Identification and genetic analysis of the *APOSPORY* locus in *Hypericum perforatum* L.

Dissertation

zur Erlangung des akademischen Grades doctor rerum naturalium (Dr. rer. nat.)

vorgelegt der

Mathematisch-Naturwissenschaftlich-Technischen Fakultät

(mathematisch-naturwissenschaftlicher Bereich)

Martin-Luther-Universität Halle-Wittenberg

von

Anna Schallau

geboren am 24. Februar 1982 in Chisinau, Moldawien

Gutachter :
1. Prof. G. Reuter
2. Prof. M. Koch
3. Prof. U. Wobus

Halle (Saale)

To my family

The microscope can see things the naked eye cannot,

but the reverse is equally true.

Hans Selye. From dream to discovery

Table of contents

Table of contents

Table of contents

List of abbreviations

ACL	*Apomixis Controlling Locus* (*Paspalum*)
AFLP	Amplified Fragment Length Polymorphism
AI	Aposporous Initials
APO	aposporous
ARI	ARIADNE gene
ASGR	*Apospory Segregating Genomic Region* (*Pennisetum*)
BAC	Bacterial artificial chromosome
BLAST	Basic Local Alignment Search Tool
bp	base pair
CAPS	Cleaved Amplified Polymorphic Sequence
cDNA	Complementary DNA
cM	centiMorgan
DAB	Decolorized aniline blue
DAPI	4',6-diamidino-2-phenylindole
DNA	Deoxyribonucleic acid
dNTPs	Deoxyribonucleotides
EDTA	Ethylenediamine tetraacetic acid
ES	Embryo sac
2N ES	Two-nucleate embryo sac
4N ES	Four-nucleate embryo sac
8N ES	Eight-nucleate embryo sac
EST	Expressed Sequence Tags
FCSS	Flow Cytometric Seed Screen
FM	Functional Megaspore
HAPPY	*Hypericum* Apospory locus
HpARI	*Hypericum* ARIADNE gene
hpt	hygromycin phospho-transferase gene
kb	kilobase pair

List of abbreviations

LB medium	Lysogeny Broth, or Luria-Bertani Broth
MES	2-(N-morpholino)ethanesulfonic acid
MMC	Megaspore Mother Cell
MS medium	Murashige and Scoog medium
NCBI	National Center for Biotechnology Information
PCR	Polymerase Chain Reaction
PFGE	Pulsed Field Gel Electrophoresis
PSL	Photostimulated luminescence
RFLP	Restriction fragment length polymorphism
RNA	Ribonucleic acid
RT-PCR	Reverse transcription polymerase chain reaction
SEX	sexual
SNP	Single Nucleotide Polymorphism
SSR	Simple Sequence Repeat
T-35S	Terminator 35S
TIGR	The Institute for Genomic Research
Ubi	Ubiquitin promoter
UV	ultraviolet radiation

1. Introduction

1. 1. Sexual reproduction

The typical life cycle of higher plants comprises a haploid phase in which multicellular gametophytes produce male and female gametes that, upon fertilization, reconstitute the diploid sporophyte. The phylogeny of land plants is characterized by an evolutionary trend towards gametophyte reduction, as described by Wilhelm Hofmeister more then a century ago (Hofmeister, 1851).

In angiosperms, the female gametophyte or embryo sac is strongly reduced and embedded in the sporophytic tissues of the ovule. The development of the female gametophyte includes two phases: megasporogenesis and megagametogenesis. The megasporogenesis starts with the differentiation of a single cell within the ovule primordium into a megaspore mother cell (MMC). The MMC undergoes meiosis and four haploid megaspores are produced. Three megaspores degenerate and the fourth megaspore develops further to become a functional megaspore (FM). During megagametogenesis the FM undergoes three mitotic divisions, producing eight nuclei. After cellularization, seven cells can be found in the mature embryo sac (ES): the egg cell flanked by two synergids, three antipodal cells and the central cell with two nuclei fused to a diploid nucleus (Figure 1). The described type of ES is present in 70% of angiosperms (Maheshawari, 1950) and is called *Polygonum* embryo sac type.

The development of the male gametophyte occurs in the anthers, where the pollen mother cell undergoes meiosis to generate four haploid microspores. The first mitotic division results in a vegetative and a generative cell. The vegetative cell does not divide further, but develops into the pollen tube, which enables the transport of the sperms to the female gametophyte. The generative cell goes through a second of mitotic division and generates two sperm cells. One of them fertilizes the egg cell to form a diploid embryo and the other one fuses with the diploid central cell to form a triploid endosperm (Nawaschin, 1898). Double fertilization initiates seed development and represents a hallmark of sexual reproduction of angiosperms (Grossniklaus and Schneitz, 1998). The offspring genome is a complex mixture of maternal and paternal alleles giving rise to variable genotypes and phenotypes.

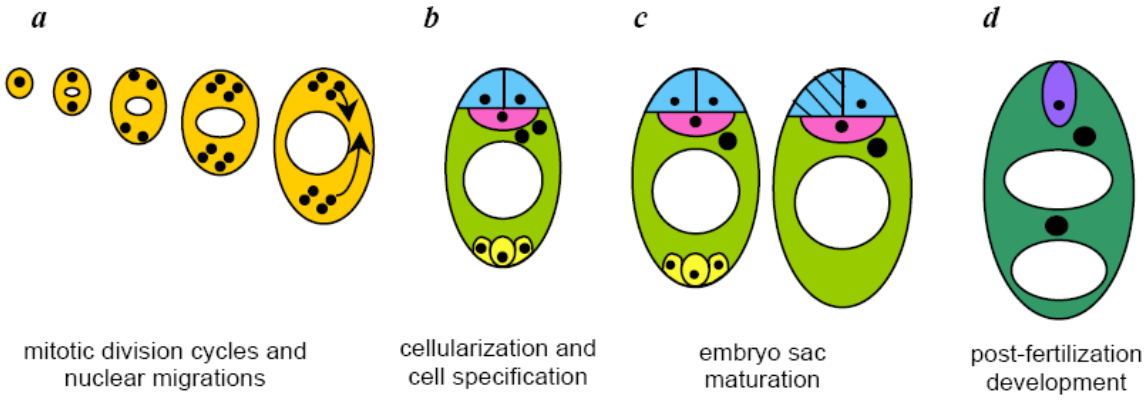


Figure 1: Schematic representation of the development of a *Polygonum* type female gametophyte (after Brukhin *et al.*, 2005)

a. Three nuclear divisions occur in a syncytium to form an 8-nucleate female gametophyte. One nucleus from each pole (the polar nuclei) migrates and will eventually be enclosed by the central cell. b. Cellularization forms the typical 7-celled, 8-nucleate female gametophyte with two synergid cells (blue), one egg cell (red), a bi-nucleate central cell (green) and three antipodal cells (yellow). c. Before fertilization, the female gametophyte matures: the two polar nuclei fuse to form the homo-diploid secondary endosperm nucleus, the antipodals undergo programmed cell death and one of the synergids degenerates. d**.** During double fertilization one sperm cell fuses with the egg cell to form the diploid zygote (purple), while the second sperm fertilizes the central cell to form the triploid endosperm (dark green). The primary endosperm nucleus divides prior to the zygote in a syncytium.

1.2. Apomixis

In contrast to the variable phenotypes of sexual offspring, asexual reproduction leads to uniform genotypes, which are copies of the mother plant. The asexual reproduction can be divided into vegetative reproduction and the reproduction with clonal seed formation (apomixis). Apomixis, the asexual reproduction through seeds, occurs in more than 400 plant species of about 40 angiosperm families (Carman, 1997; van Dijk and Vijverberg, 2005). Apomictic offspring are genetically identical clones of the mother plant. Successful engineering of apomixis into sexual crop plants is considered as a high priority goal of plant breeding with far-reaching potential for instance for the fixation of heterosis effects. The economic and social benefits are thought to exceed those of the green revolution (Spillane *et al.*, 2004). For example, it was estimated that a saving of US$ 2.5 billion per annum can be reached if it is possible to transfer apomixis to rice alone (McMeniman and Lubulwa, 1997).

Understanding the genetic and molecular mechanisms underlying apomixis traits in model plants is one of the primary steps towards transferring apomixis to crop plants.

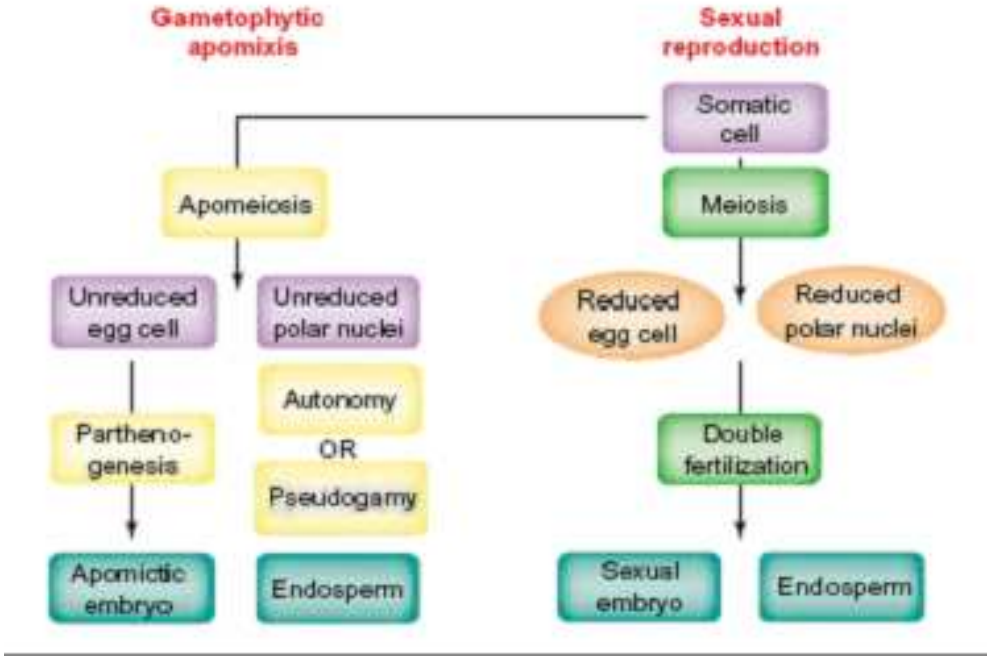


Figure 2: Comparison of the major steps in sexual and apomictic development (modified after Grimanelli *et al.*, 2001)

Three major elements of apomixis: apomeiosis, parthenogenesis and autonomous/pseudogamous endosperm development are highlighted in yellow.

There are two major types of apomixis observed in plants, (i) sporophytic apomixis also called adventitious embryony with autonomous embryo development in sporophytic tissues; and (ii) gametophytic apomixis by which the meiotically unreduced eggs develop into parthenogenetic embryos. Based on the identity of the apomixis progenitor cell type there are two subtypes of gametophytic apomixis, diplospory and apospory. In diplospory, the megaspore mother cell (MMC) is the progenitor cell that circumvents meiosis to produce unreduced embryo sac. On the contrary, unreduced aposporous embryo sacs develop from cell(s) adjacent to the MMC that completely omit meiosis. These cells are called Aposporous Initials (AI). Both diplospory and apospory are called apomeiosis, because no meiosis occurs. In both cases the non-reduced egg cells initiate embryo development independent of fertilization, a process called parthenogenesis. Concerning endosperm development natural

apomicts can be divided into two groups which differ in their requirement of fertilization for endosperm development. Whereas apomicts with autonomous endosperm formation (e.g. *Hieracium, Taraxacum*) do not depend upon fertilization of the central cell nuclei, apomicts with the pseudogamous endosperm formation require the fertilization of the central cell nuclei (e.g. *Poa pratemsis*, *Hypericum*) (Figure 2) (for reviews see Nogler, 1984; Koltunow, 1993; Grossniklaus *et al.*, 2001; Spillane *et al.*, 2004; Bicknell and Koltunow, 2004; Ozias-Akins, 2006). In all cases, the maternal genotype will be retained in the progeny developing from seeds formed without meiotic reduction and egg cell fertilization. Most apomictic species are facultative apomicts, with a peculiar balance between genetic stability determined by apomixis and flexibility caused by recombination during sexual reproduction.

1.3. Genetic control of apomixis

Despite the identification and description of a number of apomictic angiosperms, simple introgression of apomixis into related sexual crop plants by breeding seems unrealistic due to epigenetic barriers (Savidan, 2001). Besides, many taxa of agricultural importance do not have apomictic relatives that are required for introgression. Therefore, a coordinated strategy for understanding of apomixis at the molecular level is being sought by plant reproductive biologists with the ultimate goal to genetically engineer apomixis in crop plants (Spillane *et al.*, 2004). However, despite of long lasting efforts, the knowledge about the genetic control of apomixis is still rather limited.

The molecular mapping and positional cloning of apomixis genes is hindered by the lack of recombination within the chromosomal region governing apomeiosis. Earlier studies suggested a simple inheritance of apomixis by one dominant locus (Peacock, 1992; Savidan, 1980; Leblanc *et al.*, 1995; Bicknell *et al.*, 2000). Thus, for some apomictic plants it was shown, that the apomixis-specific markers segregated as a single locus. For example, in *Panicum maximum* 9 AFLP markers define the apomixis locus, estimated to be 1.4 cM long (Ebina *et al.*, 2005). In *Paspalum simplex* an extensive AFLP analysis led to the indentification of 33 apospory-specific markers. Interestingly, 8 of them hybridized only to the DNA of the apomicts, e.g. were hemizygous, suggesting that the apomixis locus lacks sequence homology and almost likely contains rearrangements with respect to the sexual

genotype. For one of the markers the hemizygosity was also confirmed cytologically (Calderini *et al.,* 2006). The apomixis controlling locus (*ACL*), defined by the markers, is synthenic with a part of rice chromosome 12, with the size of the locus estimated to be 8-10 cM (Pupilli *et al.,* 2001). Additional mapping efforts revealed that the *ACL* may contain translocations based on the linkage with a marker from rice chromosome 2 (Pupilli *et al.,* 2004). The cloning and sequencing of an apomixis linked *Paspalum* BAC revealed large-scale rearrangements due to insertion of transposable elements, and small-scale rearrangements due to deletions and single point mutations within some potential coding sequences (Calderini *et al.,* 2006). Similary, the apospory specific genomic region (*ASGR*) of *Pennisetum* has been identified to be more than 2.7 Mbp and almost half a chromosome long (Goel *et al.,* 2006). Based on several apomixis-associated RFLP markers, present only in apomicts it was shown that the locus exhibits partial hemizygosity. Hemizygosity was observed not only in the regions containing low copy-sequences, but also with regions containing repetitive sequences. Aposporous apomixis of *Pennisetum* and *Paspalum* species is linked with a non-recombining region of supernumerary chromatin (Ozias-Akins *et al.,* 1998; Roche *et al.,* 2001; Labombarda *et al.,* 2002). Apomixis was also shown to be controlled by single locus in *Brachiaria* sp. (do Valle *et al*., 1993; Miles and Escandon, 1996), *Paspalum notatum* (Martinez *et al*., 2001), *Ranunculus* sp. (Nogler, 1984). Interestingly, diplosporous apomixis of the *Boechera holboelli* complex is considered to be associated with homeologous chromosome substitutions (Kantama *et al.,* 2007). All *Boechera* apomicts share a large heterochromatic chromosome. Both the heterochromatic nature of this chromosome and its association with the apomictic phenotype hint at similarities with the supernumerary chromosome fragment of *Pennisetum* (Kantama *et al.,* 2007).

In other model systems the genetic control of apomixis was shown to be controlled by two loci controlling the two components of apomixis (apomeiosis and parthenogenesis) independently. For example apomixis in *Hieracium* was initially described as a dominant monogenic trait (Bicknell *et al.,* 2000). However, after deletion mapping had been carried out the model of apomixis control was revised. Chromosomal deletions were induced by gamma irradiation of seeds. In a simplex dominant apomixis control, deletion of a chromosome segment carrying the dominant apomixis allele resulted in reversion to sexuality (Catanach *et al.,* 2006). Phenotyping of the progeny of the deleterious plants resulted in plants that lost apospory or plants that lost parthenogenesis or plants that had lost both. After the fine

mapping of these mutants two genomic regions could be associated with the mutant phenotypes: *Loss of Apomeiosis* (*LOA*) and *Loss of Parthenogenesis* (*LOP*). Similarly, in some other plants the recombination between apospory and parthenogenesis is possible. A clearly documented case of recombination between apospory and parthenogenesis is found in *Poa pratensis* (Albertini *et al*., 2001). In *Poa*, parthenogenesis can be phenotyped in the absence of fertilization by the development of embryos after auxin treatment (Matzk, 1991). Using auxin treatment, parthenogenesis had been mapped as a single locus, encoding a qualitative trait as shown by AFLP and SCAR markers (Albertini *et al*., 2001a; Barcaccia *et al*., 1998).

Similar to the inheritance of apomixis in *Poa*, linkage between diplosporous embryo sac development and parthenogenesis can readily be broken in *Erigeron* and *Taraxacum*. Thus, in *Erigon* the genetic markers formed two loci, with 11 markers co-segregating with diplospory defining the *Diplospory Linkage Group* and 4 markers co-segregating with parthenogenesis defining the *Parthenogenesis Linkage Group* (Noyes and Rieseberg, 2000). Similarly, in *Taraxacum* two loci, encoding diplospory and parthenogenesis are known (Van Dijk and Bax-Schotman, 2004). Surprisingly, in *Taraxacum* a third locus, co-segregating with apomixis has been identified and is considered to be responsible for autonomous endosperm development (van Dijk *et al*., 2003). Thus, it was shown at least for *Taraxacum* that each component of apomixis is encoded by an independent locus.

However, the individual components of apomixis may be controlled by multiple genes. For example, apospory is typically treated as a single qualitative trait, although it is quantitative and can highly vary from plant to plant. Probably, modifiers play a role in the regulation of penetrance of apospory in some species (Koltunow and Grossniklaus, 2003). A comprehensive model for the control of apomictic seed formation is proposed for the aposporous species *Poa pratensis* and includes 5 different genes controlling apomixis and considers the role of modifiers. Mutations of genes which originally were responsible for the sexual pathway as well as newly evolved genes have been considered to explain the inheritance of apomixis (Matzk *et al.*, 2005).

Thus, comparison of many model systems makes it difficult to draw final conclusions about genetic control of apomixis. It is hard to define a model which fits all apomicts.

Whether the observed differences are due to experimental bias or to more fundamental differences in the nature of various forms of apomixis is unclear.

1.4. Apomixis and asynchronous expression of the developmental program

Apomixis is frequently associated with a high degree of heterozygosity and polyploidy (Nogler, 1982; Roy, 1995; Bicknell, 1997; Kojima and Nagato, 1997). Interestingly, there are few examples of diploid apomicts, one of them being *Boechera*, were some of the diploids are sexuals and the others are pseudogamous aposporous apomicts. However, care should be taken in assigning a particular ploidy to sexual or apomictic reproduction because an apparently diploid species might represent a diploidized ancient polyploid (Paleopolyploid) (Uwe Praeckelt and Rod Scott, 2001). Currently it is not clear whether higher ploidy levels are cause or consequence of apomixis (Koltunow and Grossniklaus, 2003). One possible explanation of frequent occurence of apomictic polyploids is that a certain apomixis factor can not be transmitted through gametes because the allele responsible for apomixis is lethal when present under haploid conditions (Nogler, 1982; Grimanelli *et al.,* 1998; Noyes and Riesberg, 2000). Another explanation suggested by Carman (1997) is that polyploidy may result in asynchronous expression of the genes belonging to the different genomes which are contained within the polyploid. The asynchronous expression of these genes, most likely involved in the sexual gametophyte development, might be responsible for the apomictic phenotype.

Another hypothesis, involving asynchronous development was proposed and supported by many authors (Moggy, 1992; Spillane *et al.*, 2001; Koltunow and Grossniklaus, 2003; Ozias-Akins, 2006). Since apomixis has evolved from a sexual background (Holsinger, 2000), it can be considered to be a short circuit of the sexual pathway of reproduction due to the omission or deregulation of parts of the sexual developmental program, which may be caused by mutations in the existing genes or the appearance of new genes (Spillane *et al.,* 2001; Tucker *et al.,* 2003; Sharbel *et al.,* 2009). Indeed, during the development of apomictic seeds, a subsequent developmental event is initiated before the previous one is completed. For example, meiosis and embryo sac formation which usually follow each other, may occur at the same time: the Aposporous Initial initiates embryo sac development without entering

meiosis or after premature meiotic abortion and nuclear restitution (Grossniklaus, 2001b). Similarly, parthenogenetic embryogenesis occurs often prior to anthesis and often before the fertilization of the central cell in pseudogamous apomicts. Therefore, apomixis can be interpreted as a relaxation of the spatial and temporal constraints on the reproductive developmental processes (Grossniklaus *et al.*, 1998). This led to the suggestion that apomixis might be triggered by reproduction specific gene expression programmes activated at the wrong time and/or at the wrong place (Koltunow and Grossniklaus, 2003). Therefore, expression and ultimate function of genes that are critical for the sexual developmental program are expected to be misregulated in apomicts.

1.5. Endosperm development of apomictic plants

The endosperm development is very important forb both apomictic and sexual seed development. The relative contributions of the maternal and paternal genomes to endosperm development in an apomictic plant differ significantly from the relative contributions in the sexual plant. For a sexual plant the ratio of the 2m:1p is characteristic: 2 maternal genomes and 1 paternal genome are combined in the triploid endosperm. In sexual plants genome dosage is often critical for seed development. For example in a crop plant like maize, the deviation from the 2m:1p ratio leads to seed abortion (Birchler, 1993). In apomictic plants the ratio can vary. Many apomicts have modifications in the structure of female gametophyte or in fertilization behavior to restore the normal balance and avoid seed abortion. For example, the *Panicum* type of embryo sac contains four nuclei instead of the typical eight, so that there is only a single unreduced polar nucleus and not two nuclei in the central cell as usual. This leads to 2m:1p ratio in endosperm after fertilization. In *Ranunculus auricomus* both sperms fertilize the unreduced central cell (Rutishauser, 1954), resulting again in 2m:1p ratio. There are also apomicts with an endosperm, which has a ratio other then 2m:1p and still produce viable seeds. For example in *Tripsacum dactyloides* a single sperm fuses with an unreduced central cell to produce 4m:1p endosperm (Haig and Westoby, 1991). Or in autonomous apomicts, such as *Hieracium* or *Taraxacum*, the violation of the maternal: paternal ratio is extreme, since there is no contribution from the male part and the ratio is 8m:0p. So, for some apomictic plants an alteration of the 2m:1p ratio seems to be uncritical, most likely due to an alteration of the imprinting system. Such alterations might be due to the modification of DNA

methylation, which remove the developmental block on endosperms with insufficient, or no, parental genomes owing to the bypass of meiosis in pseudogamous apomicts or the absence of fertilization in autonomous apomicts (Spielman *et al.,* 2003).

1.6. Epigenetic control of apomixis

In recent years, hypotheses concerning genetic control of apomixis have been extended to incorporate epigenetic mechanisms for several reasons (Lohe and Chaudhury, 2002; Koltunow and Grossniklaus, 2003). Firstly, identification of a plethora of transposons and associated repetitive sequences surrounding the apomixis loci indicated that these loci could be associated with inherent epigenetic modifications such as DNA methylation (Lohe and Chaudhury, 2002; Koltunow and Grossniklaus, 2003). Secondly, genes that cause mutants of sexual genotypes to express features of parthenogenesis and autonomous endosperm development encode essential epigenetic regulators involved in DNA and/or histone methylation (reviewed in Köhler and Makarevich, 2006). In addition, there are instances where epialleles could behave in a dominant manner (Lohe and Chaudhury, 2002). Therefore, master gene(s) controlling apomixis might be the target of epigenetic modifications or these regulatory factors could reciprocally control epigenetic marks. Taken together, an epigenetic model as a basis for apomixis would be able to combine earlier hypotheses involving mutant alleles, dominant genes, hybridization and polyploidy (Lohe and Chaudhury, 2002; Koltunow and Grossniklaus, 2003).

1.7. Towards the identification of candidate genes for apomixis

Two main approaches are being pursued to identify candidate apomixis genes. The first is mapping of the apomixis loci in apomictic plants and differential expression analysis between apomictic and sexual plants. The second approach includes the analysis of sexual reproduction in a plant model systems such as *Arabidopsis* or maize, to identify genes that, if mutated or deregulated, display elements of apomixis (Spillane *et al.,* 2004).

Mapping of the apomixis loci in apomictic plants is hindered due to the lack of recombination around the responsible loci. Thus, very few candidate genes were proposed based on this method. One of the well-characterised apomixis loci is the *ASGR* of *Pennisetum* syn. *Cenchrus*. The partial sequencing (0.5 coverage) of *ASGR* revealed several genes that may play a role in apomictic development. In total 53 unique genes were identified that displayed a wide range of functional classifications. As the most probable candidate a *BABY BOOM-like* gene was suggested (Goel *et al.,* 2006). However, no functional analysis confirming the role of the identified gene has been done. Similar, a short region of the *ACL* of *Paspalum* was sequenced. Interestingly, the *ACL* contained a truncated allele of a gene, which exhibited similarity to Arabidopsis At3g04610 (Pulpilli, personal communication). This gene was characterized to be involved in the regulation of flowering initiation. No further function connected to apomixis has been detected. A more detailed molecular mapping of the locus/loci controlling apomixis, followed by its extended characterization is necessary to identify possible candidate apomixis genes. Due to lowering of the sequencing costs this approach, although cumbersome, is perhaps the most promising strategy to identify key genes for apomictic reproduction in wild apomictic species (Calderini *et al.,* 2007).

Another approach to identify candidate genes for apomixis in apomictic plants includes comparisons of differential gene expression during sexual and apomictic development. Several studies have been performed to identify genes based on their differential expression in apomictic and sexual plants (Vielle-Calzada *et al.,* 1996; Pessino *et al.,* 2001; Akiyama *et al.,* 2005; Sharbel *et al.,* 2009). Like other approaches, also this experimental strategy meets certain difficulties. Firstly, the development of sexual and apomictic ovules within a single plant might be asynchronous. This partially excludes the direct comparison of transcriptional programmes, but might be overcome by precise staging of the ovule development and sufficient number of repetitions. Secondly, the female gametophyte is deeply imbedded in the surrounding sporophytic tissues. This might lead to the dilution of apomixis-related transcripts, a problem which can be overcome by microdissection based methods. Due to this experimental problems, only few of the mentioned above studies resulted in defined candidate genes for apomixis. Only four differentially expressed genes were identified and selected as candidate genes for apospory: APOSTART in *Poa* (Albertini *et al.,* 2005), ASG-1 in *Paspalum (Chen et al.,* 2005) and Pca21, Pca24 from *Pennisetum* (Singh *et al.,* 2007). The first three mentioned genes were

shown to be higher expressed in apomicts than in sexuals at the early stages of ovule development. Pca24 was shown to be specific for apomicts. Interestingly, both APOSTART and ASG-1 genes are also expressed in the anthers during microsporogenesis. However, no further functional analysis demonstrating the role of these genes in apomixis has been published.

In addition, apomictic and sexual pathways share certain gene expression patterns (Vielle-Calzada *et al.,* 1996; Tucker *et al.,* 2003; Albertini *et al.,* 2005). According to Matzk *et al.* (2005), once the Aposporous Initials are selected, the cascade for gametogenesis proceeds as during sexual development, expressing similar genes during the process. Thus, recent studies in aposporous apomicts such as *Paspalum* and *Panicum* provided evidence that the aposporous initial cells would show the identity of functional megaspores. However, there is a temporal and spatial shift in expression between aposporous initial cells and the megaspore mother cells (MMCs) (Tucker *et al.,* 2003; Albertini *et al.,* 2005; Chen *et al.,* 2005). Similarly, Tucker *et al.* (2003) observed that the developmental expression of some *FERTILIZATION INDEPENDENT SEED* (*FIS*) genes and tissue specific marker genes were equally shared between apomictic and sexual *Hieracium* species.

Since apomixis might be a short circuit of the sexual pathway, genes that are involved in sexual plant reproduction were considered to be good apomixis candidate genes. Following this approach, several mutants of *Arabidopsis* and maize have been identified that display components of apomixis, such as absence of meiosis, parthenogenesis and autonomous endosperm formation (Huang and Sheridan, 1996; Ohad *et al.,* 1996; Sheridan *et al.,* 1996; Chaudhury *et al.,* 1997; Kohler *et al.,* 2003b; Guitton *et al.,* 2004; Barrell and Grossniklaus, 2006). While mutations in the three FIS (fertilization independent seed) genes: *fis1/mea* (medea), *fis2* and *fis3/fie* mimic autonomous development of the central cell to some extent, none of these mutations induces full differentiation of autonomous endosperm (Vinkenoog *et al.,* 2000). In addition, there is little or no development of the egg, because *FIS* genes are acting relatively late (Chaudhury *et al.,* 1997). Recently a *dyad* mutant (Ravi *et al.,* 2008) was described to mimic apomeiosis. Unfortunately, only 0.2% of *dyad* ovules generate viable gametes, which makes it practically unusable for engineering apomixis. Another attempt to engineer apomixis was made by D'Erfurth *et al.,* 2009, where apomeiosis was partially mimicked. The mutations of three genes (*osd1, Atspo11-1, Atrec8*) involved in meiosis were

combined. However, the F2 progeny was tetraploid and the ploidy doubled with each generation. It is unclear if it is possible to completely mimic apomeiosis without the genome doubling. Therefore, in contrast to the original hypothesis that apomixis arose from sexuality due to a mutation in a regulatory gene, additive effects of several mutations could be necessary for apomixis.

1.8. *Hypericum perforatum* as a model system to study apospory

St. Johns wort (*Hypericum perforatum* L.) is a yellow-flowering perennial herb, growing from a rhizome (Figure 3). Stems are 1 to 3 feet high, erect, with numerous branches above, glabrous, somewhat 2-ridged, rust-colored and woody at the base. Leaves are opposite, sessile, entire, elliptic to oblong, about 12 mm or longer and glabrous with transparent spots throughout and black marginal dots. Flowers are 2.5 cm in diameter, bright yellow, numerous in flattopped cymes, with 5 petals with occasional minute black dots around the edges. The sepals are pointed, with glandular dots in the tissue. Stamens are numerous, arranged in 3 groups. Seed pods are 2mm long, rust-brown, 3-celled capsules, each with numerous seeds (Lomer and Douglas 1998, Whitson *et al.*, 2000). The name *Hypericum* is derived from the Greek words *hyper* (above) and *eikon* (picture), in reference to the traditional use of the plant to ward off evil, by hanging plants over a picture in the house. The species name *perforatum* refers to the presence of small oil glands in the leaves that look like windows, which can be seen when they are held against the light. *Hypericum perforatum* produces pharmaceutically important metabolites with possible antidepressant, anticancer, antiviral, antifungal and antimicrobial activities (Di Carlo *et al.*, 2001; Agostinis *et al.*, 2002; Miskovsky, 2002; Schempp *et al.*, 2002; Dulger *et al.*, 2005; Fenner *et al.*, 2005; Ferraz *et al.*, 2005; Francis, 2005; Malaty, 2005). St John's wort is today most widely known as herbal treatment for major depression.

Figure 3: *Hypericum perforatum* (Thome, O.W. 1885)

A. rhizome and stem, **B.** branch with flowers, **1.** flower bud **2.** opened flower **3.** androecium and gynoecium **4.** stamens gathered in a fascicle **5.** stamen, anther with black glands, 6. three carpels grown together, abaxial section **7.** ripe fruit **8.** fruit with seed capsule, axial section **9.** seed capsule and seeds

However, here it is of importance that *Hypericum perforatum* is a tetraploid, facultative aposporous and pseudogamous apomict (Noack, 1939; Matzk *et al.*, 2001, 2003). Apomixis in this species was first described in the pioneering work of Noack (1939), but only recently characterized in great detail at both the individual plant and the population level by Matzk *et al.* (2001) and Barcaccia *et al*. (2006). In *H. perforatum* individuals, embryo sacs may be either reduced (meiotic) or unreduced (aposporous) and both types of egg cells may be either fertilized (gamic) or develop partenogenetically (agamic), resulting in six possible categories of progeny (Matzk *et al*., 2001). Moreover, *H. perforatum* ecotypes are usually polyclonal, being not dominated by a single genotype, and characterized by different levels of differentiation among multilocus genotypes (Barcaccia *et al*., 2006). The more ubiquitous apomictic pathway is typically characterized by the megaspore mother cell entering meiosis,

followed by embryo sac degeneration. In its vicinity, in the basal part of the nucellus or, more frequently, in the deeper part of the chalaza, a somatic cell becomes an Aposporous Initial and divides mitotically to form a *Hieracium*-type unreduced embryo sac (Noack, 1941; Barcaccia *et al.*, 2006). The facultativeness of apomixis in this species indicates that sexual and aposporous embryo sacs can initially coexist, although it is unclear in *H. perforatum* whether the degeneration of the sexual embryo sac has a genetic basis or results from competition with the aposporous initial or embryo sac (see Barcaccia *et al*., 2007).

Apomixis occurs in about 97% of the plants of the natural populations (Noack, 1939), composed mainly of tetraploids (2*n*=4*x*=32), although diploid (2*n*=2*x*=16) and hexaploid (2*n*=6*x*=48) chromosome numbers have also been reported (Matzk *et al.*, 2001; Robson, 2002). The occurrence of different ploidy levels is most likely caused by the dynamic reproductive system in *Hypericum*. Haploidization and polyploidization are the consequences of parthenogenesis of meiotic egg cells and fertilization of aposporous egg cells, respectively (Barcaccia *et al*., 2007). *Hypericum* has the smallest genome amongst the studied apomicts (637Mbp, http://data.kew.org/cvalues/database1.html), which is four times bigger than *Arabidopsis* genome and 1.3 times bigger than rice. It is about 2 times smaller than the genomes of other studied apomicts such as *Taraxacum* (1250 Mbp) and *Hieracium* (≥1054 Mbp). A versatile mode of reproduction, ranging from complete sexuality to nearly obligate apomixis, along with a relatively small genome size and a relatively short generation time, make St. John's wort an interesting model system for apomixis research (Matzk *et al*., 2001; Barcaccia *et al*., 2007).

1.9. *ARIADNE* belongs to E3 ligases

As described later in the thesis, the *ARIADNE* gene which is a part of the *Hypericum* apospory-specific locus identified in this study belongs to ARIADNE family of E3 ligases. Members of this family have been described to be involved in the regulation of protein degradation. The degradation of proteins is essential for many aspects in plant physiology and development. The major proteolytic pathway for protein breakdown is the ubiquitin / 26S proteasome pathway, which requires a cascade of enzymes including ubiquitin-activating enzymes (E1), ubiquitin-conjugating enzymes (E2) and ubiquitin-protein ligases (E3; Figure

4). The initiation reaction starts with the activation of ubiquitin by the ATP-dependent ligation of E1. This activated ubiquitin is transferred to the E2 enzyme. Using the activity of E3 ligase, the ubiquitin moiety is attached to the substrate. E3s have the highest diversity, they interact directly with the target protein and are responsible for the specificity. The attachment of ubiquitin to the substrate leads to the recognition and degradation of the target protein.

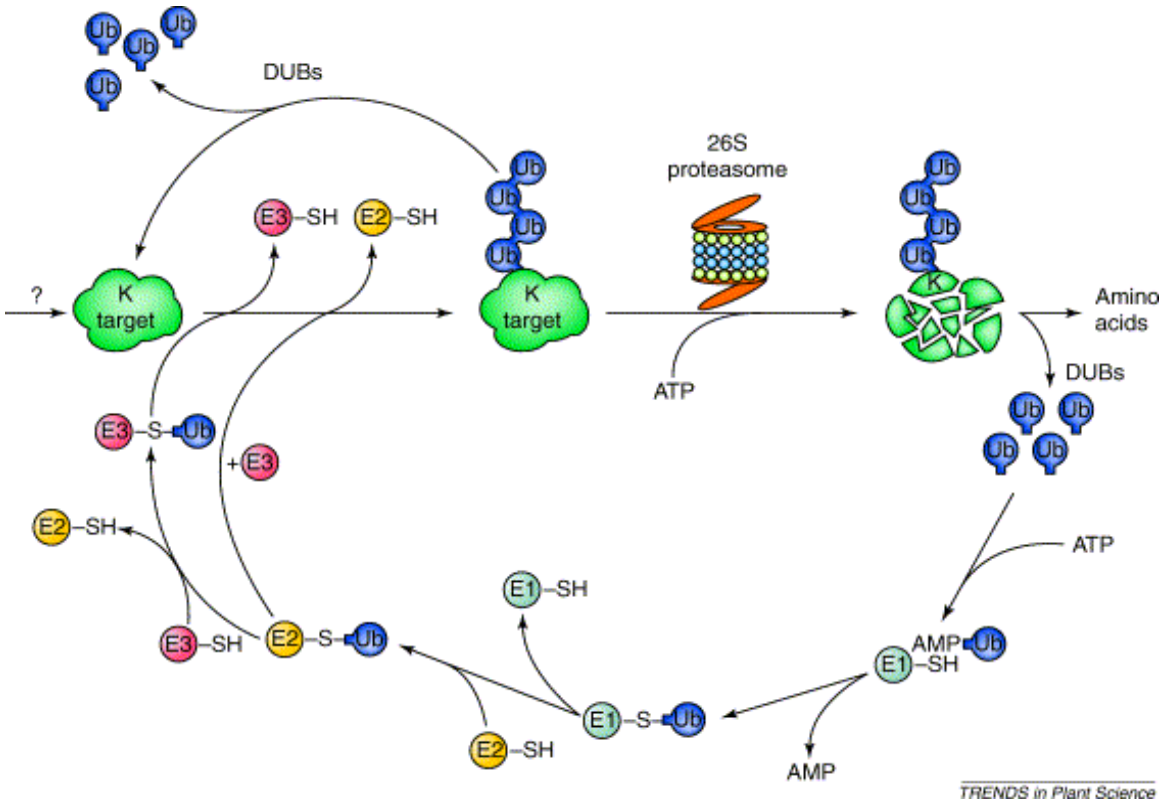


Figure 4: The ubiquitin/26S proteasome pathway (after Vierstra, 2003).

The pathway begins with the ATP-dependent activation of ubiquitin by E1, followed by transfer to an E2, and finally attachment of the ubiquitin to the target protein with the help of an E3. Once a conjugate is assembled bearing a chain of multiple ubiquitins, it is recognized by the 26S proteasome and degraded in an ATP-dependent process. Abbreviation: K - lysine, Ub – ubiquitin, DUBs - deubiquitinating enzymes.

Remarkably, in *Arabidopsis thaliana* about 5% of the proteome are involved in the ubiquitin/26S proteosome pathway, making it one of the most elaborate regulatory mechanisms in plants (Vierstra, 2003). The biggest group of E3 ligases is the RING-finger

group. Some RING-finger containing E3 proteins are part of E3 ubiquitin ligase complexes such as SKP1, Cullin/CDC53, F-box protein. These complexes were shown to be involved in many aspects of plant development, such as the regulation of auxin (Gray *et al.,* 2001) and jasmonate signaling (Xie *et al.,* 1998), flower morphogenesis (Samach *et al.,* 1999), circadian rhythms (Nelson *et al.,* 2000; Somers *et al.,* 2000; Dieterle *et al.,* 2001), leaf senescence (Woo *et al.,* 2001), R gene mediated pathogen resistance (Austin *et al.,* 2002; Azevedo *et al.,* 2002), embryogenesis (Shen *et al.,* 2002) and gametophyte development (O'Brien *et al.,* 2004; Honys *et al.,* 2004; Liu *et al.,* 2008). The ARIADNE (ARI) class of RING-finger proteins are characterized by the presence of an N-terminal acid-rich cluster, followed by a C3HC4 RING-finger motif, a central in between RING-finger (IBR) or B-box, and a second C3HC4 RING-finger structure. At the C terminus, these proteins have a potential coiled-coil domain and a leucin-rich region. The specific functions of genes belonging to the *ARIADNE* family are not known at the moment.

1.10. Aims of the project

Apomixis is one of the most intriguing questions of plant biology with an expected economic impact similar to that of the green revolution. The major aim of this project was to contribute to the understanding of apomixis, as a mode of asexual seed formation in the natural apomict *Hypericum perforatum*. The work of F. Matzk (2001) describing 6 different reproductive pathways and the availability of various populations and ecotypes served as basis for establishing *Hypericum* as a novel model system to study apomixis. In the project I concentrated the efforts on one component of apomixis: apospory.

The following aims were pursued:

1. Cytological characterization of the ovule development within aposporous and sexual plants

2. Elucidation of the genetic control of apomixis based on the segregation of the trait in several populations derived from directed crosses

3. Identification of the apospory-specific locus in those populations and in the collection of wild accessions

4. Molecular characterization of the apospory-specific locus and the corresponding sexual locus

2. Materials and methods

2.1. Material

2.1.1. Plant Material

Species: *Hypericum perforatum*

For the detection of DNA markers linked to apospory in *Hypericum perforatum*, a screening was performed by genomic AFLP analysis using 10 apomictic and 6 sexual genotypes clearly characterized for a contrasting mode of reproduction by means of FCSS (Flow Cytometric Seed Screen, see 2.1) (Matzk *et al*., 2001). The 16 genotypes originated from different germplasm sources, collection sites or botanical gardens in Europe (Table 1).

Plant code	Germplasm origin	Mode of reproduction	Ploidy
aTo	cv. Topaz (Poland)	apomictic	4n
aAn	Breeding stock (Germany)	apomictic	4n
aNo	Breeding stock (Germany)	apomictic	4n
aMü	Münster (Germany)	apomictic	4n
aSie	Siena (Italy)	apomictic	4n
aCa	Caen (France)	apomictic	4n
aWr	Wroclaw (Poland)	apomictic	4n
aWa	Warszawa (Poland)	apomictic	4n
aKs	Kaunas (Lithuania)	apomictic	4n
aV4	Vacrotot (Hungary)	apomictic	4n
sR1	Selection from cv. Topaz	sexual	2n
sP1	Padova (Italy)	sexual	2n
sP2	Padova (Italy)	sexual	2n
sV1	Vacrotot (Hungary)	sexual	2n
sV2	Vacrotot (Hungary)	sexual	2n
sV3	Vacrotot (Hungary)	sexual	2n

Table 1: Origins of the 10 apomictic (a) and 6 sexual (s) *Hypericum perforatum* accessions used for the AFLP analysis

A triploid segregating population (59 plants) was obtained by crossing completely sexual diploid plants with obligate apomictic tetraploid plants. To exclude ploidy effects, a tetraploid segregating population (72 plants) was also obtained. For this purpose, an obligate sexual plant (R1) was tetraploidized with colchicine and crossed with a tetraploid obligate apomict (To). Two resulting sexuals were further crossed with various obligate apomicts (see the crossing scheme in Figure 5).

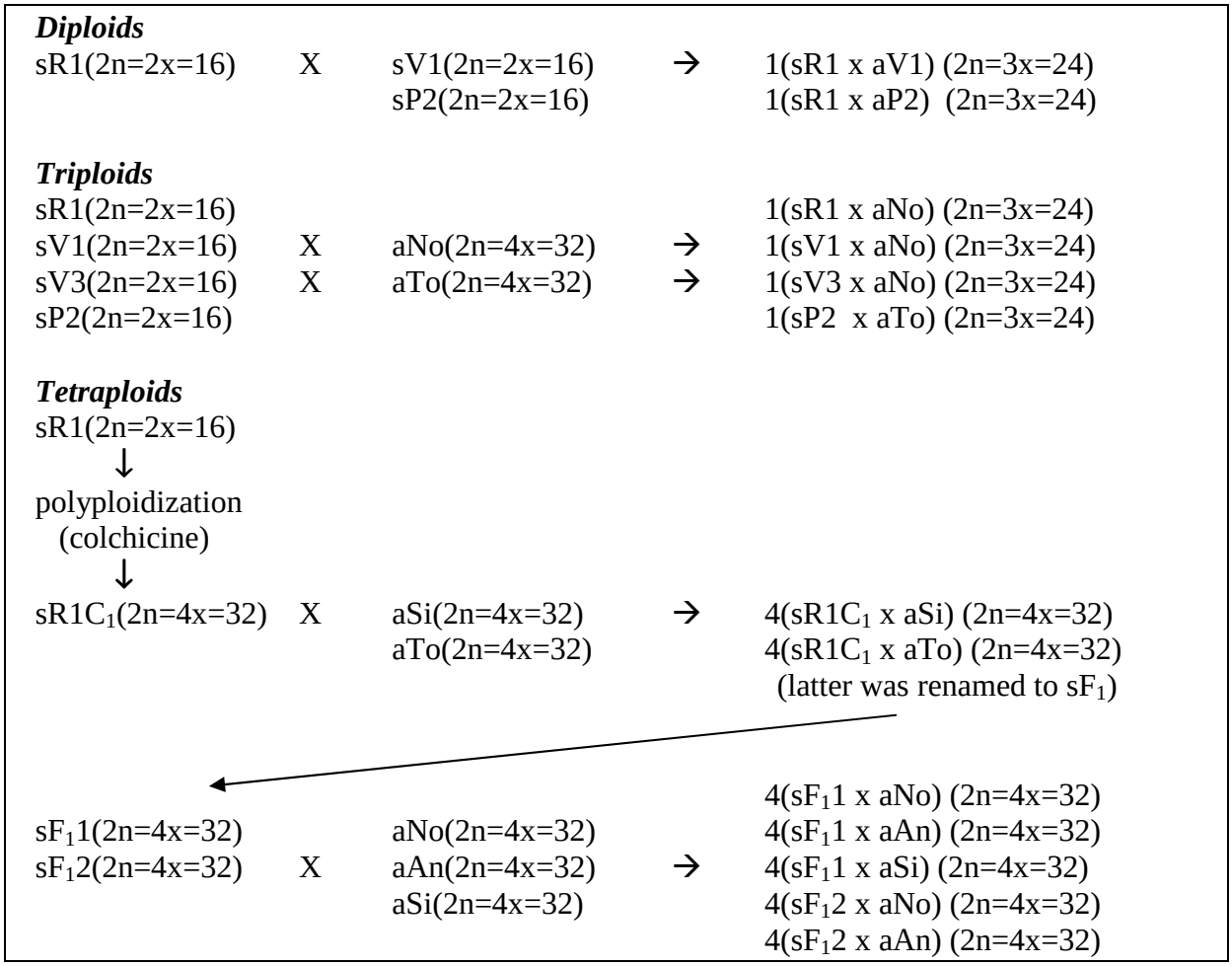


Figure 5: Generation of a population segregating for the mode of reproduction: crossing scheme

Additionally, *Hypericum* wild populations from Europe and North America containing tetraploid apomicts and diploid sexuals were screened with genomic markers. The origin of the populations and the number of plants are presented in the Table 2.

Origin	Nr of plants
Apomicts, tetraploid	
Tuscola IL US 39,8 88,28	9
Green Lake WI US 43,85 89,3	2
Point Beach WI US 44,26 87,56	2
Kewaunee MI US 44,61 88,11	2
Gillett WI US 44,84 88,61	2
Rideau River ON Canada 45 75, 62	5
Menominee MI US 45, 21 87. 75	1
Wausaukee WI US 45, 45 87, 63	2
Carney MI US 45,6 87,03	1
Iron Mountain MI US 45, 83 88, 08	4
Tecumseh MI US 42 89, 66	2

Cazadero CA US 38, 55 123, 13	4
Covelo CA US 39,78 123,25	1
Holiday CA US 40,61 122,3	4
Gasquet CA US 41,84 122,63	2
Mt Shasta CA US 41,3 122,25	2
Abrams Lake Rd CA US 41,33 122,25	2
Weed CA US 41,43 122,38	3
Corvallis OR US 44,61 123,2	3
Granera Spain 41,73 2,06	1
Bolzano Italy 46,51 12,15	1
Clapier France 44,1 7,41	3
Adliswil Switzerland 47,3 8,56	1
Praha Czech 50,08 14,46	2
Suu Ravine Kyrgizstan 42.65 74.51	1
Sexuals, diploid	
Italy	13
Hungary	7
Russia	3
Bulgaria	5
total	90

Table 2: Origin of the wild populations used for the apospory allele quantification *via* pyrosequencing

Genus: *Arabidopsis thaliana*

Arabidopsis thaliana L. cv. „*Columbia*“ diploid

Arabidopsis thaliana L. cv. „*Columbia*“ tetraploid (From Prof. Geelen, Ghent University)

2.1.2. Bacterial strains

Bacterium strain	**features**
Escherichia coli DH5α	RecA1, endA1, gyrA96, thi-1, hsdrR17, (rk-mk+), relA1, supE44, u80ΔlacZΔ15, Tn15 (Sambrook *et al.*, 2001)
Escherichia coli XL1Blue	RecA1, endA1, gyrA96, thi-1, hsdR17, relA1, lac[F'proABlacI^{q}ZΔM15, Tn10, (Tet^{r})]^{c} (Jerpseth *et al.*, 1992)
Escherichia coli DB3.1	Invitrogen
Agrobacterium tumefaciens pGV 2260 in C58C1	(Deblaere *et al.*, 1985)

2.1.3. Plasmids

Vector	features	source
pDONR 201	Kanamycin^r , Gateway donor vector	Invitrogen
pBluescript SK(+)	Ampicillin^r	Stratagene
pCR2.1	Ampicillin^r, Kanamycin^r	Invitrogen
pCC1BAC	Chloramphenicol^r	EPICENTRE Biotechnologies
pUbiH-Nos	Ampicillin^r	GeneCloningService
pNOS-ABM	Ampicillin^r	GeneCloningService
p6N	Streptomycin, Hygromycin	GeneCloningService
pGEM-T vector	Ampicillin^r	Promega

2.1.4. Media

LB **(**1 L): 10g NaCl, 5 g Tryptone, 5 g Yeast extract (pH 7.4)

TBY **(**1 L): 5 g $MgSO_4 \cdot 7H_2O$, 5 g NaCl, 10 g Tryptone, 5 g Yeast extract (pH 7.4)

YEB **(**1 L): 0.5 g $MgSO_4 \cdot 7H_2O$, 5 g Beef extract, 5 g Peptone, 5 g Sucrose, 1 g Yeast extract (pH 7.0)

All solidified media contain 1.5% Difco-agar.

Rich medium for Arabidopsis (1L): 4.3 g MS micro- and macrosalt mixture, 10 g Sucrose, 10 ml Vitamin solution, 8 g Agar (0.8%; for plates) (pH5.8)

Vitamin stock (100ml): 10 mg Thiamine x HCl, 5 mg Pyridoxine x HCl, 5 mg Nicotinic acid, 5 g MES, 1g Inositol

Infection medium (1L): 2.652 g MS micro- and macrosalt mixture, 5 g Sucrose, 10 ml *Arabidopsis* vitamin solution (pH5.7), add 0.04% Silwet L-77

2.1.5. Kits

Invitek, Berlin, Germany	Invisorb Spin Plant Mini Kit
	InviTrap Spin Plant RNA Mini Kit
GE Healthcare, London, UK	RediprimeTM II DNA Labelling Kit
	Hybond N+ membrane

Fermentas, Vilnius, Lithuania	Revert AidTM H Minus First Strand cDNA Synthesis Kit
	Restriction enzymes
	dNTP stock solutions
	Dream *Taq*, *Pfu* polymerases
	T4 DNA ligase
Qiagen, Hilden, Germany	Taq DNA polymerase
	QIAquickR Gel Extraction Kit
	QIAprepR Spin Miniprep Kit
	PCR Purification Kit
Roche, Mannheim, Germany	DNaseI
	Expand Long Template PCR System
Invitogen, Carlsbad, CA, USA	Dual Promoter TA Cloning Kit
Clontech, Palo Alto CA, USA	Universal GenomeWalker Kit
	Advantage2 Polymerase
Macherey-Nagel, Düren, Germany	NucleoSpin 96 Flash
Biotage AB, Upsala, Sweden	Pyro Gold Reagents

2.2. Methods

2.2.1. Flow Cytometric Seed Screen (FCSS)

The mode of reproduction in *Hypericum* plants was investigated using the Flow Cytometric Seed Screen (FCSS), a method which allows the discrimination of apomixis from sexual reproduction based on the seed DNA contents of embryo and endosperm (Matzk *et al*., 2000) and was performed as described previously (Matzk *et al*., 2001). Nuclei of single seeds were isolated by crushing the seeds with sandpaper in a glass petri dish containing 1 ml staining buffer (100 mM Tris-HCl, 5 mM $MgCl_2$, 85 mM NaCl, 0.1% Triton ×100, 1 mg/ml DAPI, pH 7.0). Cellular debris was removed by filtration of the released nuclei through nylon tissue of 30 μm mesh width. One ml of staining buffer was added and the tubes were stored on ice in the dark for 1-2 h. Then the DAPI-stained nuclei suspensions were processed using a Facstar Plus flow cytometer and sorter (Becton-Dickinson, San José, CA, USA) equipped with an argon ion laser in UV mode. Classification of sexual and aposporous genotypes was done on the basis of the cellular DNA contents in embryo and endosperm cell in seeds. In diploid genotypes sexual reproduction is leads to a diploid embryo and a triploid endosperm. Seeds of tetraploid genotypes exhibit four distinct ploidy ratios of embryo and endosperm. (i) 4*C* embryo and 6*C* endosperm cells are a product of sexually produced seeds from reduced, double fertilized embryo sacs; (ii) 4*C* embryo and 10*C* endosperm cells are derived from apomictically produced seeds *via* unreduced embryo sacs and pseudogamous endosperm formation; (iii) 2*C* embryo and 6*C* endosperm are a product of apomictically produced seeds from reduced and parthenogenetically developed embryo and pseudogamous endosperm formation; (iv) 6*C* embryo and 10*C* endosperm are derived from apomictically produced seeds *via* unreduced fertilized embryo and pseudogamous endosperm.

2.2.2 Cytohistological technics to analyse sporogenesis and gametogenesis

Flower buds were sampled at different developmental stages, according to the length, and divided into a total of seven stages. For the length of the flower bud, the distance between the insertion point of the external carpels to the receptacle and the apex of the flower bud was considered. Flower buds ranging from a minimum of 4 mm and a maximum length of 11 mm

were analyzed for each plant. Pistils were dissected from flower buds under a Zeiss Discovery.V20 (Carl Zeiss MicroImaging, Germany) stereomicroscope prior to subsequent staining procedures. Ovules were subsequently dissected onto a microscope slide and cleared with chloral hydrate:water:glycerol (8:2:1) prior to observation. Pistils were alternatively cleared and stained following the protocol reported by Stelly *et al.*, (1984) with some minor modifications. Briefly, the tissues were fixed in FAA (3.7% formalin, 5% acetic acid, 50% ethanol) overnight at 4°C, and then hydrated in 50%, 75% and 100% progressive water : ethanol solutions for 30 min each. Samples were stained with pure Mayer's hemalum for 48 hrs, placed in 2% acetic acid for 24 h, and then dehydrated in 25%, 50%, 70%, 95% and 100% progressive ethanol solutions for 40 min each. After dehydration, samples were cleared with absolute ethanol : methyl salicylate solutions (2:1 and 1:2, v/v) and twice with pure methyl salicylate (10 min per step). Ovules were then mounted with one drop of pure methyl salicylate and coverslipped. Cytohistological observations were made using a Zeiss Axioplan (Carl Zeiss MicroImaging, Germany) microscope under DIC optics, using a 100X objective.

Decolorized aniline blue (DAB; 0.005% w/v) was used to detect the presence of callose as described by Worrall *et al.,* (1992). Ovules were dissected from fresh pistils directly into DAB : glycerol (1:1, w/v) under a Zeiss Discovery.V20 (Carl Zeiss MicroImaging, Germany) stereomicroscope. After ovule isolation, samples were coverslipped and observed under UV light using a Zeiss Axioplan (Carl Zeiss MicroImaging, Germany) microscope with a 365-400 μm filter set.

2.2.3 Basic cloning methods and sequencing

Insert preparation: PCR products were purified using QIAquick PCR Purification Kit or were separated on agarose gel and eluted from the gel by use of PCR Product Isolation Kit. Restriction enzymes and nucleotides were removed by QIAquick Nucleotide Removal Kit. Purified fragments were used for the ligation reaction.

