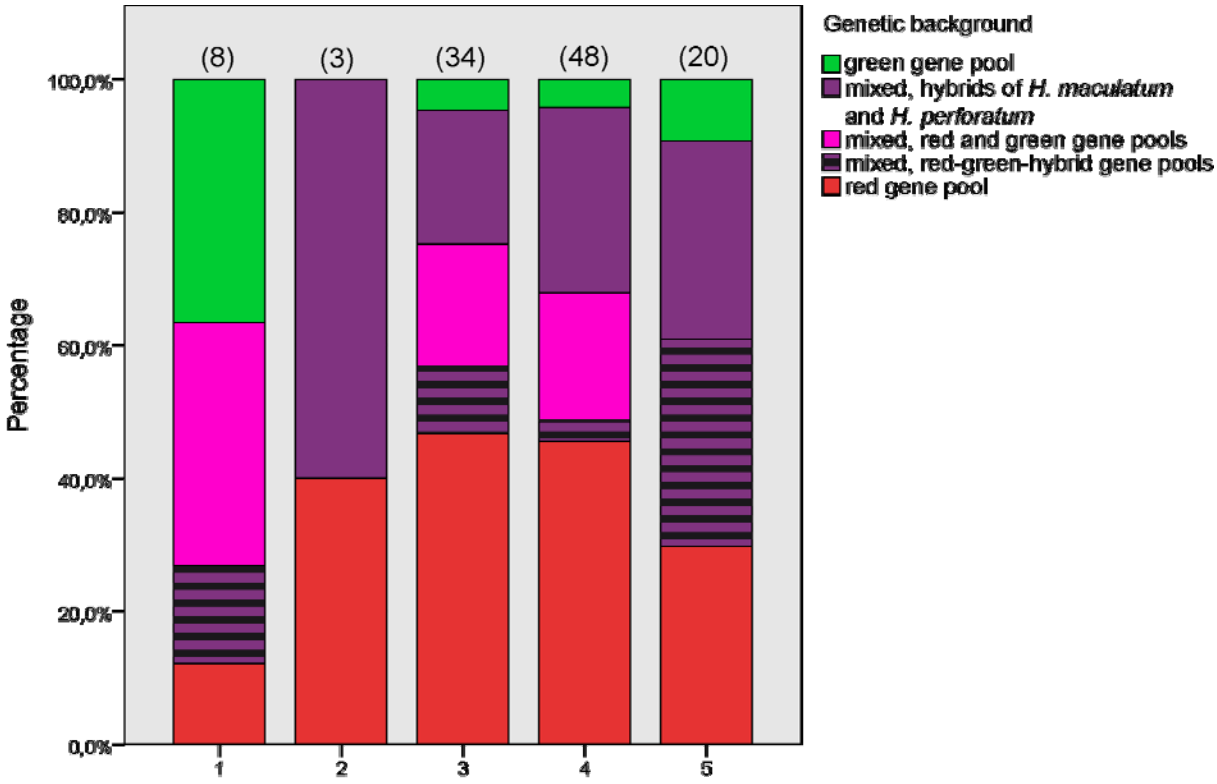


Fig. 11 ***Genetic background of five different population types in H. perforatum***. X-axis shows:**Population group1**: only diploid individuals; **population group2**: population with diploids and mixed other ploidy levels; **population group3**: mixed ploidy levels with triploid and pentaploid individuals; **population group4**: populations with tetraploids exclusively and **population group5** includes only tetraploids and hexaploids. Y-xis shows the percentage of different types of genetic background (green gene pool, mixed gene pools or red gene pool) in a population type

4.6 Obligate sexuality is restricted to diploid individuals

The breeding system of 77 individuals was analysed. 52 individuals were determined *H. perforatum* and 25 individuals were determined *H. maculatum*. 31 individuals were diploids (*H. perforatum*: 24 individuals and *H. maculatum*: 7 individuals) and 46 individuals were polyploids (*H. perforatum*: 28 individuals and *H. maculatum*: 18 individuals).

From the initial 1848 estimations (77 accessions times 24 seeds) 559 measurements provided no C-value ratio. This coud be expected since it has been already described that *Hypericum* often produces seeds without any endosperm (Robson 2002).

Of all diploid individuals reproduced via double fertilization three individuals had a low frequency (5.6-17.6 %) of apomeiotic and double fertilized seeds and four individuals had a low frequency (4.5-13.3 %) of seeds that did not develop from double fertilized eggs (a summary of these results is given in extra material 1 on CD; for detailed results see extra material 4 on CD). Polyploid individuals of both species had a wide variability of reproductive pathways, ranging from obligate apomictic to sexual with a low frequency of apomixis (see Fig. 12).

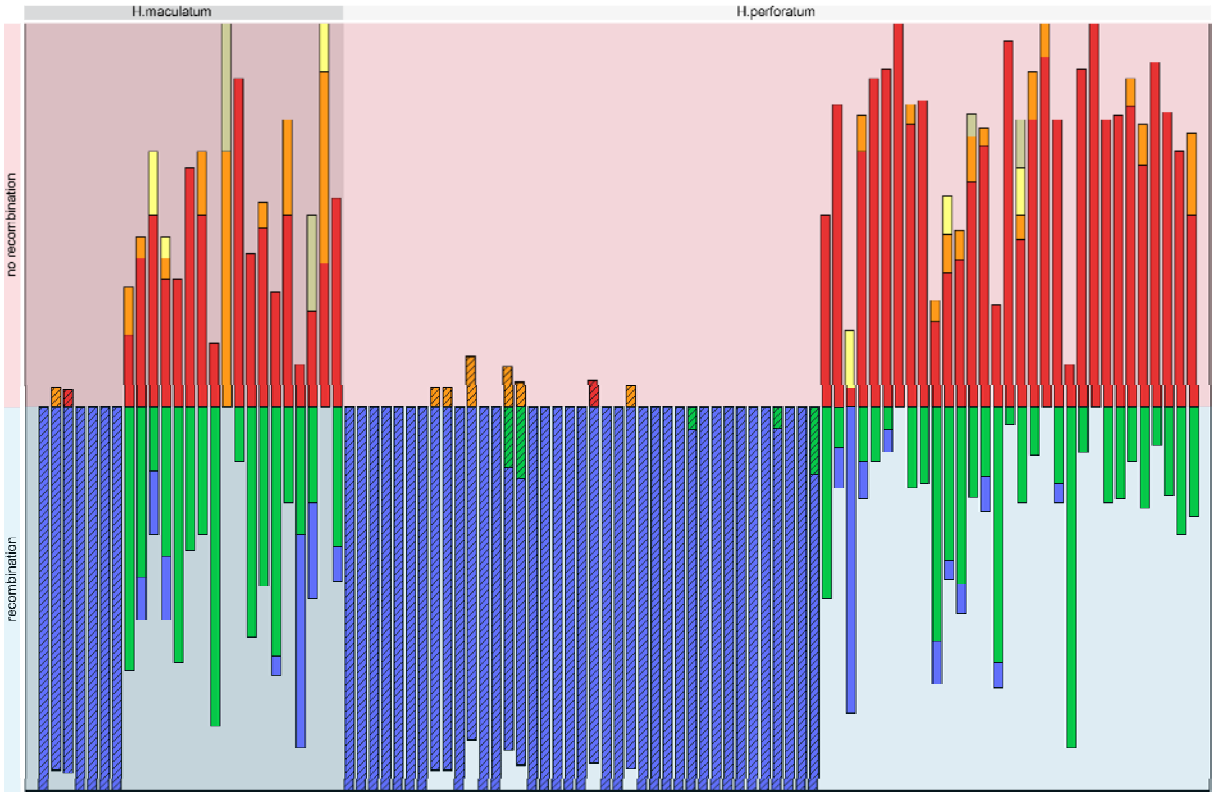


Fig. 12 Six different pathways for reproduction in *H. perforatum* and *H. maculatum*. One bar represents 100 % of the seeds analysed for one individual. **Blue** represents sexually produced seeds, **green** apomeiotic (but double fertilized) seeds, **red** apomeiotic and pseudogamous seeds, **orange** seeds that were produced apomeiotic and autonomous, **yellow** seeds of parthenogenetic and pseudogamous origin and **beige** seeds that are of parthenogenetic and autonomous origin. Striped bars stand for diploid individuals.

None of the five genetic groups (green, red, red-green, blue-red-green and blue) showed a preference for a reproductive pathway (see Suppl.Mat. 9).

4.7 Morphology

For a beginning it has to be pointed at classification difficulties among the arrangement of close relatives of *H. perforatum*. *H. attenuatum* was often found wrongly determined. It often turned out to belong to the groups of *H. maculatum* or *H. erectum* in our TCS network. Three ITS sequences of *H. attenuatum* were acquired from the NCBI genbank; two of those were *H. erectum* (HEII: AY572993; HEIII: AY572995) and one turned out to be *H. maculatum* (HMII: FJ793043). Leaf material and DNA material from potential *H. attenuatum* individuals was also acquired from collaborators in IPK, Gatersleben and HZU, Zhejiang University. Only one individual with chloroplast and ITS sequences matching a position in the network for *H. attenuatum* was found (chloroplast haplotype HA01 and ITS genotype HAV) other individuals were identified as *H. maculatum* or *H. erectum*.
We also found that the determination and genetic relationship between *H. pibairense, H. kamtschaticum, H. gracillimum, H. nakaii, H. hakonense* and *H. senanense* needs to be thouroughly clarified. Due to determination problems sequences acquired from potential *H. nakaii, H. hakonense* and *H. senanense* individuals are not included.

4.7.1 PCA based on morphological characters reveals high variability in *H. perforatum*

In a PCA, containing 450 individulas, two major groups of individuals were identified representing *H. maculatum* and *H. perforatum*, respectively, with 48,4 % of the variance explained by the first two axes (Fig. 13). Main characters to distinguish between the two taxa

are the density of transparent glands on leaves (characters c and f), differences in the size of leaves and sepals (characters a,d and e), differences in shape of petals (character m), shape of sepals (character p), number of stripes on the stem (character g) and the distribution pattern of black glands on petals (character k), for details see Suppl.Mat. 10 and Suppl.Mat. 11.

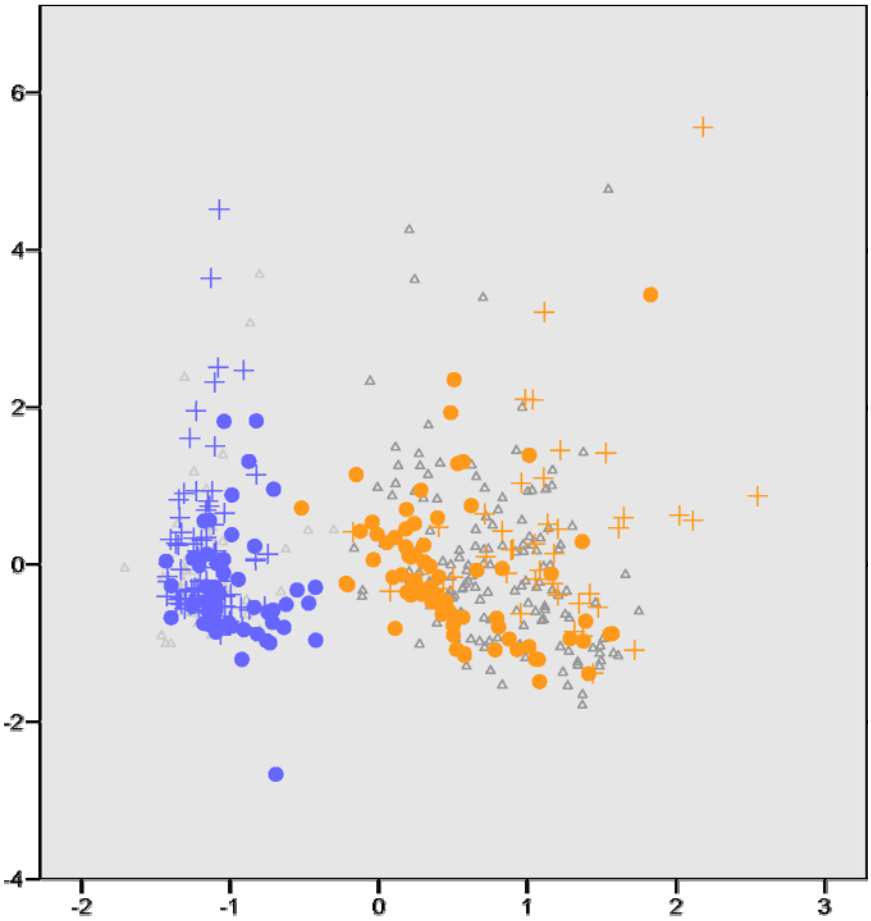


Fig. 13 ***PCA results of 15 morphological characters***. The x-axis explains 36.2 % and the y-axis explains 12.2 % of the total variation. Individuals of *H. maculatum* and *H. perforatum* with unknown ploidy level are shown as **triangles**; diploids are shown as **crosses** and polyploids are **circles**; individuals determined *H. perforatum* are **orange** and individuals determined *H. maculatum* are in **blue**.

