Grundlegende Untersuchungen zur Genetik, Züchtung und Blütenorganidentität von *Calluna vulgaris* (L.) Hull

Basic research into genetics, breeding and flower organ identity of *Calluna vulgaris* (L.) Hull

Von der Naturwissenschaftlichen Fakultät der
Gottfried Wilhelm Leibniz Universität Hannover
zur Erlangung des Grades Doktor der Gartenbauwissenschaften Dr. rer. hort.
genehmigte Dissertation von

Dipl.-Ing. (FH) Thomas Borchert geb. Schmidt
geboren am 09.01.1980 in Friedberg (Hessen)

Grundlegende Untersuchungen zur Genetik, Züchtung und Blütenorganidentität von *Calluna vulgaris* (L.) Hull

Basic research into genetics, breeding and flower organ identity of *Calluna vulgaris* (L.) Hull

Von der Naturwissenschaftlichen Fakultät der
Gottfried Wilhelm Leibniz Universität Hannover
zur Erlangung des Grades
Doktor der Gartenbauwissenschaften
Dr. rer. hort.
genehmigte Dissertation
von

Dipl.-Ing. (FH) Thomas Borchert geb. Schmidt
geboren am 09.01.1980 in Friedberg (Hessen)

(2009)

Referent: Prof. Dr. Thomas Debener

Korreferent: Prof. Dr. Traud Winkelmann

Tag der Promotion: 24. August 2009

Zusammenfassung

Calluna vulgaris (L.) Hull ist ein für den deutschen und europäischen Gartenbau bedeutendes Ziergehölz, bei dem vorrangig der Phänotyp der Blüte im Fokus der Konsumenten liegt. Als wirtschaftlich wichtigste Blütenform gelten die sogenannten „Knospenblüher", deren Blüte im Gegensatz zum Wildtyp keine Antheren ausbildet und sich nicht öffnet. Dies führt zu einer Erhaltung der Blütenfarbe und somit der Attraktivität der Blüten bis in den Winter. Das Merkmal „Knospenblütigkeit" stellt daher seit einigen Jahrzehnten das herausragende Zuchtziel für diese Art dar. Die hauptsächlich angewendeten Züchtungsmethoden der Rückkreuzung und der Sport-Selektion führten zu einer abnehmenden phänotypischen Differenzierbarkeit von Genotypen und Sorten. Hieraus wiederum resultierten zunehmend juristische Auseinandersetzungen bezüglich der Frage der Sortenableitung und des -schutzes.

In der vorliegenden Arbeit wurde daher als Grundlage weiterer Züchtungsvorhaben mit PCR-basierten „fingerprinting"-Methoden (RAPDs, ISSRs) für eine wissenschaftlich und wirtschaftlich relevante Auswahl von 74 *C. vulgaris* Genotypen die vorhandene Diversität auf Genomebene geschätzt. Es konnte ein hohes Maß an Übereinstimmung (> 0.74 Dice-Index) der getesteten Genotypen nachgewiesen werden. Ferner wurde die Anwendbarkeit eines Systems zur Identifizierung abgeleiteter Sorten geprüft und für die Verwendung in *C. vulgaris* erfolgreich modifiziert. Die Erstellung von spaltenden Rückkreuzungspopulationen diente der weiteren Aufklärung der vermuteten monogen-rezessiven Vererbung des Merkmals „Knospenblütigkeit". Molekulare Untersuchungen in diesen Populationen wurden im Hinblick auf die Etablierung eines markergestützten Selektionssystems für dieses Merkmal durchgeführt. Mit ‚bulked segregant analysis' wurden zwei merkmalsgekoppelten DNA-Segmenten (RAPD) identifiziert, mit denen die Selektion von Knospenblühern in der Ausgangspopulation möglich ist. Die Sequenz dieser Fragmente wurde untersucht und diente als Grundlage zur Etablierung von SCAR- und SSCP-Markern. In Bezug auf die Blütenbiologie der „Knospenblüher" fehlten bisher Erkenntnisse zur Blütenorganidentität sowie zur Ursache des Antherenverlusts. Daher wurden detaillierte mikroskopische, histologische und molekularbiologische Untersuchungen des normal- sowie des knospenblütigen Phänotyps durchgeführt, die darauf hindeuten, dass bei Knospenblühern die Petalen in einen zusätzlichen Kreis petaloider Sepalen umgewandelt sind, während die Antheren nicht in andere Organe umgewandelt wurden, sondern vollständig fehlen. Mit Hilfe von RACE-PCR wurden partielle Sequenzen zweier unterschiedlicher MADS-box-Transkriptionsfaktoren der Klassen B (*AP3*/*DEF*-like) und E (*SEP1*-like) gewonnen. Diese Daten wurden für quantitative Genexpressionsanalysen in Blütenorganen mit qRT-PCR genutzt. In Kombination mit den morphologischen Erkenntnissen stellt die Gesamtheit der gewonnenen molekularen Daten die Grundlage und die Voraussetzung für die weitere gezielte züchterische Bearbeitung dieser Kultur und des Merkmals „Knospenblütigkeit" im Speziellen dar.

Summary

Calluna vulgaris (L.) Hull is an important ornamental crop for German and European horticulture, and consumers focus primarily on the phenotype of the flowers. In contrast to the 'wild-type' phenotype, the 'bud-flowering' types do not develop stamens and remain closed as buds. This phenomenon means that the plants maintain their colorful flowers, and hence remain attractive, into the winter months and therefore, the 'bud-flowering' type is regarded as the economically most significant flower type. For this reason, the 'bud flowering' trait has been the prominent breeding objective for this species for several decades. The most frequently applied breeding methods of back-crossing and sport selection have led to genotypes and varieties being increasingly phenotypically indifferentiable. This circumstance has in turn resulted in an increasing number of legal disputes on the issue of variety protection and derivation.

Consequently, in the present work the existing diversity at the genome level was estimated as the foundation for further breeding projects using PCR-based 'fingerprinting' methods (RAPDs, ISSRs) for a scientifically and economically relevant selection of 74 *C. vulgaris* genotypes. A high level of similarity (> 0.74 Dice Index) was confirmed for the tested genotypes. Furthermore, the applicability of a system to identify Essentially Derived Varieties was tested and successfully modified for the application in *C. vulgaris*. The creation of segregating back-crossing populations served to further clarify the assumed monogenic-recessive inheritance of the 'bud-flowering' trait. Molecular investigations into these populations were carried out with regard to establishing a marker-assisted selection system for this trait. Using 'bulked segregant analysis', two trait-coupled DNA segments (RAPD) were identified which enabled the selection of 'bud-flowering' individuals in the initial population. Partial sequences of these fragments were identified, and formed the starting point of establishing SCAR- and SSCP-markers. With regard to the floral biology of 'bud-flowering' genotypes, knowledge regarding their flower organ identity and the cause of the loss of stamens was lacking to date. Thus, detailed microscopic, histological and molecular biological investigations into the normal and 'bud-flowering' phenotypes were carried out. These experiments indicated that with 'bud-flowering' plants, the petals are transformed into an additional whorl of petaloid sepals, while the anthers were not transformed into other organs, but are completely missing. Using RACE-PCR, it was possible to achieve partial sequence data for class B (*AP3*/*DEF*-like) and class E (*SEP1*-like) MADS-box transcription factors. This data was used for quantitative gene expression analyses in flower organs with qRT-PCR. Combined with the morphological results, the entirety of the molecular data gained forms the basis and prerequisite for further targeted breeding of this culture and, in particular, the 'bud-flowering' trait.

Keywords:

variety protection, marker-assisted selection, flower development

Schlüsselwörter:

Sortenschutz, Marker-gestützte Selektion, Blütenentwicklung

Index of Contents

[1] Borchert T, Krueger J, Hohe A (2008) **Implementation of a model for identifying Essentially Derived Varieties in vegetatively propagated *Calluna vulgaris* varieties.** BMC Genetics 9:56

[2] Borchert T, Hohe A (2009) **Identification of molecular markers for the flower type in the ornamental crop *Calluna vulgaris*.** Euphytica, in press.

[3] Borchert T, Eckart K, Fuchs J, Krueger K, Hohe A (submitted) **'Who's who' in different flower types of *Calluna vulgaris* (*Ericaceae*): morphological and molecular analyses of flower organ identity.** BMC Plant Biology.

Index of Contents (continued)

Abbreviation	Explanation
A	Androecium
AFLP	Amplified Fragment Length Polymorphism
AG	AGAMOUS
AICc	corrected Akaike Information Criterion
AiF	German Federation of Industrial Research Associations
AP3	APETALA3
BAC	Bacterial Artificial Chromosome
BLAST	Basic Local Alignment Search Tool
BMELV	Federal Ministry of Food, Agriculture and Consumer Protection
BSA	Bulked segregant analysis
Ca	Calyx
CArG-Box	CC-A rich-GG-box
CIOPORA	International community of breeders of asexually reproduced ornamental and fruit varieties
Co	Corolla
CPVO	Community Plant Variety Office
CPVR	Community Plant Variety Right
DAD1	DEFECTIVE IN ANTHER DEHISCENCE1
DEF	DEFICIENS
DNA	Deoxyribonucleic Acid
DUS	Distinctness, Uniformity, Stability
EDV	Essentially Derived Variety
FRI	FRIGIDA
FT	FLOWERING LOCUS T
G	Gynoecium
GLO	GLOBOSA
IGZ	Leibniz Institute of Vegetable and Ornamental Crops
InDel	Insertion / Deletion
IP	Intellectual Property
ISSR	Inter Simple Sequence Repeat
ITS	Internal Transcribed Spacer
LFY	LEAFY
MADS	MCM1, Agamous, Deficiens, Srf1
MAS	Marker-assisted Selection
matK	maturase within the trnK intron

Abbreviation	Explanation
MIKC	MADS – intervening – keratin-like – C-terminal
miRNA	microRNA
NADH	Nicotineamide-Adenine-Dinucleotide
NCBI	National Center for Biotechnology Information
ORF	open reading frame
PAGE	Polyacrylamide Gel Electrophoresis
PBR	Plant Breeder's Rights
PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
pers. comm.	personal communications
PI	PISTILLATA
PP	Plant Patent
PVP	Plant Variety Protection
RACE	Rapid Amplification of cDNA ends
RAPD	Randomly Amplified Polymorphic DNA
rbcL	ribulose-bisphosphate carboxylase (RuBisCO Large subunit)
RFLP	Restriction Fragment Length Polymorphism
RHS	Royal Horticultural Society
RNA	Ribonucleic Acid
rRNA	ribosomal RNA
SAGE	Serial analysis of gene expression
SCAR	Sequence Characterized Amplified Regions
SEM	scanning electron microscopy
SEP	SEPALLATA
SNP	Single Nucleotide Polymorphism
SSCP	Single Strand Conformation Polymorphism
STR	short tandem repeats
UFO	Unusal Flower Organs
UKHS	United Kingdom Heather Society
UPOV	International Union for the Protection of New Varieties of Plants
VNTR	variable number of tandem repeats
ZMP	Zentrale Markt- und Preisberichtsstelle

1. General foreword

1.1. Summary of the project objectives

The results described below mainly were part of the project “Grundlegende Untersuchungen zur Genetik und Züchtung von Knospenblühern bei *Calluna vulgaris*”, supported by grants from the German Federation of Industrial Research Associations (AiF, KP0172401BN5A, 01.08.2005 – 31.07.2007). In addition, elements of the work were conducted during the project “Nutzung neuer molekularer Methoden zur Effizienzsteigerung der Züchtung von Knospenblühern bei *Calluna vulgaris* L. (Hull.): ein interdisziplinärer Ansatz zur Stärkung von Innovationskraft und Wettbewerbsfähigkeit“, funded by the Federal Ministry of Food, Agriculture and Consumer Protection (BMELV, PGI-06.01-28-1-43.038-07, 01.03.2008 – 28.02.2011).

In detail, the objectives of this thesis were

(i) to investigate the available genetic diversity within *C. vulgaris* by applying PCR-based ‘fingerprinting’ methods for genotypes of actual interest with special regard to Essential Derivation, see Chapter 2.1.,

(ii) to develop an initial marker-assisted selection (MAS) system for the economically most important trait of ‘bud-flowering’ as described in Chapter 2.2. and finally

(iii) to contribute to the initial clarification of the vague and unknown flower organ identities, see Chapter 2.3.

1.2. Introducing *Calluna vulgaris* (*Ericaceae*)

Calluna vulgaris (L.) Hull (Taxonomy ID: 13385, Genbank common name: heather flower) is an evergreen, perennial shrub belonging to the asterids (core eudicotelydons) and it is the single species of this genus. The superior order *Ericales* comprises 24 other families including approx. 11,515 species (Stevens 2001) which range from arctic to alpine regions in both hemispheres (Hermann and Palser 2000) and therefore, this family is tolerant to varying climatic conditions (Dixon and Dutton

1987). Most of these are acid-loving woody plants (Kron 2008) and commonplace in moorlands in Europe (Figure 1A) while the plant family itself is common almost worldwide (Figure 1B). A fewer number of the species of the related genus *Erica* is present in Europe, but these reside in a broader geographic range when compared to Africa. Especially in the southernmost regions of Africa, the largest number of species of this genus can be found (McGuire and Kron 2005).

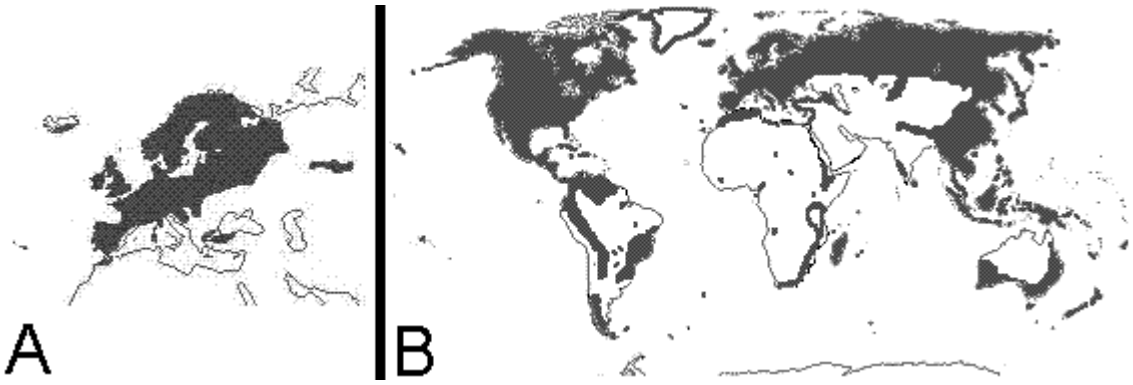


Figure 1: Distribution map of *C. vulgaris* (A, reproduced from UKHS) and the *Ericaceae* (B, reproduced from Stevens 2001).

The denomination *Calluna* presumably has developed from 'kallunein' (to cleanse, Greek) in a time in which the branches of this agricultural crop were used as brooms (UKHS). Nowadays, *C. vulgaris* is primarily sold as a potted plant for graveyards or balcony planting and the shoots are also used by florists in decorative arrangements. Nevertheless, in Australia and New Zealand, *C. vulgaris* is classified and treated as a weed because of its threatening the domestic flora due to its ability to spread quickly (see, e.g. Roy et al. 2004).

The key trade and industry data for this species are reported annually by the "Zentrale Markt- und Preisberichtsstelle" (ZMP). Since 1998, the production in Germany increased on average by 9.3% per year to more than 100 million plants in 2007. In contrast, roughly within the same timeframe (2000 – 2006), the relative market share of woody (ornamental) plants decreased from 17.5% to 15.9%. In 2007, *C. vulgaris* itself had a market value of approx. 120 Million EUR (~ 6% market share of the balcony and bedding plants). Furthermore, *C. vulgaris* is one of the biggest export articles by the German ornamental sector and is mainly shipped to countries in Scandinavia (all data from ZMP, 2007).

In Germany, *C. vulgaris* is produced in specialized horticultural companies since the production is time consuming (about 15-16 months are required from the first cuttings

to disposable plants) and both greenhouses and open-spaced areas are necessary during the different cultivation steps. In May, the first propagation of flowerless stem cuttings is conducted using stock plants that are cultivated in acidic soils (pH < 4.0). Usually in June and October of the same year, the plants are automatically pruned. Balanced fertilization is carried out weekly. Depending on the weather conditions in February of the following year, the amount of fertilizers is increased and in and around March, the young plants are pruned for the third time, immediately before potting. The potted plants are immediately transferred to open-spaced cultivation areas and typically are grown in 9-12 cm pots until May, when the fourth and last pruning takes place. Until sale in August / September, fertilization and plant protection efforts are decreased to a minimum that just maintains the plant's health status. In order to suppress typical plant pathogens (e.g. *Glomorella* / *Colletotrichum*) waterlogging has to be avoided throughout the complete propagation process. *C. vulgaris* does not require high energy inputs during the cultivation procedure, i.e. frost-free greenhouse are sufficient throughout the winter months.

Except for some studies of wild populations, see, e.g. Mahy et al. (1997), Mahy and Jacquemart (1998), Mahy et al. (1999), Meikle et al. (1999), Rendell and Ennos (2002) and a handful of database entries (only 11 nucleotide records, e.g. the partial sequence of the 18S rRNA gene, and four protein records are available in the NCBI (National Center for Biotechnology Information) database, checked: May 2009), no further molecular data is publicly available for *C. vulgaris* itself. Nevertheless, several research groups have focused on the phylogenetic classification of the order *Ericales*, see, e.g. Anderberg (1992 and 1993), Kron et al. (1999), McGuire and Kron (2005), Schönenberger et al. (2005), Duangjai et al. (2006). These studies were either conducted with morphological data e.g. pollen (Janssens 2005) or seed morphology (Fagúndez and Izco 2004), partial molecular / sequence information e.g. from 28S rRNA (Cullings 1994), 18S rRNA (Kron 1996), or ITS (internal transcribed spacer) data (Geuten et al. 2004) or the combination of these types (e.g. Kron et al. 2002). An overview of the phylogenetic relationship within the *Ericaceae* as well as the directly related species within the *Ericeae* is demonstrated in Figure 2 (next page). This strict consensus tree resulted from the combined analysis of molecular i.e. plastid (*rbcL*: large subunit of the ribulose-bisphosphate carboxylase, *matK*: maturase within the *trnK* intron) and morphological characters (mainly of the flower) and is

mostly supported with high bootstrap values. Among the *Ericeae*, *C. vulgaris* (indicated by arrow) forms a monophyletic group with *Erica* spp. (86% bootstrap probability) and *Daboecia cantabrica*. The *Ericeae* are supported by 79%.

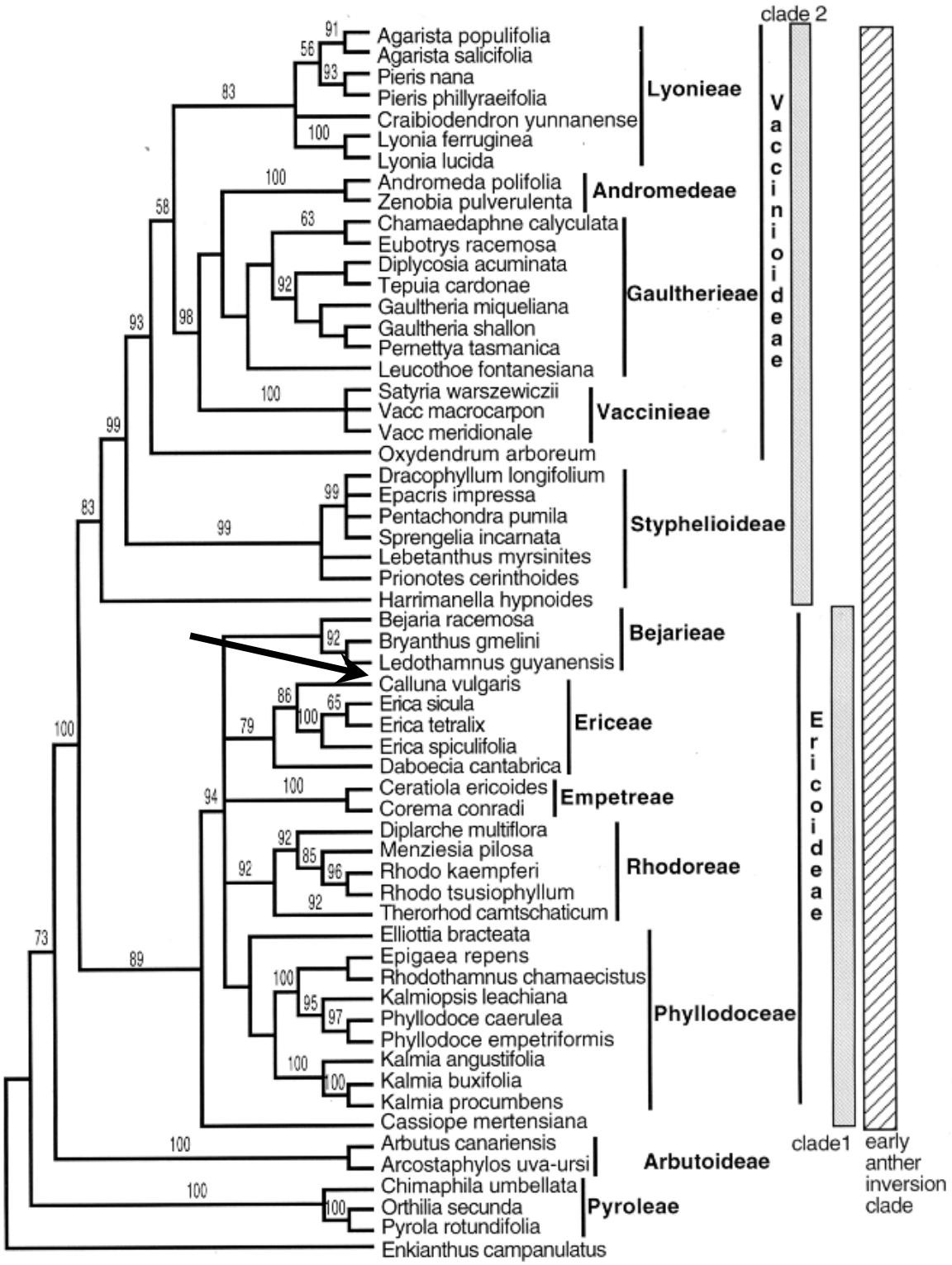


Figure 2: Strict consensus of 96 trees found in the combined analysis of morphology, *matK,* and *rbcL* data for 59 taxa of *Ericaceae*. gaps scored as missing data, bootstrap values above lines. Rhodo = *Rhododendron;* Therorhod = *Therorhodion;* Vacc = *Vaccinium* (reproduced from Kron et al. 2002, Figure 7).

1.3. Breeding *Calluna vulgaris*

1.3.1. The trait of interest: ‘bud-flowering’

The synoecious, symmetric flower of *C. vulgaris* consists of four colored leaves in each of the two outmost whorls, eight stamens in the third and four coadunate carpels in the fourth whorl. The economic importance of *C. vulgaris* for the horticultural business may be broken down to a simple but influential change of this described flower anatomy, the so-called ‘bud-flowering’ phenotype. Roughly 80% of the actual assortment of the economically relevant varieties is of this type which can be easily identified either by its lack of the male reproductive organs or more simply by the non-opening flowers i.e. by their remaining as closed bud in a stage prior to opening. Mature flowers of both the ‘wild-type’ and the ‘bud-flowering’ phenotype are shown in Figure 3.

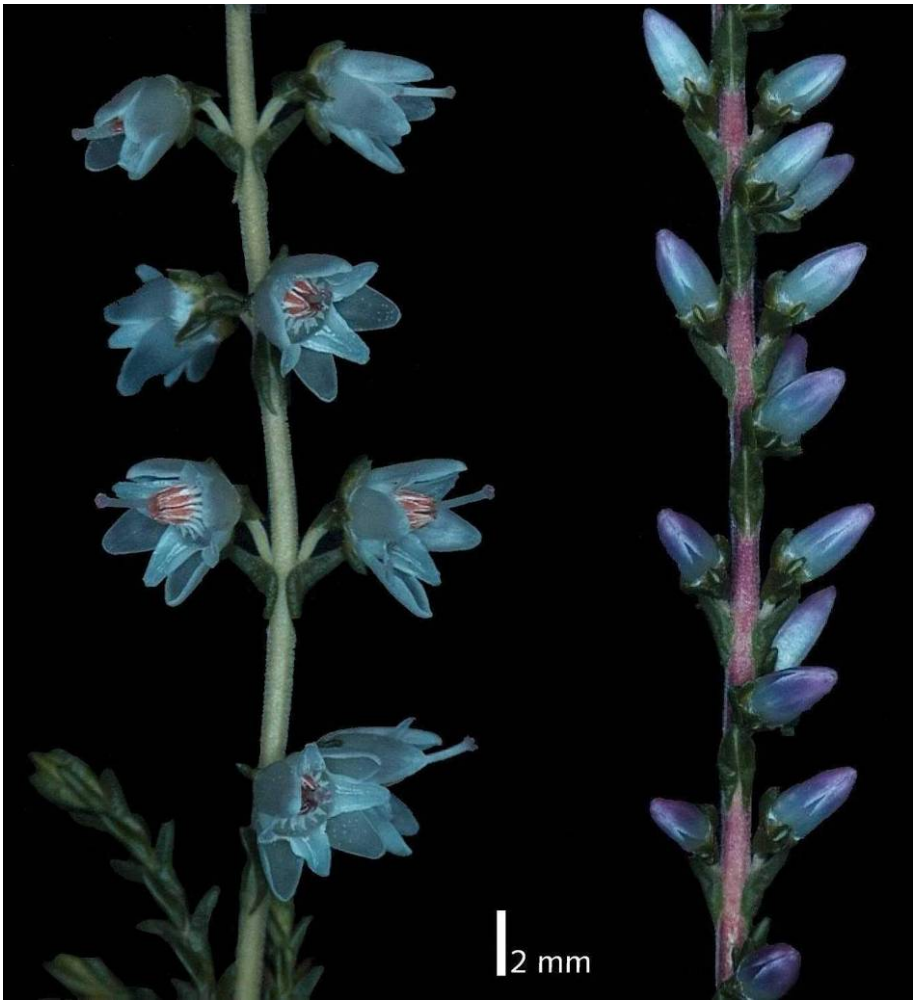


Figure 3: Shoots of *C. vulgaris* with mature flowers. Left: ‘wild-type’ variety ´White Mite´, right: ‘bud-flowering’ variety ´Anneliese´ (reproduced from Borchert et al. (2008), Figure 1).

The first reference to the 'bud-flowering' phenotype remains unclear, but several internet-based sources (e.g. http://www.gardengirls.de/heidezuechtung/, checked: May 2009) name the town Golm (Brandenburg, Germany) as the place of its first appearance in 1903. However, in 1885, Thomé portrayed only the 'wild-type' (Stueber 2007) but in 1935, three sub-forms of the 'bud-flowering' phenomenon were described (McClintock 1986) that exhibited different anatomical features in comparison to the unaltered 'wild-type' (Table 1) and that were identified in nurseries in the UK.

Table 1: Subforms of **'**bud-flowering**'** in *C. vulgaris* and example varieties, descriptions cited from McClintock (1986).

Subform	Features in comparison to the 'wild-type'	Example variety
'diplocalyx'	• eight instead of four sepals • **no stamens**, **no corolla** • deformed style	´Marleen´
'polysepala'	• similar to 'diplocalyx' i.e. **no stamens** • fatty flowers • **more than eight sepals**	´Adrie´
'clisthantes'	• **normal amount of flower parts** • flowers do not open • crouched and bent stigma	´Visser´s Fancy´

Interestingly, the 'clisthantes' subform was not classified as having lost its stamens. By investigating the example varieties mentioned in Table 1 for the 'polysepala' and the 'clisthantes' type, we were not able to confirm the descriptions. For example, the variety 'Visser´s Fancy' as it was available to us did not develop stamens (data not shown) which deviates from the normal amount of flower organs. Furthermore, Heß (1990) described the 'wild-type' as displaying 4 – 5 leaves in each whorl. Since both, McClintock (1986) and Heß (1990), delivered no justifications for the assignment of the organ identities, the available descriptions remain rather vague.

As an effect of the non-opening of the flower buds, the stigma of 'bud-flowering' phenotypes remains unpollinated throughout the whole reproduction season. In nature, insect pollination of spontaneous 'bud-flowering' mutants is therefore completely inhibited. Consequently, uncontrolled spreading of traits and especially of the 'bud-flowering' trait itself is not possible. The closed buds of 'bud-flowering' phenotypes lead to prolonged flowering times i.e. to an extended conservation of flower colors in comparison to 'wild-type' individuals. Hence, this feature is the prime reason for the economic success of this phenotype since only a few other species are capable of maintaining their flower color throughout the winter i.e. at least until December. Con-

sequently, it is a popular plant not only for consumer's balcony in autumn but especially for graveyard plantings in winter.
From the breeder's point of view, 'bud-flowering' individuals are only applicable as pollen acceptors. Therefore, crossings are required that lead to the segregation of 'bud-flowering' genotypes in their progenies. Practically, the 'bud-flowering' trait significantly increases the manual labor that is required for the production of new varieties since each flower has to be opened manually in order to pollinate the carpels.

Breeding in *C. vulgaris* started approximately three decades ago. The first released variety was a white 'bud-flowering' type (CPV.4937 2007). Today, flower colors ranging from white to dark ruby and foliage colors from dark green to lime have been introduced to the market as a result of breeding efforts. In contrast, the natural color range of *C. vulgaris* flowers is limited to lilac and rose (Witt 1996).
Due to the economical importance of the 'bud-flowering' type, breeding efforts in normal-flowering plants were significantly lower (CPV.4937 2007). Hence, only a small part of the currently protected varieties belong to the 'wild-type': only 60 of 307 varieties of the German reference collection at the German Federal Plant Variety Office, Hannover, are of the normal flowering type (CPV.4937 2007). Nevertheless, actual filings for variety protection indicate an increase of new 'wild-type' flower phenotypes, but there is no evidence for the 'bud-flowering' type to be replaced by other interesting anatomical features, e.g. filled or 'multi-bracteate' genotypes (Spellerberg, Federal Plant Variety Office, pers. comm.). Although increasing, overall filings for Plants Breeders' Rights (PBR) on the European level for a woody species like *C. vulgaris* are low (33 or ~0.26%, respectively) in comparison to the total amount of 12,300 grants (Mac Cárthaigh 2008). Today (checked: May 2009), 46 *C. vulgaris* varieties are granted Community Plant Variety Rights (CPVR) and 24 are still under procedure. These resemble ~0.43% of 16,339 grants, respectively (CPVO 2009a).
Normally, the breeding of woody plants is time-consuming due to a reduced developmental speed but the shrub *C. vulgaris* has a short juvenile phase and thereby offers the chance of faster selection cycles in comparison to e.g. *Rhododendron*. Traditionally, breeding in *C. vulgaris* started off by the undemanding selection of spontaneous mutations within the plants at a nursery or in the wild (Mac Cárthaigh 2008). This selection breeding method was applied both to seedling populations and to individual plants (e.g. for the selection of flower color sports). Today, selection by chance

is supplemented by classical cross-breeding which is of an increasing importance for the introduction of new varieties in *C. vulgaris* (Heidepflanzen de Winkel, pers. comm.; Mac Cárthaigh 2008). However, application documents of the German Federal Plant Variety Office may indicate that mutation breeding was a regularly applied breeding method in *C. vulgaris*, too. For example, the variety 'Dark Beauty' was identified after irradiation.

C. vulgaris is the solitary species within its genus and thereby, inter-specific hybridizations in order to expand the available gene pool are not possible. As described above, mutation selection and repeated backcrossing are the main breeding methods. In combination with continuing breeding for one common phenotype – 'bud-flowering' – a decrease of the genetic diversity was expectable (Vosman et al. 2004). As a consequence of these facts, German courts have already been engaged in several juridical incidents.

Since studies of the genetic diversity in this species are only available for several wild populations (see above), scientific investigations of the genetic diversity of current and economically important genotypes were an essential task for breeding companies. The knowledge of the genetic distances of genotypes that are integrated into breeding programs is a crucial prerequisite for successful breeding. Such examination is either based on morphological traits or on data gained from molecular tools, e.g. isoenzyme or DNA analysis. Modifying influences e.g. environmental effects or post-translational alterations have led to a decreasing application of non-DNA-based methods over the last few years, both for diversity and marker research (see e.g., Staub et al. 1996a, Rout and Mohapatra 2006). In contrast, molecular tools benefit from their universal applicability and the fact that quite uncomplicated changes of the methodology permit access to different genomic regions (compare RAPDs (Randomly Amplified Polymorphic DNAs) vs. ISSRs (Inter Simple Sequence Repeats) or AFLPs (Amplified Fragment Length Polymorphisms).

The respective results of the identification of genetic diversity by RAPD- and ISSR-'fingerprinting' are described in Chapter 2.1.

1.3.2. Variety protection in ornamental crops and *C. vulgaris*

Since breeding is a time-consuming and budget-intensive procedure, the development of a new variety may lead to enormous investments depending on the market

and the crop itself. For example, in maize, the development of a new hybrid requires 15 years of breeding work and cash resources of up to US$ 5,000,000 (Troyer et al. 2002). A comparable sum is not known for *C. vulgaris* but, according to CIOPORA (International community of breeders of asexually reproduced ornamental and fruit varieties) information and pers. comm. (Heidepflanzen de Winkel), German breeding companies spend approx. 10% of their annual turnover for breeding purposes. Therefore, a reliable and functional system for re-funding a breeder's investments and protecting the breeder and his invention from illegal reproduction by competitors is required. Nowadays, such structures are available since protection can either be maintained by patent protection (PP, in the USA) or by plant variety protection (PVP, in the European Union and other UPOV (International Union for the Protection of New Varieties of Plants) member states, respectively). Both types of intellectual property (IP) rights are used to counterbalance public i.e. research and private i.e. commercial interests.