Vector preparation: Plasmid extraction and purification was done using Qiagen Spin Kit or by Qiagen Plasmid MAXI Kit or according to the fast preparation method described by Holmes and Quigley (1981). The standard molecular cloning methods (restriction, digestion, ligation, DNA gel electrophoresis) were performed according to Sambrook *et al.,* (1989).

Transformation of E. coli: The transformation of *E. coli* was performed using the heat-shock procedure (Cohen *et al.*, 1972).

DNA sequencing*:* DNA sequencing was performed at the IPK-Gatersleben at BigDye terminator cycle sequencing chemistry (Applied Biosystems, Foster City, CA, USA) and the 3730x1-DNA sequencer.

2.2.4. AFLP and CAPS marker analysis

Total genomic DNA was isolated using the Invisorb Spin Plant Kit (Invitek, Berlin, Germany) and the AFLP analysis was performed as described by Potokina *et al*., (2002). Fluorescence labelled *Pst*I-anchored primers with two selective nucleotides (5′-GTAGACTGCGTACATGCAGNN-3′) and *Mse*I-anchored primers with three selective nucleotides (5′-GATGAGTCCTGAGTAANNN-3′) have been used for selective amplification. Fragment sizes were determined by comparison with a size standard (Genescan-500 Rox, Applied Biosystems Inc., Foster City, CA, USA) supplemented with five additional DNA fragments ranging from 568 to 812 bp. Gels were analyzed with the GeneScan software version 3.0 (Applied Biosystems Inc., Foster City, CA, USA).

To isolate specific amplicons they were separated on 4.5% polyacrylamide gels, visualized by silver-staining (Bassam *et al*., 1991), excised from the dried gel and eluted overnight by shaking in sterile water. An aliquot of the eluate was used to re-amplify the fragment by PCR. The following PCR conditions have been used: denaturation at 94°C for 5 min, 30 cycles at 94°C for 1 min, 55°C for 1 min and 72°C for 2 min and a final extension step at 72°C for 10 min. Amplicon sequences were either obtained directly after elution from a 1.5% agarose gel or after subcloning into a pGEM-T vector (Promega, Madison, WI, USA).

For the conversion into a CAPS marker, site-specific primers (5′-TAAAGGAGTATC TGAAAAGTGAGG-3′ and 5′-CAGGCAGCACGAGAATCTA-3′) were designed on the basis of the sequence data obtained from the excised AFLP amplicon. The PCR reaction was done in 20 µl containing 50 ng genomic DNA, 1U *Taq* DNA polymerase, 1× PCR buffer (Qiagen, Hilden, Germany), 10 pmol of each primer and 0.25 mM dNTPs. Following PCR conditions were used: initial denaturation at 94°C for 5 min; 30 cycles of 94°C for 30 s, 60°C

for 30 s, 72°C for 1 min, and a final step at 72°C for 10 min. Restriction was done in a volume of 20 µl overnight at 37°C with 6U *Eco*RI, 1× *Eco*RI buffer and 15µl of the PCR product. The CAPS marker status was analyzed on a 2% agarose gel.

2.2.5. BAC library characterization and screening

One-month-old seedlings of a diploid sexual plant were used for DNA isolation. A BAC library was generated by the company Amplicon Express (Pullman, USA) based on partial *Hind*III digestion and cloning in the vector pCC1BAC (EPICENTRE Biotechnologies, Madison, Wisconsin, USA). A total of 26,000 BAC clones were obtained representing approximately a 6-fold coverage, assuming a *Hypericum* genome size of about 630 Mbp (*i.e.*, 1*C*=0.650 pg, http://www.rbgkew.org.uk/cval; 1 pg=965 Mbp, Bennett and Leitch, 1995). Insert sizes were determined for 72 randomly selected BAC clones. The inserts were liberated by *Not*I digestion and sized by PFGE (1% agarose gel, 0.5× TBE, 12.5°C, 6V/cm, 5 s initial and 15 s final pulse time, 16 h). The BAC library was spotted on Hybond N+ membranes (GE Heathcare, London, UK) using the MicroGrid II robot (BioRobotics, UK). The marker sequence containing probe was labelled with ^{33}P-α dCTP using random hexamer priming (Feinberg and Vogelstein, 1983). Hybridization and washing conditions were according to Church and Gilbert (1984). Hybridization was performed overnight at 65°C. The membrane was washed twice with 2x SSC/0.1% SDS, twice with 1x SSC/0.1% SDS, and once with 0.5x SSC/0.1% SDS at 65°C for 15 minutes each. The signal was detected by BAS-IP MS 2325 imaging plate from Fujifilm and was read by FLA 5100 phosphoimager. Positive BAC clones were tested for the presence of the probe *via* PCR with the CAPS primers and used for fingerprinting of the 3′ and 5′ ends according to Luo *et al.*, (2003). The fingerprinting reaction was performed in triplicate using *Bam*HI, *Eco*RI, *Xba*I, *Xho*I, and *Hae*III restriction endonucleases. Capillary electrophoresis of the digested DNA, labelled with SNaPshot labeling solution was carried out on an ABI 3730 Genetic Analyzer (Applied Biosystems Inc., Foster City, CA, USA) and fragment analysis and contig assembly was done using the software packages Gene Mapper V4.0 (Applied Biosystems Inc., Foster City, CA, USA), FPPipeliner V2.0 (BioinforSoft LLC) and FPC V9.0. (Sanger Institute).

2.2.6. BAC clone sequencing

The BAC clone H25H09 was chosen for sequencing, because it contained the marker sequence approximately in the middle. BAC DNA was isolated using a Plasmid Purification Maxi Kit with the low-copy plasmid/cosmid protocol (Qiagen). For subcloning BAC DNA was randomly sheared using a GeneMachines Hydroshear (Genomic Solutions, Ann Arbor, MI, USA) and fractioned by size on 0.8% agarose gel. Fragments of 1-2 kb were purified, blunted using the Klenow fragment of DNA polymerase (Fermentas, Vilnius, Lithuania) and ligated into the pBlueskriptSK (Stratagene) vector. The sequencing was done on a 3730x1-DNA sequencer (Applied Biosystems Inc., Foster City, CA, USA). The sequences were assembled using Sequencher 4.0 (Gene Codes Corporation MI, USA) software set to a minimum overlap of 20 bp and 95% match requirement. Remaining gaps were closed with primers derived from the flanking sequence. The final assembly contained 1,253 sequences coveragingthe assembly 6.1 times on average. The accuracy of the assembly was confirmed by comparing *in silico* and experimental restriction digests for eight different restriction endonucleases (*Eco*RV, *Acc*651, *Apa*I, *Bam*HI, *Sac*II, *Spe*I, *Xho*I, *Sma*I). The size of large fragments was determined by pulse field gel electrophoresis as described above and by conventional agarose electrophoresis for smaller fragments. The band sizes were evaluated with the TotalLAB TL120 software and compared to the restriction patterns predicted based on the sequence.

To detect the collinearity between the *Hypericum* and poplar the 25H09 BAC was blasted against a database containing poplar genome assemblies using the BLASTN and BLASTX scripts adapted for PEARL (L. Altschmied, personal communication).

2.2.7. BAC clone annotation

The total BAC sequence of 141.941 kb was used for gene annotation. Gene predictions are based on consensus gene models derived from several sources of evidence. *Ab initio* gene predictions were carried out applying the programs Fgenesh+ with a dicot markov model and GeneID (Parra *et al*., 2000) using matrices specific for *Arabidopsis* and tomato. Due to the lack of a *Hypericum* EST database, TIGR transcript assemblies (Childs *et al.*, 2007) of related species of the *Saliceae* (*Populus* sp.) and the *Euphorbiaceae* (*Manihot*

esculenta and *Euphorbia esula*) as well as *Arabidopsis thaliana* proteins (TAIR version 7) were aligned as heterologous experimental evidences to the genomic sequence. Alignments were computed as optimal spliced alignments applying the program GenomeThreader (Gremme *et al*., 2005). Consensus gene models were derived by fitting initial gene models to a reference protein database (UNIREF90; Suzek *et al*., 2007). Manual inspection of the consensus gene models retained a total of 4 transposon-related and 24 protein coding genes as candidate gene set of BAC Hp25H09. The annotation of the BAC clone was done in cooperation with G. Haberer (MIPS, Munich).

General sequence data handling was performed using the programs BLASTN and TBLASTX (Altschul *et al*., 1990) and the NCBI databases. The SeqEd programme (ABI Applied Biosystems Inc., Foster City, CA, USA) and Sequencher 4.0 (Gene Codes Corporation MI, USA) software were used for sequence editing.

2.2.8. Detection of single nucleotide polymorphisms

The corresponding 23 annotated genes on the sequenced BAC were amplified from four aposporous and four sexual plants using the primers given in Table 2. The mixture of Dream *Taq* (Fermentas, Vilnius, Lithuania) and *Pfu* (Fermentas, Vilnius, Lithuania) polymerases were used to minimize the amplification mistakes. The amplicons were cloned in the pCRr2.1 vector (Invitrogen, Carlsbad, CA, USA) and at least 16 colonies per gene were sequenced to analyse single nucleotide polymorphisms and haplotype structure. The plasmid isolation was done in a 96 well-plate format using NucleoSpin 96 Flash (Macherey-Nagel, Düren, Germany). The haplotypes were evaluated for the polymorphisms with MEGA 4.1 (Tamura *et al.,* 2007) software.

Amplified region name	BAC-position	Forward primer	Reverse primer	Fragment length
ARIADNE	61782 - 64755	GGGACCCAAAAACTCTCCTC	GCAGTGAACTCAACCGCATA	2973
ARIADNE promoter	61075 - 64755	TTTCTTCCTGTCCTTGTACCG	GCAGTGAACTCAACCGCATA	3680
ARIADNE- truncated	63169 - 65015	TAAAGGAGTATCTGAAAAGTGAGG	CCATGAACCAAGGCACAAC	1846
HK1	7538 - 8192	CAGGTGGAAAAGTTGGGAGA	TTTGAGCTCCACTCGTTCCT	654
HK2	31509 - 32162	CTGTTCTTGTCGGGCACATA	CTCATTAATGTGCGCTGCTG	653
GK1, GK2	37103 - 37718	AAGCTGCTTCTAGTTGTGCACTT	GTAGCGGCAGGAGCTATGAC	615
ExP1	78262 - 79000	GACTTTCGCGTACGTCCTGT	GAAGAAGCCGCAGAGGTAGA	738
PAT1	53041 - 53774	GGAGGTCAACGATTTCAGGA	CTGCACCTCACAGCCAGATA	733
WRKY	55854 - 56394	GCATTCTCGTGCTTCTCCTC	AAGAGCATGGATCGTGGAAC	512
DGCR	60011 - 60424	CGAGCGAGGATAATGAGAGC	CCCCTTCTCACCATCCCTAT	413
NPH3	68725 - 69315	GCATCGAGTCGTTAGCCTTC	TCTCCATGGTAGCCAACTCC	591
TK	75087 - 75512	GCAGCAACCTCGTGAACTTT	GATGGAAAGATCGTGGTCGT	425
RNP	81181 - 81613	GTTTATGCCCATCCCATGAC	GGAGGTTTGGCATCAACTGT	432
AUXR	89472 - 89907	ACGTGGAGCTTGGTCTTGTT	TTGCACCACCGTCACTTCTA	435
MO25	106317 - 106804	CTGGACGATCACCAATCTCA	CCATCGTTCTATCCCACACC	487
RING H-2	125473 - 125803	TCTCCTCTCCCGAAAGTTCA	GGTCTTCTCGGTGTCGTTGT	330

Table 2: Primer sequences used for BAC Hp25H09 characterization.

Gene symbols, position within the BAC sequence, used forward and reverse primers and the resulting fragment lengths for all genes present on the BAC are given. HK1,2 - histidine kinases, GH1,2 - glycosyl hydrolases ExP1 - expressed protein, PAT1 - a member of the GRAS gene family, WRKY - a WRKY transcription factor, DGCR a hypothetical protein, ARI - an ubiquitin-mediated E3-ligase, NPH3 – a plastidal phototropic-responsive protein, TK - a thymidine kinase, RNP - a heterogeneous nuclear ribonucleoprotein, AUXR - an auxin-response protein, Mo25 - an unknown protein of the Mo-25 family, RINGH2 - RING finger protein of the H2 class .

2.2.9. Allele quantification by pyrosequencing

Primers for pyrosequencing were derived using the SNP primer Assay Design Software (version 4.0) from Biotag AB (Uppsala, Sweden). The primers used are given in Table 3. After PCR with genomic DNA, pyrosequencing was carried out according to the manufacturer's standard protocol with Pyro Gold Reagents (Biotage AB) on a Pyrosequencer PSQ 96HS96A 1.2 machine (Biotage AB) The allele frequency was estimated using the manufacturer's software.

SNP no.	Forward primer	Reverse primer (biotinilated)	Sequencing primer
SNP 06	GGACCCAAAAACTCTCCTCTCA	CTCGTCGGAATCCATAGCTC	GCTTCTGGTAATTAGGGT
SNP 24, 25	TGGGACGTGAGTAAGTTGCATGAT	TGTCCACAAGAAGAAGCCGAATA	CGACGAGGAAGCAGTG

Table 3: Pyrosequencing assays developed for *HpARI*

2.2.10. Genomic Southern

For southern hybridization 5 µg of genomic DNA were digested with *Eco*RI and blotted onto a Hybond N+ membrane (GE Heathcare, London, UK). The marker sequence containing probe was labelled with ^{33}P-α dCTP using random hexamer priming (Feinberg and Vogelstein, 1983). Prehybridization was performed at 65°C for 2 hours and hybridisation was done at 65°C overnight. After hybridisation the membrane was washed twice with 2x SSC/0.1% SDS, twice with 1x SSC/0.1% SDS, and once with 0.5x SSC/0.1% SDS at 65°C for 30 minutes each. The signal was detected by BAS-IP MS 2325 imaging plates from Fujifilm. Band intensities (PSL/mm^2) were quantified with a phosphoimager FUJIX Bas 2000 (Fuji Photo Film, Tokyo, Japan). After background subtraction the intensities of the 583 bp and 394 bp bands were added and compared to the intensity of the 978 bp band.

2.2.11. Genome walking

The Genome Walker Universal Kit (Clontech, Palo Alto, CA, USA) and Advantage 2 Polymerase mix (Clontech) was used for genomic walking following the instructions of the manufacturer. Two rounds of walking have been done using primers given in Table 4.

	Primer name	**Gene walking round 1**	**Gene walking round 2**
5' end	Gene specific primer 1	CGCCATGGAACACCAATGGCAATGATCCA	TGTCTTTGCCTAGCCATGACATGAATC
	Gene specific primer 2	GTCTGGTTGGATTAAGCCAAGAAATC	TCGGTCTGATACCCGAAGGCAGTATAG
3' end	Gene specific primer 1	CGCGTTTAACTTTGGAGTTCTTTCAGG	CTTAGTTCATGGCAAGAGCATCGTGAG
	Gene specific primer 2	TCGAAATCCTCTCACTAACACACAGGT	GACATGGAAAATGGTGCAAACCCTTAG

Table 4: Primers used for genomic walking

2.2.12. Expression analysis

Total RNA was isolated from pistils 4-5 mm and 6-8 mm and leaves using Invitek RNA Isolation Kit (Invitek, Berlin, Germany). These pistils correspond to megasporogenesis and megagametogenesis, respectively, as determined by staging analysis (Galla *et al.*, in press). For the production of the first strand cDNA 1 µg of total RNA was treated with DNAaseI (Invitrogen, Carslbad, CA, USA) followed by RT PCR with RevertAid H Minus M-

MuLV Reverse transcriptase (Fermentas, Vilnus, Lithuania) using random hexamer primers at 42°C for 60 minutes. The primers TCCGACGAGGACGTGTATTA and TTATAATGACGGAGGGGATGCT were designed to amplify both sexual and aposporous alleles, and the amplicon was digested with *Eco*RV (Fermentas, Vilnus, Estonia). The PCR step was performed using the programme: 94°C, 2 minutes, 35 cycles, at 94°C for 30 seconds, 60°C for 30 seconds and 72°C for 60 seconds with a final elongation step at 72°C for 5 minutes, with Qiagen *Taq* polymerase (Qiagen, Germany).

2.2.13. Generation of transgenic lines

The regions containing *HpARI* gene and *HpARI* promoter with *HpARI* gene were amplified using *Taq* and *Pfu* polymerases mixture (Fermentas, Vilnus, Estonia) and cloned to *Ava*I and *Hind*II restriction sites into pUbiH-Nos (GeneCloningService, Hamburg, Germany) pNOS-ABM (GeneCloningService, Hamburg, Germany) respectively. The ligation was done in 10 µl volume, with 50 ng vector, 200 ng PCR product, 2.5U T4 DNA ligase (Fermentas, Vilnus, Estonia) and 1 µl T4 DNA ligase buffer at 14°C overnight. The plasmids positive for the inserts were selected after *Ava*I and *Hind*II restriction. The cloning cassette was cut out by *Sfi*I enzyme (Fermentas Vilnus, Estonia) and ligated into the binary p6N vector. The positive clones were selected with *Sfi*I enzyme and sequenced to check the correctness of the reading frame and absence of mutations.

The constructs were introduced into the *Agrobacterium tumefaciens* strain GV2260 by freeze-thaw transformation (Chen *et al.*, 1994). Transformation of Arabidopsis was done according to the protocol of (Clough and Bent 1998). Plants of *Arabidopsis thaliana* ecotype *Columbia* were grown for three weeks under short day conditions (8 hours light, 16 hours dark) and transferred to long day (16 hours light, 8 hours dark). After three weeks, the emerging bolts were cut to induce growth of secondary bolts. Bacteria were grown till OD600>2.0, harvested by centrifugation and resuspended in three volumes of infiltration medium supplemented with 0.01% Silwet L-77 (Union Carbide Chemicals and Plastics). Inflorescences were dipped into the *Agrobacterium tumefaciens* (strain GV2260) suspension for about 1 minute. After short shaking they were incubated horizontally in covered bowls for 24 hours to keep a high humidity. After 24 hours, they were uncovered and set upright. Seeds

were harvested from dry siliques, sterilized and plated on selection plates containing MS basal medium (Murashige and Skoog, 1962) complemented with 1 % sucrose and 0.8 % agar (Sigma M-9274) and the appropriate 50 μg/mL selective antibiotic. After two weeks, viable plants were transferred to soil, grown up and their seeds collected. Transgenic plants were analyzed by PCR to confirm the successful transformation.

3. Results

3.1. Segregation of apospory and parthenogenesis

In order to examine if apospory and parthenogenesis segregate as independent traits in *Hypericum*, a crossing scheme involving diploid and tetraploid sexual and tetraploid apomictic lines was established. Four apomictic lines (An, No, Si, To) used as pollinators in segregation studies exhibited from 83% to 100% apospory and from 80% to 100% parthenogenesis (Table 5).

Line	Mode of reproduction (%)		
	Apospory	Parthenogenesis	Apomixis
aAn	95	86	85
aNo	100	~100	~100
aSi	88	81	81
aTo	83	80	78

Table 5: Selection of lines with highest expression of apomixis used in the crosses

Both the diploid sexual parents (sR1, sP1, sP2, sV1, sV2, sV3) and the tetraploid sexual parents ($sF_1$1, $sF_1$2, sR1C) were propagated separately (Table 6, for the scheme of the crosses see material and methods, Figure 5). These lines were confirmed to be obligate sexual, neither apospory nor parthenogenesis were detected by FCSS. The crossing of obligate sexuals resulted in obligate sexuals (Table 6), suggesting that the parents are most likely homozygous at the loci responsible for sexual seed formation.

In triploid F_1 plants resulting from a cross between tetraploid apomicts and diploid sexuals, the ratio of aposporous and sexual plants was approximately 3:1 (44:15, see Table 6), showing a clear dominance of apospory at the triploid level. This ratio is intermediate between 1:1 and a 5:1 gamete segregation to be expected when assuming either one or two dominant alleles (*Aaaa*, simplex *versus AAaa*, duplex genetic constitution) of one responsible locus in the tetraploid male parents, assuming both recessive alleles (*aa*) for the diploid female parents.

Among the F_1 progenies from crosses between tetraploid sexuals (R1C, $F_1$1, $F_1$2) and tetraploid apomicts (An, No, Si, To), 34 plants were aposporous with different levels of

penetrance and expressivity and 38 plants were sexual, confirming a 1:1 segregation ratio (χ^2 = 0.222) between these two modes of reproduction. Most of the aposporous plants (27) showed high expression of apospory (71% to 100%) and only a few plants (7) scored intermediate values (20% to 38%) (See table 15, attachment). The ratio close to 1:1 between sexual and aposporous progeny plants indicates a simplex genetic constitution of one responsible locus in the tetraploid male parents.

Cross	Plants	APO+/PAR+	APO+/PAR-	APO-/PAR+	SEX
Diploid					
2(sR1 x sP2)	11	0	0	0	11
2(sP2 x sV1)	13	0	0	0	13
Total	**24**	**0**	**0**	**0**	**24**
Triploid					
1/2(sR1 x aNo)	16	7	5	0	4
1(sV1 x aNo)	25	12	6	1	6
1(sV3 x aNo)	4	3	1	0	0
1(sP2 x aTo)	14	6	4	0	4
Total	**59**	**28**	**16**	**1**	**14**
Tetraploid					
4($sF_1$1 x aAn)	8	2	0	0	6
4($sF_1$1 x aSi)	4	0	1	0	3
4($sF_1$1 x aNo)	10	8	0	0	2
4($sF_1$2 x aAn)	10	3	0	1	6
4($sF_1$2 x aNo)	24	12	4	0	8
4($sR1C_1$ x aSi)	16	2	2	0	12
Total	**72**	**27**	**7**	**1**	**37**

Table 6: Segregation analysis of the directed crosses of different ploidy

In addition to apospory, the degree of parthenogenesis was determined and found to be zero or rather low (0- 53%). No intermediate or high expression of parthenogenesis was found among the triploids and only a few exceptions were observed among the tetraploid F_1 plants. The ratio of plants with and without (0%) parthenogenesis was 29 (28+1) : 30 (16+14) in triploids and 28 (27+1) : 45 (7+38) in tetraploids (Table 6 and 7), which is close to a 1:1 ratio (χ^2 = 3,556) and suggests a simplex genetic constitution of the responsible locus in the male parent. Plants with a high expression of both apospory and parthenogenesis are rare after crossing obligate sexual and highly apomictic parents. In total only three such plants were found among 155 individuals (Table 15, attachment). The occurrence of 16 triploid and 7 tetraploid F_1 plants which exhibit apospory without parthenogenesis as well as one triploid

and tetraploid plant each that exhibit parthenogenesis without apospory, suggests the independent genetic control of both apomixis components in *Hypericum*.

To test whether recombination between apospory and parthenogenesis had occurred, a joint segregation analysis was performed. This showed a significant deviation from the expected 1:1:1:1 ratio at both the triploid and tetraploid levels (Table 6). Therefore, both these traits are still associated in *Hypericum*. Furthermore, I observed that the recombination frequency between apospory and parthenogenesis was about 19% with the estimated genetic distance between the corresponding loci being 20.1 cM after correction with Kosambi´s mapping function (Table 7). Thus, the two dominant alleles responsible for apospory and parthenogenesis were linked in the coupling phase (e.g. cis-type linkage).

Ploidy of	**Apomixis**	**Recombinants**		**Sexuals**	
progenies	**Apo^+/Par^+**	**Apo^+/Par^-**	**Apo^-/Par^+**	**Apo^-/Par^-**	**Total**
3n	28	16	1	14	59
4n	27	7	1	37	72
Overall	55	23	2	51	131
Ploidy of	**Overall**		**Chi-square values**		
progenies	**Apo (+vs-)**	**Par (+vs-)**	**Apo (+vs-)**	**Par (+vs-)**	**Overall (Apo vs Par)**
3n	44 vs.15	29 vs. 30	9.577*	0.017 ns	24.864***
4n	34 vs.38	28 vs. 44	0.222 ns	3.556 ns	47.333***
Overall	78 vs. 53	57 vs. 74	4.771 ns	2.206 ns	57.061***
Recombination frequency: 19.08%					
Genetic distance (Kosambi mapping function): 20.1 cM					

Table 7: Statistical calculations of sexual and apospory-specific allele segregation in triploid and tetraploid offspring

Symbols: Apo-Apospory, Par-Partenogenesis ns- not significant at P=0.01; *-significant at P=0.01; ***- significant at P=0.001

3.2. Cytohistological observations of apospory in *Hypericum perforatum*

3.2.1. Plants selected for the cytohistological studies and their mode of reproduction

Based upon the flow cytometric profiles of single seeds, the four tetraploid (4C) accessions collected from different geographical regions were all determined to be facultative apomicts, with 79% to 98% of the analyzed seeds per accession developed as a result of an aposporous event and 73% to 90% of the seeds arising after a parthenogenesis event (Table 8).

Plant name	Origin	Mode of reproduction		Ploidy
		Apospory, %	Parthenogenesis, %	
H06_1915	Iron Mountain MI (USA)	98	90	4n
H06_2751	Bolzano, Italy	87	73	4n
H06_2849	Badia Polesine, Italy	79	82	4n
H06_2974	Cerbere, France	98	76	4n
H06_2842	Badia Polesine, Italy	0	0	2n
$(F1_1xAn)1/4$	IPK-Gatersleben[a]	0	0	4n
$(R1C_1xTo)1/1$	IPK-Gatersleben[a]	0	0	4n
$(F_12xAn)1/7$	IPK-Gatersleben[a]	0	0	4n
$(F_12xNo)1a/5$	IPK-Gatersleben[a]	0	0	4n

Table 8: Geographic origin and flow cytometric characterization of unreduced gamete formation (apomeiosis) for *H. perforatum* used in embryological analysis

[a] Induced tetraploid plants via colchicine treatment of seeds (for the scheme of crosses see Material and Methods Figure 5)

3.2.2. Ovule development and gamete formation in sexual *Hypericum*

Early developmental stages of the ovule are shown in Figure 1. As the ovule primordium reaches approximately 15 cells in length, the emerging integument undergoes periclinal division. Integument growth delineates the main funicular, chalazal and nucellar domains within the ovule (Figure 6). The *Hypericum* ovule is bigtemic as both outer and inner integuments differentiate from the middle region of the proximal - distal axis of the ovule (Figure 6 and 7). At this stage the nucellus is about five cells in length and is composed of one

epidermal layer, enclosing one to two hypodermal columns of cells (Figure 6). I found the number of hypodermal cells to be variable between ovules of the same plant, with no obvious correlation with reproductive behavior (Figure 7 C, D).

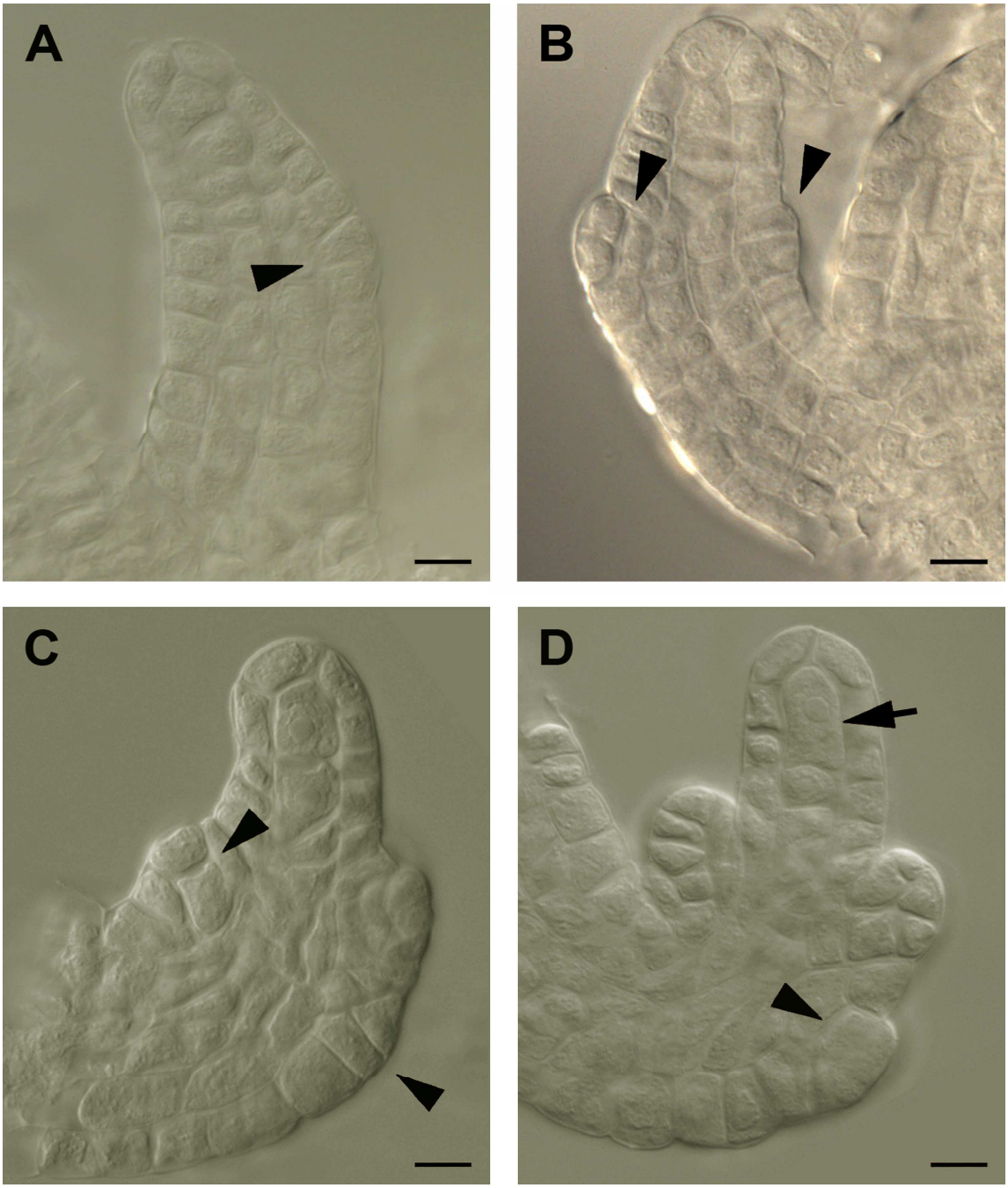


Figure 6: Early phase of ovule development in sexual and aposporous *H. perforatum* plants

A-B. Ovules at stage 1-II (staging according to Schneitz *et al.*, 1995): A- Inner integument initiation in the abaxial side of the ovule (at this stage the epidermal layer is distinguishable from the internal sub-epidermal tissue); B - The development of adaxial inner integuments is anticipated with respect to the abaxial side; C-D. - Outer integument initiation shortly follows the inner integument protrusion (C); the archesporial cell is located next to the epidermids at the tip of the nucellus (C) and forms a sub-epidermal MMC (D). Arrowheads: sites of integument formation. Arrow: MMC. All scale bars = 17 µm.

As the internal integument starts to form, the evident archesporial cell differentiates just beneath the most apical epidermal cell. No intermediate division of the archesporial cell prior to MMC differentiation was observed (Figure 6). Callose deposition patterns within sexual ovules fully resemble the pattern previously described by Rodkiewicz (1970) for the monosporic *Polygonum* type of embryo sac formation (Figure 8 B-E). The fluorescence is strongly localized at both poles of the MMC, with crescent-like accumulation that finally encompasses the complete internal surface of the cell (Figure 8 B). The first meiotic division and cytokinesis lead to marked callose deposition within the middle cell wall (Figure 8 C), and at this time little or no callose is present within the proximal-distal apex of the dyad. Moreover callose is clearly detectable within the young tetrad stage (Figure 8 D and E), being strongly accumulated among newly originated megaspores. Late tetrads are distinguishable from younger ones by the massive accumulation of polysaccharides around all but the functional megaspores. As in *Arabidopsis* and maize, only the most proximal megaspore survives and undergo further development (Figure 7 D, E). The complete degeneration of the most micropylar megaspores is accompanied by the onset of megagametogenesis which involves the enlargement of the functional megaspore to give rise to the one nucleate embryo sac (Figure 7 E).

The first mitotic nuclear division of the one nucleate embryo sac leads to the formation of a two nucleated embryo sac (early 2N ES stage). The positioning of the two nuclei within the embryo sac is strictly defined as they are always detectable in the proximal and distal sides of the cell. Prior to the second mitotic division, the embryo sac markedly increases in length, almost reaching the most micropylar side of the ovule (late 2N ES stage; Figure 7 F). The second nuclear division of the proximal and distal nuclei is highly synchronized and results in the formation of a four nucleate embryo sac with two nuclei oppositely localized at each side of the cell (Figure 7 G). A precise and conserved pattern of nuclear positioning within the cell is observed at this stage, suggesting that this is under strict control. More specifically, the nuclei are positioned one above the other with respect to the chalazal pole of the micropylar/chalazal axis, while the nuclei generally are located side by side at the micropylar end (Figure 7 G). The third nuclear division closely follows the second one to generate the final eight nucleated embryo sac (8N ES; Figure 7 H and I). As cytoplasm and

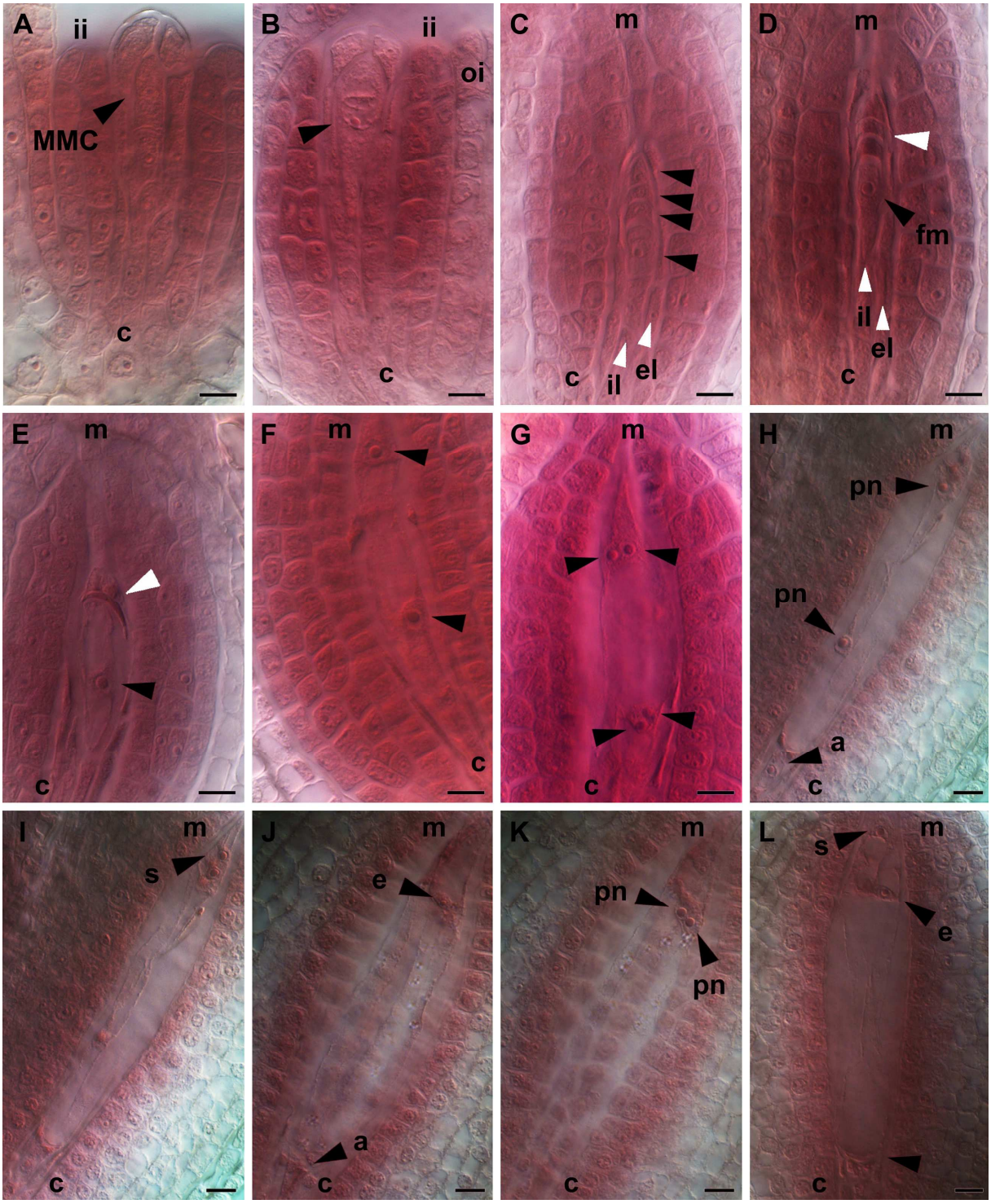


Figure 7: Megasporogenesis and megagametogenesis progression in sexual *Hypericum*

For each panel the micropylar (m) and chalazal (c) side of the ovules are reported. A-D. Megasporogenesis. A: MMC; B: Dyad; C: Tetrad; D: Functional megaspore (fM). E-L. Megagametogenesis. E: 1N ES; F: 2N ES; G: 4N ES; H-K: 8N ES spanning the time point of antipodal degeneration; L: mature ES. Degeneration of the most micropylar megaspores is marked with the white arrow head. Nuclei within megaspores and embryo sacs are marked with black arrow heads, except in J and L where the black arrow heads in the chalazal region marks the degenerating antipodals. oi: outer integument; ii: inner integument; il: inner layer; el: external layer; pn: polar nuclei; a: antipodal; s: synergid; e: egg cell. 1ES, 2ES, 4ES, 8ES: one nucleate, two nucleate, four nucleate and eight nucleate embryosac repectively. All scale bars = 8 µm.

organelle partitioning take place, two antipodals and synergids originate from the most externally localized nucleus, while all other constituents of the mature embryo sac are derived from the more centrally localized nuclei (Figure 7 H). The mature, fully cellularized embryo sac is composed of the three antipodals, a bi-nucleated central cell and the egg cell positioned beneath the two most micropylar synergids (Figure 7 L). Synergids and egg cell are clearly characterized by inverted localization of the large vacuole and nucleus one with respect to the other (Figure 7 L). Indeed, while synergids are typically characterized by a distal nucleus and a proximal vacuole, the egg cell is characterized by the presence of a distally-localized vacuole and a proximal cytoplasmic area (Figure 7 L, black arrowheads). Such organization of the egg cell results in the proximity of the large nucleus of the germ cell to the secondary nucleus of the central cell, the latter being distally localized within the large auxiliary cell (Figure 7 L). I observed that degeneration of the antipodal cells always precedes polar nuclei fusion (Figure 7 H, I) and that secondary nucleus formation always precedes synergid cell degeneration and fertilization of the embryo sac (Figure 7 J, K).

3.2.3. Aposporous embryo sac induction and development

Significant deviations from the sexual pathway were observed within aposporous ovules. Isolated ovules from apomictic individuals were typically characterized by an archesporial cell that eventually produced a MMC which apparently mimics the developmental timing observed in sexual individuals. As shown in Figure 9 (panels B-D), even if megaspore-like enlarged cells seem to be present within the hypodermic layer, meiosis frequently stops at the two to four megaspores stage. Furthermore, if distinguishable from the surrounding cells, such megaspores typically carry a small nucleus and exhibit signs of degeneration (Figure 9 A-C). Moreover, the ovules of apomictic individuals are characterized by novel callose deposition patterns compared to those of sexuals, and very rarely show dyads or tetrads (Figure 8). Furthermore, FCSS analysis confirmed that apomictic individuals retained low frequencies of sexual seed formation (Table 8), and thus the possibility that my observations of dyads and tetrads reflect meiosis rather than aposporous apomeiosis cannot be ruled out.

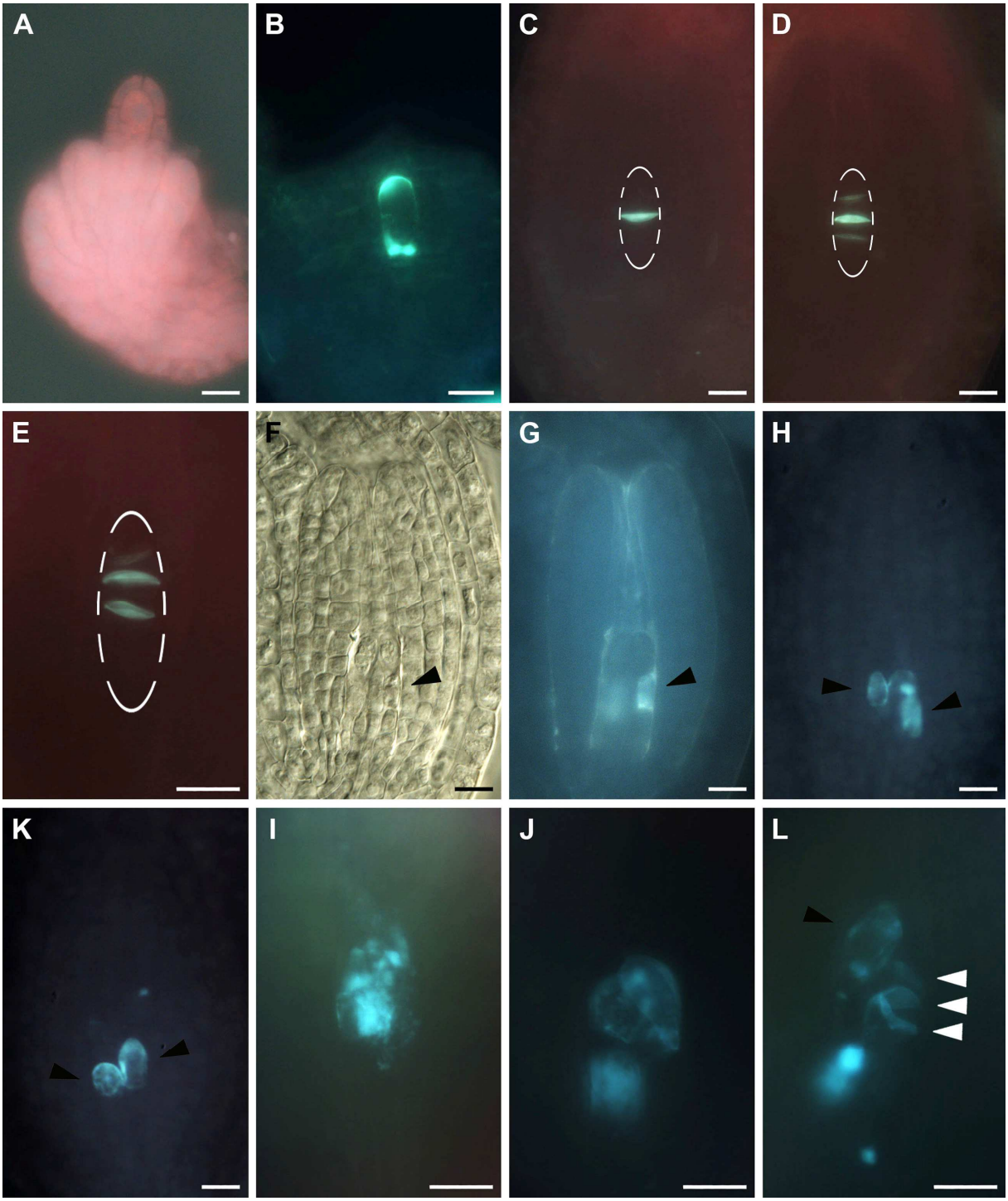


Figure 8: Callose localization detected by DAB staining of sexual (A to E) and aposporous (F to L) ovules at pre-meiotic and meiotic developmental stages

A Pre-meiotic ovule: no callose is deposited within the nucellus; B: MMC; C: Dyad; D-E Tetrads: the middle cell wall is the major site of callose accumulation at this developmental stage. Panels F and G show one ovule isolated from an apomictic individual, photographed under normal and UV light (for callose visualization): fluorescence is present in unconventional sites of deposition. Callose deposition is frequently localized in non-conventional sites (H, K), with spread or spotted distribution (I, J, L). Panels J and L furthermore show two sexual-like patterns of callose localization whereby flanking cells also accumulated callose (compare with E-L). All scale bars = 13 µm.

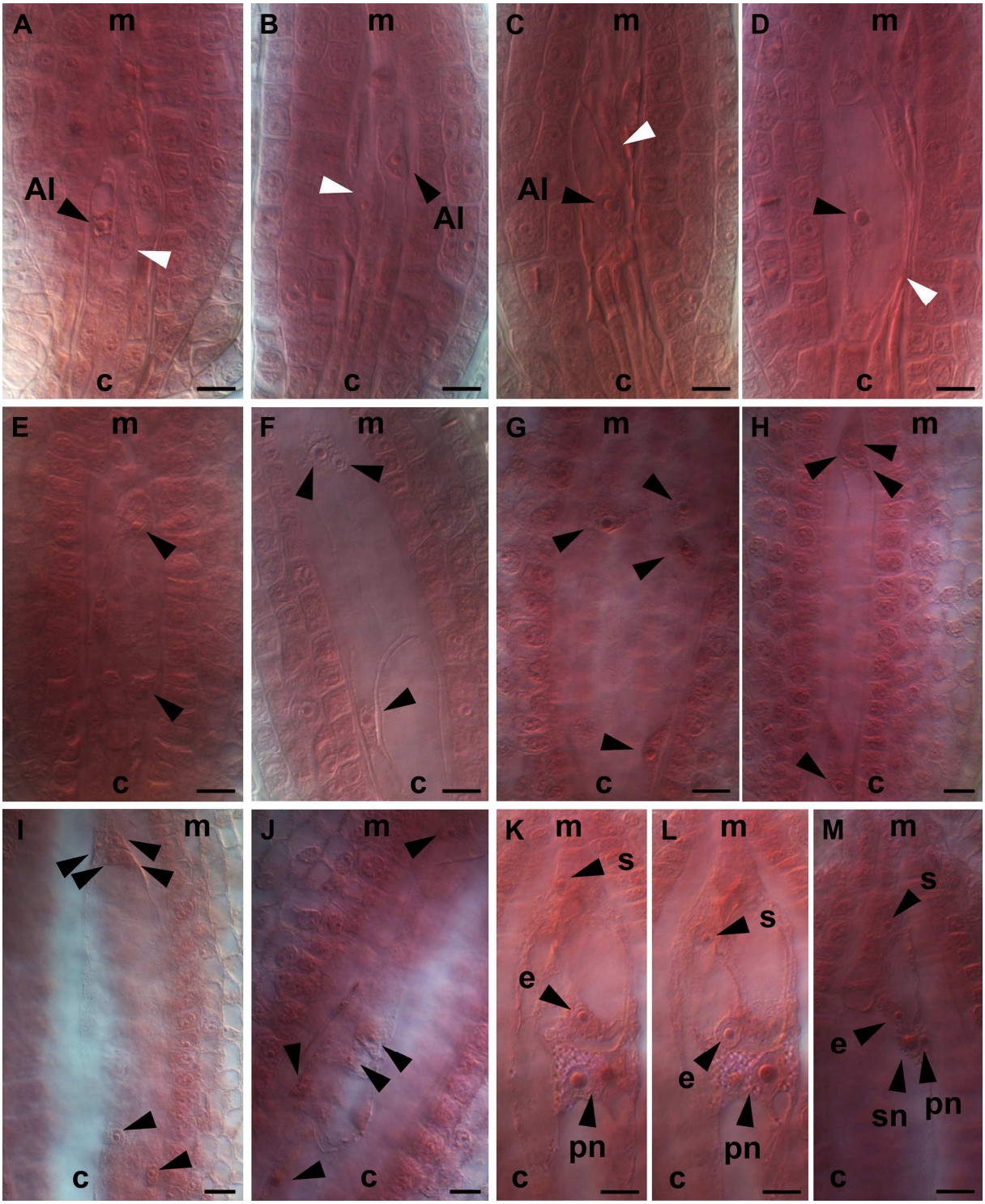


Figure 9: Embryo sac development within aposporous ovules

Micropylar (m) and chalazal (c) sides of the ovules are reported for each panel. A-C: aposporous initial differentiation and enlargement; E-J: coenocytic growth of the aposporous initial which finally gives rise to an embryo sac; K-M: egg cell apparatus with supernumerary cells (K-L) or nuclei (M). From A to F, the cellular layer in which the AI differentiates is clearly detectable. White arrows indicate the expected site of megasporores localization, while black arrows mark the nuclei detectable within coenocytic megagametophytes. s: synergid; e: egg cell; sn: secondary nucleus; pn: polar nucleus; All scale bars = 8 µm.

In contrast to sexual ovules, aposporous ovules are characterized by diffuse fluorescence signals in place of the clearly defined callose deposition normally associated with meiosis progression, and brightly-stained single cells were frequently observed in both hypodermal and epidermal areas of the nucellus, where sporogenesis is not likely to take place (Figure 8). In addition to megasporogenesis arrest, the number and morphology of cells involved with aposporous seed production deviate from that which is characteristic for normal development. These cells are clearly recognizable in the apomeiotic embryo sac, and share a number of traits: i) exclusive to apomeiotic ovules; ii) differentiation from the epidermal layer of the nucellus; iii) cell size considerably increased with respect to the neighboring cells; iv) large vacuoles frequently present along both sides of the cell and a dense cytoplasmic middle region; v) clearly defined large nucleus. Moreover, older aposporous ovules (i.e. larger ovules isolated from longer flower buds) were frequently characterized by the presence of a large developing coenocyte clearly developing from the same target area of the nucellus. Considering these morphological traits and the apparent ability to escape their conventional cell fate, these cellular types were defined as aposporous initials.

AIs are typically elliptical in shape, frequently drop-shaped, and clearly distinguishable from the square shaped neighboring epidermal cells (Figure 9 A-C). Early developmental steps of AIs are characterized by dramatic growth in length and width, and the concomitant displacement of the surrounding, mostly hypodermic, tissues (Figure 9 B, C). Furthermore, the neighboring cells are frequently arrested in development, or appear as degenerating megaspores (Figure 9 B-D). Interestingly, the increase in size and particularly in length of the AIs was comparable to that of enlarging sexual embryo sacs, achieving three to four times the length of surrounding cells prior to any nuclear division (compare Figure 2 E and Figure 9 C, D). The first nuclear division of the centrally localized nucleus of the AI results in the formation of a binucleate coenocyte, whose nuclei are localized towards the apex of the cell in a pattern similar to the early sexual 2N ES (Figure 9 E). After this nuclear migration, a second and third nuclear division usually takes places within the embryo sac (Figure 9 F-J). Frequently the second and third divisions seem to be asynchronous and lead to unconventionally-nucleated coenocytes when compared to the sexual 4N and 8N ES morphologies (compare Figure 7 and Figure 9).

Two main types of deviation from the developmental pattern of sexual embryo sacs were observed within aposporous coenocytes. First, aposporous embryo sacs may contain an odd number of nuclei, resulting in three, five and nine nucleate cells (Figure 9 F, H, K, L). Second, the positioning of nuclei within the embryo sac frequently does not resemble the *Polygonum* type, and results, for instance, in the presence of two to six nuclei in the most micropylar pole of the embryo sac (Figure 9 F-M). While these deviations in the distribution of nuclei are linked with asynchronous cell division, it is unclear whether this is a cause or consequence of it. The 4N ES stage is additionally characterized by a coenocyte which frequently possessed four nuclei localized in the four corners of the embryo sac (Figure 9 G).

The localization pattern of nuclei within the aposporous embryo sac following the third mitotic division frequently differs from that in sexuals. However, as cellularization proceeds, the identification of most cell types was possible considering their position and morphological traits. Normally, one or two antipodal cells are proximally localized in the embryo sac, but they rarely resembled the characteristic triangular morphology observed in sexual ovules (Figure 9 G-J). Furthermore, degeneration of the antipodal cells within aposporous embryo sacs frequently occurs earlier with respect to the sexual ones. The egg cell is always easily detectable soon after cellularization, and as observed in sexual ovules, it is localized at the most micropylar side of aposporous ovules and is characterized by a pronounced proximal nucleus and a large distal vacuole (Figure 9 K, L). Based on cell morphology and positioning within the megagametophyte, ovules bearing multiple egg cells in the same embryo sac were observed in aposporous but not in sexual ovules. Hence, in the most extreme cases, the embryo sac contains degenerating synergids and supernumerary nuclei (Figure 9 M) or cells (Figure 9 K and L). The possibility that eventual supernumerary egg cells are functional could not be ruled out. In contrast, synergids are frequently difficult to localize or atypically shaped and close to degeneration (Figure 9 J-M).

Apomictic individuals typically have ovules bearing large degenerating cells, and sometimes an enormously enlarged nucleus or no nucleus at all, all of which suggest embryo sac degeneration. One to four AIs were recorded within the same nucellus, which eventually led to multiple enlarged coenocytes within the same ovule. The developmental stages of multiple AIs were frequently unsynchronized. Moreover, the developmental stages of multiple aposporous embryo sacs frequently demonstrated a distal-proximal gradient, with the

larger coenocyte reaching the micropylar axis with the newly differentiated AIs enlarging in the chalazal proximal side. In contrast, no obvious adaxial - abaxial developmental gradient was observed. The formation of multiple AIs within the same ovule further enabled to identify the target area of AI differentiation as the epidermal cell layer of the nucellus.

3.3. Development of CAPS marker co-segregating with apospory

3.3.1. A genomic AFLP marker discriminates *Hypericum* genotypes with contrasting modes of reproduction

To gain insight into the genetic control of apomixis in *Hypericum*, genomic AFLP marker display was performed on apomictic and sexual genotypes from unrelated ecotypes as well as on hybrid progenies segregating for apospory. The initial screening based on 10 apomictic and 6 sexual genetically unrelated genotypes (Table 1, material and methods) clearly differentiated for their reproduction mode (aposporous and sexual), as assessed by means of FCSS analysis. Genomic DNA fingerprinting of these 16 genotypes was performed using AFLP technology with 36 *Pst*I/*Mse*I primer combinations. As expected, many polymorphic markers were observed between plants of distinct geographical origin and genetic background. Only one AFLP marker was identified as an apospory-specific candidate since it was shared by all 10 apomictic genotypes and it was absent in all 6 sexual ones (Figure 10). The reliability of this AFLP marker was further confirmed by its consistent co-segregation with the aposporous mode of reproduction in a F1 progeny set of 40 plants. Discovery of this AFLP marker formed the initial basis for characterization of the associated apospory locus.

3.3.2. Developing a CAPS assay with a marker tightly co-segregating with apospory

The AFLP marker was converted into a more easily detectable PCR-based marker. Primers derived from the sequence of the AFLP band were used to amplify the corresponding genomic region. Sequence analysis of this amplicon revealed the presence of two sequences in the amplicon distinguished by the presence or absence of an *Eco*RI site. A test digestion with *Eco*RI confirmed that the 223 bp amplicon was either split into two fragments of nearly the same sizes (105 bp and 118 bp) or it remained undigestable by *Eco*RI. This sequence polymorphism allowed the conversion of the AFLP marker into a CAPS marker. I noticed that a trace amount of sexual alleles were present in aposporous plants, as shown by a faint band of 105/118 bp along with the major band of 223 bp on gel (Figure 11). I was unable to

distinguish the origin of this faint band when the progeny test was performed. Later, when I extended the sequence of the apospory-specific locus, I noticed that the primer sequences allow a preferential amplification of the aposporous allele when present in combination with but the sexual alleles (Figure 19). Nevertheless, the identified CAPS marker allowed to faithfully distinguish between aposporous and sexual modes of reproduction.