Within *H. maculatum* there is no obvious morphological differentiation according to ploidy level variation based on the 15 characters scored.
H. perforatum did not cluster as tight as *H. maculatum*. Due to the high variability of *H. perforatum* (shown in Suppl.Mat. 10 and Suppl.Mat. 11) two morphotypes are hidden in the cloud of individuals. The position of the diploid individuals to the polyploid individuals of *H. perforatum* is significantly different. The main characters to distinguish between the two cytotypes are the density of transparent glands on leaves (characters c and f), density of black glands on leaves (characters b and e) and differences in the size of leaves (characters a and d).

4.7.2 Morphological data corresponds to genetic differentiation

To illustrate a correlation of morphology and genetic background the individuals from the PCA in Fig. 14 are coloured according to their genetic background (based on AFLP, ITS sequence and chloroplast sequence data) and their own cytology. The three gene pools correlated with the morphology. The blue gene pool was found in the *H. maculatum* cloud and red and green gene pools correlated with two morphotypes in the *H. perforatum* cloud. Individuals with an interspecific hybrid background are separated in eight individuals similar

to the habit of *H. maculatum* and 16 individuals similar to *H. perforatum*. Only two individuals show an intermediate position. Individuals with a mixed genetic background of the red and green gene pool were scattered all over the cloud of *H. perforatum*.

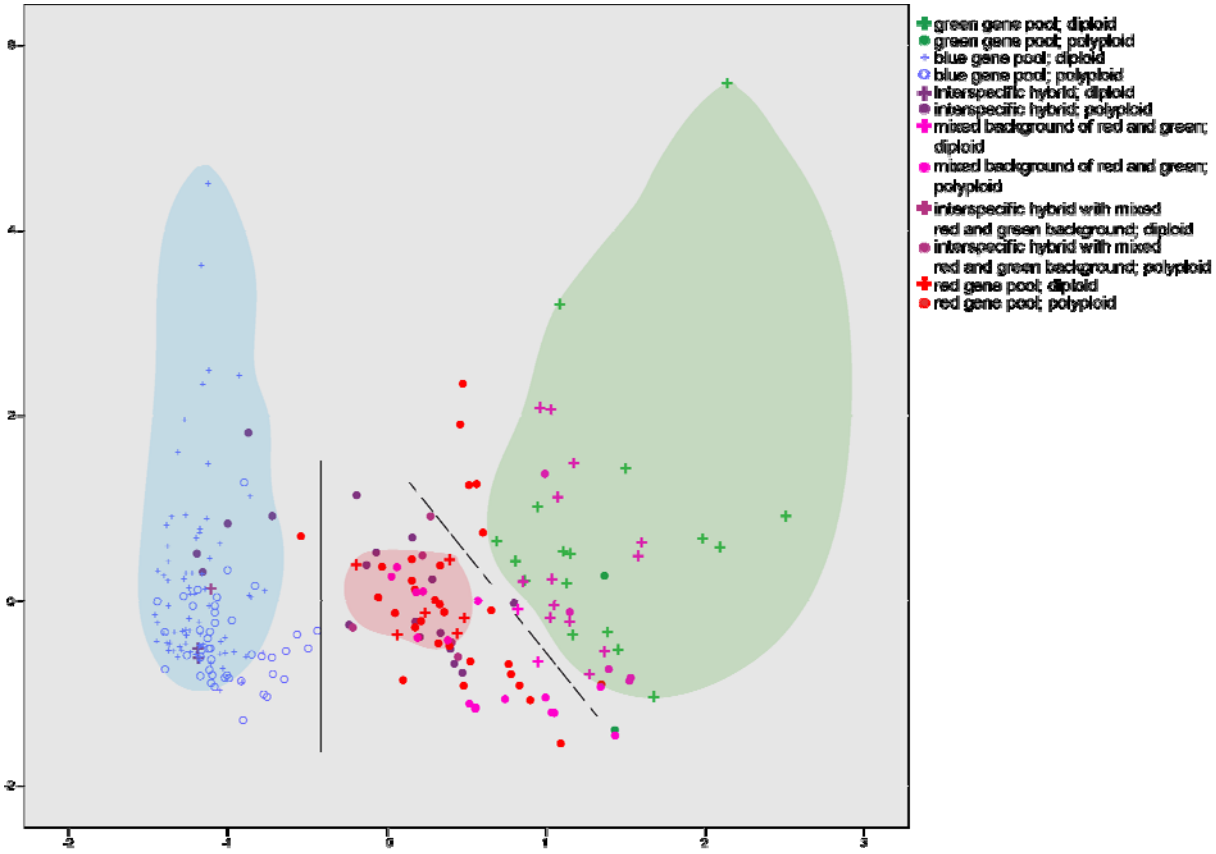


Fig. 14 ***PCA results of 15 morphological characters***. Individuals are coloured refering to their genetic background and ploidylevel. The x-axis explains 36.2 % and the y-axis explains 12.2 % of the total variation. Only individuals with known ploidy level are displayed. The continuous line represents the disjunction between *H. maculatum* and *H. perforatum* and the dotted line represents the split between red and green diploids. The clouds represent the variation in diploid *H. maculatum* (in blue), diploid *H. perforatum* with red genetic background (in red) and diploid *H. perforatum* with green genetic background (in green).

4.7.3 CDA separates two morphotypes due to their genetic background

The CDA was performed for 442 individuals, including 103 individuals serving as references (group1: 19 diploid individuals with green genotype, group2: 67 diploid individuals with blue genotype and group3: 16 polyploid individuals with a mixed genotype of *H. perforatum* and *H. maculatum*). Results are shown on CD: extra material 1.

106 individuals were assigned to group1, 107 individuals were assigned to group2 and 127 individuals were assigned to group3.

The CDA program predicted (based on morphological characters) for three individuals a higher probability to be included in group2 instead of group3, in which they were included due to their genetic background.

Seven individuals were not correctly grouped in the CDA according to the PCA (see Fig. 15). In further calculations they were regrouped according to their PCA classification.

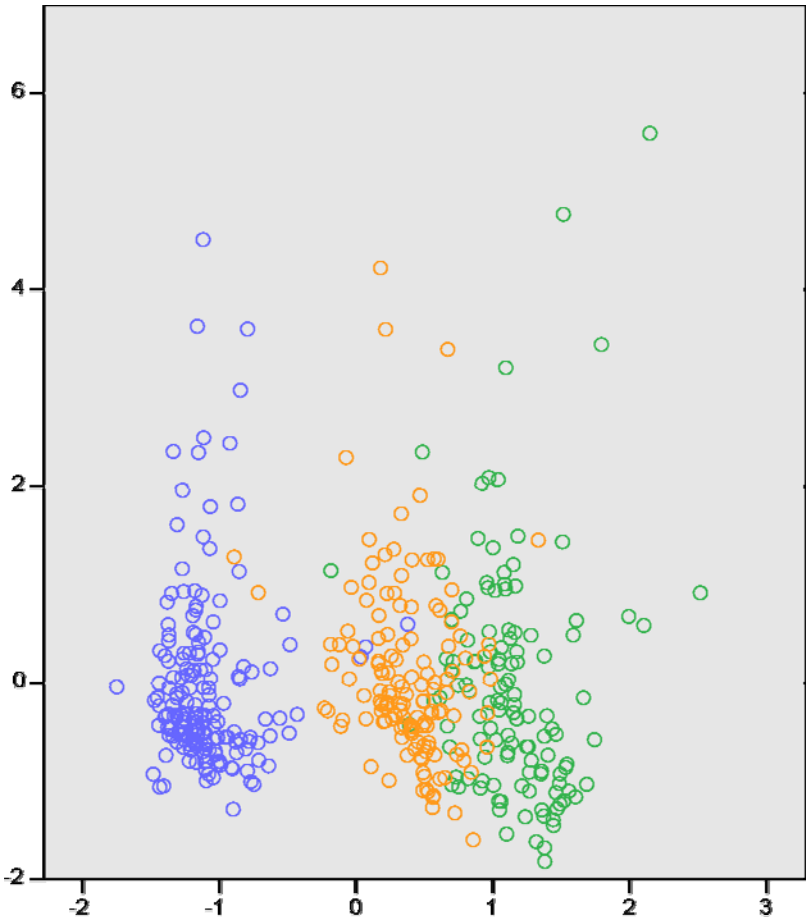


Fig. 15 ***PCA results of 15 morphological characters coloured according to results of CDA***. The x-axis explains 36.2 % and the y-axis explains 12.2 % of the total variation. Individuals are coloured due to their predicted groups. **Group1**: *H. maculatum* (blue); **group2** diploid *H. perforatum* individuals (green) and **group3** polyploid *H. perforatum* individuals (orange) with a mixed genetic background.

4.7.4 PCA shows no environmental influence on morphological characters measured for *H. perforatum* and *H. maculatum*

A substantial amount of individuals analysed morphological were cultivated in the Botanical Garden Heidelberg. Seeds were germinated directly on soil in the green house. After some months they were transfered to the outdoor territory, where they were exposed to the temperate climate of Heidelberg. The seeds for the cultivated individuals and individuals collected in the wild came to a similar percentage from different environmental influenced locations.

The results of the PCA do not change when specimens cultivated in the Botanical Garden were compared with the corresponding data set of individuals collected directly in the wild (Suppl.Mat. 12) indicating that on average the characters measured are not influenced by environmental differences.

5 Discussion

5.1 Melting gene pools: a diploid-sexual versus polyploid-apomictic reproductive system

5.1.1 Polyploidy in *H. perforatum*

Concerning *H. perforatum*, a diverse distinction can be observed. Diploids and polyploids coexist in Southern Europe, but further North only polyploids are found.

For angiosperms the majority of polyploids are found in higher altitude and latitude (Grant V., 1981). Now many scientists agree with Stebbins (1985) that these findings are true due to the dominant life form as perennials (polyploids are found more frequently in perennials than in annuals) or due to greater colonizing abilities of polyploids.
Many studies also show that polyploids often occupy habitats different from those of their diploid progenitors (reviewed in Bierzychudek P, 1985; Soltis, 2000; Stenberg, 2003 and Van Dijk, 1997). For *H. perforatum* both theories are true to some degree. *H. perforatum* is a perennial and the polyploids seem to be better colonizers of the North of Europe. From our data we can't find a differentiation of habitat between diploid and polyploid populations. We found populations of both cytotypes on disturbed road sides or openings in forests. However no diploid population was found higher than 1000 m.s.l..
Only in a few populations actual sympatry could be discovered. Nonetheless a constant gene flow between diploids and polyploids seems to be feasible, otherwise genotypes/haplotypes which are restricted to a certain cytotype could be found.
Matzk *et al.* (2001) discovered in *H. perforatum* rare diploid progeny deriving from tetraploid parents, which of course raises the questions how this is possible and how their genetical appearance looks like.
Fortunately we were able to analyse two of those diploids from two separate populations in collaboration with Marta Molins from the IPK, Gatersleben, Germany. They did not differ from the rest of the population. One was from the red gene pool and the other population was of hybrid origin. So it seems plusible that these diploids are dihaploids. The findings of rare unreduced not fertilized embryogenesis was already mentioned in Robson N. (2002) and Matzk *et al.* (2001) and underpins our seed screen results. These dihaploids nonetheless do not deliver an explanation for populations that consist of only diploid individuals. These pure diploid populations are residuals from the formerly wider distributed pure diploid species.
But it needs to be considered that polyploidy can be seen as transilience an idea already stated by Soltis D. (1999), when we see polyploidy as a recurrent incident with constant diploidization events accompanying speciation (as shown in e.g. in Brassica, Song 1995 and in Nicotiana, Leitch 1997).
Our results suggest that diploids show also a rare tendency to produce polyploids via apomeiotic embryogenesis. This theory seems to be confirmed in *H. perforatum*.