The UPOV conventions are applied in its 67 member states (as at May 12, 2009). The UPOV is a union of countries that focuses on the synchronization of PVP systems. Their primary system for PVP was initially established at the Diplomatic Conference in Paris on December 2, 1961 and is well-known as the 'UPOV Act of Convention'. Until now, it was amended in 1972, 1978 and 1991. The most important amendment in 1978 included the introduction of the 'Breeder's Exemption' which led to the possibility of the legal use of a protected variety in breeding programs by others without the requirement for authorization by the initial breeder. This exception included experimental purposes, too. In 1991, the concept of 'Essential Derivation' i.e. the definition of an Essentially Derived Variety (EDV) was incorporated into the Act in order to prevent plagiarism. With the acceptance of those concerned i.e. the breeders, the determination of an Essential Derivation was left open to themselves. Both additions were an essential prerequisite for maintaining future breeding efforts lucrative and thus promoting the introduction of new varieties.

Proving the EDV-criteria can be achieved either by phenotypic comparative testing i.e. the investigations for Distinctness, Uniformity and Stability (DUS testing) as introduced and established by the UPOV; or, moreover, distinctness can be assessed on the genetic level using molecular methods. With both techniques, phenotypic as well as molecular, dissimilarity is more easily verifiable than similarity since most of the available procedures acquire (random) sample subsets and leave behind a non-

investigated area differing in its size. In contrast, to proof exact identity between two organisms on the molecular level and thereby excluding the chance and probability for errors would mean sequencing the complete genome. Even then, the occurrence of e.g. SNPs (Single Nucleotide Polymorphisms) would complicate the analysis. Today, this is not an appropriate option for most species and most cases due to the cost factor.

On the basis of the UPOV definition of a variety ('a variety is a plant grouping within a single botanical taxon of the lowest known rank'), testing systems for different species were developed. For *C. vulgaris*, the guideline for DUS testing includes 22 characteristics which are acknowledged to discriminate between varieties (Table 2).

Table 2: Key characteristics for DUS testing in *C. vulgaris* (L.) Hull / Ling / Scots Heather according to CPVO (2009b). Characteristics no. 16 – 19 refer only to the 'wild-type', characteristics 20 – 21 refer only to 'bud-flowering' phenotypes, respectively. RHS = Royal Horticultural Society.

CPVO No.	Plant Part	Characteristic	Categories
1	Plant	Growth habit	4
2		Density	3
3		Height	3
4	Shoot tip	anthocyanin coloration (during winter)	5
5		Color of new growth	12
6		anthocyanin coloration (in middle of summer)	5
7	Shoot	Color on sunny side	10
8	Leaf	Color	12
9	Flowering shoot	Length of current season's growth	3
10		Color	5
11		Color of tip (at beginning of flowering)	6
12	Inflorescence	Density of flowers	3
13	Flower	Opening of bud	2
14		Type	2
15		Size	3
16		Calyx : Corolla (length)	3
17		Color of outer side of sepal	RHS
18		Color of outer side of petal (at beginning of flowering)	RHS
19		Color of outer side of petal (at the end of flowering)	RHS
20		Main color (at beginning of flowering)	RHS
21		Main color (at the end of flowering)	RHS
22		Time of beginning of flowering	5

Since these DUS testings are still and solely based on phenotypic classifications and are thereby prone to environmental and personal influences (of the tester) i.e. errors, the increasing demand for molecular-based assays to assist and supplement the application procedure for PBR or to completely substitute phenotype-based assays has been noticeable for several years (see, e.g. Staub et al. 1996b). Recently, the Community Plant Variety Office (CPVO) itself successfully conducted a study to evaluate the suitability of isoenzyme characteristics for integration into *C. vulgaris* DUS testing as both, a rising amount of declared varieties and the subsequent decrease in phenotypic differences due to the breeding methods applied, led to increasing efforts in morphological DUS testing (CPV.4937 2007).
Regarding the insights gained from the research into the genetic diversity in *C. vulgaris*, the initial project plan was belatedly expanded to additionally cover up the currently increasing and important field of EDV conflicts. The results were included in Chapter 2.1.

1.3.3. Marker-Assisted Selection for the 'bud-flowering' trait

As an outcome of crossing experiments at the former Institute for Ornamental Plant Breeding of the Federal Centre for Breeding Research on Cultivated Plants, Ahrensburg, Germany as well as by the cooperating breeder (pers. comm.), the inheritance of the 'bud-flowering' trait is assumed to be monogenic-recessive. These crossings revealed the segregation of *C. vulgaris* progenies in accordance to the 1st and 2nd Mendelian law (law of uniformity and law of segregation): if the female parent (the 'bud-flowering' phenotype) is crossed with a 'wild-type' plant as pollen donor, all individuals of the resulting F_1 population flower completely normal i.e. they resemble the 'wild-type'. Backcrossing the former 'bud-flowering' parent with an F_1 individual, BC_1 progenies segregate in phenotype approx. 1:1 regarding the trait of interest as expected for monogenic traits. At the Leibniz-Institute of Vegetables and Ornamental Crops, Erfurt, Germany, (IGZ) these segregation ratios could be partially (i.e. for some genotype combinations) reproduced. But, supposedly and as segregation ratios indicate, there still exist other additional factors influencing the segregation since a number of progenies displayed differing ratios, e.g. of up to 1:33 regarding the 'bud-flowering':'wild-type' ratio (see Table 3 of Chapter 2.2.1.). If these differing ratios are based on environmental influences, on a lower fitness of seedlings of one of the

flower types or on genetic conditions i.e. unknown crossing incompatibilities between individual varieties and / or modifying background genes, still remains unclear.
MAS towards the monogenic-recessive inherited 'bud-flowering' type may be established by applying 'bulked segregant analysis' (BSA, Michelmore et al. 1991). Breeders would benefit financially from a MAS system since the usual time line for the production of 'bud-flowering' progenies includes at least two crossing steps. Applying MAS, earlier selection steps could be achieved and in addition, several rating dates could be spared.
MAS systems for phenotypic traits are not frequently applied to ornamental crops (Rout and Mohapatra 2006, Byrne 2007). Instead, the first markers developed for ornamentals or fruits were of the resistance type e.g. Yang and Kruger (1994) in *Malus floribunda*. Until now, functional markers for flower architecture i.e. phenotypic traits developed for ornamentals are known e.g. in carnations (*Dianthus caryophyllus*, Scovel et al. 1998). These authors developed a marker system that discriminates the economically important 'semi-double' and 'double' flower types. In *Rosa hybrida*, Debener and Mattiesch (1999) mapped genes for the petal number and flower color. Cheghamirza et al. (2002) inter alia identified RAPD markers for flower and seed color in pea (*Pisum sativum*) by BSA. Other examples are the discovery of markers for homostyly in buckwheat (*Fagopyrum esculentum*, Aii et al. 1998) or the detection of a marker for male-sterility in apricots (*Prunus armeniaca*, Badenes et al. 2000).
The results of the establishment of a molecular marker linked to the 'bud-flowering' phenotype in *C. vulgaris* using BSA by RAPD-, SCAR- (Sequence Characterized Amplified Regions) and SSCP- (Single Strand Conformation Polymorphism) PCR are described in Chapter 2.2.

1.3.4. Common molecular features of the flower

The biological consequences for a plant missing the male reproduction organs are manifold: most simply, the genotypes are male sterile. In addition, the process of flower opening may be affected. This procedure, according to van Doorn and van Meeteren (2003), is either based on cell contraction and elongation or on different growth rates on two opposing sites on the flower. As shown for *A. thaliana* and other species (see, e.g. Ishiguro et al. 2001), the water transport into stamens and petals regulated by jasmonic acid is one fundamental possibility of inducing flower opening, since the water uptake leads to elongation of these organs and thereby, the petals

press on the sepals and the flower opens. These results were obtained using *DAD1*-mutants (*DEFECTIVE IN ANTHER DEHISCENCE1*, Ishiguro et al. 2001), for which the authors could restore the normal flower opening procedure by exogenous application of jasmonic acids on flower buds of mutant genotypes. *DAD1* plants carry a sporophytic mutation that inhibits anther dehiscence and production of viable pollen grains; *DAD1* is assumed to be involved in the production of jasmonic acid. In contrast, *C. vulgaris* exhibits a total loss of stamens which is a crucial difference to *DAD1*-mutants in that it resembles the complete loss of one of the two main transport routes of this phytohormone and water, respectively. Thus, the water uptake and the subsequent elongation of petals, which is induced by the elongation of stamens, is assumed to be less effective in 'bud-flowering' than in 'wild-type' phenotypes. However, a complete restoration of flower opening in 'bud-flowering' individuals by exogenous application of phytohormones is not desired.

Flowers of angiosperms consist of four organ types, which can be separated into the perianth (whorl I: sepals, whorl II: petals), the androeceum (whorl III: stamens) and the gynoeceum (whorl IV: carpels). Flower organ identity is defined by whorl-specific expression of homeotic genes according to the so-called ABC-model that was initially introduced by Coen and Meyerowitz (1991) resulting from studies with *Arabidopsis* and *Antirrhinum* mutants. Within this original model, the terms A, B and C were used to simultaneously denominate four different subjects: areas of the floral meristem, classes of homeotic mutants, of regulatory functions and of floral homeotic genes (Theissen 2001). For a comprehensive review of molecular explanations ('floral quartet'-model) of the genetic 'ABC+DE'-model, see e.g. Theissen (2001) or Theissen and Melzer (2007).

The homeotic genes specify the organ identity of each whorl by activating so-called 'realizator genes' (Theissen et al. 2000) and therefore have to be classified as transcription factors. They commonly belong to the MIKC-type (MADS – intervening – keratin-like – C-terminal - domain) proteins and include the conserved and name-giving MADS (*MCM1*, *Agamous*, *Deficiens*, Serum Response Factor) domain. As shown in Figure 4 (next page), A-function genes are required for sepal formation in whorl I, simultaneous expression of A- und B-function genes is required for petal formation in whorl II and synchronized expression of B- and C-genes is essential for stamen development. C-genes alone lead to the formation of carpels and expression

of D-genes is required for ovule formation. Class E genes are invariably necessary for proper organ development in all whorls since their protein products are essentially required for the formation of the multimeric DNA binding structures (Theissen and Saedler 2001) that activate the 'realizator genes'. These findings led to the formulation of the 'floral quartet'-model and to the assigning of tetrameric protein structures to each homeotic function (Figure 4 top). According to this model, shifts in the flower organ composition are usually followed by adaptations in other whorls i.e. an increase of the amount of organs in the affected whorl since these homeotic functions are mutually antagonistic. This means, for example, that in class C mutants (i.e. mutants lacking the expression of class C genes), the A function will 'expand' to the two innermost whorls and lead to an alternating and repetitive development of sepals – petals – petals – sepals (whorls 1 to 4).

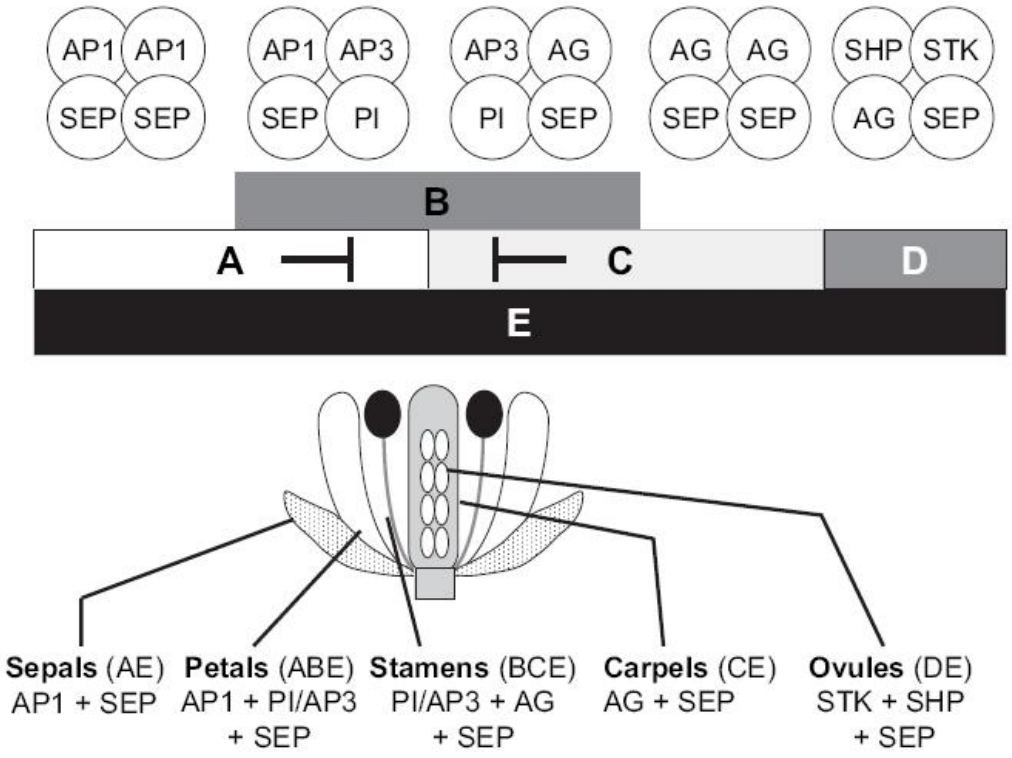


Figure 4: The genetic ABC+DE model and the quartet model. The ABC model states that A-function genes, such as *AP1* in *Arabidopsis*, are necessary for the formation of the sepals, B-function genes, which include *AP3* and *PI* in *Arabidopsis*, along with A-function genes, are necessary for the formation of the petals and B-function, along with C-function genes, which in *Arabidopsis* includes *AG*, are necessary for the formation of the stamens, and C-function genes alone are necessary for the formation of the carpels. This has been expanded to include class D- and E-function genes, which are necessary for the ovules and whorls of the flower, respectively. D-function genes in *Arabidopsis* include *SEEDSTICK* (STK) and *SHATTERPROOF1* and *SHATTERPROOF2* (*SHP1*, *SHP2*). E function *sensu lato* requires at least one of the four *SEPALLATA* (*SEP1*, *SEP2*, *SEP3*, and *SEP4*) genes. The floral quartet model expands on this idea using data from protein interaction studies (Theissen 2001; Theissen and Saedler 2001). In this figure the hypothesized quartets, based on experimentally determined dimeric or multimeric protein interactions, necessary for each floral organ are presented (reproduced from Zahn et al. 2005, Figure 2).

However, the molecular basics for the occurrence of the ‘bud-flowering’ trait in *C. vulgaris* are unknown. In addition, further flower types, e.g. filled flowers or ‘multi-bracteate’ inflorescences, are well-known. Their development remains unclear due to non-available or dissatisfactory descriptions of the organs (see above) and the marginal availability of molecular data, too. Therefore, basic examination of the flower organ identity of *C. vulgaris* was desired.

The results regarding MADS-box gene identification, expression analysis and macroscopic organ classification of flower organs in *C. vulgaris* are described in Chapter 2.3.

1.4. References

Aii J, Penner GA, Clayton CG, Adachi T (1998) **Identification of RAPD markers to the homostylar (Ho) gene in buckwheat**. Breeding Science. 48, 59-62

Anderberg AA (1992) **The circumscription of the *Ericales*, and their cladistic relationships to other families of ''higher'' dicotyledons**. Systematic Botany 17:660–675

Anderberg AA (1993) **Cladistic interrelationships and major clades of the *Ericales***. Plant Systematics and Evolution 184:207–231

Badenes ML, Hurtado MA, Sanz F, Archelos DM, Burgos L, Egea J,Llacer G (2000) **Searching for molecular markers linked to male sterility and self compatibility in apricot**. Plant Breeding 119:157-160

Borchert T, Krueger J, Hohe A (2008) **Implementation of a model for identifying Essentially Derived Varieties in vegetatively propagated *Calluna vulgaris* varieties**. BMC Genetics 9:56

Byrne DH (2007) **Molecular marker use in perennial breeding**. Acta Horticulturae 751:163-167

Cheghamirza K, Koveza O, Konovalov F, Gostimsky S (2002) **Identification of RAPD markers and their use for molecular mapping in pea (*Pisum sativum* L.).** Cellular and Molecular Biology Letters 7(4):1158

Coen ES, Meyerowitz EM (1991) **The war of the whorls: genetic interactions controlling flower development.** Nature 353:31-37

CPV.4937 (2007) **A study on the possible use of isoenzyme characteristics for the assessment of distinctness, uniformity and stability in *Calluna vulgaris* L.** Final Report, Agreement number CPV.4937

CPVO (2009a) **PVR Case Law database**. http://www.cpvo.fr/main/de

CPVO (2009b) **Guidelines for the conduct of tests for Distinctness, Uniformity and Stability**. **Ling, Scots Heather, *Calluna vulgaris* (L.) Hull**. TH/94/6 Corr. of 04/04/2001 and 27/02/2009

Cullings K (1994) **Molecular phylogeny of the *Monotropoideae* (*Ericaceae*) with a note on the placement of the *Pyroloideae*.** Journal of Evolutionary Biology 7:501-516

Debener T, Mattiesch L (1999) **Construction of a genetic linkage map for roses using RAPD and AFLP markers**. Theoretical and Applied Genetics 99:891-899

Dixon GR, Dutton KJ (1987) **Manipulation of flowering in white heather**. Acta Horticulturae 205:219-224

Duangjai S, Wallnofer B, Samuel R, Munzinger J, Chase MW (2006) **Generic delimitation and relationships in *Ebenaceae* sensu lato: Evidence from six plastid DNA regions**. American Journal of Botany 93(12):1808-1827

Fagúndez J, Izco J (2004) **Seed morphology of *Calluna salisb*. (*Ericaceae*).** Acta Botanica Malacitana 29:215-220

Geuten K, Smets E, Schols P, Yuan Y-M, Janssens S, Küpfer P, Pyck N (2004) **Conflicting phylogenies of balsaminoid families and the polytomy in *Ericales*: combining data in a Bayesian framework.** Molecular Phylogenetics and Evolution 31:711–729

Hermann PM, Palser BF (2000) **Stamen Development in the *Ericaceae*. I Anther wall, microsporogenesis, inversion and appendages**. American Journal of Botany 87(7):934-957

Heß D (1990) **Die Blüte. Einführung in Struktur und Funktion, Ökologie und Evolution der Blüten**. 2. Auflage, Stuttgart, Ulmer

Ishiguro S, Kawai-Oda A, Ueda J, Nishida I, Okada K (2001) **The DEFECTIVE IN ANTHER DEHISCENCE1 Gene Encodes a Novel Phospholipase A1 Catalyzing the Initial Step of Jasmonic Acid Biosynthesis, Which Synchronizes Pollen Maturation, Anther Dehiscence, and Flower Opening in *Arabidopsis*.** The Plant Cell 13:2191-2209

Janssens S, Lens F, Dressler S, Geuten K, Smets E, Vinckier S (2005) **Palynological Variation in Balsaminoid *Ericales*. II. *Balsaminaceae*, *Tetrameristaceae*, *Pellicieraceae* and General Conclusions**. Annals of Botany 96:1061–1073

Kron KA (1996) **Phylogenetic Relationships of *Empetraceae*, *Epacridaceae*, *Ericaceae*, *Monotropaceae*, and *Pyrolaceae*: Evidence from Nuclear Ribosomal 18s Sequence Data**. Annals of Botany 77:293-303

Kron KA (2008) ***Ericaceae* homepage: Blueberries, Heathers and Rhododendrons**. http://www.ericaceae.org/ericeae.html

Kron KA, Judd WS, Crayn DM (1999) **Phylogenetic analyses of *Andromedeae* (*Ericaceae* subfam. *Vaccinioideae*)**. American Journal of Botany 86(9):1290–1300

Kron KA, Judd WS, Stevens PF, Crany DM, Anderberg AA, Gadek PA, Quinn CJ, Luteyn JL (2002) **Phylogenetic Classification of *Ericaceae*: Molecular and Morphological Evidence**. The Botanical Review 68(3):335-423

Mac Cárthaigh D (2008) **Breeding of Woody Plants in Europe**. Acta Horticulturae 766:349-354

Mahy G, Ennos RA, Jacquemart A-L (1999) **Allozyme variation and genetic structure of *Calluna vulgaris* (heather) populations in Scotland: the effect of postglacial recolonization**. Heredity 82:654-660

Mahy G, Jacquemart A-L (1998) **Mating systems of *Calluna vulgaris*: self-sterility and outcrossing estimations.** Canadian Journal of Botany 76:37-42

Mahy G, Vekemans X, Jacquemart A-L, de Sloover JR (1997) **Allozyme diversity and genetic structure in South-Western population of heather, *Calluna vulgaris***. New Phytologist 137:325-334

McClintock D (1986) **Harmonising botanical and cultivar classification with special reference to hardy heathers.** Acta Horticulturae 182:277-283

McGuire AF, Kron KA (2005) **Phylogenetic Relationships of European and African *Ericas***. International Journal of Plant Sciences 166(2):311-318

Meikle A, Paterson S, Finch RP, Marshall G, Waterhouse A (1999) **Genetic characterization of heather (*Calluna vulgaris* (L.) Hull) subject to different managment regimes across Great Britain**. Molecular Ecology 8:2037-2047

Michelmore RW, Paran I, Kesseli RV (1991) **Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations**. Proceedings of the National Academy of Sciences USA 88:9828-9832

Rendell S, Ennos RA (2002) **Chloroplast DNA diversity in *Calluna vulgaris* (heather) populations in Europe**. Molecular Ecology 11:69-78

Rout GR, Mohapatra A (2006) **Use of molecular markers in ornamental plants: a critical reappraisal**. European Journal of Horticultural Science 71(2):53-68

Roy B, Popay I, Champion P, James T, Rahman A (2004) **An Illustrated Guide to Common Weeds of New Zealand (2nd Edition).** New Zealand Plant Protection Society

Schönenberger J, Anderberg AA, Sytsma KJ (2005) **Molecular phylogenetics and patterns of floral evolution in the *Ericales***. International Journal of Plant Science 166(2):265-288

Scovel G, Ben-Meir H, Ovadis M, Itzhaki H, Vainstein A (1998) **RAPD and RFLP markers tightly linked to the locus controlling carnation (*Dianthus caryophyllus*) flower type**. Theoretical and Applied Genetics 96:117-122

Staub JE, Serquen FC, Gupta M (1996a) **Genetic markers, map construction, and their application in plant breeding**. HortScience 31(5):729-741

Staub JE, Gabert A, Wehner TC (1996b) **Plant Variety Protection: a consideration of genetic relationship**. HortScience 31:1086-1091

Stevens PF (2001 onwards) **Angiosperm Phylogeny Website**. Version 9, June 2008 http://www.mobot.org/MOBOT/research/APweb/

Stueber K (2007) http://www.BioLib.de: Thomé OW (Gera, 1885) **Flora von Deutschland, Österreich und der Schweiz.** http://caliban.mpiz-koeln.mpg.de/~stueber/thome/index.html

Theissen G (2001) **Development of floral organ identity: stories from the MADS house.** Current Opinion in Plant Biology 4:75–85

Theissen G, Becker A, Di Rosa A, Kanno A, Kim JT, Münster T, Winter K-U, Saedler H (2000) **A short history of MADS-box genes in plants.** Plant Molecular Biology 42:115–149

Theissen G, Melzer R (2007) **Molecular mechanisms underlying origin and diversification of the angiosperm flower**. Annals of Botany 1–17

Theissen G, Saedler H (2001) **Floral quartets**. Nature 49:469-471

Troyer AF, Rocheford TR (2002) **Germplasm ownership: related corn inbreds**. Crop Science 42:3-11

UKHS: **United Kingdom Heather Society**, http://www.heathersociety.org.uk/

van Doorn WG, van Meeteren U (2003) **Flower opening and closure: a review**. Journal of Eperimantal Botany 54(389):1801-1812

Vosman B, Visser D, Voort JR van der, Smulders MJM, van Eeuwijk F (2004) **The establishment of 'essential derivation' among rose varieties, using AFLP**. Theoretical and Applied Genetics 109:1718-1725

Witt HH (1996) **Ericaceae, Heidekrautgewächse. *Callúna vulgáris* Hull., Besenheide**. In: Horn W. Zierpflanzenbau. Blackwell Wissenschafts-Verlag GmbH, Wien

Yang HY, Kruger J (1994) **Identification of RAPD marker linked to the Vf gene for scrab resistance in apples**. Euphytica 77:83-87

Zahn LM, Leebens-Mack J, dePamphilis CW, Ma H, Theissen G (2005) **To B or not to B a flower: the role of DEFICIENS and GLOBOSA orthologs in the evolution of the angiosperms**. Journal of Heredity 96(3):225–240

ZMP: **Zentrale Markt- und Preisberichtsstelle**, http://www.zmp.de

2. Publications and Manuscripts

The following Chapter includes the peer-reviewed publications and the submitted manuscript that are related to the three main topics of this thesis: estimation of genetic diversity and essential derivation (Chapter 2.1.), establishment of a molecular marker (Chapter 2.2.) and flower organ identity (Chapter 2.3.).

The manuscript **Implementation of a model for identifying Essentially Derived Varieties in vegetatively propagated *Calluna vulgaris* varieties** was published in BMC Genetics (doi: 10.1186/1471-2156-9-56). The document itself and the additional files can be accessed online at http://www.biomedcentral.com/1471-2156/9/56.

The manuscript **Identification of molecular markers for the flower type in the ornamental crop *Calluna vulgaris*** was published in Euphytica (doi: 10.1007/s10681-009-9926-3). According to the Copyright Transfer Statement of Springer Science+Business Media B.V., an author-created version i.e. the last revised document (final revision of March 9, 2009) is included within this thesis instead of the final publisher's PDF-document. This procedure is approved by Springer Science+Business Media B.V. (a written permission is existent). For subscribers, the original publication in its final version is available at http://www.springerlink.com.

The manuscript **'Who's who' in different flower types of *Calluna vulgaris* (*Ericaceae*): morphological and molecular analyses of flower organ identity** was submitted to BMC Plant Biology (http://www.biomedcentral.com/bmcplantbiol) as at May 26, 2009. The submitted document is included in Chapter 2.3.
Remark: as at July 10, 2009, the manuscript was accepted for publication in principle (revision requested).

BMC Genetics

BioMed Central

Research article

Open Access

Implementation of a model for identifying Essentially Derived Varieties in vegetatively propagated *Calluna vulgaris* varieties

Thomas Borchert[†1], Joerg Krueger[†2] and Annette Hohe*[†2]

Address: [1]Heidepflanzen Peter de Winkel, Douvenberg 34, 47547 Goch, Germany and [2]Institute of Vegetable and Ornamental Crops (IGZ), Dept. Plant Propagation, Kuehnhaeuser Str. 101, 99189 Erfurt, Germany

Email: Thomas Borchert - borchert@erfurt.igzev.de; Joerg Krueger - krueger@erfurt.igzev.de; Annette Hohe* - hohe@erfurt.igzev.de

* Corresponding author †Equal contributors

Published: 20 August 2008

BMC Genetics 2008, **9**:56 doi:10.1186/1471-2156-9-56

Received: 31 March 2008
Accepted: 20 August 2008

This article is available from: http://www.biomedcentral.com/1471-2156/9/56

Abstract

Background: Variety protection is of high relevance for the horticultural community and juridical cases have become more frequent in a globalized economy due to essential derivation of varieties. This applies equally to *Calluna vulgaris*, a vegetatively propagated species from the *Ericaceae* family that belongs to the top-selling pot plants in Europe. We therefore analyzed the genetic diversity of 74 selected varieties and genotypes of *C. vulgaris* and 3 of *Erica* spp. by means of RAPD and iSSR fingerprinting using 168 mono- and polymorphisms. The same data set was utilized to generate a system to reliably identify Essentially Derived Varieties (EDVs) in *C. vulgaris*, which was adapted from a method suggested for lettuce and barley. This system was developed, validated and used for selected tests of interest in *C. vulgaris*.

Results: As expected following personal communications with breeders, a very small genetic diversity became evident within *C. vulgaris* when investigated using our molecular methods. Thus, a dendrogram-based assay to detect Essentially Derived Varieties in this species is not suitable, although varieties are propagated vegetatively. In contrast, the system applied in lettuce, which itself applies pairwise comparisons using appropriate reference sets, proved functional with this species.

Conclusion: The narrow gene pool detected in *C. vulgaris* may be the genetic basis for juridical conflicts between breeders. We successfully tested a methodology for identification of Essentially Derived Varieties in highly identical *C. vulgaris* genotypes and recommend this for future proof of essential derivation in *C. vulgaris* and other vegetatively propagated crops.

Background

Calluna vulgaris L. (Hull.), an exclusive species within the genus *Calluna*, has increased its economic weight, and not only within the German horticultural industry over the last few decades: In 2005 almost 100 million plants were produced in Germany, of which about 30% were exported to other European countries [1] where the demand is also still increasing. Although merely a handful of breeders are commercially active in breeding *C. vulgaris*, more than 300 varieties now exist, which are or have been protected at the Bundessortenamt, Hannover (BSA) [2] and/or the Community Plant Variety Office (CPVO), Angers, France [3]. More than 50% of applications for variety protection at the CPVO date from 2003 or later, which supports the argument of the increasing importance of *C. vulgaris*.

Breeding efforts in *C. vulgaris* primarily aim at a special type of its inflorescence, the so-called bud flowers (Fig. 1).

BMC Genetics 2008, **9**:56 http://www.biomedcentral.com/1471-2156/9/56

Figure 1
Shoots of two *C. vulgaris* genotypes representing the main inflorescence types. left: normal ('White Mite'), right: bud ('Anneliese').

Flowers of these plants do not open during the entire reproduction phase from August to December which makes them appear visually attractive for a long period of time when not many other flowering ornamental outdoor plants are available in the northern hemisphere. This phenotype is closely linked with and possibly caused by a lack of anthers. This connection, in turn, has a severe impact on breeding methods because interesting bud-flowering genotypes are only applicable as the female parent in crossings. In addition, there is only sparse information and hypotheses available concerning the inheritance of this trait. Therefore – and since *C. vulgaris* is a vegetatively propagated crop – breeding in *C. vulgaris* over the past few decades was to a large extent performed by selection of spontaneous mutations, rather than by systematic crossings (personal communications with breeders). The actual variety composition in Europe offers a mixture of normal flowering and bud flowering types (state: 01/2008) with main focus on the latter (~85%). Some special forms (e.g. 'Radnor' with filled flowers or 'Peace' as a multi-bracteate type) are present as well. However, due to the problems described above, the actual gene pool used in breeding of *C. vulgaris* is presumably quite narrow.

Therefore, in this study the genetic diversity within the species *C. vulgaris* was examined with molecular DNA techniques, comprising a selection of 64 economic important and partially still-protected varieties from Germany, including varieties from other European countries and the USA, 5 genotypes resulting from crossings, as well as a selection of 5 wild plants of different origin. Moreover, 3 different genotypes of *Erica* spp. were included in this study as an anticipated outgroup [see Additional file 1].

In the case of *C. vulgaris*, variety protection assessments as executed by the CPVO and as described in the Protocol for Distinctness, Uniformity and Stability Tests for *Calluna* L. (Hull.), LING, Scots Heather (CPVO-TP/94/Final of 06/11/2003), comprises 22 phenotypic traits in total but only 18 traits for bud-flowering varieties, which are tested in 2 flowering seasons with 30 plants (replications). Herein, problems arise from continually increasing applications for protection of bud-flowering genotypes, from their partial identicalness in many of these traits and from the subjectiveness that is inherent in the measurement of phenotypic traits. Moreover, breeding of bud flowering types requires backcrossing, which is – in contrast to mutant selection or 'cosmetic breeding' – 'true breeding', but which also contributes to the narrow gene pool. Previously, these drawbacks led to some juridical disputes in the field of variety derivation in *C. vulgaris* in Germany.

The problem of variety derivation and the need for an appropriate protection system was already identified decades ago and is especially pressing in the context of global marketing. In Europe, the Act of Convention from 1991 followed on from a Convention on the Protection of Plant Varieties [4] and first introduced the term 'Plant Breeder's Rights'. Today it is acknowledged by 65 member states (http://www.upov.int, state: 01/2008). Variety protection in these member states is based on DUS-tests (distinctness, stability, uniformity: see above). Despite increasing testing efforts, these tests remain sketchy since the investigated traits may be influenced by several factors, e.g. environmental changes, and are evaluated by subjective ratings so that molecular markers have become a desirable tool [5]. But also the so-called 'fingerprinting techniques' – although widely recommended as a supplement to phenotypic tests e.g. by [6] – entail a number of problems, since only a random sample i.e. a subset of the genome, can be examined. Therefore, any statistical method applied to this problem has to be able to maintain a delicate balance in order to avoid excessive identification of false positives on the one hand, as well as false negatives on the other [6-8].

Several case studies have been recently published in the context of ED-conflicts, however, for the most part these do not concern vegetatively propagated species, since it is generally assumed that variation in such varieties does not occur, which would allow clear-cut molecular genotyping.

[9] tested essential derivation in various vegetatively propagated ornamentals (*Rhododendron*, *Rosa*, *Phalaenopsis*) by AFLP-genotyping and constructing UPGMA-dendrograms. However, these investigations relied on the assumption of total genetic stability within vegetatively propagated varieties and therefore dispensed with any statistical analysis. From our point of view, this is not appropriate for all vegetatively propagated species, because – for example with *Calluna* – phenotypic variations (sports) are well-known and are based on genotypic variation. From these experiences we support EDV-identifying systems with respect to statistical validation as with the one introduced by [10] for lettuce and barley [11], which is based on the definition of a minimum distance (threshold) for distinctness. Such procedures are necessary since proving identity is more difficult than proving distinction with molecular markers [12]. Lettuce is a self-fertilizer and consequently genetic variation within varieties can be expected to be very low. Moreover total variation between today's cultivars should be somewhat reduced due to an intensive breeding history. For this reason, [11] suggested that ED-conflicts should not be analyzed through the construction of a dendrogram visualizing hypothetical kinship relations, but instead by the examination of all pairwise genetic distances within an appropriate reference population, and then comparing these results to the distance between actual varieties in question.