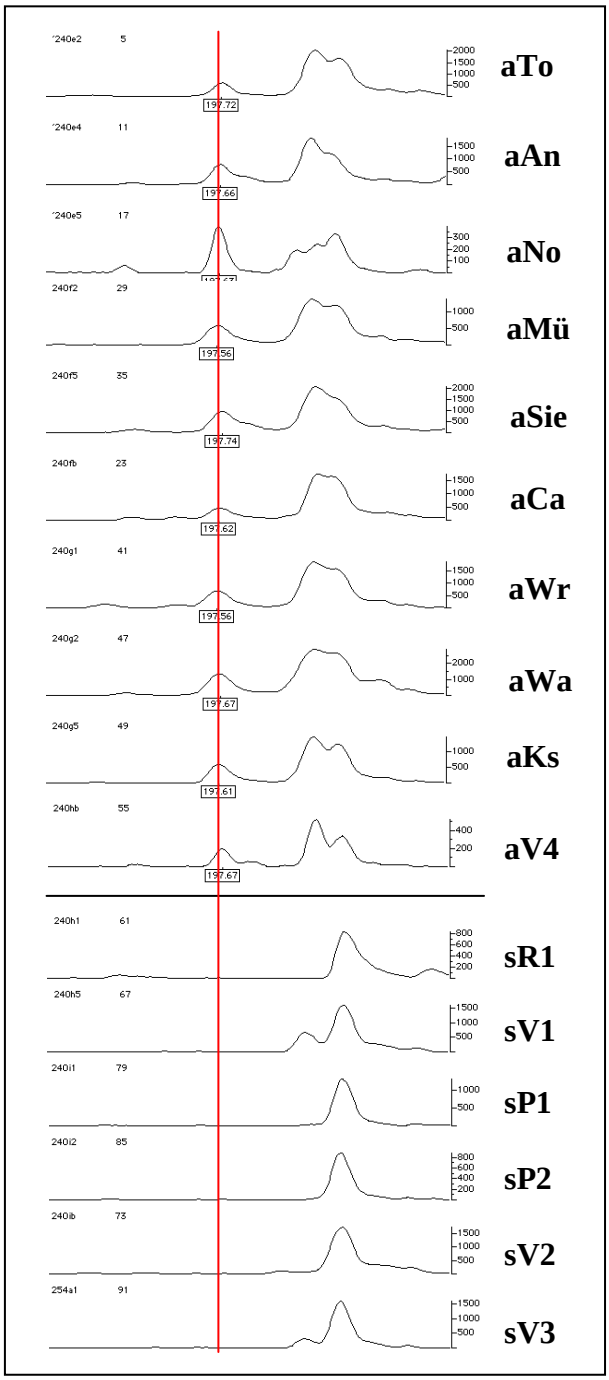
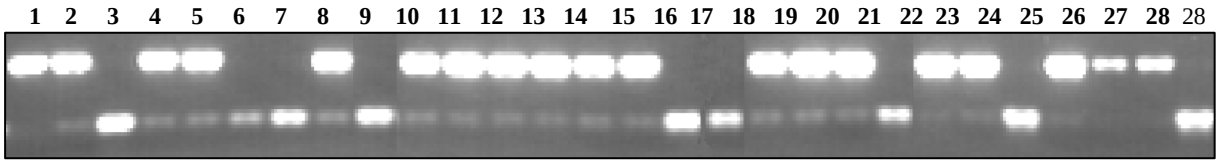


Figure 10: AFLP marker detected based on 10 apomictic and 6 sexual accessions originated from Europe.

aTo, aAn, aNo, aMü, aSie, aCa, aWr, aWa, aKs, aV4 – apomictic accesssions,
sR1, sV1, sP1, sP2, sV2, sV3 – sexuals



1	1(sR1xaNo)1/1	99%	**11** 1(sV1xaNo)1/2	94%	**21** 1(sV1xaNo)1/23	10%
2	1(sR1xaNo)1/2	99%	**12** 1(sV1xaNo)1/3	100%	**22** 1(sV1xaNo)1/25	100%
3	1(sR1xaNo)1/3	7%	**13** 1(sV1xaNo)1/4	100%	**23** 1(sV1xaNo)1/27	100%
4	1(sR1xaNo)1/4	100%	**14** 1(sV1xaNo)1/5	90%	**24** 1(sV1xaNo)1/28	0%
5	1(sR1xaNo)1/5	98%	**15** 1(sV1xaNo)1/6	100%	**25** 1(sV1xaNo)2a/1	100%
6	1(sR1xaNo)2/1	0%	**16** 1(sV1xaNo)1/7	0%	**26** 1(sV1xaNo)2b/2	100%
7	1(sR1xaNo)2/2	0%	**17** 1(sV1xaNo)1/10	0%	**27** 1(sV1xaNo)2b/3	100%
8	1(sR1xaNo)2/4	95%	**18** 1(sV1xaNo)1/17	100%	**28** 1(sV1xaNo)2b/4	0%
9	1(sR1xaNo)3/5	0%	**19** 1(sV1xaNo)1/19	100%		
10	1(sV1xaNo)1/1	99%	**20** 1(sV1xaNo)1/20	95%		

Figure 11: AFLP derived CAPS marker in 28 arbitrarily chosen sexual (S) and aposporous (A) *Hypericum* lines

The presence of the *Eco*RI undigestable 223 bp fragment and the absence of this fragment strictly co-segregate with the aposporous and sexual modes of reproduction, respectively. The individual plant identifier and the degree of apospory, measured by FSSC (%) are given for each plant individually.

In order to test the association between the CAPS marker and the mode of reproduction, the analysis was extended to all available F_1 populations of directed crosses, for a total of 157 tetraploid, triploid and diploid plants. The presence of the *Eco*RI undigestable 223 bp fragment "A" (for Apospory) and the absence of this fragment "S" (for Sexual) strictly co-segregates with the aposporous and sexual modes of reproduction, respectively. As summarized in Table 9 and given in detail in Table15 (attachment) all aposporous and sexual individuals, as characterized by FCSS analysis, were correctly classified on the basis of the CAPS marker analysis. No F_1 individual was identified with a potential genetic recombination between the apospory-linked marker and the apospory trait.

Next I asked if this apospory specific marker locus is linked with parthenogenesis too. I observed that at least 16 plants of the triploid families and 9 plants of the tetraploids which were confirmed to exhibit the marker state "A" for apospory but scored zero percent parthenogenesis in FCSS (labelled with # in Table 15 (attachment)). Therefore, the CAPS marker strictly co-segregates with apospory but not with parthenogenesis, the second important component of apomixis. Therefore, the described CAPS marker can be considered

to be an apospory-specific marker and should allow the dissection of both components of apomixis in *H. perforatum*.

Cross	No. of plants	FCSS		Marker state		
		SEX	APO	SEX	APO	n. d.
Diploid						
2(sR1 x sP2)	11	11	0	11	0	0
2(sP1 x sV1)	13	13	0	13	0	0
Triploid						
1/2(sR1 x aNo)	16	5	11	5	11	0
1(sV1 x aNo)	25	7	18	7	18	0
1(sV3 x aNo)	4	0	4	0	4	0
1(sP2 x aTo)	14	4	10	4	10	0
Tetraploid						
4($sF_1$1 x aAn)	8	6	2	6	2	0
4($sF_1$1 x aSi)	4	3	1	3	1	0
4($sF_1$1 x aNo)	10	2	8	2	6	2
4($sF_1$2 x aAn)	10	7	3	7	3	0
4($sF_1$2 x aNo)	24	8	16	8	14	2
4($sR1C_2$ x aSi)	16	12	4	12	4	0
Total	**155**	**78**	**75**	**78**	**73**	**4**

Table 9: Mode of reproduction detected based on FCSS and CAPS marker analysis in the segregating population of controlled crosses

3.4. Cloning and molecular characterization of apospory-linked genomic region

3.4.1. Cloning of the apospory-linked genomic region

While the identification of the apospory marker aided understanding of the genetic behaviour of the corresponding locus, I wanted to extend this work to fully understand the molecular nature of the locus. First, the marker-containing sequence was used as query for BLASTX analysis. This revealed a high similarity with genes coding for the ARIADNE-subclass of RING-finger proteins (Jackson *et al*., 2000; Mladek *et al*., 2003). For a detailed molecular characterization of this genomic locus, a BAC library was constructed from a diploid sexual plant. The average insert size was determined in 72 randomly picked BAC clones. The BAC DNA was digested with *Not*I and the insert size was estimated using Pulsed Field Gel Electrophoresis. The insert size ranged from 60 to 270 kb, the average size of the inserts was 121 kb. The fraction of BACs containing an insert of more than 100 kb was 83% (Figure 12). With the size of the haploid *Hypericum* genome being 635.7 Mbp, the BAC library covers the genome aproximately 6 times. The percentage of BAC clones without insert was 1.4%.

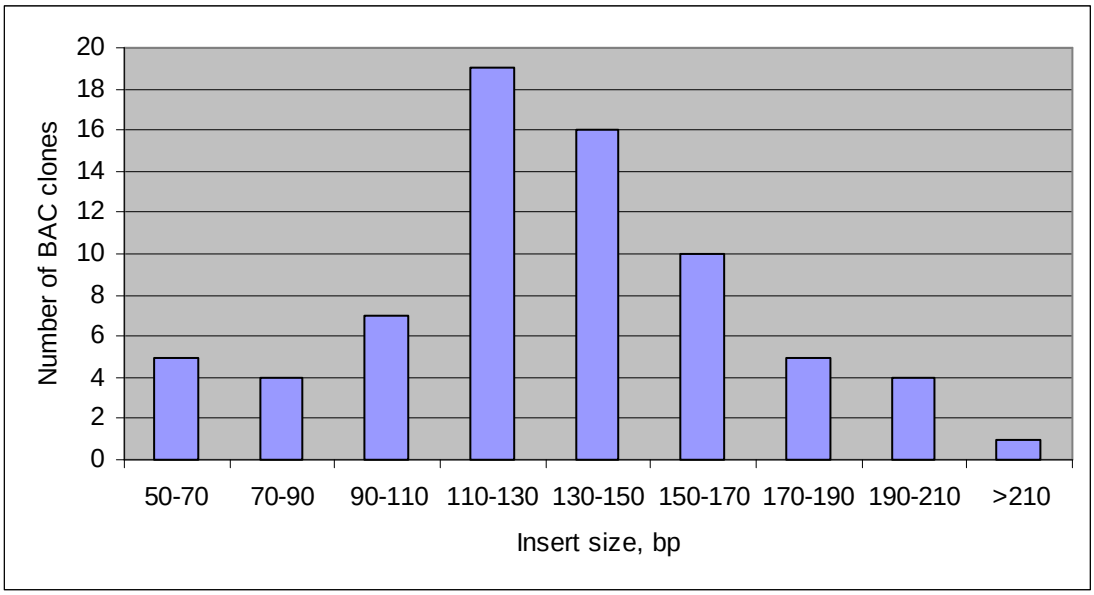


Figure 12: Distribution of insert size in the *Hypericum* BAC library

The number of clones required to anchor the target gene was estimated according to the following equation (Clarke and Carbon, 1976):
$P = 1-(1-L/G)^N$

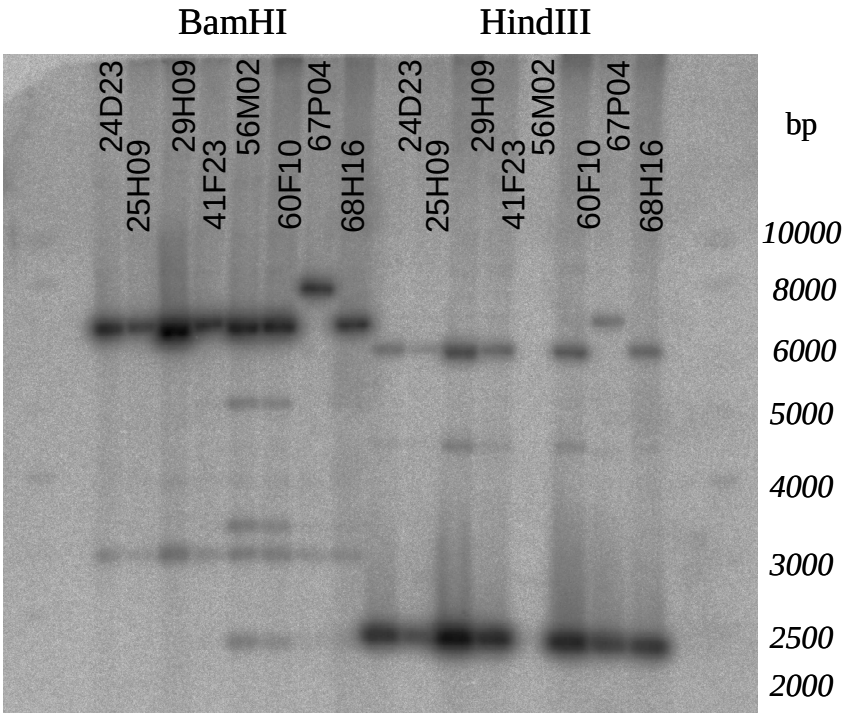
with
P- probability to obtain the target gene
N- number of clones in the library
L- length of an average clone insert in bp
G- haploid genome size in bp

For the given library with 27648 BAC clones (N), an average insert size $121x10^3$ bp (L) and for *Hypericum* with the genome size $635.7x10^6$ bp (G), the probability to obtain any single copy gene of interest is higher then 99.4%.

The stability of BAC clones is often taken as given. To test if the BAC clones from the Hypericum BAC library are stable on prolonged growth, repeated inoculation and 24h-growth during 7 days was performed for 10 randomly picked clones. The DNA-fingerprint pattern determined on growth days 1, 3 and 7 was found to be unaltered, suggesting that the BAC clones are relatively stable and thus faithfully reflect the genomic structure.

To identify BAC clones containing the marker region, the BAC library was screened with the CAPS marker amplicon as a probe. Amongst the 24 detected BAC clones only the eight BACs with the highest signal in hybridization were confirmed to yield a PCR amplicon with CAPS marker primers. To characterize the positive BAC clones the BAC DNA was digested with a single restriction enzyme and hybridized with the CAPS marker amplicon. Similar restriction pattern proved that they belong to the same group and thus the same genomic region (Figure 13 A). Alternatively, the positive BACs were characterized by fingerprinting, which showed the relative position of the BACs to each other (Figure 13 B). The fingerprinting analysis demonstrated that all of them belong to a single contig, suggesting that they represent a single genome region. Based on the fingerprinting pattern and taking into account that all of the BACs contain the CAPS marker the approximate position of the CAPS marker within the genomic region represented by the BACs can be deduced. It is situated in the region flanked by the 5' end of BAC 24D23 and 3' end of BAC 41F23 (Figure 8B). The BACs that contain the CAPS marker in the middle are 25H09, 29H09. Additionally, the length of the positive BAC clones was estimated with PFGE (Table 10). I chose to sequence BAC 25H09 with an insert size of 140 kb, since its insert is much longer than that of BAC 29H09 (97kb).

A



B

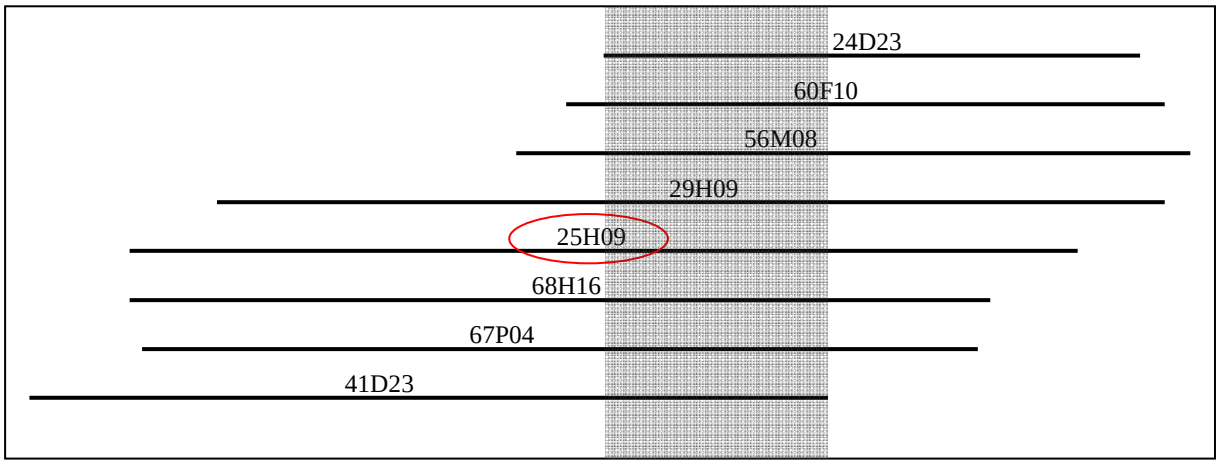


Figure 8: Characterization of the positive BAC clones

A. Hybridization of the positive BAC clones after single enzyme restriction with the probe containing CAPS marker. B. Fingerprinting analysis of the positive BACs. All eight BACs positive for CAPS marker belong to the same genomic region. The region flanked by the 5' end of BAC 24D23 and 3' end of 41D23 contain the CAPS marker and is marked in grey. The BAC 25H09 marked with the red circle was selected for sequencing.

BAC address	Length, kb
24D23	97
2509	140
29H09	97
41F23	130
56M02	100
60F10	80
67P04	155
68H16	160

Table 10: Length and hybridization signal intensity of the BAC clones positive for the marker region

3.4.2. Sequencing of BAC 25H09

For the sequencing of the BAC 25H09 the BAC DNA was sheared into 1-2kb fragments and subcloned. A PCR screening of the resulted sublibrary revealed that 91.2% of the clones contained an insert. Clones with inserts larger then 700bp were selected for sequencing. In total 1253 reads were obtained with an average read length of 700bp after quality trimming and vector removal. The complete BAC sequence is 141,941 bp long. On average the assembled sequence is covered 6.1 times with a minimum coverage of 3 fold for 3.25% of its length (Figure 14). The sequence coverage is uniformly distributed through the BAC suggesting the correctness of assembly and absence of duplication within the BAC.

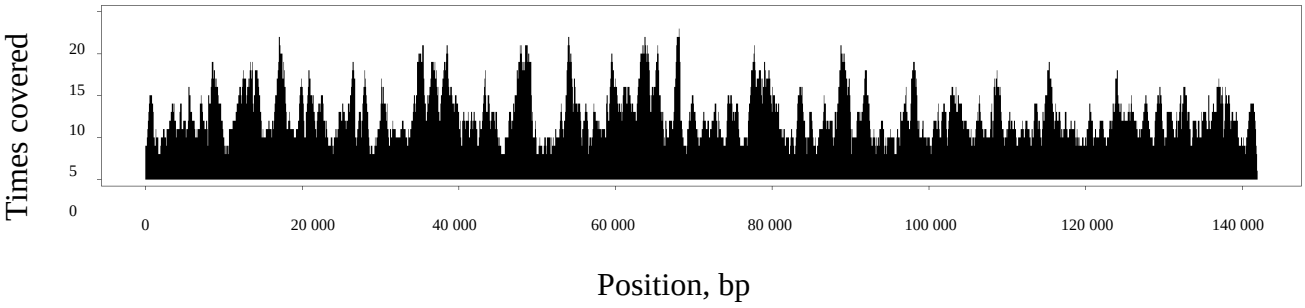


Figure 14: Sequence coverage of the BAC 25H09

To further evaluate the correctness of the assembly, the BAC was digested with eight restriction endonucleases (*Eco*RV, *Acc*65I, *Apa*I, *Bam*HI, *Sac*II, *Spe*I, *Xho*I, *Sma*I). The generated restriction patterns were identical to the patterns predicted from the assembled

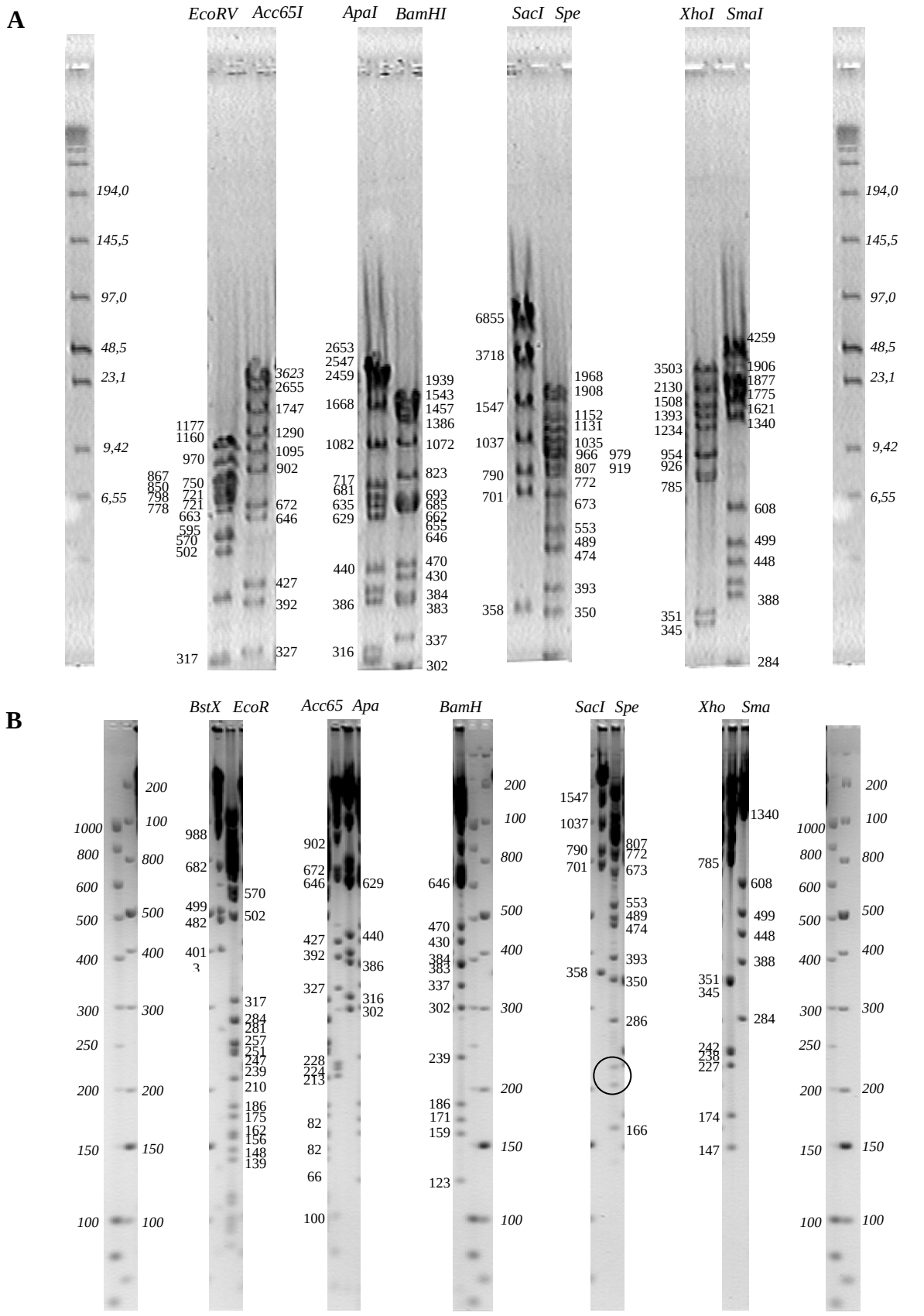


Figure 10: Restriction patterns of the assembled BAC sequence

A. PFGE gel for high weight fragments, B. Agarose gel for low weight fragments. Two fragments which could not be explained by in silico digestion are marked with circle.

sequence (Figure 15). Out of 179 predicted fragments all 179 could be confirmed experimentally. Two extra fragments, which could not be explained by *in silico* digestion were obtained after restriction with *Spe*I. They show lower intensity in comparison to other bands and are probably due to partial digestion. Moreover, the sequences of the BAC ends of the all positive BACs were included in the assembly and confirmed the fingerprinting data.

Annotation revealed the presence of 24 predicted genes on the BAC clone. Assuming 26.000 genes for an average angiosperm genome and a *Hypericum* genome size of about 635.7 Mbp (Matzk *et al.*, 2001), the average number of genes expected on the 25H09 BAC would be only 5.8. This indicates that the genomic region represented in BAC 25H09 is gene rich with one gene per 6.17 kb. The average GC content of the total BAC is rather low with 38.8%. GC content is higher in the coding regions (41.2 %). The main sequence characteristics and functional classification of the annotated genes are given in Table 11. The predicted genes on the BAC include for instance the CAPS-marker containing gene for an E3 ligase of the ARIADNE family surrounded by genes which encode proteins with similarity to DGCR14, NPH3, PAT1, WRKY transcription factor, Auxin-response protein, EMB2761, threonine-tRNA ligase, histidine kinases as well as four retrotransposons (Figure 16). The list of predicted genes on the BAC 25H09 together with the most similar genes based on BLAST hits are given in Table 12.

25H09 BAC sequence characteristics	
Number of genes	23
Gene density	1 gene/6.17kbp
Average exon length	1328bb
Average intron length	230bp
Average gene length	1883bp
Exons per genes	3.75 (range 1 to 10)
Microsatellites number	86
Retrotransposon number	4

Functional classification of the annotated genes on BAC 25H09	
Transcription factor	3
Binding protein	4
Enzyme	5
Retrotransposon	4
Unknown	6
Pseudogene	3

Table 11: The main sequence characteristics of the BAC 25H09 and functional classification of the annotated genes

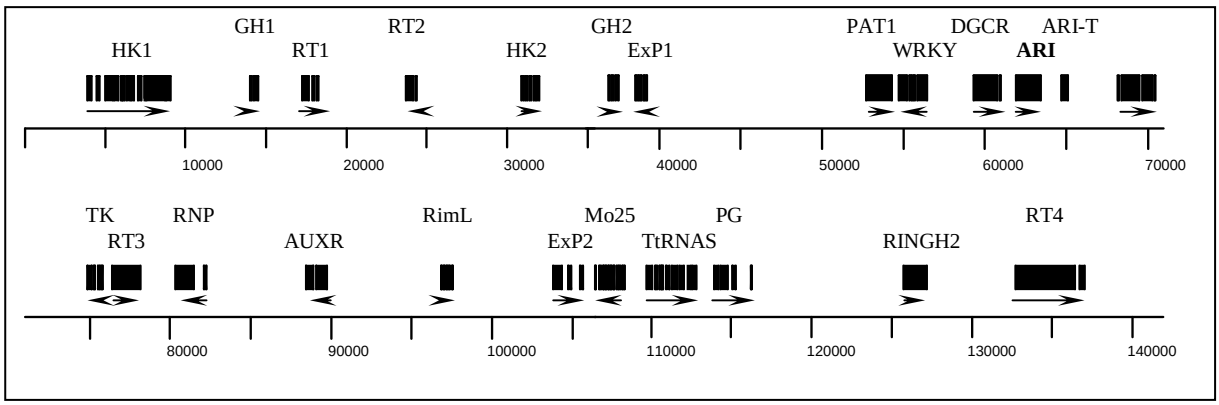


Figure 16: Annotation scheme of the ARIADNE-marker containing genomic region of a diploid sexual plant

The line represents the insert sequence of 25H09 BAC (141 941 bp) with encoding exons shown as black rectangles. The annotated genes include histidine kinases (HK1,2), glycosyl hydrolases (GH1,2), retrotransposons (RT1,2,3,4), an expressed protein (ExP1), a member of the GRAS gene family (PAT1), a WRKY transcription factor (WRKY), a hypothetical protein (DGCR), an ubiquitin-mediated E3-ligase (ARI), a truncated version of the E3-ligase (ARI-T), a plastidal phototropic-responsive protein (NPH3), a thymidine kinase (TK), a heterogeneous nuclear ribonucleoprotein (RNP), an auxin-response protein (AUXR), a GCN5-related N-acetyltransferase domain containing protein (RimL), a membrane protein (ExP2), an unknown protein of the Mo-25 family (Mo25), a threonine-tRNA ligase also known as embryo defective gene 2761 (TtRNAS), a polygalaturonase (PG) and a RING finger protein of the H2 class (RINGH2).

Symbol	Description	Best hit	E value	Best hit in *Arabidopsis*
HK1,2	Histidine kinase AHK2	*Arabidopsis thaliana*	0.0	At5G35750
GH1, 2	Glycosyl hydrolase family protein 17	EF144611 *Populus trichocarpa*	1.00E-40	At5G35740
RT1	Ty-1 copia retrotransposon	AC151801 *Solanum demissum*	3.00E-23	-
RT2	Retrotransposon	AP006368 *Lotus japonicus*	2.00E-47	-
ExP1	Expressed protein	XM_002320476 *Populus*	3.00E-60	At2G04900
PAT1	PAT1 like protein	CR955006 *Medicago truncatula*	0.0	At5G48150
WRKY	WRKY related	*Arabidopsis thaliana*	1.00E-70	At2G04880
DGCR	hypothetical protein	AM478160 *Vitis vinifera*	1.00E-173	At3G07790
HpARI	ARIADNE protein	*Arabidopsis thaliana*	0.0	At2G31510
ARI-T	ARIADNE protein (truncated)	XM_002310407 Populus trichocarpa	2.00E-69	At2G31510
NPH3	Similar to NPH3 family protein	*Arabidopsis thaliana*	0.0	At5G48130
TK	Thymidine kinase	AM492793 *Zea mays*	3.00E-187	At3G07800
RT3	Retrotransposon	AP008240 *Lotus japonicus*	2.00E-24	-
RNP	Putative hnRNP	BT013750 Solanum l*ycopersicon*	2.00E-142	At2G31860
AUXR	Auxin-response related	AM478160 *Vitis vinifera*	6.00E-167	At2G04850
RimL	RimL/ Acetyltransferase domain protein	*Arabidopsis thaliana*	8.00E-55	At2G04845
ExP2	Membrane protein, acethyltransferase	AJ309300 *Solanum tuberosum*	3.00E-111	At3G07810
Mo25	Mo-25 like protein	BT013201 Solanum l*ycopersicon*	6.00E-116	At5G47540
TtRNAS	EMBRYO DEFECTIVE 2761) Thre-tRNA ligase	*Arabidopsis thaliana*	0.0	At2G04842
PG	Polygalacturonase	*Arabidopsis thaliana*	3.00E-123	At3G14060
RINGH2	RING-H2 finger protein	AM484170 *Vitis vinifera*	3.00E-76	At5G05810
RT4	Retroelement polyprotein	AC135396 *Orysa sativa*	7.00E-29	-

Table 12: Used symbols, putative functional description, best BLAST hit, E-value and most similar Arabidopsis genes of the annotated genes on BAC HP25H09

Simple Sequence Repeats (SSRs) were identified in the sequence of BAC 25H09 with the aim to develop microsatellite-based molecular markers which would distinguish the traits of apomictic and sexual reproduction. Moreover, microsatellites may also be used to study gene dosage, to identify duplications or deletions in a particular genomic region). Microsatellites with a minimum length of 8 (mononucleotide and dinucleotide), 9 (trinucleotide), 12 (tetranucleotide) or 15 (pentanucleotide) nucleotides were considered. In total 86 microsatellites were detected, the majority of them were located in intergenic regions and introns. Only 4 microsatellites were located within the genes. Nearly half of the microsatellites are of the mononucleotide T/A type (Table 13). This type is approximately two times more frequent in *Hypericum* BAC than in the *Arabidopsis* genome (Casacuberta *et al.,* 2000). The dinucleotide microsatellites in contrary are two times rarer in *Hypericum* compared with *Arabidopsis*. More then half of the SSRs (51) are less then 12 nucleotides long, however 26 SSRs are 12 - 19 nucleotides long and 9 SSRs are longer then 20 nucleotides. An octamer repeat of 32 nucleotides and 17mer, 24mer, 37mer and 110mer tandem repeats consisting of two repetitions each are also present within the BAC. The six longest microsatellites were selected for further analysis. Although all of them were polymorphic, none of them were diagnostic for the mode of reproduction.

MS type	Nr of MS	kb/MS		MS type	Nr of MS	kb/MS
Mononucleotide (T/A)	47	3.0		trinucleotide	13	10.9
dinucleotide	25	5.7		TCC/GGA	3	47.3
GA/TC	11	12.9		TAA/TTA	2	71.0
TA/AT	11	12.9		AAT/ATT	1	141.9
CA/GT	3	47.3		GTT/AAC	2	71.0
tetranucleotide	0			CAT/ATG	1	141.9
pentanucleotide	1	141.9		AGA/TCT	2	71.0
				TTG/CAA	1	141.9
				TGA/TCA	1	141.9

Table 13: Microsatellites in *Hypericum* BAC 25H09

3.4.3. Collinearity between Poplar LGXIV and *Hypericum* 25H09 BAC

The most closely related plant species to *Hypericum perforatum* with a completely sequenced genome is *Populus trichonocarpa.* Both belong to the order *Malphigiales*, so it is of great interest to determine, if synteny between *Hypericum* BAC and poplar genome can be detected since this may aid further mapping studies in *Hypericum*. Poplar has a genome size comparable to *Hypericum* (~520Mbp) distributed over 19 chromosomes. Using BLASTN for sequence comparison, significant similarities were found for two regions on the BAC, the largest of them being over 23 kb long (Figure 17). These regions correspond to two regions on LGXIV. The first synthenic region is situated in opposite orientation compared to the BAC sequence. Moreover, the order of the genes on LGXIV is the same as within the *Hypericum* BAC: *HK2*, *ExPr1*, *PAT1*, *WRKY*, *DGCR*, *ARI*, and *NPH3*. Interestingly, the marker containing the *HpARI* gene is located within the second synthenic region. The second region has the same orientation as the Hypericum BAC and contains *Auxresp*, *RimL*, *ExPr2*, *Mo25* and *TtRNAS*. Thus, the most of the genes on the BAC are also presented on the LGXIV of poplar. No significant collinearity to *Arabidopsis thaliana* has been found, which can be explained by a bigger evolutionary distance.

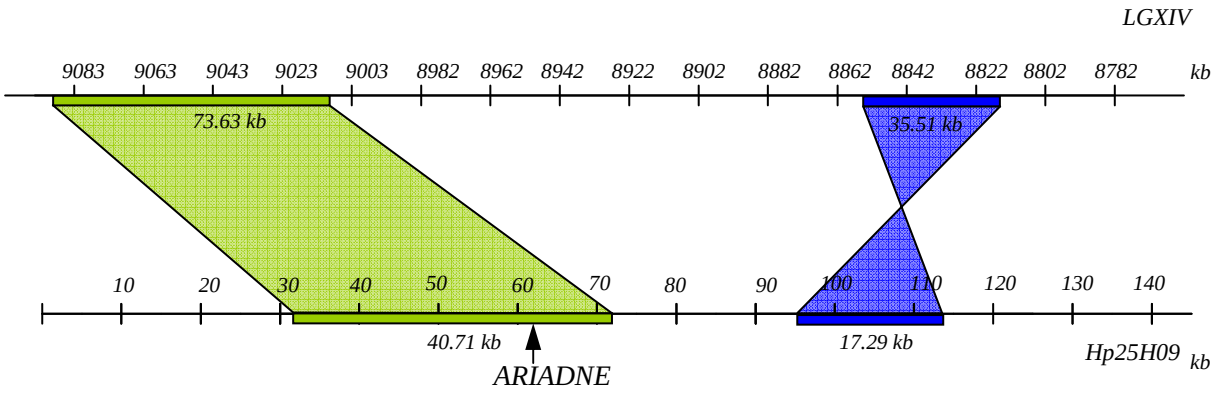


Figure 17: Microcollinearity between Poplar LG XIV and *Hypericum* 25H09 BAC

3.5. *Hypericum HpARI* gene in aposporous and sexual plants

3.5.1. Aposporous *Hypericum* plants contain an apospory-linked truncated allele of an Hp*ARI* gene

In a first attempt to analyze the molecular function of the *Hypericum* Apospory locus (*HAPPY* locus), I investigated the structure of sexual and aposporous alleles of the *HpARI* gene. The sexual allele in the BAC 25H09, encodes a protein of 519 amino acids without any introns. The translated sequence shows strong similarity with the *Arabidopsis ARIADNE 7 (ARI7)* protein. To determine the main differences between for aposporous and sequal alleles, the genomic region containing *HpARI* was amplified from 4 obligatory aposporous and 4 sexual tetraploid plants. The length of amplicon was 1,596 bp . Two additional primers in the middle of the *HpARI* were used for sequencing, thus only doublestranded sequences covering the whole gene were considered for the analysis. For each cloned amplicon at least 16 plasmid inserts were sequenced, corresponding to a probability of 96.7% for detecting all present alleles of a simplex locus structure in a tetraploid plant (Simko, 2004).

Eleven haplotypes of the *HpARI* gene were identified, which fall into two distinct groups (Figure 18 B). One group consists of haplotypes 300 and 311, with the latter being exclusively found in aposporous plants. Among the ten haplotypes occuring in sexual plants, 73 single nucleotide polymorphisms were identified resulting in 40 amino acid changes. In addition a 6 bp deletion was observed in haplotype 200, which removes two amino acid codons, but leaves the reading frame of *HpARI* intact (Figure 18 A). 39 polymorphisms proved to be specific for the apospory-specific haplotype 311, some leading to changes in the amino acid sequence. Furthermore, two 2 bp deletions and a 1bp insertion were detected. The 2 bp deletion removing the bases 146 and 147 of the coding sequences of *HpARI* causes a shift in the reading frame. As a result, the encoded protein of the apospory-specific allele consists of only 48 N-terminal amino acids with similarity to the sexual alleles and 19 aberrant amino acids added to it, until the novel reading frame ends in a stop codon (Figure 19). The other 2 bp deletion and 1 bp insertion occur after the stop codon. In summary, detailed sequencing analysis identified several haplotypes belonging to the sexual and apomictic alleles; the apospory-specific haplotype seems to code for a truncated mutant protein of *HpARI.*

A

	34	38	39	41	47	52	65	68	69	80	92	94	104	105	120	131	146	147	153	155	176	206	240	267	268	307	312	318	320	341	364	367	378	389	477	484	488	521
100	A	A	G	T	T	G	T	C	A	G	T	C	G	G	G	G	A	G	A	T	T	C	G	G	T	G	G	A	C	G	C	C	G	T	T	C	T	C
133	A	A	G	C	C	G	G	T	A	G	T	C	G	G	G	G	A	G	A	T	C	C	A	G	T	G	G	A	T	G	C	C	A	T	C	C	C	G
132	A	A	G	T	T	G	T	C	A	G	T	C	G	G	G	G	A	G	A	T	T	C	G	G	T	G	G	A	C	G	C	C	G	T	T	C	T	C
131	A	A	G	T	T	G	T	C	A	G	T	C	G	G	G	G	A	G	A	T	T	C	G	G	T	G	G	A	C	G	C	C	G	T	T	C	T	C
130	A	A	G	C	T	A	T	T	A	G	T	C	A	G	G	G	A	G	A	T	T	C	G	G	T	G	G	A	C	G	C	C	G	T	T	C	T	C
120	A	A	G	C	T	G	T	C	A	G	T	C	G	G	G	G	A	G	A	T	T	C	G	G	T	G	G	A	C	G	T	C	G	T	T	T	T	C
111	A	A	G	T	T	G	T	C	A	G	T	C	G	G	G	G	A	G	A	T	T	C	G	G	T	G	G	A	C	G	C	C	G	T	T	T	T	C
110	A	A	G	C	T	G	T	C	A	G	T	C	G	G	G	G	A	G	A	T	T	T	G	G	T	G	G	A	C	G	C	C	G	T	T	C	T	C
200	A	A	G	T	T	G	T	C	A	G	T	C	G	G	G	G	A	G	A	T	T	C	G	G	T	G	G	A	C	G	C	C	G	T	T	C	T	C
300	A	A	G	C	T	G	T	T	G	G	T	C	A	G	T	G	A	G	A	T	T	C	G	G	T	A	A	G	C	G	C	T	G	T	C	C	C	G
311	G	C	C	T	C	G	G	T	A	A	C	G	G	A	G	A	-	-	T	C	C	C	G	-	-	A	A	G	C	A	C	C	G	C	C	C	T	C

	524	540	572	599	617	626	630	649	653	673	674	701	741	750	762	787	799	800	801	843	949	960	977	979	984	990	996	1020	1059	1070	1079	1087	1093	1095	1100	1113	1115	1116
100	A	A	T	C	G	C	G	A	G	G	C	C	A	T	-	G	A	T	G	G	T	T	G	G	A	G	C	G	A	G	C	C	G	A	C	G	T	G
133	T	T	T	C	G	C	G	A	G	G	C	C	A	A	-	G	A	T	G	G	T	T	G	G	A	C	T	A	A	G	C	C	G	T	C	G	T	G
132	A	T	T	C	G	C	G	A	G	G	C	C	A	T	-	G	A	T	G	G	T	T	G	G	A	C	T	G	A	G	C	C	G	T	C	G	T	G
131	A	T	T	C	G	C	G	A	G	G	C	C	A	T	-	G	A	T	G	G	T	T	G	G	A	C	T	G	A	G	C	C	G	T	C	G	T	G
130	A	A	T	C	G	C	G	A	G	G	C	C	A	T	-	G	A	T	G	G	T	T	G	G	A	C	T	G	A	G	C	C	G	T	C	G	T	G
120	A	T	T	C	G	C	G	A	G	G	C	C	A	T	-	G	A	T	G	G	T	T	G	G	A	G	C	G	A	G	C	C	G	T	C	G	T	G
111	A	T	T	C	G	C	G	A	G	G	C	C	T	T	-	G	A	T	G	A	T	T	G	G	A	C	T	G	A	G	C	C	G	T	C	G	T	G
110	A	T	T	C	G	C	G	G	G	A	C	C	A	T	-	G	A	T	G	G	T	T	G	G	A	C	T	G	A	G	C	C	A	T	C	G	T	G
200	A	T	T	C	G	C	G	A	G	G	C	C	A	T	-	G	A	T	G	G	T	T	G	G	A	C	T	G	A	G	C	C	G	T	C	G	-	-
300	T	T	A	C	G	T	T	A	T	A	A	C	A	T	-	G	C	T	T	G	G	C	A	G	G	G	C	G	A	A	C	G	G	T	C	G	T	G
311	A	T	T	A	A	C	G	A	G	G	C	G	A	T	T	T	C	C	T	G	T	C	A	C	G	G	C	G	G	A	T	G	G	T	G	T	T	G

	1117	1118	1119	1120	1126	1150	1161	1162	1163	1180	1222	1236	1255	1260	1267	1268	1269	1279	1284	1285	1286	1287	1297	1300	1316	1323	1384	1385	1387	1394	1398	1455	1498	1510	1513	1530
100	C	A	A	A	A	C	T	T	T	A	T	C	T	C	G	A	T	G	C	T	T	T	C	G	G	C	A	A	G	A	C	G	A	A	C	A
133	C	A	A	A	A	C	T	G	T	A	C	C	C	C	G	A	T	G	C	T	T	T	C	G	G	C	A	A	G	G	C	A	A	A	C	A
132	C	A	A	A	A	C	T	G	T	G	C	C	C	C	G	A	T	G	C	T	T	T	C	G	G	C	A	A	G	A	C	G	A	A	C	A
131	A	A	A	A	A	C	T	G	T	A	T	C	C	C	G	A	T	G	C	T	T	T	C	G	G	C	A	A	G	A	C	G	A	A	C	A
130	C	A	A	A	A	C	T	G	T	G	C	C	C	C	G	A	T	G	C	T	T	T	C	G	G	C	A	A	G	A	C	G	A	A	C	A
120	C	A	A	A	A	C	T	T	T	A	T	C	T	C	G	A	T	G	C	T	T	T	C	G	G	C	A	A	G	A	C	G	A	A	C	A
111	C	A	A	A	A	C	T	G	T	A	T	C	C	C	G	A	T	G	C	T	T	T	C	G	G	C	A	A	G	A	C	G	A	A	C	A
110	C	A	A	A	A	C	T	G	T	A	T	C	C	C	G	A	T	G	C	T	T	T	C	G	G	C	A	A	G	A	C	G	A	A	C	A
200	-	-	-	-	A	C	T	G	T	A	T	C	C	C	G	A	T	G	C	T	T	T	C	G	G	C	A	A	G	A	C	G	A	A	C	A
300	C	A	A	A	T	C	G	G	C	A	C	T	C	G	C	G	C	C	C	T	T	T	C	A	G	C	A	C	G	A	G	G	A	A	C	A
311	C	A	A	A	A	T	G	G	C	A	C	T	C	G	C	G	C	C	A	A	A	G	A	A	T	G	T	C	A	A	G	G	T	G	A	G

B

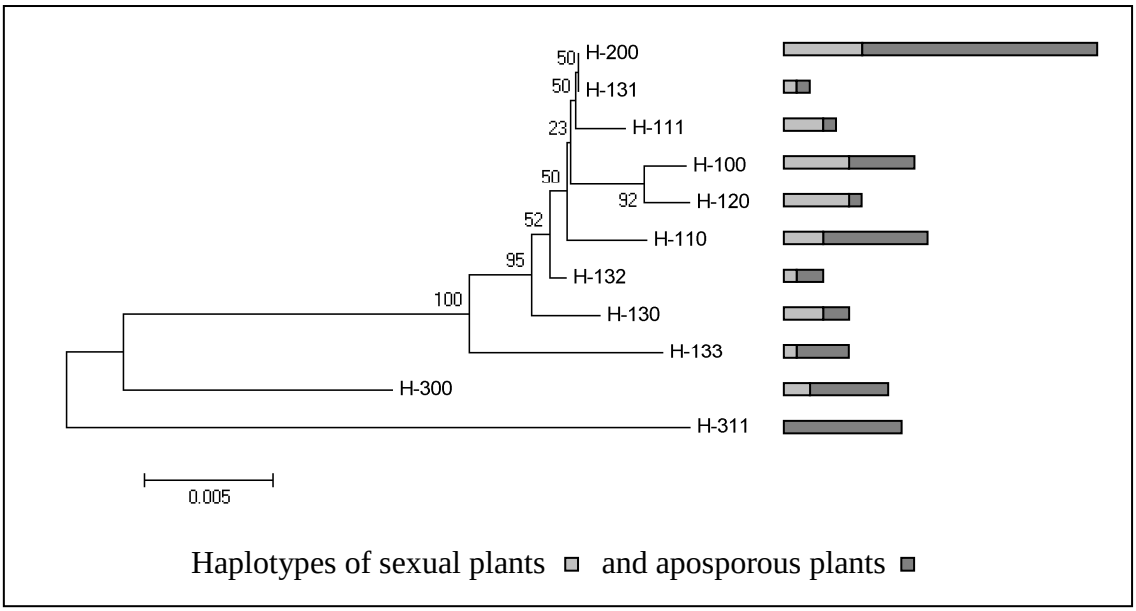


Figure 18: *HpARI* haplotypes

A. Polymorphisms within the detected haplotypes. The apospory-specific polymorphisms are marked in light grey. The sexual haplotype 300 and the apospory-specific haplotype 311 are marked in dark grey and black respectively. B. The similarity tree of *ARIADNE* haplotypes and their occurrence in sexual and aposporous plants. The bootstrap values are placed at the nodes.

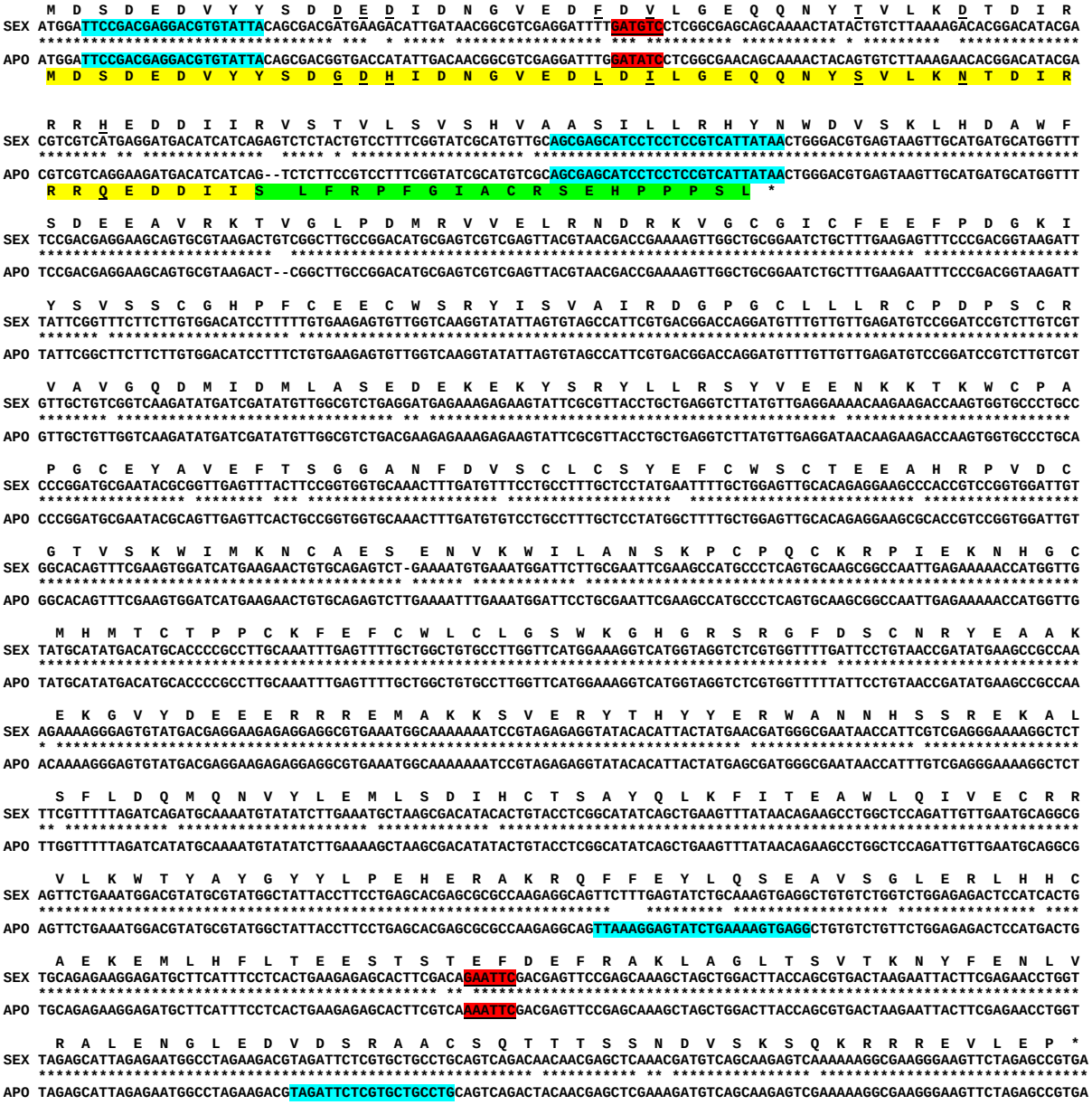


Figure 19: Alignment of the sexual haplotype 300 and aposporous haplotype 311

The amino acid sequence of the truncated apospory-specific HpARI protein is highlighted in yellow, the 19 aberrant aminoacids in green. The position of the primers of the CAPS marker, as well as of the primers used in the expression analysis is highlighted in blue. The diagnostic *Eco*RI and *Eco*RV sites used in these analyses are underlined and highlighted in red.

3.5.2. Quantification of aposporous alleles in the population of directed crosses reveals simplex genomic constitution

Using the *HpARI* locus as a land mark for the *HAPPY* locus, I examined the quantitative nature of the *HpARI* alleles. In sexual and aposporous tetraploid plants the allele

constitution was analyzed by Southern blot hybridization (Figure 20). Three aposporous plants showed both the aposporous (A) allele represented by the *Eco*RI uncut 978 bp fragment and the sexual (S) allele represented by the two *Eco*RI fragments of 583 bp and 395 bp. The ratio of intensities of both short fragments to the uncut large fragment is close to 1:3 indicating the presence of three S alleles and one A allele in apomicts (simplex locus) and four S alleles in tetraploid sexuals (quadruplex locus). The Southern analysis with enzymes (*Hind*III and *Bam*HI) not cutting in the *HpARI* gene region resulted in only one hybridized fragment, indicating the presence of a single copy of the *HpARI* gene in the *Hypericum* genome. In addition to the genomic hybridization, the *HpARI* gene was PCR-amplified and subcloned from four aposporous plants. Using the informative *Eco*RI site, the bacterial plasmid clones were classified into two distinct groups, giving 15 aposporous and 49 sexual alleles, a ratio which again is close to 1:3. Taken together, the data demonstrate the simplex allelic composition of the apospory-related locus in tetraploid plants.

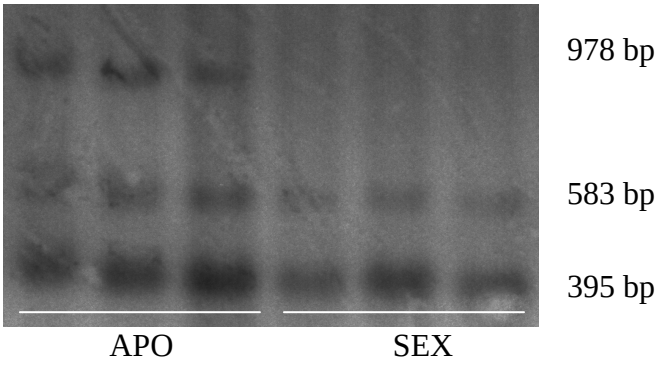
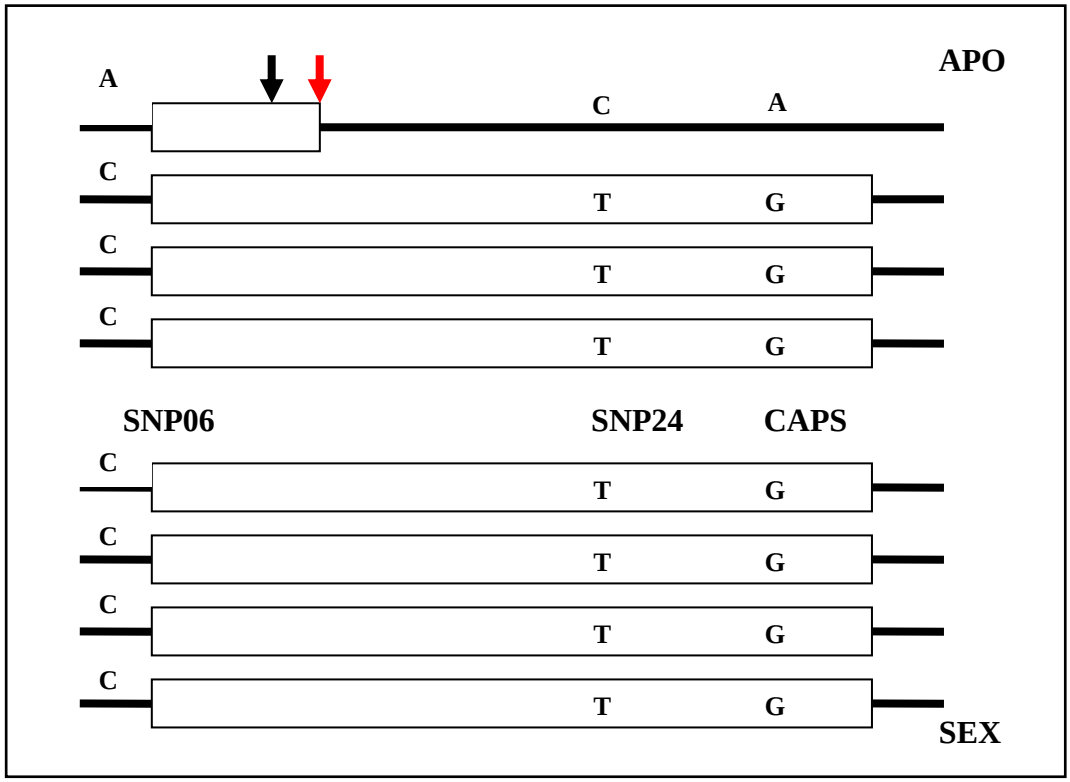


Figure 20: Detection of the HpARI allele ratio via genomic southern

APO: identifier 4($F_1$2xNo)1a/1, 4($F_1$1xNo)1/8, 4($F_1$2xAn)1/10
SEX: identifier 4($F_1$2xNo)1a/10, 4($F_1$2xNo)1a/4, 4($F_1$2xNo)1/5

To prove the presence of the apospory-specific allele and to estimate its frequency in the population of tetraploid plants, pyrosequencing markers were established (Figure 21 A). Allele quantification revealed that aposporous plants contain 20-25% of the apospory-specific allele and 70-75% of the sexual ones (Figure 21 B, in detail Table 16, supplement). In contrast, sexual plants contain from 96% to 100% of the sexual alleles. A minor proportion of apospory-specific alleles detectable in sexual plants is within the range of errors of the method. No recombination events between the markers and the apospory trait were detected

A



B

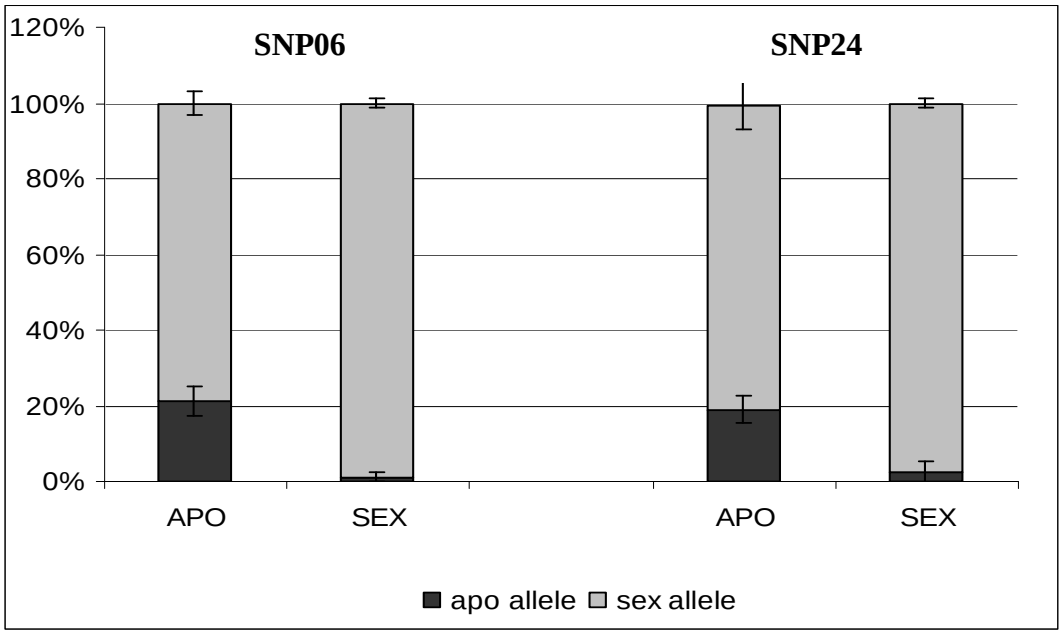


Figure 21: Determination of sexual and aposporous allele frequencies in the segregating population of directed crosses by pyrosequencing

A. The positions of the pyrosequencing primers SNP06 and SNP24. The CAPS marker is shown for orientation. The boxes represent the coding sequence. The black and the red arrow indicate the shift in the reading frame and the stop codon in the aposporous haplotype respectively. B. Quantification of HpARI alleles in the tetraploid population of directed crosses via pyrosequencing.

in the plants tested. The allele distribution suggests a 1:3 ratio of aposporous and sexual alleles. This allele distribution within the tetraploid aposporous progeny of the various crosses is expected, if all the aposporous parents have a simplex allelic constitution.