5.1.2 Apomixis in *H. perforatum*

In *H. perforatum* sexuals are found in Southern Europe and facultative apomicts in all parts of Europe.
It is a frequently used and convenient way to view the different forms of a species as sexuals and asexuals, but this is seldom the whole truth. In plants and animals asexuality is often accompanied with polyploidy and hybridization (Bierzychudek P, 1985; Kearney M, 2005 and Lundmark M. *et al.*, 2006). It is not clear which of these phenomena are responsible for the success of species which incorporate them. In *H. perforatum* we can definitely say that the diploid obligate sexual populations are not as successful as the polyploid facultative apomicts in respect to their distribution in Europe. In connection with hybridization as driving force, we find populations with pure genetic background (in fact the red gene pool) as successful as populations with mixed genetic background (intra- and interspecific mixture). Both groups are mainly polyploids, which gives polyploidy over hybridization credit as a driving force.
A common assumption is that asexual lineages are evolutionary younger than sexual outcrossers (Holsinger K., 2000). Of course there are exceptions like in the long-term asexual evolution in rotifera (Welch J. *et al.*, 2003). It is also debated that apomixis is reversible. Actually most sexual plants have the capability of asexual reproduction and most asexual plants retain some degree of sexuality (Richards A, 2003). D'Souza states that occasional sex is as advantageous as obligate sex (D'Souza T. *et al.*, 2010), but rare sex seems to be

evolutionary unstable. If sexuality is too rare benefits of sex vanish (Peck J. *et al*., 2000) and organism might dismiss sex completely, but also the opposite is possible, that more and more the reproduction turns to obligate sex. It seems that plants maintain a certain degree of apomixis and sexuality and due to reasons not known (e.g. environmental disturbance or genetic drift) the degree from one reproductive system can rise or fall.
In the case of *H. perforatum* it seems likely that the sexual diploid populations in Southern Europe are the remainings of the progenitors of polyploid apomicts. In Southern Europe the diploids do characterize two distinct gene pools corresponding to cp DNA haplotypes, which provide evidence for an ancestral state of diploids.
However *H. perforatum* is not stuck within asexuality, due to a remainig low sexuality. Maybe we can see apomixis like polyploidy as a helpful companion during speciation under conditions that are not easy to manage for sexuals, i.e. in disturbed environments, be it due to quaternary glaciation cycles or due to human impact.

5.2 Inter- and intraspecific hybridization in H. perforatum and H. maculatum

5.2.1 Intraspecific hybridization

To distinguish between interspecific and intraspecific hybridization it has to be considered, that it is not always easy to define a species. Is there a threshhold of genetic distance or a percentage of high reproductive isolation from which we can read the species delimination?
Mallet J. (1995) stated that if distinguishable genotypes remain distinct even while gene flow is possible you may call it species. This leads to the conclusion that the two gene pools in *H. perforatum* could be considered as two species. Here crossing experiments with respect to the genetic background are of interest.
The questions remaining are for how long will they stay distinct or will they melt together in one species different from their two parental lineages, is one going to conquer the distribution range of the other or is this a rather stable phase in which they can stay, till more persuavive environmental disturbances push them forward in evolution.
Looking at the morphology of H. perforatum we already see that the two lineages do not seem to stay that much distinct anymore. This will be highlighted in chapter 5.2.3.

5.2.2 Interspecific hybridization

We did not find any evidence for an interspecific hybrid origin of *H. perforatum*. However our data provides clear evidence that *H. perforatum* and *H. maculatum* are in close genetic contact, but are separated for already a long time.
Hybridization is most prevalent in rapidly radiating groups (e.g. Brassicaceae, Asteraceae, British ducks (Anatidae)), but hybridization and introgression may often persist for millions of years after initial divergence (Mallet J., 2005) even when the two diverged populations share the same environment (Barton N., 2001).
In *H. perforatum* and *H. maculatum* we encountered two species which can rapidly colonize open niches via polyploid and apomictic populations, but *H. maculatum* with a prefered habitat (rather humid woods, scrub, grassland, moorland, meadows and ditches) and *H. perforatum* not biased by this (open woodland, meadows, grassland and steppes, riverbanks, stony and grassy slopes, roadsides, in dry or well-drained habitats). In all contact zones, that we analysed we found individuals with hybridogen background. It seems, where ever those two lineages meet they form hybrids.
Perennial habit, outcrossing breeding system and clonal reproduction are advantageous to formation and stabilizition of hybrid taxa (Ellstrand *et al.* 1996). This is partly true for our study we find many hybrids in contact zones, but stabilization is so far not reached, otherwise they would not be restricted to the contact zone of their parents.

Many authors, also Soltis *et al.* in 2008 cautioned to use nrDNA like ITS sequence data, that undergoes homogenization. As for the "biggest pit" due to misinterpreting of polymorphic sites in chromatograms we were aware of the polyploidy in *H. perforatum* and due to our back and forward sequencing it can be said with absolut certainty that the polymorphic sites we found are valid.
However, the problem that a complete homogenization of an old hybridization event took already place, without being noticed is possible as it would then also be too old to be recognized in the AFLP analysis.
Hybridization and Apomixis are often associated with environmental disturbance (Arnold M. et al. 1999 and Bierzychudek P. 1985), but at least for apomicts not statistically significant. We also did not find an association between hybridization and apomixis with todays disturbed environment, but our data was not collected with an emphasis on that (many accessions missed a detailed report about their habitat).

5.2.3 Indications from morphological hybrids

In our morphological analyses we detected not many individuals with intermediate habits of *H. perforatum* and *H. maculatum*.
There is a long record in systematics to successfully identify hybrids primarely by morphological intermediacy. It can be used in the early stages of genetic introgression to indicate the hybrid degree. In general its application is limited due to dominance or epistasis effects or backcrossing (Wilson P, 1992). Also showed Rieseberg L. (1993) in artificial hybrids, that 10 % of the morphological characters measured in the F1 hybrids and 30 % in the later generation hybrids were novel or extreme relative to the parental species.
Already Noak (1939) showed in his crossing experiments between *H. maculatum* and *H. perforatum* that, F1 hybrids with *H. perforatum* as their mother, look not different from *H. perforatum*. He found intermediate looking hybrids only in plants with *H. maculatum* as mother. From our data we can see that most individuals with interspecific hybrid background had *H. perforatum* as their mother, that means that our results may reflect rather recent hybridization events in North and East Europe. On the other hand we found individuals with interspecific hybrid background in the morphomatrix-PCA also in proximity of *H. maculatum* detected, but those were the few individuals with *H. maculatum* chloroplast haplotype and restricted to the region between Munich, Germany and Salzburg, Austria.
To discuss the hidden morphotypes in *H. perforatum* we have to consider the fact that we found only one population with diploids and a pure red genetic background. Which led to an underestimation of them as a discrete genetic and cytological group.
We can conclude anyway that the two morpho types found correspond to two subspecies recognized by Robson N (2002). The diploid green morphotype can be correlated to the *H. perforatum* subsp. *veronense* and the red morphotype correlates to *H. perforatum* subsp. *perforatum.* Individuals with a green or red genetic background are not only genetically closer, but seem to be also morphologically closer. There was either not yet enough time to separate them properly (reproductive isolation) or which seems more conclusive, they are melting together again, by reduced reproductive isolation via polyploidy and apomixis.
For future studies a dense sampling in the eastern distribution would demonstrate, if the other two subspecies (*songaricum* and *chinense*) can also be identified as residual populations between the prominent red-green polyploids.

5.3 Biogeographic implications: centre of origin

5.3.1 Hypericum perforatum

H. perforatum is widely distributed in Eurasia and introduced even world wide (Robson N., 2002). Populations occur continuously in lowland habitats, but also in mountain regions. To

our knowledge, fossil records of *H. perforatum* are lacking. However it is probably due to its wide ecological range, that *H. perforatum* could have survived in lowland steppe and tundra vegetation during cold stages of the ice-ages. This would mean that there was a potential for gene flow across the three european refuge areas (Iberian Peninsula, Apennine Peninsula and Balkan region) in glacials or at least a rather fast recovery of gene flow in interglacials. Rapid re-colonization is known for many species (see Taberlet P. *et al.*, 1998, Hewitt G., 1999 and Bhagwat S. *et al*., 2008)
This could lead one to the conclusion, that despite the potential gene flow two lineages were formed in glacial refugia. From the chloroplast data it can be assumed that the two lineages (red and green) are not recently formed in glacial refugia. This split may have started already earlier maybe even together with splitting from *H. maculatum,* during an radiation event Nürk N. *et al.* (2011) discovered in the late Miocene to early Pliocene of the 'core *Hypericum*'.
From our data, refugias in Iberian and Apennine Peninsula can not be rejected or confirmed for *H. perforatum*, due to rather scares sampling in those regions. But our data suggests, that *H. perforatum* had refugia in all three potential refuge areas of Europe.
In this study it was not possible to detect the centre of origin for the species *H. perforatum* as such. We could detect residual diploid sexual populations in Southern Europe reflecting two old 'species', which gave rise to the today polyploid populations recognized commonly as *H .perforatum*. The origin of the diploid sexual populations lays hidden rather in the East of Europe, but not connected with hybridization events of *H. maculatum* and *H attenuatum*.

5.3.2 Hypericum maculatum

In our chloroplast data we find four distinct lineages of *H. maculatum.* Three of those four are restricted to mountain regions in the Eastern Alps. An exception is haplotype HM03, which was also found in northern Germany. This may indicate that due to our focus in this study on *H. perforatum, H. maculatum* was undersampled. Nevertheless does this data imply a refugium or even origin in the Eastern Alps and a fast colonization afterwards to the West of Euope. Here further investigation with *H. maculatum* in focus is necessary.

6 Conclusions

The provided work wittnesses a very remarkable phase during the evolution of *H. perforatum*. On the one hand we can observe the splitting proccess from its sister species (*H. maculatum*, or maybe it is a fusion proccess, enforced by polyploids in both species?), on the other hand *H. perforatum* simultaneously splits into two new lineages (red and green gene pool). The question remains whether the two lineages in *H. perforatum* continue to rather melt together again, due to their polyploids or if they manage to keep separated, due to remaining sexual diploids. We can confirm that H. perforatum is an autopolyploid (refering to the two lineages as on species *H. perforatum*).
We were able to demonstrate that hybridization is common in *H. perforatum*, but interspecific hybridization is not responsible for the high morphological variability. Whereas intraspecific hybridization seems to be the main reason of morphological variability in *H. perforatum*.

7 Supplementary Material

species	subspecies	leaves	Inflorescence	Petals	Capsel vittae	sepals	stem lines
H. perforatum	ssp perforatum	usually petiolate; leaf base rounded to broadly cuneate; usually broadly to narrowly oblong or ovate, rarely elliptic or orbicular or obovate; not glaucous (grey) beneath; tertiary venetion lax or scarcely	not usually congested, branches relatively short, straight	laminar glands all pale to mostly black	lateral vittae linear, narrow or rarely swollen distally but not interupted, vesicles forming a regular row, not scattered irregularly		stem is 2-lined
	ssp veronense	sessile at least on main stem; leaf base rounded to cuneate; usually linear but sometimes narrowly triangular-lanceolate to linear-oblongor occasionally broadly ovate to elliptic or obovate, then short (c5-10mm long); paler but not glaucous (grey) beneath; tertiary venetion lax or scarcely	occasionally congested branches relatively short, straight or curved-scending	laminar glands all pale to rarely mostly black	lateral vittae swollen at base and/or interupted to punctiform and irregular (vesicular) not forming a regular distal row	acute	stem is 2-lined
	ssp songaricum	sessile at least on main stem; leaf base cordate to amplexicaul; oblong to oblong-ovate; often glaucous (grey) beneath; tertiary venetion lax or scarcely	not congested, branches relatively short, straight	laminar glands all pale	lateral vittae linear, narrow or slightly swollen overall, rarely interupted	finally acuminate	stem is 2-lined
	ssp chinense	petiolate; leaf base rather narrowly cuneate to subcordate-amplexicaul; narrowly elliptic to narrowly oblong or linear; maybe glaucous beaneath; tertiary venetion lax or scarcely	congested, branches relatively long, curved-ascending	laminar glands all black linear to punctiform or often absent	lateral vittae linear narrow or somewhat interupted or shortened, but not swollen		stem is 2-lined
H. maculatum	ssp maculatum	pale glands usually absent, but sometimes few to numerous; densely reticulate tertiary venation	narrow angle c. 30°, narrowly ascending	entire, without or rarely with 1-3 black marginal glands, laminar glands all or mostly black glands: dots/punctiform or proximally short streaks/striiform		broadly ovate to broadly elliptic or elliptic-oblong, apex entire or rarely slightly eroded-denticulate, laminar glands pale and black glands, punctiform	internodes always complete subsidiary lines
	ssp immaculatum	laminar gland dots pale, usually ±dense, occadionally with few black ; densely reticulate tertiary venation	narrow angle c. 30°, narrowly ascending	entire, without or rarely with 1-3 black marginal glands, laminar glands pale: linear to streaks/striiform		broadly ovate, rounded to obtuse or somewhat acute, apex entire or minutely denticulate, laminar glands pale, striiform and punctiform	internodes always complete subsidiary lines