As a result, another aim of our study, drawing on the publication by [11], was to implement a comparable concept of identification of EDVs in *C. vulgaris* based on molecular data resulting from RAPD and iSSR techniques. Our system proposal is critically evaluated with regard to essential premises e.g. variation and stability [13], its success in *C. vulgaris*, and its practicability in the future.

The results presented here were obtained during a BMWi-(Federal Ministry of Economics and Technology) funded cooperation between the IGZ and a German breeding company (Heidepflanzen Peter de Winkel, http://www.dewinkel.de). Thus, variety denotation is ciphered in cases where the breeder's interests may be affected.

Results

Estimation of genetic diversity and kinship relations within C. vulgaris

Using RAPD- and iSSR-techniques, we achieved a total of 129 (RAPD) and 39 (iSSR) distinguishable and reproducible bands. This corresponds to 9.9 bands/RAPD primer and 7.8 bands/iSSR primer. The combined results of RAPD and iSSR studies are shown in the dendrogram in Fig. 2. While the three *Erica* genera do cluster as an outgroup, all tested genotypes from the *Calluna* species cluster to the right of one node. Interestingly, the wild-types from Thuringia (Ruhla) and from the Italian Alps (San Remo) cluster as an additional outgroup within the *Calluna* species while the other wild-types available (Löhnstein, Niederohe, Tiefenthal, all from the Lüneburger Heide in Germany) are grouped within the rest of the *Calluna* genotypes.

The statistical significance of our data was investigated with the resampling method of bootstrapping as initially described by [14] using the software Winboot and n = 10,000 replications. Those few nodes with moderate support (50% < p < 85%), as well as strong support (p ≥ 85%), which appeared both in the NTSYSpc-constructed tree as well as in the majority-rule consensus tree of Winboot, are marked with * and **, respectively, in Fig. 2. The linked genotypes to the right of these nodes may be considered to be linked in real kinship. Despite the high number of analyzed bands, all other linkages are statistically unconfirmable within the present data set.

Identifying EDVs in C. vulgaris

Due to former juridical conflicts concerning property rights of varieties in the genus *Calluna* we endeavored to develop a reliable statistical system for identifying EDVs in this species based on the results from the first part of our study. Since the dendrogram analysis did not support statistically significant decisions on kinship relations and probably would not do so even after analysis with a clearly expanded data set, we decided to implement a method based on a procedure published by [11] for similar analyses in lettuce. We therefore created appropriate Reference Sets of 25 varieties (Table 1) for each pair of tested genotypes (Test Set) in question and then computed primer-wise and pair-wise similarity values within each set. The Test Sets were chosen to represent non-ambiguous EDV or clear non-EDV cases for proof of concept, as well as several cases of interest in *Calluna* (Table 2). This non-ambiguousness was derived from personal communications with the involved breeding company in case of the EDV-pair. The test of a BC_1 against the parents as a clear non-EDV case was performed with our own crossings.

After extensive testing we selected a threshold provided by the highest Dice value of the 98% lowest values of all pairwise comparisons within the reference set (Fig. 3). This threshold was chosen in order to prevent the BC_1 individual from being categorized as essentially derived from the backcross parent which constitutes an essential prerequisite for validation of our test since backcrossing is the normal breeding system in bud-flowering *Calluna*. The 98% thresholds in both Reference Sets differ due to the necessary adjustment of the reference set according to the test in question (exchange of wild-type genotypes against varieties from the upper cluster of the dendrogram): 98%-Set A: 0.865 Dice similarity value, 98%-Set B: 0.893 Dice similarity value.

BMC Genetics 2008, **9**:56 http://www.biomedcentral.com/1471-2156/9/56

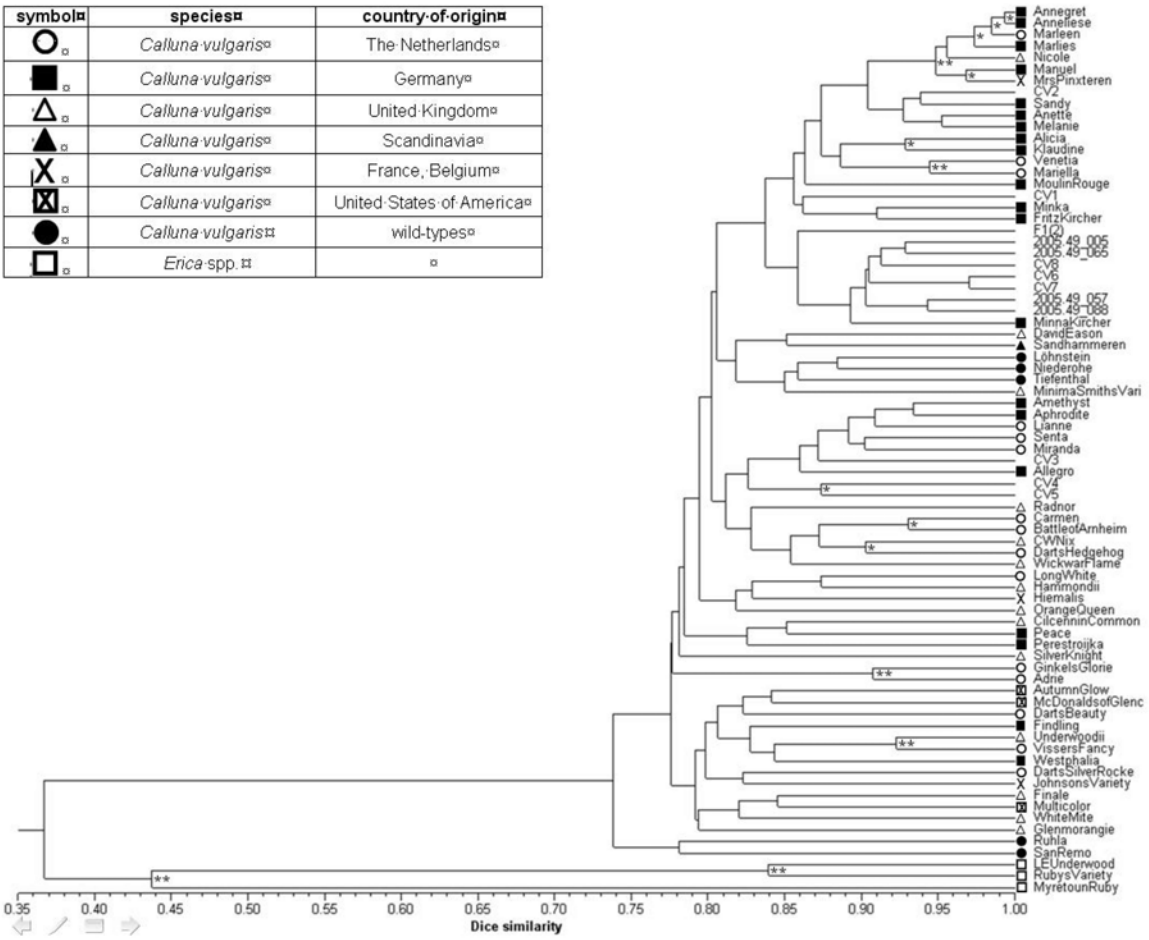


Figure 2

Dendrogram consisting of 74 *C. vulgaris* and 3 *Erica* spp. genotypes. Constructed from 168 mono- and polymorphisms amplified from 13 RAPD and 5 iSSR-primers and based on the Dice/Nei and Li coefficient with subsequent UPGMA-clustering. Nodes with strong support (> 85%) by bootstrapping (n = 10.000, PHYLIP) are marked with **, moderately supported groups (50% – 85%) are marked with *, varieties of interest for the involved company are ciphered by CV# where # is replaced by increasing numbers. Variety encryption is known to the authors and the company, respectively. For purposes of clarity and according to their regional provenance, genotypes have been classified by symbols as indicated.

For proof of concept we tested, on the one hand, one pair of individuals ('Maria' and *Maria Hell*), from which it was known that the latter was derived from the first one. On the other hand, an individual from a backcross progeny was tested against both parents, which should result in the categorization of being non-derived. As expected, the first result was positive and the second one negative, using the threshold as given above (Fig. 3). Moreover, similarity between the BC_1 individual and the backcross parent was clearly higher than between the BC_1 individual and the second parent. The Dice value of the comparison with the backcross parent was actually slightly above the threshold; however, overlapping error bars indicated that the similarity was nevertheless not sufficiently high for these two genotypes to be categorized as essentially derived.

Regarding the 'true tests', the results were negative for several pairs of morphologically similar cultivars from different breeders (Fig. 3), as well as for wild genotypes of different origin (Fig. 4). In contrast, when testing the cultivars 'Melanie' and 'Anette', their genetic similarity was found to exceed the threshold, thus confirming the public

Table 1: Identifying essential derivation in *C. vulgaris*.

	Reference Sets	
	Set A	**Set B**
1	Niederohe	'Sandy'
2	San Remo	'Annegret'
3	'Adrie'	
4	'Allegro'	
5	'Boskoop'	
6	'Carmen'	
7	'C. W. Nix'	
8	'Dark Beauty'	
9	'Findling'	
10	'Glenmorangie'	
11	'Johnson's Variety'	
12	'Long White'	
13	'Mariella'	
14	'Marlies'	
15	'McDonalds of Glencoe'	
16	'Minima Smith's Variety'	
17	'Mrs. Pinxteren'	
18	'Multicolor'	
19	'Orange Queen'	
20	'Peace'	
21	'Radnor'	
22	'Sandhammeren'	
23	'Silver Knight'	
24	'Underwoodii'	
25	'Wickwar Flame'	

Identifying essential derivation in *C. vulgaris*: Reference-Sets A and B chosen for the Test-Sets. Set B differs from A only in the exchange of two wild-type genotypes (Niederohe, San Remo) versus 2 varieties ('Sandy', 'Annegret') as indicated by summarizing both columns of Set A and B for the residual 23 varieties.

data supplied by the BSA according to which 'Anette' is a sport of 'Melanie'. This was also confirmed for 'Melanie' vs. 'Sandy' and 'Annegret' vs. 'Anneliese' (Fig. 3).

The last test concerned a pair of cultivars ('Fritz Kircher' vs. CV7) which have in a former, non-public study been characterized as being essentially derived from one another using dendrogram analysis. In our investigation, however, their genetic similarity is lower than the threshold, thus clearly indicating an absence of essential derivation.

Discussion

Until now, molecular data on genetic diversity within the species *C. vulgaris* was only available for regionally restricted wild-type populations [15-17], not for varieties used in commercial breeding. An actual and urgent necessity for a comprehensive study in *C. vulgaris* can be deduced from several points: the number of applications for variety protection is currently increasing considerably, whereas the information given in the registration schedules is at least occasionally unreliable or equivocal (e.g. a bud flowering variety is said to be the result of selfing of another bud-flowering genotype, which is biologically impossible due to the total loss of anthers in bud-flowering genotypes). This leads to an ambiguous situation with regard to variety derivation. Additionally, molecular data are needed for concerted breeding works and the elimination of coincidence in this process.

Since it is technically simple to accomplish and requires no a priori sequence information, iSSR- and RAPD-PCR [18-20] are widely used techniques in different species; but RAPDs in particular may be 'considered the practice of PCR without a clue' [21]. All the same, both techniques

Table 2: Identifying essential derivation in *C. vulgaris*.

#	Test-Sets		selection criteria	Reference Set	Hypothesis	Result
1	'Maria'	*Maria Hell*	Maria Hell = known sport of 'Maria' according to information from a breeder	A	yes	yes
2	'Maria'	BC_1-individual	progeny testing	A	no	no
3	'Roter Oktober'	BC_1-individual	progeny testing	A	no	no
4	'Melanie	'Anette'	'Anette' = sport of 'Melanie' according to information given by BSA doc	A	yes	yes
5	'Melanie'	'Sandy'	'Sandy' = sport of 'Melanie' according to information given by BSA doc	A	yes	yes
6	'Annegret'	'Anneliese'	'Anneliese' = sport of 'Annegret' according to information given by BSA doc	A	Yes	yes
7	'Fritz Kircher'	CV7	re-testing results from former investigations	A	yes	no
8	'Karla'	'Venetia'	similar cultivars from different breeders	A	no	no
9	'Minka'	'Miranda'	similar cultivars from different breeders	A	no	no
10	SanRemo	Ruhla	wild-type testing	B	no	no
11	Niederohe	Löhnstein	wild-type testing	B	no	no
12	Niederohe	SanRemo	wild-type testing	B	no	no

Test Sets and their criteria for selection, the applied Reference Set and our initial hypothesis with regard to whether or not ED was to be expected. The first three tests were used as proof of concept, meaning consistency of hypothesis and validation of the eligibility of the method; tests 4–12 are true testings.

BMC Genetics 2008, **9**:56 http://www.biomedcentral.com/1471-2156/9/56

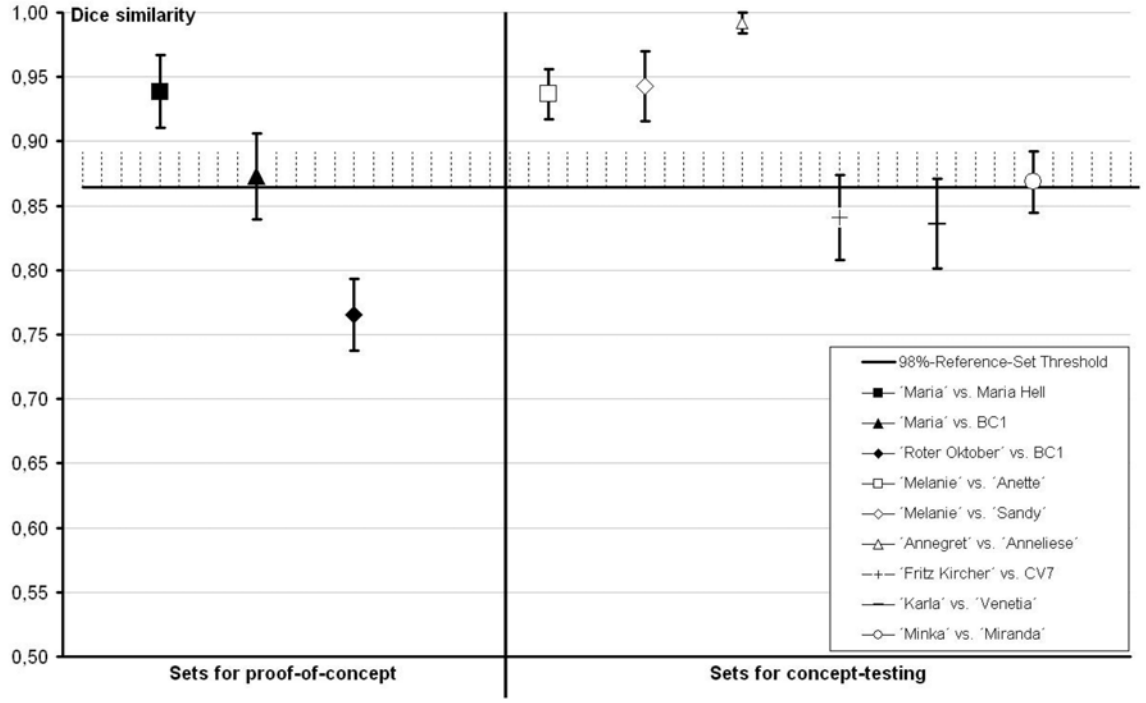


Figure 3
Identifying essential derivation in *C. vulgaris* I. Validation of the method using three sets for proof-of-concept: One set of a known essentially derived variety pair and two sets of genotypes involved in backcrossing, marked by black symbols. Additionally six pairs of varieties of interest have been tested against the chosen threshold of 0.865 Dice similarity value, which was derived from Reference Set A.

provide a uniformly distributed amplification of DNA fragments throughout the genome of eukaryotic organisms due to the nature of their origin, and were shown to be an adequate molecular tool for studying DNA polymorphisms (e.g. [22,23]). The same was true for our investigations as we observed very robust inner-laboratory reproducibility: here, a value as low as 0.46% of missing data within the 77 × 168 similarity matrix was achieved.

The dendrogram resulting from the combined computation of both RAPD and iSSR banding patterns showed a low genetic variability within the species *C. vulgaris*: almost all tested varieties and genotypes are grouped at a Dice/Nei & Li similarity value of 0.80, or even higher. This confirmed our hypothesis of a narrow gene pool, which was expected by the breeding experiences and methods applied of the participating company (personal communications) and its competitors. Moreover, one has to bear in mind that *C. vulgaris* is the only species within the genus *Calluna* and that crossing with other genera of the *Ericaceae* is thus impossible, thereby assisting in the conservation in nature, too, of a slender genetic diversity. We esteem the clear discrimination between *Erica* and *Calluna* as one argument of reassurance for our methodological approach and consider the dendrogram to be unbiased in the sense of an essential prerequisite for picturing genetic data [11]. The fact that wild type genotypes from the Lüneburger Heide are grouped this near to economically important varieties is another piece of evidence for our line of argument in respect of a significantly narrow gene pool in *C. vulgaris*. In addition, to our knowledge, breeding in *C. vulgaris* began in exactly this area of Germany by collecting incidentally originated bud-flowering genotypes. Our results might thus confirm this hypothesis, especially since the wild types from Thuringia and the Italian Alps do not cluster within this group.

Another interesting feature of the resulting dendrogram is that the data were insufficient to support more than the few marked nodes (marked with * or ** in Fig. 2) as statistically significant. However, we do not consider the amount of bands i.e. mono-/polymorphisms from our

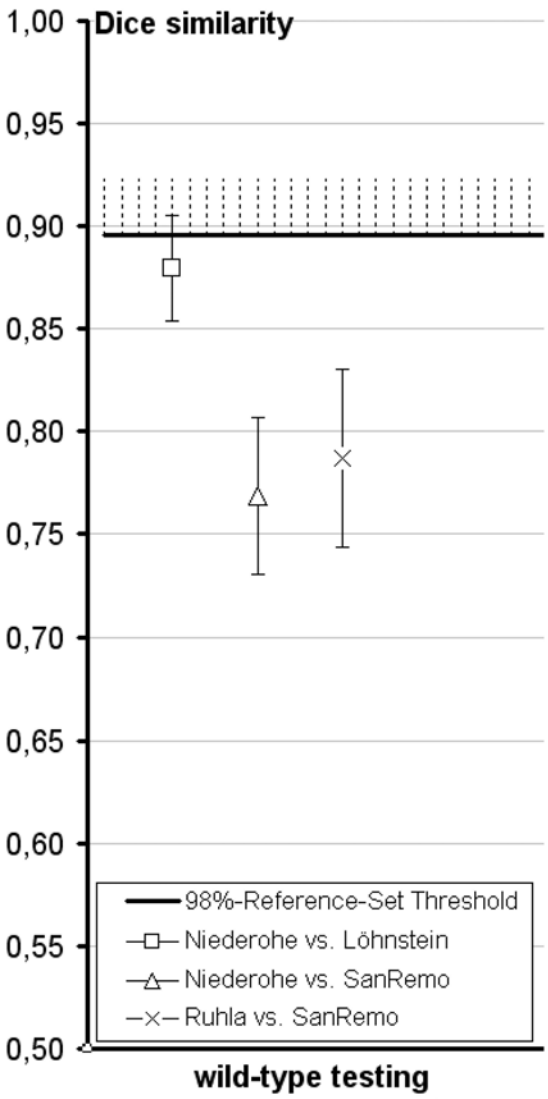


Figure 4
Identifying essential derivation in *C. vulgaris* II. Test of three pairs of wild types of different origin using Reference Set B (0.893 Dice similarity value).

data as generally too sparse, since [24] showed that an estimation of diversity within one population using approx. 200 dominant (i.e. AFLP) markers is as efficient as using 50 codominant (i.e. microsatellite) markers. Therefore, it is our suspicion, that the dendrogram method is not suitable for EDV identification in species with narrow gene pools.

ED issues arise for varieties that successfully passed DUS testing. An EDV is (i) predominantly derived from an initial variety, (ii) clearly distinguishable from it and except for these differences (iii) conforms to the initial variety in the expression of essential characteristics [5]. We consider it to be of paramount importance to apply a well adjusted system for identification of these EDVs for each species, and in our case for *C. vulgaris*, since the range of similarities presented in Fig. 2 proved the hypothesis of some breeders that the economically important varieties (and the genus *Calluna* in general) are closely related and thereby may readily lead to ED disputes, as has already been the case in the past.

As explained above, construction of a dendrogram proved to be no satisfactory tool for EDV identification in *C. vulgaris* – contrary to the results obtained for other vegetatively propagated species presented by [9] for *Phalaenopsis*, *Rosa* and *Rhododendron*. Another example is given by [25]. Using AFLPs, they proved that *Rosa* × *hybrida* original varieties are not more closely linked than 0.80 Jaccard's index. In contrast, the genetic similarities in so-called mutant groups were always higher than 0.96 (but not 1.0). Their dendrogram assay is therefore correctly rated as a suitable method to unambiguously distinguish rose EDVs from their initial variety. In addition, the detection of polymorphisms between sports and the original variety may be considered somewhat coincidental since molecular markers only cover a small portion of the target organism's genome. [26] demonstrated, that in cut roses RAPD-polymorphisms between a variety and its sports did occur in two varieties, but were not reproducible. Using AFLPs the authors were even able to amplify stable polymorphisms in sports of another variety. However, they were still able to distinguish vegetatively and sexually propagated progenies, since amplification in seedlings constantly resulted in a higher number of polymorphisms.

We ascribe our differing results to the coincidence of two phenomena in *C. vulgaris*. First, stable genetic conditions – which could be reasonably anticipated for vegetatively propagated species – are worthy of discussion in the context of *Calluna*, since the phenomenon of sport/reversion (a type of somatic mutation) is well-known by breeders.

Moreover, the very narrow gene pool in *C. vulgaris* gives rise to high genetic similarities, even if a new variety was obtained through crossing, due to the fact that even quite different individual plants, e.g. a wild type from the Lüneburger Heide and a bud flowering variety, show a considerable proportion of monomorphic bands in RAPD and iSSR analyses. Such lack of genetic diversity is our main reason for focusing on a system for EDV-identification involving a reference-set, as this is the important difference to e.g. the rose cases mentioned above: even in

Rosa × *hybrida* more than 10,000 varieties exist, resulting from some 150 years of breeding efforts [25], and they are still clearly distinguishable. The opposite situation is, in fact, the result of the differing breeding methods applied in *C. vulgaris*: breeding for a common phenotype (bud-flowering) and repeated back-crossing are generally accepted reasons that promote the development of narrow gene pools [25].

By working with a system similar to that described for lettuce and barley by [11], we were successful in both, identifying well-known essentially derived genotypes as well as discriminating between a genotype resulting from back-crossing and its parents (Fig. 3). We considered these results as a proof of concept for our method and additionally analyzed other test-sets whose information of origin we regarded to be unreliable, questionable or simply of interest. Here, information on variety derivation was primarily confirmed by our method as outlined in table 2. Moreover, the system discriminated phenotypically similar varieties from different breeders as well as wild genotypes of different origin, thus also confirming the hypotheses.

Conclusion

As a result of these findings, we would like to suggest the outlined method as an appropriate system for EDV-testing in *C. vulgaris*. Applicability to other vegetatively propagated crops should be tested, as well as the combined use of 'fixed' and 'random/unmapped markers' as suggested by [11]. Moreover, we recommend the inclusion of at least three independent gDNA isolations of different individuals per genotype, since inner-varietal identity cannot be presumed and is hard to verify, even in vegetatively propagated crops.

Methods

DNA techniques: isolation of genomic DNA (gDNA)

gDNA of *C. vulgaris* genotypes was isolated according to [27]. About 200 mg young leaf tissue (stored over night and frozen in liquid nitrogen) was homogenized in 2 ml tubes in a mixer mill (MM301, Retsch) using 2 stainless steal balls (Ø = 5 mm). The tissue was resuspended in buffer A (50 mM Tris-HCl pH 8.0, 5 mM EDTA, 350 mM sorbitol, 1% β-mercaptoethanole, 10% PEG-6000) and centrifuged for 1 min at 4°C and 8,000 rpm (Sigma 3K30, rotor-no. 12148). The resulting pellet was again resuspended in buffer B (50 mM Tris-HCl pH 8.0, 5 mM EDTA, 350 mM sorbitol, 1% β-mercaptoethanole, 1% sodium-sarcosyle, 0.1% CTAB, 710 mM NaCl) and incubated for 30 min at 60°C. After adding 0.8 volumes chloroform-isoamyl alcohol 24:1, the samples were centrifuged for 15 min at 4°C at 15,300 rpm. The supernatant was transferred to a new 2 ml reaction tube and incubated at -20°C for 30 min after adding 0.75 volumes isopropanole. After centrifugation (5 min at 4°C at 5,000 rpm) the pellet was washed with 70% ethanol, air-dried and resuspended in 500 µl TE buffer. To each sample, 1 ng RNAse (Carl Roth GmbH) was added, followed by incubation for 15 min at 37°C. Subsequently, phenol-chloroform extraction step was performed twice and the resulting supernatant containing purified gDNA was pelleted at -20°C for 60 min after addition of 0.1 volumes 3 M sodium acetate and 0.75 volumes isopropanole. The precipitated gDNA was washed twice with 70% ethanol, air-dried and resuspended in 100 µl TE. Long-time storage was achieved at -20°C.

DNA techniques: PCR amplification and electrophoresis

Amplifications of RAPD-fragments generated from random decamer primers (Carl Roth GmbH) were performed in a Primus 96 advanced thermocycler (peqlab GmbH) using the following protocol: 5 min at 95°C, [1 min at 95°C, 1 min at 35°C, 1 min at 72°C]$_{35\times}$, 10 min at 72°C. The reaction mixture for a total volume of 25 µl contained 1× reaction buffer, 2.5 mM $MgCl_2$, 1 U Taq.-DNA Polymerase (recombinant, Invitrogen), 0.2 mM of each dNTP (Invitrogen), 0.5 µM primer (Carl Roth GmbH, MWG Biotech AG), 10 ng gDNA and the adequate amount of sterile deionized H_2O.

Amplification of iSSR-Fragments was performed following the same protocol as described for RAPDs, with the altered annealing temperatures according to primer length. Table 3 provides an overview of primers used in this study; these were chosen after screening 60 decamer primers for reproducibility. Decamer primers were obtained as random primer kits from Carl Roth GmbH; iSSR primers were synthesized by MWG Biotech AG. iSSR primers given in Table 3 were chosen by referring to the common di- and trinucloetide motifs in plants $(AC/TG)_n$ and $(AAG/TTC)_n$ (e.g. [28]).

Electrophoretic separation of the amplification products was performed in 23 × 25 cm 1.5% agarose gels by applying 7 V/cm for 2.5 hours. The gel contained ethidium bromide for visualization of fragments at 254 nm. Documentation was carried out with a digital imaging system (Biostep GmbH).

Reactions were repeated at least twice before fragments were used for distance calculations.

Statistics: gel analysis and phylogenetic calculations

Gel analysis (band detection, noise reduction, size calibration, fragment matching) was performed with the Phoretix 1D Advanced software (Nonlinear Dynamics). The selection of bands derived from each primer was performed by objective criteria (e.g. thresholds for routine band detection and matching and recommendations to ensure reproducibility, e.g. the exclusion of fragments of very high and very low size). The banding data were trans-

Table 3: List of iSSR- and RAPD-primers.

type	denomination	sequence (5' → 3')	source
iSSR	17898B	$(CA)_6$-gT	according to [21]
	17898C	$(CA)_6$-AC	
	17899	$(CA)_6$-gg	
	17901B	$(gT)_6$-TT	
	P02	$(AAg)_6$-Cg	according to [28]
RAPD	RX13	ACgggAgCAA	random primer kits
	RX14	ACAggTgCTg	
	RY01	gTggCATCTC	
	RY13	gggTCTCggT	
	RY15	AgTCgCCCTT	
	RY16	gggCCAATgT	
	RY17	gACgTggTgA	
	RY18	gTggAgTCAg	
	RZ04	AggCTgTgCT	
	RZ05	TCCCATgCTg	
	RZ07	CCAggAggAC	
	RZ12	TCAACgggAC	
	RZ17	CCTTCCCACT	

List of primers used for PCR-amplification of mono- and polymorphic fragments within gDNA of *C. vulgaris*, their sequence, and the source according to which the sequences were selected.

formed to a computable 0/1 matrix in the common [OTU × band] layout.

Phylogenetic as well as dendrogram calculations were conducted with the NTSYSpc 2.20 L software (©1986–2006, Applied Biostatistics Inc.). Qualitative banding values were computed using the *SimQual* module with the similarity coefficient of Dice [29] and Nei and Li [30], respectively. Subsequent UPGMA clustering was conducted within the *Sahn* module, while the module *Treeview* was used to visualize the data set as a dendrogram.

For Bootstrapping using the Dice coefficient, Winboot [31] was used (replications given in the text) which finally constructs a majority-rule consensus tree based on the *Consense* module of the PHYLIP software.

EDV-testing by application of the tail principle

A system first published by [11] for EDV identification in lettuce and barley was adapted for *C. vulgaris* as follows. Since a priori pedigree information is unavailable for *C. vulgaris* and the application of the *pedigree principle* – a threshold selection based upon inclusion of 'identity by descent' probabilities – was not possible, we selected the *tail principle* for the identification of a threshold from a distribution of pair-wise similarities from a reference-set. The configuration of the Reference Set (25 varieties) for each Test Set (2 varieties) was adapted by using the information gained from our phylogenetic results which matches the integration of the *calibration principle*. In the context of [11], we decided not to include known EDVs in this set, e.g. mutation-derived varieties or sports, since these would only represent extremely high values within the reference-set and would complicate data interpretation. Both polymorphic and monomorphic markers were analyzed. Detached primer-wise computation of similarity values becomes applicable by assuming a genetic independence between primers and a uniform distribution of primer binding sequences throughout the genome. Arithmetic means, standard deviations, standard errors as well as its medians were calculated according to [10] and [11] in analyzing the inner- and inter-set-similarities and also to define a threshold for identifying EDVs in *Calluna*. This threshold was positioned in such a way that varieties known to be ED exceed the threshold and non-EDVs do not.

As carried out before for RAPD and iSSR values, these data were computed for their similarity values with the coefficient of Dice (Nei & Li) by NTSYSpc 2.20 L. Only in case of Dice similarity values above the threshold and non-overlapping error bars, a pair of genotypes is categorized as being essentially derived. In a case where similarity values are below the threshold or where they are above but with overlapping error bars, a pair of genotypes is categorized as not essentially derived.

Abbreviations

AFLP: Amplified Fragment Length Polymorphism; ASSINSEL: 'International Association of Plant Breeders for the Protection of Plant Varieties'; EDV: Essentially Derived Variety/Varieties; BSA: Bundessortenamt; CPVO: Community Plant Variety Office; gDNA: genomic DNA; ISF: Inter-

national Seed Federation; iSSR: inter Simple Sequence Repeats; RAPD: Randomly Amplified Polymorphic DNA; UPOV: International Union for the Protection of new Varieties of Plants; UPGMA: Unweighted Pair Group Method with Arithmetic Mean

Authors' contributions

After methodological setup, TB carried out the complete RAPD section from laboratory work to analysis and drafted the manuscript. JK performed the complete iSSR part from laboratory work to analysis. AH designed the study and participated in drafting the manuscript.

All authors read and approved the final manuscript.

Additional material

Additional file 1

Complete list of included varieties and genotypes, their country of origin and pedigree information where known (source is given in the last column). *Column 3 defines the flower type either as normal, bud, multibracteate (multi) or filled. Sources of information are either the Bundessortenamt (BSA:doc), the appropriate website (web) of the 'The International Register of Heather Names' http://www.heathersociety.org.uk/handy_guide.html or personal communications (personal contact).*
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2156-9-56-S1.xls]

Acknowledgements

This study was carried out as part of the BMWi-funded project (KP0172401BN5A) between the IGZ and Heidepflanzen Peter de Winkel, Goch http://www.dewinkel.de. The authors wish specifically to thank the company owner for his support and especially Katja Krüger for her sustained and motivating assistance throughout the daily laboratory work. Additionally, we would like to thank Prof. Dr. T. Debener (University of Hannover) for his support concerning the experimental design and the manuscript review.