3.5.3. Quantification of aposporous alleles in plants from wild populations confirms simplex genomic constitution

To prove that the *HAPPY* locus is not restricted to the generated population of the segregating crosses, the presence and the frequency of the apospory-specific allele were determined in the collection of Hypericum accessions collected from over the world. Using the established pyrosequencing markers (Figure 21 A), 21 sexual diploid and 62 aposporous tetraploid plants were analyzed from various populations (see Table 17, materials and methods).

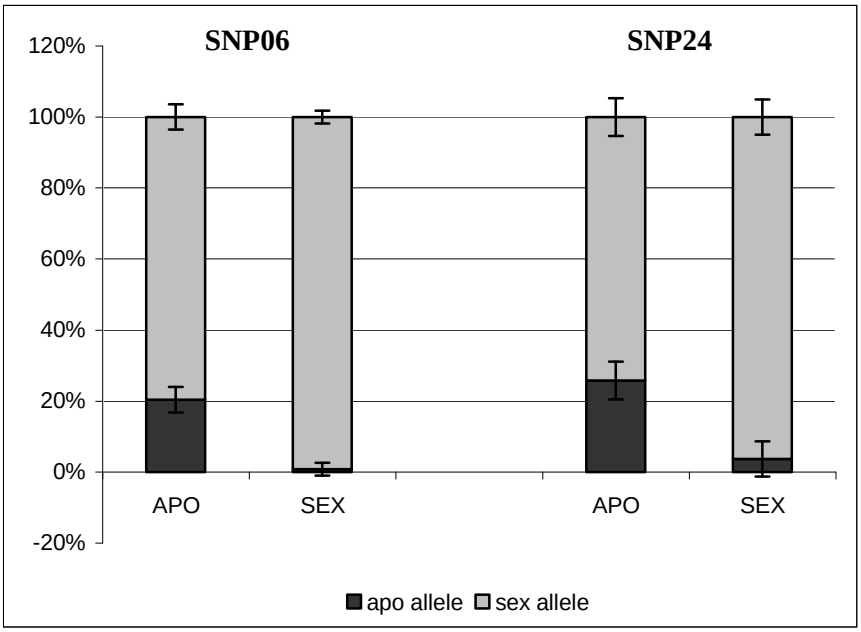


Figure 22: Quantification of *HpARI* alleles in tetraploid apomictic plants and diploid sexual plants. Individuals from several natural populations from all over the world were analyzed by pyrosequencing

Similar to the results obtained with tetraploid population of the directed crosses, the quantification of *HpARI* alleles revealed that aposporous plants contain 20-25% of the apospory-specific allele and 70-75% of the sexual alleles (Figure 22, in detail Supplement,

Table 17 A). The diploid sexual plants contain 96% to 100% of the sexual alleles (Figure 22, in detail Supplement, Table 17 B). Again a minor proportion of the apospory-specific allele detected in sexual plants (up to 3%) is within the range of errors of the analytical method. Both markers co-segregated with the apospory in each of the plants tested. The 1:3 ratio of aposporous vs. sexual alleles is retained in all of the tetraploid plants from wild population. The fact, that the apospory-specific markers are co-segregating with the apospory-trait in all of the tested populations suggests that the *HAPPY* locus is present exclusively in the apomictic plants independent of their origin.

3.5.4. Genomic walking extends the apospory-specific genomic locus

In order to extend the *HAPPY* locus beyond *HpARI*, I cloned and sequenced parts of exons of 13 other genes on the BAC 25H09 in the same manner as for *HpARI* (see 5.1). The analysed region was in total 6,252 bp covered by 1,702 doublestranded sequences. Suprisingly, no alleles specific for the aposporous plants could be detected for any of the genes. As an example the haplotype trees of the genes flanking *HpARI* are given in Figure 23. Similar to *HpARI* two haplotype groups and ten different haplotypes were detected for *DGCR* and for *NPH3*. However, none of the haplotypes was strictly apospory-specific.

Although this result suggested the unique nature of the aposporous *HpARI* allele, I was aware of the technical problems in recovering novel aposporous alleles. In case the sequence of an aposporous-specific allele diverged substancially from the corresponding allele(s) in the sexual plants, it may not be possible to amplify both alleles with the same primer pair. Therefore, I also undertook a genome walking approach to extend the *HAPPY* locus beyond *HpARI*. Upstream of the *HpARI* the apospory-specific locus was extended to sequences reaching into the *DGCR14* gene and downstream sequence was extended beyond the *ARI-T* gene (Figure 24). As shown in Figures 24 and 25, the sequence similarity between sexual and apospory-specific alleles is restricted to the close proximity of the *HpARI* gene. The similarity abruptly disappears in the 3'-terminal part of the *DGCR* gene 72 bp upstream of the predicted stop codon (Figure 24). Downstream of the aposporous *HpARI* sequence, similarity between sexual and apospory-specific alleles disappears within the *ARI-T* gene 102 bp upstream of the predicted stop codon (Figure 25). Therefore, I believe that *HpARI* gene is an important

member of the *HAPPY* locus and it is likely that most of the sexual loci in the vicinity of *HpARI* might not have a corresponding apospory-specific alleles.

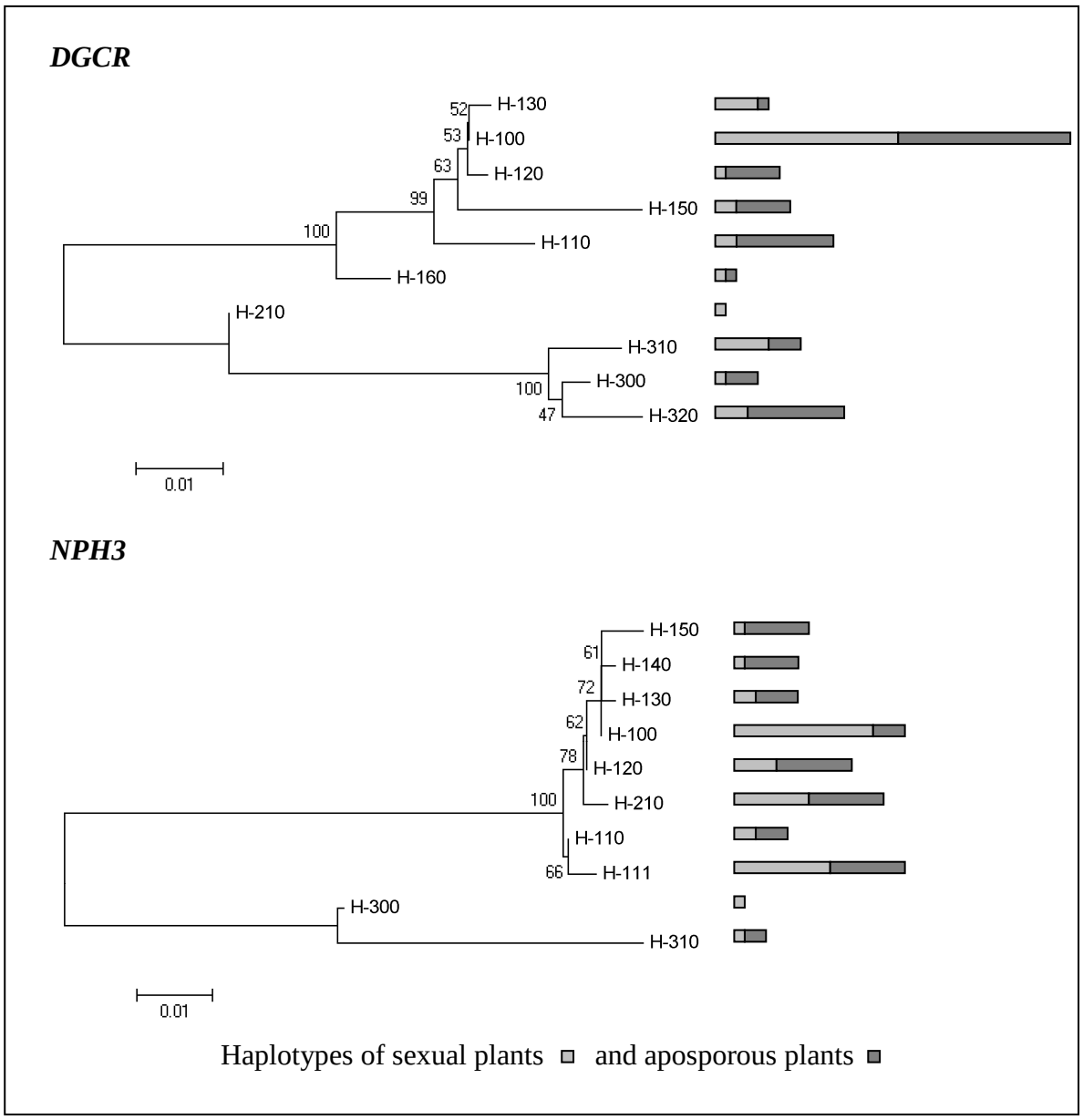


Figure 23: Similarity tree for the sexual haplotypes of the two ARI-flanking genes *DGCR* and *NPH3*

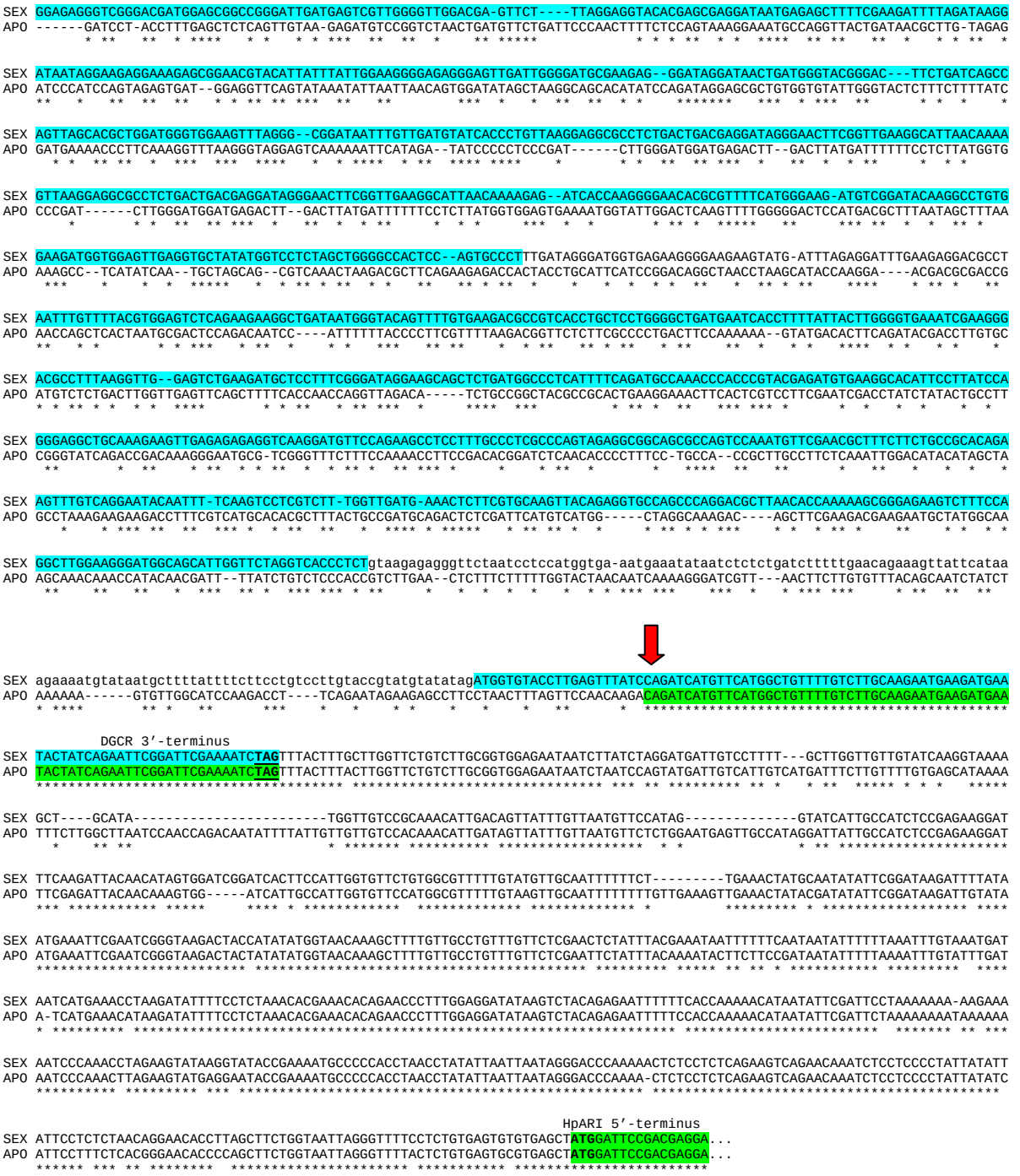


Figure 24: Nucleotide alignment of the sexual and aposporous *HpARI* sequence and that of its 5' gene neighbour *DGCR*

The predicted intron in *DGCR* is marked by lower case lettering. The arrow indicates where homology between the sexual and aposporous sequence is first disrupted. No homology was detected beyond this point.

Figure 25: Nucleotide alignment of the sexual and aposporous *HpARI* sequence and that of its 3' gene neighbour *ARI-T*

The arrow indicates where homology between the sexual and aposporous sequence is first disrupted.

3.6. Towards the functional analysis of *HpARI* gene

3.6.1. Both alleles of *HpARI* are expressed

To analyze whether both sexual and aposporous alleles are expressed, an RT-PCR assay was performed. The *Hypericum* pistils at the megasporogenesis and megagametogenesis stages (as staged by Galla *et al.*, in press) were used. Primers were designed to uniformly amplify both alleles within the coding region. The distinction between the alleles was possible due to a *Eco*RV site, uniquely present in the apospory-specific sequence. Due to the intron-less nature of the *HpARI* gene in sexual and apomictic plants, the absence of genomic DNA in the generated cDNA samples was demonstrated by the analysis of an adjacent gene encoding *HISTIDINE KINASE 2* (*HK2*) with primers located in intron-separated exons (Figure 26). As shown in Figure 15, temporal expression of both sexual and apomictic *HpARI* alleles was indistinguishable in floral buds at developmental stages during which megasporogenesis and megagametogenesis do occur, as well as in leaves. This leaves the possibility that the expressed *HpARI* allele of the *HAPPY* locus and the derived truncated ARIADNE protein form might have a function in the control of apospory in *H. perforatum*.

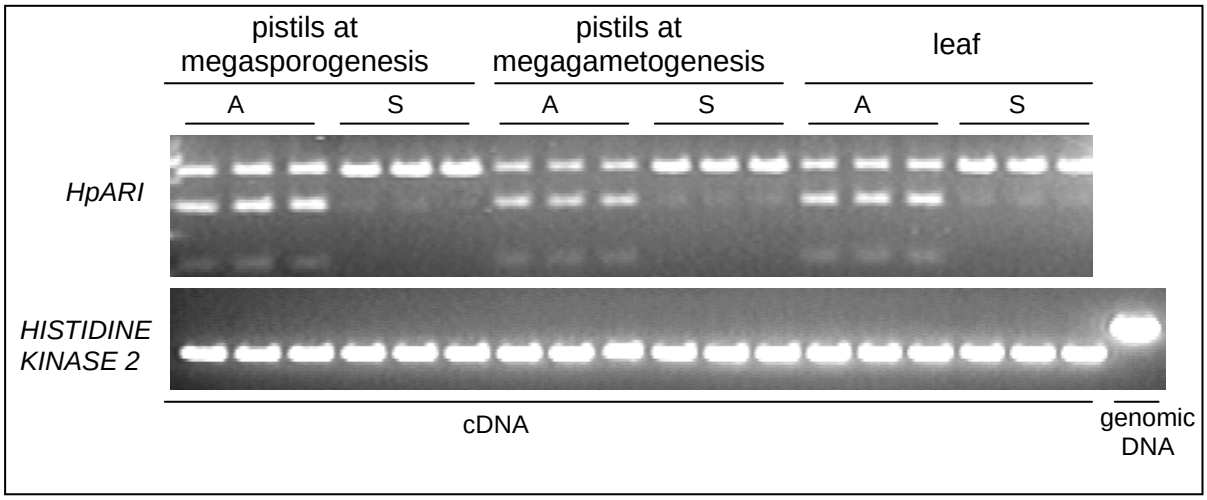


Figure 26: Expression analysis of *HpARI* gene

Both, sexual and apospory-specific alleles are expressed in pistils and leaves of aposporous plants. The exclude genomic DNA contamination an intron containing gene *HSK2* was amplified in the same RNA samples. Abbreviations: A- aposporous, S - sexual

3.6.2. *HpARI* gene might be involved in the flowering time control.

To define the functional role of the expressed apospory-specific *HpARI* two approaches were employed: (a) generation of plant lines expressing the *Hypericum HpARI* gene and (b) characterization of T-DNA insertion mutants of the most related *Arabidopsis* genes. Following the first approach constructs carrying the sexual and apospory-specific *Hypericum HpARI* gene under control of the ubiquitin promoter of maize were generated (Figure 27 A). The ubiquitin promoter and not the 35S promoter was chosen since the latter might have some effect in gametophyte development (A. Johnston, personal communication). Two constructs carrying the apospory-specific *HpARI* allele and the corresponding sexual *HpARI* allele were generated (Figure 27 A). In case of the apospory-specific *HpARI* allele I included all the sequence downstream of the truncated coding sequence, so that the length of the both inserts was the same. Two additional constructs were generated, carrying the native *Hypericum* sexual and apospory-specific promoters instead of the ubiquitin promoter. *Arabidopsis thaliana* plants were transformed with all four constructs. Since apomixis is often associated with polyploidy, not only the most common diploid *Columbia* accession was used, but also a tetraploidized *Columbia*. Additionally the homologous host plant *Hypericum perforatum* was used for transformation. A transformation method of *Hypericum* using *Agrobacterium* is under development (Kumlehn *et al.,* unpublished). Moreover, a transformation complementation of a line generated by deletional mutagenesis in apomictic *Hieracium* (Catanach *et al.,* 2006) is also planned. The diploid transformed *Arabidopsis* lines of did not show any phenotype. The FCSS analysis also did not reveal any differences in the ploidy ratios embryo and endosperm. Tetraploid *Arabidopsis* lines containing the aposporous *HpARI* under the control of ubiquitin promoter show a late-flowering phenotype. After 28 days under long day conditions none of the transformed plants showed any signs of flowering (Figure 27 B). After 53 days of growth they could be divided into two groups according to the length of the delay. At this time 16 transformants out of 25 did not flower (Figure 27 C). The remaining 9 plants flowered and developed viable seeds. All transformants carrying the sexual allele flowered after 22 days. These findings suggest that the *HpARI* might be involved in controlling of transition from vegetative to reproductive development in a gene dosage dependent manner. Since some of the genes involved in flowering time control were shown to have pleiotropic effects (Bäurle *et al.,* 2007; Kinoshita *et al.,* 2004), one might speculate that the *HpARI* has more than one function and is also involved in sexual and/or apomictic

reproduction. However, a more detailed analysis of the transgenic lines will be required in the future.

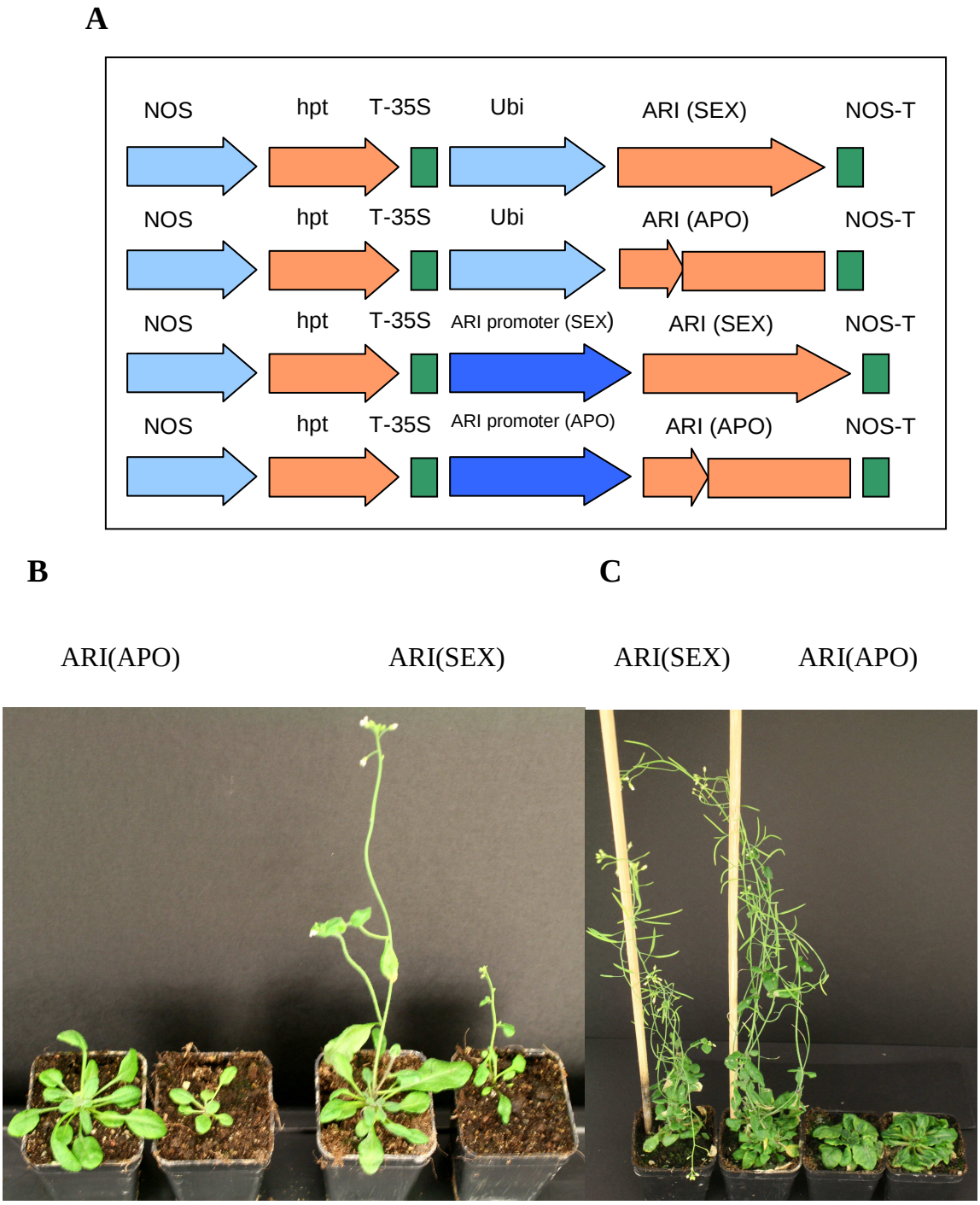


Figure 27: Expression cassette used for plant transformation and phenotype of *Arabidopsis* transformants

A. Schematic diagram of the cassette used for ectopic expression of *HpARI* gene. B. Phenotype of *Arabidopsis* (Col, 4n) transformants containing sexual and apomictic ARIADNE alleles under the control of the maize ubiquitin promoter - 28 days old. C. Phenotype of Arabidopsis (Col, 4n) transformants 53 days old.

Following the second approach, T-DNA insertion mutants of *Arabidopsis thaliana* were analyzed (Table 14). I was most interested in *ARI7*, since the sexual allele showed the highest sequence similarity with this Arabidopsis gene (AT2G31510). Additionally, five *ARIADNE* genes, which show the highest BLAST similarity to *ARI7* were analyzed.

Locus Name	**Gene ID**	**Sequence Similarities**	**Availability of ESTs and cDNAs (tissue type)**	**T-DNA mutant allele**	**Genomic location of insertion**	**Mutant plant phenotype**
ARI7	AT2G31510	ARI5, ARI8, ARI9, ARI10, ARI11	yes (callus)	SALK_027620 SALK_073106 SALK_116408 SALK_025210 SALK_082541 SALK_132004 SALK_029122 SALK_069374	intron intron intron intron intron exon 3' UTR 3' UTR	none observed
ARI5	AT1G05890	ARI7 ARI8 ARI9 ARI10 ARI11	yes (vegetative tissues)	Not isolated yet	NA	NA
ARI8	AT1G65430	ARI7 ARI11 ARI8 ARI9 ARI10	yes (vegetative tissues)	Not isolated yet	NA	NA
ARI9	At2G31770	ARI10, ARI11, ARI7, ARI5, ARI8	yes (vegetative tissues)	SALK_027269 SALK_049257	exon intron	none observed
ARI10	AT2G31760	ARI9, ARI11, ARI7, ARI5, ARI8	yes (vegetative tissues)	SALK_014075 SALK_068103	exon promoter	None observed Unlinked phenotype with a gametophytic mutation carrying 51% infertile ovules (N=493); ovules arresting at one and two-nucleated stages
ARI11	AT2G31780	ARI9, ARI10, ARI7, ARI5, ARI8	yes (vegetative tissues)	Not isolated yet	NA	NA

Table 14: Isolation of T-DNA mutant alleles that disrupt ARIADNE (ARI) RING finger genes in Arabidopsis

The next similar gene is clearly distinct suggesting that *ARI7* is most likely the homologous gene. However, no discernable phenotype was detected either for *ARI7* mutant lines or for the lines of other genes analyzed. This indicates that the *AtARI* genes may retain redundant functions. Currently, double mutants are being generated to search for possible distortions in gametophyte development.

4. Discussion

4.1. Aposporous initials arise from the external cellular layer of the nucellus

The performed analysis of female sporogenesis and gametogenesis in sexually and aposporously reproducing individuals defined the major morphological traits of all structures playing a role in embryo sac formation. Present data might be considered as necessary upgrade of our knowledge of *Hypericum* ovule development traced back over sixty years ago by the pioneer publications of Noack (1939, 1941) by tissue stain clearing and high resolution imaging. The sexual ovule is anatropous, bitegmic (i.e. with 2 integuments) and tenuinucellate, and encloses at maturity a monosporic, *Polygonum* type embryo sac. The aposporous ovule retains all major morphological characteristics reported herein, but fails to develop a reduced embryo sac. In contrast to the sexual pathway, the aposporous embryo sac develops through a series of free nuclear divisions from a single sporophytic cell, which belongs to the epidermal cell layer of the nucellus. My observations imply that the major developmental features characteristic of aposporous ovule formation include: i) miss-expression of the meiotic program, ii) failure or delay in degeneration of the nucellus epidermal layer, iii) differentiation of Aposporous Initials (AI) from the nucellar tissue, and iv) development into an alternative coenocytic embryo sac. Furthermore, low frequencies of sexual developing ovules are retained within the aposporous ovary of facultative apomicts. Interestingly, according to Noack (1939, 1941) AIs were reported to differentiate from the hypodermal cell layer of the nucellus. In contrast, the data I obtained suggests that most AIs differentiate from the external nucellar cellular layer. However, the general importance of the nucellus in AIs as the site of differentiation is confirmed.

The importance of AI positioning with respect to the site of differentiation of the meiotic products has been reported for several aposporous species. In *Hieracium piloselloides* and *H. aurantiacum*, the AIs were reported to differentiate in close connection to the megaspores (Koltunow *et al*., 1998). Moreover, aposporous gametophyte development is influenced by AI mis-positioning resulting from the *loa1* mutation (Okada *et al*., 2007). Similarly, the correct positioning of AIs in *Hypericum* ovules seems to be necessary for further development, as no AIs or embryo sacs were ever detected distantly from the site of megasporogenesis within the nucellus. Since both MMCs and AIs originate from the same

site of differentiation, the nucellus, it is likely that the two processes depend on the expression of specific factors restricted to this area of the ovule. Furthermore, the fact that megasporogenesis is limited to the hypodermal inner area and never occurs in the epidermal cell layer, where aposporous initials are confined, suggests an additional regional level of regulation within the nucellus background.

Callose deposition is frequently lacking or follows unconventional patterns within the aposporous ovules of *Hypericum,* while within the sexual ovules it is consistent with the pattern described for other plants. It is generally recognized that callose deposition during megasporogenesis progression might play a role in the regulation of the fate of meiotically derived megaspores by physically limiting symplastic communication with adjacent cells (Rodkiewicz, 1970; Haig, 1986). The pattern of callose deposition in ovules of sexual species is not conserved in apomeiotic pathways, as has been shown in *Elymus rectisetus* (Carman *et al*., 1991), *Hieracium* spp. (Tucker *et al*., 2001) *Poa pratensis* (Naumova *et al*., 1993), *Tripsacum dactyloides* (Leblanc *et al*., 1993) and *Medicago falcata* (Barcaccia *et al*., 1996), where meiosis is bypassed or disturbed. In this species, the lack of callose accumulation has been reported for both Aposporous Initials (Peel *et al*., 1997; Tucker *et al*., 2001) and MMCs undergoing diplosporic development (Carman *et al*., 1991). It is possible to speculate, that the altered pattern of callose deposition within aposporous ovules of *Hypericum* indicates upon a deregulation in normal cell-to-cell communication. It is likely that developing megaspores either directly or indirectly regulate epidermal layer proliferation whereby apomeiosis leads to abnormal signaling between closely neighboring cells.

Aposporous and sexual embryo sacs share certain features. For example, the development of the female gametophyte by means of nuclear division without cytoplasm repartition (i.e. unequal allocation of cytoplasm to daughter cells after nuclear division) is similar in both embryo sacs. Embryo sac formation, both in sexual and aposporous ovules always proceeds through successive nuclear divisions and concludes with a final cellularization event, in contrast to sporophytic cells, in which karyokinesis is immediately followed by cytokinesis. The differences between sexual and aposporous embryo sac development are dramatic. In addition to changes in cell shape and developmental timing, the main differences observed include the number of nuclei and the organization of embryo sac following cellularization. Nuclear divisions within the embryo sac are frequently

asynchronous, leading to unconventional numbers of nuclei that are miss-positioned within the same coenocytic structure. The elevated variation in aposporous gametophyte development (relative to sexual ones; see Figure 2 and 3, and Table 4, results) could reflect stochastic processes, for example perturbations of signaling pathways in the ovule (Okada *et al*., 2007), or mutation accumulation in independently-derived apomictic clones (Koltunow *et al*., 1993). Similarly, stochastic gametophyte development has been hypothesized in aposporous *Hieracium aurantiacum* (Koltunow *et al*., 1998), where one to multiple Aposporous Initials were demonstrated to form and grow without any precise developmental patterns. The degeneration of synergids prior to fertilization can be interpreted as a physical barrier against fertilization. It is planned, that the identity of the cells within the aposporous embryo sac will be demonstrated with marker lines, which need to be established in *Hypericum*.

Recent reports support the hypothesis that deviation of the fate of cells enclosed in the aposporous embryo sac might be related to miss-positioning of nuclei within the embryo sac, prior to cellularization. Thus, the characterization of Arabidopsis gametophytic mutations such as *eostre* (Pagnussat *et al*., 2007) and *lachesis* (Gross-Hardt *et al*., 2007) might help elucidate the gametophyte development in *Hypericum*, for example whether apospory involves mis-regulation of the above mentioned genes. Moreover, specification of cell fate within the *Arabidopsis* embryo sac appears to rely on a position-based mechanism, as the switch from synergid to egg cell is accompanied by mis-positioning of the nucleus during early developmental stages (Pagnussat *et al*., 2007). Similarly, the auxiliary and gametic cell fates within the embryo sac of *Arabidopsis* were recently reported to be affected in the *lachesis* mutant (Gross-Hardt *et al*., 2007), in which functional supernumerary egg cells differentiate from accessory cells, pointing to a mechanism that prevents accessory cells from adopting gametic cell fate. The present investigation supports the hypothesis that the nucellus is a highly organized tissue, composed of two morphologically and functionally defined cellular layers. During the sexual development the cellular division processes within the nucellus are likely to be highly controlled as, with the exception of female sporogenesis and gametogenesis, no divisions were reported in the tissue as meiosis begins. During the aposporous development the control over the nucellus is probably loosened and mitotic divisions leading to AI formation are possible. The AI could be an epidermal cell that re-enters the cell cycle to undergo mitotic division, a process that could be influenced by the

specific positioning of the cell within the ovule. This could be achieved, for instance, by mis-regulation of genetic factors involved in epidermal cell identity (Sieber *et al*., 2004), although nucellus identity would have to be maintained in the other parts of the nucellus. In summary, factors required for both aposporous and sexual gametophyte development may be modulated by some regulative elements localized in hypodermal and epidermal nucellar cell domains.

4.2. Apospory and parthenogenesis do co-segregate but can recombine in *Hypericum*

Extensive research aiming to unravel the genetic control of apomixis in several apomictic and sexual model systems has not resulted in a generally accepted scenario, suggesting the existence of different pathways in distinct species of interest. This hypothesis is compatible with the independent *in planta* evolution of apomixis in general and each apomictic component in particular (Van Dijk and Vijverberg, 2005). Nonetheless, individual genetic approaches focused on each model system greatly facilitate global efforts to isolate the involved genes. Mutations of genes which predominantly control sexual pathways as well as newly evolved genes have been considered to explain the inheritance of the two apomixis components, apomeiosis and parthenogenesis. However, whether these two developmental traits are under the control of a single master gene or the influence of a few closely related genes is an unsolved puzzle for most of the apomicts. Studies aiming to understand the genetic and molecular factors underlying apomixis have been limited since the asexual mode of reproduction is often associated with polyploidy and high degrees of heterozygosity, traits which make genetic and genomic analyses very difficult.

Independent genetic studies investigating apomixis in *Hypericum* plants have indicated that apomixis in this species occurs facultatively only in a polyploid state (Martonfi *et al.,* 1996; Matzk *et al.,* 2003; Mayo und Landridge, 2003; Barcaccia *et al.,* 2006). So far, the identification of *Hypericum* genotypes that are able to produce embryos either from aposporous fertilized egg cells or by parthenogenesis from meiotically reduced egg cells suggests that two distinct genetic factors control apospory and parthenogenesis, and that apospory and parthenogenesis may be developmentally uncoupled, an idea previously hypothesized by Noack (1939). This hypothesis is further supported by the finding of

Hypericum genotypes which almost exclusively express only one component of apomixis or suppress the other (Matzk *et al.*, 2001; Barcaccia *et al.*, 2006).

In this study I compared the expressivity and penetrance of apospory and parthenogenesis in experimental populations of *Hypericum* segregating for the mode of reproduction *i.e.* apomixis *versus* sexuality. The data demonstrated that parthenogenetic capacity is preferentially expressed by aposporous egg cells. I also documented the occurrence of aposporous egg cells in non-parthenogenetic individuals as well as the parthenogenetic development of meiotic egg cells. While apospory and parthenogenesis seem to be loosely linked and preferentially inherited together, recombination is not suppressed between these two traits. The observed 1:1 ratio between sexual and aposporous or parthenogenetic plants in the triploid and tetraploid segregating populations indicates a simplex genetic constitution of one responsible genomic locus in the polyploid male parents. The analysis of segregation data reveals that the putative apospory and parthenogenetis loci are associated in a chromosome window where recombination is possible. However, it cannot be ruled out that the individual components may be controlled by multiple, tightly linked genes of chromosome blocks, as it occurs for Quantitative Trait Loci (QTLs). Although the question remains open, apomixis in *Hypericum* seems to be the result of two linked genetic determinants rather than the combination of two processes, apospory and parthenogenesis, determined by independent genetic factors.

Taken together, the present study has clearly demonstrated that for *Hypericum* aposporous apomeiosis and parthenogenesis can be functionally separated by recombination, and that they are independently controlled by dominant alleles which act as single dose genes. While this scenario is in contrast to some studies in other systems that suggested a simple inheritance of apomixes by one dominant gene (Savidan 1980; Leblanc *et al.*, 1995; Bicknell *et al.*, 2000), several other studies support our data and the underlying hypothesis regarding the genetic control of these distinct traits (Noyes and Riesenberg 2000; Matzk *et al.*, 2001; Van Baarlen *et al.*, 2002; Van Dijk and Bakx-Schotman 2004; Catanach *et al.*, 2006; Noyes *et al.*, 2006).

4.3. Aposporous plants are heterozygous and contain both sexual and aposporous alleles

The apospory-specific marker and the corresponding genomic locus *HAPPY* are present in plants with both low and high degree of apospory, which suggests that the gene(s) necessary to switch on the apospory trait are located within the *HAPPY* locus, and possibly recombination could separate an additional apospory preventer locus. This model is similar to what was proposed previously for the control of apomixis in *Poa*, wherein apomixis is thought to be controlled by five different loci including one that dominantly controls initiation of apospory and a recessive apospory preventer that influences the penetrance of this trait (Matzk, 2005). According to the segregation analysis the *HAPPY* locus in *Hypericum* is dominantly inherited and the apomictic parents contain both sexual and aposporous alleles in 3:1 ratio. The same ratio was proposed in the pioneering work of Nogler (1982) for *Ranunculus auricomus.* Nogler also suggested that the allele A conferring bipolar apospory (*Hieracium* type) behaves as a lethal factor when it is present in the gametes as A or AA, in such a way only the genotypes AAaa and Aaaa are possible. This is also true for *Hypericum*, the aposporous plants resulted from the directed cross as well as the aposporous plants from different populations from all over the world have a simplex genomic constitution, as measured by pyrosequencing. This is only possible if all aposporous parents used for the crosses also have a simplex genomic constitution. Thus, only one apospory allele is sufficient for apospory generation. The plant still carries the sexual allele and can switch back to partial sexual seed production (facultative apomixis) if the apospory preventer is present.

The sexual counterpart of the *HAPPY* locus is collinear to *Populus trihonocarpa* LGXIV. *Populus* is the closest relative of *Hypericum* with available genome sequences (Tuskan *et al.,* 2006), both belonging to *Malphigiales*. Interestingly, the apomixis-segregating loci of a number of other apomictic plants were shown to have regions collinear with well studied rice or maize. For example, *ASGR* of *Pennisetum* has been found to have regions highly microsyntenic to centromeric-proximal genomic region of *Oryza* chromosome 11 (Gualtieri *et al.,* 2006) and and *ACL* of *Paspalum* to the telomeric region of the long arm of *Oryza* chromosome 12 respectively (Pupilli *et al.,* 2004). The apospory region in *Brachiaria* is collinear with markers mapped on the short arm of maize chromosome 5 (Pessino *et al.,* 1997), and diplosporous apomixis in *Tripsacum* showed linkage with markers located on the

long arm of maize chromosome 6 (Leblanc *et al.*, 1995). I plan to use the collinearity between *Hypericum* and *Populus* to develop new markers characterizing the *Hypericum* locus of interest and determining its size. Assuming that one possible mechanism governing the apospory phenotype could be the alteration of the function of a gene that normally acts during sexual reproduction (Koltunow and Grossniklaus, 2003), i.e. through mutation or by a regulatory shift, the use of existing *Populus* sequence data spanning regions of microsynteny with the aposporous *Hypericum* can well be a source of apospory candidate genes.

4.4. The *HAPPY* locus is hemizygous

Our work has established that *HpARI* is part of the *HAPPY* locus in *Hypericum* that likely controls the aposporous reproduction. The extended sequence upstream and downstream of the marker gene *HpARI* neither has any similarity with the sexual *Hypericum* alleles nor gives any hit when blasted against all known databases. Given that aposporous alleles of 13 other genes on the sexual BAC could not be detected, I propose that the genomic organization of the aposporous alleles within the *HAPPY* locus is very distinct from that of the corresponding sexual alleles, and that the hemizygosity is retained throughout this genomic region (Figure 1). Hemizygosity was first described for *APOSPORY-SPECIFIC GENOMIC REGION* (*ASGR*) locus in *Pennisetum,* whereby a part of the locus has no allelic form in sexual genotypes, while another part has a partial similarity with the corresponding sexual locus (Ozias-Akins *et al.,* 1998). Similarly, the hemizygosity was also described in *Apomixis Controlling Locus (ACL)* of *Paspalum* (Labombarda *et al.,* 2002). Such a tremendous difference between two alleles within the same plant was described only in case of apomictic genotypes. However, an extensive sequence difference including deletions, rearrangements and missing genes has also been found by comparison of two haplotypes of two maize inbred lines (Song and Messing, 2003). This unusual finding was reported to be associated with the heterosis effect. Indeed, the heterozygous environment, which is a necessary condition for the manifestation of heterosis, is also common for apomicts, since many apomicts are allopolyploids. It can be speculated that specific allelic combinations could produce an interaction that causes deregulation of gene expression leading to heterosis and/or changes in the mode of reproduction such as apomixis.

Interestingly, our results are not compatible with the model of Carman (1997), which suggests that the mechanism for apomixis, including apospory, is based on asynchronous expression of many duplicate genes in polyploid angiosperms. From this hypothesis, the segregation of a trait with a region of hemizygosity in the *Hypericum* genome in all of the analysed apomicts (segregating population of directed crosses as well as natural accessions were tested) can not be predicted. In summary, hemizygosity of the *HAPPY* locus suggests that the aposporous locus underwent extensive sequence divergence including large deletions, inversions and substitutions.

4.5. The *HAPPY* locus includes a mutated *ARIADNE* gene *HpARI*

Given that the truncated gene product of *HpARI* is part of the hemizygous and dominant *HAPPY* locus, it is possible to speculate that this apospory-specific ARIADNE mutant protein functions in a dominant negative fashion in a simplex dosage by interacting with three functional sexual alleles. ARIADNE proteins belong to E3 ligases that are conserved from yeast to plant kingdom and are thought to control ubiquitin-dependent protein degradation (reviewed in Vierstra, 2003). The apospory specific truncated *HpARI* gene is ubiquitously expressed but the corresponding protein product is truncated in contrast to the presumably functional sexual allele. Currently it is hypothesized that the truncated apomictic HpARI might represent a “mutant ARIADNE” which might disrupt the expected Ring/U-box complexes formed by the corresponding sexual forms, thus the normal course of meiotic gametophytic development is impaired in the apomicts.

4.6. ARIADNE and flowering time control

Interestingly, the overexpression lines containing the apomictic *HpARI* allele showed a late-flowering phenotype in tetraploid *Arabidopsis* plants. It will be interesting to establish an inducible system for the *HpARI* gene and check if it plays a role during gametophyte development. Recently, the RRM-domain proteins FCA and FPA which have previously been characterized as flowering-time regulators in *Arabidopsis* were shown to affect female gametophytic development and early embryonic development. It was demonstrated, that these

proteins are required for RNA-mediated chromatin silencing of a range of loci in the genome. It was demonstrated, that *FCA* and *FPA* regulate chromatin silencing of single and low-copy genes and interact in a locus-dependent manner with the canonical small interfering RNA–directed DNA methylation pathway to regulate common targets (Bäurle *et al.,* 2007). Similarly, the *FLOWERING LOCUS A* (*FWA*), a homeobox gene with a pleiotropic function is implicated in flowering time control and plays a role in imprinting during female gametophyte and endosperm development (Kinoshita *et al.,* 2004). Moreover, genes having pleiotropic effects and acting both during flowering time gametophyte and seed development are well known, for example *MULTICOPY SUPPRESSOR OF IRA 1* (*MSI1*), *METHYLTRANSFERASE 1* (*MET1*) and *FERTILIZATION INDEPENDENT ENDOSPERM* (*FIE*) (Kankel *et al.,* 2003; Katz *et al.,* 2004; Bouveret *et al.,* 2006; Xiao *et al.,* 2006). This findings support the hypothesis that the *HpARI* gene might act pleiotropically.

The late flowering phenotype of *HpARI* transformants can be observed only in the tetraploid state. That implies the possible involvement of epigenetic mechanisms in the way of action of the *HpARI* gene. It is known that epialleles in tetraploid plants (but not in diploids) interact in *trans* and lead to heritable gene silencing persisting after segregation from the inactivating allele. This mechanism, resembling paramutation, leads to the establishment of functional epigenetic homozygosity and, thus, to conversion of new recessive alleles into traits expressed in polyploid generations (Scheid *et al.,* 2003). Probably, the phenotypic manifestation of action of the *HpARI* gene is only possible in a tetraploid background with corresponding epigenetic modifications. Since the initial experiments were done in a tetraploidized *Columbia*, an additional transformation of tetraploid *Lansberg* is planned to check if the same phenotype can be obtained in a native, long time established tetraploid.

Interestingly, the *Arabidopsis* gene *AT3G04610,* which is orthologous to an apomixis marker C 454 on the *ACL* in *Paspalum,* was recently described as an important regulator of flowering (Mockler *et al.,* 2004). Briefly, *FLK* (flowering locus K homology (*KH*) domain) encodes the K-homology domain RNA binding protein. The *flk* mutations cause a delay of flowering time of up to 40 days after germination. The *KH* domain is involved in different processes such as floral morphogenesis regulation in Arabidopsis (*HUA ENHANCER 4* (Lorkovic and Barta, 2002; Cheng *et al.,* 2003)). In *Drosophila*, mutations within a KH domain protein *FRAGILE X MENTAL RETARDATION 1 (FMR1)* lead to defects in the rapid

nuclear division cycles during embryo development (Deshpande *et al.*, 2006). Another KH gene *BICAUDAL-D* was implicated in patterning the *Drosophila* egg chamber in mid-oogenesis (Swan and Suter, 1996). In summary, it is possible to speculate that future functional studies could demonstrate if the *Hypericum HpARI* gene and the *Paspalum FLK-like* gene have a pleiotropic effect and play an additional role during sexual and/or apomictic reproduction. A more detailed analysis should also be carried out with the existing tetraploid *HpARI* lines to understand the mechanism of the flowering time delay. Currently, *Agrobacterium*-mediated tissue-culture based transformation methods in *H. perforatum* are under development (J. Kumlehn and H. Bäumlein, ongoing work) aimed at producing stable sexual *Hypericum* transformants carrying mutant (aposporous) alleles of *HpARI* in order to examine whether or not this mutant form will be required and/or sufficient to induce apospory.

Except of the recent results indicating upon the function of *HpARI* in the control of flowering time, the developmental function of Arabidopsis *ARIADNE* genes remains very obscure (Mladek *et al*., 2003). Considering the ubiquitous expression of *ARIADNE* gene products in *Arabidopsis*, it is anticipated that these factors play a pleitropic role during plant development. *ARIADNE* genes were primarily cloned and characterized in *Drosophila* and mammals, and an essential developmental role of *ARIADNE* was for instance demonstrated by identification of a fully penetrant *ariadne-2* lethal mutant in *Drosophila* (Aguilera *et al.*, 2000). Unfortunately, both in animals and plants gene redundancy within the *ARIADNE* clade has impeded understanding of the corresponding developmental function of these genes. Amongst the *Arabidopsis* 16 *ARIADNE* proteins I was mostly interested in *ARI7* (AT2G31510), the Arabidopsis homologue of the *HpARI*. In particular, *ARI7* in Arabidopsis is a tandemly duplicated gene unlike the single, structurally intact copy *HpARI* that was reported in this study. The genes which show high sequence similarity to *ARI7* are *ARI9/10/11/12*. As several other *ARIADNE* members, *ARI7* possesses a complete RING1-IBR-RING2 domain, and it has a conserved nuclear localization sequence in the highly conserved region of RING2. It has polypeptide sequences enriched in proline (P), glutamic acid (E), serine (S) and threonine (T) (PEST) at the N-terminus, a motif implicated for proteins that undergo rapid destruction (Rogers *et al.*, 1986). Although mutant alleles that disrupted the Arabidopsis *ARI7* and some of its homologues were identified (Table 14), none of them showed developmental aberrations in the mutant plants possibly due to gene

redundancy, suggesting the need for combining several mutant alleles of the *ARIADNE* family of genes. Interestingly, the expression profile of *ARI7* detected using AtGenExpress microarray data (Schmid *et al.,* 2005) is different from the other *ARIADNE* genes. *ARI7* is relatively high expressed in all sporophytic tissues in comparison to the co-expressed *ARIADNE* members ARI9/10/11/12 and low expressed in mature pollen in comparison to other tissues. The other *ARIADNE* genes, in contrary, show the highest levels of expression in the mature pollen. So ARI7 has a unique pattern of expression among other family members. Recently, Spencer *et al.,* (2007) reported that the *pARI7*:*GUS* showed low level of activity in the globular embryo, and it is necessary to extend this analysis to the rest of sporophytic and gametophytic stages. Additionally, the expression analysis of *ARI7* in the mega-sporophyte and mega-gametophyte, and both *in situ* hybridization- and promoter fusion analysis should be carried out to shed more light on the possible function of this gene.

4.7. Structure of the *HAPPY* locus

An extensive population genetic analysis followed by mapping and sequencing approaches led to identification of at least a part of the apospory controlling locus in *Hypericum*. A scheme reflecting the structure of the the *HAPPY* locus is presented in Figure 28. It was shown that the locus is simplex, dominant and hemizygous. The known sequence of the apospory controlling locus extends from *DGCR* to *ARI-T* and is 5.221 kb long.

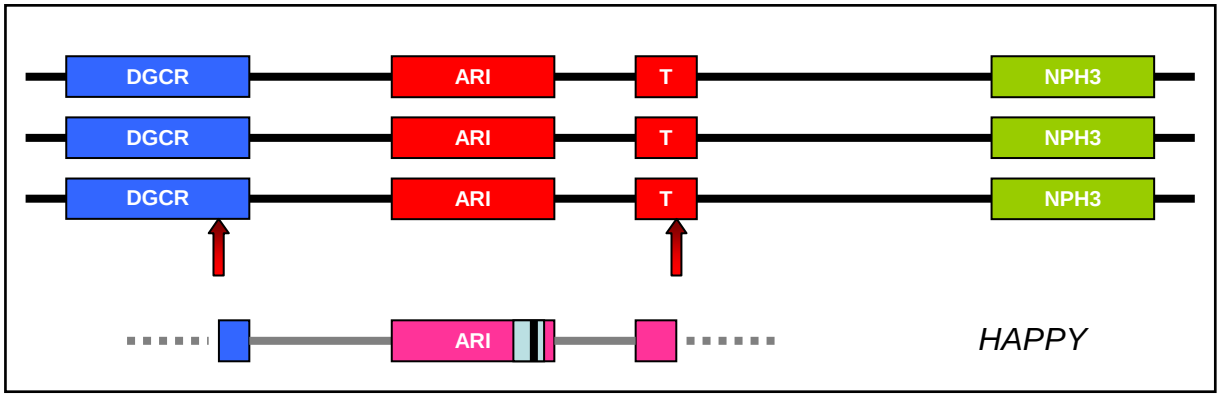


Figure 28: Schematic structure of sexual and aposporous alleles of the *HpARI* gene containing locus of *Hypericum perforatum*

ARI indicates the marker-containing *HpARI* gene, T stands for the truncated version of the ARI gene and DGCR and NPH3 represent the 5`and 3´-flanking genes, respectively. Tetraploid apomicts possess three sexual alleles and one aposporous allele. The red arrows indicate the positions where the sequence homology between sexual and aposporous alleles is interrupted. The dashed lines stand for the BLAST negative sequences found in the flanking region of the aposporous allele. The light blue box with black bar label marks the position of the original AFLP/CAPS marker fragment and the informative *Eco*RI site, respectively.

4.8. Conclusions

Comprehensive histological, genetic and molecular studies have been performed to analyse aposporous and sexual reproduction in the St. John´s Wort (*Hypericum perforatum*). From the described results I draw the following main conclusions:

- Apospory is inherited as a dominant trait.
- Apospory and parthenogenesis do co-segregate, but can recombine.
- An aposporous initial cell (AI) differentiates from an outer cell layer of the nucellus and outcompetes the development of the sexual megaspore mothercell (MMC).
- A reliable CAPS marker, co-segregating with apospory, but not with parthenogenesis, has been established and further confirms that apospory and parthenogenesis are independently controlled.
- A genomic region of 141.941 kb containing the apospory-marker has been identified in sexual plants by BAC cloning and designated as *HAPPY* (*Hypericum* APOSPORY).
- Allele quantification revealed that the hemizygous aposporous allele is present in simplex form both in segregating populations as well as in *Hypericum* accessions collected from all over the world.
- The *HAPPY* locus includes a gene with similarity to the Arabidopsis gene ARI7 encoding the RING finger protein ARIADNE7, known to be required for various regulatory processes, like ubiquitin-mediated protein degradation.
- The apospory-specific *HpARI* allele contains a premature stop codon and is predicted to encode a truncated protein.
- Both, the apospory-specific and the sex-specific *HpARI* alleles are expressed ubiquitously at the transcript level.
- Tetraploid Arabidopsis plants transformed with the truncated *HpARI* allele exhibit a late flowering phenotype.
- A current working hypothesis assumes that *HpARI* encodes a gene product with pleiotropic functions, also required to control the mode of reproduction.
- The partial cloning and characterisation of the *HAPPY* locus is a first step towards the elucidation of the molecular mechanisms underlying apomixis in *Hypericum*.
- Ongoing research deals with the extension of the *HAPPY* locus based BACs from aposporous plants as well as initial experiments for *Hypericum* transformation and complementation.