	ssp obtusiusculum	laminar gland dots pale, dense to very sparse or absent; densely reticulate tertiary venation (somewhat laxer than in ssp maculatum	broad angle c. 50°, widely ascending, flowers: 25-30mm diam.	usually distally unilaterally creanate, marginal black glands in depressions?,with laminar glands pale and black or rarely only pale, mainly lines to streaks sometimes punctiform		broadly to narrowly ovate, usually finaly eroded denticulate or rarely entire	internodes complete or incomplete subsidiary lines
H. attenuatum		sessile; leaf base subcordate to cuneate; oblong to oblong-ovate; paler beneath; tertiary reticulation dense, but often obscure and lax		entire; laminar glands black, punctiform to striiform, scattered , marginal glands black	densely longitudinally vittate, occasionally with a few black longitudinal glandular streaks	ovate to lanceolate, acute to subacuminate	Stems 2-lined
hybrid	desetangsii nothossp carinthiacum nothoforma perforatiforme	often deflexed,laxly reticulate-veined denser than in perforatum but obscure, sparsely pale-gland-dotted; elliptic to rather narrowly oblong		without or relatively numerous dark glands, laminar glands all or some black, black marginal glands present	lateral vittae linear to absent, not shortened	broadly elliptic to linear lanceolate, acute to acuminate, +-eroded-denticulate to glandular-ciliate to subfimbriate or entire	internodes 2-lined (from leaf to bases) or with weaker and usually incomplete subsidiary lines (from between leaves)
	desetangsii nothossp carinthiacum nothoforma maculatiforme	relatively densely translucent reticulate-veined, without or almost without pale gland dots; ovate to broadly elliptic or elliptic-oblong, apex obtuse, not acute, base cuneate to rounded		without or relatively numerous dark glands, laminar glands all or some black, black marginal glands present	all linear, dorsal and lateral oblique, the latter regular or sometimes irregular	apex obtuse or subacute, entire or finely denticulate, without distinct hair-point, without or relatively numerous black glands	completely 4-lined
	desetangsii nothossp desetangsii	very laxly to rather densely translucent reticulate-veined, usually rather densely pale-gland-dotted (dense especially in upper leaves); ovate or broadly elliptic to narrowly oblong to linear		withblack laminar glands usually few, linear to elongate-punctiform, and some marginal	all linear and vertical or some lateral, oblique and linear to elongate-vesicular	narrowly oblong to linear-lanceolate, finally acuminate or apiculate, eroded-denticulate to entire, with fine hair-point, with sparse dark glands	internodes 2-lined (from leaf to bases), or with weaker and usually incomplete subsidiary lines (from between leaves)or complete 4-lined, usually partially 4-lined
	desetangsii nothossp balcanicum	relatively densely translucent reticulate-veined, without or almost without pale gland dots; apex obtuse or sometimes acute, base cuneate to cordate-amplexicaul		without black laminar glands, laminar glands all pale		apex obtuse or subacute, entire or finely denticulate, without distinct hair-point, without or relatively numerous dark glands	completely 4-lined

Suppl.Mat. 1 ***Summary of characters for subspecies of H. perforatum and H. maculatum and their hybrids*** (changed from Robson N. 2002).

P_{nb}	N_i	PV	Nei's gene diversity	DW
27	1	na	na	na
28	1	na	na	na
48	3	0.0392	0.0261	20.94
49	3	0.2598	0.1732	10.52
50	3	0.0245	0.0163	15.53
56	3	0.3186	0.2124	13.69
68	3	0.2892	0.1928	5.99
73	3	0.2745	0.183	13.79
74	3	0.174	0.116	10.59
76	3	0.2353	0.1569	8.32
78	3	0.201	0.134	11.71
85	3	0.1029	0.0686	6.77
91	5	0.0515	0.0221	5.63
96	7	0.4191	0.1763	19.28
99	3	0.2672	0.1781	11.23
102	3	0.2181	0.1454	8.39
110	3	0.1691	0.1127	9.02
116	3	0.0784	0.0523	15.17
124	3	0.1103	0.0735	13.88
131	3	na	na	na
132	3	na	na	na
133	3	0.2108	0.1405	8.96
137	3	0.1005	0.067	8.67
140	3	0.2181	0.1454	7.94
143	1	na	na	na
146	3	0.0319	0.0212	9.06
154	4	0.4363	0.2406	11.66
158	5	0.3578	0.1828	13.22
164	5	0.3039	0.152	23.61
175	5	0.1765	0.0907	8.13
177	3	0.0417	0.0278	12.29
180	5	0.3971	0.1922	13.23
181	5	0.4363	0.2142	16.26
182	5	0.2525	0.123	8.54
188	3	0.1716	0.1144	5.59
190	5	0.2255	0.1044	9.61
196	3	0.1814	0.1209	12.64
197	2	na	na	na
200	5	0.3235	0.1613	12.94
201	4	0.2475	0.136	10.42
207	1	na	na	na
210	3	0.2819	0.1879	9.83
216	3	0.2157	0.1438	8.84
218	3	0.2083	0.1389	9.54
223	3	0.1176	0.0784	6.41
227	3	0.098	0.0654	7.36
230	2	na	na	na
240	3	0.299	0.1993	8.77
300	1	na	na	na

305	3	0.1446	0.0964	8.39
307	3	0.049	0.0327	15.04
310	1	na	na	na
312	3	0.0931	0.0621	13.26
316	1	na	na	na
317	3	0.1225	0.0817	20.39
319	1	na	na	na
320	3	0.0441	0.0294	15.16
322	1	na	na	na
333	1	na	na	na
334	1	na	na	na
336	1	na	na	na
337	1	na	na	na
338	1	na	na	na
340	1	na	na	na
343	1	na	na	na
346	1	na	na	na
359	1	na	na	na
362	1	na	na	na
365	1	na	na	na
366	1	na	na	na
367	1	na	na	na
666	1	na	na	na
996	1	na	na	na
997	1	na	na	na
998	1	na	na	na
999	1	na	na	na

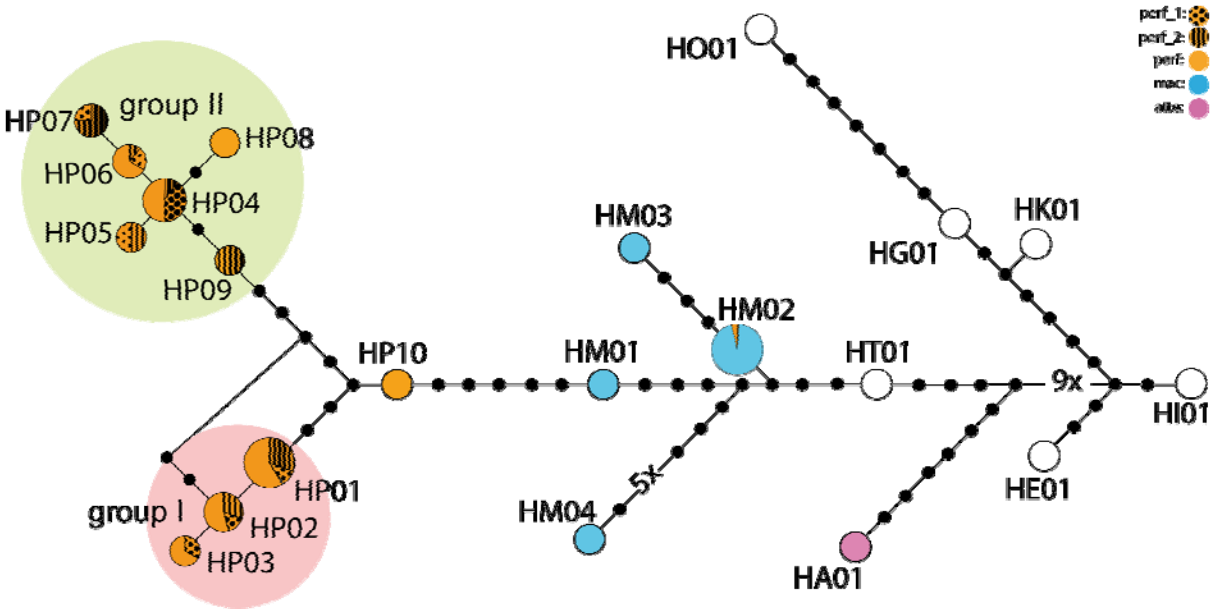
Suppl.Mat. 2 Population identification number (P_{nb}), number of individuals (N_i), proportion of variable markers (PV), Nei's gene diversity and the frequency down-weighted marker index (DW) for each population with three or more individuals.

cpDNA haplotype	GenBank accession codes
HA01	JF734846, JF734855, JF734868
HE01	JF734847, JF734856, JF734869
HG01	JF734848, JF734857, JF734870
HI01	JF734849, JF734858, JF734871
HK01	JF734848, JF734859, JF734870
HM01	JF734850, JF734860, JF734872
HM02	JF734850, JF734860, JF734873
HM03	JF734852, JF734860, JF734873
HM04	JF734850, JF734860, JF734874
HO01	JF734851, JF734861, JF734875
HP01	JF734852, JF734862, JF734872
HP02	JF734852, JF734862, JF734876
HP03	JF734852, JF734862, JF734877
HP04	JF734853, JF734863, JF734878
HP05	JF734853, JF734864, JF734878
HP06	JF734853, JF734863, JF734879
HP07	JF734853, JF734865, JF734879
HP08	JF734853, JF734866, JF734880
HP09	JF734853, JF734862, JF734878
HP10	JF734862, JF734850, JF734881
HT01	JF734854, JF734867, JF734882

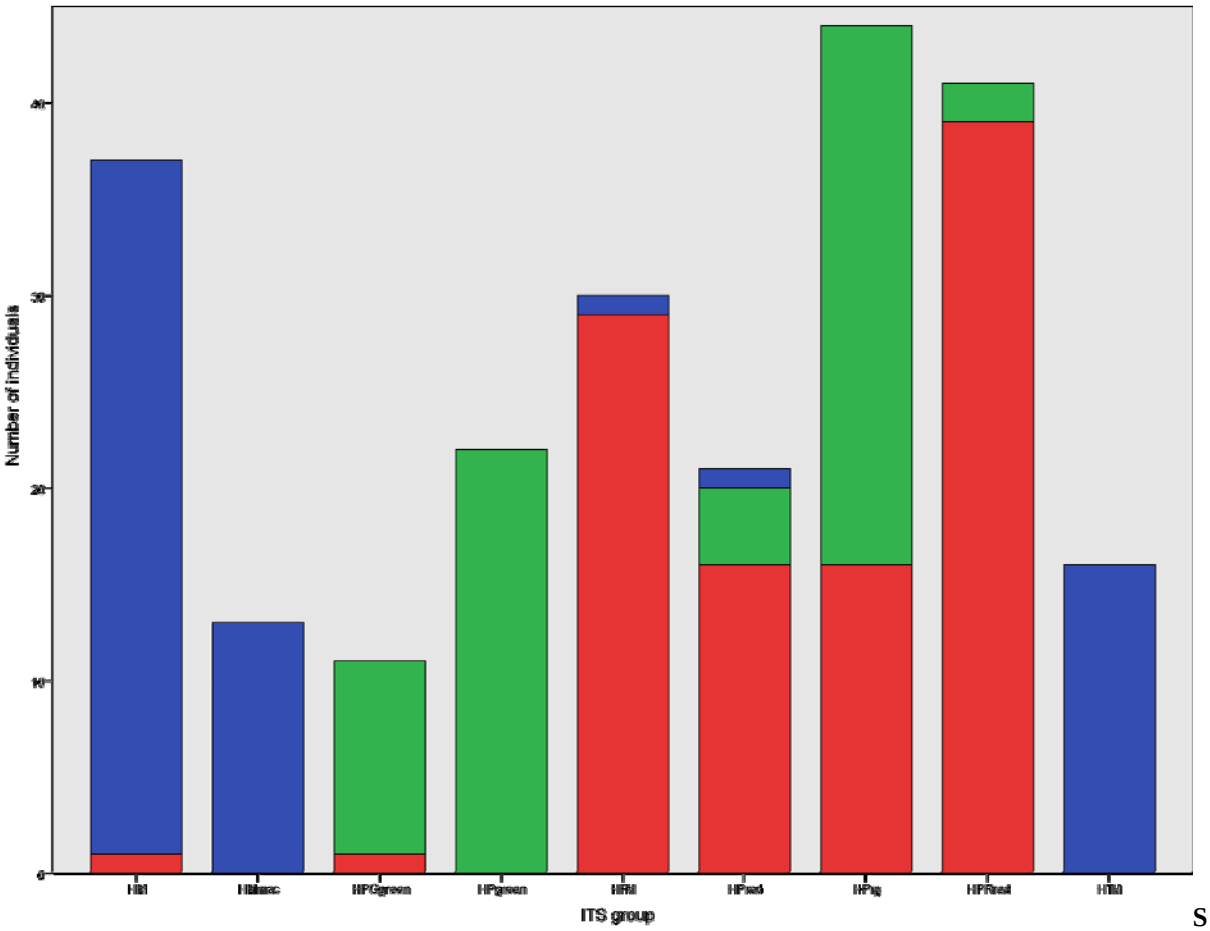
Suppl.Mat. 3 ***GenBank accession codes of chloroplast haplotypes***