References

1. **ZMP: Zentrale Markt- und Preisberichtstelle für Erzeugnisse der Land-, Forst- und Ernährungswirtschaft GmbH** [http://www.zmp.de]
2. **BSA doc** [http://www.bundessortenamt.de]
3. **CPVO** [http://www.cpvo.fr]
4. **UPOV: International convention for the protection of new varieties of plants** [http://www.upov.int]
5. Staub JE, Gabert A, Wehner TC: **Plant Variety Protection: a consideration of genetic relationship.** *HortScience* 1996, **31:**1086-1091.
6. **ASSINSEL: Position Paper on DUS testing: Phenotype vs. Genotype** [http://www.worldseed.org/en-us/international_seed/on_intellectual_property.html]
7. Lombard V, Dubreuil P, Dillmann C, Baril C: **Genetic distance estimators based on molecular data for plant registration and protection: a review.** *Acta Hort* 2001, **546:**55-63.
8. Eeuwijk FA, Baril CP: **Conceptual and statistical issues related to the use of molecular markers for distinctness and essential derivation.** *Acta Hort* 2001, **546:**35-53.
9. de Riek J, Dendauw J, Leus L, de Loose M, van Bockstaele E: **Variety protection by use of molecular markers: some case studies.** *Plant Biosystems* 2000, **135(1):**107-113.
10. **ASSINSEL/ISF: Guidelines for the Handling of a Dispute on EDV in Lettuce** [http://www.worldseed.org/en-us/international_seed/on_intellectual_property.html]
11. Eeuwijk FA, Law JR: **Statistical aspects of essential derivation, with illustrations based on lettuce and barley.** *Euphytica* 2004, **137:**129-137.
12. Ibañez J: **Mathematical analysis of RAPD data to establish reliability of varietal assignment in vegetatively propagated species.** *Acta Hort* 2001, **546:**73-79.
13. Heckenberger M, Bohn M, Ziegle JS, Joe LK, Hauser JD, Hutton M, Melchinger AE: **Variation of DNA fingerprints among accessions within maize inbred lines and implications for identification of essentially derived varieties. I. Genetic and technical sources of variation in SSR data.** *Molecular Breeding* 2002, **10:**181-191.
14. Efron B: **Bootstrap methods: another look at the jackknife.** *The Annals of Statistics* 1979, **7:**1-26.
15. Mahy G, Ennos RA, Jacquemart AL: **Allozyme variation and genetic structure of *Calluna vulgaris* (heather) populations in Scotland: the effect of postglacial recolonization.** *Heredity* 1999, **82:**654-660.
16. Meikle A, Paterson S, Finch RP, Marshall G, Waterhouse A: **Genetic characterization of heather (*Calluna vulgaris* (L.) Hull) subject to different management regimes across Great Britain.** *Molecular Ecology* 1999, **8:**2037-2047.
17. Rendell S, Ennos RA: **Chloroplast DNA diversity in *Calluna vulgaris* (heather) populations in Europe.** *Molecular Ecology* 2002, **11:**69-78.
18. Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV: **DNA polymorphisms amplified by arbitrary primers are useful as genetic markers.** *Nucleic Acid Research* 1990, **18:**6531-6535.
19. Welsh J, McClelland M: **Fingerprinting genomes using PCR with arbitrary primers.** *Nucleic Acid Research* 1990, **18:**7213-7218.
20. Zietkiewicz E, Rafalski A, Labuda D: **Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification.** *Genomics* 1994, **20:**176-183.
21. Wolfe AD, Liston A: **Contributions of PCR-based methods to plant systematics and evolutionary biology.** In *Molecular systematics of plants II – DNA sequencing* Edited by: Soltis DE, Soltis PS, Doyle JJ. Kluwer Academic Publishers, Boston-London; 1998.
22. Bagley MJ, Anderson SL, May B: **Choice of methodology for assessing genetic impacts of environmental stressors: polymorphism and reproducibility of RAPD and AFLP fingerprints.** *Ecotoxicology* 2001, **10:**239-244.
23. Garcia AFA, Bechnimol LL, Barbosa AMM, Geraldi IO, Souza CL Jr, de Souza AP: **Comparison of RAPD, RFLP, AFLP and SSR markers for diversity studies in tropical maize inbred lines.** *Genetics and Molecular Biology* 2004, **27(4):**579-588.
24. Mariette S, Lecorre V, Kremer A: **Sampling within the genome for measuring within-population diversity: trade-offs between markers.** *Molecular Ecology* 2002, **11(7):**1154-1156.
25. Vosman B, Visser D, Voort JR van der, Smulders MJM, van Eeuwijk F: **The establishment of 'essential derivation' among rose varieties, using AFLP.** *Theor Appl Genet* 2004, **109:**1718-1725.
26. Debener T, Janakiram T, Mattiesch L: **Sports and seedlings of rose varieties analysed with molecular markers.** *Plant Breeding* 2000, **119:**71-74.
27. Kobayashi N, Horikoshi T, Katsuyama H, Handa T, Takayanagi K: **A simple and efficient DNA extraction method for plants, especially woody plants.** *Plant Tiss Cult Biotech* 1998, **4:**76-80.
28. Wolfe AD, Xiang QY, Kephart SR: **Assessing hybridization in natural populations of *Penstemon* (Scrophulariaceae) using hypervariable intersimple sequence repeats (ISSR) bands.** *Molecular Ecology* 1998, **7:**1107-1125.
29. Dice LR: **Measures of the amount of ecologic association between species.** *Ecology* 1945, **26:**297-302.
30. Nei M, Li WH: **Mathematical model for studying genetic variation in terms of restriction endonucleases.** *Proc Natl Acad Sci USA* 1979, **76:**5269-5273.
31. Yap IV, Nelson RJ: **Winboot: a program for performing bootstrap analysis of binary data to determine the confidence limits of UPGMA-based dendrograms.** *Manual* 1996.

2.1.1. Annex I: Genetic Diversity

Due to the clustering of genotypes above a genetic similarity value of ~ 0.74%, only few statistically supported subgroups with relevance for the breeding of *C. vulgaris* were identifiable within the dendrogram (compare Figure 2, Borchert et al. 2008). Only the three *Erica* spp. and the wild-types from San Remo and Ruhla were clearly recognizable as distinct outgroups.

Thus, alternative modes of data presentation e.g. of genetic distance / similarity values are required, e.g. two- or three-dimensional patterns resulting from Principal Component Analysis (PCA) which may allow the extraction of additional informative data. Due to the high genetic identities, the three-dimensional plot is barely interpretable (data not shown) Thus, Figure 5 resulted from the transformation of the original distance matrix used for dendrogram construction into only two units-less dimensions by application of the Pearson coefficient (NTSYSpc).

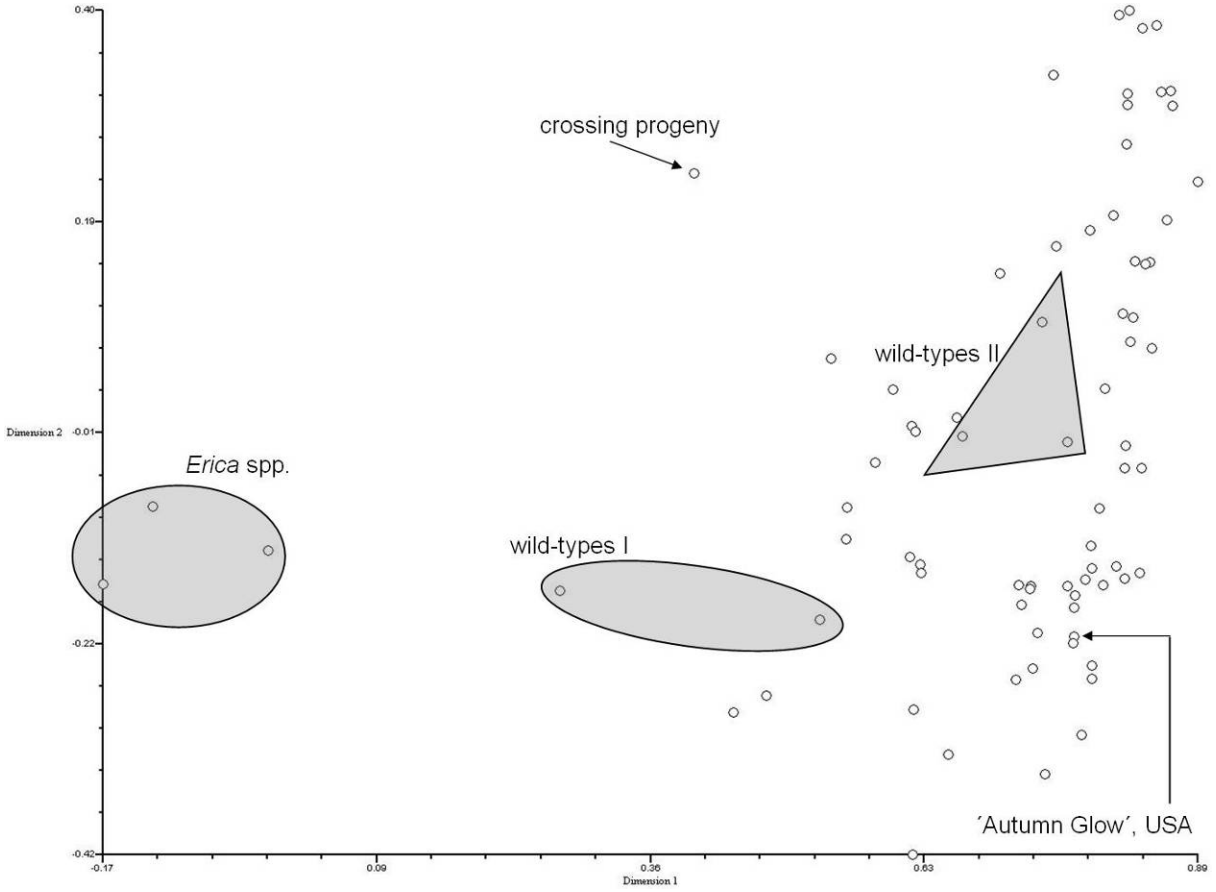


Figure 5: two-dimensional PCA of the *C. vulgaris* RAPD- and ISSR-dataset, formerly used for dendrogram construction. Coefficient: Pearson product-moment correlation (NTSYSpc) *Erica* spp.: *E. carnea*, *E. tetralix*; wild-types I: Ruhla, San Remo; wild-types II: Löhnstein, Niederohe, Tiefenthal (Lüneburger Heide)

Likewise to the dendrogram, the included *C. vulgaris* genotypes cluster once again. The set ‘wild-types I’ comprises Ruhla and San Remo, genotypes that were collected in Thuringia, Germany and near San Remo, Italy, respectively. The ‘wild-type II’ group includes three genotypes from the Lüneburger Heide: Löhnstein, Niederohe and Tiefenthal. The *Erica* spp. outgroup is characteristically separated from all other genotypes.

In addition, a crossing progeny resulting from a cross of parent varieties that are included within this analysis is visibly separated from the *C. vulgaris* cluster. This is contradictory to the original dendrogram, in which this distinct individual is grouped near other progenies of the same crossing. In addition, other dyads that were moderately supported by bootstrapping in the dendrogram are no longer arranged together after PCA. Furthermore, the single variety from the USA cannot be distinguished from other regional i.e. European subgroups.

Thus, the transformation of the data set by PCA did not result in a considerable increase of data resolution.

2.1.2. Annex II: Essential Derivation

Since the date of publication of this manuscript, several more variety pairs in question were tested according to the method initially described by Eeuwijk and Law (2004) and adapted to *C. vulgaris* as illustrated in Borchert et al. (2008).

Figure 6 shows several test sets of two independent back-crossing populations. For each population, the 'bud-flowering' and the 'wild-type' parent was tested vs. the F_1 genotype used for back-crossing. In addition, two randomly chosen individuals from the back-crossing population were tested vs. each of the 'bud-flowering' and the 'wild-type' parent. As a result and in accordance to the sets used for proof-of-concept, all sets of these samples were classified as non-ED-cases, since their mean values (including the error bars) were either below or within the threshold and its error margin.

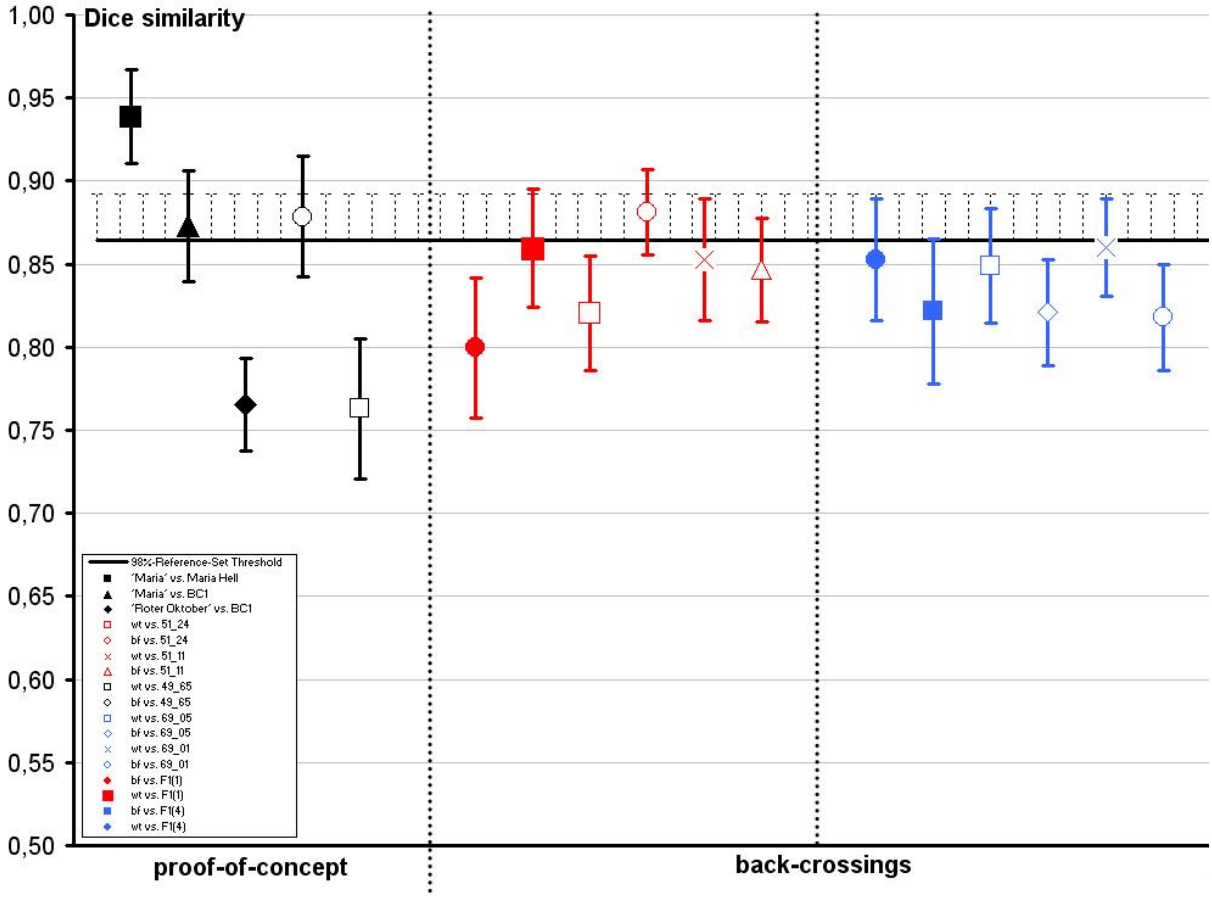


Figure 6: Identifying essential derivation in *C. vulgaris* III. Randomly selected individuals of two independent back-crossing populations, the parents of the crossing and the associated F_1 individuals were tested according to Borchert et al. (2008). bf: 'bud-flowering' parent; wt: 'wild-type' parent; F_1(x): individual of the F1 population used for back-crossing, x: internal number; back-crossing individuals are described by 'x_y', referring to an internal crossing number x and the consecutive individual number y; black, filled measurements: proof-of-concept sets as described in Borchert et al. (2008).

Figure 7 presents additional results of variety test sets, either of known EDV-cases or known non-EDV-cases. For almost all samples, the expectation was verified by the EDV test system. On the one hand, the means of known EDV-cases were located above the threshold lacking overlapping error bars. In contrast, the means of the test sets selected to represent known non-EDV-cases were positioned in the same area similar to the back-crossing test sets (Figure 6) and therefore have to be esteemed to be not essentially derived.

The test set 'Fritz Kircher' vs. 'Minna Kircher' resulted in overlapping error bars of the test set and the threshold. This case was included as 'known EDV' due to pers. comm. (Heidepflanzen de Winkel). Therefore, our results do not allow a definite status assignment for this variety pair.

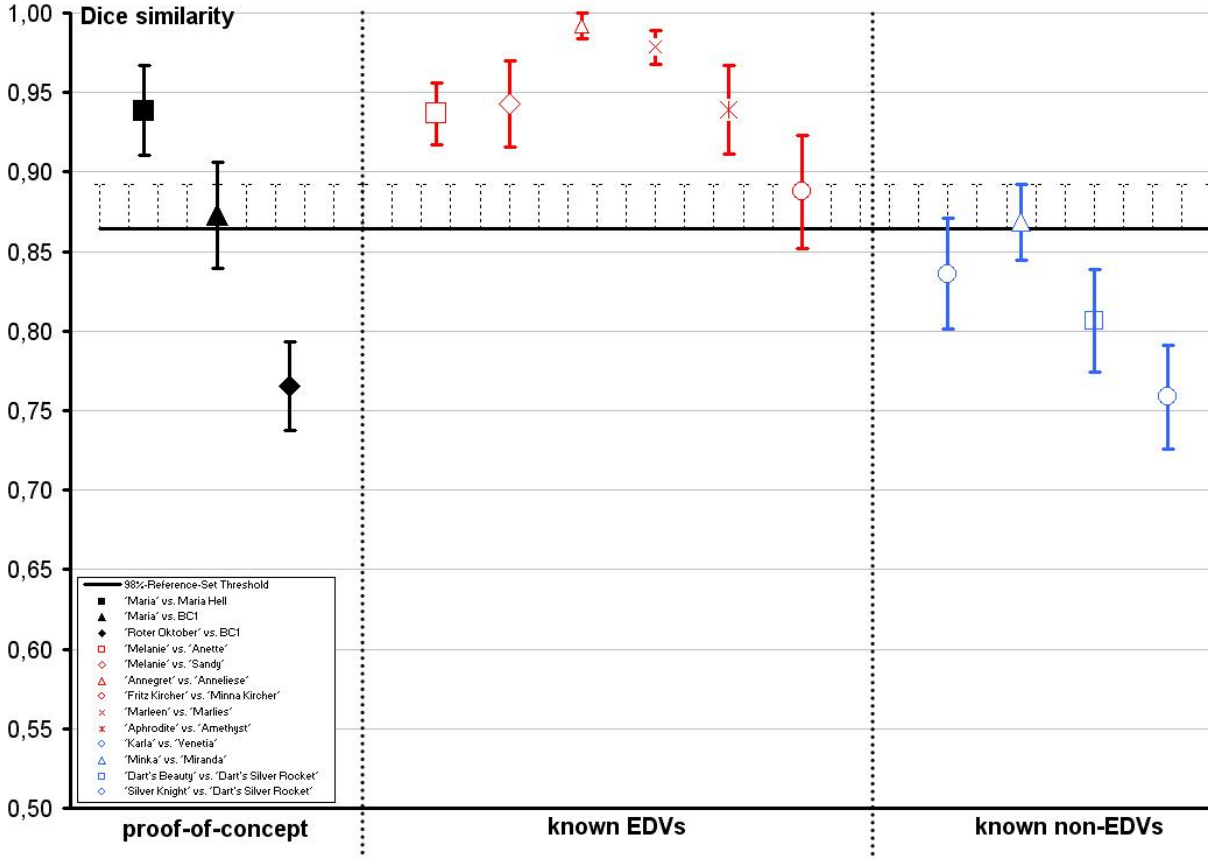


Figure 7: Identifying essential derivation in *C. vulgaris* IV. Known EDV-cases (identified from official documents, websites or pers. comm.) and known non-EDV-cases (as known from official documents) were tested according to Borchert et al. (2008). Black, filled measurements: proof-of-concept sets as described in Borchert et al. (2008).

Nevertheless, these additional results furthermore support the system of Eeuwijk and Law (2004) to be applicable to *C. vulgaris*, since progenies resulting from 'true breeding' i.e. backcrossing are distinguishable from EDVs.

Identification of molecular markers for the flower type in the ornamental crop *Calluna vulgaris*

Borchert T, Hohe A*

Leibniz-Institute of Vegetable and Ornamental Crops (IGZ)

Department of Plant Propagation

Kuehnhaeuser Str. 101

99189 Erfurt

Germany

www.igzev.de

*corresponding author:

E-Mail: hohe@erfurt.igzev.de

Phone: +49 (0) 36201 785 210

Fax: +49 (0) 36201 785 250

ABSTRACT

The establishment of a marker-assisted selection system for the economically important 'bud-flowering' phenotype in the ornamental crop *Calluna vulgaris* is of great interest to practical breeding companies, as it would allow selection at the juvenile stage.

Segregation analyses revealed a monogenic recessive inheritance of the bud flowering trait. Since in *C. vulgaris* only sparse molecular data are available, the search for molecular markers in a segregating backcross progeny was accomplished using PCR techniques based on random primers. Two candidate RAPD markers in coupling of the trait of interest were identified. Results on their applicability in different populations and independent varieties are presented. Their transformation capability to sequence characterized amplified region and single strand conformation polymorphism markers are described and discussed in the context of marker-assisted selection strategies in breeding of ornamental crops.

KEYWORDS:

- *Ericaceae*
- bulked segregant analysis
- marker-assisted selection
- ISSR
- RAPD
- PCR-SSCP

ABBREVIATIONS:

AFLP – amplified fragment length polymorphism
BLAST – basic local alignment search tool
BSA – bulked segregant analysis
CAPS – cleaved amplified polymorphic site
df – degree of freedom
EDV – essentially derived variety
ISSR – inter simple sequence repeat
MAS – marker-assisted selection
ORF – open reading frame
PAGE – polyacrylamide gel electrophoresis
RAPD – randomly amplified polymorphic DNA
RFLP – restriction fragment length polymorphism
SCAR – sequence characterized amplified region
SSCP – single strand conformational polymorphism

INTRODUCTION

Calluna vulgaris L. (Hull.) belongs to the *Ericaceae* family, and is the only species of this genus. It is an evergreen shrub common to Europe, South Africa and North America. The so-called 'bud-flowering' type (Fig. 1) is most important from an economic perspective. This phenotype is completely lacking stamens, which is assumed to be the reason why the flowers do not open, as reported for mutants of *A. thaliana* that have only poorly developed stamens (Ishiguro et al. 2001). This mutation in flower morphology causes a remarkable elongation of the flowering time compared to wild type genotypes, which makes *C. vulgaris* an attractive alternative to *Erica gracilis* as a bedding plant in autumn. Consequently, the economic importance of *C. vulgaris* has steadily increased lasting recent years. In Germany, *C. vulgaris* has a market volume of approx. 120 million EUR, and is third placed on the top ten selling list of the balcony and bedding plants in Germany (ZMP 2007).

Based on previous knowledge from crossing experiments, the 'bud-flowering' trait is assumed to be inherited in a Mendelian fashion as a monogenic and recessive trait. F_1 populations resulting from a crossing of a 'bud-flowering' genotype as the mother and a 'normal-flowering' genotype as the father only consist of 'normal-flowering' plants. Backcrossing an F_1 individual with the recessive, 'bud-flowering' parent results in segregation ratios of approximately 1 : 1 for the flower type in the backcross populations. *C. vulgaris* only flowers once a year in autumn from August to September, depending on the variety. The juvenile phase until flowering lasts approximately 4-6 months. For this reason, the breeding process in *C. vulgaris*, as described above, takes at least 2.5 years from the first crossing until the flowering of BC_1 individuals and the first possibility to identify and select 'bud-flowering' genotypes. In addition, 'bud-flowering' genotypes cannot be used as male parents, since they do not possess stamens, further restricting breeding possibilities.

Since selection for the flowering trait cannot be performed before the backcross individuals flower, all progenies have to be grown until flowering, which is expensive and time-consuming for breeding companies. The introduction of a marker-assisted selection (MAS) system in *C. vulgaris* that already allows selection for the 'bud-flowering' trait in the juvenile phase is therefore a promising approach. Since no sequence data are available in *C. vulgaris* to date, we chose the non-sequence-based marker systems of RAPDs (Randomly Amplified Polymorphic DNA, Williams et al. 1990, Welsh and McClelland 1990) and ISSRs (Inter Simple Sequence Repeats, Zietkiewicz et al. 1994) for the initial establishment of the MAS. In addition, RAPDs and ISSRs are easily established and facilitate the detection of polymorphisms with efficiency comparable to that of AFLPs (Amplified Fragment Length Polymorphisms, Powell et al. 1996).

Based on a segregating BC_1 population with 68 individuals, we developed two RAPD markers, which were additionally tested in independent cultivars and a different backcross population. Transformability into SCAR (sequence characterized amplified regions) and SSCP (single strand conformation polymorphism) markers was assessed.

The results presented here were obtained in the course of a joint project by the IGZ and a German breeding company (Heidepflanzen Peter de Winkel, www.dewinkel.de), funded by the BMWi (Federal Ministry of Economics and Technology). Thus, primer sequences are either ciphered or not named at all, respecting the economic interests of the funding breeding company.

MATERIALS AND METHODS

Plant material

The individual plants involved in primer screening were derived from the population ciphered by BC_1(i). This population results from a backcross of the varieties ´Melanie´ and ´Roter Oktober´ and comprises 40 'normal-flowering' and 28 'bud-flowering' genotypes, corresponding to a segregation ratio of 1 : 1.4. Population BC_1(ii), segregating 1 : 1.1 (62 'normal-flowering' and 57 'bud-flowering' individuals), was used for marker testing and results from a backcross between 'Maria' and 'Boskoop'. In order to assess the mode of inheritance, further populations were created: BC_1(iii) is a backcross of the varieties ´Anette´ and ´Roter Oktober´, and BC_1(iv) results from backcrossing ´Mariella´ and ´Boskoop´. Segregation ratios were tested against the anticipated segregation ratio for a monogenic-recessive inheritance using a X^2 goodness-of-fit test.

Isolation of genomic DNA

Genomic DNA of *C. vulgaris* genotypes was isolated by applying a modified CTAB protocol adapted from Kobayashi et al. (1998). This protocol and its modifications are described in detail in Borchert et al. (2008).

PCR amplification and standard electrophoresis

Amplifications of RAPD fragments generated from random decamer primers (Carl Roth GmbH) were performed in a Primus 96 advanced thermocycler (peqlab GmbH) using the following protocol: 5 min at 95°C, [1 min at 95°C, 1 min at 35°C, 1 min at 72°C]$_{35X}$, 10 min at 72°C. The reaction mixture for a total volume of 25 µl contained 1X reaction buffer, 2.5 mM $MgCl_2$, 1 U Taq.-DNA Polymerase (recombinant, Invitrogen), 0.2 mM of each dNTP (Invitrogen), 0.5 µM primer (Carl Roth GmbH, MWG Biotech AG), 10 ng genomic DNA and the adequate amount of sterile deionized H_2O. Electrophoretic separation of the amplification products was performed in 1.5% agarose gels, which contained ethidium bromide for the visualization of fragments at 254 nm. Documentation was carried out with a digital imaging system (Biostep GmbH). PCR using ISSR primers that refer to common di- and trinucleotide motifs in plants, e.g. $(AC/TG)_n$ (Wolfe et al. 1998), and SCAR primer combinations were conducted with adapted annealing temperatures. All other conditions were kept identical.

Bulked segregant analysis

Bulked segregant analysis (BSA) was carried out as initially described by Michelmore et al. (1991). Two DNA bulks of phenotypically defined individuals ('normal-flowering' and 'bud-flowering') were created from the backcross population BC_1(i) and PCR-screened with RAPD and ISSR primers. Primers that led to polymorphic fragments between both pools were tested using the single plants of the BC_1(i) population.

Fragment sequencing & primer design

Polymorphic DNA fragments resulting from amplification with 10 bp primers were excised and extracted from the gel using the original protocol of the Avgene Gel/PCR DNA Fragments Extraction Kit. These fragments were controlled for size consistency with gel electrophoresis and subsequently cloned using the original protocol from the Promega pGEM®-T Easy Vector Systems Kit (0.5 volumes of each component). Due to the low fragment concentration a ratio of 1:3 for vector : insert was applied. Positive clones were identified by standard blue/white screening on LB/Amp/IPTG/X-Gal plates. Sequencing reactions were performed by MWG Eurofins AG, Martinsried.

Using these sequences, SCAR primers were designed using the Primer3Plus design tool (Untergasser et al. 2007) and checked for consistency using the Sequence Manipulation Suite (Stothard 2000).

Single Strand Conformation Polymorphism (PCR-SSCP)

PCR samples i.e. SCAR fragments were separated on polyacrylamide gels. The gels were run at 18°C for 8 - 14 hours in 0.6X TBE buffer on a Li-Cor 4300 DNA Analyzer using standard AFLP plates, 0.25 mm spacers and 240 V, 20 mA, 2W. 20 ml of a 0.5X SSCP gel (1.2 ml Tris-borate-EDTA-buffer, 12 ml H_2O, 5 ml MDE Solution (mutation detection enhancement, Biozym), 1.7 ml 60% glycerine, 100 µL 10% ammoniumpersulfate and 12.5 µL N,N,N',N'-tetramethylethylenediamine) was poured between the glass plates, and hardened over night. An equal volume of loading dye (95% formamide, 0.01 M NaOH, 0.25% xylencyanole, 0.25% bromphenolblue) was added to each PCR sample, and the mixture was denatured for 5 min at 95°C directly prior loading onto the gel. The gels were silver-stained (45 min fixing: 7.5% acetic acid in H_2O; 2x 5 min washing in H_2O; 45 min staining: 1 g/L silver nitrate, 1.5 ml/L formaldehyde; 1x 5 min washing in H_2O; development for 2x 5 min in pre-cooled (4°C) developing solution: 30 g/L sodium carbonate, 1.5 ml formaldehyde, 1.6 mg sodium thiosulfate; stopping the reaction for 3 min in fixing solution; 5 min washing in H_2O). Image collection was conducted by scanning the gels at 360 dpi (Epson Perfection Scanner).

RESULTS

In order to identify the mode of inheritance of the trait of interest (‘bud-flowering’), four different back-crossing populations were analyzed regarding their segregation ratios. The results are presented in Table 1. The deviation from the expected 1 : 1 segregation ratio was not statistically significant in three of the four tested BC_1 populations ($df = 1$, $\alpha = 0.05$). The population BC_1(iii) showed a deviating segregation ratio presumably due to its low total amount of individuals. Thus, the anticipated monogenic-recessive inheritance of the ‘bud-flowering’ trait was not disproved.

For identification of molecular markers, a BSA (Michelmore et al. 1991) was applied to screen the segregating population BC_1(i). The DNA of ‘normal’ and ‘bud-flowering’ individuals of this population were pooled and subsequently screened for polymorphisms with 120 10 bp primers and 10 14-20 bp ISSR primers. Due to a loss of plants, only 15 (out of originally 40) ‘normal-flowering’ individuals could be tested, but all 28 ‘bud-flowering’ were available for screening purposes.

With this approach, two polymorphisms amplified by two different RAPD primers were identified. Both polymorphic fragments were amplified in DNA samples from ‘normal-flowering’, not from ‘bud-flowering’ genotypes. In contrast, all amplification products of the tested ISSR primers were monomorphic. From this point onwards, fragments, i.e. PCR-products derived from the RAPD primers, are described as ‘A-marker’ and ‘B-marker’, respectively. The A-marker can easily be identified due to its size of ~ 2.5 kb (Fig. 2). In contrast, it is more difficult to determine the B-marker (~ 1.2 kb) since the majority of ´negative´ individuals either display a fragment of only ~ 30 bp less than the marker. We therefore assume the A-marker to be of higher practical impact. By evaluating the single plants from test population BC_1(i), we confirmed that only one ‘bud-flowering’ genotype (2.3% of the total 43 genotypes) displayed the A-marker, thus representing a recombinant plant. No ‘bud-flowering’ genotype exhibited the B-marker and all ‘normal-flowering’ individuals were positive for both the A- and the B-marker.

In addition, we checked the amplification of both RAPD markers in 49 randomly selected *C. vulgaris* varieties, including wild-type genotypes, of which a total of 25 were of the 'normal' and 24 of the 'bud-flowering' phenotype (Table 2A, 2B). A coincidental distribution was observed here, since approx. 50% of the 'normal-flowering' varieties amplified both markers. Regarding the 'bud-flowering' varieties, only four of these (16.7%) amplified the A-marker, whereas nine (37.5%) amplified the B-marker. As expected, it was more difficult to analyze the B-marker than the A-marker. Therefore, using standard 1.5% agarose gels, false positive scoring results cannot be completely excluded in the case of the B-marker. Additionally, the A-marker was checked for its amplification in the backcross population BC_1(ii), which consists of 119 individuals. For this experiment, 49 individuals each of the 'normal' and the 'bud-flowering' type, respectively, were randomly chosen. However, only three 'normal-flowering' (6.1%) and two 'bud-flowering' (4.1%) individuals amplified the marker fragment. Both RAPD markers are therefore only functional within the initial backcross population in which they were originally detected.

To improve the marker quality i.e. to eliminate unspecific and irrelevant fragments from the PCR process, we attempted to convert the RAPD markers to Sequence Characterized Amplified Region (SCAR) markers. Both marker fragments were extracted from the gel (Fig. 3), cloned and sequenced. Two differing sequences were identified for clones of the B-marker, which may be the result of sequencing the two nearby fragments described above. One consensus alignment was identified for the A-marker, which includes the 5'-flanking RAPD binding site. BLASTX search (Basic Local Alignment Search Tool, Altschul et al. 1997) for the A-marker revealed the inclusion of two Open Reading Frames (ORFs), the first being 210 and the second 119 amino acids in length. The 210 amino acid ORF encoded a highly conserved Zn-Finger domain. These specialized types of Zn-fingers are assumed to be involved in mediating protein-protein interactions (see, e.g. Freemont 1993). The 119 amino acid ORF included the sequence of an NADH dehydrogenase I. One of the sequences for the B-marker included an ORF of 226 amino acids. However, no putative function could be attributed to this ORF. In summary, the three ORFs do not seem to be related to the trait of interest of this study.

Based on these sequences, 16 different 20 to 27 bp SCAR primers were designed to amplify fragments between 0.4 and 1.1 kb in length using the Primer3Plus tool. For both RAPD markers, the DNA bulks and the individual plants of the original test population BC_1(i) all displayed the SCAR fragment with the appropriate length. Thus, the original RAPD polymorphism was lost, as expected for SCAR markers that are amplified with primers nested within the original RAPD fragment. Interestingly, this was also the case when one primer covered the 5'-RAPD binding site of the A-marker.