4.9. Future perspectives

It is likely that apomixis is not a single gene phenomenon, and that possibly several genes involved in meiosis and reproduction could show linkage to apomixis (Grossniklaus, 2001). While I concentrate on the functional role of *HpARI* in the control of apospory, I believe that the extension of the *HAPPY* locus beyond *HpARI* will be required to understand the full control of apospory and its putative modifiers in *Hypericum*. I have initiated cloning of the *HAPPY* locus by identifying and sequencing BACs with 8-fold genome-coverage from the tetraploid genotypes that are fully penetrant for apospory. The fine mapping of the *HAPPY* locus with genomic markers is also planned. This would enable the estimation of the size of the locus and would facilitate the sequencing of the locus. Once the sequences of the *HAPPY* locus are available, the detailed characterization of the locus, including candidate gene identification and examination of the *HAPPY* locus for epigenetic features should be performed. To fully understand the mechanisms of apomixis I also plan to concentrate my efforts on isolating the candidate locus for parthenogenesis from the segregating *Hypericum* population. These projects will provide valuable insights for the understanding of apomixis in *Hypericum* and in plants in general.

Summary

The introduction of apomixis - the seed formation without meiotic recombination and reduction (apomeiosis) and without fertilization (parthenogenesis) - into crop plants is an exciting problem of plant developmental biology and a high priority goal of breeding research, since it would allow the fixation of hybrid effects. However, the genetic and molecular basis of apomixis is still poorly understood.

St. John´s wort (*Hypericum perforatum* L.), a plant with a small genome and a short generation time, can reproduce by apospory and represents an interesting new experimental model for apomixis research. Here, we have combined histological, genetic and molecular analyses to study sexual and aposporous reproduction of *Hypericum*.

Histological analyses compared female sporogenesis and gametogenesis between sexual and aposporous accessions and led to the major phenotypic differences in embryo sac formation. In aposporous ovules the aposporous initial cells (AI) differentiate from the external nucellar cell layer. The AI development outcompetes and suppresses the megaspore mother cell of the sexual pathway.

Genetic analyses both at the triploid and tetraploid level reveal the dominant inheritance of apospory and indicate that apospory and parthenogenesis do co-segregate but can recombine in *Hypericum*.

The genetic data were extended by molecular analyses. AFLP profiling led to the identification of a CAPS marker which co-segregates with apospory but not with parthenogenesis, further confirming that both components of apomixis - apospory and parthenogenesis - are controlled independently.

The CAPS marker was used to isolate BAC clones with the corresponding genomic region, designated as *HAPPY* (*Hypericum APOSPORY*). Allele quantification revealed that the hemizygous apospory-specific allele is present in simplex constellation both in segregating crossing populations as well as in *Hypericum* accessions collected from all over the world.

The marker-containing part of the *HAPPY* locus exhibits sequence similarity to the Arabidopsis gene *ARI7* encoding the RING finger protein ARIADNE7. These proteins are

known to be involved in various regulatory processes, including ubiquitin-mediated protein degradation. The apospory-specific *HpARI* allele is predicted to encode a truncated protein. Both alleles are ubiquitously co-expressed at the transcript level. Tetraploid *Arabidopsis* lines over-expressing the truncated *HpARI* exhibit a late flowering phenotype, suggesting a putative role of *HpARI* in regulation of flowering time. Currently, it is speculated that *HpARI* represents a pleiotropic gene and might be also involved in the control of sexual or/and aposporous reproduction.

The partial cloning and characterisation of the *HAPPY* locus is a first step towards the elucidation of the molecular mechanisms underlying apomictic seed formation in *Hypericum perforatum*.

Zusammenfassung

Apomixis ist Samenbildung ohne meiotische Rekombination und Reduktion (Apomeiosis) und ohne Befruchtung (Parthenogenese). Die Übertragung von Apomixis aus natürlichen Apomikten auf Kulturpflanzen ist ein zentrales Problem pflanzlicher Entwicklungsbiologie und von hoher Priorität für die Pflanzenzüchtung. Unter anderem erhofft man sich davon die Fixierung von Hybrideffekten. Die genetischen und molekularen Grundlagen von Apomixis sind aber noch weitgehend unverstanden.

Johanniskraut (*Hyperiucm perforatum*), eine Pflanze mit kleinem Genom und kurzer Generationszeit, kann sich durch Aposporie und Parthenogenese reproduzieren. Dadurch eignet es sich als ein interessantes neues experimentelles Modell für die Apomixisforschung.

Die vergleichende histologische Analyse der weiblichen Sporogenese und Gametogenese zwischen sexuellen und aposporen Pflanzen führte zur Identifizierung der wesentlichen phänotypischen Unterschiede in der Embryosackbildung. In aposporen Ovulen differenziert die apospore Initialzelle (AI) von der äußeren Zellschicht des Nucellus. Die AI entwickelt sich sehr schnell und unterdrückt die weitere Entwicklung der Megasporen Mutterzelle des sexuellen *pathways*.

Genetische Analysen auf triploidem und tetraploidem Niveau zeigen die dominante Vererbung von Apoporie sowie die Kosegregation von Aposporie und Parthenogenese, auch wenn Rekombination zwischen beiden Komponenten möglich ist.

Die genetischen Befunde wurden durch molekulare Analysen erweitert. AFLP-Techniken führten zur Identifizierung eines CAPS-Markers, der wohl mit Aposporie, nicht aber mit Parthenogenese kosegregiert. Dies bestätigt erneut, dass beide Komponenten von Apomixis - Aposporie und Parthenogenese - unabhängig voneinander kontrolliert werden.

Der CAPS-Marker wurde für Isolierung von BAC-Klonen mit der entsprechenden, als *HAPPY*-Locus (Hypericum APOSPORY) bezeichneten genomischen Region benutzt. Allel-Quantifizierung zeigte, dass das hemizygote apospore Allel in Simplex-Konstellation vorliegt. Dies gilt sowohl für die untersuchten segregierenden Kreuzungsnachkommen als auch für weltweit gesammelte *Hypericum* Akzessionen.

Zusammenfassung

Der den Marker enthaltende Teil des *HAPPY*-Locus besitzt Sequenzähnlichkeit mit dem Arabidopsis-Gen *ARI7*, welches für das RING-Finger Protein ARIADNE7 kodiert. Diese Proteine sind an verschiedenen regulatorischen Prozessen, einschließlich dem Ubiquitin-vermittelten Proteinabbau beteiligt. Das Aposporie-spezifische *HpARI* Allel kodiert für ein verkürztes Protein. Beide Allele werden auf Transkriptebene ubiquitär exprimiert. Tetraploide *Arabidopsis*-Pflanzen, die mit dem verkürzten *HpARI* Allel transformiert wurden, blühen verspätet. Dies weist auf eine pleiotrope, möglicherweise die sexuelle und apospore Reproduktion kontrollierende Funktion des *HpARI* Gens hin.

Die partielle Klonierung und Charakterisierung des *HAPPY*-Locus ist ein wesentlicher erster Schritt zur Aufklärung der molekularen Mechanismen, die der apomiktischen Samenbildung bei *Hypericum perforatum* zu Grunde liegen.

References

Agostinis, P., Vantieghem, A., Merlevede, W. and de Witte, P.A. (2002) Hypericin in cancer treatment: more light on the way. *Int. J. Biochem. Cell Biol.*, **34**, 221-241.

Aguilera, M., Oliveros, M., Martinez-Padron, M., Barbas, J.A. and Ferrus, A. (2000) Ariadne-1: a vital Drosophila gene is required in development and defines a new conserved family of ring-finger proteins. *Genetics*, **155**, 1231-1244.

Akiyama, Y., Conner, J.A., Goel, S., Morishige, D.T., Mullet, J.E., Hanna, W.W. and Ozias-Akins, P. (2004) High-resolution physical mapping in Pennisetum squamulatum reveals extensive chromosomal heteromorphism of the genomic region associated with apomixis. *Plant Physiol*, **134**, 1733-1741.

Akiyama, Y., Hanna, W.W. and Ozias-Akins, P. (2005) High-resolution physical mapping reveals that the apospory-specific genomic region (ASGR) in Cenchrus ciliaris is located on a heterochromatic and hemizygous region of a single chromosome. *Theor Appl Genet*, **111**, 1042-1051.

Albertini, E., Porceddu, A., Ferranti, F., Reale, L., Barcaccia, G., Romano, B. and Falcinelli, M. (2001) Apospory and parthenogenesis may be uncoupled in Poa pratensis: a cytological investigation. *Sex Plant Reprod*, **14**, 213–217.

Albertini, E., Porceddu, A., Marconi, G., Barcaccia, G., Pallottini, L. and Falcinelli, M. (2003) Microsatellite-AFLP for genetic mapping of complex polyploids. *Genome*, **46**, 824-832.

Albertini, E., Marconi, G., Barcaccia, G., Raggi, L. and Falcinelli, M. (2004) Isolation of candidate genes for apomixis in Poa pratensis L. *Plant Mol Biol*, **56**, 879-894.

Albertini, E., Marconi, G., Reale, L., Barcaccia, G., Porceddu, A., Ferranti, F. and Falcinelli, M. (2005) SERK and APOSTART. Candidate genes for apomixis in Poa pratensis. *Plant Physiol*, **138**, 2185-2199.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol*, **215**, 403-410.

Austin, M.J., Muskett, P., Kahn, K., Feys, B.J., Jones, J.D. and Parker, J.E. (2002) Regulatory role of SGT1 in early R gene-mediated plant defenses. *Science*, **295**, 2077-2080.

Azevedo, C., Sadanandom, A., Kitagawa, K., Freialdenhoven, A., Shirasu, K. and Schulze-Lefert, P. (2002) The RAR1 interactor SGT1, an essential component of R gene-triggered disease resistance. *Science*, **295**, 2073-2076.

Barcaccia, G., Mazzucato, A., Falcinelli, M. and Veronesi, F. (1996) Callose localization in cell walls during meiotic and apomeiotic megasporogenesis in diploid alfalfa (Medicago spp.). *Caryologia*, **49**, 45-56.

Barcaccia, G., Mazzucato, A., Albertini, E., Zethof, J., Gerats, A., Pezzotti, M. and Falcinelli, M. (1998) Inheritance of parthenogenesis in Poa pratensis L.: auxin test and AFLP linkage analyses support monogenic control. *TAG*, **97**, 74-82.

Barcaccia, G., Albertini, E., Rosellini, D., Tavoletti, S. and Veronesi, F. (2000) Inheritance and mapping of 2n-egg production in diploid alfalfa. *Genome*, **43**, 528-537.

Barcaccia, G., Arzenton, F., Sharbel, T.F., 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., Baumlein, H. and Sharbel T. (2007) Apomixis in St. Johns wort (*Hypericum perforatum* L.): an overview and glimpse towards the future In *Apomixis: evolution, mechanisms and perspectives* (Hörandl, E., Grossniklaus, U., Van Dijk, P. and Sharbel, T. eds). Vienna: International Association for Plant Taxonomy, Institute of Botany, University of Vienna, pp. 259-280.

Barrell, P.J. and Grossniklaus, U. (2006) Confocal microscopy of whole ovules for analysis of reproductive development: the elongate1 mutant affects meiosis II. *Plant J*, **43**, 309-320.

Bassam, B.J., Caetano-Anolles, G. and Gresshoff, P.M. (1991) Fast and sensitive silver staining of DNA in polyacrylamide gels. *Anal Biochem*, **196**, 80-83.

Baurle, I., Smith, L., Baulcombe, D.C. and Dean, C. (2007) Widespread role for the flowering-time regulators FCA and FPA in RNA-mediated chromatin silencing. *Science*, **318**, 109-112.

Bennett, S.T., Leitch, I.J. and Bennett, M.D. (1995) Chromosome identification and mapping in the grass Zingeria biebersteiniana (2n = 4) using fluorochromes. *Chromosome Res*, **3**, 101-108.

Berger, F., Grini, P.E. and Schnittger, A. (2006) Endosperm: an integrator of seed growth and development. *Curr Opin Plant Biol*, **9**, 664-670.

Bicknell, R.A. (1997) Isolation of a diploid apomictic plant of Hieracium aurantiacum. *Sex Plant Reprod*, **10**, 168-172.

Bicknell, R.A., Borst, N.K. and Koltunow, A.M. (2000) Monogenic inheritance of apomixis in two Hieracium species with distinct developmental mechanisms. *Heredity*, **84 (Pt 2)**, 228-237.

Bicknell, R.A. and Koltunow, A.M. (2004) Understanding apomixis: recent advances and remaining conundrums. *Plant Cell*, **16 Suppl**, S228-245.

Birchler, J.A. (1993) Dosage analysis of maize endosperm development. *Annu Rev Genet*, **27**, 181-204.

Boutilier, K., Offringa, R., Sharma, V.K., Kieft, H., Ouellet, T., Zhang, L., Hattori, J., Liu, C.M., van Lammeren, A.A., Miki, B.L., Custers, J.B. and van Lookeren Campagne, M.M. (2002) Ectopic expression of BABY BOOM triggers a conversion from vegetative to embryonic growth. *Plant Cell*, **14**, 1737-1749.

Bouveret, R., Schonrock, N., Gruissem, W. and Hennig, L. (2006) Regulation of flowering time by Arabidopsis MSI1. *Development*, **133**, 1693-1702.

Brukhin, V., Curtis, M. and Grossniklaus, U. (2005) The angiosperm female gametophyte: no longer the forgotten generation. *Current Science*, **89**, 1844-1852.

Calderini, O., Chang, S.B., de Jong, H., Busti, A., Paolocci, F., Arcioni, S., de Vries, S.C., Abma-Henkens, M.H., Lankhorst, R.M., Donnison, I.S. and Pupilli, F. (2006) Molecular cytogenetics and DNA sequence analysis of an apomixis-linked BAC in Paspalum simplex reveal a non pericentromere location and partial microcolinearity with rice. *Theor Appl Genet*, **112**, 1179-1191.

Calderini, O., Chang, S.B., de Jong, H., Busti, A., Paolocci, F., Arcioni, S., de Vries, S.C., Abma-Henkens, M.H., Lankhorst, R.M., Donnison, I.S. and Pupilli, F. (2007) Molecular cytogenetics of the apomixis controlled locus in Paspalum simplex. In *Apomixis: evolution, mechanisms and perspectives* (Hörandl, E., Grossniklaus, U., Van Dijk, P. and Sharbel, T. eds). Vienna: International Association for Plant Taxonomy, Institute of Botany, University of Vienna, pp. 125-136.

Carman, J., Crane, C. and Riera-Lizarazu, O. (1991) Comparative histology of cell walls during meiotic and apomeiotic megasporogenesis in two hexaploid Australasian Elymus species. *Crop Sci*, **31**, 1527-1532.

Carman, J.G. (1997) Asynchronous expression of duplicate genes in angiosperms may cause apomixis, bispory, tetraspory, and polyembryony. *Biol J Linn Soc*, **61**, 51-94.

Casacuberta, E., Puigdomenech, P. and Monfort, A. (2000) Distribution of microsatellites in relation to coding sequences within the Arabidopsis thaliana genome. *Plant Sci*, **157**, 97-104.

Catanach, A.S., Erasmuson, S.K., Podivinsky, E., Jordan, B.R. and Bicknell, R. (2006) Deletion mapping of genetic regions associated with apomixis in Hieracium. *Proc Natl Acad Sci U S A*, **103**, 18650-18655.

Chaudhury, A.M., Ming, L., Miller, C., Craig, S., Dennis, E.S. and Peacock, W.J. (1997) Fertilization-independent seed development in Arabidopsis thaliana. *Proc Natl Acad Sci U S A*, **94**, 4223-4228.

Chen, H., Nelson, R.S. and Sherwood, J.L. (1994) Enhanced recovery of transformants of Agrobacterium tumefaciens after freeze-thaw transformation and drug selection. *Biotechniques*, **16**, 664-668, 670.

Chen, L., Guan, L., Seo, M., Hoffmann, F. and Adachi, T. (2005) Developmental expression of ASG- 1 during gametogenesis in apomictic guinea grass (Panicum maximum). *J Plant Physiol*, **162**, 1141-1148.

Cheng, Y., Kato, N., Wang, W., Li, J. and Chen, X. (2003) Two RNA binding proteins, HEN4 and HUA1, act in the processing of AGAMOUS pre-mRNA in Arabidopsis thaliana. *Dev Cell*, **4**, 53-66.

Childs, K.L., Hamilton, J.P., Zhu, W., Ly, E., Cheung, F., Wu, H., Rabinowicz, P.D., Town, C.D., Buell, C.R. and Chan, A.P. (2007) The TIGR Plant Transcript Assemblies database. *Nucleic Acids Res*, **35**, D846-851.

Church, G.M. and Gilbert, W. (1984) Genomic sequencing. *Proc Natl Acad Sci U S A*, **81**, 1991-1995.

Clarke, L. and Carbon, J. (1976) A colony bank containing synthetic Col El hybrid plasmids representative of the entire E. coli genome. *Cell*, **9**, 91-99.

Clough, S.J. and Bent, A.F. (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J*, **16**, 735-743.

Cohen, S.N., Chang, A.C. and Hsu, L. (1972) Nonchromosomal antibiotic resistance in bacteria: genetic transformation of Escherichia coli by R-factor DNA. *Proc Natl Acad Sci U S A*, **69**, 2110-2114.

Conner, J.A., Goel, S., Gunawan, G., Cordonnier-Pratt, M.M., Johnson, V.E., Liang, C., Wang, H., Pratt, L.H., Mullet, J.E., Debarry, J., Yang, L., Bennetzen, J.L., Klein, P.E. and Ozias-Akins, P. (2008) Sequence analysis of bacterial artificial chromosome clones from the apospory-specific genomic region of Pennisetum and Cenchrus. *Plant Physiol*, **147**, 1396-1411.

d'Erfurth, I., Jolivet, S., Froger, N., Catrice, O., Novatchkova, M. and Mercier, R. (2009) Turning meiosis into mitosis. *PLoS Biol*, **7**, e1000124.

Deshpande, G., Calhoun, G. and Schedl, P. (2006) The drosophila fragile X protein dFMR1 is required during early embryogenesis for pole cell formation and rapid nuclear division cycles. *Genetics*, **174**, 1287-1298.

Di Carlo, G., Borrelli, F., Ernst, E. and Izzo, A.A. (2001) St John's wort: Prozac from the plant kingdom. *Trends Pharmacol. Sci.* , **22**, 292-297.

Dieterle, M., Zhou, Y.C., Schafer, E., Funk, M. and Kretsch, T. (2001) EID1, an F-box protein involved in phytochrome A-specific light signaling. *Genes Dev*, **15**, 939-944.

do Valle, C.B., Glienke, C. and Leguizamon, G.O.C. (1993) Towards defining the inheritance of apomixis in Brachiaria. *Apomixis Newsletter*, 24-25.

Dujardin, M. and Hanna, W.W. (1983) Apomictic and sexual pearl millet X Pennisetum squamulatum hybrids.*Heredity*, **74**, 277-279.

Dulger, B. and Gonuz, A. (2005) Antibacterial activity of the endemic Hypericum kazdaghensis. *Fitoterapia*, **76**, 237-239.

Ebina, M., Nakagawa, H., Yamamoto, T., Araya, H., Tsuruta, S.-i., Takahara, M. and Nakajima, K. (2005) Co-segregation of AFLP and RAPD markers to apospory in Guineagrass (Panicum maximum Jacq.). *Grassland Science*, **51**, 71 - 78.

Feinberg, A.P. and Vogelstein, B. (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal Biochem*, **132**, 6-13.

Fenner, R., Sortino, M., Rates, S.M., Dall'Agnol, R., Ferraz, A., Bernardi, A.P., Albring, D., Nor, C., von Poser, G., Schapoval, E. and Zacchino, S. (2005) Antifungal activity of some Brazilian Hypericum species. *Phytomedicine*, **12**, 236-240.

Ferraz, A.B., Grivicich, I., von Poser, G.L., Faria, D.H., Kayser, G.B., Schwartsmann, G., Henriques, A.T. and da Rocha, A.B. (2005) Antitumor activity of three benzopyrans isolated from Hypericum polyanthemum. *Fitoterapia*, **76**, 210-215.

Francis, A.J. (2005) Antidepressant action of St. John's Wort, Hypericum perforatum: a test of the circadian hypotheses. *Phytomedicine*, **12**, 167-172.

Friedman, W.E. (2001) Developmental and evolutionary hypotheses for the origin of double fertilization and endosperm. *C R Acad Sci III*, **324**, 559-567.

Goel, S., Chen, Z., Conner, J.A., Akiyama, Y., Hanna, W.W. and Ozias-Akins, P. (2003) Delineation by fluorescence in situ hybridization of a single hemizygous chromosomal region associated with aposporous embryo sac formation in Pennisetum squamulatum and Cenchrus ciliaris. *Genetics*, **163**, 1069-1082.

Goel, S., Chen, Z., Akiyama, Y., Conner, J.A., Basu, M., Gualtieri, G., Hanna, W.W. and Ozias-Akins, P. (2006) Comparative physical mapping of the apospory-specific genomic region in two apomictic grasses: Pennisetum squamulatum and Cenchrus ciliaris. *Genetics*, **173**, 389-400.

Gray, W.M., Kepinski, S., Rouse, D., Leyser, O. and Estelle, M. (2001) Auxin regulates SCF(TIR1)-dependent degradation of AUX/IAA proteins. *Nature*, **414**, 271-276.

Gremme, G., Brendel, V., Sparks, M.E. and Kurtz, S. (2005) Engineering a software tool for gene structure prediction in higher organisms. *Information and Software Technology*, **47**, 965-978.

Grimanelli, D., Leblanc, O., Espinosa, E., Perotti, E., Gonzalez de Leon, D. and Savidan, Y. (1998) Non-Mendelian transmission of apomixis in maize-Tripsacum hybrids caused by a transmission ratio distortion. *Heredity*, **80 (Pt 1)**, 40-47.

Grimanelli, D., Leblanc, O., Perotti, E. and Grossniklaus, U. (2001) Developmental genetics of gametophytic apomixis. *Trends Genet*, **17**, 597-604.

Gross-Hardt, R., Kagi, C., Baumann, N., Moore, J.M., Baskar, R., Gagliano, W.B., Jurgens, G. and Grossniklaus, U. (2007) LACHESIS restricts gametic cell fate in the female gametophyte of Arabidopsis. *PLoS Biol*, **5**, e47.

Grossniklaus, U. and Schneitz, K. (1998) The molecular and genetic basis of ovule and megagametophyte development. *Seminars in Cell & Developmental Biology*, **9**, 227-238.

Grossniklaus, U., Nogler, G.A. and van Dijk, P.J. (2001) How to avoid sex: the genetic control of gametophytic apomixis. *Plant Cell*, **13**, 1491-1498.

Grossniklaus, U. (2001b) From Sexuality to apomixis: Molecular and genetic approaches. In *The flowering of apomixis: From mechanisms to genetic engineering* (Savidan, Y., Carman, J.G. and Dresselhaus, T. eds). Mexico: CIMMYT, IRD, European Commission DG VI (FAIR)), pp. 168-202.

Gualtieri, G., Conner, J.A., Morishige, D.T., Moore, L.D., Mullet, J.E. and Ozias-Akins, P. (2006) A segment of the apospory-specific genomic region is highly microsyntenic not only between the apomicts Pennisetum squamulatum and buffelgrass, but also with a rice chromosome 11 centromeric-proximal genomic region. *Plant Physiol*, **140**, 963-971.

Guitton, A.E., Page, D.R., Chambrier, P., Lionnet, C., Faure, J.E., Grossniklaus, U. and Berger, F. (2004) Identification of new members of Fertilisation Independent Seed Polycomb Group pathway involved in the control of seed development in Arabidopsis thaliana. *Development*, **131**, 2971-2981.

Haig, D. (1986) Conflicts among megaspores. *Journal of Theoretical Biology*, **123**, 471-480

Haig, D. and Westoby, M. (1991) Genomic Imprinting in Endosperm: Its Effect on Seed Development in Crosses between Species, and between Different Ploidies of the Same Species, and Its Implications for the Evolution of Apomixis. *Philosophical Transactions: Biological Sciences*, **333**, 1-13.

Hofmeister, W. (1851) *Vergleichende Untersuchungen der Keimung. Entfaltung und Fruchtbildung höherer Kryptogamen (Moose, Farrn, Equisetaceen, Rhizocarpeen und Lycopodiaceen) und die Samenbildung der Coniferen* Leipzig, Germany.

Holmes, D.S. and Quigley, M. (1981) A rapid boiling method for the preparation of bacterial plasmids. *Anal Biochem*, **114**, 193-197.

Holsinger, K.E. (2000) Reproductive systems and evolution in vascular plants. *Proc. Natl. Acad. Sci. USA*, **97**, 7037-7042.

Honys, D. and Twell, D. (2004) Transcriptome analysis of haploid male gametophyte development in Arabidopsis. *Genome Biol*, **5**, R85.

Huang, B.Q. and Sheridan, W.F. (1996) Embryo Sac Development in the Maize indeterminate gametophyte1 Mutant: Abnormal Nuclear Behavior and Defective Microtubule Organization. *Plant Cell*, **8**, 1391-1407.

Jackson, P.K., Eldridge, A.G., Freed, E., Furstenthal, L., Hsu, J.Y., Kaiser, B.K. and Reimann, J.D. (2000) The lore of the RINGs: substrate recognition and catalysis by ubiquitin ligases. *Trends Cell Biol*, **10**, 429-439.

Jessup R.W., Burson B.L., Burow O., Wang Y.W., Chang C., Li Z., Paterson A.H. and M.A., H. (2003) Segmental allotetraploidy and allelic interactions in buffelgrass (Pennisetum ciliare (L.) Link syn. Cenchrus ciliaris L.) as revealed by genome mapping. *Genome*, **46**, 304-313.

Kankel, M.W., Ramsey, D.E., Stokes, T.L., Flowers, S.K., Haag, J.R., Jeddeloh, J.A., Riddle, N.C., Verbsky, M.L. and Richards, E.J. (2003) Arabidopsis MET1 cytosine methyltransferase mutants. *Genetics*, **163**, 1109-1122.

Kantama, L., Sharbel, T.F., Schranz, M.E., Mitchell-Olds, T., de Vries, S. and de Jong, H. (2007) Diploid apomicts of the Boechera holboellii complex display large-scale chromosome substitutions and aberrant chromosomes. *Proc Natl Acad Sci U S A*, **104**, 14026-14031.

Katz, A., Oliva, M., Mosquna, A., Hakim, O. and Ohad, N. (2004) FIE and CURLY LEAF polycomb proteins interact in the regulation of homeobox gene expression during sporophyte development. *Plant J*, **37**, 707-719.

Kinoshita, T., Miura, A., Choi, Y., Kinoshita, Y., Cao, X., Jacobsen, S.E., Fischer, R.L. and Kakutani, T. (2004) One-way control of FWA imprinting in Arabidopsis endosperm by DNA methylation. *Science*, **303**, 521-523.

Kohler, C., Hennig, L., Bouveret, R., Gheyselinck, J., Grossniklaus, U. and Gruissem, W. (2003) Arabidopsis MSI1 is a component of the MEA/FIE Polycomb group complex and required for seed development. *Embo J*, **22**, 4804-4814.

Kohler, C., Hennig, L., Spillane, C., Pien, S., Gruissem, W. and Grossniklaus, U. (2003b) The Polycomb-group protein MEDEA regulates seed development by controlling expression of the MADS-box gene PHERES1. *Genes Dev*, **17**, 1540-1553.

Kohler, C. and Makarevich, G. (2006) Epigenetic mechanisms governing seed development in plants. *EMBO Rep*, **7**, 1223-1227.

Kojima, A. and Nagato, Y. (1997) Discovery of highly apomictic and highly amphimictic dihaploids in Allium tuberosum. *Sex Plant Reprod*, **10**, 8-12.

Koltunow, A.M. (1993) Apomixis: Embryo Sacs and Embryos Formed without Meiosis or Fertilization in Ovules. *Plant Cell*, **5**, 1425-1437.

Koltunow, A.M., Johnson, S.D. and Bicknell, R.A. (1998) Sexual and apomictic development in Hieracium. *Sex Plant Reprod*, **11**, 213-230.

Koltunow, A.M. and Grossniklaus, U. (2003) Apomixis: a developmental perspective. *Annu Rev Plant Biol*, **54**, 547-574.

Labombarda, P., Busti, A., Caceres, M.E., Pupilli, F. and Arcioni, S. (2002) An AFLP marker tightly linked to apomixis reveals hemizygosity in a portion of the apomixis-controlling locus in Paspalum simplex. *Genome*, **45**, 513-519.

Leblanc, O., Peel, M., Carman, J. and Savidan, Y. (1993) Megasporogenesis in sexual and apomictic Trypsacum species using interference contrasts and fluorescence. *Apomixis Newsletter*, **6**, 14-17.

Leblanc, O., Grimanelli, D., Gonzalez de Leon and D. and Savidan, Y. (1995) Detection of the apomictic mode of reproduction in maize-Tripsacum hybrids using maize RFLP markers. *Theor. Appl. Genet.* , **90**, 1198-1203.

Liu, J., Zhang, Y., Qin, G., Tsuge, T., Sakaguchi, N., Luo, G., Sun, K., Shi, D., Aki, S., Zheng, N., Aoyama, T., Oka, A., Yang, W., Umeda, M., Xie, Q., Gu, H. and Qu, L.J. (2008) Targeted degradation of the cyclin-dependent kinase inhibitor ICK4/KRP6 by RING-type E3 ligases is essential for mitotic cell cycle progression during Arabidopsis gametogenesis. *Plant Cell*, **20**, 1538-1554.

Lohe, A.R. and Chaudhury, A. (2002) Genetic and epigenetic processes in seed development. *Curr Opin Plant Biol*, **5**, 19-25.

Lomer, F. and Douglas, G.W. (1998) Clusiaceae. In *Illustrated flora of British Columbia* (Douglas, G.W., Straley, G.B., Meidinger, D. and Pojar, J. eds). British Columbia: Ministry of Environment, Lands and Parks, Ministry of Forests, pp. 336-342.

Lorkovic, Z.J. and Barta, A. (2002) Genome analysis: RNA recognition motif (RRM) and K homology (KH) domain RNA-binding proteins from the flowering plant Arabidopsis thaliana. *Nucleic Acids Res*, **30**, 623-635.

Lotan, T., Ohto, M., Yee, K.M., West, M.A., Lo, R., Kwong, R.W., Yamagishi, K., Fischer, R.L., Goldberg, R.B. and Harada, J.J. (1998) Arabidopsis LEAFY COTYLEDON1 is sufficient to induce embryo development in vegetative cells. *Cell*, **93**, 1195-1205.

Luo, M.C., Thomas, C., You, F.M., Hsiao, J., Ouyang, S., Buell, C.R., Malandro, M., McGuire, P.E., Anderson, O.D. and Dvorak, J. (2003) High-throughput fingerprinting of bacterial artificial chromosomes using the snapshot labeling kit and sizing of restriction fragments by capillary electrophoresis. *Genomics*, **82**, 378-389.

Maheshawari (1950) *An introduction to the embryology of angiosperms* New-York: McGraw-Hill.

Malaty, W. (2005) St. John's wort for depression. *Am Fam Physician*, **71**, 1375-1376.

Martinez, E.J., Urbani, M.H., Quarin, C.L. and Ortiz, J.P. (2001) Inheritance of apospory in bahiagrass, Paspalum notatum. *Hereditas*, **135**, 19-25.

Mártonfi, P., Brutovská, R., Čellárová, E. and Repčák, M. (1996) Apomixis and Hybridity in Hypericum perforatum. *Folia Geobotanica & Phytotaxonomica*, **31**, 389-396.

Matzk, F. (1991) New efforts to overcome apomixis in Poa pratensis L. *Euphytica*, **55**, 65-72.

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

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

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

Matzk, F., Prodanovic, S., Baumlein, H. and Schubert, I. (2005) The Inheritance of apomixis in Poa pratensis confirms a five locus model with differences in gene expressivity and penetrance. *Plant Cell*, **17**, 13-24.

Mayo, G.M. and Langridge, P. (2003) Modes of reproduction in Australian populations of Hypericum perforatum L. (St. John's wort) revealed by DNA fingerprinting and cytological methods. *Genome*, **46**, 573-579.

McMenimen, S. and Lubulwa, G. (1997) *Project development assessment: an economic evaluation of the potential benefits of integrating apomixis into hybrid rice* Canberra: Australian Centre for International Agricultural Research.

Miles, J.W. and Escandon, M.L. (1996) Further evidence on the inheritance of reproductive mode in Brachiaria. *Can. J. Plant Sci*, **77**, 105-107.

Miskovsky, P. (2002) Hypericin--a new antiviral and antitumor photosensitizer: mechanism of action and interaction with biological macromolecules. *Curr Drug Targets*, **3**, 55-84.

Mladek, C., Guger, K. and Hauser, M.T. (2003) Identification and characterization of the ARIADNE gene family in Arabidopsis. A group of putative E3 ligases. *Plant Physiol*, **131**, 27-40.

Mockler, T.C., Yu, X., Shalitin, D., Parikh, D., Michael, T.P., Liou, J., Huang, J., Smith, Z., Alonso, J.M., Ecker, J.R., Chory, J. and Lin, C. (2004) Regulation of flowering time in Arabidopsis by K homology domain proteins. *Proc Natl Acad Sci U S A*, **101**, 12759-12764.

Moggie, M. (1992) *The evolution of asexual reproduction in plants* London: Chapman and Hall.

Murashige, T. and Skoog, F. (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiol. Plant*, **15**, 473-497.

Naumova, T., Den Nijs, A. and Willemse, M. (1993) Quantitative analysis of aposporous parthenogenesis in Poa pratensis genotypes. *Acta Bot Neerl*, **42**, 299-312.

Nawaschin, S.G. (1898) *Resultate einer Revision der Befruchtungsvorgaenge bei Lilium martagon und Fritillaria tenella* St Petersburg: Bull Acad Imp Sci

Nelson, D.C., Lasswell, J., Rogg, L.E., Cohen, M.A. and Bartel, B. (2000) FKF1, a clock-controlled gene that regulates the transition to flowering in Arabidopsis. *Cell*, **101**, 331-340.

Noack, K.L. (1939) Über Hypericum - Kreuzungen VI: Fortpflanzungsverhältnisse und bastarde von Hypericum perforatum L. *Z. Indukt Abstamm. Vererbungsl.*, **76**, 569-601.

Noack, K.L. (1941) Geschlechtsverlust und bastardierung beim Johanniskraut. *Vorsch. Vortschr.*, **17**, 13-15.

Nogler, G.A. (1982) How to obtain diploid apomictic Ranunculus auricomus plants not found in the wild state. *Bot Helv*, **91**, 13-22.

Nogler, G.A. (1984) *Gametophytic apomixis* Berlin: Springer-Verlag.

Noyes, R.D. and Rieseberg, L.H. (2000) Two independent loci control agamospermy (Apomixis) in the triploid flowering plant Erigeron annuus. *Genetics*, **155**, 379-390.

Noyes, R.D., Baker, R. and Mai, B. (2006) Mendelian segregation for two-factor apomixis in Erigeron annuus (Asteraceae). *Heredity*, **98**, 92-98.

O'Brien, M., Major, G., Chantha, S.-C. and Matton, D.P. (2004) Isolation of S-RNase binding proteins from Solanum chacoense: identification of an SBP1 (RING finger protein) orthologue. *Sex Plant Reprod*, **17**, 81-87.

Ohad, N., Margossian, L., Hsu, Y.C., Williams, C., Repetti, P. and Fischer, R.L. (1996) A mutation that allows endosperm development without fertilization. *Proc Natl Acad Sci U S A*, **93**, 5319-5324.

Okada, T., Catanach, A.S., Johnson, S.D., Bicknell, R.A. and Koltunow, A.M. (2007) An Hieracium mutant, loss of apomeiosis 1 (loa1) is defective in the initiation of apomixis. *Sex Plant Reprod*, **20**, 199-211.

Ozias-Akins, P., Roche, D. and Hanna, W.W. (1998) Tight clustering and hemizygosity of apomixis-linked molecular markers in Pennisetum squamulatum implies genetic control of apospory by a divergent locus that may have no allelic form in sexual genotypes. *Proc Natl Acad Sci U S A*, **95**, 5127-5132.

Ozias-Akins, P., Akiyama, Y. and Hanna, W.W. (2003) Molecular characterization of the genomic region linked with apomixis in Pennisetum/Cenchrus. *Funct Integr Genomics*, **3**, 94-104.

Ozias-Akins, P., Akiyama, Y. and Hanna, W.W. (2003) Molecular characterization of the genomic region linked with apomixis in Pennisetum/Cenchrus. *Funct Integr Genomics*, **3**, 94-104.

Ozias-Akins, P. (2006) Apomixis: Developmental Characteristics and Genetics. *Crit. Rev. in Pl. Sci.*, **25**, 199-214.

Ozias-Akins, P. and van Dijk, P.J. (2007) Mendelian genetics of apomixis in plants. *Annu Rev Genet*, **41**, 509-537.

Pagnussat, G.C., Yu, H.J. and Sundaresan, V. (2007) Cell-fate switch of synergid to egg cell in Arabidopsis eostre mutant embryo sacs arises from misexpression of the BEL1-like homeodomain gene BLH1. *Plant Cell*, **19**, 3578-3592.

Parra, G., Blanco, E. and Guigo, R. (2000) GeneID in Drosophila. *Genome Res*, **10**, 511-515.

Peacock, J. (1992) Plant biology. Twenty-first century crops. *Nature*, **357**, 358.

Peel, M., Carman, J. and Leblanc, O. (1997) Megasporocyte callose in apomictic buffelgrass, Kentucky bluegrass, Pennisetum squamulatum Fresen, Tripsacum L, and weeping lovegrass. *Crop Sci*, **37**, 724-732

Pessino, S.C., Ortiz, J.P.A., Leblanc, O., do Valle, C.B., Evans, C. and Hayward, M.D. (1997) Identification of a maize linkage group related to apomixis in Brachiaria *Theor Appl Genet*, **94**, 439-444.

Pessino, S.C., Espinoza, F., Martinez, E.J., Ortiz, J.P., Valle, E.M. and Quarin, C.L. (2001) Isolation of cDNA clones differentially expressed in flowers of apomictic and sexual Paspalum notatum. *Hereditas*, **134**, 35-42.

Potokina, E., Blattner, R., Alexandrova, T. and Bachmann, K. (2002) AFLP diversity in the common vetch (*Vicia sativa* L.) on the world scale. *Theor Appl Genet*, **105**, 58-67.

Praeckelt, U. and Scott, R. (2001) Introduction of apomixis in sexual plants by mutagenesis. In *The flowering of apomixis: From mechanisms to genetic engineering* (Savidan, Y., Carman, J.G. and Dresselhaus, T. eds). Mexico: CIMMYT, IRD, European Commission DG VI (FAIR)), pp. 212-227.

Pupilli, F., Labombarda, P., Caceres, M.E., Quarín, C.L. and Arcioni, S. (2001) The chromosome segment related to apomixis in Paspalum simplex is homoeologous to the telomeric region of the long arm of rice chromosome 12 *Mol Breed*, **8**, 53-61.

Pupilli, F., Martinez, E.J., Busti, A., Calderini, O., Quarin, C.L. and Arcioni, S. (2004) Comparative mapping reveals partial conservation of synteny at the apomixis locus in Paspalum spp. *Mol Genet Genomics*, **270**, 539-548.

Ravi, M., Marimuthu, M.P. and Siddiqi, I. (2008) Gamete formation without meiosis in Arabidopsis. *Nature*, **451**, 1121-1124.

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

Roche, D., Hanna, W.W. and Ozias-Akins, P. (2001) Is supernumerary chromatin involved in gametophytic apomixis of polyploid plants? . *Sex. Pl. Reprod.*, **13**, 343-349.

Rodkiewitz, B. (1970) Callose in cell walls during megasporogenesis in angiosperms. *Planta*, **93**, 39-47.

Rogers, S., Wells, R. and Rechsteiner, M. (1986) Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. *Science*, **234**, 364-368.

Roy, B.A. (1995) The Breeding Systems of Six Species of Arabis (Brassicaceae). *American Journal of Botany*, **82**, 869-877.

Rutishauser, A. (1954) [Developmental stimulation of the egg cell in pseudogamous species of Ranunculus genus.]. *Bull Schweiz Akad Med Wiss*, **10**, 491-512.

Samach, A., Klenz, J.E., Kohalmi, S.E., Risseeuw, E., Haughn, G.W. and Crosby, W.L. (1999) The UNUSUAL FLORAL ORGANS gene of Arabidopsis thaliana is an F-box protein required for normal patterning and growth in the floral meristem. *Plant J*, **20**, 433-445.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: a Laboratory Manual* NY: 2nd edn. Cold Spring Harbor, Cold Spring Harbor Laboratory

Savidan, Y. (1980) Chromosomal and embryological analyses in sexual x apomictic hybrids of Panicum maximum Jacq. . *TAG* **58**, 153-156.

Savidan, Y. (2000) Apomixis: genetics and breeding. *Plant Breed. Rev*, **18**, 13-85.

Savidan, C., Dresselhaus (2001) *The Flowering of apomixis: from mechanisms to genetic ingineering* Mexico: CIMMYT, IRD, European Commision DG VI.

Scheid, O.M., Afsar, K. and Paszkowski, J. (2003) Formation of stable epialleles and their paramutation-like interaction in tetraploid Arabidopsis thaliana. *Nat. Genet.*, **34**, 450-454.

Schempp, C.M., Muller, K.A., Winghofer, B., Schopf, E. and Simon, J.C. (2002) [St. John's wort (Hypericum perforatum L.). A plant with relevance for dermatology]. *Hautarzt*, **53**, 316-321.

Schmid, M., Davison, T.S., Henz, S.R., Pape, U.J., Demar, M., Vingron, M., Scholkopf, B., Weigel, D. and Lohmann, J.U. (2005) A gene expression map of Arabidopsis thaliana development. *Nat Genet*, **37**, 501-506.

Schneitz, K., Hülskamp, M. and Pruitt, R.E. (2002) Wild-type ovule development in Arabidopsis thaliana: a light microscope study of cleared whole-mount tissue. *Plant J.*, **7**, 731 - 749.

Sharbel, T.F., Voigt, M.L., Corral, J.M., Thiel, T., Varshney, A., Kumlehn, J., Vogel, H. and Rotter, B. (2009) Molecular signatures of apomictic and sexual ovules in the Boechera holboellii complex. *Plant J*, **58**, 870-882.

Shen, W.H., Parmentier, Y., Hellmann, H., Lechner, E., Dong, A., Masson, J., Granier, F., Lepiniec, L., Estelle, M. and Genschik, P. (2002) Null mutation of AtCUL1 causes arrest in early embryogenesis in Arabidopsis. *Mol Biol Cell*, **13**, 1916-1928.

Sheridan, W.F., Avalkina, N.A., Shamrov, II, Batygina, T.B. and Golubovskaya, I.N. (1996) The mac1 gene: controlling the commitment to the meiotic pathway in maize. *Genetics*, **142**, 1009-1020.

Sherwood, R.T., Berg, C.C. and Young, B.A. (1994) Inheritance of Apospory in Buffelgrass. *Crop Sci*, **34**, 1490-1494.

Sieber, P., Gheyselinck, J., Gross-Hardt, R., Laux, T., Grossniklaus, U. and Schneitz, K. (2004) Pattern formation during early ovule development in Arabidopsis thaliana. *Devep. Biol.*, **273**, 321-334.

Simko, I. (2004) One potato, two potato: haplotype association mapping in autotetraploids. *Trends Plant Sci*, **9**, 441-448.

Singh, M., Burson, B.L. and Finlayson, S.A. (2007) Isolation of candidate genes for apomictic development in buffelgrass (Pennisetum ciliare). *Plant Mol Biol*, **64**, 673-682.

Somers, D.E., Schultz, T.F., Milnamow, M. and Kay, S.A. (2000) ZEITLUPE encodes a novel clock-associated PAS protein from Arabidopsis. *Cell*, **101**, 319-329.

Song, R. and Messing, J. (2003) Gene expression of a gene family in maize based on noncollinear haplotypes. *Proc Natl Acad Sci U S A*, **100**, 9055-9060.

Spencer, M.W., Casson, S.A. and Lindsey, K. (2007) Transcriptional profiling of the Arabidopsis embryo. *Plant Physiol*, **143**, 924-940.

Spielman, M., Vinkenoog, R. and Scott, R.J. (2003) Genetic mechanisms of apomixis. *Philos Trans R Soc Lond B Biol Sci*, **358**, 1095-1103.

Spillane, C., Vielle-Calzada, J.P. and Grossniklaus, U. (2001) APO2001: A sexy apomixer in como. *Plant Cell*, **13**, 1480-1491.

Spillane, C., Curtis, M.D. and Grossniklaus, U. (2004) Apomixis technology development-virgin births in farmers' fields? *Nat Biotechnol*, **22**, 687-691.

Stelly, D.M., Peloquin, S.J., Palmer, R.G. and Crane, C.F. (1984) Mayer's hemalum-methyl salicylate: a stain-clearing technique for observations within whole ovules. *Stain Technol*, **59**, 155-161.

Stone, S.L., Kwong, L.W., Yee, K.M., Pelletier, J., Lepiniec, L., Fischer, R.L., Goldberg, R.B. and Harada, J.J. (2001) LEAFY COTYLEDON2 encodes a B3 domain transcription factor that induces embryo development. *Proc Natl Acad Sci U S A*, **98**, 11806-11811.

Suzek, B.E., Huang, H., McGarvey, P., Mazumder, R. and Wu, C.H. (2007) UniRef: comprehensive and non-redundant UniProt reference clusters. *Bioinformatics*, **23**, 1282-1288.

Swan, A. and Suter, B. (1996) Role of Bicaudal-D in patterning the Drosophila egg chamber in mid-oogenesis. *Development*, **122**, 3577-3586.

Tamura, K., Dudley, J., Nei, M. and Kumar, S. (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol*, **24**, 1596-1599.

Thome, O.W. (1885) *Flora von Deutschland, Österreich und der Schweiz- in Wort und Bild für Schule und Haus*. Gera, p.572

Tucker, M.R., Paech, N.A., Willemse, M.T. and Koltunow, A.M. (2001) Dynamics of callose deposition and beta-1,3-glucanase expression during reproductive events in sexual and apomictic Hieracium. *Planta*, **212**, 487-498.

Tucker, M.R., Araujo, A.C., Paech, N.A., Hecht, V., Schmidt, E.D., Rossell, J.B., De Vries, S.C. and Koltunow, A.M. (2003) Sexual and apomictic reproduction in Hieracium subgenus pilosella are closely interrelated developmental pathways. *Plant Cell*, **15**, 1524-1537.

Tuskan, G.A., Difazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., Putnam, N., Ralph, S., Rombauts, S., Salamov, A., Schein, J., Sterck, L., Aerts, A., Bhalerao, R.R., Bhalerao, R.P., Blaudez, D., Boerjan, W., Brun, A., Brunner, A., Busov, V., Campbell, M., Carlson, J., Chalot, M., Chapman, J., Chen, G.L., Cooper, D., Coutinho, P.M., Couturier, J., Covert, S., Cronk, Q., Cunningham, R., Davis, J., Degroeve, S., Dejardin, A., Depamphilis, C., Detter, J., Dirks, B., Dubchak, I., Duplessis, S., Ehlting, J., Ellis, B., Gendler, K., Goodstein, D., Gribskov, M., Grimwood, J., Groover, A., Gunter, L., Hamberger, B., Heinze, B., Helariutta, Y., Henrissat, B., Holligan, D., Holt, R., Huang, W., Islam-Faridi, N., Jones, S., Jones-Rhoades, M., Jorgensen, R., Joshi, C., Kangasjarvi, J., Karlsson, J., Kelleher, C., Kirkpatrick, R., Kirst, M., Kohler, A., Kalluri, U., Larimer, F., Leebens-Mack, J., Leple, J.C., Locascio, P., Lou, Y., Lucas, S., Martin, F., Montanini, B., Napoli, C., Nelson, D.R., Nelson, C., Nieminen, K., Nilsson, O., Pereda, V., Peter, G., Philippe, R., Pilate, G., Poliakov, A., Razumovskaya, J., Richardson, P., Rinaldi, C., Ritland, K., Rouze, P., Ryaboy, D., Schmutz, J., Schrader, J., Segerman, B., Shin, H., Siddiqui, A., Sterky, F., Terry, A., Tsai, C.J., Uberbacher, E., Unneberg, P., Vahala, J., Wall, K., Wessler, S., Yang, G., Yin, T., Douglas, C., Marra, M., Sandberg, G., Van de Peer, Y. and Rokhsar, D. (2006) The genome of black cottonwood, Populus trichocarpa (Torr. & Gray). *Science*, **313**, 1596-1604.

Van Baarlen, P., de Jong, J.H. and van Dijk, P.J. (2002) Comparative cyto-embryological investigations of sexual and apomictic dandelions (Taraxacum) and their apomictic hybrids *Sex Plant Reprod*, **15**, 31–38.

Van Dijk, P.J., Tas, I.C., Falque, M. and Bakx-Schotman, T. (1999) Crosses between sexual and apomictic dandelions (Taraxacum). II. The breakdown of apomixis. *Heredity*, **83 (Pt 6)**, 715-721.

Van Dijk, P.J., van Baarlen, P. and de Jong, H. (2003) The occurrence of phenotypically complementary apomixis-recombinants in crosses between sexual and apomictic dandelions (Taraxacum officinale). *Sex Plant Reprod*, **16**.

Van Dijk, P.J. and Bakx-Schotman, J.M. (2004) Formation of unreduced megaspores (diplospory) in apomictic dandelions (Taraxacum officinale, s.l.) is controlled by a sex-specific dominant locus. *Genetics*, **166**, 483-492.

Van Dijk P.J. and Vijverberg, K. (2005) *The significance of apomixis in the evolution of the angiosperms: a reappraisal. In Plant species-level systematics: new perspectives on pattern & process* Ruggell, Lichtenstein: Gantner Verlag.

Vielle-Calzada, J.P., Nuccio, M.L., Budiman, M.A., Thomas, T.L., Burson, B.L., Hussey, M.A. and Wing, R.A. (1996) Comparative gene expression in sexual and apomictic ovaries of Pennisetum ciliare (L.) Link. *Plant Mol Biol*, **32**, 1085-1092.

Vierstra, R.D. (2003) The ubiquitin/26S proteasome pathway, the complex last chapter in the life of many plant proteins. *Trends Plant Sci*, **8**, 135-142.

Vinkenoog, R., Spielman, M., Adams, S., Fischer, R.L., Dickinson, H.G. and Scott, R.J. (2000) Hypomethylation promotes autonomous endosperm development and rescues postfertilization lethality in fie mutants. *Plant Cell*, **12**, 2271-2282.

West, M. and Harada, J.J. (1993) Embryogenesis in Higher Plants: An Overview. *Plant Cell*, **5**, 1361-1369.

Whitson, T.D., Burrill, L.C., Dewey, S.A., Cudney, D.W., Nelson, B.E., Lee, R.D. and Parker, R. (2000) *Weeds of the West* Laramie, Wyoming.: The Western Society of Weed Science in cooperation with the Western United States Land Grant Universities, Cooperative Extension Services.

Woo, H.R., Chung, K.M., Park, J.H., Oh, S.A., Ahn, T., Hong, S.H., Jang, S.K. and Nam, H.G. (2001) ORE9, an F-box protein that regulates leaf senescence in Arabidopsis. *Plant Cell*, **13**, 1779-1790.

Worrall, D., Hird, D.L., Hodge, R., Paul, W., Draper, J. and Scott, R. (1992) Premature dissolution of the microsporocyte callose wall causes male sterility in transgenic tobacco. *Plant Cell*, **4**, 759-771.

Xiao, W., Custard, K.D., Brown, R.C., Lemmon, B.E., Harada, J.J., Goldberg, R.B. and Fischer, R.L. (2006) DNA methylation is critical for Arabidopsis embryogenesis and seed viability. *Plant Cell*, **18**, 805-814.

Xie, D.X., Feys, B.F., James, S., Nieto-Rostro, M. and Turner, J.G. (1998) COI1: an Arabidopsis gene required for jasmonate-regulated defense and fertility. *Science*, **280**, 1091-1094.

Yamada-Akiyama, H., Akiyama, Y., Ebina, M., Xu, Q., Tsuruta, S., Yazaki, J., Kishimoto, N., Kikuchi, S., Takahara, M., Takamizo, T., Sugita, S. and Nakagawa, H. (2009) Analysis of expressed sequence tags in apomictic guineagrass (Panicum maximum). *J Plant Physiol*, **166**, 750-761.

Zuo, J., Niu, Q.W., Frugis, G. and Chua, N.H. (2002) The WUSCHEL gene promotes vegetative-to-embryonic transition in Arabidopsis. *Plant J*, **30**, 349-359.

Curriculum Vitae

Personal Data

Name	**Anna Schallau,** *nee* Sutova
Address	Leibniz Institute of Plant Genetics and Crop Plant Research Corrensstrasse 3, 06466 Gatersleben, Germany
E-mail	schalla@ipk-gatersleben.de
Telephone	0394825440
Citizenship	Republic of Moldova
Marital status	Married
Date of birth	24.02.1982
Languages	English, German, Rumanian, Russian

Employment History/ Research Experience

2006 – present **Institute of Plant Genetics and Crop Plant Research,**
Gatersleben, Germany**.**
PhD student, Gene Regulation Group, Head Dr Habil H. Baumlein
Indentification and genetic analysis of the APOSPORY locus in Hypericum perforatum L.

2005 – 2006 **Institute of Plant Genetics and Crop Plant Research,**
Gatersleben, Germany
Visiting scientist (DAAD Scholarship), Gene Regulation Group, Head Dr Habil H. Baumlein
Evolutionary conservation of tissue specific gene regulation between ferns and higher plants

2004 – 2005 **Institute of Genetics,** Academy of Sciences of Moldova
Junior research scientist, Lab. of Genome Instability and Gene Engineering
Characterization of Mutator transposon family in Vicia sp.

2002 – 2004 **State University of Moldova**
Diploma work in the Lab. of Chemistry of Protein
Limited proteolisys of seed storage proteins of Vicia sativa

Educational Background

2004 – 2005 **Diploma of Master of Science**
Specialization in Molecular Biology, State University of Moldova

2000 – 2004 **Diploma of Licentiate in Biology** (equal to Bachelor degree)
Specialization in Biochemistry, State University of Moldova

1989 – 2000 **School Leaving Sertificat**
Specialization chemistry and biology, Lyceum D. Cantemir

Journal Publications

SCHALLAU A., KAKHOVSKAYA I., TEWES A., CZIHAL A., TIEDEMANN J., MOHR M., GROSSE I., MANTEUFFEL R. and BÄUMLEIN H. (2008) Phylogenetic footprints in fern spore- and seed-specific gene promoters. *Plant Journal,* V. 53, 3, p414-424.

IVANOV R., TIEDEMANN J., CZIHAL A., **SCHALLAU A.**, LE HONG D., MOCK H., CLAUS B., TEWES A. and BÄUMLEIN H. (2008) EFFECTOR OF TRANSCRIPTION2 is involved in xylem differentiation and includes a functional DNA single strand cutting domain. *Developmental Biology*, V. 313, 1 p 93-106.

SCHALLAU A., ARZENTON F., JOHNSTON A., HÄHNEL U., KOSZEGI D., BLATTNER F., ALTSCHMIED L., HABERER G., BARCACCIA G., and BAUMLEIN H. Identification and genetic analysis of the APOSPORY locus in *Hypericum perforatum L*. *Plant Journal*, V.62, 5, p773-784.

GALLA G., BARCACCIA G., **SCHALLAU A.**, PUENTE MOLINS M., BAUMLEIN H., and SHARBEL T.F. The cytohistological basis of apospory in *Hypericum perforatum* L. Sex *Plant Reprod*., DOI 10.1007/s00497-010-0147-7.