ITS genotype	GenBank accession number
HAV	JN811118
HEI	Nürk N. *et al.* 2011
HEII	AY572993
HEIII	AY572995
HEIV	JN811119
HGI	Nürk N. *et al.* 2011
HGII	Nürk N. *et al.* 2011
HII	Nürk N. *et al.* 2011
HIII	Nürk N. *et al.* 2011
HKI	Nürk N. *et al.* 2011
HMI	JN811120
HMII	FJ793043
HMIII	JN811121
HMIV	JN811122
HMV	JN811123
HMVI	JN811124
HOI	Nürk N. *et al.* 2011
HPI	JN811125
HPII	JN811126
HPIII	JN811127
HPIV	JN811128
HPV	JN811129
HPIX	JN811130
HPX	JN811131
HPXI	JN811132
HPXII	JN811133
HPXIII	JN811134
HPXIV	JN811135
HPXV	JN811136
HTI	JN811137
HTII	JN811138
HUI	JN811139
HPVI	Nürk N. *et al.* 2011
HPVII	Nürk N. *et al.* 2011
HPVIII	Nürk N. *et al.* 2011

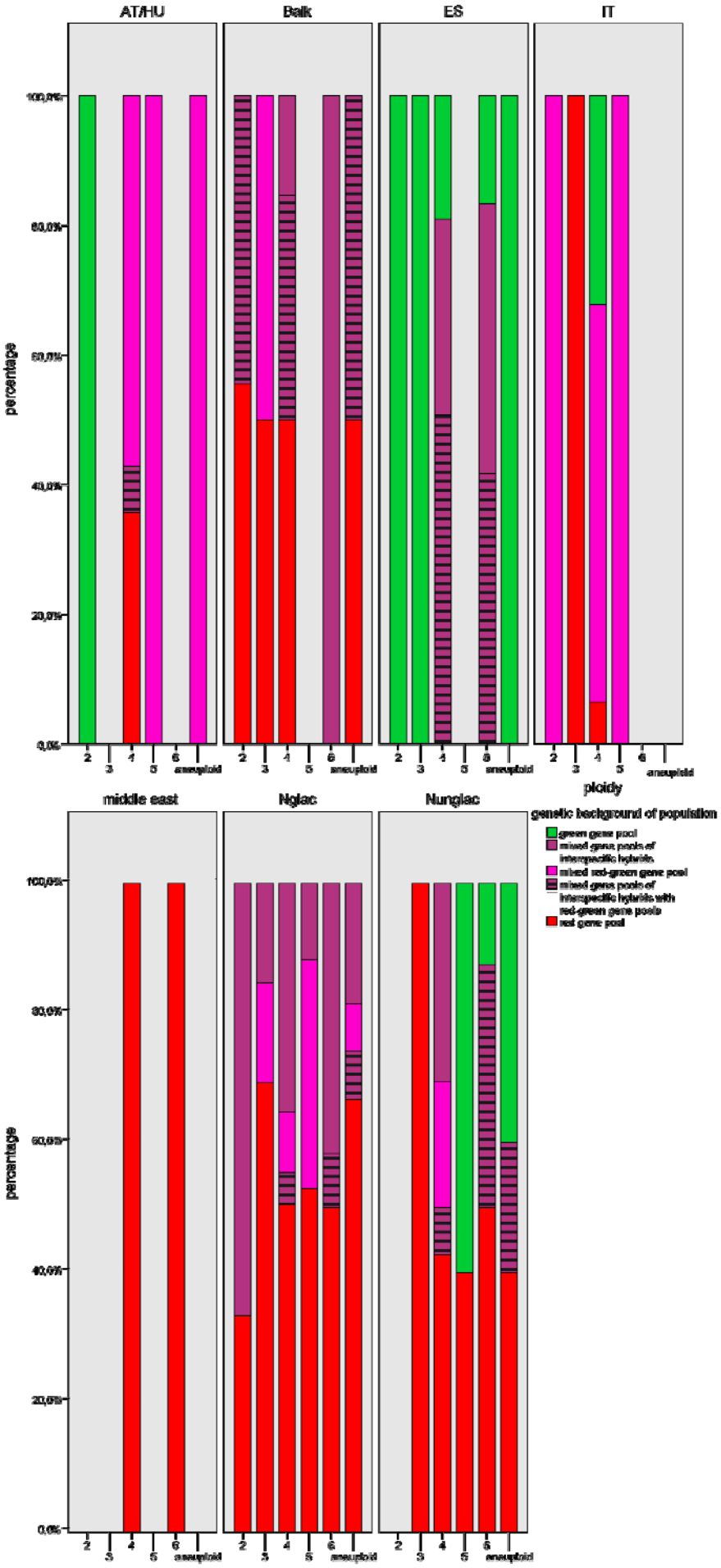
Suppl.Mat. 4 ***Genbank accession codes for 25 ITS genotypes***. Genbank accession codes for sequences acquired from collaboration with Nürk N. see Nürk N *et al.* 2011.



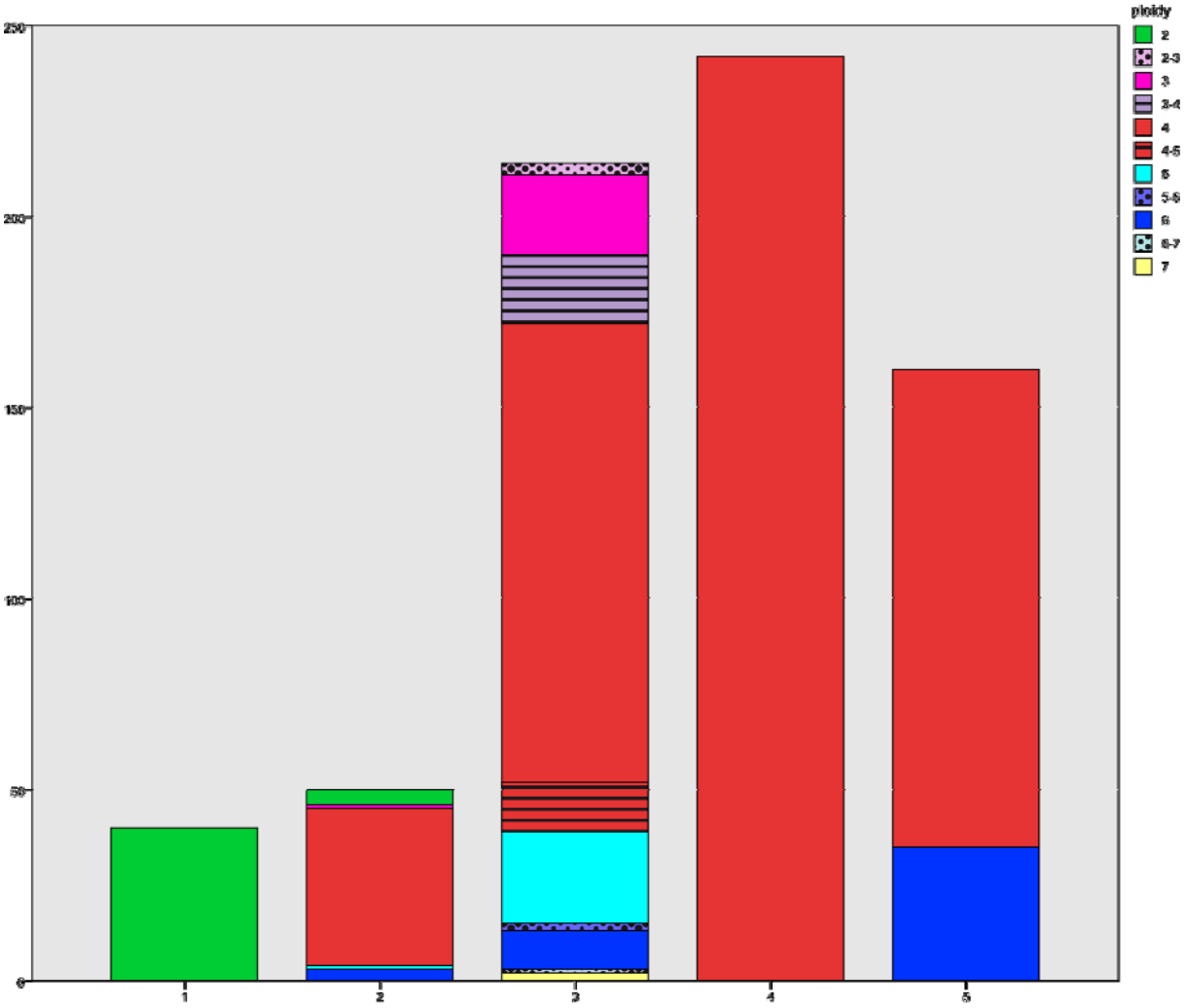
Suppl.Mat. 5 ***TCS network of 21 cpDNA haplotypes: each circle represents one haplotype***; **HP01-10**: *H. perforatum* and **HM01-HM04**: *H. maculatum*.coloures indicate species determination: **orange** indicates *H. perforatum*, **blue** colours indicate *H. maculatum* and **pink** indicates *H. attenuatum*. *H. perforatum* is also differentiated in individuals that were analysed by CDA: **stripes** are group1 individuals and **dots** represent individuals of group3; individuals with no pattern were not analysed by CDA. Circles with no colour indicate outgroup species that are close related to *H .perforatum*. **HT01**: *H. tetrapterum*, **HE01**: *H. erectum*, **HI01**: *H. pibariense*, **HK01:** *H. kamtschaticum*, **HG01**: *H. gracillimum* and **HO01**: *H. oliganthum*.



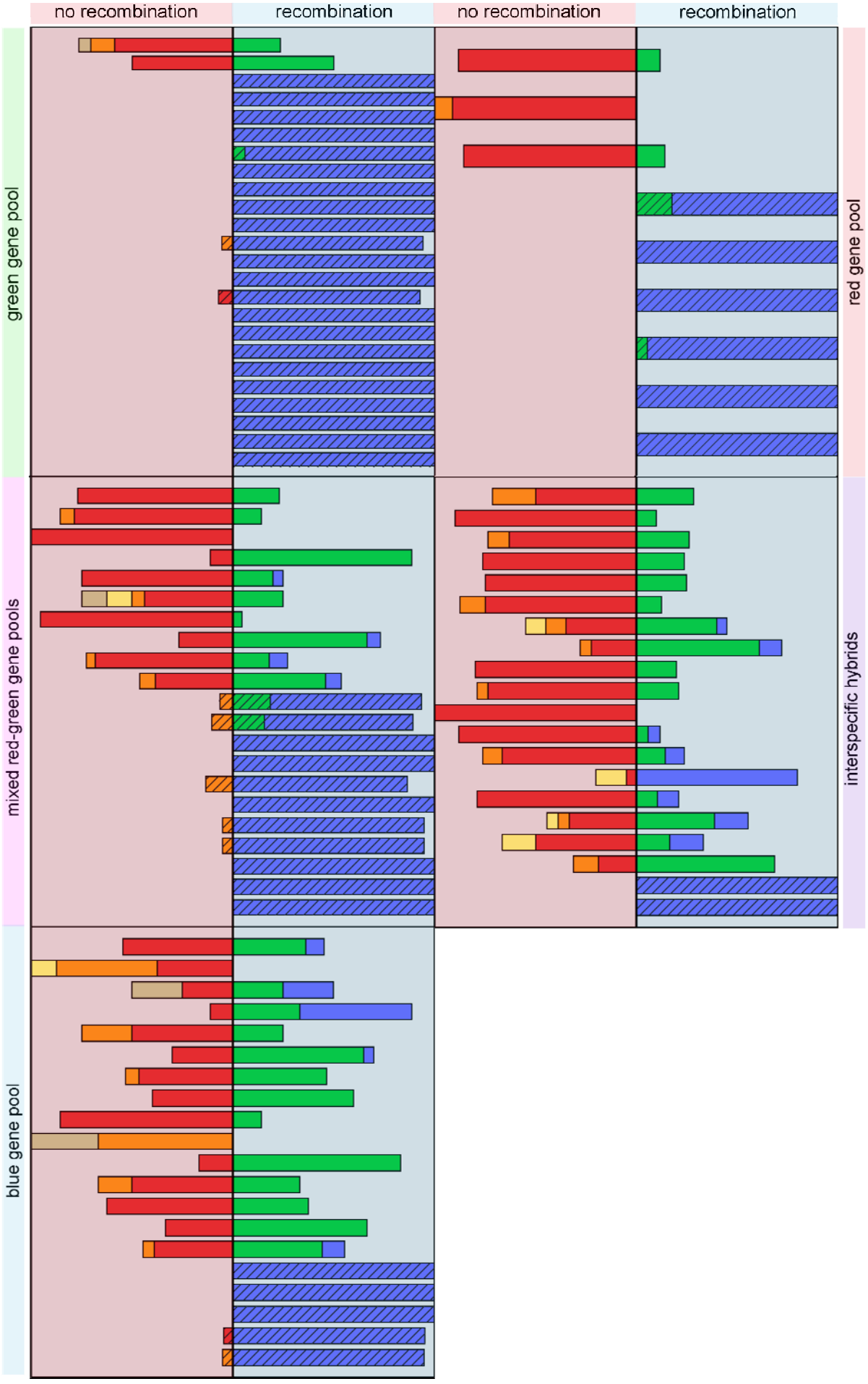
uppl.Mat. 6 Columns represent individuals carrying certain ITS genotypes (**HM**: individuals carried one genotype of *H. maculatum* gene pool HMI to HMVI; **HMmac**: individuals carried two genotypes of *H. maculatum* gene pool HMI to HMVI; **HPGgreen**: individuals carried two genotypes of *H. perforatum* green gene pool HPXI to HPXV; **Hpgreen**: individuals carried one genotype of *H. perforatum* green gene pool HPXI to HPXV; **HPM**: individuals carried two genotypes of *H. perforatum* and *H. maculatum* gene pool; **Hpred**: individuals carried one genotype of *H. perforatum* red gene pool HPI to HPX; **HPrg**: individuals carried two genotypes of *H. perforatum* red and green gene pool HPI to HPXV; **HPRred**: individuals carried two genotypes of *H. perforatum* red gene pool HPI to HPX; **HTM**: individuals carried two genotypes one of *H. maculatum* and one of *H. tetrapterum*) colours represent chloroplast sequence haplotypes: **red** represents haplotypes of group I, **green** represents haplotypes of group II and **blue** represents haplotypes characteristic for *H. maculatum*.