The detection of a new polymorphism within the now monomorphic SCAR marker is an essential prerequisite for further practical applications of these markers. As one of several possible approaches, we chose to investigate the sequenced and amplified SCARs for Single Strand Conformation Polymorphisms (SSCP). Since the optimum fragment length for SSCP analysis is in the range below 500 bp, three SCAR fragments required restriction enzyme digestion previous to neutral polyacrylamide gel electrophoresis (PAGE). Digestion was conducted using the enzymes *Pst*I or *Eco*RI, depending on the number of cutting sites within the sequence and the length of the resulting fragments. Previous to PAGE, successful restriction digestion was controlled on agarose gels (data not shown). All SCAR products of the appropriate lengths were subsequently separated using the DNA Analysis System (0.25 mm gel thickness).

Only one primer combination from nine selected SCAR products (amplified from all available 43 individuals of the BC_1(i) population and separated using SSCP-PAGE) revealed banding patterns differing between 'normal'

and 'bud-flowering' individuals: in this case, silver staining visualized one additional fragment for the 'normal-flowering' genotypes (Fig. 4). This ~ 400 bp fragment is amplified between bases 900 to 1.300 of the RAPD A-marker sequence. None of the digested fragments revealed differing banding patterns. Since only one band appears on agarose gels and at least nine bands appear on SSCP gels, both primers were elongated from 20 bp to 27 bp in order to increase the specificity of the PCR process. Interestingly, this did even increase the absolute number of fragments, including one additional, shorter, polymorphic fragment that was again unique for 'normal-flowering' genotypes (Fig. 5). The amplification of the initial fragment was maintained. With regard to the individuals tested, we estimate the amount of deviating genotypes for the fragment amplified by the 20 bp SCAR primers at 13.9%. This value decreases to 9.3% if the elongated 27 bp SCAR primers are used. The additional fragment deviates in 6.9% of the 43 individuals of population BC_1(i).

DISCUSSION

The main target of this study was to initially establish an MAS system for the faster and cheaper selection cycles of the 'bud-flowering' trait in *C. vulgaris*. Our results prove the applicability of marker technologies for the potential introduction of such a system, since we identified two markers linked to the 'normal-flowering' trait in one backcross population. Both of these markers are linked in coupling phase to the 'normal flower' phenotype. We were also able to confirm the assumed monogenic and recessive inheritance of this character by analyzing the segregation ratios of four backcross populations: in three cases, no significant difference to the expected values was calculated. These four populations were chosen by reasons of their flowering rate at the time of assessment and the total population size. Populations displaying flowering rates below 90.0% were excluded just as populations that consisted of 45 individuals or less e.g. an available population of only 10 individuals that segregated exactly 1 : 1 was rejected to analyze.

Although a lower reproducibility of RAPDs compared to other techniques, e.g. AFLPs, is often reported (e.g. Jones et al. 1997, Heckenberger et al. 2003, Garcia et al. 2004), the technique proved to be sufficiently reliable in our hands (Borchert et al. 2008). This is in line with results of Bagley et al. (2001), who demonstrated that inner-laboratory RAPD reproducibility can be enhanced if some basic requirements in the course of daily work and in gel analysis are met. Moreover, reproducibility issues of RAPDs mainly result from heritability and not repeatability problems (Bagley et al. 2001).

Nevertheless, one approach to deal with common RAPD-associated problems, i.e. multiple binding sites or the dominance of the marker system, is the conversion to SCAR markers, as described by Paran and Michelmore (1993). Since their introduction, SCAR markers have been used in several plant breeding studies (e.g. Corrêa et al. 2000, Zhang and Stommel 2001, Scheef et al. 2003, Rugienius et al. 2006). Although they are valuable markers, their creation is known to be laborious (e.g. Bradeen and Simon 1998, Shan et al. 1999). In fact, the loss of the original polymorphism, which occurred in our study, is not an unusual event when working with SCARs (e.g. Paran and Michelmore 1993, Scovel 1998), because the SCAR primers are usually nested within the original RAPD primers. Since the polymorphism of RAPDs likely results from alterations in the primer binding sites, the original polymorphism may be lost by using the SCAR primers. Usually, the chance for reproducing the original polymorphism is higher if RAPD fragments are cloned rather than sequenced directly (Hernández et al. 2003), since cloning leads to the RAPD- primer binding sites being included in the sequenced fragment. In our study, this was the case for the 5'-RAPD binding site of the A-marker. However, the polymorphism could not be

maintained even if this binding site was also used to amplify the SCAR marker. Nonetheless, the cloning assay for this RAPD fragment revealed a primer mismatch of one base in the binding site. Although modifying the annealing temperature during PCR is considered to be one of the most effective tools to eliminate amplification from one of the alleles as a result from mismatching, this assay did not work for our SCARs (data not shown).

In addition, regarding the B-marker, two different RAPD fragments have obviously been cloned. As described above, a fragment of little size difference to the proper marker band was amplified in 25 of the 28 'bud-flowering' individuals of population BC_1(i). Since the resolution power of agarose gels is limited, we cannot exclude the possibility that both bands are amplified in positive ('normal-flowering') individuals of the segregating population BC_1(i), which were then cloned and sequenced simultaneously. This would lead to mixed sequencing reactions and alignment problems as described above. In addition, RAPD fragments of the same molecular size do not necessarily have to share the same sequence. Although the gel resolution could be increased using PAGE, in our point of view this would be inapplicable for a small-sized business with limited lab equipment as is the case for *Calluna* breeders. We therefore consider the A-marker to be of a greater practical importance.

The detection of a new and screenable polymorphism within the given SCAR sequence is a prerequisite for the applicability of the MAS system in practical breeding processes. Two common approaches are widely used for this problem: SCAR-CAPS conversion and PCR-SSCP analysis. The development of CAPS marker is a considerable matter of expense, i.e. it requires either in-depth sequence data from cloning experiments for the pre-experimental analysis of potential restriction sites or a large set of available restriction enzymes for the 'random cutting approach'. For example, Fang et al. (1997) did not recover restriction site mutations by screening 20 different restriction enzymes.

We therefore applied SSCP analysis (Orita et a. 1989) to detect polymorphisms in monomorphic PCR products amplified by the SCAR primers. The SSCP technique, i.e. its capability to detect polymorphisms, is based on mobility differences of single stranded DNA because conformation variations resulting from different base compositions of these separated DNA fragments lead to differing migrating rates in the gel matrices used (usually mutation detection enhancement gels, Hauser et al. 1998). Until now, SSCP analysis has frequently been used in medical research, although only a few studies have been published in population genetics (Sunnucks et al. 2000). Sunnucks et al. (2000) reported the existence of two banding systems, a faster migrating, lower one and a slower migrating, upper one. Polymorphic bands may occur in both banding systems, although the upper one is usually multibanded. In fact, separation of our monomorphic SCAR markers reproducibly resulted in two polymorphic bands in the uppermost banding system between pools of 'normal' and 'bud-flowering' genotypes for one primer combination in the backcross population BC_1(i) (Fig. 5). The amount of individuals with a deviating genotype was higher when compared to the RAPD markers but decrease if primers are elongated from 20 bp to 27 bp. The reasons for multiple bands in PCR-SSCP analysis, especially in the slow-migrating banding system, are manifold and range from the expectable sequence differences of denatured single-strand DNA to double-stranded homoduplex DNA as a result of a lack of denaturation (Sunnucks et al. 2000). We therefore assume that these polymorphisms are coupled to the 'flower type' trait and are based on differing sequences of the single strands of the A-marker PCR fragments The practical applicability of this marker is restricted due to the PAGE participation.

From an economic point of view, two recently published surveys (Rout and Mohapatra 2006, Byrne 2007) of scientific publications and of the activities of ornamental breeding companies regarding the application of mo-

lecular marker techniques revealed that, besides SSRs and AFLPs markers, RAPDs remain one of the most commonly applied marker systems in ornamental crops. The study by Rout and Mohapatra (2006) revealed that only a small minority of breeding companies actually use MAS in their breeding process. The main applications are the identification of parental relationships as well as cultivar discrimination (Essentially Derived Varieties (EDVs)). Byrne´s questionnaire (2007) revealed similar results, i.e. that almost 40% of 79 fruit and ornamental breeders applied molecular markers for their breeding program, mostly using SSRs, AFLPs and RAPDs. Only 14% were engaged in MAS and only 3% of these were actually using this method to aid a selection process– the remaining 11% of the research programs were still at the developmental stage. Yet again, the main application of molecular markers in ornamental crops was shown to be the analysis of genetic diversity (44%).

MAS only brings about real progress in the breeding process if the trait of interest is difficult to identify – as is the case for resistances – or if it requires long cultivation periods of the progenies before selection is possible – as is the case for flower characteristics, especially in woody plants with long juvenile periods. The costs and efforts required to integrate an MAS system into an existing breeding system depend on the technique, and should not be underestimated. However, in 2001, Yousef and Juvik demonstrated for *Zea mays* that MAS can directly compete with phenotypic selection with regard to both the costs of the process and the amount of manual labor involved.

For this reason, our current research aims to identify a closer linked marker based on segregating populations of 100 – 1000 individuals using AFLPs. In so doing, we expect to detect tightly linked markers on both sides of the anticipated gene, which would form the basis for map-based cloning.

ACKNOWLEDGEMENT

This study was carried out jointly by the IGZ and the participating company Heidepflanzen Peter de Winkel (Goch) within the BMWi-funded project (KP0172401BN5A). The authors expressly thank the company owner for his support during the project, Katja Krueger and Anke Mueller (IGZ) for their excellent technical assistance in the lab as well as Monika Spiller, Anja Biber and Marcus Linde (Leibniz-University of Hannover) for cloning and sequencing, and their supplementary academic support.

REFERENCES

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997), **"Gapped BLAST and PSI-BLAST: a new generation of protein database search programs"**, Nucleic Acids Res. 25:3389-3402.

Bagley MJ, Anderson SL, May B (2001) **Choice of methodology for assessing genetic impacts of environmental stressors: polymorphism and reproducibility of RAPD and AFLP fingerprints.** Ecotoxicology 10:239-244

Borchert T, Krueger J, Hohe A (2008) **Implementation of a model for identifying Essentially Derived Varieties in vegetatively propagated *Calluna vulgaris* varieties.** BMC Genetics 9:56

Bradeen JM, Simon PW (1998) **Conversion of an AFLP fragment linked to the carrot *Y2* locus to a simple, codominant, PCR-based marker form.** Theor. Appl. Genet. 97:960-967.

Byrne DH (2007) **Molecular marker use in perennial plant breeding.** Acta Hort 751:163-167.

Corrêa RX, Costa MR, Good-God PI, Ragagnin VA, Faleiro FG, Moreira MA, de Barros EG (2000) **Sequence Characterized Amplified Regions linked to Rust Resistance Genes in the common bean.** Crop Sci 40:804-807

Fang DQ, Federici CT, Roose ML (1997) **Development of molecular markers linked to a gene controlling fruit acidity in citrus**. Genome 40:841-849

Freemont PS (1993) **The RING finger. A novel protein sequence motif related to the zinc finger.** Ann. N.Y. Acad. Sci. 684:174-192

Garcia AAF, Benchimol LL, Barbosa AMM, Geraldi IO, Souza Jr. CL, de Souza AP (2004) **Comparison of RADP, RFLP, AFLP and SSR markers for diversity studies in tropical maize inbred lines.** Genetics and Molecular Biology 27(4):579-588

Hauser MT; Adhami F, Dorner M, Fuchs E, Glossl J (1998) **Generation of co-dominant PCR-based markers by duplex analysis on high resolution gels**. Plant J 16:117-125.

Heckenberger M, van der Voort JR, Melchinger A, Peleman J, Bohn M (2003) **Variation of DNA fingerprints among accessions within maize inbred lines and implications for identification of essentially derived varieties: II. Genetic and technical sources of variation in AFLP data and comparison with SSR data.** Mol. Breed. 12:97-106

Hernández P, Dorado G, Ramírez MC, Laurie DA, Snape JW, Martín A (2003) **Development of cost-effective *Hordeum chilense* DNA markers: molecular aids for marker-assisted cereal breeding.** Hereditas 138:54-58

Ishiguro S, Kawai-Oda A, Ueda J, Nishida I, Okada K (2001) **The *Defective in Anther Dehiscence1* gene encodes a novel Phospholipase A1 catalyzing the initial step of jasmonic acid synthesis, which synchronizes pollen maturation, anther dehiscence and flower opening in Arabidopsis.** The Plant Cell 13:2191-2209.

Jones CJ, Edwards KJ, Castaglione S, Winfield MO, Sala F, Vandewiel C et al. (1997) **Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories.** Mol. Breed. 3:381-390

Kobayashi N, Horikoshi T, Katsuyama H, Handa T, Takayanagi K (1998) **A simple and efficient DNA extraction method for plants, especially woody plants.** Plant Tiss Cult Biotech 4:76-80.

Michelmore RW, Paran I, Kesseli RV (1991) **Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions using segregating populations.** Proc. Natl. Acad. Sci. USA 88:9828-9832

Orita M, Iwahana H, Kanazawa H, Hayashi K, Sekiya T (1989) **Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphism.** Proc. Natl. Acad. Sci. 86:2766-2770.

Paran I, Michelmore RW (1993) **Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce.** Theor. Appl. Genet. 85:985-993.

Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S, Rafalski A (1996) **The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis**. Molecular Breeding 2:225-238

Rout GR, Mohapatra A (2006) **Use of molecular markers in ornamental plants: a critical reappraisal.** Europ. J. Hort Sci. 71(2):53-68.

Rugienius R, Siksnianas T, Stanys V, Gelvonauskiene D, Bendokas V (2006) **Use of RAPD and SCAR markers for identification of strawberry genotypes carrying red stele (*Phytophtora fragariae*) resistance gene *Rpf1*.** Agronomy Research 4:335-339

Scheef EA, Casler MD, Jung G (2003) **Development of species-specific SCAR markers in bentgrass. Crop Sci.** 43:345-349.

Scovel G, Ben-Meir H, Ovadis M, Itzhaki H, Vainstein A (1998) **RAPD and RFLP markers tightly linked to the locus controlling carnation (Dianthus caryophyllus) flower type.** Theoretical and Applied Genetics 96(1): 117-122.

Shan X, Blake TK, Talbert LE (1999) **Conversion of AFLP markers to sequence-specific PCR markers in barley and wheat.** Theor. Appl. Genet. 98:1072-1078.

Stothard P (2000) **The Sequence Manipulation Suite: JavaScript programs for analyzing and formatting protein and DNA sequences.** Biotechniques 28:1102-1104.

Sunnucks P, Wilson ACC, Beheregaray LB, Zenger K, French J, Taylor AC (2000) **SSCP is not so difficult: the application and utility of single-stranded conformation polymorphism in evolutionary biology and molecular ecology.** Molecular Ecology 9:1699-1710.

Untergasser A, Nijveen H, Rao X, Bisseling T, Geurts R, Leunissen JAM (2007): **Primer3Plus, an enhanced web interface to Primer3** Nucleic Acids Research 35:W71-W74

Welsh J, McClelland M (1990) **Fingerprinting genomes using PCR with arbitrary primers.** Nucleic Acid Research 18:7213-7218.

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

Wolfe AD, Xiang QY, Kephart SR (1998) **Assessing hybridizsation in natural populations of *Penstemon* (Scrophulariaceae) using hypervariable intersimple sequence repeats (ISSR) bands.** Molecular Ecology 7:1107-1125.

Yousef GG, Juvik JA (2001) **Comparison of phenotypic and marker-assisted selection for quantitative traits in sweet corn.** Crop Sci. 41:645-655

Zhang Y, Stommel JR (2001) **Development of SCAR and CAPS markers linked to the *Beta* gene in tomato.** Crop Sci. 41:1602-1608.

Zietkiewicz E, Rafalski A, Labuda D (1994) **Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification.** Genomics 20:176-183.

ZMP (2007) **Zentrale Markt- und Preisberichtsstelle für Erzeugnisse der Land-, Forst- und Ernährungswirtschaft.**

Table 1: Segregation analysis of the 'bud flowering' trait.

The absolute amount of 'normal' and 'bud-flowering' genotypes from four back-crossing populations, segregation ratios with regard to the trait of interest and the corresponding X^2-values (H_0: segregation ratios do not differ from 1:1; $X^2_{0.95, 1} = 3.84$). Parent varieties used for backcrossing are listed under the subheading ´plant material´. The anticipated monogenic-recessive inheritance of the 'bud-flowering' trait was not disproved in three of the four of the populations ($\alpha = 0.05$).

Population denomination	Mother Variety	Father Variety	'Bud-flowering' individuals	Sum	Segregation ratio	X^2
BC_1(i)	´Melanie´	´Roter Oktober´	28	68	1,43	2,117
BC_1(ii)	´Maria´	´Boskoop´	57	119	1,09	0,210
BC_1(iii)	´Anette´	´Roter Oktober´	16	46	1,86	4,260
BC_1(iv)	´Mariella´	´Boskoop´	45	101	1,24	1,198

Table 2: Amplification of RAPD A- and B-markers in independent varieties / genotypes (n = 24 'bud-flowering' varieties, n = 25 'normal-flowering' varieties).

A: Overview. A positive result for 'normal-flowering' varieties implies the amplification of the marker fragment. On the contrary, a positive result for 'bud-flowering' varieties means that the marker fragment is not amplified by PCR.

Marker	Flower type	Positive result
A	normal	44.0%
	bud	83.3%
B	normal	56.0%
	bud	62.5%

B: List is tested varieties / genotypes.

flower type	denomination	flower type	denomination
'wild-type'	´Allegro´	'bud-flowering'	´Adrie´
	´Battle of Arnhem´		´Alicia´
	´Boskoop´		´Amethyst´
	´C. W. Nix´		´Anette´
	´Carmen´		´Annegret´
	´Dart´s Beauty´		´Anneliese´
	´Dart´s Hedgehog´		´Aphrodite´
	´Dart´s Silver Rocket´		´Barbara´
	´Finale´		´Cilcennin Common´
	´Glenmorangie´		´David Eason´
	´Hiemalis´		´Fritz Kircher´
	´Long White´		´Ginkel´s Glorie´
	´Manuel´		´Johnson´s Variety´
	´McDonalds of Glencoe´		´Karla´
	´Multicolor´		´Klaudine´
	´Orange Queen´		´Lianne´
	´Roter Oktober´		´Mariella´
	´Silver Knight´		´Marleen´
	´Westphalia´		´Marlies´
	´White Mite´		´Melanie´
	Löhnstein		´Minka´
	Niederohe		´Minna Kircher´
	Ruhla		´Sandy´
	Tiefenthal		´Senta´
	San Remo		

Figure 1: Shoots of *C. vulgaris*.

Left: ‘Normal-flowering’ genotype (wild type), right: ‘bud-flowering’ genotype (´Annegret´)

Figure 2: RAPD A-marker amplified in a backcross population.

Separation of RAPD products (1.5% agarose gel, left: 10 'normal-flowering' genotypes, right: 10 'bud-flowering' genotypes). The marker fragment (~ 2.5 kb) is clearly visible as the uppermost band. The 20 individuals belong to population BC_1(i). The first 'bud-flowering' genotype is the individual with marker amplification (recombination). The three most intensive bands in the size marker lanes resemble 0.5, 1.0 and 3.0 kb, respectively.

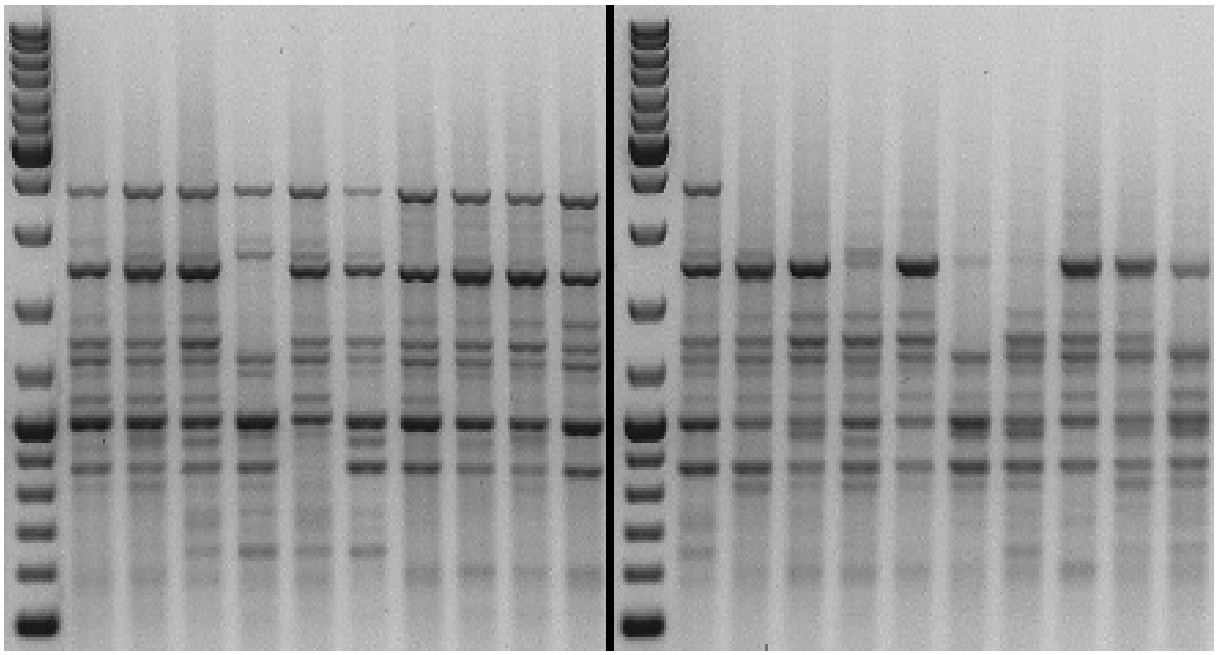


Figure 3: RAPD markers 'A' (lanes 2-5) and 'B' (lanes 6-9).

One positive and one negative genotype (lanes 2, 3 and 6, 7) as well as the purified fragment after gel-extraction (lanes 4 and 8) are shown for each marker. Marker fragments are highlighted by arrows. A faint band beneath the marker band is recognizable in lane 7, making it difficult to analyse data on agarose gels. Lanes 5 and 9: H_2O; lanes 1 and 10: size marker. The three most intensive bands in the size marker lanes resemble 0.5, 1.0 and 3.0 kb, respectively.

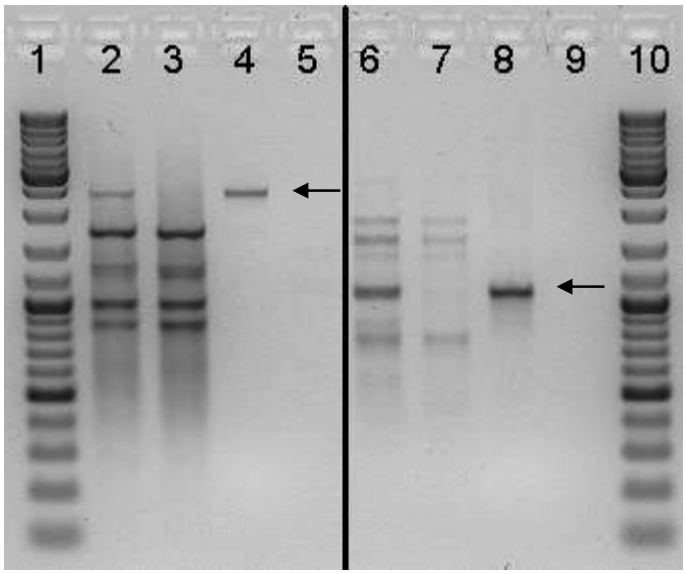


Figure 4: SSCP analysis (neutral PAGE, silver staining) of SCAR fragments (A-marker)

Direct separation of PCR products (20 bp SCAR primers, left: 3 ‘normal-flowering’, right: 3 ‘bud-flowering’ individuals, population: BC_1(i)). The white arrow indicates the polymorphic fragment for ‘normal-flowering’ genotypes. Corresponding fragments between both pools (same molecular size) are linked with dotted lines for better interpretation.

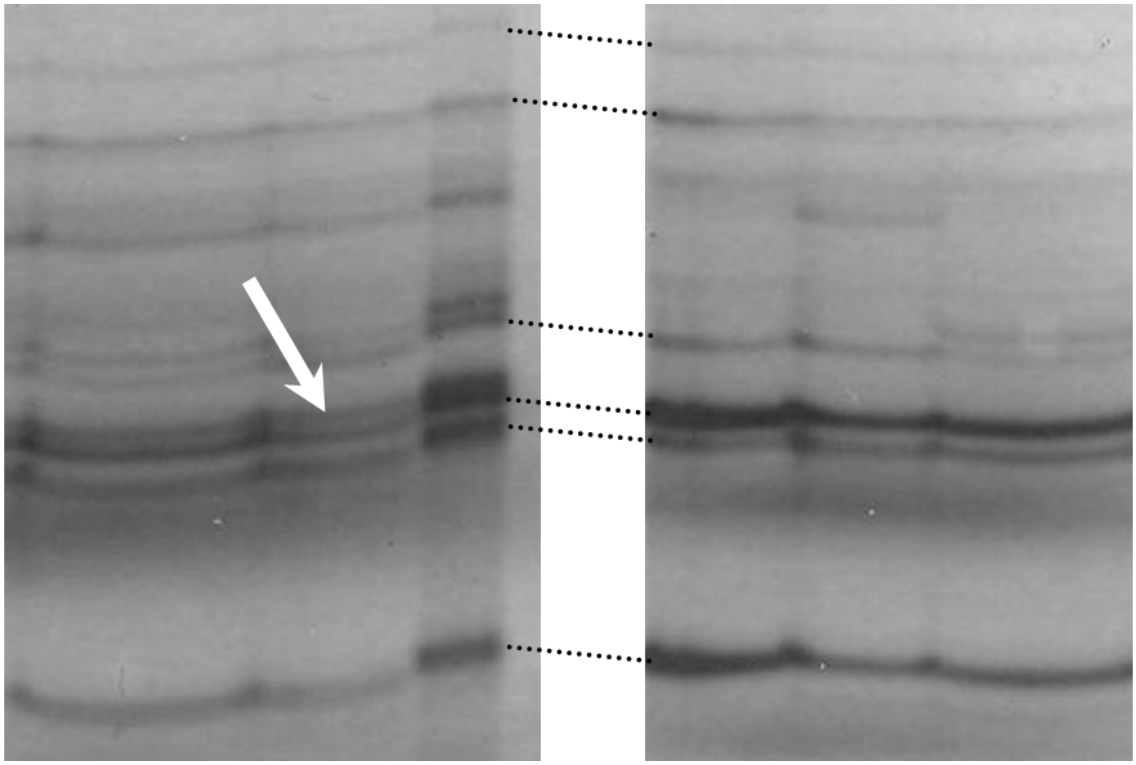
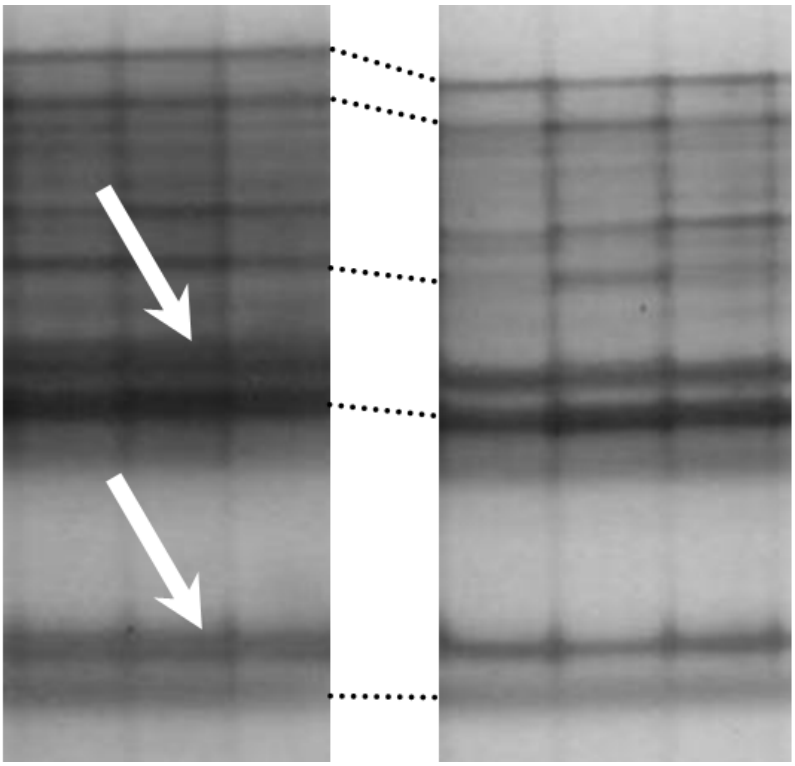


Figure 5: SSCP analysis (neutral PAGE, silver staining) of SCAR fragments (A-marker)

Direct separation of PCR products (27 bp SCAR primers, left: 3 ‘normal-flowering’, right: 3 ‘bud-flowering’ individuals, population: BC_1(i)). The upper white arrow indicates the maintained polymorphic fragment (compare Fig. 4). The lower white arrow indicates the additional fragment for ‘normal-flowering’ genotypes resulting from primer elongation. Corresponding fragments between both pools (same molecular size) are linked with dotted lines for better interpretation.



2.2.1. Annex I: Additional segregation data

In addition to the given data in Table 1 of Chapter 2.2, further segregating populations were generated in 2005 and 2006. In Table 3, the segregation ratios of other selected populations are presented. By ignoring the critical factor of population size, all populations with a 1st year flowering rate of 0.5 or above are shown. In addition, populations are shown that were produced by crossing F_1 individuals to randomly chosen and potentially homozygous recessive genotypes in order to test the hypothesis of monogenic-recessive inheritance of the 'bud-flowering' trait independent of a specific crossing combination.

Table 3: Segregation analysis of the 'bud-flowering' trait in BC_1 populations. The absolute amount of 'bud-flowering' and 'wild-type' progenies is shown for six different populations. Segregation ratios and χ^2 values are presented (H_0: segregation ratios do not differ from 1:1; $\chi^2_{0:95;\,1}$= 3.84).

crossing	parent varieties		BC_1 individuals		segregation	
#	bf	wt	bf	wt	ratio	χ^2
1	'Aphrodite'	Maria x Boskoop	12	20	0,60	2,00
2	'Lianne'	Maria x Boskoop	6	13	0,46	2,58
3	'Maria'	Maria x Boskoop	1	33	0,03	30,12
4	'Maria'	Maria x Boskoop	4	7	0,57	0,82
5	'Maria'	Maria x Boskoop	0	6	0,00	6,00
6	'Venetia'	Maria x Boskoop	5	4	1,25	0,11

The anticipated monogenic-recessive inheritance was not disproved in four (#1, #2, #4, #6) of these six populations (α = 0.05), these included 'false' back-crossing populations. In population #3, a very unusual low amount of 'bud-flowering' individuals were observed, although the same genotypes were used as parent individuals as e.g. in population BC_1(ii) (see Table 1 of Chapter 2.2.). Furthermore, population #5 consists of only six individuals in total which were derived from 150 seeds (germination rate 6.67%, 1st year flowering rate 0.6) and therefore, such a population is in principle unsuitable for analysis of its segregation ratio.

These diverse segregation ratios may indicate additional factors that influence the segregation ratio of back-crossing populations.

‘Who’s who’ in different flower types of *Calluna vulgaris* (*Ericaceae*): morphological and molecular analyses of flower organ identity

Thomas Borchert[1*], Katrin Eckardt[2], Jörg Fuchs[3], Katja Krüger[1*], Annette Hohe[1*§]

1: Institute of Vegetable and Ornamental Crops (IGZ)
Department Plant Propagation
Kuehnhaeuser Str. 101
99189 Erfurt
Germany
2: University of Applied Sciences Dresden
Faculty for Agriculture and Landscape Management
Pillnitzer Platz 2
01326 Dresden
Germany
3: Leibniz-Institute of Plant Genetics and Crop Plant Research (IPK)
Department of Cytogenetics and Genome Analysis
Corrensstrasse 3
06466 Gatersleben
Germany

*: these authors contributed equally to this work
§: Corresponding author

TB: borchert@erfurt.igzev.de
KE: katrin.eckardt@gmx.net
JF: fuchs@ipk-gatersleben.de
KK: k.krueger@erfurt.igzev.de
AH: hohe@erfurt.igzev.de

KEYWORDS: MADS transcription factors, Genome size, Ericales, Loss of stamen, Bracts

ABSTRACT

Background

The ornamental crop *C. vulgaris* is of increasing importance to the horticultural industry in the northern hemisphere due to a quite simple flower organ mutation: the flowers of the 'bud-flowering' phenotype remain closed i.e. as buds throughout the total flowering period and thereby maintain more colorful flowers for a longer period of time than the 'wild-type'. This feature is accompanied and presumably caused by the complete lack of stamens. Descriptions of this botanical particularity are inconsistent and partially conflicting. In order to clarify basic questions of flower organ identity in general and stamen loss in detail, a comparative study of the diverse flower types of *C. vulgaris* was initiated.

Results

Flowers were examined by macro- and microscopic techniques. Organ development was investigated comparatively in both the 'wild-type' and the 'bud-flowering' type by histological analyses. Analysis of cell surface structures of the perianth epidermis using scanning electron microscopy revealed that in 'wild-type' flowers the outer whorls of colored leaves can be identified as sepals. The uppermost whorls of green leaves represent bracts. In the 'bud-flowering' type, two whorls of sepals are directly followed by the gynoecium, which is why both petals and stamens are completely missing.

In addition, two MADS-box genes (homologs of *AP3/DEF* and *SEP1*) were identified in *C. vulgaris* using RACE-PCR. Expression analysis by qRT-PCR was conducted for both genes in vegetative tissues, bracts, sepals and petals. These experiments revealed a gradual decrease of B-gene expression levels from petals to bracts and vegetative tissue, which supports our previous organ classification based on morphological characteristics.