Oral presentations

SCHALLAU, A., ARZENTON F., HÄHNEL U., KOSZEGI D., MATZK F., BARCACCIA G., BLATTNER F.R., ALTSCHMIED L., SHARBEL T.F., H. BÄUMLEIN. Indentification of the apospory-locus in *Hypericum perforatum*. – 2nd Annual Hypericum meeting, IPK Gatersleben, 24.11.2008

SCHALLAU, A., ARZENTON F., HÄHNEL U., KOSZEGI D., MATZK F., BARCACCIA G., BLATTNER F.R., ALTSCHMIED L., SHARBEL T.F., H. BÄUMLEIN. Indentification of the apospory-locus in *Hypericum perforatum*. – 4th Plant Science Student Conference Halle/S., 23. -26.06.09

SCHALLAU, A., HÄHNEL U., ARZENTON F., KOSZEGI D., MATZK F., BARCACCIA G., BLATTNER F.R., ALTSCHMIED L., SHARBEL T.F., H. BÄUMLEIN. Marker-assisted selection of apospory in *Hypericum performatum* L. – 4th Plant Science Student Conference, IPK Gatersleben, 01.-04.07.2008

SUTOVA A., KAKHOVSKAYA I., TEWES A., CZIHAL A., TIEDEMANN J., MOHR M., GROSSE I., MANTEUFFEL R. AND BÄUMLEIN H. Evolutionary conservation of tissue specific gene regulation between ferns and higher plants.-12th European meeting of PhD students in Evolutionary Biology, St. Andrews/UK, 04.-09.09.2006

Posters

SCHALLAU, A., F. ARZENTON, U. HÄHNEL, D. KOSZEGI, F.R. BLATTNER, L. ALTSCHMIED, F. MATZK, G. BARCACCIA, T.F. SHARBEL & H. BÄUMLEIN: Indentification of the apospory-locus in *Hypericum perforatum L*. – Institutstag IPK Gatersleben, 12.-14.10.2009.

SCHALLAU, A., F. ARZENTON, U. HÄHNEL, D. KOSZEGI, F.R. BLATTNER, L. ALTSCHMIED, F.MATZK, G. BARCACCIA, T.F. SHARBEL & H. BÄUMLEIN. Indentification of the apospory-locus in *Hypericum perforatum L*. – 2nd International PhD School on Plant Development Retzbach, 7.-9.10. 2009

SCHALLAU, A., F. ARZENTON, U. HÄHNEL, D. KOSZEGI, F.R. BLATTNER, L. ALTSCHMIED, F. MATZK, G. BARCACCIA, T.F. SHARBEL & H. BÄUMLEIN. A putative apospory marker in *Hypericum perforatum L*. – Institutstag IPK Gatersleben, 29.-30.09.2008.

SCHALLAU A., HÄHNEL U., KOSZEGI D., ARCENTON F., BLATTNER L., ALTSCHMIED L., MATZK F., SHARBEL T.F. , BÄUMLEIN H. A putative apospory marker in *Hypericum perforatum L*.-3rd Plant Science Student Conference, IPB, Halle/S., 05.-08.06.07.

SCHALLAU A., HÄHNEL U., KOSZEGI D., ARCENTON F., BLATTNER L., ALTSCHMIED L., MATZK F., SHARBEL T.F. , BÄUMLEIN H. A putative apospory marker in *Hypericum perforatum L*.-3rd International Conference on Apomixis, Wernigerode, 27.06.-01.07.2007.

PUENTE MOLINS M., CORRAL GARCIA J.M., ALIYU O.M., GALLA G., BARCACCIA G., SCHALLAU A., BÄUMLEIN H., MARON J., SHARBEL T.F. Biogeographic variation for apomixis expression in St. John's wort (*Hypericum perforatum L*.).- 6th Plant Genomics European Meeting, Tenerife/Spain, 03. -06.10.2007

PUENTE MOLINS M., CORRAL GARCIA J.M., ALIYU O.M., GALLA G., BARCACCIA G., SCHALLAU A., BÄUMLEIN H., MARON J., SHARBEL T.F. Biogeographic variation for apomixis expression in St. John's wort (*Hypericum perforatum L*.).- Institutstag IPK, Gatersleben, 22. -23.10.2007

Acknowledgements

This work was done in the Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) in Gatersleben. I would like to thank all the co-workers of the Institute who helped in this work, especially from the groups of Gene Regulation, Apomixis, Expression Mapping and Experimental Taxonomy.

I am very grateful to my supervisor Helmut Baumlein for giving me the opportunity to work on this project, for guiding me and for encouraging my ideas. I would also like to thank the people who suppoted me with ideas and discussions: Tim Sharbel, Lothar Altschmied, Frank Blattner, David Koszegi, Konrad Bachman and Amal Johnston.

I am very much obliged to Fritz Matzk who performed the genetic crosses used for this study. Thanks to Tim Sharbel for shearing with his collection of natural apomictic accessions and Marcus Koch for giving me the rare diploid accessions.

Since this project required a lot of techniques that were not available in our group, this work would not be possible without tight cooperation within the IPK. I would like to express my deep gratitude to Inge Matties, Anja Hanemann and Marion Röder for introducing me into the pyrosequencing, Urs Hähnel for helping dealing with the BAC assembly and managing the pipeting robots, Lothar Altschmied for shearing PEARL scripts for the sequence analysis, Daniela Schulte and Niels Stein for helping me with the fingerprinting analysis.

I am very much exited with the fact that such an interesting plant as *Hypericum* deserves attention from many researches. I would like to acknowledge the people who also work with this plant for sharing the accessions and experience: Francesco Arcenton who was involved in the AFLP analysis, Marta Puente Molins and Leane Börner who performed the FCSS of plants belonging to wild populations, Jose Garcia Maria Coral and Nicolai Nürk. Many very warm and deep thanks to Gulio Galla who participated in the cytological analysis of the Hypericum ovules.

Special thanks to Andreas Czihal, for helping in hybridization experiments and general advice about molecular work.

I would like to thank for the excellent technical support Elke Liemann, Monika Gottowick, Anett Busching, Sabine Skeibe, Ingelore Dommes and Leane Börner. I would like also to acknowledge the gardeners for taking care over my plants.

Acknowledgements

I would also like to thank our external cooperation partners: Gianni Barcaccia (University of Padua), who was involved in the generation of the AFLP marker and Georg Haberer (MIPS, Munich) who performed the annotation of the sequenced BAC.

I would like to acknowledge all members of Gene Regulation Group for a nice working atmosphere: David Koszegi, Amal Johnston, Astrid Junker, Corina Gryczka, Maria Mildner and Diep Le Hong.

Special thanks to Lothar Altschmied for the critical reading of the manuscript.

I would also like to express my gratitude to all my friends within IPK, especialy to Astrid and Bjorn, Anja and Jose, Kasia and Young Min, Marta and Thomas, Marie and Marco, Miroslava and Volodja and many others. We shared a very good time together. Special thanks to all the members of the IPK Student Board.

Last, but not least I would like to thank my parents, who encouraged me to be a scientist and my husband Kai for the endless help and support.

ERKLÄRUNGEN

GEMÄß § 5 ABS. 2 DER PROMOTIONSORDNUNG VOM 17.09.1998

Ich erkläre hiermit, daß ich mich mit der vorliegenden wissenschaftlichen Arbeit erstmals um die Erlangung des Doktorgrades bewerbe. Ich habe die Arbeit selbständig und ohne fremde Hilfe verfaßt, und nur die von mir angegebenen Quellen und Hilfsmittel benützt. Wörtlich oder inhaltlich entnommene aus anderen Werken habe ich Stellen als solche kenntlich gemacht.

Gatersleben, Januar 2010

Schallau, Anna

Supplement

Plant Identifier	Apospory %	Parthenogenesis %	Marker state
Tetraploids			
4(sF$_1$1 x aAn)1/1	0	0	S
4(sF$_1$1 x aAn)1/2	0	0	S
4(sF$_1$1 x aAn)1/3	100	42	A
4(sF$_1$1 x aAn)1/4	0	0	S
4(sF$_1$1 x aAn)1/5	0	0	S
4(sF$_1$1 x aAn)1/6	0	0	S
4(sF$_1$1 x aAn)1/7	100	24	A
4(sF$_1$1 x aAn)1/8	0	0	S
4(sF$_1$1 x aSi)1/1	38	0 #	A
4(sF$_1$1 x aSi)1/2	0	0	S
4(sF$_1$1 x aSi)1/3	0	0	S
4(sF$_1$1 x aSi)1/4	0	0	S
4(sF$_1$1 x aNo)1/1	28	4	A
4(sF$_1$1 x aNo)1/2	0	0	S
4(sF$_1$1 x aNo)1/3	100	42	A
4(sF$_1$1 x aNo)1/4	0	0	S
4(sF$_1$1 x aNo)1/5	100	10	A
4(sF$_1$1 x aNo)1/6	100	30	A
4(sF$_1$1 x aNo)1/7	99	39	A
4(sF$_1$1 x aNo)1/8	100	18	A
4(sF$_1$1 x aNo)1/9	71	3	A
4(sF$_1$1 x aNo)1/10	100	31	A
4(sF$_1$2 x aAn)1/1	92	77	A
4(sF$_1$2 x aAn)1/2	92	42	A
4(sF$_1$2 x aAn)1/3	0	0	S
4(sF$_1$2 x aAn)1/4	0	0	S
4(sF$_1$2 x aAn)1/5	0	0	S
4(sF$_1$2 x aAn)1/6	0	0	S
4(sF$_1$2 x aAn)1/7	0	0	S
4(sF$_1$2 x aAn)1/8	0	6	S
4(sF$_1$2 x aAn)1/9	0	0	S
4(sF12 x aAn)1/10	100	24	A
4(sF$_1$2 x aNo)1a/1	99	46	A
4(sF$_1$2 x aNo)1a/2	0	0	S
4(sF$_1$2 x aNo)1a/3	0	0	S
4(sF$_1$2 x aNo)1a/4	0	0	S
4(sF$_1$2 x aNo)1a/5	0	0	S
4(sF$_1$2 x aNo)1a/6	0	0	S
4(sF$_1$2 x aNo)1a/7	98	53	A
4(sF$_1$2 x aNo)1a/8	96	0#	A
4(sF$_1$2 x aNo)1a/9	0	0	S
4(sF$_1$2 x aNo)1a/10	0	0	S
4(sF$_1$2 x aNo)1a/11	96	0#	A
4(sF$_1$2 x aNo)1a/12	29	0#	A
4(sF$_1$2 x aNo)1a/13	0	0	S
4(sF$_1$2 x aNo)1a/14	100	4	A

4(sF$_1$2 x aNo)1a/15	96	4	A
4(sF$_1$2 x aNo)1a/16	88	8	A
4(sF$_1$2 x aNo)1a/17	100	0#	A
4(sF$_1$2 x aNo)1a/20	100	4	A
4(sF$_1$2 x aNo)1a/24	100	0#	A
4(sF$_1$2 x aNo)1a/25	92	3	A
4(sF$_1$2 x aNo)1a/29	100	24	A
4(sF$_1$2 x aNo)1a/31	100	0#	A
4(sF$_1$2 x aNo)1a/33	100	4	A
4(sF$_1$2 x aNo)1a/43	20	4	A
4(R1C2 x aSi)1a/1	29	0#	A
4(R1C2 x aSi)1a/2	0	0	S
4(R1C2 x aSi)1a/3	0	0	S
4(R1C2 x aSi)1a/4	0	0	S
4(R1C2 x aSi)1c/2	96	33	A
4(R1C2 x aSi)1c/3	24	0#	A
4(R1C2 x aSi)1c/7	0	0	S
4(R1C2 x aSi)1c/8	0	0	S
4(R1C2 x aSi)1c/9	0	0	S
4(R1C2 x aSi)1d/2	0	0	S
4(R1C2 x aSi)1d/4	0	0	S
4(R1C2 x aSi)1d/5	0	0	S
4(R1C2 x aSi)1d/6	0	0	S
4(R1C2 x aSi)1d/7	0	0	S
4(R1C2 x aSi)1d/8	0	0	S
4(R1C2 x aSi)1d/9	31	18	A
Triploids			
1(sR1 x aNo) 1/1	99	0#	A
1(sR1 x aNo) 1/2	99	4	A
1(sR1 x aNo) 1/3	0	0	S
1(sR1 x aNo) 1/4	100	3	A
1(sR1 x aNo) 1/5	98	2	A
1(sR1 x aNo) 1/6	100	3	A
1(sR1 x aNo) 2/1	0	0	S
1(sR1 x aNo) 2/2	0	0	S
1(sR1 x aNo) 2/3	100	4	A
1(sR1 x aNo) 2/4	95	0#	A
1(sR1 x aNo) 3/5	0	0	S
2(sR1 x aNo)1/2	97	13	A
2(sR1 x aNo)2a/3	100	0#	A
2(sR1 x aNo)2a/2	100	0#	A
2(sR1 x aNo)2a/7	100	33	A
2(sR1 x aNo)2a/14	100	0#	A
1(sV1 x aNo) 1/1	99	8	A
1(sV1 x aNo) 1/2	94	8	A
1(sV1 x aNo) 1/3	100	0#	A
1(sV1 x aNo) 1/4	100	0#	A
1(sV1 x aNo) 1/5	90	5	A
1(sV1 x aNo) 1/6	100	3	A
1(sV1 x aNo) 1/7	0	0	S
1(sV1 x aNo) 1/8	0	0	S
1(sV1 x aNo) 1/10	0	15	S
1(sV1 x aNo) 1/12	0	0	S

1(sV1 x aNo)1/13	100	3	A
1(sV1 x aNo) 1/17	100	7	A
1(sV1 x aNo) 1/19	100	7	A
1(sV1 x aNo) 1/20	95	7	A
1(sV1 x aNo) 1/22	100	11	A
1(sV1 x aNo) 1/23	0	0	S
1(sV1 x aNo) 1/25	100	3	A
1(sV1 x aNo) 1/27	100	0#	A
1(sV1 x aNo) 1/28	0	0	S
1(sV1 x aNo) 1/29	100	14	A
1(sV1 x aNo) 2a/1	100	0#	A
1(sV1 x aNo) 2b/2	100	7	A
1(sV1 x aNo) 2b/3	100	0#	A
1(sV1 x aNo) 2b/4	0	0	S
1(sV1 x aNo) 2b/5	100	0#	A
1(sV3 x aNo) 1/1	75	66	A
1(sV3 x aNo) 1/2	100	0#	A
1(sV3 x aNo) 1/4	98	3	A
1(sV3 x aNo) 1/6	100	12	A
2(sP2 x aTo)1/1	0	0	S
2(sP2 x aTo)1/2	100	0#	A
2(sP2 x aTo)1/11	100	66	A
2(sP2 x aTo)1/13	98	8	A
2(sP2 x aTo)1/14	100	3	A
2(sP2 x aTo)1/15	0	0	S
2(sP2 x aTo)1/16	95	14	A
2(sP2 x aTo)1/18	100	6	A
2(sP2 x aTo)1/19	95	0#	A
2(sP2 x aTo)1/20	0	0	S
2(sP2 x aTo)1/21	100	0#	A
2(sP2 x aTo)1/34	0	0	S
2(sP2 x aTo)1/26	100	5	A
2(sP2 x aTo)1/23	100	0#	A
Diploids			
2(sP1 x sV1)1a/3	0	0	S
2(sP1 x sV1)1a/4	0	0	S
2(sP1 x sV1)1a/6	0	0	S
2(sP1 x sV1)1a/7	0	0	S
2(sP1 x sV1)1a/8	0	0	S
2(sP1 x sV1)1a/9	0	0	S
2(sP1 x sV1)1a/11	0	0	S
2(sP1 x sV1)1a/13	0	0	S
2(sP1 x sV1)1b/1	0	0	S
2(sP1 x sV1)1b/6	0	0	S
2(sP1 x sV1)1b/7	0	0	S
2(sP1 x sV1)1b/8	0	0	S
2(sP1 x sV1)1b/11	0	0	S
2(sR1 x sP2)1/3	0	0	S
2(sR1 x sP2)1/8	0	0	S
2(sR1 x sP2)1/9	0	0	S
2(sR1 x sP2)1/13	0	0	S
2(sR1 x sP2)1/16	0	0	S
2(sR1 x sP2)2/2	0	0	S

2(sR1 x sP2)2/3	0	0	S
2(sR1 x sP2)2/4	0	0	S
2(sR1 x sP2)2/8	0	0	S
2(sR1 x sP2)2/18	0	0	S
2(sR1 x sP2)2/25	7	0	S

Table 15: Flow cytometric seed screen and marker analysis in tetraploid, triploid and diploid offspring of directed crosses

Plant identifier, percent apospory, percent of parthenogenesis and marker state are given.

Plant	Apospory	Partheno-	SNP06		SNP24	
Identifier	%	genesis%	A	C	A	T
4($sF_1$1 x aAn)1/1	0	0	0.80%	99.20%	2.50%	97.50%
4($sF_1$1 x aAn)1/2	0	0	0.90%	99.10%	2.50%	97.50%
4($sF_1$1 x aAn)1/3	100	42	23.00%	77.00%	24.90%	75.10%
4($sF_1$1 x aAn)1/4	0	0	0.00%	100.00%	4.10%	95.90%
4($sF_1$1 x aAn)1/5	0	0	1.20%	98.80%	0.00%	100.00%
4($sF_1$1 x aAn)1/6	0	0	0.70%	99.30%	1.90%	98.10%
4($sF_1$1 x aAn)1/7	100	24	25.00%	75.00%	20.10%	79.90%
4($sF_1$1 x aAn)1/8	0	0	0.60%	99.40%	0.00%	100.00%
4($sF_1$1 x aSi)1/1	38	0	19.10%	80.90%	13.90%	86.10%
4($sF_1$1 x aSi)1/2	0	0	0.80%	99.20%	0.00%	100.00%
4($sF_1$1 x aSi)1/3	0	0	1.10%	98.90%	2.30%	97.70%
4($sF_1$1 x aSi)1/4	0	0	0.00%	100.00%	8.40%	91.60%
4($sF_1$1 x aNo)1/1	28	4	21.70%	78.30%	19.30%	80.70%
4($sF_1$1 x aNo)1/2	0	0	0.70%	99.30%	0.00%	100.00%
4($sF_1$1 x aNo)1/3	100	42	19.80%	80.20%	14.40%	85.60%
4($sF_1$1 x aNo)1/4	0	0	0.80%	99.20%	3.60%	96.40%
4($sF_1$1 x aNo)1/5	100	10	20.00%	80.00%	20.70%	79.30%
4($sF_1$1 x aNo)1/6	100	30	24.50%	75.50%	20.70%	79.30%
4($sF_1$1 x aNo)1/7	99	39	22.70%	77.30%	17.20%	82.80%
4($sF_1$1 x aNo)1/8	100	18	20.60%	79.40%	18.20%	81.80%
4($sF_1$1 x aNo)1/9	71	3	21.80%	78.20%	17.30%	82.10%
4($sF_1$1 x aNo)1/10	100	31	24.50%	75.50%	22.00%	78.00%
4($sF_1$2 x aAn)1/1	92	77	11.80%	88.20%	16.00%	84.00%
4($sF_1$2 x aAn)1/2	92	42	16.20%	83.80%	18.70%	81.30%
4($sF_1$2 x aAn)1/3	0	0	3.00%	97.00%	5.80%	94.20%
4($sF_1$2 x aAn)1/4	0	0	0.90%	99.10%	4.40%	95.60%
4($sF_1$2 x aAn)1/5	0	0	4.60%	95.40%	4.70%	95.30%
4($sF_1$2 x aAn)1/6	0	0	0.60%	99.40%	4.50%	95.50%
4($sF_1$2 x aAn)1/7	0	0	1.20%	98.80%	0.10%	99.90%
4($sF_1$2 x aAn)1/8	0	6	0.60%	99.40%	0.00%	100.00%
4($sF_1$2 x aAn)1/9	0	0	0.70%	99.30%	0.00%	100.00%
4(sF12 x aAn)1/10	100	24	23.90%	76.10%	21.60%	78.40%
4($sF_1$2 x aNo)1a/1	99	46	20.70%	79.30%	22.20%	77.80%
4($sF_1$2 x aNo)1a/2	0	0	2.70%	97.30%	5.20%	94.80%
4($sF_1$2 x aNo)1a/3	0	0	0.00%	100.00%	1.00%	99.00%

4(sF$_1$2 x aNo)1a/4	0	0	0.00%	100.00%	1.10%	98.90%
4(sF$_1$2 x aNo)1a/5	0	0	2.50%	97.50%	7.00%	93.00%
4(sF$_1$2 x aNo)1a/6	0	0	0.60%	99.40%	1.10%	98.90%
4(sF$_1$2 x aNo)1a/7	98	53	25.80%	74.20%	27.10%	72.90%
4(sF$_1$2 x aNo)1a/8	96	0	20.50%	79.50%	22.30%	77.70%
4(sF$_1$2 x aNo)1a/9	0	0	0.00%	100.00%	2.60%	97.40%
4(sF$_1$2 x aNo)1a/10	0	0	0.90%	99.10%	0.00%	100.00%
4(sF$_1$2 x aNo)1a/11	96	0	21.80%	78.20%	18.30%	81.70%
4(sF$_1$2 x aNo)1a/12	29	0	24.90%	75.10%	18.60%	81.40%
4(sF$_1$2 x aNo)1a/13	0	0	2.00%	98.00%	2.30%	97.70%
4(sF$_1$2 x aNo)1a/14	100	4	24.60%	75.40%	24.40%	75.60%
4(sF$_1$2 x aNo)1a/15	96	4	19.90%	80.10%	17.60%	82.40%
4(sF$_1$2 x aNo)1a/16	88	8	21.30%	78.70%	16.60%	83.60%
4(sF$_1$2 x aNo)1a/17	100	0	16.60%	83.40%	14.50%	85.50%
4(sF$_1$2 x aNo)1a/20	100	4	19.10%	80.90%	22.60%	77.40%
4(sF$_1$2 x aNo)1a/24	100	0	18.90%	81.10%	15.30%	84.70%
4(sF$_1$2 x aNo)1a/25	92	3	21.30%	78.70%	16.50%	83.50%
4(sF$_1$2 x aNo)1a/29	100	24	20.90%	79.10%	14.50%	85.50%
4(sF$_1$2 x aNo)1a/31	100	0	19.10%	80.90%	17.30%	82.70%
4(sF$_1$2 x aNo)1a/33	100	4	17.50%	82.50%	12.60%	87.40%
4(sF$_1$2 x aNo)1a/43	20	4	22.60%	77.40%	16.60%	83.40%
4(R1C2 x aSi)1a/1	29	0	24.40%	75.60%	15.90%	84.10%
4(R1C2 x aSi)1a/2	0	0	0.70%	99.30%	0.00%	100.00%
4(R1C2 x aSi)1a/3	0	0	0.00%	100.00%	1.00%	99.00%
4(R1C2 x aSi)1a/4	0	0	1.00%	99.00%	0.10%	99.90%
4(R1C2 x aSi)1c/2	96	33	22.50%	77.50%	20.60%	79.40%
4(R1C2 x aSi)1c/3	24	0	20.10%	79.90%	21.00%	79.00%
4(R1C2 x aSi)1c/7	0	0	3.10%	96.90%	0.00%	100.00%
4(R1C2 x aSi)1c/8	0	0	1.30%	98.70%	0.00%	100.00%
4(R1C2 x aSi)1c/9	0	0	1.20%	98.80%	0.00%	100.00%
4(R1C2 x aSi)1d/2	0	0	0.00%	100.00%	3.80%	96.20%
4(R1C2 x aSi)1d/4	0	0	1.20%	98.80%	0.00%	100.00%
4(R1C2 x aSi)1d/5	0	0	1.20%	98.80%	8.00%	92.00%
4(R1C2 x aSi)1d/6	0	0	0.80%	99.20%	0.00%	100.00%
4(R1C2 x aSi)1d/7	0	0	0.80%	99.20%	0.00%	100.00%
4(R1C2 x aSi)1d/8	0	0	0.70%	99.30%	3.50%	96.50%
4(R1C2 x aSi)1d/9	31	18	23.10%	76.90%	24.90%	75.10%

Table 16: Pyrosequencing analysis of the tetraploid population of directed crosses

Plant identifier, percent apospory, percent of parthenogenesis and quantification of allele A and C for SNP 06 and alleles A and T and SNP 24 are given. The sexuals are marked in light grey and the apomicts in dark grey respectively.

A

Plant Identifier	Origin	Apospory %	Partheno-genesis %	SNP 06		SNP 24	
				A	C	A	T
H06-1358	Tuscola IL US 39,8 88,28	87.5	68.8	19.90%	80.10%	23.60%	76.40%
H06-1359	Tuscola IL US 39,8 88,28	90	85	24.50%	75.50%	30.20%	69.80%
H06-1362	Tuscola IL US 39,8 88,28	89.5	89.5	21.20%	78.80%	27.70%	72.30%
H06-1363	Tuscola IL US 39,8 88,28	100	93.3	19.50%	80.50%	22.60%	77.40%
H06-1366	Tuscola IL US 39,8 88,28	95.8	87.5	21.20%	78.80%	21.70%	78.30%
H06-1367	Tuscola IL US 39,8 88,28	91.7	91.7	21.60%	78.40%	26.10%	73.90%
H06-1369	Tuscola IL US 39,8 88,28	91.3	95.7	24.80%	75.20%	30.10%	69.90%
H06-1372	Tuscola IL US 39,8 88,28	95	75	18.00%	82.00%	25.80%	74.20%
H06-1376	Tuscola IL US 39,8 88,28	95.7	86.9	24.20%	75.80%	25.80%	74.20%
H06-1379	Green Lake WI US 43,85 89,3	84.2	84.2	21.60%	78.40%	26.80%	73.20%
H06-1441	Green Lake WI US 43,85 89,3	91.7	95.8	24.80%	75.20%	32.10%	67.90%
H06-1447	Point Beach WI US 44,26 87,56	100	84.2	18.00%	82.00%	31.10%	68.90%
H06-1483	Point Beach WI US 44,26 87,56	85.7	85.7	24.20%	75.80%	29.60%	70.40%
H06-1489	Kewaunee MI US 44,61 88,11	85.7	100	24.00%	76.00%	30.00%	70.00%
H06-1563	Kewaunee MI US 44,61 88,11	82.6	69.6	20.20%	79.80%	32.60%	67.40%
H06-1569	Gillett WI US 44,84 88,61	95.2	90.5	18.20%	81.80%	20.00%	80.00%
H06-1638	Gillett WI US 44,84 88,61	95.5	90.9	21.10%	78.90%	28.30%	71.70%
H06-1643	Rideau River ON Canada 45 75, 62	95.8	91.7	24.10%	75.90%	28.20%	71.80%
H06-1650	Rideau River ON Canada 45 75, 62	100	95.5	22.60%	77.40%	28.90%	71.10%
H06-1654	Rideau River ON Canada 45 75, 62	71.4	66.7	18.10%	81.90%	27.00%	73.00%
H06-1657	Rideau River ON Canada 45 75, 62	80.9	71.4	24.90%	75.10%	28.40%	71.60%
H06-1661	Rideau River ON Canada 45 75, 62	66.7	61.9	22.50%	77.50%	27.00%	73.00%
H06-1743	Menominee MI US 45, 21 87. 75	72.7	63.6	22.60%	77.40%	32.70%	67.30%
H06-1749	Wausaukee WI US 45, 45 87, 63	86.9	82.6	24.40%	75.60%	20.10%	79.90%
H06-1807	Wausaukee WI US 45, 45 87, 63	87.5	87.5	22.80%	77.20%	26.70%	73.30%
H06-1877	Carney MI US 45,6 87,03	95.8	87.5	21.70%	78.30%	26.50%	73.50%
H06-1883	Iron Mountain MI US 45,83 88,08	100	95.2	18.20%	81.80%	33.30%	66.70%
H06-1886	Iron Mountain MI US 45, 83 88, 08	73.7	52.6	25.00%	75.00%	26.90%	73.10%
H06-1898	Iron Mountain MI US 45, 83 88, 08	86.4	81.8	17.90%	82.10%	31.50%	68.50%
H06-1960	Iron Mountain MI US 45, 83 88, 08	95.2	90.5	20.30%	79.70%	31.10%	68.90%
H06-1964	Tecumseh MI US 42 89, 66	52.6	47.4	17.00%	83.00%	25.90%	74.10%
H06-1992	Tecumseh MI US 42 89,66	85	75	22.80%	77.20%	30.60%	69.40%
H06-1994	Cazadero CA US 38,55	86.7	93.3	20.50%	79.50%	32.80%	67.20%

	123,13						
H06-1998	Cazadero CA US 38,55 123,13	84.2	94.7	19.20%	80.80%	24.40%	75.60%
H06-2006	Cazadero CA US 38, 55 123, 13	65	50	19.50%	80.50%	27.90%	72.10%
H06-2013	Cazadero CA US 38, 55 123, 13	63.6	59.1	18.90%	81.10%	24.90%	75.10%
H06-2114	Covelo CA US 39,78 123,25	75	70	20.70%	79.30%	29.60%	70.40%
H06-2118	Holiday CA US 40,61 122,3	95.5	81.8	21.50%	78.50%	27.20%	72.80%
H06-2144	Holiday CA US 40,61 122,3	91.3	91.3	11.50%	88.50%	18.20%	81.80%
H06-2148	Holiday CA US 40,61 122,3	75	58.3	18.90%	81.10%	24.70%	75.30%
H06-2160	Holiday CA US 40,61 122,3	100	86.9	20.50%	79.50%	23.80%	76.20%
H06-2167	Gasquet CA US 41,84 122,63	86.4	86.4	15.50%	84.50%	23.80%	76.20%
H06-2208	Gasquet CA US 41,84 122,63	65	60	23.40%	76.60%	27.40%	72.60%
H06-2209	Mt Shasta CA US 41,3 122,25	65.2	47.8	20.70%	79.30%	23.60%	76.40%
H06-2241	Mt Shasta CA US 41,3 122,25	95.2	85.7	27.10%	72.90%	26.50%	73.50%
H06-2245	Abrams Lake Rd CA US 41,33 122,25	95.7	78.3	29.00%	71.00%	21.40%	78.60%
H06-2279	Abrams Lake Rd CA US 41,33 122,25	86.9	78.3	17.00%	83.00%	23.70%	76.30%
H06-2280	Weed CA US 41,43 122,38	55	65	17.00%	83.00%	20.10%	79.90%
H06-2331	Weed CA US 41,43 122,38	89.5	89.5	18.10%	81.90%	26.50%	73.50%
H06-2345	Weed CA US 41,43 122,38	95	75	33.30%	66.70%	23.50%	76.50%
H06-2412	Corvallis OR US 44,61 123,2	81.8	68.2	28.00%	72.00%	28.70%	71.30%
H06-2425	Corvallis OR US 44,61 123,2	77.3	72.3	23.70%	76.30%	22.10%	77.90%
H06-2458	Corvallis OR US 44,61 123,2	95.7	86.9	18.80%	81.20%	43.70%	56.30%
H06-2464	Granera Spain 41,73 2,06	90.9	81.8	18.00%	82.00%	20.10%	79.90%
H06-2761	Bolzano Italy 46,51 12,15	81.8	86.4	19.40%	80.60%	31.90%	68.10%
H06-2941	Clapier France 44,1 7,41	90	95	14.50%	85.50%	18.00%	82.00%
H06-2943	Clapier France 44,1 7,41	86.9	78.3	19.20%	80.80%	26.50%	73.50%
H06-2957	Clapier France 44,1 7,41	100	95.5	19.90%	80.10%	28.40%	71.60%
H06-3062	Adliswil Switzerland 47,3 8,56	66.7	55.6	15.50%	84.50%	21.50%	78.50%
H06-3243	Praha Czech 50,08 14,46	30.4	56.5	19.60%	80.40%	28.40%	71.60%
H06-3299	Praha Czech 50,08 14,46	82.6	82.6	15.10%	84.90%	20.20%	79.80%
H06-3386	Suu Ravine Kyrgizstan 42.65 74.51	95.65	82.6	20.60%	79.40%	33.00%	67.00%

B

plant Identifier	Origin	Apospory,%	Partheno-	SNP 06		SNP 24	
			genesis,%	A	C	A	T
110-IT-13-5-Hyp0187	Italy	0	0	0.50%	99.50%	1.60%	98.40%
110-IT-13-5-Hyp0773	Italy	0	0	0.50%	99.50%	0.00%	100.00%
110-IT-13-5-Hyp0775	Italy	0	0	0.00%	100.00%	0.80%	99.20%
111-IT-8-5-Hyp0185	Italy	0	0	0.80%	99.20%	4.50%	95.50%
111-IT-8-5-Hyp0765	Italy	0	0	0.40%	99.60%	2.20%	97.80%
111-IT-8-5-Hyp0768	Italy	0	0	4.60%	95.40%	4.10%	95.90%
140-IT-13-5-Hyp0849	Italy	0	0	0.30%	99.70%	0.00%	100.00%
140-IT-13-5-Hyp0850	Italy	0	0	0.40%	99.60%	3.90%	96.10%
207-HU-13-5-Hyp0753	Hungary	0	0	0.40%	99.60%	0.00%	100.00%
207-HU-13-5-Hyp0754	Hungary	0	0	0.00%	100.00%	2.10%	97.90%
207-HU-13-5-Hyp0756	Hungary	0	0	0.00%	100.00%	4.90%	95.10%
210-RS-12-5-Hyp0650	Russia	0	0	0.50%	99.50%	0.50%	99.50%
210-RS-12-5-Hyp0651	Russia	0	0	0.00%	100.00%	1.80%	98.20%
210-RS-12-5-Hyp0652	Russia	0	0	0.50%	99.50%	3.90%	96.10%
218-HU-13-5-Hyp0726	Hungary	0	0	0.00%	100.00%	0.00%	100.00%
218-HU-13-5-Hyp0727	Hungary	0	0	0.00%	100.00%	7.10%	92.90%
218-HU-13-5-Hyp0728	Hungary	0	0	0.00%	100.00%	7.60%	92.40%
227-BG-13-5-Hyp0262	Bulgaria	0	0	0.00%	100.00%	0.60%	99.40%
227-BG-13-5-Hyp1411	Bulgaria	0	0	0.00%	100.00%	7.30%	92.70%
227-BG-13-5-Hyp1413	Bulgaria	0	0	0.00%	100.00%	0.00%	100.00%
227-BG-13-5-Hyp1414	Bulgaria	0	0	0.70%	99.30%	7.10%	92.90%

Table 17: Pyrosequencing analysis of the collection of wild accessions

A. Collection of apomictic populations from Europe and North America; B. Collection of sexual populations from Europe. Plant identifier, accession origin, percent apospory, percent of parthenogenesis and quantification of allele A and C for SNP 06 and alleles A and T and SNP 24 are given.

Most of the results presented in the thesis were published in

SCHALLAU A., ARZENTON F., JOHNSTON A., HÄHNEL U., KOSZEGI D., BLATTNER F., ALTSCHMIED L., HABERER G., BARCACCIA G., and BAUMLEIN H. Identification and genetic analysis of the APOSPORY locus in *Hypericum perforatum L*. *Plant Journal*, V.62, 5, p773-784.

GALLA G., BARCACCIA G., **SCHALLAU A.**, PUENTE MOLINS M., BAUMLEIN H., and SHARBEL T.F. The cytohistological basis of apospory in *Hypericum perforatum* L. Sex *Plant Reprod*., DOI 10.1007/s00497-010-0147-7.

The first paper is attached.

Identification and genetic analysis of the *APOSPORY* locus in *Hypericum perforatum* L.

Anna Schallau[1], Francesco Arzenton[2], Amal J. Johnston[1], Urs Hähnel[1], David Koszegi[1], Frank R. Blattner[1], Lothar Altschmied[1], Georg Haberer[3], Gianni Barcaccia[2] and Helmut Bäumlein[1*#]

[1]Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany

[2]Department of Environmental Agronomy and Crop Science, University of Padova, Legnaro, Italy

[3]Munich Information Center for Protein Sequences (MIPS), Institute of Bioinformatics, Neuherberg, Germany

[*]Corresponding author: H. Bäumlein (fax +49 39482 5500; baumlein@ipk-gatersleben.de)

[#] This paper is dedicated to our former colleague Dr. Fritz Matzk

SUMMARY

The introduction of apomixis -seed formation without fertilization- into crop plants is a long-held goal of breeding research, since it would allow for the ready fixation of heterozygosity. The genetic basis of apomixis, whether of the aposporous or the diplosporous type is still only poorly understood. *Hypericum perforatum* (St. John's wort), a plant with a small genome and a short generation time, can be aposporous and/or parthenogenetic, and so represents an interesting model dicot for apomixis research. Here we describe a genetic analysis which first defined and then isolated a locus (designated *HAPPY* for *Hypericum APOSPORY*) associated with apospory. AFLP profiling was used to generate a CAPS marker for *HAPPY* which co-segregated with apospory but not with parthenogenesis, showing that these two components of apomixis are independently controlled. Apospory was inherited as a dominant simplex gene at the tetraploid level. Part of the *HAPPY* sequence is homologous to the *Arabidopsis thaliana* gene *ARI7* encoding the ring finger protein ARIADNE7. This protein is predicted to be involved in various regulatory processes, including ubiquitin-mediated protein degradation. While the aposporous and sexual alleles of the *HAPPY* component *HpARI* were co-expressed in many parts of the plant, the gene product of the apomict's allele is truncated. Cloning *HpARI* represents the first step towards the full characterization of *HAPPY* and the elucidation of the molecular mechanisms underlying apomixis in *H. perforatum*.

Keywords: *Hypericum perforatum*, St. John’s wort, apomixis, apospory-linked marker, *HAPPY*, sexuality and apospory specific alleles, *ARIADNE7*

INTRODUCTION

Apomixis, a form of asexual reproduction through seed, is found in more than 400 angiosperm species belonging to about 40 families (Carman, 1997; van Dijk and Vijverberg, 2005). The genotype of the apomictic seed is identical to that of the mother plant. The engineering of apomixis into sexual crop plants has long been considered highly desirable, since it would allow for the fixation of favourable heterozygosity, and in particular, of heterosis. The economic and social benefits of such a technology are likely to exceed those of the Green Revolution (Spillane *et al.*, 2004). Achieving this goal requires a full understanding of the genetic and molecular mechanisms underlying apomixis. Apomixis occurs in two distinct forms - sporophytic (or adventitious embryony with autonomous embryo development in sporophytic tissues) and gametophytic (where a meiotically unreduced, non-fertilized egg develops into an embyo). The gametophytic type is recognized to include two subtypes, diplospory and apospory. In the former, the progenitor is the megaspore mother cell (MMC), which bypasses meiosis to produce an unreduced embryo sac; while in the latter embryo sacs develop from cell(s) adjacent to the MMC. In both cases, the development of a parthenogenetic embryo proceeds from an unreduced embryo sac, but endosperm development can be either autonomous, or if fertilization of the polar nucleus is required, pseudogamous (for reviews see Nogler, 1984; Koltunow, 1993; Savidan *et al*., 2001; Koltunow and Grossniklaus; 2003, Ozias-Akins, 2006; Hörandl *et al.*, 2007). Apomictic species which maintain a balance between the genetic stability ensured by apomixis and the flexibility offered by segregation and/or recombination are referred to “facultative apomicts”.

Despite many years of apomixis research, the genetic control of apomixis remains in its infancy. Examples are known where apomixis is determined by the action of a single dominant gene (Savidan, 1980; Leblanc *et al*., 1995; Bicknell *et al*., 2000), but in others, the pattern of inheritance is more complex. Apospory in *Pennisetum* and *Paspalum* spp. is associated with the presence of a non-recombining region of supernumerary chromatin (Ozias-Akins *et al*., 1998; Roche *et al*., 2001; Labombarda *et al*., 2002; Conner *et al*., 2008). Diplospory in the *Boechera holboelli* complex involves homoeologous chromosome substitutions (Kantama *et al*., 2007). A growing body of evidence supports the notion that some of the components of apomixis (in particular, apospory and parthenogenesis) are independently, but rather simply inherited (Noyes and Rieseberg, 2000; Barcaccia *et al*., 2000; Albertini *et al*., 2001; Matzk *et al*., 2001; Van Baarlen *et al*., 2002; van Dijk and Bakx-Schotman, 2004; Catanach *et al*., 2006; Noyes, 2006). However, some of these components

can also be under polygenic control (Matzk *et al*., 2005). Apomixis has evolved independently a number of times (Holsinger, 2000). It represents a short circuiting of the sexual reproduction pathway, presumably *via* the omission or deregulation, as a result of either mutation or *de novo* gene evolution (Tucker *et al.*, 2003; Albertini *et al*., 2005; Ozias-Akins and van Dijk, 2007; Sharbel *et al*., 2009) As proposed by Nogler (1984), apomixis may be triggered by reproduction-specific gene expression activated at the wrong time and/or place (reviewed in Koltunow and Grossniklaus, 2003). Several mutants of sexual species display developmental heterochrony and apomixis-like meiotic non-reduction, parthenogenesis and autonomous endosperm formation (Huang and Sheridan, 1996; Ohad *et al*., 1996; Chaudhury *et al*., 1997; Guitton *et al.*, 2004; Barrell and Grossniklaus, 2006, Ravi *et al*., 2008; d´Erfurth *et al*., 2009). Thus the expectation is that the expression and ultimate function of the genes critical for sexual development are perturbed in apomicts, although whether such changes are primary or causal effects has yet to be established.

Apomixis is often associated with extensive heterozygosity and polyploidy, although some diploid apomicts have been described (Nogler, 1982; Roy, 1995; Bicknell, 1997; Kojima and Nagato, 1997). Whether the characteristically high ploidy levels are the cause or the consequence of apomixis remains unclear (Koltunow and Grossniklaus, 2003). Chromosomal regions associated with some apomixis factors have been identified in several species, and molecular markers for diplospory and parthenogenesis have been identified (Barcaccia *et al*., 1998; Noyes and Rieseberg, 2000). Recombination in the region of the loci linked to apomeiosis tends to be strongly suppressed, although this was not the case for either *Taraxacum* (van Dijk *et al*., 2004) or *Poa pratensis* (Barcaccia *et al*., 1998). More recently, the possibility has been raised that epigenetic mechanisms can also be involved in the determination of apomixis (Lohe and Chaudhury, 2002; Koltunow and Grossniklaus, 2003). Many transposons and other repetitive sequences lie in the region of the apomixis loci mapped so far, and this has been taken to represent indirect evidence for an elevated extent of localized DNA methylation (Lohe and Chaudhury, 2002; Koltunow and Grossniklaus, 2003). A second line of evidence flows from the recognition that mutations in the direction of parthenogenesis and autonomous endosperm development involve epigenetic regulators of DNA and/or histone methylation (reviewed in Köhler and Makarevich, 2006). Some examples of the dominance of epialleles have been presented (Lohe and Chaudhury, 2002). The argument is therefore that a master gene(s) controlling apomixis can be epigenetically modified, or that regulatory factors reciprocally control epigenetic marks. The epigenetic

model combines prior hypotheses surrounding mutant alleles, dominant genes, hybridization and polyploidy (Lohe and Chaudhury, 2002; Koltunow and Grossniklaus, 2003).

Hypericum perforatum L. (St. John´s wort) produces pharmaceutically important metabolites with possible antidepressant, anticancer and antiviral/fungal/microbial activities (for a review seed Barnes *et al*., 2001). Moreover, the occurrence of variable ploidy levels, facultative apospory and pseudogamy in *Hypericum perforatum* L. has attracted the attention of apomixis researchers (Matzk *et al.*, 2001, 2003). Wild populations are predominantly tetraploid (2*n*=32), although both diploid (2*n*=16) and hexaploid (2*n*=48) forms are also known (Matzk *et al.*, 2001; Robson, 2002; Barcaccia *et al*., 2006). The variation in ploidy level is thought to reflect a dynamic reproductive system. Haploidization and polyploidization are the consequences of parthenogenesis of egg cells and fertilization of non-reduced aposporous egg cells, respectively (Barcaccia *et al.*, 2007). The result is that the species has a versatile mode of reproduction, ranging from completely sexual to nearly obligate apomictic. Along with a relatively small genome size and a short generation time, this feature has promoted *H. perforatum* to be a leading model for apomixis research (Matzk *et al*., 2001; Barcaccia *et al*., 2007). Here, we report the mapping and cloning of at least a part of a gene responsible for apospory in *H. perforatum*. We **s**uggest that apospory in *H. perforatum* is most probably controlled by dominant factors contained within the cloned sequence. This advance represents the initial step necessary for the isolation of the *Hypericum APOSPORY* locus *(HAPPY)*, which should provide a major insight into the molecular control of the aposporous mode of reproduction**.**

RESULTS

Apospory and parthenogenesis are under independent control

The four apomictic accessions (aAn, aNo, aSi, aTo) used as pollinators exhibited between 83% and 100% apospory and 80-100% parthenogenesis (Table S1). Both the diploid (sR1, sP1, sP2, sV1, sV2 and sV3) and the tetraploid ($sF_1$1, $sF_1$2, sR1C) sexual parents were confirmed as obligate sexuals, because FCSS analysis failed to detect either apospory or parthenogenesis among their selfed progeny. Hybrids between the obligate sexuals produced uniformly obligate sexual progeny (Table 1), suggesting that the parents were all homozygous for the genes responsible for sexual reproduction. Among the triploid F_1 (4x apomictic x 2x sexual) individuals, the ratio of aposporous to sexual plants was ~3:1, inferring the dominance of apospory at the triploid level (Table 1). This ratio is intermediate between the expected simplex 1:1 and duplex 5:1 ratios. In the F_1 populations 4x sexual ($F_1$1, $F_1$2, R1C) x 4x apomictic (aAn, aNo, aSi, aTo), 34 progeny were aposporous at least to some degree (Table S2), and 38 were sexual, consistent with a 1:1 segregation (χ^2 = 0.222). This 1:1 ratio was taken to indicate that apospory in the tetraploid male parent was a consequence of a simplex allelic constitution (for statistical details see Table S5).

In contrast to the apospory, the level of parthenogenesis was low (0-53%). None of the triploid progeny expressed an intermediate or a high level of parthenogenesis, but a small number of the tetraploid progeny did so. The ratio of plants with and without parthenogenesis was 29:30 among the triploids and 28:44 among the tetraploids (Table 1), which is close to a 1:1 ratio, so suggesting a simplex genetic constitution at the critical gene in the apomictic parent. Plants showing high levels of both apospory and parthenogenesis were rare (3/131) among the progeny of sexual x apomictic crosses. The occurrence of 16 triploid and seven tetraploid progeny which expressed apospory but not parthenogenesis, along with one triploid and one tetraploid which was parthenogenetic but not aposporous shows that these two apomixis components must be independently controlled in *H. perforatum.* To test whether recombination between apospory and parthenogenesis had occurred, a joint segregation analysis was performed. This showed a significant deviation from the expected 1:1:1:1 ratio at both the triploid and tetraploid levels (Table 1), demonstrating the presence of linkage, estimated to be 20.1cM (Table S5). We concluded that the two dominant alleles responsible for apospory and parthenogenesis were linked in the coupling phase.

An AFLP marker discriminates between accessions with contrasting modes of reproduction

An AFLP-based analysis was applied to a sample of ten apomictic and six sexual individuals (Figure 1, Table S1), employing 36 *Pst*I/*Mse*I primer combinations. Although there were many polymorphisms among the 16 DNAs, only one fragment was present in all of the apomictic but in none of the sexual templates. The same fragment also co-segregated with apospory among a sample of 40 F_1 progeny. The AFLP marker was convertible into a CAPS marker, since the sequences of the amplicon derived from template of aposporous and sexual accessions differed from one another by the presence/absence of an *Eco*RI site, which splits the 223bp sequence into two similar sized fragments (105bp and 118bp) (Table 2). DNA extracted from aposporous plants did appear to contain traces of the sexual allele, as shown by the appearance of a faint 105/118bp product along with the well amplified 223bp one (Table 2). However, analysis of the extended sequence of the apospory-specific allele showed that primer mismatching was responsible for the preferential amplification of the aposporous over the sexual allele (Figure S3). The association between the CAPS marker and the mode of reproduction was further tested by an analysis of all the segregating material available (157 tetraploid, triploid and diploid plants). The undigested 223bp fragment co-segregated perfectly with the aposporous mode of reproduction, allowing the CAPS marker to be used as a replacement for FCSS analysis (Table 2 and Table S2). As expected from the joint segregation between apospory and parthenogenesis, the CAPS marker was not predictive of parthenogenesis.

Cloning of the apospory-linked region

A BlastX analysis (Altschul *et al*., 1990) of the CAPS marker sequence revealed significant hits with genes encoding the ARIADNE-subclass of RING-finger proteins (Jackson *et al*., 2000; Mladek *et al*., 2003). The screen of the BAC library with the CAPS marker produced eight positive clones, which, upon fingerprinting, all belonged to a single contig (Figure S1). The longest BAC clone (25H09) containing the marker in its central region was sequenced (Genbank accession number *will be provided*). The GC content of the BAC sequence was 38.8%, rising to 41.2% in the coding regions (Figure 2). The 24 predicted genes in 25H09 are depicted in Figure 2, along with their closest BlastX hits (Table S4). These include, *inter alia*, a variety of transcription factors and four retrotransposons.

The *ARIADNE*-like gene containing the CAPS marker was a prime candidate for the *HAPPY* locus. In 25H09 (which contains the sexual allele), the coding region is of length 1,556bp with no introns. Its predicted translation product shares strong homology with the 562 residue product of the *A. thaliana ARIADNE 7 (ARI7)* gene. The relevant genomic region was amplified from four obligate aposporous and four sexual plants to investigate sequence variation in the candidate gene. To ensure capture of a simplex locus, at least 16 independent clones of these amplicon inserts were sequenced to have a 96% chance to clone any allele (Simko, 2004). Eleven haplotypes were identified, falling into two distinct groups (Figure 3). One group contained haplotypes 300 and 311, with only the latter being specific to aposporous plants. Among the ten haplotypes distributed among the sexual plants, 73 SNPs relative to the 25H09 sequence were identified, resulting in 40 amino acid changes. The apospory specific haplotype 311 contained 39 SNPs, some of which were responsible for a change in the translation sequence. A 6bp deletion in haplotype 200 removes two amino acid codons, but leaves the reading frame intact. Three other indels – two 2bp deletions and a 1bp insertion - were detected. The 2bp deletion in haplotype 311 caused a frame-shift which resulted in a truncated gene product sharing 48 residues at the N-terminus with the sexual alleles, followed by 19 unique residues up to the stop codon (Figure S3).

Allele-specific RT-PCR was performed to determine whether both the sexual and the aposporous alleles are expressed. As the target gene lacks introns, we first needed to successfully demonstrate the absence of contaminating genomic DNA in the cDNA template by a parallel assay of two adjacent genes (*HISTIDINE KINASE 2* and *Mo25*) using primers which amplify across an intron. The expression of the sexual and apomictic *HpARI* alleles in the floral buds at megasporogenesis was indistinguishable from that at megagametogenesis (Figure 4).

Using the *HpARI* locus as a landmark for the *HAPPY* locus, we also examined the quantitative nature of the *HpARI* alleles, by comparing hybridization intensities displayed by Southern blotting (Figure 5). The aposporous allele is marked by a 978bp *Eco*RI fragment lacking the internal *Eco*RI restriction site, while the sexual allele generated a pair of fragments (583bp and 395bp) because of the presence of the internal *Eco*RI site. The ratio of hybridization intensity in three aposporous accessions between the 978bp fragment and the 583bp/395bp pair was ~1:3, suggesting the presence of one apospory allele and three sexual alleles in the apomictic types, and four sexual alleles in the 4x sexual types. A similar Southern analysis based on either *Hin*dIII of *Bam*HI digestion produced only a single hybridizing fragment, indicating that the *HpARI* locus is single copy (data not shown). As a

further test, the *HpARI* sequence was cloned from a set of aposporous plants, and the clones digested with *Eco*RI. Of these, 15 lacked the restriction site (aposporous allele) and 49 possessed it (sexual allele), a ratio which conforms closely to 1:3. Thus we were able to conclude that the allelic constitution of the apospory locus in tetraploid plants is simplex. Pyrosequencing within *HpARI* was then used to determine the allelic constitution in the full collection of accessions. This revealed that the aposporous plants contained 20-25% of the apospory-specific *HpARI* allele and 70-75% of the sexual ones, while the sexual plants contain from 96-100% of the sexual allele (Figure 5). The very small proportion of apospory-specific alleles detected in the sexual plants lies within the range of error of the pyrosequencing method.

Genomic walking extends the apospory-specific genomic locus

When the alleles at 13 neighbouring genes along 25H09 in aposporous and sexual accessions were compared at the sequence level, no apospory-specific alleles were detected (Figure S2). Genome walking was then employed to investigate the immediate upstream and downstream regions of the aposporous *HpARI* allele. The former extended the sequence as far as the *DGCR14* orthologue, while the latter reached beyond a truncated second copy of *ARI* (*ARI-T*). From this extended sequence, it was possible to conclude that homology between the sexual and aposporous sequences is restricted to a region closely surrounding *HpARI*, and disappears abruptly in the C-terminal region of *DGCR14*, 72bp upstream of its predicted stop codon (Figure S4). At the downstream end, homology is lost within *ARI-T*, 102bp upstream of its predicted stop codon (Figure S5). Therefore, it seems likely that *HpARI* is a critical component of the *HAPPY* locus.

DISCUSSION

Apospory and parthenogenesis can recombine in *H. perforatum*

Extensive research aimed at unravelling the genetic control of apomixis has not resulted in a consensus picture, suggesting that a number of distinct pathways have evolved in various species. This idea fits well with the notion that apomixis in general, and each apomictic trait in particular, evolved independently (Van Dijk and Vijverberg, 2005). Both the mutation of genes involved in the control of sexual reproduction, and the *de novo* evolution of genes have been proposed as underlying the determination of apospory and parthenogenesis. Whether these two traits are under the control of a single master gene or the influence of a few closely related genes remains a puzzle for most apomicts. Studies aiming to understand the genetic and molecular factors underlying apomixis have been limited since the asexual mode of reproduction is often associated with polyploidy and high degrees of heterozygosity, traits which make genetic and genomic analyses very difficult. Apomixis in *H. perforatum* appears to be facultative to various extents, and restricted to polyploid types (Martonfi *et al*., 1996; Matzk *et al.*, 2001; Mayo and Langridge, 2003). Some of these apomicts are able to develop embryos from aposporous fertilized egg cells, and others by parthenogenesis from meiotically reduced egg cells. This has been taken to suggest that distinct genetic factors control apospory and parthenogenesis, and that the two traits may be developmentally uncoupled, a possibility previously raised by Noack (1939). This hypothesis is further supported by the finding of *H. perforatum* genotypes which almost exclusively express only one component of apomixis or suppress the other (Matzk *et al*., 2001; Barcaccia *et al.*, 2006).

The current experimental data have demonstrated that parthenogenetic capacity is preferentially expressed by aposporous egg cells. We have also documented the occurrence of aposporous egg cells in non-parthenogenetic individuals, as well as parthenogenetic development in meiotic egg cells. While apospory and parthenogenesis seem to be genetically linked (by ~20cM) and so tend to be co-inherited, recombination can separate them. The analysis of segregation data reveals that the putative apospory and parthenogenesis loci are associated in a chromosome window where recombination is possible. Both traits are determined at the tetraploid level by a dominant simplex allele. While this genetic model differs from the idea that the inheritance of apomixis is governed by the allelic state at a single, fully dominant gene (Savidan 1980, Leblanc *et al*., 1995, Bicknell *et al*., 2000), it is consistent with conclusions reached in other studies (Noyes and Riesenberg 2000, Matzk *et*

al., 2001, Van Baarlen *et al.*, 2002, Van Dijk and Bakx-Schotman, 2004, Catanach *et al*., 2006, Noyes *et al*., 2006).

Aposporous plants are heterozygous and contain both sexual and aposporous alleles

The apospory-specific CAPS marker and the corresponding genomic locus *HAPPY* are both present in plants which express varying degrees of apospory. This suggests that genes required for the apospory trait are located within the HAPPY locus, but additional modifier genes might govern the level of trait expression. This model is similar to what has been proposed for the genetic basis of apomixis in *Poa*, where five different loci are involved, including a dominant gene which controls the initiation of apospory and a recessive one which acts to modulate the expression of the trait (Matzk *et al*., 2005). The segregation analysis performed here has shown that the *HAPPY* locus is dominantly inherited, and the apomictic parents tested are all simplex, just as is the case in *Ranunculus auricomus* (reviewed in Koltunow and Grossniklaus, 2003). The retention of the sexual allele allows the plant to revert to sexual seed production if apospory is prevented in some way.