AT/HU
Balk
ES
IT
middle east
Nglac
Nunglac
percentage
ploidy
aneuploid
genetic background of population
green gene pool
mixed gene pools of interspecific hybrids
mixed red-green gene pool
mixed gene pools of interspecific hybrids with red-green gene pools
red gene pool

Suppl.Mat. 7 Ploidy levels and genetic background of H. perforatum in seven different geographical regions. Seven graphs correspond to seven geographical regions (**AT/HU**: unglaciated parts of Austria and Hungary, **Balk**: unglaciated parts in the Balkan region, **ES**: unglaciated parts on the Iberian Peninsula, **IT**: unglaciated parts on the Apennine Peninsula, **middle East**: region between Saudi Arabia and Pakistan, **Nglac:** glaciated regions in North Europe **Nunglac**: unglaciated regions in North Europe).Y- axis explains in percent how many individuals had the same genetic background (green, red or mixed gene pools). Each bar represents the sum of individuals with a certain ploidy level.



Suppl.Mat. 8 Amount of cytotypes in different types of populations; **population group1**: only diploid individuals; **population group2**: population with diploids and mixed other ploidy levels; **population group3**: mixed ploidy levels with triploid and pentaploid individuals; **population group4**: includes only populations with tetraploids exclusively and **population group5** includes only tetraploids and hexaploids


no recombination
recombination
no recombination
recombination
green gene pool
red gene pool
mixed red-green gene pools
interspecific hybrids
blue gene pool

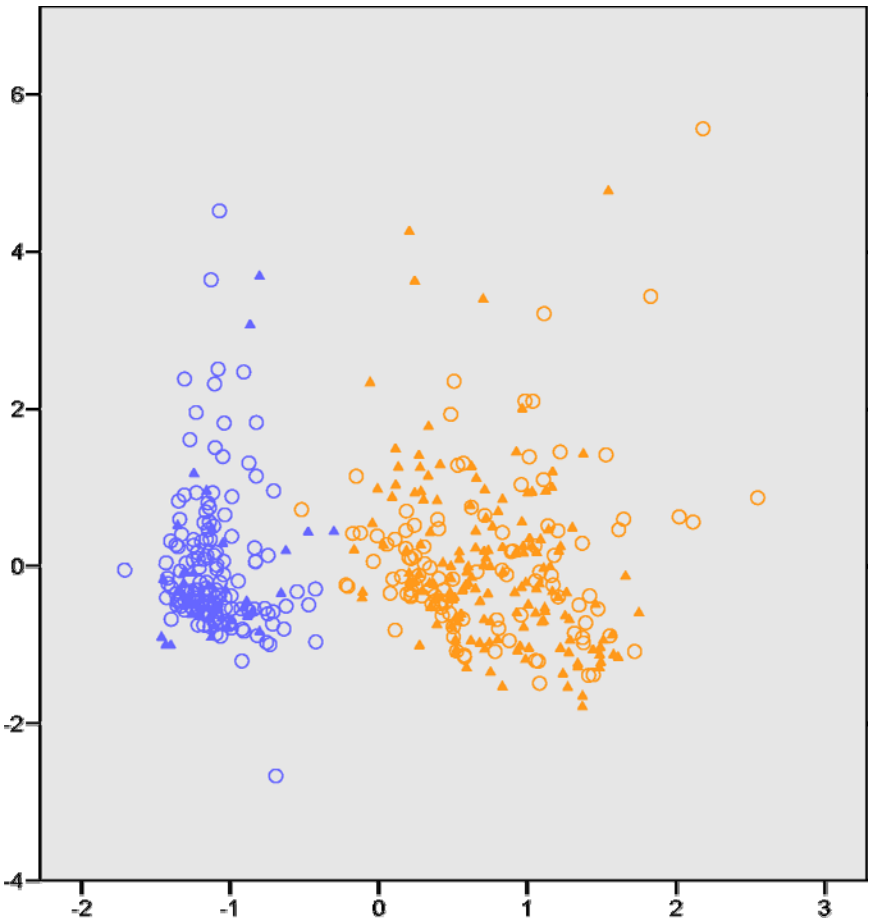
Suppl.Mat. 9 Six different pathways for reproduction are ilustrated in the way that one bar represents 100 % of the seeds analysed for one individual. **Blue** represents sexually produced seeds, **green** apomeiotic (but double fertilized) seeds, **red** apomeiotic and pseudogamous seeds, **orange** seeds that were produced apomeiotic and autonomous, **yellow** seeds of parthenogenetic and pseudogamous origin and **beige** seeds that are of parthenogenetic and autonomous origin. Striped bars stand for diploid individuals. Grouped according to genetic background of individual, from green gene pool, red gene pool, mixed red-green gene pools, mixed interspecific gene pools or from blue gene pool.

		H. maculatum	***H. perforatum***	***H. perforatum* group1**	***H. perforatum* group3**
c	min	0.00000	0.00000	0.00000	0.00000
	max	10.00000	32.25000	32.50000	17.00000
	average	**1.38352**	**7.27632**	**9.64000**	**5.18085**
	sd	1.78784	5.25516	5.99516	3.32108
a	min	1.38925	1.41873	1.82679	1.41873
	max	2.79018	4.90533	4.90533	3.42503
	average	**1.96218**	**2.90729**	**3.42261**	**2.45045**
	sd	0.26951	0.71640	0.62955	0.41609
b	min	0.00000	0.00000	0.00000	0.00000
	max	0.05177	0.06159	0.06159	0.05285
	average	**0.00488**	**0.00686**	**0.00742**	**0.00636**
	sd	0.00838	0.00956	0.01068	0.00845
f	min	0.00000	0.00000	0.00000	0.00000
	max	12.00000	54.00000	54.00000	20.00000
	average	**1.25568**	**8.40977**	**11.17600**	**5.95745**
	sd	2.01068	6.78409	8.03326	4.13154
d	min	1.55527	1.86265	2.33280	1.86265
	max	4.20482	6.36667	6.36667	5.41951
	average	**2.65655**	**3.61214**	**4.14706**	**3.13793**
	sd	0.45696	0.90056	0.83943	0.65502
e	min	0.00000	0.00000	0.00000	0.00000
	max	0.12532	0.24493	0.24493	0.10917
	average	**0.00852**	**0.01070**	**0.01163**	**0.00988**
	sd	0.02120	0.02154	0.02647	0.01601
i	min	25.00000	25.00000	25.00000	29.00000
	max	59.00000	70.00000	70.00000	65.20000
	average	**41.25354**	**48.00849**	**47.33317**	**48.62456**
	sd	9.53624	7.22079	7.54578	6.88653
l	min	1.50000	1.47137	1.53591	1.47137
	max	3.00214	3.39007	3.34018	3.39007
	average	**2.07864**	**2.14288**	**2.16544**	**2.12288**
	sd	0.27914	0.30718	0.32673	0.28844
o	min	1.22595	2.93478	3.64771	2.93478
	max	4.04808	11.17689	9.24891	11.17689
	average	**2.36056**	**5.30628**	**5.80928**	**4.86035**
	sd	0.43541	1.24737	1.18582	1.12849
q	min	1.08859	1.25712	1.50975	1.25712
	max	2.41137	2.71349	2.63205	2.71349
	average	**1.56981**	**1.78533**	**1.92773**	**1.68689**
	sd	0.22184	0.29578	0.28251	0.26429

Suppl.Mat. 10 Minimum (min); maximum (max); average; and standard deviation of continous characters in *H. maculatum*, H .perforatum (entire species) and in the two morphotypes in *H. perforatum* group1 (green gene pool) and group3 (red gene pool). Character abbrivations are explained in extra material 6 on CD.

	H. maculatum	*H. perforatum*	*H. perforatum* group1	*H. perforatum* group3
g	***1/2/3***	***1/2/3***	***1/2***	***1/2/3***
h	0/1	0/1	0/1	0/1
j	1/2/3	2/3	2/3	2/3
k	***0/1/2***	***0/1***	***0/1***	***0/1***
m	1/2	1/2	1/2	1/2
n	1/2/3	1/2/3	1/2/3	1/2/3
p	***2/4/5/6***	***1/2/3/4***	***1/2/3/4***	***1/2/3/4***
r	1/4	1/2/3	1/2/3	1/3

Suppl.Mat. 11 types of discrete characters measured in *H. maculatum*, *H .perforatum* (entire species) and in the two morphotypes in *H. perforatum* group1 (green gene pool) and group3 (red gene pool). Character abbrivations are explained in extra material 6 on CD.



Suppl.Mat. 12 Graph shows the results of the PCA of 15 morphological characters with the x-axis explaining 36.2 % and the y-axis explaining 12.2 % of the total variation. In **orange** are individuals determined as *H. perforatum* and in **blue** are individuals determined as *H. maculatum*; **triangles** represent individuals sampled in the field and **circles** stand for individuals grown in the botanical garden Heidelberg.

8 References

Abbott R, Brochmann C 2003. History and evolution of the arctic flora: in the footsteps of Eric Hultén. Molecular Ecology 12: 299-313

Alsos I, Eidesen P, Ehrich D, Skrede I, Westergaard K, Jacobsen, Landvik J, 2007. Frequent Long-Distance Plant Colonization in the Changing Arctic. Science 316: 1606-1609.

APG III, 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. Botanical Journal of the Linnean Society 161: 105–121.

Arnold M, Bulgar M, Burke J, Hempel A and Williams J 1999. How Low Can You Go and Still Be Important? Ecology 80:2: 371-381.

Avise J, 2000. In: Phylogeography: The History and Formation of Species. Harvard University Press, Cambridge, MA.

Barcaccia G, Arzenton F, Sharbel T, Varotto S, Parrini P and Lucchin M 2006. Genetic diversity and reproductive biology in ecotypes of the facultative apomict *Hypericum* perforatum L. Heredity 96: 322–334.

Barcaccia G, Bäumlein H and Sharbel T 2007. Apomixis in St. John's wort (*Hypericum* perforatum L.): an overview and glimpse towards the future. Hörandl E., Grossniklaus U., van Dijk P. and Sharbel T. (ed.) In: Apomixis: Evolution, Mechanisms and Perspectives. International Association for Plant Taxonomy - Koeltz Scientific Books. pp 259-280.

Barringer B 2007. Polyploidy and Self-fertilization in flowering plants. American Journal of Botany 94:9: 1527-1533

Barton N 2001. The role of hybridization in evolution. Molecular Ecology 10: 551-568.

Bell G 1982. The masterpiece of nature. Croom Helm, London.

Bhagwat S and Willis K 2008. Species persistence in northerly glacial refugia of Europe: a matter of chance or biogeographical traits? Journal of Biogeography 35: 464-482.

Bicknell R, Koltunow A 2004. Understanding apomixis: Recent advances and remaining conundrums. The Plant Cell 16: S228-S245.

Bierzychudek P. 1985. Patterns in Plant Parthenogenesis. Experientia 41:10 1255-1264.

Birks H 2008. The late-quaternary history of arctic and alpine plants. Plant Ecology & Diversity 1: 135-146.