Conclusions

Organ identity in both 'wild-type' and 'bud-flowering' *C. vulgaris* was clarified using a combination of microscopic and molecular methods. Bract, sepal and petal organ identity is supported by a modified 'ABCDE model'. However, parallel loss of stamens in the 'bud-flowering' phenotype is an exceptional flower organ modification that cannot be explained by modified spatial expression of known organ identity genes.

BACKGROUND

Calluna vulgaris L. (Hull.) (Fig. 1A) belongs to the order *Ericales*, which comprises 25 families including 346 genera with more than 11,500 species in total [1]. The *Ericales* incorporate about 5.9% of core eudicot diversity, one third of which is made up of the *Ericaceae* alone [1]. This apparent diversity particularly affects the flower morphology and organization since the only shared feature of all *Ericales* is the occurrence of tenuincellate ovules, which however is common to all asterids [2]. In addition, a persistent corolla is the only synapomorphy for *C. vulgaris* and the *Ericeae* (except *Daboecia*) [3].

The economic significance of *C. vulgaris* to the horticultural industry in Europe and North-America is continually increasing [4]. The current market share in Germany for instance, amounts to approximately 120 million EUR, or > 100 million plants per year, respectively [4]. In principal, this economic significance is the results of a single but considerable change in the flower morphology: the loss of stamens. In contrast to 'wild-type' flowers (Fig. 1A) that are only attractive from August to October the resulting 'bud-flowering' phenotype (Fig. 1B) preserves its unpollinated stigmas within the never-opening buds and has an extended flowering period up to December. For this reason, it is the most valuable flower type of this species to the horticultural business. In contrast, other forms, such as the 'filled' (Fig. 1CD) or the 'multi-bracteate' types (Fig. 1E) are less important. Previous investigations revealed the monogenic recessive inheritance of the 'bud-flowering' trait [5].

The synoecious flower of 'wild-type' *C. vulgaris* is of radial symmetry, posing with two outer whorls with four to five colored leaves in each whorl as the perianth, two whorls of four to five stamens and four to five coadunate carpels [6, 7, 8]. The sepals are grouped into two distinct whorls of 2x 2 sepals by [9]. The colored leaves of the whorl II are fused at the receptacle to form a corolla tube [6, 10]. Two surrounding whorls of at least six green leaves were described by [11]. These uppermost whorls of green leaves do not match the perianth symmetry, since they are aligned with the sepal whorl instead with the petal whorl (Fig. 1F).

In contrast, the 'bud-flowering' type completely lacks the male reproductive organs, which is probably at least one of the reasons for its developmental arrest in the bud stage (Fig. 1B). In 1986, [12] described three different subforms of the 'bud-flowering' type: *f. diplocalyx* ('[…] eight instead of four sepals and usually neither stamens nor corolla […]', page 281), *f. polysepala* ('[…] similar to *f. diplocalyx* but […] there are indeed many sepals, more than eight.', page 281) and *f. clistanthes* ('[…] flower parts are present in the normal number, but the corolla never, or hardly, opens.', page 281). Evidences or justifications for this classification of organs e.g. of the colored leaves either as sepals or petals are missing in [6] as well as in [12].

Moreover, no explanation is given by [9] for the grouping of the sepals into two whorls and by [6] for the grouping of stamen in two whorls.

According to the 'ABCDE'-model of flower organ identity, changes in flower morphology are the results of expression shifts of different classes of floral homeotic genes encoding transcription factors in the corresponding whorls (see, e.g. [13, 14 or 15]): class A gene function in the outmost whorl leads to the formation of sepals; combined expression of class A and B genes in the second whorl leads to the formation of petals; class B and C gene function in whorl three promotes the development of stamens, and expression of class C genes in the innermost whorl leads to the development of carpels. Additionally, class D gene function is required for ovule formation, whereas class E gene function is required for the development of all organs, respectively (see. e.g. [16, 17, 18, 19]). Two different approaches are commonly applied to identify organ characteristics in the perianth of angiosperms: morphological comparisons and gene expression studies (see, e.g. [20]). The molecular procedure mainly investigates the expression of the floral homeotic genes. On the other hand, several studies demonstrated that the perianth organs can be distinguished by the assessment of their epidermal cell surface structure by scanning electron microscopy (SEM), see, e.g. [16] in *Arabidopsis thaliana* or [21] in *Ericales* (*Impatiens*, *Marcgravia*). Both assays – the morphological and the molecular assay – have to be regarded as complementary [20].

Regarding the vague descriptions and the lack of current in-depth studies and molecular data in *C. vulgaris*, several uncertainties still exist on the topic of the flower organ identity in this species. On the one hand, questions arise regarding the discrete identity of the two outer whorls of colored leaves. On the other hand, the lack of the androecium in the 'bud-flowering' type has not been ascertained either. Until now, it is even uncertain, whether stamen development is been initiated or whether the initiation of primordia is already inhibited.

The determination of the flower organ identity and the understanding of the development of the 'bud-flowering' mutation itself are of importance for future breeding efforts in *C. vulgaris* since the 'bud-flowering' phenotype is the most important breeding target in this species. We therefore initiated histological, microscopic and molecular examinations to clarify the identity of flower organs and the differences between 'wild-type' and 'bud-flowering' phenotypes.

RESULTS

In order to elucidate the unknown organ identities of different flower phenotypes in the ornamental crop *C. vulgaris* flower development was monitored histologically. In addition, the

perianth organs were examined and made successfully distinguishable by SEM. In order to achieve a better understanding of mutations in flower morphology in this crop, an initial cloning of two MADS-box genes was realized in addition to preliminary expression analyses. The genome size was determined in order to evaluate the chances of future cloning of new unknown genes by chromosome walking.

Morphological perianth organ analysis of the 'wild-type' and the 'bud-flowering' type

In 'wild-type' phenotypes, the whorl II organs are commonly fused at their base and are more delicate compared to whorl I organs, which are clearly separated and appear quite robust. In contrast, 'bud-flowering' organs of whorl II are not fused and resemble the whorl I organs in shape, color and stability. SEM of whorl I and whorl II organs in both the 'wild-type' and the 'bud-flowering' phenotype was carried out to identify, whether whorl I organs in the 'wild-type' can be identified as sepals or petals and in order to clarify the identity of the whorl II organs in the 'bud-flowering' phenotype. The abaxial and adaxial epidermis structures (n = 3 varieties each) were compared for both outer whorls (Fig. 2). Cells of the outermost whorls of the 'wild-type' phenotype are flat and stretched (Fig. 2AB). In contrast, cells of the second whorl appear bloated ('dome-shaped'), are shorter in diameter and length and are striated with papillate structures (Fig. 2CD). On the contrary, the cell surfaces of the 'bud-flowering' perianth organs are indistinguishable from each other, since both whorls consist of the flat and stretched cell type (Fig. 2E-H), comparable to the outmost whorl of the 'wild-type'. In particular, the second whorl leaves are not 'dome-shaped'. Thus, concerning whorl I organs of the 'wild-type' phenotype, both their position and their cell surface structure indicate a sepaloid identity, whereas their color suggests a petaloid identity. Regarding whorl II organs, all these three criteria investigated may be a hint to petaloid identity. In contrast, all organs in both perianth whorls of the 'bud-flowering' phenotypes are morphologically not distinguishable and show the same characteristics as whorl I organs of the 'wild-type' phenotype. Therefore, they are likewise presumably to be identified as sepals, although again their coloring suggests a petaloid identity.

As a consequence of the identification of whorl I organs as sepals the uppermost green leaves (Fig. 1F) can be classified either as foliage leaves or as bracts. This is further supported by the occurrence of stomata in these organs which we did not observe in any colored perianth organ. However, differentiation between bracts and vegetative leaves by morphological characteristics only appears vague.

Cloning of MADS-box genes

MADS-box transcription factors were identified using RACE. Our initial 3'-RACE experiments resulted in the cloning of two gene fragments, one putative *AP3/DEF*-like gene we named *CvAP3* [GenBank:GQ202026], and one *SEP1/2*-like gene we named *CvSEP1* [GenBank:GQ202027]. For *CvAP3*, the sequence data resulted from three independent experimental PCR and cloning assays. *CvSEP1* was cloned by chance since the primer was designed to amplify B-genes. Thus, *CvSEP1* could not be verified independently until now. Both partial genes were obtained by cloning a fragment of the anticipated size of approximately 950 bp. Using the BLAST conserved domain database [22], the K-box and the (partial) MADS-box were identified in *CvAP3*, and the K-box in *CvSEP1*. Furthermore, both the EuAP3 motif and the PI derived motif [23] were identified within *CvAP3*, whereas the *CvSEP1* gene included the SEP I and SEP II motif, as described by [24]. The latter motif, also termed as *AGL2/SEP1* terminal motif [25], may be used to discriminate *SEP1/2* (the *LOFSEP* clade) and *SEP3* genes: *SEP3* genes are missing this motif, but instead, they contain either the *AGL9/SEP3* or the *ZmM7* motif [25]. As additional validation of gene identification, Table 1 includes relevant BLASTx hits [26] for *CvSEP1*. These hits furthermore confirmed the *SEP1/2* identity, since (i) these hits are retrievable in [25] as belonging to the *LOFSEP* clade and (ii) E-values for any *SEP3* hit were clearly lower.

Molecular perianth organ analysis of the 'wild-type' and the 'bud-flowering' type

Results of relative expression analyses of the *C. vulgaris AP3/DEF*- and *SEP1/2*-like genes are presented in Fig. 3. ΔΔCt-values have been calculated to compare expression levels between the different flower tissues and the vegetative tissue of the corresponding flower type, since the expression of both genes was lowest (albeit present, compared to the normalizer) in the vegetative tissue ($\Delta Ct_{CvAP3,\text{'bud-flowering'}} = 12.12$; $\Delta Ct_{CvAP3,\text{'wild-type'}} = 12.96$, $\Delta Ct_{CvSEP1,\text{'bud-flowering'}} = 14.79$; $\Delta Ct_{CvSEP1,\text{'wild-type'}} = 14.72$). The expression levels of *CvAP3* (Fig. 3A) gradually and constantly decreased from whorl II over whorl I to the uppermost green leaves in both flowering types. While in whorl I organs of both flowering types, the expression was comparable, whorl II organs differed significantly with 'wild-type' petals showing an almost double fold expression level in comparison to the 'bud-flowering' type. For *CvSEP1* (Fig. 3B), expression was highest in both perianth organs of both flower types, while in the uppermost green leaves the expression level was significantly lower. Interestingly, expression was significantly higher in the uppermost green leaves of the 'bud-flowering' phenotype in com-

parison to the corresponding 'wild-type' organs. However, expression of *CvSEP1* did not differ markedly between whorl I and whorl II organs.

Due to the clear differences in expression of both genes between vegetative tissue and the uppermost green leaves, the latter are to be identified as bracts.

For both target genes and the 18S rRNA normalizer, the unusual expression in vegetative tissue was confirmed in foliage samples of different age of three independant genotypes. However, cloning and sequencing of PCR products both from 'wild-type' and 'bud-flowering' samples confirmed the identity of the amplified transcripts.

Floral formula of different flower types

Since we were not able to decide whether organs of the same identity were arranged in one or several whorls, we uniformly speak of one whorl per organ type, except for flower types with changes in organ identity. Thus, the floral formulas presented consecutively are based on the described morphological results (e.g. cell surface structure) and not on positional information of the organs.

In contrast to the 'wild-type' ($Ca^{4}Co^{(4)}A^{8}G^{(4)}$, Fig. 4A; Ca: calyx; Co: corolla; A: androecium; G; gynoecium), the 'bud-flowering' phenotype completely lacks stamens whereas its petals are transformed to sepals ($Ca^{4+4}Co^{0}A^{0}G^{(4)}$, Fig. 4B). This corresponds to the 'diplocalyx' type described by McClintock (1986). In contrast, the 'bud-flowering' variety 'David Eason' develops carpel-like structures within whorl III which at least partially are adnate to the carpels of whorl IV ($Ca^{4+4}Co^{0}A^{0}G^{(8+4)}$, Fig. 4C). 'Filled' flowers ('Radnor', Fig. 1D) present only indistinguishable colored leaves of varying amount in each of their whorls ($CaCo^{\infty}A^{0}G^{0}$).

Flower organ development

Flower organ development of the 'wild-type' and the 'bud-flowering' type were investigated comparatively by histological analysis. Samples were derived from the uppermost part of shoots for which the initiation of flower development could undoubtedly be reported. Fig. 5 shows three equal stages of both the 'wild-type' and the 'bud-flowering' type in parallel.

Generative meristems of both flower types did not differ anatomically (Fig. 5AD). The formation of the petals in the 'wild-type' flower is the first clear morphological difference to be detected (Fig. 5BE). When the carpels are clearly recognizable as such (Fig 5CF), the comparison of 'wild-type' and 'bud-flowering' types reveals there is no indication of stamen formation in the latter type. We observed no evidence for a rudimentary development or a subsequent reduction of stamens.

Interestingly, tissue of petal and stamen bases in the 'wild-type' (Fig. 5C) display the same staining pattern. The same is true for the two whorls of sepals in the 'bud-flowering' type (Fig. 5F). This becomes even more obvious in opened, mature flowers of each type, using SGL instead of FCA staining (Fig. 6AB).

Estimation of the genome size

The genome size of *C. vulgaris* was estimated by laser-based flowcytometry since the knowledge of this parameter is essential for future genetic applications. We compared seven 'wild-type', two 'bud-flowering', one 'filled' and one 'multi-bracteate' genotype from different countries (Table 2). Three to six replications of each sample led to an overall average genome size of 1.1799 +/- 0.0028 pg/2C (mean +/- standard error, n = 50). According to the equation given by [27], from this the total DNA length of *C. vulgaris* can be calculated to be approximately 1,154 Mbp.

DISCUSSION

The vague and differing descriptions of the flower anatomy ([6, 7, 9, 12]) of *C. vulgaris* necessitated more in-depth investigations regarding the flower organ identity. We combined different microscopic and molecular analyses, since both approaches are complementary (see, e.g. [20] and references therein). In so doing, we appointed the organ numbers for *C. vulgaris* inflorescences as summed up in the floral formulas. 'Wild-type' flowers are complete synoecious flowers, since both gynoecium and androecium are present. In addition, sepals and petals are clearly distinguishable. The borderline between perianth and bracts was drawn *inter alia* as a result of stomatal occurrence: corolla and calyx of *C. vulgaris* do not show stomata as the bracts do as was demonstrated earlier ([11, 28]). This hypothesis on the borderline between perianth and bracts is further supported by the phenotype of 'filled' flowers. 'Filled' flowers presumably originate from a total loss of class C gene expression (e.g. [29]) in whorls III and IV, which is substituted by additional expression of class A genes, resulting in flowers consisting only of sepals and petals. Since in *C. vulgaris*, 'filled' flowers consist of colored leaves only, this phenotype supports the argumentation of drawing the flower borderline between the green and the colored leaves, since otherwise, green sepals would be noticeable in 'filled' flowers. Regarding the 'bud-flowering' type our anatomical analyses revealed that second whorl organs are macroscopically indistinguishable from the first whorl organs.

These morphological and anatomical data were supported by our gene expression analyses. We detected expression of *CvAP3* in both perianth organs of both the 'wild-type' and the 'bud-flowering' type in *C. vulgaris*. Thus, in contrast to the classical 'ABCDE'-model, significant expression of a B-class gene was detected in the sepals. This resembles the 'shifting/sliding boundary' model known e.g. from *Tulipa* [30] and other monocots. However, this model also fits to basal eudicots, e.g. *Ranunculales* or *Aquilegia* ([31] and references therein). For this reason, an outward shift of the B-function border could be a reason for the petaloid appearance of sepals in 'wild-type' flowers of *C. vulgaris*. Nevertheless, expression of *CvAP3* was significantly higher in petals of the 'wild-type' than in its sepals. In contrast, no difference of expression was observed between these organs in the 'bud-flowering' type, which confirmed our morphological (SEM) data suggesting an additional whorl of petaloid sepals and the coincidental loss of petals in this flowering type. Thus, differential expression of *CvAP3* consistently reflects changes and similarities in the morphology of whorl I and whorl II flower organs in the 'wild-type' as well as in the 'bud-flowering' type. However, these differences are of a quantitative and not of a qualitative nature. Thus, a resemblance with the 'fading border' model as proposed by [32] for *Amborella trichopoda* (*Amborellaceae*) is subject to conjecture. This hypothesis may be supported by the morphological resemblance (i.e. staining and cell structure) of 'bud-flowering' petal and stamen bases; again, the resemblance of the cell structure of stamens and petals was already shown for *A. trichopoda* [33]. However, the more variable expression patterns of MADS-box genes described by the 'fading border' model is usually restricted to basal angiosperms, e.g. *Nymphaeaceae*, *Illicium* [32] or *Eupomatia* [34]. So far, fading borders are unknown in eudicots species. However, also [21] showed a gradual decrease of *DEF*-like gene expression from stamens to petals and sepals in *Impatiens hawkeri* (*Ericales*).

Regarding the identification of bracts, our expression analyses also confirmed the morphological argumentation. On the one hand, the combined expression of *CvAP3* and *CvSEP1* in these organs indicates a clear difference to vegetative leaves, especially for *CvAP3*. On the other hand, besides the clear morphological dissimilarity, expression of *CvSEP1* was obviously repressed in these uppermost green leaves compared to the sepals. Therefore, we identified these leaves as bracts. Again, this result is in line with [21], who also showed that a *DEF*-like gene is quantitatively lower expressed in bracteoles/sepals of *Marcgravia umbellata* (*Ericales*) than in petals or stamens of the same species.

In the 'multi-bracteate' type of *C. vulgaris*, an indeterminate increase of green (and not colored) bracteate leaves can be observed, a phenomenon also known from *Antirrhinum*. Here,

low expression levels of the MADS-box *squamosa* gene were identified to participate in the induction of so-called 'bracteomania', a phenotype equivalent to the *C. vulgaris* 'multi-bracteate' type due to its excessive development of bracts [35]. Thus, a homolog of *squamosa* might be a candidate gene for molecular analysis of the 'multi-bracteate' flower type in *C. vulgaris*.

Expression of *CvAP3* and *CvSEP1* was detected in vegetative leaves independent of tissue age for both flowering genes. According to the classical 'ABCDE model' and its modifications, we anticipated expression of the *AP3/DEF*-like gene to be restricted to whorls II and whorls III [13]. In contrast, expression of *SEP1/2*-like genes was expectable for whorls II-IV, although 'expression in sepals is common but not universal' ([25], page 431). However, expression of floral organ identity genes in non-floral tissues is already known from other species. In *Gerbera*, the *SEP1/2* gene *GhGRCD2* is expressed in vegetative tissues and *SEP3*, usually restricted to the inner three whorls, is described to be expressed in vegetative tissues in more than one species, too ([25]and references therein). Likewise, in Rose expression of the *AP3*-like gene *MASAKO euB3* was detected in vegetative leaves [36].

Regarding our results, it has to be borne in mind that, according to the floral quartet model, floral organ identity genes concertedly regulate the organ identity [37]. Our analyses therefore necessarily remain incomplete. Nevertheless, we were able to differentiate all organs in question by expression analyses of just two putative MADS-Box transcription factors.

The 'wild-type' flower of *C. vulgaris* is synoecious, while the 'bud-flowering' type is unisexually female due to a total loss of stamens. Although the androecium is often modified during plant evolution, a complete loss of stamen whorls or the alteration of stamen function towards secondary floral functions e.g. the production of nectar are usually absent in *Ericaceae* [38]. For this reason, no parallel analyses in akin species are available. According to [39], every unisexual flower that has been investigated until then showed a certain degree of initial hermaphroditic characteristics. However, only four years later, stamen primordia have not been identified in *Thalictrum diocium* [40]. Thus, 'bud-flowering' *C. vulgaris* rather resembles *T. diocium*, since we did not observe an abortion of initiated stamen in the 'bud-flowering' type. Instead of a downstream regulation (which would result in a belatedly abortion of once initiated stamens) these authors therefore suggest an upstream regulation of class B MADS-box genes regarding the loss of stamens in *T. diocium*. An upstream regulation of *AP3* expression may involve multiple genes e.g. *LEAFY* and *AP1* [41]. However, putative regulation does not

either provide a hypothesis on the reason for the simultaneous transformation of petals into petaloid sepals in 'bud-flowering' phenotypes that accompanies stamen loss in *C. vulgaris*. Moreover, investigating a deviating 'bud-flowering' type − the cultivar 'David Eason' − gives an additional indication of the genetics of this flower morphology. Here, flower anatomy resembles a typical B gene deficiency mutant, since the stamen seem to be transformed to additional carpels partially coadunated to the original whorl IV carpels. As in the 'diplocalyx' type, the B gene function is presumably extended to whorl I, since both outmost whorls consist of eight colored sepals. Thus, the transformation of petals into petaloid sepals, that is true for both 'bud-flowering' types, is not necessarily linked to one or the other mode of loss of stamens. Hence, genotypes with flowers that only display the transformation of petals into petaloid sepals without a loss of stamens might also exist. However, we have not yet been able to identify any genotype that presents this change in perianth organ identity without simultaneous the loss of stamen

Within the *Ericaceae*, the genome size is only known in seven *Vaccinium* species [42]. Here, the nuclear DNA content ranged from 1.20 - 7.20 pg/2C. Knowledge of the genome size is an essential prerequisite for prospective genomic applications in this species including mapping and genome walking for isolation of putative genes responsible for the 'bud-flowering' genotype. Although the measured value of 1.18 pg/2C is low, it is still approx. four times higher than in *Arabidopsis* (0.3 pg/2C, [43]. Nevertheless, it facilitates the construction of a BAC (Bacterial Artificial Chromosome) library and subsequent chromosome walking.

CONCLUSIONS

Our study presents a first step on the analyses of flower organ identity in *C. vulgaris*. We confirmed the identity of petals, sepals and bracts in 'wild-type' as well as in the 'bud-flowering' type. From these results we deduced hypotheses on modifications of the 'ABCDE'-model in *C. vulgaris*. Further investigations should include additional cloning (full length clones) of floral organ identity gene homologs as well as studies of gene expression by qRT-PCR in all floral organs of different flower types.

The loss of stamen in the standard 'bud-flowering' type cannot be explained by modifications of the 'ABCDE'-model. Neither can clear candidate genes be deduced from comparison with other plant species. However, our analysis showed once again that too strict homology concepts are often not applicable in diversifying plants groups.

Since a comprehensive understanding of the genetics of the 'bud-flowering' phenotype is a prerequisite for future breeding of this economically important ornamental crop, mapping of this trait with subsequent chromosome walking will be the next step to identify candidate genes, since the relatively small genome size of *C. vulgaris* allows efficient construction of a BAC library.

METHODS

Histological Techniques and Microscopy

Tissues were fixed in AFE (10.4 : 1 :1 96% ethanol : formalin : acetic acid) or in Bouin-Allen's compound (14 : 5: 1 picronitric acid : formol : acetic acid + 1.48% (w/v) CrO_3), dehydrated by ethanol, infiltrated and embedded in paraffin under low air pressure conditions, and sectioned at varying µm-intervals using a Leica RM2155 microtome. The sections were stained with either FCA (fuchsin CI42520, chryosidine CI11270, astral blue CI48048) or SGL (safranine CI 50240, pyoctanin blue CI 42535, acid green CI 42095) and photographed by a Zeiss Axio Imager.A1. The macroscopical analysis of the flower morphology was performed using a Leica Wild MZ3 stereo microscope.

Scanning Electron Microscopy

Samples were fixed over night in FAEG (ethanol (65%), acetic acid (5%), 37% formaldehyde (3.2%), 50% glutaraldehyde (0.2%), Tween-20 (0.1%), H_2O) and dehydrated by an ethanol series: 15 min 80% ethanol, 15 min 90% ethanol, 15 min 96% ethanol, 3x 20 min 100% ethanol. The samples were then transferred to 100% acetone (3x 20 min) and subsequently critical point dried using liquid CO_2 in an EMITECH K850. The leaves were mounted on Leit-Tabs and gold-coated (sputter-coater: EMITECH K500). Observations of the abaxial and adaxial sides of the perianth organs of each three genotypes were performed using a Philips XL30 ESEM (at the Institute of Systematic Zoology and Evolutionary Biology, University of Jena) with an voltage of 10 kV.

Cloning of MADS-box genes

Total RNA of 'wild-type' *C. vulgaris* 'Roter Oktober' flower buds was isolated using a modified protocol of the RNeasy Plant Mini Kit ([44], Qiagen) and subsequently reverse transcribed to first strand cDNA (Reverse Transcription System, Promega) using a standard oligo(dT) primer: GACTCGAGTCGACATCTG$(T)_{14}$. 3'-RACE-PCR [45] was performed

using a 5'-B-gene-MADS-box-specific primer and the corresponding 3'-nested primer derived from the oligo(dT) primer. Amplified fragments of appropriate size were gel-extracted (Nucleo Spin Extract II kit, Macherey-Nagel), ligated into the pDRIVE vector and transformed into EZ cells (Qiagen PCR Cloning plus kit) by heat-shock. Cells were plated on standard LB/Amp/IPTG/X-Gal plates. Plasmid DNA from positive clones (blue/white selection plus colony-PCR testing) was extracted (E.Z.N.A. Plasmid mini kit II, Omega bio-tek) and sequenced (MWG Biotech AG, JenaGen GmbH, AGOWA GmbH).
The alignments of derived sequences was accomplished by ClustalW2 [46] or T-Coffee [47]. BLASTx 2.2.19+ [26] and BLASTn 2.2.19+ [48] were used to check the *C. vulgaris* sequences for matching hits at the protein or nucleotide level. Cloned genes were named using the abbreviation of the species name and the gene class, respectively, and uploaded to the GenBank database via Sequin.

Expression analysis (qRT-PCR)

Total RNA of different tissues was isolated as described above, but the tissues were stored in liquid RNAlater (Qiagen) at 4°C until RNA extraction. cDNA was reverse transcribed using the original protocol of the QuantiTect Reverse Transcription Kit (Qiagen). qRT-PCR primers (Table 3) were designed to target the *AP3/DEF*- and *SEP1/2*-like genes using Primer3Plus [49]. The partial sequence of *C. vulgaris* 18S rRNA [GenBank:AF419797] was used to design normalizing primers. PCR reactions (3 independent runs with each 3 technical replicates of a 'bud-flowering' and a 'wild-type' genotype) were performed with 0.5 ng cDNA on a Stratagene MX3000P thermocycler (qPCR MxPro v4.01) using the Absolute qPCR SYBR green ROX mix (ABgene). Gene expression analysis was normalized vs. *C. vulgaris* 18S rRNA. ΔΔCt, i.e. the fold change was calculated according to $Ratio = 2^{-\Delta\Delta Ct}$ (see, e.g. [50]), whereas the mean ΔCt of the vegetative tissue was subtracted from the normalized ΔCt-values of bracts, sepals and petals, respectively. Prior to realtime PCR experiments, all primer combinations were tested for their optimum concentration and the prerequisite of PCR-product-free non-template controls. The PCR products were additionally and successfully verified for length (by electrophoretic separation) and for sequence (AGOWA GmbH) identity with the predicted amplicons from different tissues and genotypes.

Estimation of nuclear genome size by flow cytometry

Fresh young foliage from samples and internal reference standards (0.5 cm^2 each) were co-chopped with a sharp razor blade in a Petri dish containing 500 µL nuclei isolation buffer ac-

cording to [51], supplemented with 1% polyvinylepyrrolidone 25, 0.1% Triton X-100, 50 µg/ml RNAse and 50 µg/ml propidium iodide, incubated for at least 30 sec and filtered through a 35 µm mesh. The relative fluorescence intensities of stained nuclei were measured on a FACStar^{PLUS} (BD Biosciences, San Jose, CA, USA) equipped with an INNOVA 90-C argon laser (Coherent, Santa Clara, CA, USA). Propidium iodide was excited at 514 nm and measured in FL1 channel using a 630 nm band-pass filter. At least three plants of each *C. vulgaris* sample were used for absolute DNA content estimation together with *Glycine max* (L.) Merr. convar. max var. max ('Cina 5202', 2C = 2.23 pg; Genebank Gatersleben, accession number: SOJA 392) as an internal standard. The nuclear DNA amount of the standard was determined based on the value of 0.32 pg / 2C for *Arabidopsis thaliana* 'Columbia' [52]. Usually 10,000 nuclei per sample were analyzed. The absolute DNA amounts of the samples were calculated based on the values of the G1 peak means. ANOVA HSD Post-hoc test for unequal N, which is a modification of the Tukey HSD test, was used to determine significant differences between group means ($p = 0.05$).

LIST OF ABBREVIATIONS

A – Androecium

AP – APETALA

BAC – Bacterial Artificial Chromosome

Ca – Calyx

Co – Corolla

G – Gynoecium

MADS – mini-chromosome maintance1, Agamous, Deficiens, serum response factor

RACE – rapid amplification of cDNA ends

SEM – Scanning Electron Microscopy

SEP - SEPALLATA

AUTHOR'S CONTRIBUTIONS

TB carried out the molecular experiments, performed the sequence alignments and all other genetic data analysis, captured the macroscopic, histological and SEM images and drafted the manuscript. KE participated in the qRT-PCR experiments including method establishment. KK established and carried out the complete histological analyses. JF established and carried out the flowcytometric methods and experiments. AH participated in the experimental design and critically revised the manuscript. All authors read and approved this final manuscript version.

ACKNOWLEDGEMENTS

The authors specifically thank Dajana Lobbes and Pia Nutt, Rainer Melzer, Andrea Härter, Mariana Mondragón-Palomino and Günter Theissen (University of Jena) for their academic and practical support. Our special thanks are extended to Hans Pohl (University of Jena) for his assistance in SEM analysis. In addition, we would like to thank Claudia Hönemann, Anke Müller, Luisa Hiese and Jörg Krüger (IGZ) for their continued assistance.

Part of this work was conducted within a BMWi (German Federal Ministry of Economics and Technology) funded joint project (project number KP0172401BN5A) undertaken by the Leibniz Institute of Vegetable and Ornamental Crops, Erfurt, Germany and a German breeding company (Heidepflanzen Peter de Winkel, http://www.dewinkel.de). Furthermore, some results were obtained within the enterprise funded by the BMBF (German Federal Ministry of Education and Research, support code PGI-06.01-28-1-43.038-07).

REFERENCES

[1] Stevens PF (2001) **Angiosperm Phylogeny Website.** Version 8, June 2007 (and more or less continuously updated since) [http://www.mobot.org/MOBOT/research/apweb]
[2] Schönenberger J, Anderberg AA, Sytsma KJ: **Molecular phylogenetics and patterns of floral evolution in the *Ericales*.** Int. J. Plant. Sci. 2005, 166(2):265-288.
[3] Kron KA, Judd WS, Stevens PF, Crayn DM, Anderberg AA, Gadek PA, Quinn CJ, Luteyn JL: **Phylogenetic classification of *Ericaceae:* Molecular and Morphological Evidence.** The Botanical Review 2002, 68(3): 335-423.
[4] ZMP: **Zentrale Markt- und Preisberichtsstelle** [http://www.zmp.de]
[5] Borchert T, Hohe A: **Identification of molecular markers for the flower type in the ornamental crop *Calluna vulgaris*.** Euphytica, in press.
[6] Heß D: **Die Blüte. Einführung in Struktur und Funktion, Ökologie und Evolution der Blüten.** 2. Auflage, Stuttgart, Ulmer, 1990:42-43.
[7] Strasburger E, Noll F, Schenk H, Schimpe AFW: **Übersicht des Pflanzenreichs, Eukaryota, Angiospermae, Dicotyledonae**. In Lehrbuch der Botanik. 32. Auflage, Gustav Fischer Verlag Jena; 1983:878-879.
[8] Jacquemart AL: **Floral traits of Belgian *Ericaceae:* are they good indicators to assess the breeding system?** Belg. Journ. Bot. 2003, 136(2):154-164.
[9] Dixon GR, Dutton KJ: **Manipulation of flowering in white heather.** Achta Horticulturae 1987, 205:219-224.
[10] Stueber K (2007) [http://www.BioLib.de] Thomé OW: **Flora von Deutschland, Österreich und der Schweiz.** 1885. [http://caliban.mpiz-koeln.mpg.de/~stueber/thome/index.html]
[11] Stevens PF: ***Calluna, Cassiope* and *Harrinamella:* a taxonomic and evolutionary problem.** New Phytol. 1970, 69:1131-1148.
[12] McClintock: **Harmonising botanical and cultivar classification with special reference to hardy heathers.** Acta Hort. 1986, 182:277-283.
[13] Coen ES, Meyerowitz EM: **The war of the whorls: genetic interactions controlling flower development.** Nature 1991, 353:31-37.
[14] Theissen G: **Development of floral organ identity: stories from the MADS house.** Current Opinion in Plant Biology 2001, 4:75-85.
[15] Theissen G, Melzer R: **Molecular Mechanisms underlying origin and diversification of the angiosperm flower.** Annals of Botany 2007, 1-17
[16] Pelaz S, Ditta GS, Baumann E, Wisman E, Yanofsky MF: **B and C floral organ identity functions require SEPALLATA MADS-box genes.** Nature 2000, 405:200-203.
[17] Pelaz S, Gustafson-Brown C, Kohalmi SE, Crosby WL, Yanofsky MF: **APETALA1 and SEPALLATA3 interact to promote flower development**. Plant Journal 2001a, 26:385-394
[18] Pelaz S, Tapia-Lopez R, Alvarez-Buylla ER, Yanofsky MF: **Conversion of leaves into petals in *Arabidopsis.*** Current Biology 2001b, 11:182-184.
[19] Ditta G, Pinyopich A, Robles P, Pelaz S, Yanofsky MF: **The SEP4 gene of *Arabidopsis thaliana* functions in floral organ and meristem identity.** Current Biology 2004, 14:1935-1940.
[20] de Craene LPR: **Are petals sterile stamens or bracts? The origin and evolution of petals in core eudicots.** Annals of Botany 2007, 100:621-630.
[21] Geuten K, Becker A, Kaufmann K, Caris P, Janssens S, Viaene T, Theissen G, Smets E: **Petaloidy and petal identity in MADS-box genes in the balsaminoid genera *Impatiens* and *Marcgravia*.** The Plant Journal 2006, 47:501-518.
[22] Marchler-Bauer A, Anderson JB, Derbyshire MK, DeWeese-Scott C, Gonzales NR, Gwadz M, Hao LN, He SQ, Hurwitz DI, Jackson JD, Ke ZX, Krylov D, Lanczycki CJ, Liebert CA, Liu CL, Lu F, Lu SN, Marchler GH, Mullokandov M, Song JS, Thanki N, Yamashita RA, Yin JJ, Zhang DC, Bryant SH: **CDD: a conserved domain database for interactive domain family analysis.** Nucleic Acids Res. 2007, 35(D)237-240.