It seems clear that the allelic constitution at the *HpARI* homoeologues is the primary genetic basis of aposporous reproduction in *H. perforatum*. Only a restricted region surrounding the apospory-specific *HpARI* allele shows sequence similarity to the sexual alleles. Sequences beyond this limited region do not show detectable similarity between aposporous and sexual alleles nor give any hit when blasted against all known databases. The aposporous *HAPPY* sequence is rather distinct from the sexual one, so that a degree of hemizygosity is associated with this region in the simplex state (Figure 6). Hemizygosity has also been described at the *ASGR* locus in *Pennisetum*, in which the sexual types lack entirely the sequence corresponding to a part of the locus, while there is sequence homology along the remainder of the locus (Ozias-Akins *et al*., 1998).

The *HAPPY* locus includes a mutated *ARIADNE* orthologue

The truncated *HpARI* gene is a part of the *HAPPY* locus, making it possible to suggest that its product acts in a dominant negative fashion in a simplex dosage via an interaction with the gene product of the three remaining sexual sexual alleles. The ARIADNE proteins belong to a family of E3 ligases present in yeast, plants and animals, and thought to be involved in the control of ubiquitin-dependent protein degradation (reviewed in Vierstra, 2003). The

truncated *HpARI* gene is expressed, so its product may act to disrupt the normal Ring/U-box complexes formed by the non-truncated products, thereby impairing the progress of gametophytic development. This hypothesis will of course need testing by mutation and biochemical analyses of the *HpARI* protein. The function of *ARIADNE* genes, including *ARI7* (*At2g31510*), the *A. thaliana* orthologue of *HpARI*), remains obscure (Mladek *et al*., 2003). In *A. thaliana* a battery of *ARIADNE* genes is ubiquitously expressed, so it is likely their products play an important role during development. *ARIADNE* genes were first cloned and characterized in *Drosophila* and mammals. The *Drosophila ariadne-2* mutant is lethal, thus showing that it plays some essential developmental role (Aguilera *et al*., 2000). However, because these genes are present in the form of a multi-member gene family in both animals and plants, progress in defining the role of individual members has been necessarily slow. *A. thaliana ARI7* is a tandemly duplicated gene, unlike the single, structurally intact *HpARI*. Although mutant alleles which disrupt *A. thaliana ARI7* have been isolated (Table S6), none have as yet shown any evidence of developmental aberration, probably because of gene redundancy. A priority in our research programme is to optimize transformation methods in *H. perforatum* so that the effect of introducing the aposporous *HpARI* allele into a sexual *H. perforatum* host can be examined.

While we are currently concentrating on the functional role of *HpARI* in apospory, it is likely that further extension of the *HAPPY* locus sequence will be required to fully understand the control of apospory and its putative modifiers in *H. perforatum*. We have initiated the cloning of the *HAPPY* locus by sequencing BACs from fully aposporous 4x accessions. This line of research, along with a complementary project aimed at isolating candidate genes for parthenogenesis, should provide profound insights into the molecular basis of apomixis in *H. perforatum* and, more generally, in other plant species.

EXPERIMENTAL PROCEDURES

Plant materials

The 16 accessions of *H. perforatum* used were obtained from a variety of collection sites and botanical gardens in Europe (Table S1). Ten are apomictic, and the other six sexually reproducing, characteristics which are clearly distinguishable by flow cytometry (Matzk *et al*., 2000). A segregating population of 59 triploid individuals was obtained by pollinating sexual diploid plants with obligate apomictic tetraploid accessions. A second population of 72 tetraploid individuals was obtained by crossing a chromosome doubled sexual accession (R1) with tetraploid obligate apomicts. Two sexual progeny were taken as pollinators in crosses with a number of the obligate apomictic accessions (for the crossing scheme see Table 1). The resulting tetraploid progeny from these crosses were scored for the mode of reproduction using the flow cytometry seed screen (FCSS), and for the presence/absence of a putative apospory-linked marker.

Flow Cytometric Seed Screen (FCSS)

The FCSS screen is based on the measurement of DNA content in the embryonic and endosperm cells (Matzk *et al*., 2000, 2001). In diploids, sexual reproduction generates a diploid embryo and a triploid endosperm, whereas the seed of a tetraploid plant can include one of: 4C embryo and 6C endosperm cells (the product of sexually produced seeds from reduced, double-fertilized embryo sacs), 4C embryo and 10C endosperm cells (apomictic seeds from unreduced embryo sacs, and pseudogamous endosperm formation), 2C embryo and 6C endosperm (reduced, parthenogenetic amphihaploid progeny), or 6C embryo and 10C endosperm (unreduced, double-fertilized (poly)triploid B_{III} progeny).

Genetic analysis

Genetic analysis treated each mapping family as a BC_1 population. The estimation of linkage between apospory and parthenogenesis was conducted separately in the two populations. The pseudo-testcross strategy employed assumed the existence of a single dose allele (simplex locus) in the polyploid progeny. Since the simplex allele is inherited by half of the gametes, when crossed to a nulliplex individual, the segregation ratio is expected to be

1:1. Similarly a cross with a diploid generates a 1:1 segregation at the triploid level. Observed segregation ratios were tested by the χ^2 statistic. Genetic distances were converted from recombination frequencies to centiMorgans (cM) using the Kosambi (1944) mapping function (for details see Table S5).

AFLP and cleaved amplified polymorphic sequence (CAPS) analysis

Total genomic DNA was isolated with the Invisorb Spin Plant kit (Invitek, Berlin, Germany) and AFLP analysis was performed following Potokina *et al.* (2002). Selective amplification was achieved using fluorescently labelled *Pst*I-anchored primers with two selective nucleotides and *Mse*I-anchored ones with three selective nucleotides. The amplicons were resolved by capillary electrophoresis and fragment sizes estimated by comparison with a size standard (Genescan-500 Rox, Applied Biosystems Inc., Foster City, CA, USA) supplemented with five additional DNA fragments of known length (ranging from 568 to 812bp). The data were analysed using GeneScan software v3.0 (Perkin-Elmer ABI). To isolate specific fragments, the amplicons were re-electrophoresed through 4.5% polyacrylamide gels, visualized by silver-staining (Bassam *et al*., 1991), excised from the dried gel and eluted overnight in sterile water. An aliquot of the eluate represented the template for a further PCR, and these amplicons were sequenced either following their elution from a 1.5% agarose gel and QIAquick purification (Qiagen, Hilden, Germany), or were first cloned into a pGEM-T vector (Promega, Madison, WI, USA). For conversion into a CAPS marker, primers (Table S3) were designed from the sequence of the AFLP fragment, and the resulting amplicons digested overnight with *Eco*RI, using standard conditions for restriction. CAPS genotype was assessed by electrophoresis through 1.5% agarose gels.

Bacterial artificial chromosome (BAC) library and BAC clone characterization

DNA was extracted from the leaves of one month old progeny of a single diploid sexual plant, and used to construct a partial *Hin*dIII BAC library (Amplicon Express, Pullman, USA). Based on a genome size of ~630Mbp (http://www.rbgkew.org.uk/cval; Bennett *et al*., 1995; Barcaccia *et al*., 2007), the resulting 26,000 clone library represented a ~6-fold genome coverage. The inserts were released from a sample 20 random clones by *Not*I digestion and sized by PFGE (1% agarose gel, 0.5x TBE, 12.5°C, 6V/cm, 5s initial and 15s final pulse time, run time 16h). This produced an estimated mean insert size of 110kb. The

BAC library was spotted onto Hybond N+ membranes (GE Heathcare, London, UK) using a MicroGrid II robot (BioRobotics, UK). A probe containing the CAPS marker sequence was labelled with ^{33}P-α dCTP by random hexamer priming (Feinberg and Vogelstein, 1983), and hybridized to the library following Church and Gilbert (1984). Positive BAC clones retained after a further PCR test for the presence of the probe sequence were fingerprinted in triplicate by *Bam*HI, *Eco*RI, *Xba*I, *Xho*I, and *Hae*III restriction (Luo *et al.*, 2003). The digested DNA was labelled with SNaPshot labeling solution, and separated by capillary electrophoresis. Fragment analysis and contig assembly were facilitated by the software packages GeneMapper v4.0 (Applied Biosystems), FPPipeliner v.2.0 (BioinforSoft LLC) and FPC (Sanger Institute).

One BAC clone (25H09) was taken forward for sequencing. BAC DNA was isolated using a Plasmid Purification Maxi kit applying the low-copy plasmid/cosmid protocol (Qiagen). The BAC DNA was randomly sheared (Hydroshear, Genomic Solutions, Ann Arbor, MI, USA) and size-fractionated by agarose gel electrophoresis. Fragments of size 1-2kb were treated with Klenow DNA polymerase (Fermentas, Vilnius, Lithuania), blunt-end ligated into pBluescriptSK (Stratagene) and sequenced. The resulting sequences were assembled using Sequencher v4.0 software (Gene Codes Corporation) set to an overlap minimum of 20bp with 95% identity. Remaining gaps were closed by amplification based on primers derived from the flanking sequence. The final assembly consisted of 1,253 sequences, giving about six fold coverage. The accuracy of the assembly was confirmed by comparing the predicted and actual restriction digest profiles for a range of restriction enzymes. The final 141,941bp sequence was annotated, based on the software packages Fgenesh+ with a dicot Markov model and GeneID (Parra *et al*., 2000), applying matrices specific for *Arabidopsis thaliana* and tomato. Refinement of these predictions were achieved by aligning the *H. perforatum* genomic sequence with TIGR transcript assemblies of *Populus* spp., *Manihot esculenta*, *Euphorbia esula* (Childs *et al*., 2007) and *A. thaliana* proteins (TAIR version 7), using GenomeThreader software (Gremme *et al*., 2005). Consensus gene models were derived by comparing the gene models to a reference protein database (UNIREF90; Suzek *et al*., 2007). Manual inspection of the consensus gene models retained four transposon-related and 20 protein encoding genes.

Detection of single nucleotide polymorphisms (SNPs)

All 24 genes within 25H09 were amplified from four aposporous and four sexual plants (primer sequences given in Table S3). The resulting amplicons were cloned into

pCR2.1 (Invitrogen, Carlsbad, CA, USA), and sequenced to detect any haplotype variation present.

Allele-specific expression analysis

Total RNA was isolated from both 4-5mm and 6-8mm long pistils as well as young leaves using an Invitek RNA isolation kit (Invitek, Berlin, Germany). These pistil lengths correspond to the onset of, respectively, megasporogenesis and megagametogenesis (Galla *et al*., in press). For first strand cDNA synthesis, total RNA was treated with DNaseI (Invitrogen, Carlsbad, CA, USA) followed by RT-PCR with RevertAid H Minus M-MuLV Revertase transcriptase (Fermentas, Vilnius, Lithuania) using random hexamer priming. The primers were designed to both amplify both alleles uniformly and to enable them to be distinguished from one another by the presence of an *Eco*RV site in the aposporous allele (for primer sequences, see Table S3).

Allele quantification by pyrosequencing

Primers for pyrosequencing were derived using the SNP primer design v4.0 software (Biotag AB, Uppsala, Sweden). Pyrosequencing reactions were carried out using primers listed in Table S3, following the manufacturer's standard protocols. Allele frequencies were estimated using Biotage AB software.

Southern hybridization

A 5µg sample of genomic DNA was digested with *Eco*RI, separated by agarose gel electrophoresis and transferred onto a Hybond N+ membrane (GE Healthcare, London, UK). The membrane was hybridized with a probe consisting of a radioactively labelled 220bp sequence. Signal intensities (PSL/mm^2) were quantified by a phosphoimager FUJIX Bas 2000 (Fuji Photo Film, Tokyo, Japan). After background subtraction, the intensities of the 583bp and a 394bp fragments were summed, and compared with that of the 977bp fragment lacking the internal *Eco*RI site.

Genome walking

The genomic sequence of the target locus was extended using Genome Walker Universal kit (Clonetech, Palo Alto, CA, USA) and Advantage2 Polymerase mix (Clontech), following the manufacturer’s instructions.

ACKNOWLEDGEMENTS

This study was primarily supported by a research innovation grant “Pakt für Forschung und Innovation 2007” awarded to H.B. A financial contribution towards the cost of constructing the BAC library was provided by IPK-Gatersleben. We are grateful for the skilled technical assistance of Andreas Czihal, Elke Liemann, Annett Busching and Sabine Skiebe. We thank Drs. Daniela Schulte and Inge Matthies for help with, respectively, BAC fingerprinting and pyrosequencing. We also wish to thank Dr. Tim Sharbel and Prof. Konrad Bachmann for critical discussions. We thank www.smartenglish.co.uk for linguistic help in the preparation of this manuscript.

REFERENCES

Aguilera, M., Oliveros, M., Martinez-Padron, M., Barbas, J.A., and Ferrus, A. (2000) Ariadne-1: a vital *Drosophila* gene is required in development and defines a new conserved family of ringfinger proteins. *Genetics* **155**, 1231-1244.

Albertini, E., Porceddu, A., Ferranti, F., Reale, L., Barcaccia, G., Romano, B. and Falcinelli, M. (2001) Apospory and parthenogenesis may be uncoupled in *Poa pratensis*: a cytological investigation. *Sex. Plant Reprod.*, **14**, 213–217.

Albertini, E., Marconi, G., Reale, L., Barcaccia, G., Porceddu, A., Ferranti, F. and Falcinelli, M. (2005) SERK and APOSTART. Candidate genes for apomixis in *Poa pratensis*. *Plant Physio.l*, **138**, 2185-2199.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.*, **215**, 403-410.

Barcaccia, G., Mazzucato, A., Albertini, E., Zethof, J., Gerats, A., Pezzotti, M. and Falcinelli, M. (1998) Inheritance of parthenogenesis in *Poa pratensis* L.: auxin test and AFLP linkage analyses support monogenic control. *Theor. Appl. Genet.*, **97**, 74-82.

Barcaccia, G., Albertini, E., Rosellini, D., Tavoletti, S. and Veronesi, F. (2000) Inheritance and mapping of 2n-egg production in diploid alfalfa. *Genome*, **43**, 528-537.

Barcaccia, G., Arzenton, F., Sharbel, T.F., 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.F. (2007) Apomixis in St. John´s wort (*Hypericum perforatum* L.): An overview and glimpse towards the future. *In:* Hörandl, E., Grossniklaus, U. Sharbel, T., and van Dijk, P. (eds.) Apomixis: Evolution, Mechanism and Perspectives, Ganter Verlag, Ruggell, Liechtenstein, Regnum Vegetabile.

Barnes, J., Anderson, L.A. and Phillipson, J.D. (2001) St. John´s wort (*Hypericum perforatum* L.): a review of its chemistry, pharmacology and clinical properties. *J. Pharm. Pharmacol.* 53, 583-600.

Barrell, P.J. and Grossniklaus, U. (2005) Confocal microscopy of whole ovules for analysis of reproductive development: the elongate1 mutant affects meiosis II. *Plant J.*, **43**,

309-320.

Bassam, B.J., Caetano-Anolles, G. and Gresshoff, P.M. (1991) Fast and sensitive silver staining of DNA in polyacrylamide gels. *Anal. Biochem.*, **196**, 80-83.

Bicknell, R.A. (1997) Isolation of a diploid apomictic plant of *Hieracium aurantiacum*. *Sex. Plant Reprod.*, **10**, 168-172.

Bicknell, R.A., Borst, N.K. and Koltunow, A.M. (2000) Monogenic inheritance of apomixis in two *Hieracium* species with distinct developmental mechanisms. *Heredity*, **84 (2)**, 228-237.

Carman, J.G. (1997) Asynchronous expression of duplicate genes in angiosperms may cause apomixis, bispory, tetraspory, and polyembryony. *Biol. J. Linn. Soc.*, **61**, 51-94.

Catanach, A.S., Erasmuson, S.K., Podivinsky, E., Jordan, B.R. and Bicknell, R. (2006) Deletion mapping of genetic regions associated with apomixis in *Hieracium*. *Proc. Natl. Acad. Sci. U. S. A.*, **103**, 18650-18655.

Chaudhury, A.M., Ming, L., Miller, C., Craig, S., Dennis, E.S. and Peacock, W.J. (1997) Fertilization-independent seed development in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U. S. A.*, **94**, 4223-4228.

Childs, K.L., Hamilton, J.P., Zhu, W., Ly, E., Cheung, F., Wu, H., Rabinowicz, P.D., Town, C.D., Buell, C.R. and Chan, A.P. (2007) The TIGR Plant Transcript Assemblies database. *Nucl. Acids Res.*, **35**, D846-851.

Church, G.M. and Gilbert, W. (1984) Genomic sequencing. *Proc. Natl. Acad. Sci. U. S. A.*, **81**, 1991-1995.

Conner, J.A., Goel, S., Gunawan, G., Cordonnier-Pratt, M.M., Johnson, V.E., Liang, C., Wang, H., Pratt, L.H., Mullet, J.E., Debarry, J., Yang, L., Bennetzen, J.L., Klein, P.E. and Ozias-Akins, P. (2008) Sequence analysis of bacterial artificial chromosome clones from the apospory-specific genomic region of *Pennisetum* and *Cenchrus*. *Plant Physiol.*, **147**, 1396-1411.

d'Erfurth, I., Jolivet, S., Froger, N., Catrice, O., Novatchkova, M. and Mercier, R. (2009) Turning meiosis into mitosis. *PLoS Biol.*, **7**, e1000124.

Feinberg, A.P. and Vogelstein, B. (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.*, **132**, 6-13.

Gremme, G., Brendel, V., Sparks, M.E. and Kurtz, S. (2005) Engineering a software tool for gene structure prediction in higher organisms. *Information and Software Technology*, **47**, 965-978.

Guitton, A.E., Page, D.R., Chambrier, P., Lionnet, C., Faure, J.E., Grossniklaus, U. and

Berger, F. (2004) Identification of new members of Fertilisation Independent Seed Polycomb Group pathway involved in the control of seed development in *Arabidopsis thaliana*. *Development*, **131**, 2971-2981.

Hörandl, E., Grossniklaus, U., van Dijk, P.J. and Sharbel, T. (eds.) (2007) Apomixis: evolution, mechanisms and perspectives. Regnum Vegetabile vol. 147, Gantner Verlag, Rugell, Liechtenstein.

Holsinger, K.E. (2000) Reproductive systems and evolution in vascular plants. *Proc. Natl. Acad. Sci. U. S. A.*, **97**, 7037-7042.

Huang, B.Q. and Sheridan, W.F. (1996) Embryo sac development in the maize indeterminate gametophyte1 mutant: abnormal nuclear behaviour and defective microtubule organization. *Plant Cell*, **8**, 1391-1407.

Jackson, P.K., Eldridge, A.G., Freed, E., Furstenthal, L., Hsu, J.Y., Kaiser, B.K. and Reimann, J.D. (2000) The lore of the RINGs: substrate recognition and catalysis by ubiquitin ligases. *Trends Cell Biol.*, **10**, 429-439.

Kantama, L., Sharbel, T.F., Schranz, M.E., Mitchell-Olds, T., de Vries, S. and de Jong, H. (2007) Diploid apomicts of the *Boechera holboellii* complex display large-scale chromosome substitutions and aberrant chromosomes. *Proc. Natl. Acad. Sci. U. S. A.*, **104**, 14026-14031.

Köhler, C. and Makarevich, G. (2006) Epigenetic mechanisms governing seed development in plants. *EMBO Rep.*, **7**, 1223-1227.

Kojima, A. and Nagato, Y. (1997) Discovery of highly apomictic and highly amphimictic dihaploids in *Allium tuberosum*. *Sex. Plant Reprod.*, **10**, 8-12.

Koltunow, A.M. (1993) Apomixis: Embryo Sacs and Embryos Formed without Meiosis or Fertilization in Ovules. *Plant Cell*, **5**, 1425-1437.

Koltunow, A.M. and Grossniklaus, U. (2003) Apomixis: a developmental perspective. *Annu. Rev. Plant Biol.*, **54**, 547-574.

Kosambi, D. (1944) The estimation of map distances from recombination values. *Ann. Eugen.*, **12**, 172-175.

Labombarda, P., Busti, A., Caceres, M.E., Pupilli, F. and Arcioni, S. (2002) An AFLP marker tightly linked to apomixis reveals hemizygosity in a portion of the apomixis-controlling locus in *Paspalum* simplex. *Genome*, **45**, 513-519.

Leblanc, O., Grimanelli, D., Gonzalez de Leon, D. and Savidan, Y. (1995) Detection of the apomictic mode of reproduction in maize-*Tripsacum* hybrids using maize RFLP markers. *Theor. Appl. Genet*. 90, 1198-1203.

Lohe, A.R. and Chaudhury, A. (2002) Genetic and epigenetic processes in seed development. *Curr. Opin. Plant Biol.* 5, 19-25.

Luo, M.C., Thomas, C., You, F.M., Hsiao, J., Ouyang, S., Buell, C.R., Malandro, M., McGuire, P.E., Anderson, O.D. and Dvorak, J. (2003) High-throughput fingerprinting of bacterial artificial chromosomes using the snapshot labelling kit and sizing of restriction fragments by capillary electrophoresis. *Genomics*, **82**, 378-389.

Martonfi, P., Brutovska, R., Cellarova, E. and Repcak, M. (1996) Apomixis and hybridity in *Hypericum perforatum*. *Folia Geobot. Phytotax.* **31**, 389-396.

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

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

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

Matzk, F., Prodanovic, S., Baumlein, H. and Schubert, I. (2005) The Inheritance of apomixis in *Poa pratensis* confirms a five locus model with differences in gene expressivity and penetrance. *Plant Cell*, **17**, 13-24.

Mladek, C., Guger, K. and Hauser, M.T. (2003) Identification and characterization of the *ARIADNE* gene family in *Arabidopsis*. A group of putative E3 ligases. *Plant Physiol.*, **131**, 27-40.

Mayo, G.M. and Langridge, P. (2003) Modes of reproduction in Australian populations of *Hypericum perforatum* L. (St. John´s wort) revealed by DNA fingerprinting and cytological methods. *Genome*, **46**, 573-579.

Noack, K.L. (1939) Über *Hypericum* - Kreuzungen VI: Fortpflanzungsverhältnisse und bastarde von *Hypericum perforatum* L. *Z. Indukt Abstamm. Vererbungsl.*, **76**, 569-601.

Nogler, G.A. (1982) How to obtain diploid apomictic *Ranunculus auricomus* plants not found in the wild state. *Bot. Helv.*, **91**, 13-22.

Nogler, G.A. (1984) *Gametophytic apomixis,* Berlin: Springer-Verlag.

Noyes, R.D. and Rieseberg, L.H. (2000) Two independent loci control agamospermy (apomixis) in the triploid flowering plant *Erigeron annuus*. *Genetics*, **155**, 379-390.

Noyes, R.D., Baker, R. and Mai, B. (2006) Mendelian segregation for two-factor apomixis in *Erigeron annuus* (*Asteraceae*). *Heredity*, **98**, 92-98.

Ohad, N., Margossian, L., Hsu, Y.C., Williams, C., Repetti, P. and Fischer, R.L. (1996) A

mutation that allows endosperm development without fertilization. *Proc. Natl. Acad. Sci. U. S. A*, **93**, 5319-5324.

Ozias-Akins, P., Roche, D. and Hanna, W.W. (1998) Tight clustering and hemizygosity of apomixis-linked molecular markers in *Pennisetum squamulatum* implies genetic control of apospory by a divergent locus that may have no allelic form in sexual genotypes. *Proc. Natl. Acad. Sci. U. S. A.*, **95**, 5127-5132.

Ozias-Akins, P. (2006) Apomixis: Developmental haracteristics and genetics. *Crit. Rev. Plant Sci.*, **25**, 199-214.

Ozias-Akins, P. and van Dijk, P.J. (2007) Mendelian genetics of apomixis in plants. *Annu. Rev. Genet.*, **41**, 509-537.

Parra, G., Blanco, E. and Guigo, R. (2000) GeneID in *Drosophila*. *Genome Res.*, **10**, 511-515.

Potokina, E., Blattner, R., Alexandrova, T. and Bachmann, K. (2002) AFLP diversity in the common vetch (*Vicia sativa* L.) on the world scale. *Theor. Appl. Genet.*, **105**, 58-67.

Ravi, M., Marimuthu, M.P. and Siddiqi, I. (2008) Gamete formation without meiosis in *Arabidopsis*. *Nature*, **451**, 1121-1124.

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

Roche, D., Hanna, W.W. and Ozias-Akins, P. (2001) Is supernumerary chromatin involved in gametophytic apomixis of polyploid plants? *Sex. Plant Reprod.*, **13**, 343-349.

Roy, B.A. (1995) The breeding systems of six species of *Arabis* (*Brassicaceae*). *Am. J. Bot.*, **82**, 869-877.

Savidan, Y. (1980) Chromosomal and embryological analyses in sexual x apomictic hybrids of *Panicum maximum* Jacq. . *Theor. Appl. Genet.* **58**, 153-156.

Savidan, Y., Carman, J.G. and Dresselhaus, T. (2001) The Flowering of apomixis: From Mechanisms to Genetic Engineering Mexico: CIMMYT, IRD, European Commision DG VI. (FAIR).

Sharbel, T.F., Voigt, M.L., Corral, J.M., Thiel, T., Varshney, A., Kumlehn, J., Vogel, H. and Rotter, B. (2009) Molecular signatures of apomictic and sexual ovules in the *Boechera holboellii* complex. *Plant J.*, **58**, 870-882.

Simko, I. (2004) One potato, two potato: haplotype association mapping in autotetraploids. *Trends Plant Sci.*, **9**, 441-448.

Spillane, C., Curtis, M.D. and Grossniklaus, U. (2004) Apomixis technology development-virgin births in farmers' fields? *Nature Biotechnol.*, **22**, 687-691.

Suzek, B.E., Huang, H., McGarvey, P., Mazumder, R. and Wu, C.H. (2007) UniRef: comprehensive and non-redundant UniProt reference clusters. *Bioinformatics*, **23**, 1282-1288.

Tucker, M.R., Araujo, A.C., Paech, N.A., Hecht, V., Schmidt, E.D., Rossell, J.B., De Vries, S.C. and Koltunow, A.M. (2003) Sexual and apomictic reproduction in *Hieracium* subgenus *pilosella* are closely interrelated developmental pathways. *Plant Cell*, **15**, 1524-1537.

van Baarlen, P., de Jong, J.H. and van Dijk, P.J. (2002) Comparative cyto-embryological investigations of sexual and apomictic dandelions (*Taraxacum*) and their apomictic hybrids *Sex. Plant Reprod.*, **15**, 31–38.

Van Dijk, P.J. and Bakx-Schotman, J.M. (2004) Formation of unreduced megaspores (diplospory) in apomictic dandelions (*Taraxacum officinale*) is controlled by a sex-specific dominant locus. *Genetics*, **166**, 483-492.

Van Dijk P.J. and Vijverberg, K. (2005) The significance of apomixis in the evolution of the angiosperms: a reappraisal. In Plant species-level systematics: new perspectives on pattern & process Ruggell, Lichtenstein: Gantner Verlag.

Vierstra, R.D. (2003) The ubiquitin/26S proteasome pathway, the complex last chapter in the life of many plant proteins. *Trends Plant Sci.*, **8**, 135-142.

TABLE AND FIGURE LEGENDS

Table 1. Crossing scheme (upper panel) and progeny analysis (lower panel) undertaken to determine the mode of reproduction of apomixis in *H. perforatum.* Phenotype with respect to apospory scored by flow cytometry, supported by genotyping.

Table 2. CAPS analysis of 28 sexual (S) and aposporous (A) *H. perforatum* individuals. The presence of the 223bp amplicon lacking an internal *Eco*RI site correlates perfectly with apospory, whereas the amplificon derived from templates of plants having the sexual mode of reproduction are digested by *Eco*RI into two closely migrating fragments of size ~110bp. The level of apospory of each plant is shown in the upper panel. Progeny analysis from various crosses with respect to apospory and CAPS marker state is shown in the lower panel.

Figure 1. AFLP profiling of ten apomictic (A) and six sexual accessions (S) of *H. perforatum*. The vertical line indicates the fragment present exclusively in the profile of apomictic individuals.

Figure 2. Gene annotation of the 142kb BAC clone 25H09 containing *HpARI. HK1,2*: histidine kinase, *GH1,2*: glycosyl hydrolase, RT1,2,3,4: retrotransposon, *ExP1*: expressed protein, *PAT1*: a member of the GRAS family, *WRKY*: WRKY transcription factor, *DGCR*: hypothetical protein, *ARI*: ubiquitin-mediated E3-ligase, *ARI-T*: a truncated version *ARI*, *NPH3*: plastidal phototropic-responsive protein, *TK*: thymidine kinase, *RNP*: heterogeneous nuclear ribonucleoprotein, *AUXR*: auxin-response protein, *RimL*: GCN5-related N-acetyltransferase domain containing protein, *Mo25*: membrane protein of the Mo-25 family, *TtRNAS*: threonine-tRNA ligase, *RINGH2*: polygalcturonase and a RING finger protein of the H2 class. Other details of the annotation given in Fig. S4.

Figure 3. *HpARI* haplotypes of a panel of aposporous and sexual accessions of *H. perforatum* (upper panel) and an analysis of haplotype frequency (lower panel).

Figure 4. Expression of contrasting *HpARI* alleles, as estimated by RT-PCR. Both alleles are expressed in the pistils, but not in the leaves of aposporous plants.

Figure 5. Determination of allele frequencies via pyrosequencing (upper panel). Southern hybridization to determine the ratio of sexual and aposporous alleles in tetraploid *H. perforatum* (lower panel).

Figure 6. Schematic structure of the genomic region surrounding *HpARI* in sexual and aposporous types. *ARI*: the marker-containing sequence, *T*: truncated version of *HpARI*, *DGCR* and *NPH*: respectively the upstream and downstream genes,. Tetraploid apomicts possess three sexual alleles and one aposporous allele. The red arrows indicate where sequence homology between the sexual and aposporous alleles is interrupted. BlastX negative sequence in the flanking region of the aposporous allele is shown by dashed lines. The light blue box indicates the position of the AFLP locus used to generate the CAPS marker, and the black bar within it the informative *Eco*RI site.

SUPPORTING INFORMATION

Table S1. Origin of the ten apomictic (A) and six sexual (S) *H. perforatum* accessions profiled by AFLP (upper panel). The mode of reproduction of the four apomictic accessions (An, No, Si, To) used as male parents exhibit 83-100% apospory and 80-100% parthenogenesis. No apospory or parthenogenesis was detected in the obligate sexual lines sR1, sP1, sP2, sV1, sV2 and sV3 (lower panel).

Table S2. Plant identifier, % apospory, % parthenogenesis and marker state in tetraploid, triploid and diploid F_1 progeny of sexual x apomictic crosses.

Table S3. Primer sequences used. The upper panel shows gene symbols, primer position within BAC clone 25H09 and predicted amplicon lengths. The lower panel shows the primer sequences used for SNP pyrosequencing.

Table S4. BlastX analysis of the genes present on BAC clone 25H09.

Table S5. Statistical analysis of segregation ratios among the F_1 progeny of sexual x apomictic crosses.

Table S6. T-DNA mutants which disrupt ARI ring finger genes in *A. thaliana.*

Figure S1. Contig of BAC clones containing *HpARI*. Clone 25H09 was the one chosen for shotgun sequencing. Sequencing coverage varied from three to 17 fold.

Figure S2. Haplotype analysis and allele frequencies within *DGCR* and *NPH3*, the genes flanking *HpARI*.

Figure S3. Nucleotide and peptide sequences of the sexual haplotype "300" and the apospory-specific haplotype "311". The alleles are similar for the first 48 residues (red shading). Residue exchanges are underlined. A 2bp deletion in the aposporous allele generates a frame shift which leads to translation of 19 abnormal residues (blue background). The position of the primer pair used for the CAPS marker (shaded in green) and the informative *Eco*RI site (underlined) are shown.

Figure S4. Nucleotide alignment of the sexual and aposporous *HpARI* sequence and that of its 5' gene neighbour *DGCR*. The predicted intron in *DGCR* is marked by lower case lettering. The arrow indicates where homology between the sexual and aposporous sequence is first disrupted. No homology was detected beyond this point.

Figure S5. Nucleotide alignment of the sexual and aposporous *HpARI* sequence and that of its 3' gene neighbour *ARI-T*. The arrow indicates where homology between the sexual and aposporous sequence is first disrupted.

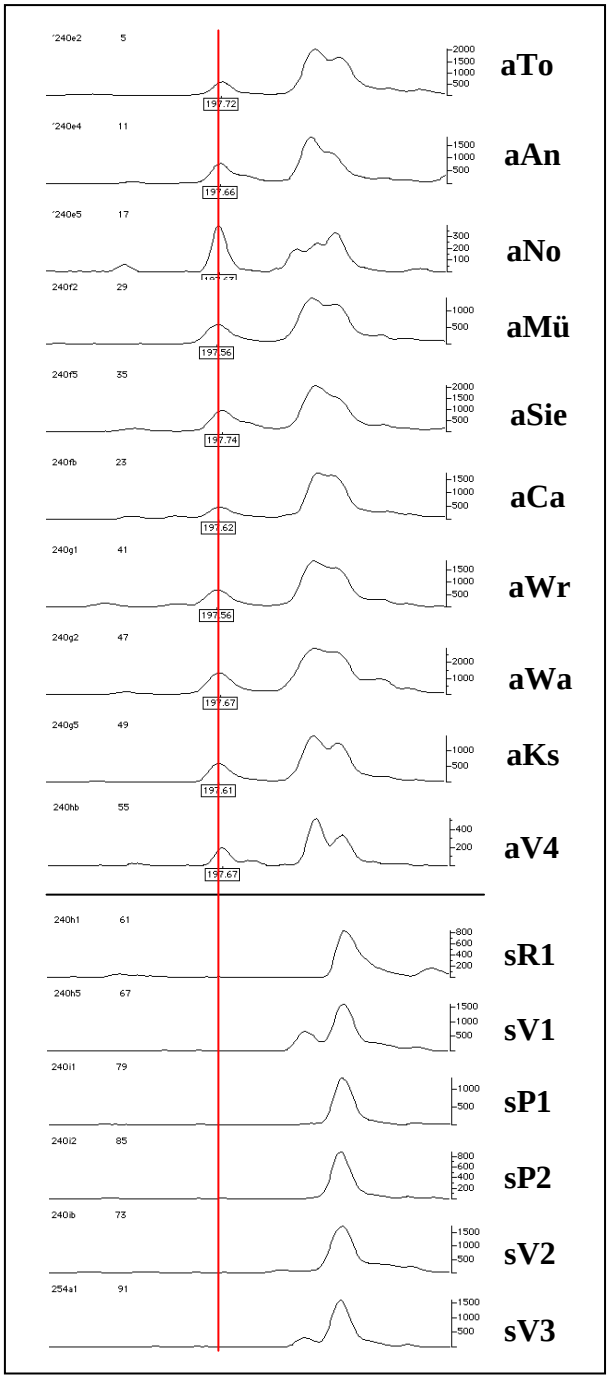


Figure 1

Diploids

sR1(2n=2x=16)	X	sV1(2n=2x=16)	→	1(sR1 x aV1) (2n=3x=24)
		sP2(2n=2x=16)		1(sR1 x aP2) (2n=3x=24)

Triploids

sR1(2n=2x=16)				1(sR1 x aNo) (2n=3x=24)
sV1(2n=2x=16)	X	aNo(2n=4x=32)	→	1(sV1 x aNo) (2n=3x=24)
sV3(2n=2x=16)	X	aTo(2n=4x=32)	→	1(sV3 x aNo) (2n=3x=24)
sP2(2n=2x=16)				1(sP2 x aTo) (2n=3x=24)

Tetraploids

sR1(2n=2x=16)
↓
polyploidization
(colchicine)
↓

$sR1C_1$(2n=4x=32)	X	aSi(2n=4x=32)	→	4($sR1C_1$ x aSi) (2n=4x=32)
		aTo(2n=4x=32)		4($sR1C_1$ x aTo) (2n=4x=32) (latter was renamed to sF_1)

				4($sF_1$1 x aNo) (2n=4x=32)
$sF_1$1(2n=4x=32)		aNo(2n=4x=32)		4($sF_1$1 x aAn) (2n=4x=32)
$sF_1$2(2n=4x=32)	X	aAn(2n=4x=32)	→	4($sF_1$1 x aSi) (2n=4x=32)
		aSi(2n=4x=32)		4($sF_1$2 x aNo) (2n=4x=32)
				4($sF_1$2 x aAn) (2n=4x=32)

Cross	Plants	APO+/PAR+	APO+/PAR-	APO-/PAR+	SEX
Diploid					
2(sR1 x sP2)	11	0	0	0	11
2(sP2 x sV1)	13	0	0	0	13
Total	24	0	0	0	24
Triploid					
1/2(sR1 x aNo)	16	7	5	0	4
1(sV1 x aNo)	25	12	6	1	6
1(sV3 x aNo)	4	3	1	0	0
1(sP2 x aTo)	14	6	4	0	4
Total	59	28	16	1	14
Tetraploid					
4($sF_1$1 x aAn)	8	2	0	0	6
4($sF_1$1 x aSi)	4	0	1	0	3
4($sF_1$1 x aNo)	10	8	0	0	2
4($sF_1$2 x aAn)	10	3	0	1	6
4($sF_1$2 x aNo)	24	12	4	0	8
4($sR1C_1$ x aSi)	16	2	2	0	12
Total	72	27	7	1	38

Table 1

1	1(sR1xaNo)1/1 99%	**11**	1(sV1xaNo)1/2 94%	**21**	1(sV1xaNo)1/23 10%
2	1(sR1xaNo)1/2 99%	**12**	1(sV1xaNo)1/3 100%	**22**	1(sV1xaNo)1/25 100%
3	1(sR1xaNo)1/3 7%	**13**	1(sV1xaNo)1/4 100%	**23**	1(sV1xaNo)1/27 100%
4	1(sR1xaNo)1/4 100%	**14**	1(sV1xaNo)1/5 90%	**24**	1(sV1xaNo)1/28 0%
5	1(sR1xaNo)1/5 98%	**15**	1(sV1xaNo)1/6 100%	**25**	1(sV1xaNo)2a/1 100%
6	1(sR1xaNo)2/1 0%	**16**	1(sV1xaNo)1/7 0%	**26**	1(sV1xaNo)2b/2 100%
7	1(sR1xaNo)2/2 0%	**17**	1(sV1xaNo)1/10 0%	**27**	1(sV1xaNo)2b/3 100%
8	1(sR1xaNo)2/4 95%	**18**	1(sV1xaNo)1/17 100%	**28**	1(sV1xaNo)2b/4 0%
9	1(sR1xaNo)3/5 0%	**19**	1(sV1xaNo)1/19 100%		
10	1(sV1xaNo)1/1 99%	**20**	1(sV1xaNo)1/20 95%		

Cross	No. of plants	FCSS		Marker state		
		SEX	APO	SEX	APO	n. d.
Diploid						
2(sR1 x sP2)	11	11	0	11	0	0
2(sP1 x sV1)	13	13	0	13	0	0
Total	24	24	0	24	0	0
Triploid						
1/2(sR1 x aNo)	16	4	12	4	12	0
1(sV1 x aNo)	25	7	18	7	18	0
1(sV3 x aNo)	4	0	4	0	4	0
1(sP2 x aTo)	14	4	10	4	10	0
Total	59	15	44	15	44	0
Tetraploid						
4($sF_1$1 x aAn)	8	6	2	6	2	0
4($sF_1$1 x aSi)	4	3	1	3	1	0
4($sF_1$1 x aNo)	10	2	8	2	6	2
4($sF_1$2 x aAn)	10	7	3	7	3	0
4($sF_1$2 x aNo)	24	8	16	8	14	2
4(sR1C_2 x aSi)	16	12	4	12	4	0
Total	72	38	34	38	30	4

Table 2

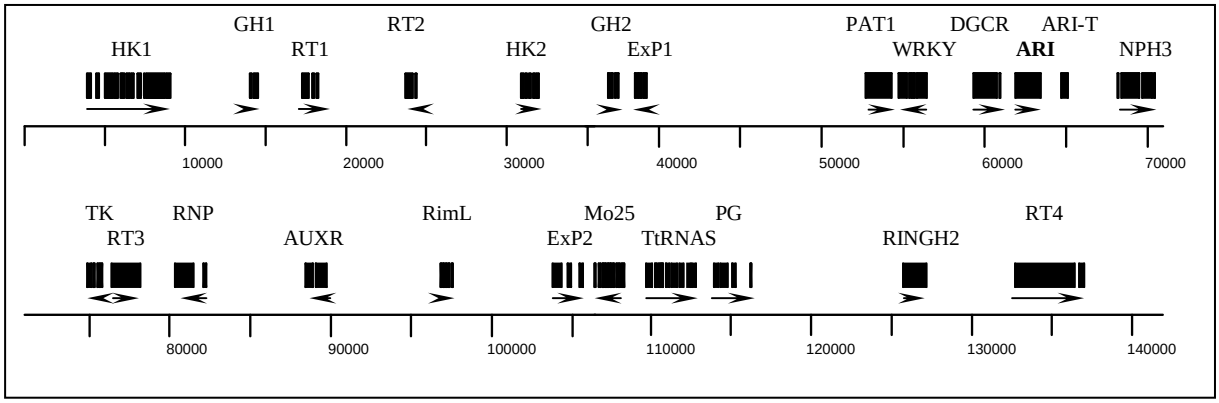

GH1
RT2
GH2
PAT1
DGCR
ARI-T
HK1
RT1
HK2
ExP1
WRKY
ARI
NPH3
10000
20000
30000
40000
50000
60000
70000
TK
RNP
RimL
Mo25
PG
RT4
RT3
AUXR
ExP2
TtRNAS
RINGH2
80000
90000
100000
110000
120000
130000
140000

Figure 2

	34	38	39	41	47	52	65	68	69	80	92	94	104	105	128	131	146	147	153	155	176	206	248	267	268	307	312	318	320	341	364	367	378	389	477	484	488	521
100	A	A	G	T	T	G	T	C	A	G	T	C	G	G	G	G	A	G	A	T	T	C	G	G	T	G	G	A	C	G	C	C	G	T	T	C	T	C
133	A	A	G	C	C	G	G	T	A	G	T	C	G	G	G	G	A	G	A	T	C	C	A	G	T	G	G	A	T	G	C	C	A	T	C	C	C	G
132	A	A	G	T	T	G	T	C	A	G	T	C	G	G	G	G	A	G	A	T	T	C	G	G	T	G	G	A	C	G	C	C	G	T	T	C	T	C
131	A	A	G	T	T	G	T	C	A	G	T	C	G	G	G	G	A	G	A	T	T	C	G	G	T	G	G	A	C	G	C	C	G	T	T	C	T	C
130	A	A	G	C	T	A	T	T	A	G	T	C	A	G	G	G	A	G	A	T	T	C	G	G	T	G	G	A	C	G	C	C	G	T	T	C	T	C
120	A	A	G	C	T	G	T	C	A	G	T	C	G	G	G	G	A	G	A	T	T	C	G	G	T	G	G	A	C	G	T	C	G	T	T	T	T	C
111	A	A	G	T	T	G	T	C	A	G	T	C	G	G	G	G	A	G	A	T	T	C	G	G	T	G	G	A	C	G	C	C	G	T	T	T	T	C
110	A	A	G	C	T	G	T	C	A	G	T	C	G	G	G	G	A	G	A	T	T	T	G	G	T	G	G	A	C	G	C	C	G	T	T	C	T	C
200	A	A	G	T	T	G	T	C	A	G	T	C	G	G	G	G	A	G	A	T	T	C	G	G	T	G	G	A	C	G	C	C	G	T	T	C	T	C
300	A	A	G	C	T	G	T	T	G	G	T	C	A	G	T	G	A	G	A	T	T	C	G	G	T	A	A	G	C	G	C	T	G	T	C	C	C	G
311	G	C	C	T	C	G	G	T	A	A	C	G	G	A	G	A	-	-	T	C	C	C	G	-	-	A	A	G	C	A	C	C	G	C	C	C	T	C

	524	540	572	599	617	626	630	649	653	673	674	701	741	750	762	787	799	800	801	843	949	960	977	979	984	990	996	1020	1059	1070	1079	1087	1093	1095	1100	1113	1115	1116
100	A	A	T	C	G	C	G	A	G	G	C	C	A	T	-	G	A	T	G	G	T	T	G	G	A	G	C	G	A	G	C	C	G	A	C	G	T	G
133	T	T	T	C	G	C	G	A	G	G	C	C	A	A	-	G	A	T	G	G	T	T	G	G	A	C	T	A	A	G	C	C	G	T	C	G	T	G
132	A	T	T	C	G	C	G	A	G	G	C	C	A	T	-	G	A	T	G	G	T	T	G	G	A	C	T	G	A	G	C	C	G	T	C	G	T	G
131	A	T	T	C	G	C	G	A	G	G	C	C	A	T	-	G	A	T	G	G	T	T	G	G	A	C	T	G	A	G	C	C	G	T	C	G	T	G
130	A	A	T	C	G	C	G	A	G	G	C	C	A	T	-	G	A	T	G	G	T	T	G	G	A	C	T	G	A	G	C	C	G	T	C	G	T	G
120	A	T	T	C	G	C	G	A	G	G	C	C	A	T	-	G	A	T	G	G	T	T	G	G	A	G	C	G	A	G	C	C	G	A	C	G	T	G
111	A	T	T	C	G	C	G	A	G	G	C	C	T	T	-	G	A	T	G	A	T	T	G	G	A	C	T	G	A	G	C	C	G	T	C	G	T	G
110	A	T	T	C	G	C	G	G	G	A	C	C	A	T	-	G	A	T	G	G	T	T	G	G	A	C	T	G	A	G	C	C	A	T	C	G	T	G
200	A	T	T	C	G	C	G	A	G	G	C	C	A	T	-	G	A	T	G	G	T	T	G	G	A	C	T	G	A	G	C	C	G	T	C	G	-	-
300	T	T	A	C	G	T	T	A	T	A	A	C	A	T	-	G	C	T	T	G	G	C	A	G	G	G	C	G	A	A	C	G	G	T	C	G	T	G
311	A	T	T	A	A	C	G	A	G	G	C	G	A	T	T	T	C	C	T	G	T	C	A	C	G	G	C	G	G	A	T	G	G	T	G	T	T	G

	1117	1118	1119	1120	1126	1150	1161	1162	1163	1180	1222	1236	1255	1260	1267	1268	1269	1279	1284	1285	1286	1287	1297	1300	1316	1333	1384	1385	1387	1394	1398	1455	1498	1510	1513	1530
100	C	A	A	A	A	C	T	T	T	A	T	C	T	C	G	A	T	G	C	T	T	T	C	G	G	C	A	A	G	A	C	G	A	A	C	A
133	C	A	A	A	A	C	T	G	T	A	C	C	C	C	G	A	T	G	C	T	T	T	C	G	G	C	A	A	G	G	C	A	A	A	C	A
132	C	A	A	A	A	C	T	G	T	G	C	C	C	C	G	A	T	G	C	T	T	T	C	G	G	C	A	A	G	A	C	G	A	A	C	A
131	A	A	A	A	A	C	T	G	T	A	T	C	C	C	G	A	T	G	C	T	T	T	C	G	G	C	A	A	G	A	C	G	A	A	C	A
130	C	A	A	A	A	C	T	G	T	G	C	C	C	C	G	A	T	G	C	T	T	T	C	G	G	C	A	A	G	A	C	G	A	A	C	A
120	C	A	A	A	A	C	T	T	T	A	T	C	T	C	G	A	T	G	C	T	T	T	C	G	G	C	A	A	G	A	C	G	A	A	C	A
111	C	A	A	A	A	C	T	G	T	A	T	C	C	C	G	A	T	G	C	T	T	T	C	G	G	C	A	A	G	A	C	G	A	A	C	A
110	C	A	A	A	A	C	T	G	T	A	T	C	C	C	G	A	T	G	C	T	T	T	C	G	G	C	A	A	G	A	C	G	A	A	C	A
200	-	-	-	-	A	C	T	G	T	A	T	C	C	C	G	A	T	G	C	T	T	T	C	G	G	C	A	A	G	A	C	G	A	A	C	A
300	C	A	A	A	T	C	G	G	C	A	C	T	C	G	C	G	C	C	C	T	T	T	C	A	G	C	A	C	G	A	G	G	A	A	C	A
311	C	A	A	A	A	T	G	G	C	A	C	T	C	G	C	G	C	C	A	A	A	G	A	A	T	G	T	C	A	A	G	G	T	G	A	G

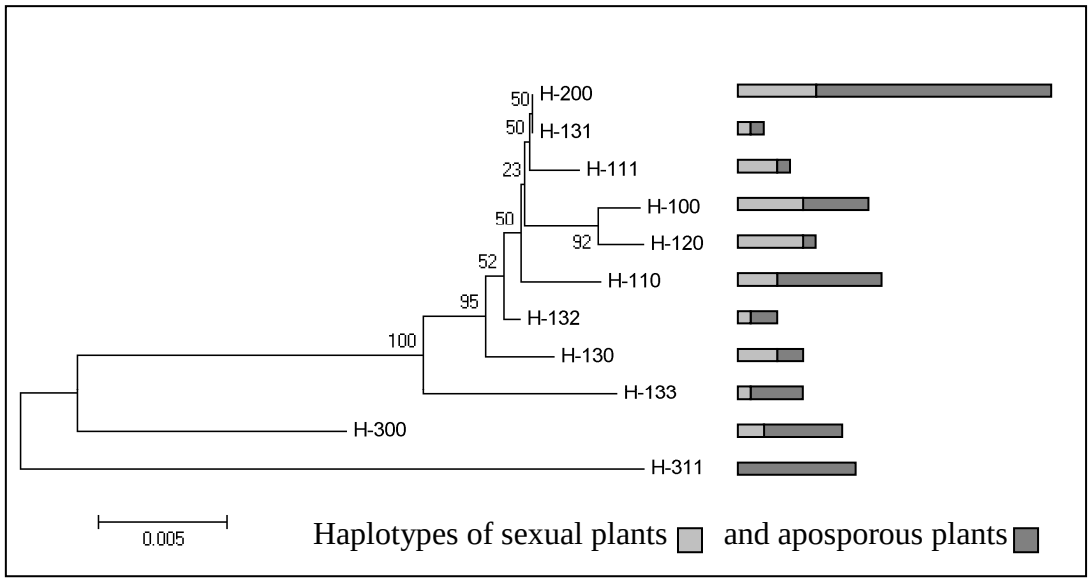


Figure 3

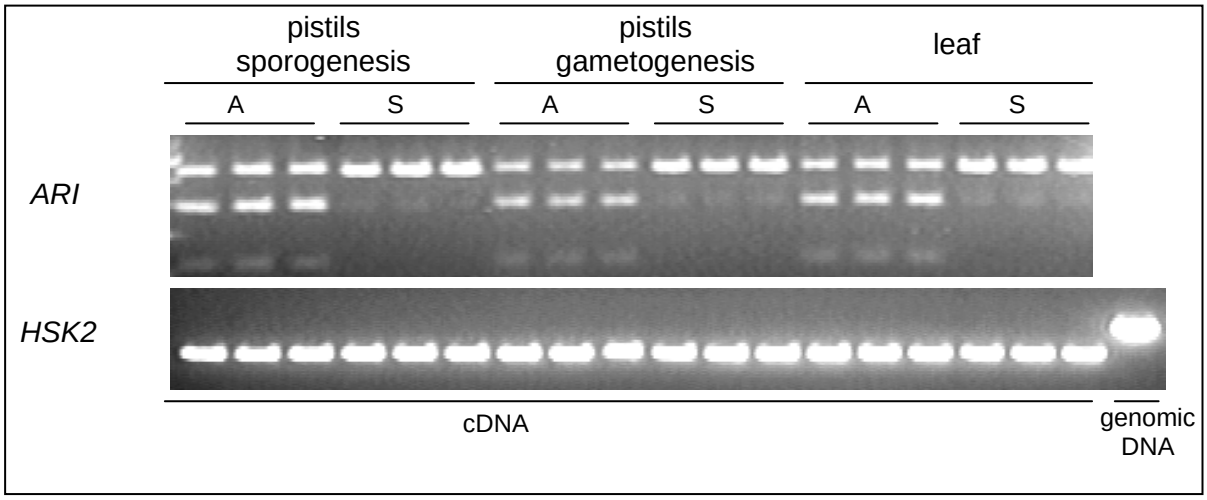
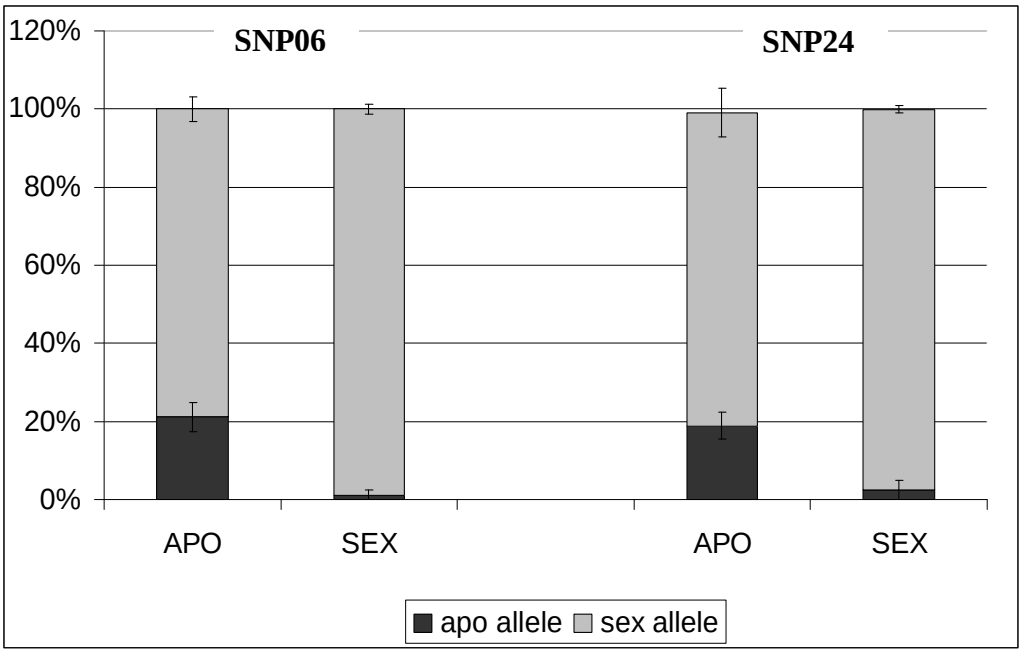
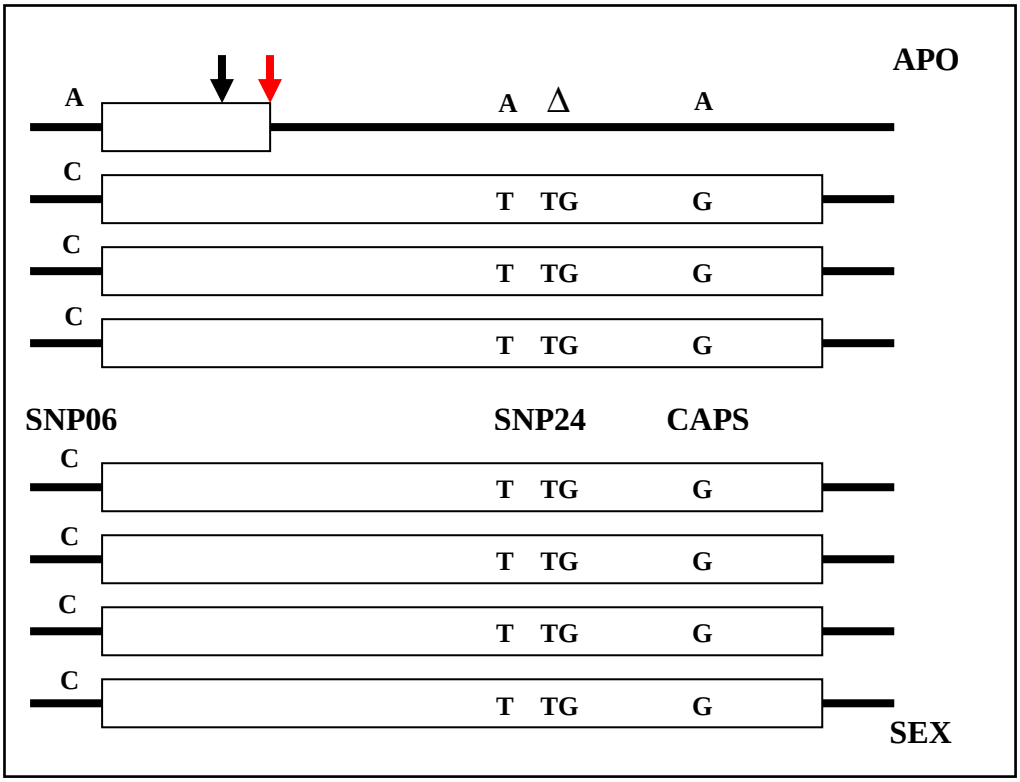

pistils
sporogenesis
pistils
gametogenesis
leaf
A
S
A
S
A
S
ARI
HSK2
cDNA
genomic
DNA