Bonin A, Bellemain E, Eidesen P, Pompanon F, Brochmann C, and Taberlet P 2004. How to track and assess genotyping errors in population genetics studies. Molecular Ecology, 13(11):3261-3273

Bretagnolle F. and Lumaret R. 1995. Bilateral polyploidization in *Dactylis glomerata* L. Subsp. lusitanica: occurence, morphological and genetic characteristics of first polyploids. Euphytica 84: 197-207

Bretagnolle F. and Thompson J. 1995 Gametes with the somatic chromosome number: mechanisms of their formation and role in th evolution of autopolyploid plants. New Phytologist 129: 1-22.

Brutovská R, Cellárová E and Schubert I 2000a. Cytogenetic characterization of three *Hypericum* species by in situ hybridization. Theoretical and Applied Genetics 101: 46–50.

Brutovská R Kušniriková P, Bogyiová E and Cellárová E. 2000b. Karyotype analysis of *Hypericum* perforatum L. Biologia Plantarum 43: 133–136.

Ciccarelli D, Andreucci A. and Pagni A. 2001. Translucent glands and secretory canals in *Hypericum* perforatum L. (Hypericaceae): morphological, anatomical and histochemical studies during the course of ontogenesis. Annals of Botany 88: 637–644.

Clement M, Posada D, and Crandall K 2000. TCS: a computer program to estimate gene genealogies. Mol Ecol 9,1657-1659

Comes H, and Kadereit J 1998. The effect of Quaternary climatic changes on plant distribution and evolution. Trends in Plant Sciences 3: 432- 438.

Comes H, and Kadereit J. 2003. Spatial and temporal patterns in the evolution of the flora of the European alpine system. Taxon 52: 451-462.

Crockett S., Schaneberg B. and Khan I. 2005. Phytochemical profiling of New and Old World *Hypericum* (St. John's wort) species. Phytochemical Analysis 16: 479–485.

Curtis J and Lersten N 1990. Internal secretory structures in *Hypericum* (Clusiaceae): *H. perforatum* L. and H. balearicum L. New Phytologist 114: 571–580.

Doležel J., Sgorbati S., Lucretti S., 1992 Comparison of three DNA fluorochromes for flow cytometric estimation of nuclear DNA content in plants. Physiologia Plantarum 85: 625-631

Doležel J., Greilhuber J., Lucretti S., Meister A., Lysák M., Nardi L. And Obermayer R. 1998. Plant Genome Size Estimation by Flow Cytometry: Inter-laboratory Comparison. Annals of Botany 82:1 17-26

Doyle, J. J. and J. L. Doyle 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bulletin, 19:11-15.

Doyle, J. J. 1991. DNA protocols for plants. pp. 283-293 in: G. Hewitt, A. W. B. Johnson, and J. P. W. Young (eds.), Molecular Techniques in Taxonomy. Cell Biology Vol. 57.

D'Souza T. and Michiels N. 2010. The Costs and Benefits of Occasional Sex: Theoretical Predictions and a Case Study. Journal of Heredity. 101 (Supplement I): S34-S41

Ehrich, D. 2006. AFLPdat: a collection of R functions for convenient handling of AFLP data. Mol. Ecol. Notes, 6:603-604.

Ehrich, D., M. Gaudeul, A. Assefa, M. Koch, K. Mummenhof, S. Nemomissa, Intrabiodiv. Consortium, and C. Brochmann 2007. Genetic consequences of Pleistocene range shifts: contrast between the Arctic, the Alps and the East African mountains. Molecular Ecology, 16:2542-2559.

Evanno, G., S. Regnaut, and J. Goudet 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. Molecular Ecology, 14(8):2611-2620.

Falush, D., M. Stephens, and J. K. Pritchard 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. Molecular Ecology Notes, 7(4):574-578.

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

Gong W, Chuan C, Dobes C, Fu C.-X., Koch M. 2008. Phylogeography of a living fossil: Pleistocene glaciations forced Ginkgo biloba L. (Ginkgoaceae) into two refuge areas in China with limited subsequent postglacial expansion. Mol. Phyl. Evol. 48: 1094-1105

Grant V. 1981. Plant speciation. New York, USA: Columbia University Press.

Green J. 1884. On the organs of secretions in the Hypericaceae. Journal of the Linnean Society of London (Botany) 20: 451–464.

Greilhuber J., Doležel J., Lysak M. and Bennett M. 2005. The Origin, Evolution and Proposed Stabilization of the Terms 'Genome Size' and 'C-Value' to Describe Nuclear DNA Contents. Annals of Botany 95: 255-260

Guerra M. 2008. Chromosome numbers in plant cytotaxonomy: concepts and implications. Cytogenetic and Genome Research 120: 339-350

Hamilton W. 1980. Sex versus non-sex versus parasite. Oikos 35: 282-290.

Hamrick J. and Godt M. 1996. Effects of Life History Traits on Genetic Diversity in Plant Species. Phil. Trans. R. Soc. London. B 351: 1291-1298

Harlan J. and deWet J. 1975. On Ö.Winge and a prayer: the origin of polyploidy. Botanical Review 41: 361-390.

Harris S. A., Ingram R. 1991. Chloroplast DNA and biosystematics: The effects of interspecific diversity and plastid transmission. TAXON 40: 393-412.

Heinze, B. 2007. A database of PCR primers for the chloroplast genomes of higher plants. Plant Methods, 3:4.

Hewitt G M. 1999. Post-glacial re-colonization of European biota. Biol. J. Linn. Soc. 68: 87-112.

Hewitt G. M. 2000. The genetic legacy of the Quaternary ice ages. Nature 405: 907-913.

Hewitt G. M. 2004. Genetic consequences of climatic oscillations in the Quaternary. Phil. Trans. R. Soc. Lond. B. 359: 183-195.

Hohmann N. 2010. Zytogenetik und Evolution eines Apomikten: *Hypericum* perforatum L. Bachlor thesis at University Heidelberg

Holsinger K. 2000. Reproductive systems and evolution in vascular plants. Proc. Natl. Acad. Sci. USA 97:13: 7037-7042.

Huck S, Büdel B, Kadereit J and Prinzten C. 2009. Range-wide phylogeography of the European temperate-montane herbaceous plant Meum athamanticum Jacq.: evidence for periglacial persistence. Journal of Biogeography 36: 1588-1599.

Husband B. 2004. The role of triploid hybrids in the evolutionary dynamics of mixed-ploidy populations. Biological Journal of the Linnean Society 82: 537-546.

Jordon-Thaden I, Hase I, Al-Shehbaz I, and Koch M. 2010. Molecular phylogeny and systematics of the genus Draba (Brassicaceae) and identification of its closest related genera. Molecular Phylogenetics and Evolution 55 (2): 524-540.

Kadereit, J.W., Westberg, E. 2007: Determinants of phylogeographic structure: a comparative study of seven coastal flowering plant species across their European range. – Watsonia 26: 229 – 238.

Kearney M. 2005. Hybridization, glaciation and geographical parthenogenesis. Trends in Ecology and Evolution 20: 495-502

Koch M, Bernhardt K.-G. 2004: Comparative Biogeography of Different Cytotypes of the European Annual Microthlaspi perfoliatum (Brassicaceae): Refugia, Diversity Centres and Post Glacial Colonization. American Journal of Botany 91 (1): 115-124.

Kogi M . 1984. A karyomorphological study of the genus *Hypericum* (Hypericaceae) in Japan. Journal of Plant Research 97: 333–343.

Kondrashov AS. 1988. Deleterious mutations and the evolution of sexual reproduction. Nature 336 435-440.

Landau A. 2009. Geographic differentiation of Central European *Hypericum* perforatum L. by means of AFLP analysis. Diploma thesis at Heidelberg University.

Leins P. and Erbar C. 2008. Blüte und Frucht: Aspekte der Morphologie, Entwicklungsgeschichte, Phylogenie, Funktion und Ökologie. Auflage: 2. Stuttgart: Schweizerbart Science Publishers.

Leitch I. and Bennett M. 1997. Polyploidy in angiosperms. Trends Plant Sci. 2: 470-476.

Lewis W. 1980. Polyploidy in species populaions. In: Lewis W. (ed.) Polyploidy: biological relevance. New York, USA: Plenum Press, 103-144.

Lewis W. 1987. The cost of sex. In S. C. Stearns (ed.) The evolution of sex and its consequences. Birkhauser Verlag, Basel, Switzerland. Pp. 33-57.

Lundmark M., Saura A. 2006. Asexuality alone does not explain the success of clonal forms in insects with geographical parthenogenesis. Hereditas 143: 24-33

Maheshwari P. 1950. An introduction to the embryology of the angiosperms. McGraw-Hill, New York.

Mallet J., 1995 A species definition for the modern synthesis. TRENDS in Ecology and Evolution. Vol.10 No.7 294-299

Mallet J., 2005. Hybridization as an invasion of the genome. TRENDS in Ecology and Evolution. Vol.20 No.5 229-237.

Manel S, Poncet B, Legendre P, Gugerili F, and Holderegger R. 2010. Common factors drive adaptive genetic variation at different scales in Arabis alpina. Molecular Ecology 19: 3824–3835.

Mártonfi P., Repcák M. and Zanvit P. 2006. Secondary metabolites variation in *Hypericum maculatum* and its relatives. Biochemical Systematics and Ecology 34: 56–59.

Masterson J. 1994. Stomatal size in fossil plants: evidence for polyploidy in majority of angiosperms. Science 264: 421-423

Matzk, F., Meister A and Schubert I. 2000. An efficient screen for reproductive pathways using mature seeds of monocots and dicots. The Plant Journal:21(1): 97-108.

Matzk, F., Meister, A., Brutovská, R., Schubert, I. 2001. Reconstruction of reproductive diversity in *Hypericum* perforatum L. opens novel strategies to manage apomixis. Plant J. 26(3)275-282.

Matzk, F., Hammer K., Schubert I., 2003. Coevolution of apomixis and genome size within the genus *Hypericum*. Sex Plant Reprod. 16:51-58

Möller M, Gao LM, Mill RR, Li DZ, Hollingworth ML, Gibby M. 2007. Morphometric analysis of the Taxus wallichiana-complex based on herbarium material. Bot. J. Linn. Soc. 155: 307-335.

Molins Puente M. 2011. unpublished

Muller H. 1932. Some genetic aspects of sex. American Naturalist 8: 118-138.

Müller K, Quandt D, Müller J, Neinhuis C. 2005. PhyDE 0.9971: Phylogenetic Data Editor. www.phyde.de.

Nahrstedt A. and Butterweck V. 2010. Lessons Learned from Herbal Medicinal Products: The Example of St. John's Wort. Journal of Natural Products. 73:5: 1015-1021

Noak K, 1939. Über *Hypericum*-Kreuzungen VI. Fortpflanzungsverhältnisse und Bastarde von *Hypericum perforatum* L. Zeitschrift für induktive Abstammungs- und Vererbungslehre 76: 569-601

Noirot M., Couvet D, Hamon S. 1997. Main role od self-pollination rate on reproductive allocations in pseudogamous apomicts. Theor. Appl. Genet. 95:479-483.

Nürk N and Blattner F. 2010. Cladistic analysis of morphological characters in *Hypericum* (Hypericaceae). Taxon 59:5 1495-1507.

Nürk N., Madriñán S., Carine M., Chase M. and Blattner F. 2011 Molecular phylogeny and character evolution in St. John's wort (*Hypericum*). Molecular Phylogenetics and Evolution submitted

Otto S. and Whitton J. 2000. Polyploid incidence and evolution.Annu. Rev. Genet. 34: 401-437.

Otto S. and Lenormand T. 2002. Resolving the paradox of sex and recombination. Nature Reviews Genetics 3: 252-261.

Ownbey M. 1950. Natural hybridization and amphiploidy in the genus *Tragopogon*. American Journal of Botany 40: 788-796

Parolly G, Nordt B, Bleeker W, and Mummenhoff K. 2010. Heldreichia Boiss. (Brassicaceae) revisited: A morphological and molecular study. Taxon 59: 187-202.

Pauwels M, Saumitou-Laprade P, Holl AC, Petit D, Bonnin I. 2005. Multiple origin of metallicolous populations of the pseudometallophyte Arabidopsis halleri (Brassicaceae) in central Europe: the cpDNA testimony. Mol. Ecol. 14: 4403-4414.

Peck J. and Waxman D. 2000. What's wrong with a little sex? J. Evol. Biol. 13: 63-69.

Petit RJ, Bialozyt R.; Brewer S.; Cheddadi R.; Comps B., 2001. From spatial patterns of genetic diversity to postglacial migration processes in forest trees. In: Integrating ecology and evolution in a spatial context. (ed. Silvertown J.; Antonovics J.). Blackwell Sciences, Oxford, pp. 295-318.

Petit R, Aguinagalde I, De Beaulieu J.-L, Bittkau C, Brewer S, Cheddadi R, Ennos R, et al. 2003. Glacial refugia: hotspots but not melting pots of genetic diversity. Science 300: 1563-1565.