[23] Kramer EM, Dorit RL, Irish VF: **Molecular Evolution of Genes Controlling Petal and Stamen Development: Duplication and Divergence Within the APETALA3 and PISTILLATA MADS-Box Gene Lineages.** Genetics 1998, 149: 765-783.
[24] Zahn LM, Kong H, Leebens-Mack JH, Kim S, Soltis PS, Landherr LL, Soltis DE, dePamphilis CW, Ma H: **The Evolution of the SEPALLATA Subfamily of MADS-Box Genes: A Preangiosperm Origin With multiple Duplications Throughout Angiosperm History.** Genetics 2005, 169: 2209–2223.
[25] Malcomber ST, Kellog EA: **SEPALLATA gene diversification: brave new whorls.** Trends in Plant Science 2005, 10(9):427-435.
[26] Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: **"Gapped BLAST and PSI-BLAST: a new generation of protein database search programs"** Nucleic Acids Res. 1997, 25:3389-3402.
[27] Dolezel J, Bartos J, Voglmayr H, Greilhuber J: **Nuclear DNA content and genome size of trout and human.** Cytometry Part A 2003, 51A:127-128.
[28] Watson L: **The taxonomic significance of certain anatomical variations among *Ericaceae*.** J. Linn. Soc. (Bot.) 1964, 59(377):111-125.
[29] Yanofsky MF, Ma H, Bowman JL, Drews GN, Feldmann KA, Meyerowitz EM: **The protein encoded by the *Arabidopsis* homeotic gene AGAMOUS resembles transcription factors.** Nature 1990, 346:35-39.
[30] Kanno A, Saeki H, Kameya T, Saedler H, Theissen G: **Heterotopic expression of class B floral homeotic genes supports a modified ABC model for tulip.** Plant Molecular Biology 2003, 52:831-841.
[31] Soltis DE, Chanderbali AS, Kim S, Buzgo M, Soltis PS: **The ABC model and its applicability to basal angiosperms.** Annals of Botany 2005, 100:155-163.
[32] Kim S, Koh J, Ma H, Hu Y, Endress PK, Hauser BA, Buzgo M, Soltis PS, Soltis DE: **Sequence and expression studies of A-, B-, and E-class MADS-box homologues in *Eupomatia* (*Eupomatiaceae*): support for the bracteate origin of the calyptra.** International Journal of Plant Science 2005b, 166(2):185-198.
[33] Buzgo M, Soltis DE, Soltis PS: **Floral developmental morphology of *Amborella trichopoda* (*Amborellaceae*).** International Journal of Plant Sciences 2004, 165:925-947.
[34] Kim S, Koh J, Yoo MJ, Kong H, Hu Y, Ma H, Soltis PS, Soltis DE: **Expression of floral MADS-box genes in basal angiosperms: implications for the evolution of floral regulators.** Plant Journal 2005a, 43:724-744.
[35] Hujiser P, Klein J, Lönnig WE, Meijer H, Saedler H, Sommer H: **Bracteomania, an inflorescence anomaly, is caused by the loss of function of the MADS-box gene *squamosa* in *Antirrhinum majus*.** The EMBO Journal 1992, 11(4):1239-1249.
[36] Foucher F, Chevalier M, Corre C, Soufflet-Freslon V, Legeai F, Hibrand-Saint Oyant L: **New resources for studying the rose flowering process.** Genome 2008, 51:827-837
[37] Theissen G, Saedler H: **Floral Quartets.** Nature 2001, 409:469-471.
[38] Walker-Larsen J, Harder LD: **The evolution of staminodes in angiosperms: patterns of stamen reduction, loss, and functional re-invention.** American Journal of Botany 2000, 87(10):1367-1384.
[39] Smyth DR: **Flower development.** Current Biology 2001, 10(3):82-84.
[40] di Stilio V, Kramer EM, Baum DA: **Floral MADS box genes and homeotic gender dimorphism in *Thalictrum dioicum* (*Ranunculaceae*) – a new model for the study of dioecy.** The Plant Journal 2005, 41:755-766.
[41] Lamb RS, Hill TA, Tan QKG, Irish VF: **Regulation of APETALA3 floral homeotic gene expression by meristem identity genes.** Development 2002, 129:2079-2086.
[42] Costich DE, Ortiz R, Meagher TR, Bruederle LP, Vorsa N: **Determination of ploidy level and nuclear DNA content in blueberry by flow cytometry.** Theoretical and Applied Genetics 1993, 86:1001-1006.

[43] Bennett MD, Leitch IJ: **Angiosperm DNA C-values Database** (release 6.0, Oct. 2005) [http://www.kew.org/cval/homepage.html]
[44] Dhanaraj AL, Slovin JP, Rowland LJ: **Analysis of gene expression associated with cold acclimatisation in blueberry floral buds using expressed sequence tags.** Plant Science 2004, 166: 863-872.
[45] Frohman MA: **Rapid amplification of complementary DNA ends for generation of full-length complementary DNAs: thermal RACE.** Methods in Enzymology 1993, 218:340-356.
[46] Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG: **Clustal W and clustal X version 2.0.** Bioinformatics 2007, 23(21):2947-2948.
[47] Notredame C, Higgins DG, Heringa J: **T-Coffee: A novel method for multiple sequence alignments.** J. Mol Biol. 2000, 302:205-217.
[48] Zhang Z, Schwartz S, Wagner L, Miller W: **A greedy algorithm for aligning DNA sequences.** J Comput Biol 2000, 7(1-2):203-14.
[49] Untergasser A, Nijveen H, Rao X, Bisseling T, Geurts R, Leunissen JAM: **Primer3Plus, an enhanced web interface to Primer3.** Nucleic Acids Res 2007, 35:W71–W74.
[50] Yuan JS, Reed A, Chen F, Stewart Jr CN: **Statistical Analysis of real-time PCR data.** BMC Bioinformatics 2006, 7:85.
[51] Galbraith DW, Harkins KR, Maddox JM, Ayres NM, Sharma DP, Firoozabady E: **Rapid flow cytometric analysis of the cell cycle in intact plant tissues.** Science 1983, 220:1049-1051.
[52] Bennett MD, Leitch IJ, Price HJ and Johnston JS: **Comparisons with *Caenorhabditis* (~100 Mb) and *Drosophila* (~ 175 Mb) using flow cytometry show genome size in *Arabidopsis* to be ~157 Mb and thus ~25 % larger than the *Arabidopsis* Genome Initiative estimate of ~125 Mb.** Annals of Botany 2003, 91: 547-557.

Table 1: Selection of BLASTx hits of *CvSEP1*.

Gene name	Species	Accession number	E-value
FBP5	*Petunia x hybrida*	AAK21248	1^{-55}
VvMADS2	*Vitis vinifera*	AAM21342	2^{-53}
AmDEFH49	*Antirrhinum majus*	CAA64741	1^{-50}
MdMADS9	*Malus x domestica*	CAA04920	1^{-47}
AtSEP1	*Arabidopsis thaliana*	NP_568322	6^{-46}
AtSEP2	*Arabidopsis thaliana*	NP_186880	1^{-43}

Table 2: Flow cytometric estimation of the absolute DNA content of *C. vulgaris*.

The table indicates the flower type, the denomination of the genotype or variety, the country of origin (if known) and the amount of measured replicates n. Genotypes in italics are samples collected in the wild.

Flower type	Denomination	Origin	pg DNA / 2C	n
'Wild-type'	*Löhnstein*	Germany	1.16 +/- 0.006	4
	Niederohe	Germany	1.17 +/- 0.008	4
	San Remo	Italy	1.20 +/- 0.006	5
	Kvam	Norway	1.19 +/- 0.015	4
	'Long White'	The Netherlands	1.18 +/- 0.011	6
	'Multicolor'	USA	1.18 +/- 0.016	5
	'Silver Knight'	UK	1.18 +/- 0.011	5
'Bud-flowering'	'Karla'	Germany	1.17+/- 0.014	5
	'Sandhammeren'	Sweden	1.20 +/- 0.018	3
'Filled'	'Radnor'	UK	1.20 +/- 0.018	4
'Multi-bracteate'	'Perestroijka'	Germany	1.15 +/- 0.015	5

Table 3: qRT-PCR primers. Primer sequences were designed to amplify products < 200 bp.

Target sequence	Primer sequence	Product size [bp]
18S rRNA [GenBank:AF419797]	Forward: GGGATGAGCGGATGTTACTT Reverse: CCCTTCCGTCAATTCCTTTA	116
CvAP3 [GQ202026]	Forward: TCGACGAGCTGAATAGTCTTGA Reverse: TCGACTAGCCCATAGTGTGGAT	190
CvSEP1 [GQ202027]	forward: AGCATCATCCTCAATCCCAG Reverse: GATCATTCCGCTCACGTTTT	143

Fig. 1: Flower types of *C. vulgaris*.

A: ‘wild-type’;

B: ‘bud-flowering’ (‘Amethyst’);

C: ‘filled’ (‘Radnor’);

D: ‘filled’ (‘Radnor’), cross-section;

E: ‘multi-bracteate’ (‘Peace’);

F: bottom of ‘wild-type’ inflorescence;

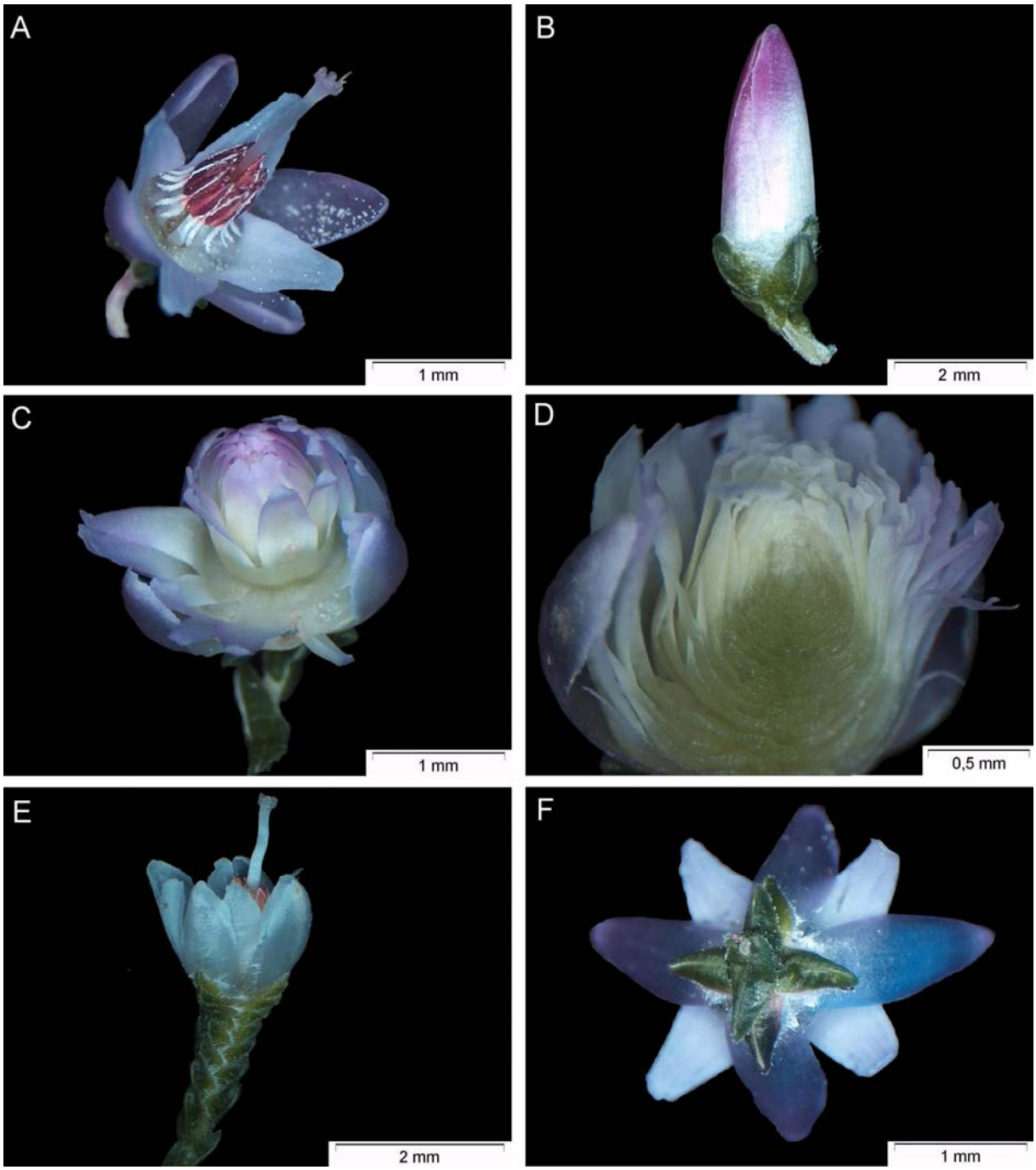


Fig. 2: Comparative SEM observations of abaxial and adaxial epidermal surface structures of *C. vulgaris* perianth organs.

'wild-type' whorl I, abaxial (A), adaxial (B);

'wild-type' whorl II, abaxial (C), adaxial (D);

'bud-flowering' whorl I, abaxial (E), adaxial (F);

'bud-flowering' whorl II, abaxial (G), adaxial (H);

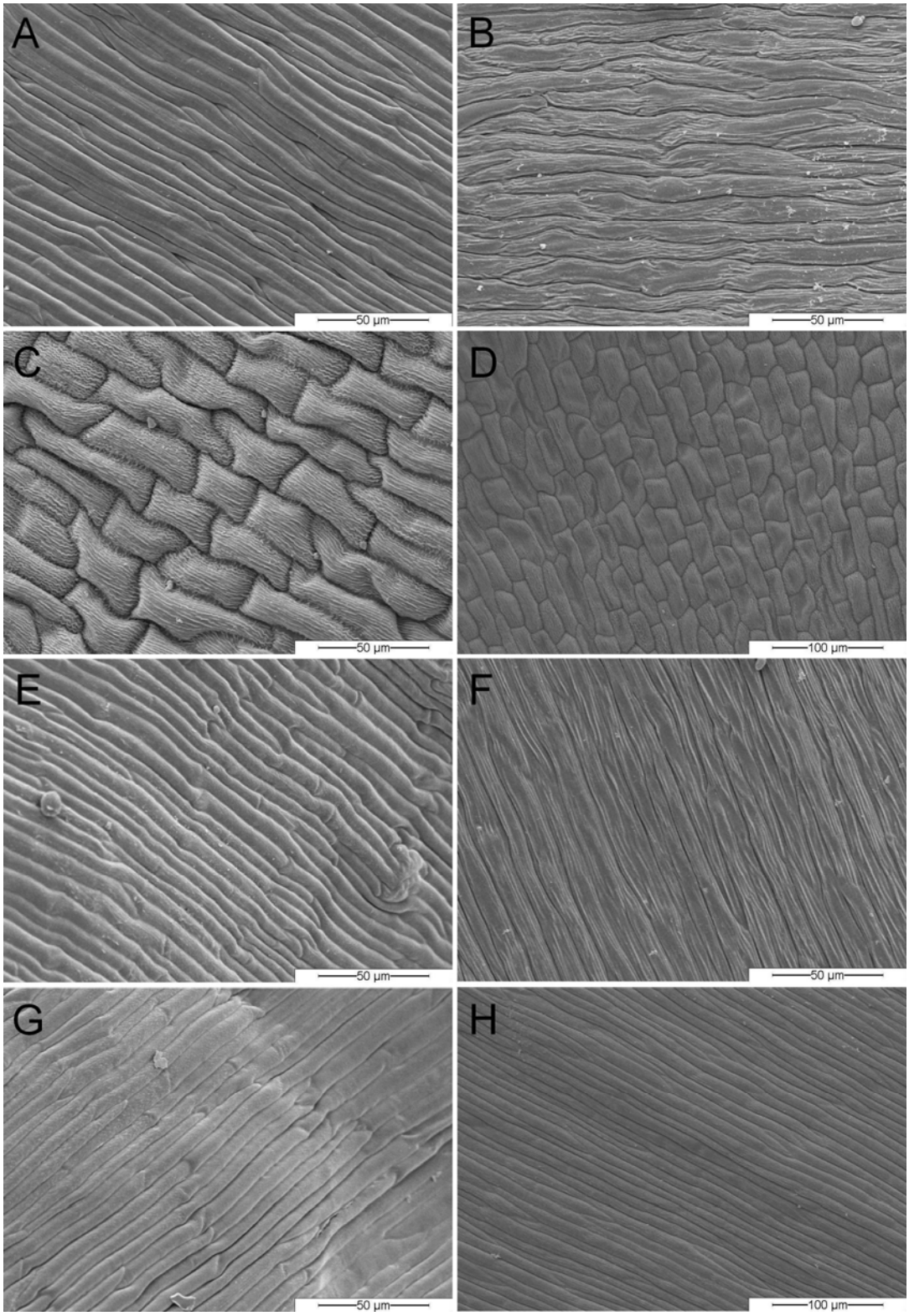


Fig. 3: Expression analysis of *CvAP3* (A) and *CvSEP1* (B) in *C. vulgaris* flower tissues. Normalized (vs. 18S rRNA) expression is presented for both the 'wild-type' and the 'bud-flowering' type as fold change (ΔΔCt) of arbitrary units vs. the reference tissue (vegetative foliage).

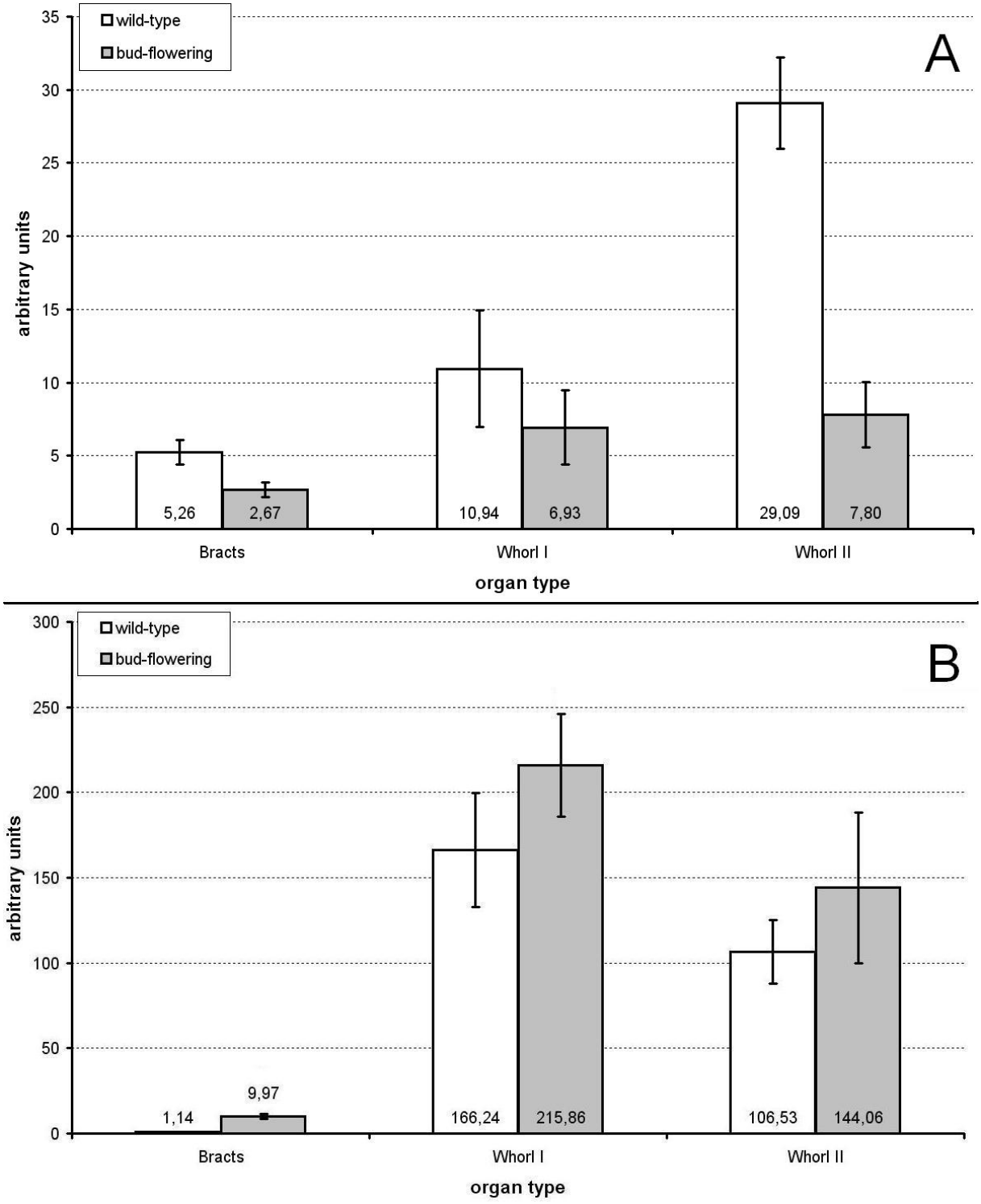


Fig. 4: Sagittal slices of mature flower buds.

A: ‘wild-type’ phenotype (A): $Ca^{4}Co^{(4)}A^{8}G^{(4)}$;

B: ‘bud-flowering’ phenotype (B, ´Anneliese´): $Ca^{4+4}Co^{0}A^{0}G^{(4)}$;

C: ‘bud-flowering’ phenotype (‘David Eason’): $Ca^{4+4}Co^{0}A^{0}G^{(8+4)}$;

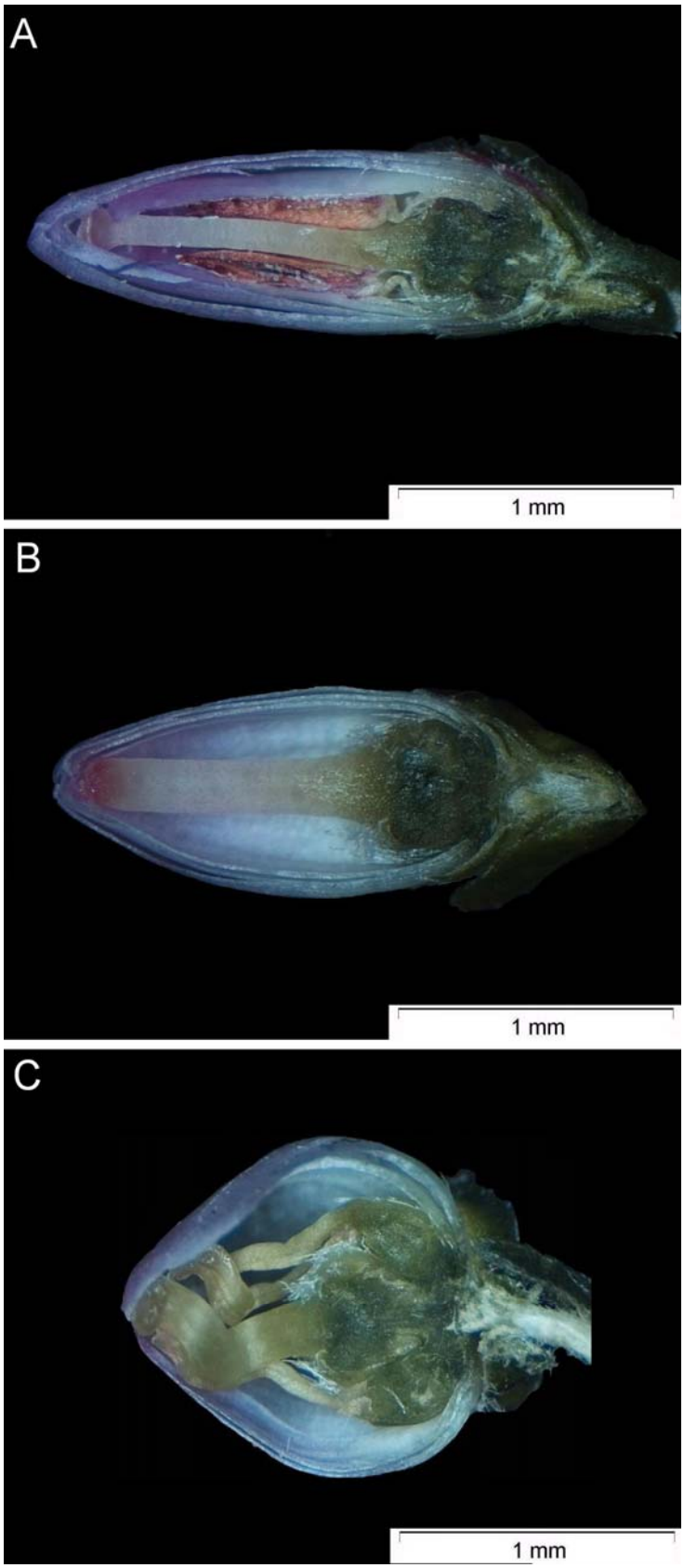


Fig. 5: Comparative investigation of *C. vulgaris* flower development.

Histological slices of 5 µm intervals were fixed in AFE and stained by FCA. Organs and tissues are labelled by veg (vegetative tissue), br (bracts), sep I or sep II (sepals, whorl no.), pet (petals, if available), sta (stamens, if available), ne (nectaroids), car (carpels), ov (ovules) and me (flower meristem), respectively.

A-C: different stages of a 'wild-type' inflorescence;

D-E: different stages of a 'bud-flowering' inflorescence;

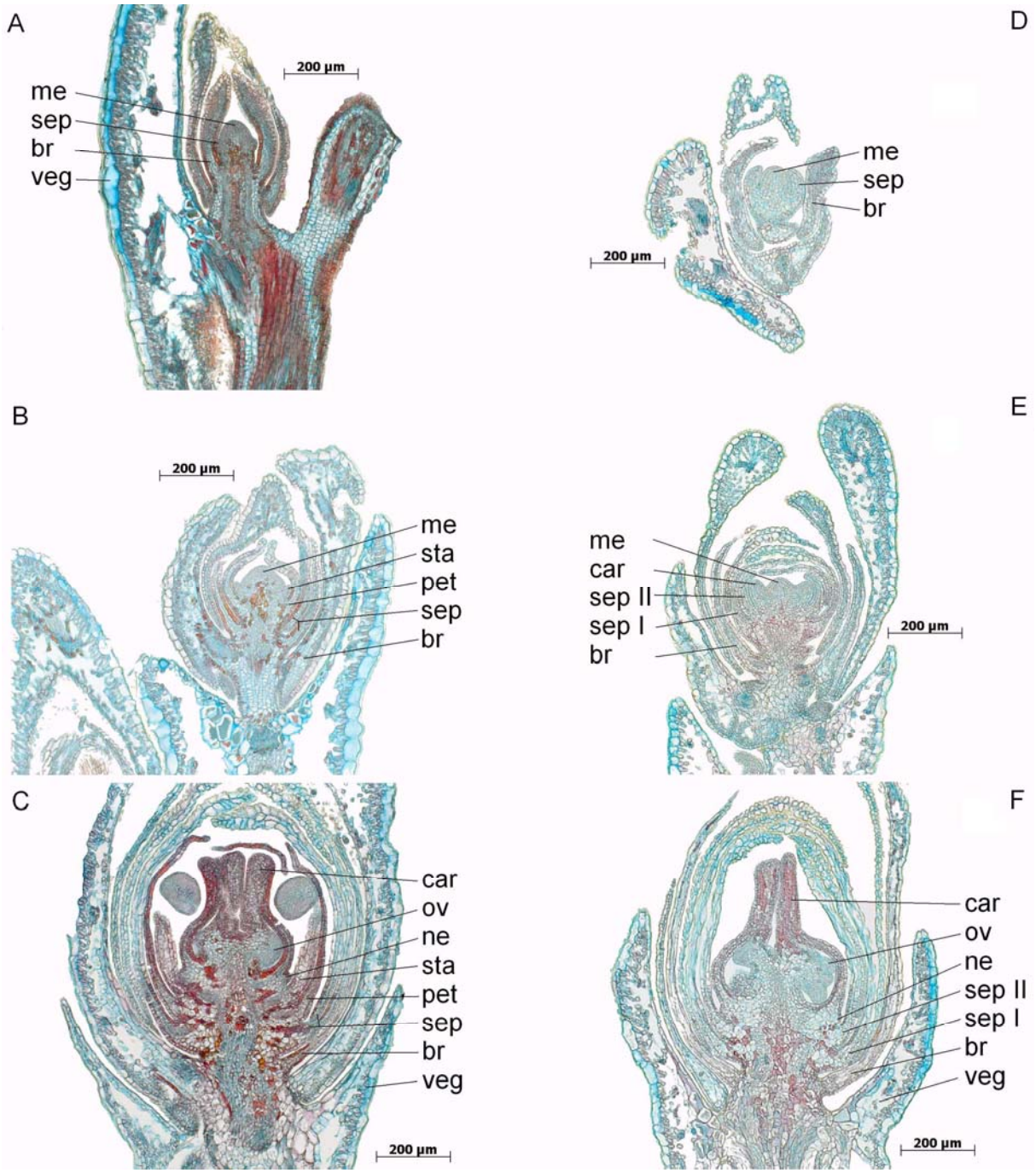
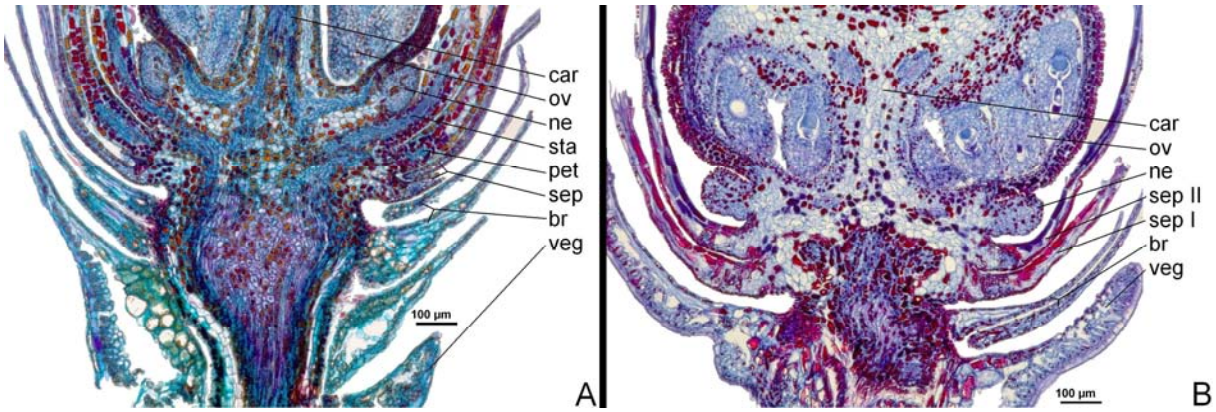


Fig. 6: Mature flowers of *C. vulgaris*.

Histological slices of 8 µm intervals were fixed in Bouin-Allen´s compound and stained by SGL. Organs and tissues are labelled by veg (vegetative tissue), br (bracts), sep I or sep II (sepals, whorl no.), pet (petals, if available), sta (stamens, if available), ne (nectaroids), car (carpels), ov (ovules) and me (flower meristem), respectively.

A: ‘wild-type’;

B: ‘bud-flowering’;



2.3.1. Annex I: Sequence Motifs

Homeotic MADS-box genes contain, beneath the familiar MADS-box, several more conserved domains and motifs that are not as highly conserved as the MADS-box but nevertheless, are common to genes of the same transcription factor class. Thus, these motifs can be used to identify a gene i.e. its partial gene sequence and to confirm its membership to a specific transcription factor class.

Kramer et al. (1998) described, among others, the EuAP3 and the PI-derived (PISTILLATA) motif for class B homeotic genes, Zahn et al. (2005) described the SEP I and the SEP II motif as already mentioned in the manuscript of Chapter 2.3.

The EuAP3 Motif as shown in Figure 8 is a synapomorphy for the euAP3 lineage of AP3 genes and is usually used for phylogenetic classification of this gene class (Kramer et al. 1998). The motif derived from the PI motif of PI genes is usually detected in AP3 genes and therefore is an additional evidence of the paralogy of AP3 und PI lineages (Kramer et al. 1998).

Regarding the SEP subfamily, their C-termini are usually less conserved than the equivalent regions of DEF/GLO (DEFICIENS / GLOBOSA) and AG (AGAMOUS) MADS-box subfamilies. In addition, SEP I and SEP II do structurally not resemble any motif with known function (Zahn et al. 2005). However, they are located in the same region as the PI-derived and other AP3-motifs in AP3 proteins (Kramer et al. 1998, Zahn et al. 2005) and therefore, they are deemed to be functionally and equivalently important for protein function.

These motifs are highlighted within the ClustalW2 protein alignments of the *AP3/DEF*- and *SEP1*/*2*-like genes, respectively (Figure 8 and Figure 9, next page). The partial *C. vulgaris* sequences of both the B- and E-like MADS-box genes each included the mentioned motifs. The coding sequence of *CvAP3* included the K-box and the partial MADS-box, the shorter *CvSEP1* sequence included the K-box, which is located downstream of the MADS-box (data not shown).