Figure 4


120%
100%
80%
60%
40%
20%
0%
SNP06
SNP24
APO
SEX
APO
SEX
apo allele
sex allele


APO
A
A Δ
A
C
T TG G
C
T TG G
C
T TG G
SNP06
SNP24
CAPS
C
T TG G
C
T TG G
C
T TG G
C
T TG G
SEX

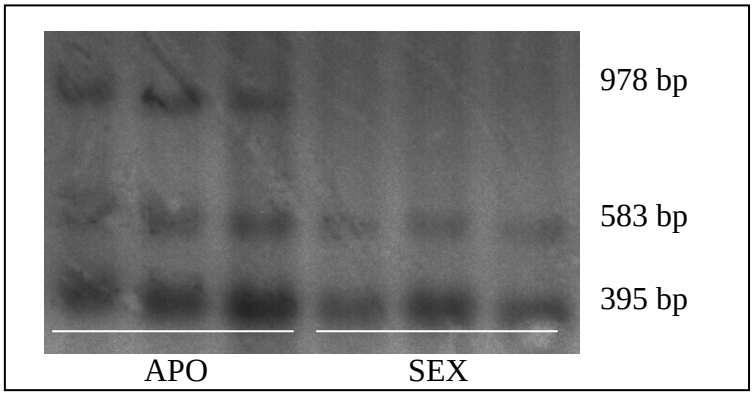

978 bp
583 bp
395 bp
APO
SEX

Figure 5

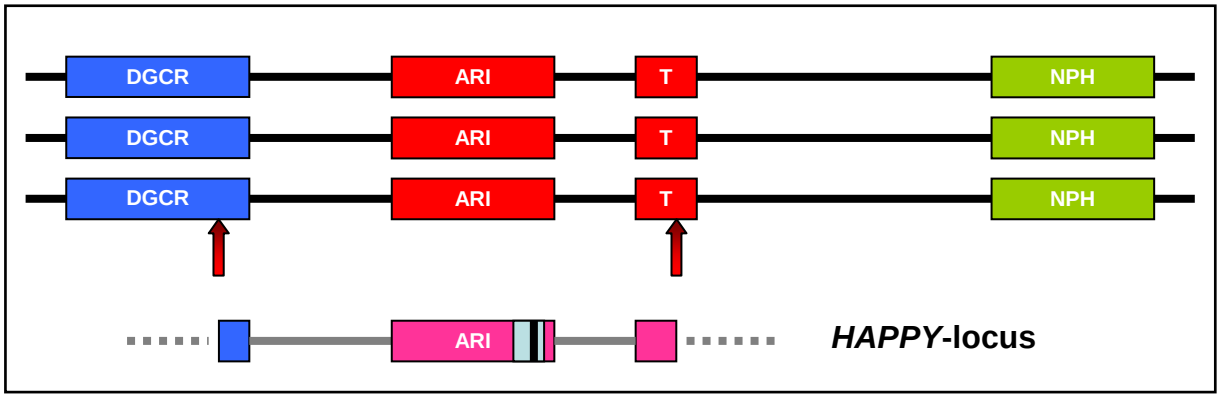


Figure 6

Plant code	Germplasm origin	Mode of reproduction	Ploidy
aTo	cv. Topaz (Poland)	apomictic	4n
aAn	Breeding stock (Germany)	apomictic	4n
aNo	Breeding stock (Germany)	apomictic	4n
aMü	Münster (Germany)	apomictic	4n
aSie	Siena (Italy)	apomictic	4n
aCa	Caen (France)	apomictic	4n
aWr	Wroclaw (Poland)	apomictic	4n
aWa	Warszawa (Poland)	apomictic	4n
aKs	Kaunas (Lithuania)	apomictic	4n
aV4	Vacrotot (Hungary)	apomictic	4n
sR1	Selection from cv. Topaz	sexual	2n
sP1	Padova (Italy)	sexual	2n
sP2	Padova (Italy)	sexual	2n
sV1	Vacrotot (Hungary)	sexual	2n
sV2	Vacrotot (Hungary)	sexual	2n
sV3	Vacrotot (Hungary)	sexual	2n

Line	Mode of reproduction (%)		
	Apospory	Parthenogenesis	Apomixis
aAn	95	86	85
aNo	100	~100	~100
aSi	88	81	81
aTo	83	80	78

Table S1

Plant Identifier	Apospory %	Parthenog. %	Marker state
Tetraploids			
4(sF$_1$1 x aAn)1/1	0	0	S
4(sF$_1$1 x aAn)1/2	0	0	S
4(sF$_1$1 x aAn)1/3	100	42	A
4(sF$_1$1 x aAn)1/4	0	0	S
4(sF$_1$1 x aAn)1/5	0	0	S
4(sF$_1$1 x aAn)1/6	0	0	S
4(sF$_1$1 x aAn)1/7	100	24	A
4(sF$_1$1 x aAn)1/8	0	0	S
4(sF$_1$1 x aSi)1/1	38	0 #	A
4(sF$_1$1 x aSi)1/2	0	0	S
4(sF$_1$1 x aSi)1/3	0	0	S
4(sF$_1$1 x aSi)1/4	0	0	S
4(sF$_1$1 x aNo)1/1	28	4	A
4(sF$_1$1 x aNo)1/2	0	0	S
4(sF$_1$1 x aNo)1/3	100	42	A
4(sF$_1$1 x aNo)1/4	0	0	S
4(sF$_1$1 x aNo)1/5	100	10	A
4(sF$_1$1 x aNo)1/6	100	30	A
4(sF$_1$1 x aNo)1/7	99	39	A
4(sF$_1$1 x aNo)1/8	100	18	A
4(sF$_1$1 x aNo)1/9	71	3	A
4(sF$_1$1 x aNo)1/10	100	31	A
4(sF$_1$2 x aAn)1/1	92	77	A
4(sF$_1$2 x aAn)1/2	92	42	A
4(sF$_1$2 x aAn)1/3	0	0	S
4(sF$_1$2 x aAn)1/4	0	0	S
4(sF$_1$2 x aAn)1/5	0	0	S
4(sF$_1$2 x aAn)1/6	0	0	S
4(sF$_1$2 x aAn)1/7	0	0	S
4(sF$_1$2 x aAn)1/8	0	6	S
4(sF$_1$2 x aAn)1/9	0	0	S
4(sF12 x aAn)1/10	100	24	A
4(sF$_1$2 x aNo)1a/1	99	46	A
4(sF$_1$2 x aNo)1a/2	0	0	S
4(sF$_1$2 x aNo)1a/3	0	0	S
4(sF$_1$2 x aNo)1a/4	0	0	S
4(sF$_1$2 x aNo)1a/5	0	0	S
4(sF$_1$2 x aNo)1a/6	0	0	S
4(sF$_1$2 x aNo)1a/7	98	53	A
4(sF$_1$2 x aNo)1a/8	96	0#	A
4(sF$_1$2 x aNo)1a/9	0	0	S
4(sF$_1$2 x aNo)1a/10	0	0	S
4(sF$_1$2 x aNo)1a/11	96	0#	A
4(sF$_1$2 x aNo)1a/12	29	0#	A
4(sF$_1$2 x aNo)1a/13	0	0	S
4(sF$_1$2 x aNo)1a/14	100	4	A
4(sF$_1$2 x aNo)1a/15	96	4	A
4(sF$_1$2 x aNo)1a/16	88	8	A
4(sF$_1$2 x aNo)1a/17	100	0#	A
4(sF12 x aNo)1a/20	100	4	A
4(sF12 x aNo)1a/24	100	0#	A
4(sF$_1$2 x aNo)1a/25	92	3	A
4(sF$_1$2 x aNo)1a/29	100	24	A

4($sF_1$2 x aNo)1a/31	100	0#	A
4($sF_1$2 x aNo)1a/33	100	4	A
4($sF_1$2 x aNo)1a/43	20	4	A
4(R1C2 x aSi)1a/1	29	0#	A
4(R1C2 x aSi)1a/2	0	0	S
4(R1C2 x aSi)1a/3	0	0	S
4(R1C2 x aSi)1a/4	0	0	S
4(R1C2 x aSi)1c/2	96	33	A
4(R1C2 x aSi)1c/3	24	0#	A
4(R1C2 x aSi)1c/7	0	0	S
4(R1C2 x aSi)1c/8	0	0	S
4(R1C2 x aSi)1c/9	0	0	S
4(R1C2 x aSi)1d/2	0	0	S
4(R1C2 x aSi)1d/4	0	0	S
4(R1C2 x aSi)1d/5	0	0	S
4(R1C2 x aSi)1d/6	0	0	S
4(R1C2 x aSi)1d/7	0	0	S
4(R1C2 x aSi)1d/8	0	0	S
4(R1C2 x aSi)1d/9	31	18	A
Triploids			
1(sR1 x aNo) 1/1	99	0#	A
1(sR1 x aNo) 1/2	99	4	A
1(sR1 x aNo) 1/3	0	0	S
1(sR1 x aNo) 1/4	100	3	A
1(sR1 x aNo) 1/5	98	2	A
1(sR1 x aNo) 1/6	100	3	A
1(sR1 x aNo) 2/1	0	0	S
1(sR1 x aNo) 2/2	0	0	S
1(sR1 x aNo) 2/3	100	4	A
1(sR1 x aNo) 2/4	95	0#	A
1(sR1 x aNo) 3/5	0	0	S
2(sR1 x aNo)1/2	97	13	A
2(sR1 x aNo)2a/3	100	0#	A
2(sR1 x aNo)2a/2	100	0#	A
2(sR1 x aNo)2a/7	100	33	A
2(sR1 x aNo)2a/14	100	0#	A
1(sV1 x aNo) 1/1	99	8	A
1(sV1 x aNo) 1/2	94	8	A
1(sV1 x aNo) 1/3	100	0#	A
1(sV1 x aNo) 1/4	100	0#	A
1(sV1 x aNo) 1/5	90	5	A
1(sV1 x aNo) 1/6	100	3	A
1(sV1 x aNo) 1/7	0	0	S
1(sV1 x aNo) 1/8	0	0	S
1(sV1 x aNo) 1/10	0	15	S
1(sV1 x aNo) 1/12	0	0	S
1(sV1 x aNo)1/13	100	3	A
1(sV1 x aNo) 1/17	100	7	A
1(sV1 x aNo) 1/19	100	7	A
1(sV1 x aNo) 1/20	95	7	A
1(sV1 x aNo) 1/22	100	11	A
1(sV1 x aNo) 1/23	0	0	S
1(sV1 x aNo) 1/25	100	3	A
1(sV1 x aNo) 1/27	100	0#	A
1(sV1 x aNo) 1/28	0	0	S

1(sV1 x aNo) 1/29	100	14	A
1(sV1 x aNo) 2a/1	100	0#	A
1(sV1 x aNo) 2b/2	100	7	A
1(sV1 x aNo) 2b/3	100	0#	A
1(sV1 x aNo) 2b/4	0	0	S
1(sV1 x aNo) 2b/5	100	0#	A
1(sV3 x aNo) 1/1	75	66	A
1(sV3 x aNo) 1/2	100	0#	A
1(sV3 x aNo) 1/4	98	3	A
1(sV3 x aNo) 1/6	100	12	A
2(sP2 x aTo)1/1	0	0	S
2(sP2 x aTo)1/2	100	0#	A
2(sP2 x aTo)1/11	100	66	A
2(sP2 x aTo)1/13	98	8	A
2(sP2 x aTo)1/14	100	3	A
2(sP2 x aTo)1/15	0	0	S
2(sP2 x aTo)1/16	95	14	A
2(sP2 x aTo)1/18	100	6	A
2(sP2 x aTo)1/19	95	0#	A
2(sP2 x aTo)1/20	0	0	S
2(sP2 x aTo)1/21	100	0#	A
2(sP2 x aTo)1/34	0	0	S
2(sP2 x aTo)1/26	100	5	A
2(sP2 x aTo)1/23	100	0#	A
Diploids			
2(P1 x sV1)1a/3	0	0	S
2(P1 x sV1)1a/4	0	0	S
2(P1 x sV1)1a/6	0	0	S
2(P1 x sV1)1a/7	0	0	S
2(P1 x sV1)1a/8	0	0	S
2(P1 x sV1)1a/9	0	0	S
2(P1 x sV1)1a/11	0	0	S
2(P1 x sV1)1a/13	0	0	S
2(P1 x sV1)1b/1	0	0	S
2(P1 x sV1)1b/6	0	0	S
2(P1 x sV1)1b/7	0	0	S
2(P1 x sV1)1b/8	0	0	S
2(P1 x sV1)1b/11	0	0	S
2(sR1 x sP2)1/3	0	0	S
2(sR1 x sP2)1/8	0	0	S
2(sR1 x sP2)1/9	0	0	S
2(sR1 x sP2)1/13	0	0	S
2(sR1 x sP2)1/16	0	0	S
2(sR1 x sP2)2/2	0	0	S
2(sR1 x sP2)2/3	0	0	S
2(sR1 x sP2)2/4	0	0	S
2(sR1 x sP2)2/8	0	0	S
2(sR1 x sP2)2/18	0	0	S
2(sR1 x sP2)2/25	7	0	S

Table S2

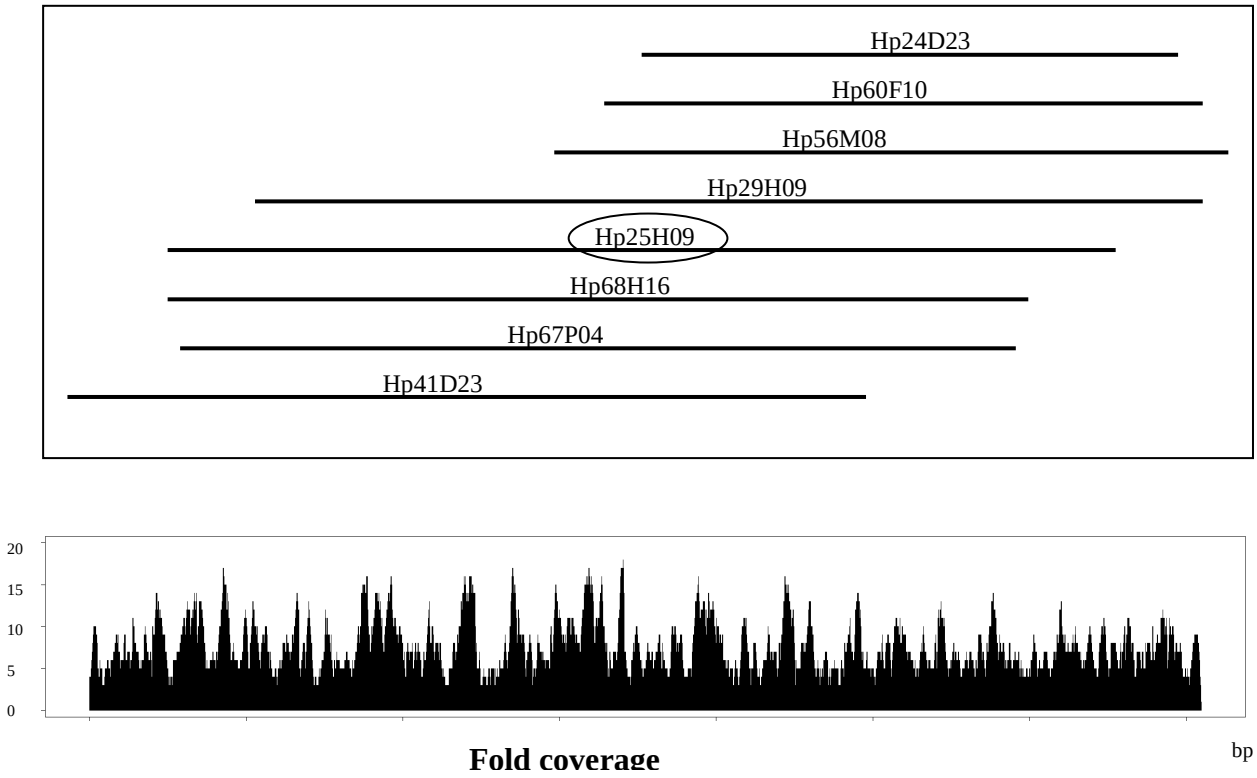

Hp24D23
Hp60F10
Hp56M08
Hp29H09
Hp25H09
Hp68H16
Hp67P04
Hp41D23
20
15
10
5
0
Fold coverage
bp

Figure S1

Amplified region name	BAC-position	Forward primer	Reverse primer	Fragment length
CAPS (*ARIADNE*)	63169-63391	TAAAGGAGTATCTGAAAAGTGAGG	CAGGCAGCACGAGAATCTA	222
Expression *ARIADNE*	61910-62109	TCCGACGAGGACGTGTATTA	TTATAATGACGGAGGGGATGCT	199
ARIADNE	61782 - 64755	GGGACCCAAAAACTCTCCTC	GCAGTGAACTCAACCGCATA	2973
ARIADNE promoter	61075 - 64755	TTTCTTCCTGTCCTTGTACCG	GCAGTGAACTCAACCGCATA	3680
ARIADNE- truncated	63169 - 65015	TAAAGGAGTATCTGAAAAGTGAGG	CCATGAACCAAGGCACAAC	1846
HK1	7538 - 8192	CAGGTGGAAAAGTTGGGAGA	TTTGAGCTCCACTCGTTCCT	654
HK2	31509 - 32162	CTGTTCTTGTCGGGCACATA	CTCATTAATGTGCGCTGCTG	653
GK1, GK2	37103 - 37718	AAGCTGCTTCTAGTTGTGCACTT	GTAGCGGCAGGAGCTATGAC	615
ExP1	78262 - 79000	GACTTTCGCGTACGTCCTGT	GAAGAAGCCGCAGAGGTAGA	738
PAT1	53041 - 53774	GGAGGTCAACGATTTCAGGA	CTGCACCTCACAGCCAGATA	733
WRKY	55854 - 56394	GCATTCTCGTGCTTCTCCTC	AAGAGCATGGATCGTGGAAC	512
DGCR	60011 - 60424	CGAGCGAGGATAATGAGAGC	CCCCTTCTCACCATCCCTAT	413
NPH3	68725 - 69315	GCATCGAGTCGTTAGCCTTC	TCTCCATGGTAGCCAACTCC	591
TK	75087 - 75512	GCAGCAACCTCGTGAACTTT	GATGGAAAGATCGTGGTCGT	425
RNP	81181 - 81613	GTTTATGCCCATCCCATGAC	GGAGGTTTGGCATCAACTGT	432
AUXR	89472 - 89907	ACGTGGAGCTTGGTCTTGTT	TTGCACCACCGTCACTTCTA	435
MO25	106317 - 106804	CTGGACGATCACCAATCTCA	CCATCGTTCTATCCCACACC	487
RING H-2	125473 - 125803	TCTCCTCTCCCGAAAGTTCA	GGTCTTCTCGGTGTCGTTGT	330

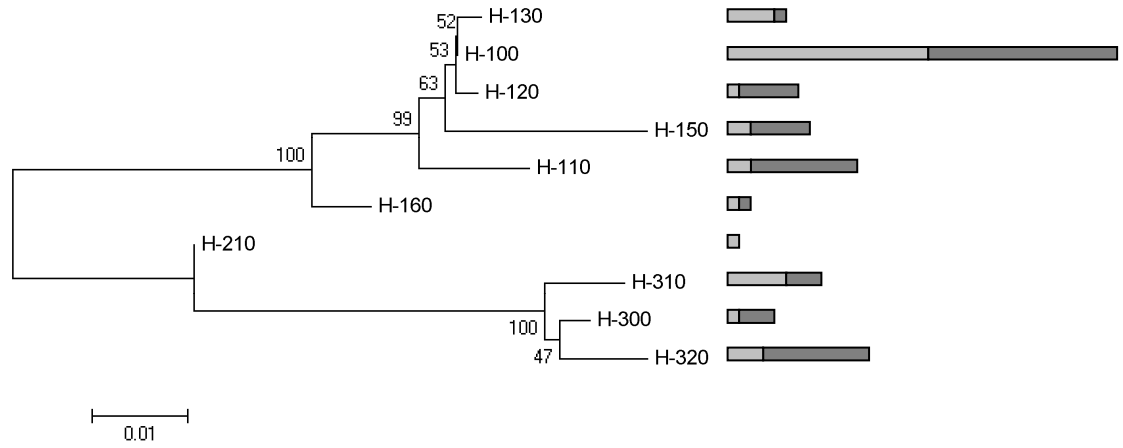
SNP no.	Forward primer	Reverse primer (biotinilated)	Sequencing primer
SNP 06	GGACCCAAAAACTCTCCTCTCA	CTCGTCGGAATCCATAGCTC	GCTTCTGGTAATTAGGGT
SNP 24, 25	TGGGACGTGAGTAAGTTGCATGAT	TGTCCACAAGAAGAAGCCGAATA	CGACGAGGAAGCAGTG

Table S3

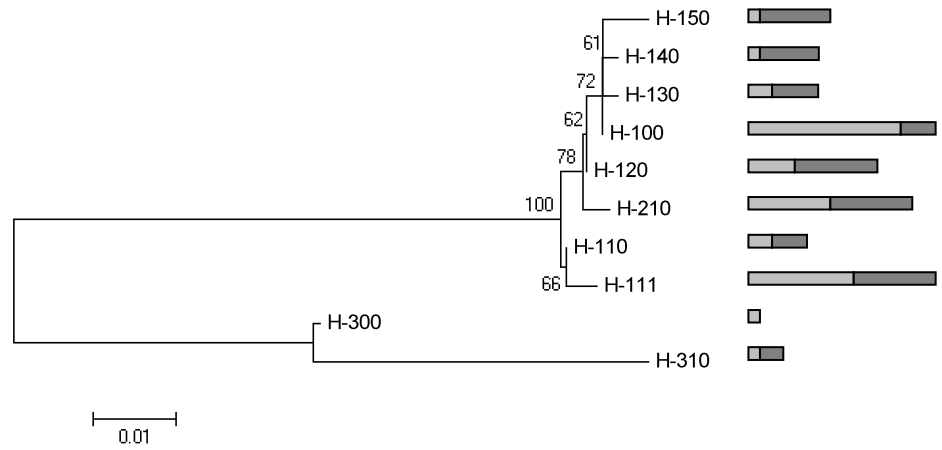
Symbol	Description	Best hit	E value	Arabidopsis
HK1,2	Histidine kinase AHK2	*Arabidopsis*	0.0	At5G35750
GH1, 2	Glycosyl hydrolase family protein 17	EF144611 *Populus*	1.00E-40	At5G35740
RT1	Ty-1 copia retrotransposon	AC151801 *Solanum demissum*	3.00E-23	-
RT2	Retrotransposon	AP006368 *Lotus japonicus*	2.00E-47	-
ExP1	Expressed protein	XM_002320476 *Populus*	3.00E-60	At2G04900
PAT1	PAT1 like protein	CR955006 *Medicago*	0.0	At5G48150
WRKY	WRKY related	*Arabidopsis*	1.00E-70	At2g04880
DGCR	hypothetical protein	AM478160 *Vitis vinifera*	1.00E-173	At3G07790
ARI	ARIADNE protein	*Arabidopsis*	0.0	At2G31510
ARI-T	ARIADNE protein (truncated)	XM_002310407 Populus	2.00E-69	At2g31510
NPH3	Similar to NPH3 family protein	*Arabidopsis*	0.0	At5G48130
TK	Thymidine kinase	AM492793 *Zea mays*	3.00E-187	At3G07800
RT3	Retrotransposon	AP008240 *Lotus japonicus*		-
RNP	Putative hnRNP	BT013750 *Lycopersicon*	2.00E-142	At2g31860
AUXR	Auxin-response related	AM478160 *Vitis vinifera*	6.00E-167	At2G04850
RimL	RimL/ Acetyltransferase domain protein	*Arabidopsis*	8.00E-55	At2G04845
ExP2	Membrane protein, acethyltransferase	AJ309300 *Solanum tuberosum*	3.00E-111	At3g07810
Mo25	Mo-25 like protein	BT013201 *Lycopersicon*	6.00E-116	At5g47540
TtRNAS	EMBRYO DEFECTIVE 2761) Thre-tRNA ligase	*Arabidopsis*	0.0	At2G04842
PG	Polygalacturonase	*Arabidopsis*	3.00E-123	At3g14060
RINGH2	RING-H2 finger protein	AM484170 *Vitis vinifera*	3.00E-76	At5G05810
RT4	Retroelement polyprotein	AC135396 *Orysa sativa*	7.00E-29	-

Table S4

DGCR



NPH3



Haplotypes of sexual plants and aposporous plants

Figure S2

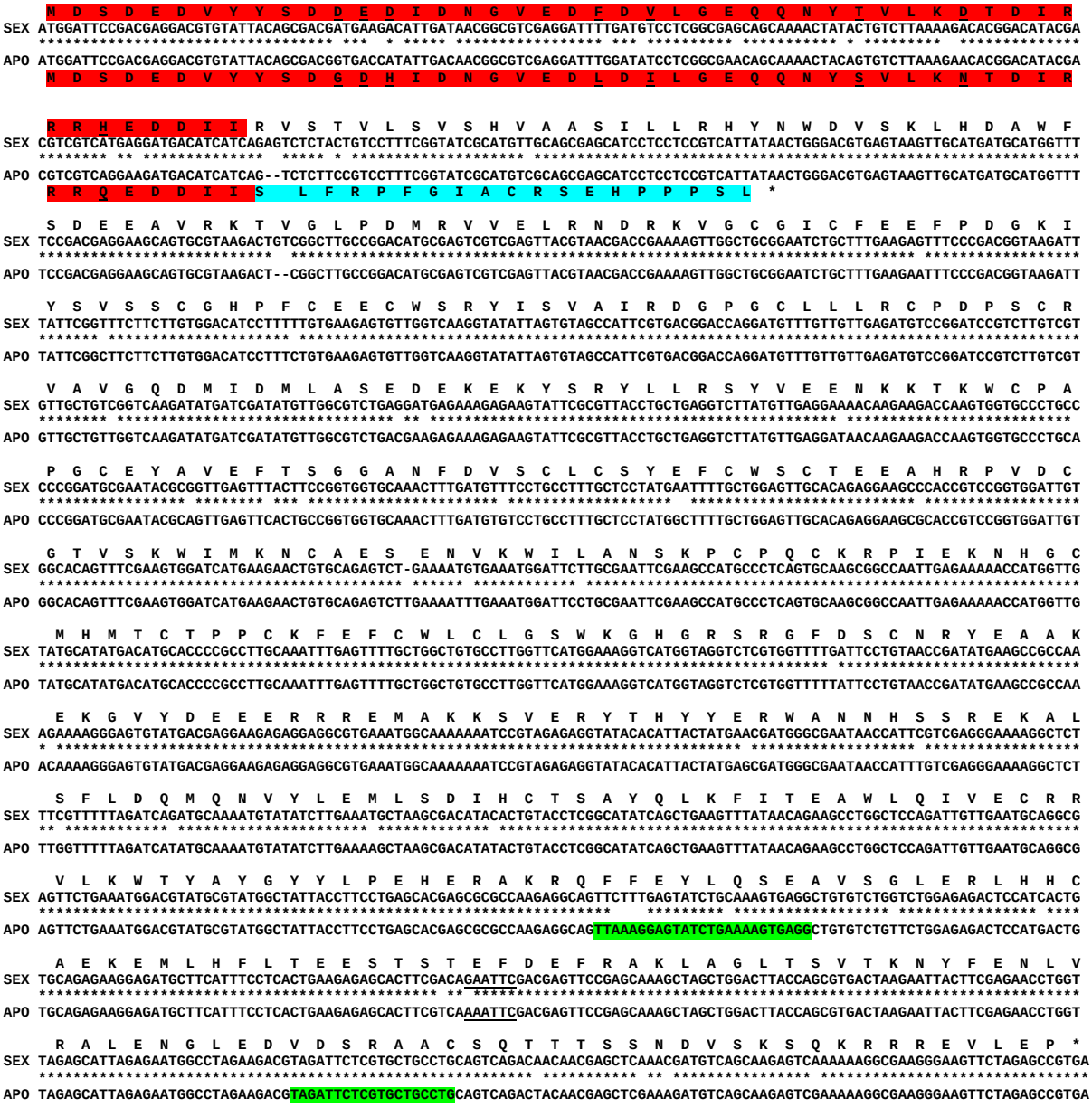
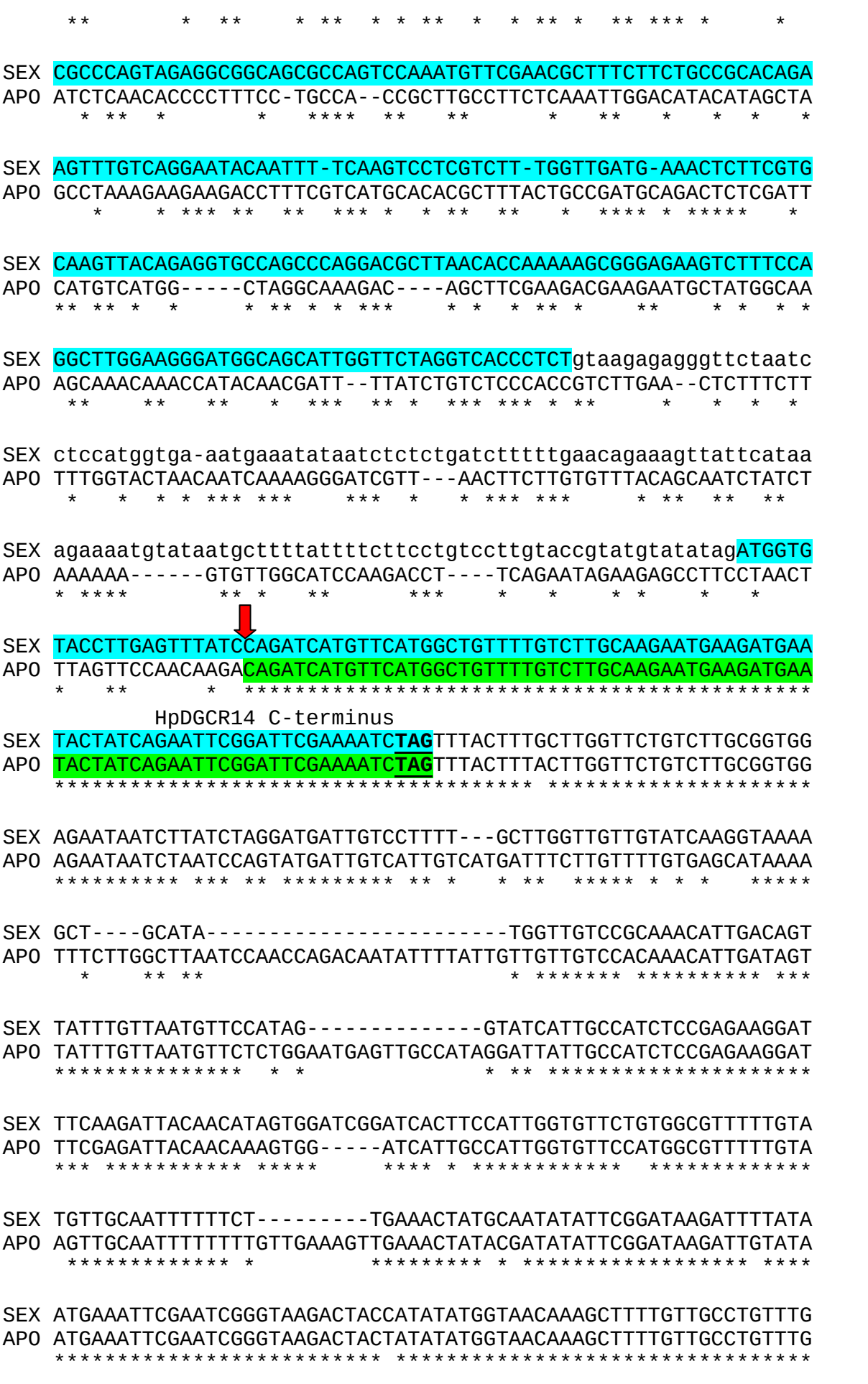

M D S D E D V Y Y S D D E D I D N G V E D F D V L G E Q Q N Y T V L K D T D I R
SEX ATGGATTCCGACGAGGACGTGTATTACAGCGACGATGAAGACATTGATAACGGCGTCGAGGATTTTGATGTCCTCGGCGAGCAGCAAAACTATACTGTCTTAAAAGACACGGACATACGA
APO ATGGATTCCGACGAGGACGTGTATTACAGCGACGGTGACCATATTGACAACGGCGTCGAGGATTTGGATATCCTCGGCGAACAGCAAAACTACAGTGTCTTAAAGAACACGGACATACGA
M D S D E D V Y Y S D G D H I D N G V E D L D I L G E Q Q N Y S V L K N T D I R
R R H E D D I I R V S T V L S V S H V A A S I L L R H Y N W D V S K L H D A W F
SEX CGTCGTCATGAGGATGACATCATCAGAGTCTCTACTGTCCTTTCGGTATCGCATGTTGCAGCGAGCATCCTCCTCCGTCATTATAACTGGGACGTGAGTAAGTTGCATGATGCATGGTTT
APO CGTCGTCAGGAAGATGACATCATCAG--TCTCTTCCGTCCTTTCGGTATCGCATGTCGCAGCGAGCATCCTCCTCCGTCATTATAACTGGGACGTGAGTAAGTTGCATGATGCATGGTTT
R R Q E D D I I S L F R P F G I A C R S E H P P P S L *
S D E E A V R K T V G L P D M R V V E L R N D R K V G C G I C F E E F P D G K I
SEX TCCGACGAGGAAGCAGTGCGTAAGACTGTCGGCTTGCCGGACATGCGAGTCGTCGAGTTACGTAACGACCGAAAAGTTGGCTGCGGAATCTGCTTTGAAGAGTTTCCCGACGGTAAGATT
APO TCCGACGAGGAAGCAGTGCGTAAGACT--CGGCTTGCCGGACATGCGAGTCGTCGAGTTACGTAACGACCGAAAAGTTGGCTGCGGAATCTGCTTTGAAGAATTTCCCGACGGTAAGATT
Y S V S S C G H P F C E E C W S R Y I S V A I R D G P G C L L L R C P D P S C R
SEX TATTCGGTTTCTTCTTGTGGACATCCTTTTTGTGAAGAGTGTTGGTCAAGGTATATTAGTGTAGCCATTCGTGACGGACCAGGATGTTTGTTGTTGAGATGTCCGGATCCGTCTTGTCGT
APO TATTCGGCTTCTTCTTGTGGACATCCTTTCTGTGAAGAGTGTTGGTCAAGGTATATTAGTGTAGCCATTCGTGACGGACCAGGATGTTTGTTGTTGAGATGTCCGGATCCGTCTTGTCGT
V A V G Q D M I D M L A S E D E K E K Y S R Y L L R S Y V E E N K K T K W C P A
SEX GTTGCTGTCGGTCAAGATATGATCGATATGTTGGCGTCTGAGGATGAGAAAGAGAAGTATTCGCGTTACCTGCTGAGGTCTTATGTTGAGGAAAACAAGAAGACCAAGTGGTGCCCTGCC
APO GTTGCTGTTGGTCAAGATATGATCGATATGTTGGCGTCTGACGAAGAGAAAGAGAAGTATTCGCGTTACCTGCTGAGGTCTTATGTTGAGGATAACAAGAAGACCAAGTGGTGCCCTGCA
P G C E Y A V E F T S G G A N F D V S C L C S Y E F C W S C T E E A H R P V D C
SEX CCCGGATGCGAATACGCGGTTGAGTTTACTTCCGGTGGTGCAAACTTTGATGTTTCCTGCCTTTGCTCCTATGAATTTTGCTGGAGTTGCACAGAGGAAGCCCACCGTCCGGTGGATTGT
APO CCCGGATGCGAATACGCAGTTGAGTTCACTGCCGGTGGTGCAAACTTTGATGTGTCCTGCCTTTGCTCCTATGGCTTTTGCTGGAGTTGCACAGAGGAAGCGCACCGTCCGGTGGATTGT
G T V S K W I M K N C A E S E N V K W I L A N S K P C P Q C K R P I E K N H G C
SEX GGCACAGTTTCGAAGTGGATCATGAAGAACTGTGCAGAGTCT-GAAAATGTGAAATGGATTCTTGCGAATTCGAAGCCATGCCCTCAGTGCAAGCGGCCAATTGAGAAAAACCATGGTTG
APO GGCACAGTTTCGAAGTGGATCATGAAGAACTGTGCAGAGTCTTGAAAATTTGAAATGGATTCCTGCGAATTCGAAGCCATGCCCTCAGTGCAAGCGGCCAATTGAGAAAAACCATGGTTG
M H M T C T P P C K F E F C W L C L G S W K G H G R S R G F D S C N R Y E A A K
SEX TATGCATATGACATGCACCCCGCCTTGCAAATTTGAGTTTTGCTGGCTGTGCCTTGGTTCATGGAAAGGTCATGGTAGGTCTCGTGGTTTTGATTCCTGTAACCGATATGAAGCCGCCAA
APO TATGCATATGACATGCACCCCGCCTTGCAAATTTGAGTTTTGCTGGCTGTGCCTTGGTTCATGGAAAGGTCATGGTAGGTCTCGTGGTTTTTATTCCTGTAACCGATATGAAGCCGCCAA
E K G V Y D E E E R R R E M A K K S V E R Y T H Y Y E R W A N N H S S R E K A L
SEX AGAAAAGGGAGTGTATGACGAGGAAGAGAGGAGGCGTGAAATGGCAAAAAAATCCGTAGAGAGGTATACACATTACTATGAACGATGGGCGAATAACCATTCGTCGAGGGAAAAGGCTCT
APO ACAAAAGGGAGTGTATGACGAGGAAGAGAGGAGGCGTGAAATGGCAAAAAAATCCGTAGAGAGGTATACACATTACTATGAGCGATGGGCGAATAACCATTTGTCGAGGGAAAAGGCTCT
S F L D Q M Q N V Y L E M L S D I H C T S A Y Q L K F I T E A W L Q I V E C R R
SEX TTCGTTTTTAGATCAGATGCAAAATGTATATCTTGAAATGCTAAGCGACATACACTGTACCTCGGCATATCAGCTGAAGTTTATAACAGAAGCCTGGCTCCAGATTGTTGAATGCAGGCG
APO TTGGTTTTTAGATCATATGCAAAATGTATATCTTGAAAAGCTAAGCGACATATACTGTACCTCGGCATATCAGCTGAAGTTTATAACAGAAGCCTGGCTCCAGATTGTTGAATGCAGGCG
V L K W T Y A Y G Y Y L P E H E R A K R Q F F E Y L Q S E A V S G L E R L H H C
SEX AGTTCTGAAATGGACGTATGCGTATGGCTATTACCTTCCTGAGCACGAGCGCGCCAAGAGGCAGTTCTTTGAGTATCTGCAAAGTGAGGCTGTGTCTGGTCTGGAGAGACTCCATCACTG
APO AGTTCTGAAATGGACGTATGCGTATGGCTATTACCTTCCTGAGCACGAGCGCGCCAAGAGGCAGTTAAAGGAGTATCTGAAAAGTGAGGCTGTGTCTGTTCTGGAGAGACTCCATGACTG
A E K E M L H F L T E E S T S T E F D E F R A K L A G L T S V T K N Y F E N L V
SEX TGCAGAGAAGGAGATGCTTCATTTCCTCACTGAAGAGAGCACTTCGACAGAATTCGACGAGTTCCGAGCAAAGCTAGCTGGACTTACCAGCGTGACTAAGAATTACTTCGAGAACCTGGT
APO TGCAGAGAAGGAGATGCTTCATTTCCTCACTGAAGAGAGCACTTCGTCAAAATTCGACGAGTTCCGAGCAAAGCTAGCTGGACTTACCAGCGTGACTAAGAATTACTTCGAGAACCTGGT
R A L E N G L E D V D S R A A C S Q T T T S S N D V S K S Q K R R R E V L E P *
SEX TAGAGCATTAGAGAATGGCCTAGAAGACGTAGATTCTCGTGCTGCCTGCAGTCAGACAACAACGAGCTCAAACGATGTCAGCAAGAGTCAAAAAAGGCGAAGGGAAGTTCTAGAGCCGTGA
APO TAGAGCATTAGAGAATGGCCTAGAAGACGTAGATTCTCGTGCTGCCTGCAGTCAGACTACAACGAGCTCGAAAGATGTCAGCAAGAGTCGAAAAAGGCGAAGGGAAGTTCTAGAGCCGTGA

Figure S3

Comparison between APO and SEX allele at the 5´border

```
SEX GGAGAGGGTCGGGACGATGGAGCGGCCGGGATTGATGAGTCGTTGGGGTTGGACGA-GTT
APO ------GATCCT-ACCTTTGAGCTCTCAGTTGTAA-GAGATGTCCGGTCTAACTGATGTT
           * **   **  * ****   * *   * * ***  **  **  *    ** ***

SEX CT----TTAGGAGGTACACGAGCGAGGATAATGAGAGCTTTTCGAAGATTTTAGATAAGG
APO CTGATTCCCAACTTTTCTCCAGTAAAGGAAATGCCAGGTTACTGATAACGCTTG-TAGAG
    **             * * * **  * *  ****  ** **   **  *   * * **  *

SEX ATAATAGGAAGAGGAAAGAGCGGAACGTACATTATTTATTGGAAGGGGAGAGGGAGTTGA
APO ATCCCATCCAGTAGAGTGAT--GGAGGTTCAGTATAAATATTAATTAACAGTGGATATAG
    **   *   **  **  **   * * ** ** ***  **   **        ***  *

SEX TTGGGGATGCGAAGAG--GGATAGGATAACTGATGGGTACGGGAC---TTCTGATCAGCC
APO CTAAGGCAGCACATATCCAGATAGGAGCGCTGTGGTGTATTGGGTACTCTTTCTTTTATC
     *  **  **  * *    *******   ***  * ***  **      * *  *    *

SEX AGTTAGCACGCTGGATGGGTGGAAGTTTAGGG--CGGATAATTTGTTGATGTATCACCCT
APO GATGAAAACCCTTCAAAGGTTTAAGGGTAGGAGTCAAAAAATTCATAGA--TATCCCCCT
      * *  ** **  *  ***  ***  ****   *  * ****  * **  **** ****

SEX GTTAAGGAGGCGCCTCTGACTGACGAGGATAGGGAACTTCGGTTGAAGGCATTAACAAAA
APO CCCGAT------CTTGGGATGGATGAGACTT--GACTTATGATTTTTTCCTCTTATGGTG
        *       * *  **  ** ***  *   **  *  * **     *  * *

SEX GAG--ATCACCAAGGGGAACACGCGTTTTCATGGGAAG-ATGTCGGATACAAGGCCTGTG
APO GAGTGAAAATGGTATTGGACTCAAGTTTTGGGGGACTCCATGACGCTTTAATAGCTTTAA
    ***  *  *       * ** *  *****   **     *** **  *  *  ** *

SEX GAAGATGGTGGAGTTGAGGTGCTATATGGTCCTCTAGCTGGGGCCACTCC--AGTGCCCT
APO AAAGCC--TCATATCAA--TGCTAGCAG--CGTCAAACTAAGACGCTTCAGAAGAGACCA
     ***    *    *  *  *****   *  * ** * **  * *   **   ** * **

SEX TTGATAGGGATGGTGAGAAGGGGAAGAAGTATG-ATTTAGAGGATTTGAAGAGGACGCCT
APO CTACCTGCATTCATCCGGACAGGCTAACCTAAGCATACCAAGGA----ACGACGCGACCG
     *    *   *  *  * *  **   *  ** * **    ****    * ** *   **

SEX AATTTGTTTTACGTGGAGTCTCAGAAGAAGGCTGATAATGGGTACAGTTTTGTGAAGACG
APO AACCAGCTCACTAATGCGACTCCAGACAATCC----ATTTTTTACCCCTTCGTTTTAAGA
    **   * *       * * ***   * **  *    * *   ***   ** **    *

SEX CCGTCACCTGCTCCTGGGGCTGATGAATCACCTTTTATTACTTGGGGTGAAATCGAAGGG
APO CGGTTCTCTTCGCCCCTGACTTCCAAAAAA--GTATGACACTTCAGATACGACCTTGTGC
    * **   ** * **   * **    **  *   * *   ****  * *   * *    *

SEX ACGCCTTTAAGGTTG--GAGTCTGAAGATGCTCCTTTCGGGATAGGAAGCAGCTCTGATG
APO ATGTCTCTGACTTGGTTGAGTTCAGCTTTTCACCAACCAGGTTAGACA-----TCTGCCG
    * * ** * *  * *  ****       * * **   * ** ***  *     ****  *

SEX GCCCTCATTTTCAGATGCCAAACCCACCCGTACGAGATGTGAAGGCACATTCCTTATCCA
APO GCTACGCCGCACTGAAGGAAACTTCACTCGTCCTTCGAATCGACCTATCTATACTGCCTT
    **         * ** *  **   *** *** *      *  *   *  *    *  *

SEX GGGAGGCTGCAAAGAAGTTGAGAGAGAGGTCAAGGATGTTCCAGAAGCCTCCTTTGCCCT
APO CGGGTATCAGACCGACAAAGGGAATGCG-TCGGGTTTCTTTCCAAAACCTTCCGACACGG
```


HpDGCR14 C-terminus


SEX TTCTCGAACTCTATTTACGAAATAATTTTTTCAATAATATTTTTTAAATTTGTAAATGAT
APO TTCTCGAATTCTATTTACAAAATACTTCTTCCGATAATATTTTTAAAATTTGTATTTGAT
SEX AATCATGAAACCTAAGATATTTTCCTCTAAACACGAAACACAGAACCCTTTGGAGGATAT
APO A-TCATGAAACATAAGATATTTTCCTCTAAACACGAAACACAGAACCCTTTGGAGGATAT
SEX AAGTCTACAGAGAATTTTTTCACCAAAAACATAATATTCGATTCCTAAAAAAA-AAGAAA
APO AAGTCTACAGAGAATTTTTCCACCAAAAACATAATATTCGATTCTAAAAAAAATAAAAAA
SEX AATCCCAAACCTAGAAGTATAAGGTATACCGAAAATGCCCCCACCTAACCTATATTAATT
APO AATCCCAAACTTAGAAGTATGAGGAATACCGAAAATGCCCCCACCTAACCTATATTAATT
SEX AATAGGGACCCAAAAACTCTCCTCTCAGAAGTCAGAACAAATCTCCTCCCCTATTATATT
APO AATAGGGACCCAAAA-CTCTCCTCTCAGAAGTCAGAACAAATCTCCTCCCCTATTATATC
SEX ATTCCTCTCTAACAGGAACACCTTAGCTTCTGGTAATTAGGGTTTTCCTCTGTGAGTGTG
APO ATTCCTTTCTCACGGGAACACCCCAGCTTCTGGTAATTAGGGTTTTACTCTGTGAGTGCG
HpARI N-terminus
SEX TGAGCTATGGATTCCGACGAGGA...
APO TGAGCTATGGATTCCGACGAGGA...

Figure S4

Comparison between APO and SEX allele at the 3´border

```
SEX  ----GCAGCCAACTTG--TAGGCCATGCCCCCGTATGCGAATATGCGGTTGAGTTCACTG
APO  AATAACAATCAAATCACTTTAGTCTTTCACACA-ATACATATACGGGGT--ATCACATTT
           **  *** *    *  * * * * * *  ** *  *** * ***  *   ** *

SEX  CCGGTGGTGCAAACTTTGATGTTTCCTGCCTT-TGCTCCCGTGGCTTTTTGCTGGAGTTG
APO  TCAAA-----TTACTTAAACCTTTCACACAATACATATACGGGTATCATATTTGGAGTTG
      *          ****  *  ****   *  *       ** *  *  *   ********

SEX  CA--GAGGAAGCCACCGCCCTGTGGATTGTGGCACAATTTCGAAGTGGATCATGAAGAAC
APO  CACAAAGGAAACCACCGCCCTATGGATTGTGGCACAATTTTGAAGTGGATCATGAAGAAC
     **   ***** ********** ****************** ********************

SEX  TGCACAGAGTCTGAAAA-TGTGAATTGGATTCTGGCGAATTTGAAGCCATGCCCGAGGCC
APO  TATGTAGTCTCTAAAAAATGTGAATTGGATTCTAGCAAATTCAAAGCCATGCCCAAGGCC
     *    **  *** **** ************** ** ****  ********** *****

SEX  AAATGAGAAAACTCATGGGTGTATGCATATGACATGCACCCTGCCTTGCAACTTTGAGTT
APO  AATTGCGAAAACTCATGTGTGTATGCATATGACATGCACCCTGCCTTGCAAATTTGAGTT
     ** ** ********** ********************************** ********

SEX  TTGCTGGTTGTGCCTTGGTTCATGGCAAGAGCATCGTGAGCGGACGGGTAGTTTTTATGC
APO  TTGCTAGTTGTGCCTTAGTTCATGGCAAGAGCATCGTGAGTAGACCAGTAGGTTTTATGC
     ***** ********** ************************  ***  **** ********

SEX  CTGTAACGGGTATGTAGTGGCCAGACAAGTAGTAGTGTTCGACGAGGCAGAGAGGACACG
APO  ATGTAATGGTTATGGAGTCGCCAGACAAGTAATAGTATACGACGAGGTAGAGAGGACATG
      ***** ** **** *** ************ **** * ******** ********** *

SEX  CGAAATGGTG----------AGGTATACACAAGACTATGAACGATGAGCTTGTAACGAT
APO  GAAAATGGTGCAAACCCTTAGAGGTATACACATTACTATGAATGATGAGTTTGTAACGAT
       ********           ***********  ********* ****** **********

SEX  TCGTCGAGGCGAAAAGCATCGTCGAATGCAGGCGAGTTCTGAAATGGACGAATGTGTACG
APO  TCGTCGAGGCAAAAAGCTCTTTTAGATTGACATCAGAAGAAAAATGTACATCTCAAAATG
     ********** ******    *   **  *    **     ***** **   *    * *

SEX  GC-TATTAC-CTTCCCGATCACGAGCATGCTAAGAAGCAGTTCTTTGATTATGTGCAAGG
APO  CTATATGACACATACCATACCTCTGCATATCAGGCT-TAGATTTATGACAGTAGCTTGAC
        *** ** * * **   *    ****   * *    ** * * ***   *

SEX  TTGAGGCTTAGTCTAGTCTGGAGAGACTCCATCAATGTGCAGAGAATGAGCTGTTACATT
APO  TCCAGATT--GTTGAATATAG-GCGAGTTCTGAAATG---GAAAAATATGTACGAGTATT
     *  **  *  **  * * * * * ** * *   ****     * ***  *       ***

SEX  TCCTTAGTGAAGAGTGCACTTCAAAAGATATTGTTTTACTGAATAACTTTGTTTTGATTT
APO  ACTGTCACGAC--CCAAATCTGGGTTGTTAATGCGGTAACTGATAAATCTTGACCCAACT
      *  *   **       *  *     * ** **   **    **** * *      *  *

SEX  TTTTTTTTATATATATGACATTTCGTTAATTGTTCTGGTCAGTTGAATTTGATTAATAAA
APO  AGCCTTATACGAACAT-ACATAATGCAAAGTATAG-----AAATATACATAAATATGGAA
       ** **   * ** ****   *  ** * *       *  *  *  * * **   **

SEX  AACCGTAATTTTTTCTATTTTTTATATCTTGTCTATTTTGATACTGGAAAAATTGTTCTA
```

```
APO G--CGTA-TTTCATATATATAT-ATATATATATTACATATACATT---ACAATCTTCTTA
       **** ***  * *** * * **** *    **  *  * * *   * ***  *  **

SEX GAAAAAATTATGATTAAGAGGTTAATTGGGTAGGCGACCTTGGTCACATTCGTTTAATAT
APO TTCAAGAT----ACTAAAA--TTCACCAGCCGGGC--CGTCGACCACGCGTGCCCTATAG
       ** **    * *** *  ** *   *   ***  * * *  ***    *    ***

SEX TTATTGACTGCTAAATTTAAATT
APO TAA------GCCGAATTCCAGCA
    * *      **  ****  *
```

Figure S5

Ploidy of	Apomixis	Recombinants		Sexuals	
progenies	Apo^+/Par^+	Apo^+/Par^-	Apo^-/Par^+	Apo^-/Par^-	Total
3x	28	16	1	14	59
4x	27	7	1	37	72
Overall	55	23	2	51	131
Ploidy of	**Overall**		**Chi-square values**		
progenies	**$Apo^{(+ vs -)}$**	**$Par^{(+ vs -)}$**	**$Apo^{(+ vs -)}$**	**$Par^{(+ vs -)}$**	**Overall $^{(Apo vs Par)}$**
3x	44 vs. 15	29 vs. 30	9,557 *	0,017 ns	24,864 ***
4x	34 vs. 38	28 vs. 44	0,222 ns	3,556 ns	47,333 ***
Overall	78 vs. 53	57 vs. 74	4,771 ns	2,206 ns	57,061 ***
Symbols: Apo=Apospory; Par=Parthenogenesis ns, not significant at *P*=0,01; *, significant at *P*=0,01; *** significant at *P*=0,001					
Recombination frequency: 19.08%					
Genetic distance (Kosambi mapping function): 20.1 cM					

Table S5

Locus Name	Gene ID	Sequence Similarities	Availability of ESTs and cDNAs (tissue type)	T-DNA mutant allele	Genomic location of insertion	Mutant plant phenotype
ARI7	At2G31510	ARI5, ARI8, ARI9, ARI10, ARI11	yes (callus)	SALK_027620 SALK_073106 SALK_116408 SALK_025210 SALK_082541 SALK_132004 SALK_029122 SALK_069374	intron intron intron intron intron exon 3' UTR 3' UTR	none observed
ARI5	At1G05890	ARI7 ARI8 ARI9 ARI10 ARI11	yes (vegetative tissues)	Not isolated yet	NA	NA
ARI8	At1G65430	ARI7 ARI11 ARI8 ARI9 ARI10	yes (vegetative tissues)	Not isolated yet	NA	NA
ARI9	At2G31770	ARI10, ARI11, ARI7, ARI5, ARI8	yes (vegetative tissues)	SALK_027269 SALK_049257	exon intron	none observed
ARI10	AT2G31760	ARI9, ARI11, ARI7, ARI5, ARI8	yes (vegetative tissues)	SALK_014075 SALK_068103	exon promoter	None observed Unlinked phenotype with a gametophytic mutation carrying 51% infertile ovules (N=493); ovules arresting at one and two-nucleated stages
ARI11	AT2G31780	ARI9, ARI10, ARI7, ARI5, ARI8	yes (vegetative tissues)	Not isolated yet	NA	NA

Table S6