Petit RJ, Hampe A, Cheddadi R. 2005. Climate changes and tree phylogeography in the Mediterranean. Taxon 54: 877- 885.

Pritchard, J. K., M. Stephens, and P. Donnelly 2000. Inference of Population Structure Using Multilocus Genotype Data. Genetics, 155(2):945-959.

Qu L, Widrlechner M and Rigby S. 2010. Analysis of breeding systems, ploidy, and the role of hexaploids in three *Hypericum perforatum* L. populations. Industrial Crops and Products 32: 1–6.

Ramsey J., Schemske D. 1998. pathways, mechanisms, and rates of polyploid formation in flowering plants. Annual Review of Ecology and Systematics 29: 467-501

Randi E, 2007. Phylogeography of South European mammals. In: Weiss S, Ferrand N (eds). Phylogeography of Southern European Refugia. Springer, Wien, Heidelberg; pp 101–126.

Rebernik, C. A., H. Weiss-Schneeweiss, G. M. Schneeweiss, P. Schönswetter, R. Obermeayer, J. L. Villaseñor, T. F. Stuessy. 2010. Quaternary range dynamics and polyploid evolution in an arid brushland plant species (Melampodium cinereum, Asteraceae). Mol. Phyl. Evol. 54: 594-606

Richards A. 1997. Plant breeding systems, 2nd ed. (London: Chapman and Hall)

Richards A., 2003. Apomixis in flowering plants: an overview. Phil. Trans. R. Soc. Lond. B 358: 1085-1093

Rieseberg L. and Ellstrand N. 1993. What can molecular and morphological markers tell us about plant hybridization? Critical Reviews in Plant Sciences. 12:3: 213-241.

Rieseberg L. and Willis J. 2007. Plant speciation. *Science* 317: 910-914.

Robson N. and Adams P. 1968. Chromosome numbers in *Hypericum* and related genera. Brittonia 20: 95–106.

Robson N. 1977. Studies in the genus *Hypericum* L. (Guttiferae): 1. Infrageneric classification. Bulletin of the British Museum (Natural History), Botany 5: 291–355.

Robson N. 1981, Studies in the genus *Hypericum* L. (Guttiferae). 2. Characters of the genus. Bull. Br. Mus. Nat. Hist. , 8, 55–226.

Robson, N. 2002. Studies in the genus *Hypericum* L. (Guttiferae) 4(2). Section 9. *Hypericum* senso lato (part 2): subsection 1. *Hypericum* series 1. *Hypericum*. Bull. Nat. Hist. Mus. Lond. Bot., 32(2):61-123.

Robson, N. 2003. *Hypericum* botany. In Ernst E. (ed.) *Hypericum*: the genus *Hypericum* Taylor and Francis: London, 1-22

Rosenberg, N, Pritchard J., Weber J., Cann H., Kidd K., Zhivotovsky L. and Feldman M. 2002. Genetic Structure of Human Populations. Science, 298(5602):2381-2385.

Schallau A, Arzenton F, Johnston A, Hahnel U, Koszegi D, Blattner F, Altschmied L, Haberer G, Barcaccia G and Bäumlein H 2010. Identification and genetic analysis of the APOSPORY locus in *Hypericum perforatum* L. The Plant Journal 62: 773–784.

Schiemann K, T. Tyler, B. Widén. 2000. Allozyme diversity in relation to geographic distribution and population size inLathyrus vernus (L.) Bernh. (Fabaceae). Pl. Syst. Evol. 225: 119-132.

Schmitt T. 2009. Biogeographical and evolutionary importance of the European high mountain systems. Frontiers in Zoology 6: e9. doi: 10.1186/1742-9994-6-9.

Schönswetter P., Tribsch A., H. Niklfeld. 2004. Amplified Fragment Length Polymorphism (AFLP) reveals no genetic divergence of the Eastern Alpine endemic Oxytropis campestris subsp. tiroliensis (Fabaceae) from widespread subsp. campestris. Pl. Syst. Evol. 244: 245-255.

Schönswetter P, Stehlik I., Holderegger R., Tribsch A. 2005. Molecular evidence for glacial refugia of mountain plants in the European Alps. – Molecular Ecology, 14, 3547–3555.

Schubert I. and Lysak M. 2011. Interpretation of karyotype evolution should consider chromosome structural constraints. Trends in Genetics 27:6: 207-216.

Sharbel T. F., Haubold, B., Mitchell-Olds T. 2000. Genetic isolation by distance in Arabidopsis thaliana: biogeography and postglacial colonization of Europe. Mol. Ecol. 9: 2109–2118.

Shaw J., Lickey E.B., Beck J. T., Farmer S. B., Liu W., Miller J, Siripun K. C., Winder C. T., Schilling E. E., Small R. L. 2005. The Tortoise and The Hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. Am. J. Bot. 92(1): 142-166.

Sirvent T., Krasnoff S. and Gibson D. 2003. Induction of hypericins and hyperforins in *Hypericum perforatum* in response to damage by herbivores. Journal of Chemical Ecology 29: 2667–2681.

Stebbins G. 1985. Polyploidy, hybridization, and the invasion of new habitats. Annals of Missouri Botanical Garden 72:2:824-832.

Stenberg P., Lundmark M., Knutelski S., Saura A.. 2003. Evolution of clonality and polyploidy in weevil system. Molecular Biology and Evolution 20: 1626-1632.

Soelberg J, Jorgensen L. and Jager A. 2007. Hyperforin accumulates in the translucent glands of *Hypericum perforatum*. Annals of Botany 99: 1097–1100.

Soltis D. and Soltis P. 1993. Molecular data and the dynamic nature of polyploidy. Critical reviews in Plant Sciences. 12:3: 243-273.

Soltis D, Soltis P. 1999. Polyploidy: recurrent formation and genome evolution. TREE 14: 328-352.

Soltis D., and Tate J. 2003. Advances in the study of polyploidy since plant speciation. New Phytologist 161: 173-191.

Soltis D, Morris A, Lachlan J, Manos P and Soltis P. 2006. Comparative phylogeography of unglaciated eastern North America. Mol Ecol. 15 , 4261–4293.

Soltis D, Mavrodiev E.; Doyle J.; Rauscher J.; Soltis P. 2008. ITS and ETS sequence data and phylogeny reconstruction in allopolyploids and hybrids. Systematic Botany 33: 1 7-20

Soltis D., Albert A., Leebens-Mack J., Bell C., Paterson A., Zheng C., Sankoff D., dePamphilis C., Wall P., Soltis P. 2009. Polyploidy and angiosperm diversification. American Journal of Botany 96:1: 336-348.

Soltis P. 2000. The role of genetic and genomic attributes in success of polyploids. Proc. Natl. Acad. Sci. USA 97: 7051-7057

Song K, Lu P, Tang K, and Osborn T. 1995. Rapid genome change in synthetic polyploids of Brassica and its implications for polyploid evolution. Proc. Natl. Acad. Sci. USA 92: 7719-7723.

Sønstebø JH, Gielly L, Brysting AK, Elven R, Edwards M, Haile J, Willerslev E, Coissac E, Rioux D, Sannier J, Taberlet P, Brochmann C 2010. Using next-generation sequencing for molecular reconstruction of past arctic vegetation and climate. Molecular Ecology Resources 10: 1009-1018.

Spring J. 2002. Genome duplication strikes back. Nature Genetics 31: 123-129

Stehlik I. 2003. Resistance or emigration? Response of alpine plants to the ice ages. Taxon 52: 499-510.

Stevens P. 2007. Hypericaceae. Pp 194–201. In: K. Kubitzki [ed]. The Families and Genera of Vascular Plants. Vol. XI. Berlin, Heidelberg: Springer Verlag

Sun X. 2010. Evolution of *Hypericum maculatum* Crantz, *Hypericum* perforatum L. and *Hypericum* X *desetangsii* by means of morphological analysis. Diploma thesis at the University of Heideleberg

Taberlet, P., Gielly, L., Pautou, G., Bouvet, J. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol Biol 17:1105-1109.

Taberlet P, Fumagalli L, Wust-Saucy A.-G and Cosson J.-F. 1998. Comparative phylogeography and postglacial colonization routes in Europe. Molecular Ecology 7: 453-464.

Templeton, A. R., Crandall, K. A., and Sing, C. F. 1992. A cladistic analysis of Phenotypic Associations With Haplotypes Inferred From Restriction Endonuclease Mapping and DNA Sequence Data. III. Cladogram Estimation. Genetics 132:619-633.

Tribsch A, and Schönswetter P 2003. Patterns of endemism and comparative Phylogeography confirm palaeo-environmental evidence for Pleistocene refugia in the Eastern Alps. Taxon 52: 477-497.

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van de Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, and M. Kuiper 1995. AFLP: a new technique for DNA -Fingerprinting. Nucleic Acids Res. 1995 November 11; 23(21):4407 -4414.

Valente, L. M., V. Savolainen, and P. Vargas. 2010. Unparalleled rates of species diversification in Europe. Proc. R. Soc. B 277: 1489-1496.

Van Dijk PJ, Bakx-Schotman t. 1997 Chloroplast DNA phylogeography and cytotype geography in autopolyploid Plantago media. Molecular Ecology 6:4: 345-352.

Welch J., Welch D. and Meselson M. 2003. Cytogenetic evidence for asexual evolution of bdelloid rotifers. Proc. Natl. Acad. Sci. USA 101:6: 1618-1621.

Wendel J. and Doyle J. 1998. Phylogenetic incongruence: Window into genome history and molecular evolution. Soltis D., Soltis P. and Doyle J. (ed.) In: Molecular Systematics of plants II DNA sequencing. Kluwer Academic Publisher Group. pp. 265-

Westberg, E. Kadereit, J.W. 2009: The influence of sea currents, past disruption of gene flow and species biology on the phylogeographic structure of coastal flowering plants. - Journal of Biogeography 36: 1398-1410.

White S. and Doebley J. 1998. Of genes and genomes and the origin of maize. Trends in Genetics 14:8: 327-332

WhiteT., Bruns T.; Lee S. and Taylor J. 1990 Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In Innis M., Gelfand D., Sninsky J. And White T. (eds) PCR Protocols: A Guide to Methods and Applications. Academic Press; 1 edition. 315-322.

Willis K.J., and R. J. Whittaker 2000. The refugial debate. Science 287: 1406-1407.

Wilson P, 1992 On Inferring Hybridity from Morphological Intermediacy. Taxon 41: 1 11-23

ZINK, R. M. 2002. Methods in comparative phylogeography, and their application to studying evolution in the North American aridlands. Integrative Comparative Biology 42: 953-959.

Zobayed S, Afreen F, Goto E. and Kozai T. 2006. Plant-environment interactions: accumulation of hypericin in dark glands of *Hypericum* perforatum. Annals of Botany 98: 793–804.

9 List of figures

10 List of tables

11 List of supplementary material

12 List of extra material on CD

Extra material 1 summary of results
Extra material 2 alignment of 21 chloroplast haplotypes
Extra material 3 alignment of 35 ITS genotypes
Extra material 4 detailed results of seed screen
Extra material 5 Summary table of 28 scored characters
Extra material 6 Summary table of 15 characters used in PCA and CDA
Extra material 7 PDF of doctoral thesis

13 Acknowledgements

This research project was based on many people´s help and advice, and I express my deep gratitude to all of them. Special thanks go to:

Prof. Marcus Koch, my supervisor, for freedom in research, scientific guidance, confidence, and encouragement. It was a great pleasure to work on the "*Hypericum*" project.

Dr. Timothy Sharbel for the agreement of being my second supervisor and for inspiration in many discussions.

All members of the AG Koch and AG Erbar for a great working atmosphere, especially Rosi for all the discussions and endless help in everything, Liza, for helping me wih my struggle for clear sentences, Peter, Paola, Robert, Markus, Susanne, Florian, Fabian, Myriam and Ingrid, Christiane, Constanze, Juraj and Ute who have left already.

Anja Landau (Diploma thesis,), Xiaorui Sun (Diploma thesis,), Norah Hohmann (Bachelor thesis,) for working on various aspects of this project.

My family, for all the years of support. Especially Renata, for taking care of my daughter and also taking care of me, my husband for everything and Laura for trying to put some sense in my sentences.

My best friend Astrid, who saved me already during school time in so many ways and also now helped me to get through this thesis.

Many thanks to all who exchanged plant material or DNA with me: Marta Molins, Nicolai Nürk, Anna Schallau, Christian Bräuchler, Christoph Dobes, Juraj Paule, Yunpen Zhao. Thanks to Herbarium München (MSB), Staatsherbarium München (M), the Natural History Museum of London and the Hungarian Natural History Museum for letting me do destructive sampling on numerous herbarium specimen.

The team of the Botanical Garden Heidelberg, especially Torsten Jakob, for taking very good care of the *Hypericum* plants.

DFG, for funding this project