The identified *C. vulgaris* sequences for *CvAP3* [GQ202026] and *CvSEP1* [GQ202027] were used for phylogram construction (Chapter 2.3.2.).

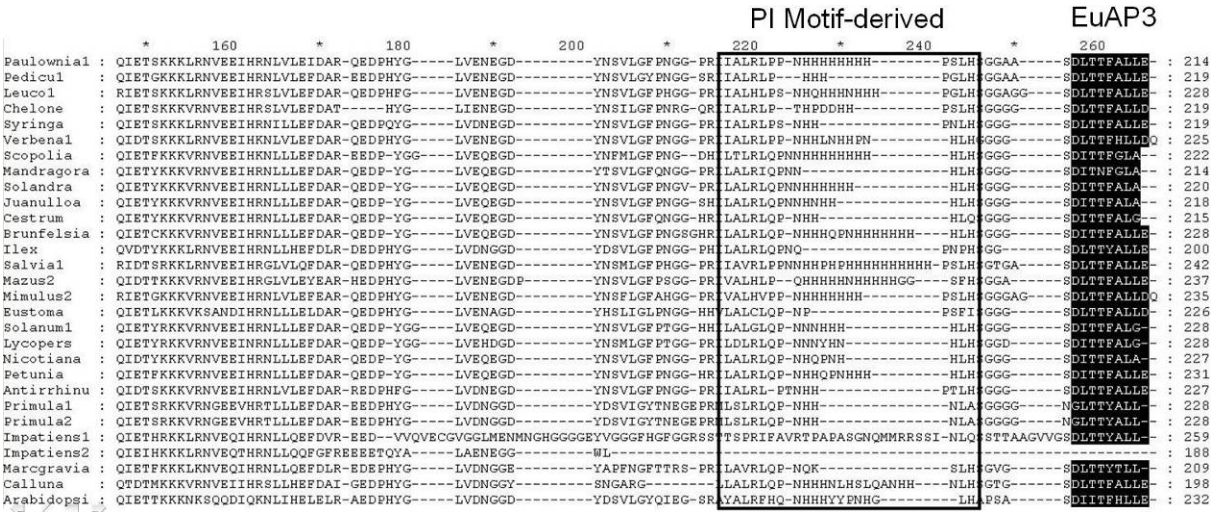


Figure 8: Alignment of the C-terminal region of *C. vulgaris AP3* and its homologs, used for tree construction. The regions designated as motifs are inverted or boxed and labeled as PI Motif-Derived and EuAP3, respectively (according to Kramer et al. 1998).

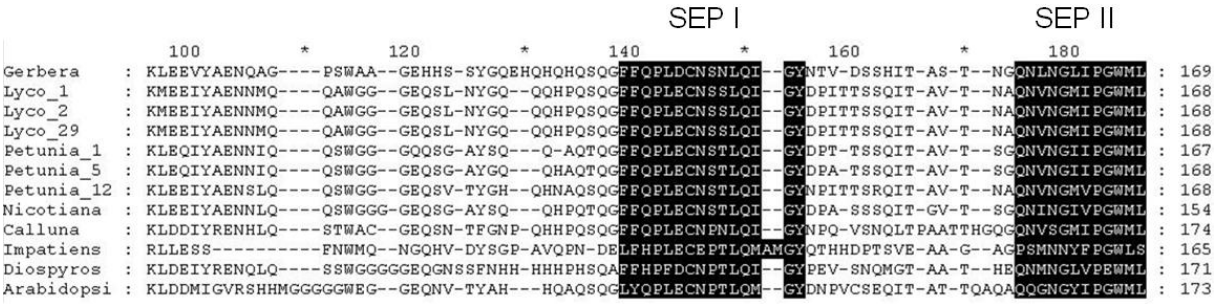


Figure 9: Alignment of the C-terminal region of *C. vulgaris SEP1* and its homologs, used for tree construction. The regions designated as motifs are inverted and labeled as SEP I and SEP II, respectively (according to Zahn et al. 2005).

2.3.2. Annex II: Phylogeny of *C. vulgaris* MADS-box genes

Phylogenetic analysis of partial gene sequences i.e. the comparison of *C. vulgaris* B- and E-like MADS-box genes vs. B- and E-MADS-box genes of related species (asterids) was performed using PaupUp (for determination of substitution model) and MrBayes (for consensus tree construction). Thus, the phylograms presented below (Figure 10, Figure 11) have to be esteemed to mainly represent the gene's phylogeny rather than the evolutionary pathway of *C. vulgaris*.

For phylogenetic analysis, matched sequences (by BLAST (Basic Local Alignment Search Tool) in NCBI) of the *Ericales* (if available) and the asterids were chosen and aligned to *C. vulgaris* sequences. The coding sequence of *A. thaliana* (rosids) *AP3* (*APETALA3*, NM_115294) and *SEP1* (*SEPALLATA1*, NM_121585) was included as anticipated outgroup. Although the DEF genes of *Impatiens* encoded an unusually long C-terminal domain (*IhDEF1*) or an unusual short C-terminal domain (*IhDEF2*), both were integrated.

Figure 10 (next page) shows the unrooted tree resulting for *CvAP3*, the class B-like MADS-box gene of *C. vulgaris* and other B-MADS-box genes of the asterids and Figure 11 (next but one page) provides the unrooted tree for *CvSEP1* (*C. vulgaris*), the class E-like MADS-box gene and other E-MADS-box genes of the asterids. The anticipated outgroup genes *AP3* and *SEP1* of *A. thaliana* were recognized as such in both cases. A shared feature of both trees is the underlying data basis, since only partial sequences of the coding region of the genes were available as the result of 3'-RACE-PCR: the *C. vulgaris* sequence was the shortest one in both cases. Therefore, it is not unexpected that terminal results i.e. highly resolved nodes with high confidence / significance levels were not achieved for *C. vulgaris* genes. In contrast, for *CvAP3*, the connection between the outgroup gene *Arabidopsis* (rosids) and the asterids remained as unclear as the relation between *CvAP3* and the genes from the same family. In case of *CvSEP1*, a low significance i.e. bootstrapping value for its node additionally decreases the informative value of the phylograms. Nevertheless, the *C. vulgaris* sequence is positioned near *Diospyros kaki* (DQ412058) which is, beneath *Impatiens hawkeri* (DQ493928), the only sample from the *Ericales*.

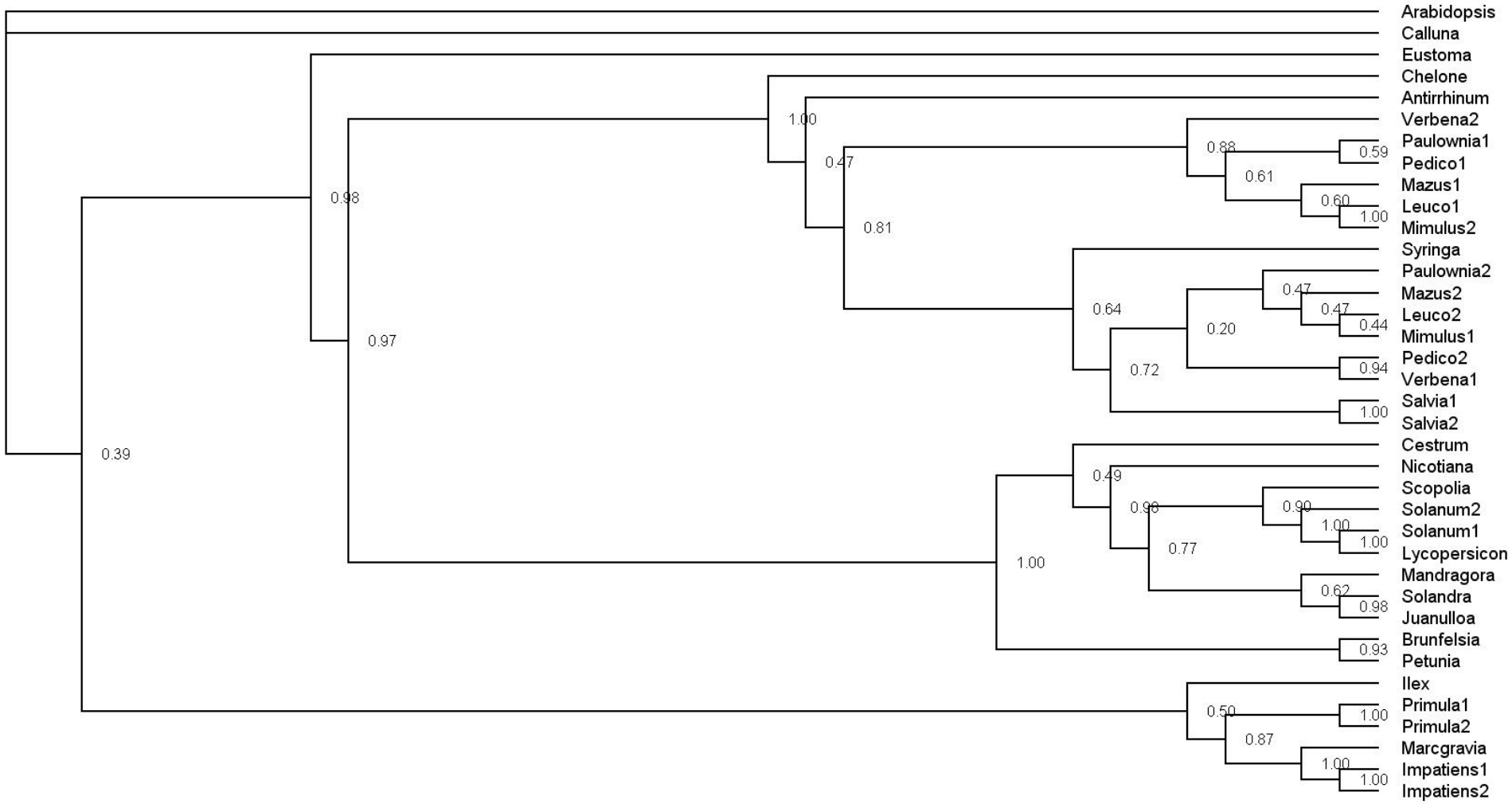


Figure 10: Consensus tree of nucleotide alignments (ClustalW2) of *CvAP3*. Best-fit model GTR+I+G selected by AICc (corrected Akaike Information Criterion, PaupUp), base frequencies 0.3048 (A), 0.2147 (C), 0.2521 (G), 0.2284 (T), burnin = 8500. Internal edge labels (MrBayes) are equivalent to posterior probability values.

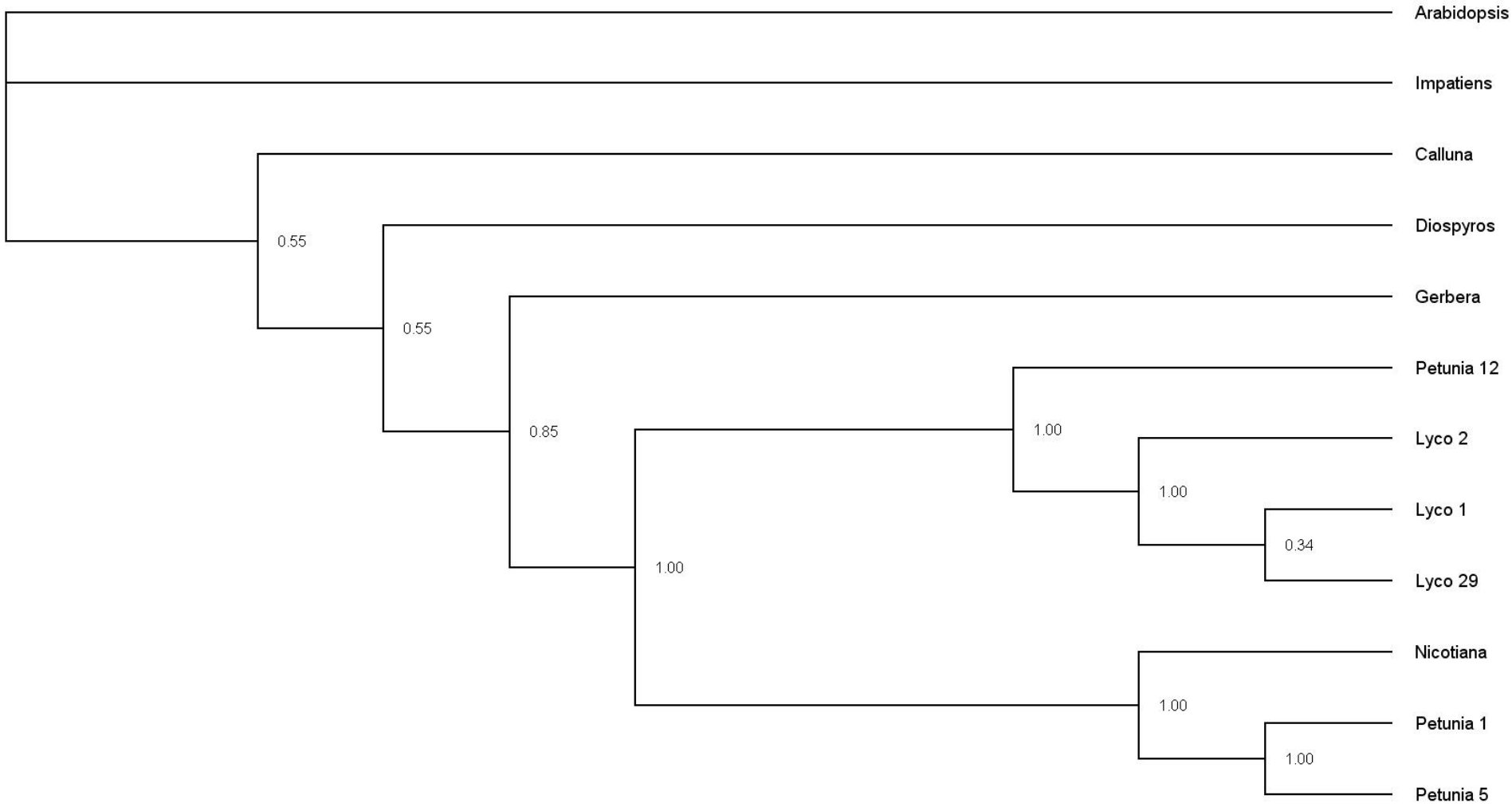


Figure 11: Consensus tree of nucleotide alignments (ClustalW2) of *CvSEP1*. Best-fit model HKY+I+G selected by AICc (PaupUp), base frequencies 0.3364 (A), 0.2244 (C), 0.2278 G, 0.2113 (T), Ti/tv ratio = 1.3025, Burnin = 700. Internal edge labels (MrBayes) are equivalent to posterior probability values.

3. General Discussion

3.1. Genetic Diversity & Variety Protection

3.1.1. Summary of Results

The selection of genotypes for the investigation of genetic diversity in *C. vulgaris* was agreed upon with the cooperating breeding company. The aim was to represent a wide assortment of the genus. For this reason, it included 62 varieties of current economic significance, older varieties from different countries and other continents (USA), five genotypes resulting from crossings, five wild-type samples and *Erica* genotypes (as an anticipated outgroup). The intention was to gain a comprehensive overview of the genetic diversity available in *C. vulgaris* and applicable for prospective breeding efforts from this mixture.

After having screened over 120 primers, 168 mono- and polymorphic RAPD (129) and ISSR (39) fingerprint fragments were applied to the above-mentioned genotypes by random PCR. The bootstrapped dendrogram is shown in Fig. 2 in Chapter 2.1. A low genetic diversity was revealed: all *Calluna* genotypes clustered with a Dice similarity of ~ 0.74 or higher. If the wild types from Thuringia and Italy are treated as an internal *C. vulgaris* outgroup, this value increases to ~ 0.78 Dice similarity. We experienced a robust inner-laboratory reproducibility of the RAPD- and ISSR 'fingerprinting' techniques. This is proved by a very low amount of missing values (0.47%), i.e. non-reproducible PCR-fragments within the distance matrix. In addition, the selected primers were able to significantly discriminate between *E. carnea* and *E. tetralix*, genotypes included to prove the concept of the dendrogram construction method. The dendrogram is therefore considered to provide an adequate estimation of the genetic diversity in *C. vulgaris*.

The small number of statistically significant nodes indicates a low bootstrapping validation within the *Calluna* group of the dendrogram. Due to this minor statistical reliability and the simultaneous and close arrangements of known non-related individuals, the dendrogram was deemed insufficient to be applied when dealing with variety derivation in *C. vulgaris*. The lack of an adequate method led to the alternative computation of the available dataset on the basis of the statistical system proposed for lettuce (Eeuwijk and Law 2004). This assay compared genetic distances of the varie-

ties in question vs. a genetic distance threshold which was derived from an adapted reference set. For the selected test sets of *C. vulgaris*, this method proved to be applicable, since known EDV cases can be distinguished from cases relating to back-crossing (Figures 3 and 4 and Table 2 in Chapter 2.1., Figure 6 and Figure 7 in Chapter 2.1.2.). Thus, based on the same dataset that resulted in the dendrogram, the typical case of an EDV (i.e. sport selection) can reliably be differentiated from 'true breeding' (i.e. back-crossing), which is a basic demand of a system dedicated to EDV identification in *C. vulgaris*.

3.1.2. Discussion & Conclusions

Genetic diversity data are a prerequisite for successful breeding programs, i.e. 'molecular phylogeny-assisted' breeding (Handa et al. 2006). In *C. vulgaris*, additional facts, such as the combination of an increasing amount of filings for variety protection and diminishing phenotypic differences among the varieties, led to the demand for molecular investigations by breeders. Nevertheless, apart from studies on the genetic diversity of wild type populations (Mahy et al. (1997), Mahy and Jacquemart (1998), Mahy et al. (1999), Meikle et al. (1999), Rendell and Ennos (2002)), molecular data is completely lacking in *C. vulgaris*.

Our results proved a low genetic diversity in *C. vulgaris*. We assume that not only the crossing incompatibility of the monotypic species *C. vulgaris* with other genera of the *Ericaceae* but also the applied key breeding methods (back-crossing and sport selection) are responsible for this development, since both are known to negatively affect the gene pool (Vosman et al. 2004). As shown in Figure 5, the occurrence of outliers within the tested genotypes is rather coincidental, with the exception of the *Erica* spp. and the wild-type groups; a crossing progeny, the parents of which are well placed within the large *C. vulgaris* cluster, is itself positioned outside by PCA. In contrast, the single variety enclosed from the USA ('Autumn Glow') is indistinguishable from the European genotypes. Nevertheless, it is not known whether or not this variety resulted mainly from European genotypes simply exported from Europe to the USA. Thus, the application of these genetic 'fingerprinting' data, i.e. its exploitation to identify novel breeding material, may be a valuable tool, despite the fact that *C. vulgaris* breeding programs cannot be adapted towards the effective cross-breeding of regional subgroups, for instance. However, it can be assumed that concerted breeding

efforts towards desired traits offside the ‘bud-flowering’ trait, e.g. foliage or flower color, will be successful. Up to now, the small gene pool available has not been targeted systematically enough.

Breeding by methods well-known for their negative effect on the gene pool within species with high genetic similarities gives rise to problems regarding the phenotypic discrimination of genotypes. As a consequence, juridical variety protection issues resulted in *C. vulgaris.* With respect to this set of problems, even the German Federal Plant Variety Office recently searched for possibilities to complement phenotypic DUS testing in *C. vulgaris* by implementing marker techniques, i.e. isoenzyme analyses (CPV.4937 2007). In general, the introduction of molecular methods into Plant Variety Protection (PVP) methodology was a common development in recent decades (see, e.g. Staub and Meglic 1993, Staub et al. 1996b, Rout and Mohapatra 2006).

In Europe, the first PVP / Intellectual Property (IP) rights for plants were introduced in the Netherlands (1938) and in Germany (1941) in terms of the Plant Breeder’s Rights (PBR). Furthermore, the Plant Variety and Protection Law was enacted in Germany on December 11, 1985, and applications for national plant protection are addressed to the Federal Plant Variety Office. Since 1994, applications for PVP throughout the EU can be filed to the CPVO, Angers, France, on the basis of one of the first supraregional agreements: the European Regulation on the Protection of Plant Varieties (Le Buanec 2004). In 1930, the Plant Patent (PP) Act led to the possibility of variety protection in the USA, which was adjusted in 1973 and 1994 by the introduction of additional PVP features. While PBR are based on distinctness (of the declared variety to every known variety of the same species), uniformity (with respect to the propagation method) and stability (over the reproduction cycles) and are applied to all species, PPs as available in the USA are only legal to asexually reproduced varieties. Further important and systematic differences are that PPs prohibit claims on sexually-derived progenies of the declared variety and that no deposit of material is required at an independent institution. In summary, a notably less rigorous practical application of PP compared to PBR is reported (Ghijsen (Bayer CropScience), personal communication). However, the PP system is a much more stringent protection system.

The use of molecular markers in such legal environments requires consensus from appropriate scientific and business authorities, especially regarding suitable definitions of genetic distance and statistical significance. Figure 12 indicates the uncertainty related to the detection of essential derivation, independent of the method applied. A 'zone of uncertainty' remains between areas of non-ambiguity regarding essential derivation or independence (Heckenberger 2004). Among other things, the size of this zone is affected by the 'fingerprinting' methods that are applied to identify essential derivation. As a result, a universally applicable system to identify EDVs in all plant species and crops of economic interest will presumably not be available due to the differing genetic premises in the different species. Thus, a crop-by-crop approach that accounts for the current circumstances, i.e. the availability of methods, may be a more realistic advance for issues related to essential derivation.

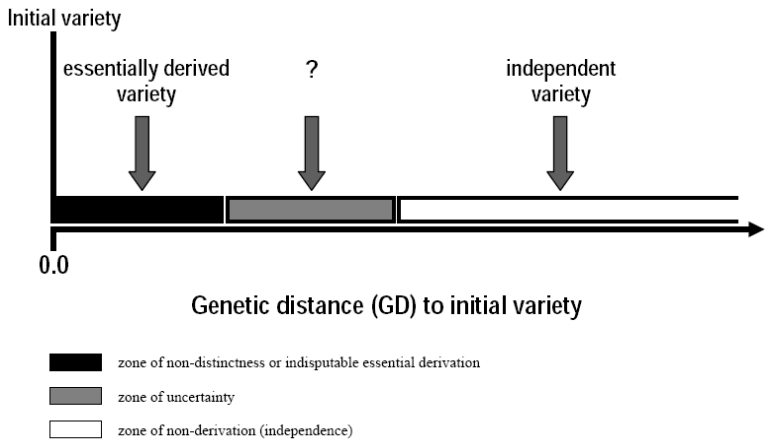


Figure 12: The concept of different thresholds to discriminate between essentially derived and independent varieties based on their GD (genetic distance) to an initial variety. (reproduced from Heckenberger 2004, Figure 1).

In *C. vulgaris*, the mere application of a dendrogram did not prove to be an adequate method for identifying EDVs. This is a critical difference when compared with other vegetatively propagated crops. De Riek et al. (2000) presented a number of case studies regarding this aspect for *Phalaenopsis, Rosa* and *Rhododendron.* With these species, the authors were able to successfully discriminate clones of a variety vs. non-related, i.e. independent varieties, using UPGMA dendrograms. For *C. vulgaris,* however, it could exaggeratedly be claimed that every variety was derived from the Lüneburger Heide wild types, since these are clustered amid all other samples with a very high identity percentage (Figure 5). Although plant material was repeatedly derived from this area (Heidepflanzen de Winkel, pers. comm.) for breeding purposes,

this does not seem to be a pragmatic solution for the pending and upcoming EDV issues.

Two explanations can be given for this elementary difference. First, the high genetic similarities in *C. vulgaris* give rise to high genetic similarities of any crossing progeny, independent of the parental genotypes. Even if assumed non-related individuals, e.g. a 'wild-type' from Germany and a 'bud-flowering' type from the USA, were used for crossing, the progeny will share a high amount of monomorphic PCR fragments with its parents, as their parents did before them. Second, the somatic mutation type of sport/reversion is a common, i.e. well-known, phenomenon in *C. vulgaris*. This may be concluded from pers. comm. (Heidepflanzen de Winkel) and from filings for variety protection kindly provided by the German Federal Plant Variety Office.

As early as in 1988, Klekowski indicated the combination of (i) vegetative propagation and (ii) the perennial character of a species to promote high mutation rates, which could be a possible explanation for the frequent occurrence of sports and reversions in *C. vulgaris*. However, individual plants of *C. vulgaris* do not reach high generation numbers during a typical production cycle in a horticultural company. Furthermore, the selection of spontaneously occurring mutants, i.e. sports, is a breeding method restricted not only to *C. vulgaris*. In contrast, such mutants have been exploited frequently for the introduction of new varieties in ornamentals breeding in all commercially relevant sectors, e.g. cut flowers and pot plants (Schum 2003).

For this reason, the efforts to solve the EDV issues in *C. vulgaris* focused on the implementation of a statistically reliable method. The system proposed for and applied to lettuce (Eeuwijk and Law 2004) was chosen for testing and adaptation since the underlying biological features were comparable to the circumstances of *C. vulgaris*. First, the applied marker system (AFLPs) was comparable to our already available RAPD- and ISSR-dataset with respect to random sampling characteristics (PCR priming). Although AFLPs have a higher performance in generating marker bands, the origin of AFLP fragments may be comparable to RAPDs and ISSRs. Second and analogue to *C. vulgaris*, a low genetic diversity could be assumed for lettuce, at least within the subgroups of economical relevance (Eeuwijk and Law 2004 and references therein). However, the threshold described in Chapter 2.1. should not be interpreted as an universally applicable threshold for *C. vulgaris*, since each adaptation of the reference set with regard to a new set of varieties in question will lead to a slightly different threshold.

The term ‘fingerprinting’ was apparently derived from the human medical and forensic sectors. In and for humans, a fingerprint is capable of unambiguously identifying a human being by comparing its epidermal ridge structure vs. other people’s or a database. Since no two identical fingerprints are known, this technique is accepted to be highly discriminative. The ‘genetic fingerprint’, mainly applied for certificates of parentage, is usually either founded on the PCR-based amplification of conserved VNTRs (variable number of tandem repeats) or STRs (short tandem repeats) since, e.g. the amount of STRs is a unique feature for each human being. Thus, the comparison of a sufficient number (usually 16) of maternal and paternal repetitions vs. the STRs of an assumed child facilitates the confirmation of maternity and paternity with a sufficiently statistical likelihood (see, e.g. Benecke 1997). In Germany, use of such diagnostic methods is regulated within the ‘GenDG’ (GenDG 2009).

In contrast, ‘fingerprinting’ plant genomes is generally a different approach. Some types of random sampling techniques, e.g. RAPDs and AFLPs, only analyze a very small range of a usually non-defined and unknown region of a plant’s genome. Thus, the absolute amount of (mono- and) polymorphisms generated by one primer set, i.e. a single PCR reaction, is frequently insufficient to discriminate two plant genotypes in a statistically significant manner. By contrast, this is the case for human fingerprints. For this reason, several more primer combinations usually have to be incorporated and statistical tools and algorithms have to be applied to increase the probability of discrimination and to decrease the error rates when investigating plant genomes. However, in contrast to *Homo sapiens*, the possibilities to reproduce a plant species are manifold and range from selfing to back-crossing. Furthermore, depending on the technique, somatic mutations may lead to an alteration of the ‘fingerprint’ over generations. In addition, the underlying question for ‘fingerprinting’ plants usually deviates from paternal relationships in humans, since, e.g. for EDV issues, dendrograms are constructed to investigate the relationship between two varieties at a defined point in time. The question regarding a common ancestor is irrelevant in this context. Hence, the premises and requirements of ‘fingerprinting’ are different in humans and plants.

In order to emphasize these differences between fingerprints of the medical / forensic sector and plant biotechnology, I propose instead to call the snap-reading method of the ‘plant fingerprinting’ technique and its results a ‘genomic snapshot’ rather than ‘fingerprinting’. This would reflect the capability of some methods to provide only an

outline of a current genomic situation more effectively. Besides, the connotation of ‘fingerprinting’ and absolute reliability would be circumvented.
However, the disadvantage described does not hold for all ‘fingerprinting’ methods applied in plants. Simple Sequence Repeats (SSRs) in plants are presumably comparable to the forensic ‘fingerprinting’ techniques in humans, since they are based upon definite sequence repeats (< 6 bp tandem nucleotide repeats) that are prevalently existent in plant genomes (every 7 kb to 1.5 Mb) and since common motifs were identified, e.g. $(AAG/TTC)_n$ (see Wolfe and Liston 1998 and references therein).

3.2. Molecular Marker

3.2.1. Summary of Results

Previous results and experiences of breeders indicated a monogenic-recessive inheritance of the economically important ‘bud-flowering’ trait in a Mendelian fashion. This theory evolved from crossing experiments at the former Federal Centre for Breeding Research on Cultivated Plants, Ahrensburg. Several back-crossing progenies approximately showed a 1 : 1 segregation of the flower types (‘wild-type’ : ‘bud-flowering’). The production of several new back-crossing populations at the IGZ did not disprove this theory, since they typically segregated as expected (1 : 1 regarding the trait of interest, Table 1 in Chapter 2.2.). No independently segregating traits, e.g. foliage of flower color, could be identified within the larger populations. Nevertheless, the assumed ‘bud-flowering’ major gene is presumably influenced by one or more background genes and/or by unknown environmental effects, since segregation ratios of up to 1 : 33 were detected (Table 3 in Chapter 2.2.1.).
Within the single population (n = 68 individuals, segregation ratio 1 : 1) available at the start of the project, screening for molecular markers by RAPD (n = 120) and ISSR (n = 10) primers resulted in the identification of two polymorphisms closely linked to the economically most important trait of ‘bud-flowering’. With regard to the economic interests of the cooperating company, the markers are ciphered ‘A’ and ‘B’, respectively. Both markers were only applicable in the original population used for primer screening and not in independent varieties (n = 49 individuals tested, Table 2 in Chapter 2.2.) or another back-crossing population (n = 119 individuals tested).

Cloning and sequencing of the ‘A’ marker resulted in the identification of a highly conserved Zn-Finger domain and the partial open reading frame (ORF) of a Nicotinamide-Adenine-Nucleotide (NADH)-dehydrogenase I. In addition, the 5’-flanking RAPD primer binding site was recognized. No putative function could be assigned to the sequence of the ‘B’ marker. Thus, sixteen 20 to 27 bp SCAR primers were derived from the available sequences and combined to amplify several fragments between 0.4 and 1.1 kb in length. All nested primer combinations amplified monomorphic fragments in DNA bulks and individual plants of the screening population and, therefore, the initial polymorphism was lost.
Pragmatically, further application of MAS in *C. vulgaris* would only be possible if novel markers were generated or new polymorphisms were detected within the monomorphic SCARs. As an initial step, therefore, the SCARs were analyzed for SSCPs: single-strand DNA was separated by polyacrylamide gel electrophoresis (PAGE). This assay resulted in the re-establishment of two polymorphisms within one of the 400 bp SCAR fragments: the additional fragments were traceable in the ‘wild-type’ genotypes of the initial back-crossing population and not in the ‘bud-flowering’ genotypes.

3.2.2. Discussion & Conclusions

A genetic marker is a trait of at least two configurations (alleles) inherited in a Mendelian manner, frequently based on point mutations, e.g. SNPs or larger InDels that can be verified within one population of a species. Molecular markers. e.g. RAPDs or AFLPs, have fundamental advantages over morphological marker systems, e.g. phenotypic traits, since they are (i) not prone to environmental effects and – depending on the genome size of the organism and the technique applied – (ii) are theoretically available in a nearly infinite amount (Kumar 1999).

The financial relevance of investigations regarding the establishment of molecular markers is considered the major aspect in *C. vulgaris*, since at least two years of breeding effort, i.e. crossing and selection attempts, are necessary to build up adequate populations for the production of ‘bud-flowering’ genotypes. The ability to use molecular techniques as a diagnostic tool to identify the flower type before natural flowering occurs would save production space in the greenhouse for at least six months. Thus, the introduction of MAS in *C. vulgaris* is of great interest to breeders

since it may have the potential to reduce the funds required to breed ‘bud-flowering’ genotypes.
Only few comparable examples exist in ornamental crops for functional molecular MAS systems aiming at a distinct phenotypic trait of both ornamental and economic relevance, e.g. a flower architectural trait (like ‘bud-flowering’ in *C. vulgaris*). Thus far, results have been provided by Scovel et al. (1998). These authors successfully converted a cloned RAPD marker for flower doubleness in *Dianthus caryophyllus* to a functional RFLP (Restriction Fragment Length Polymorphism) marker. In 1999, Dunemann et al. published results regarding linkage maps and quantitative trait loci (QTLs) for leaf chlorosis and flower color in *Rhododendron* using RAPDs, RFLPs and microsatellite markers. Debener and Matthiesch (1999) presented results regarding the RAPD- and AFLP-based mapping of genes for petal numbers and flower colors in *Rosa hybrida*.

Due to the postulated monogenic mode of inheritance of the ‘bud-flowering’ trait in *C. vulgaris*, BSA is considered an appropriate method to screen for molecular markers (Michelmore et al. 1991). This technique implies the creation of DNA bulks in each group – the ‘wild-type’ and the ‘bud-flowering’ flower type – and the screening of these bulks using, e.g. a random PCR assay (here: RAPDs, ISSRs).
Since ‘bud-flowering’ is assumed to be inherited as the recessive allele in Mendelian fashion, detection of a PCR fragment in coupling phase to the ‘wild-type’ was more probable. Thus, the two bulks were differentiated by a single band: the dominant group (‘wild-type’) showed the band while the fragment was absent in the recessive group (‘bud-flowering’). In repulsion phase linkage, both bulks would have shown the fragment and, using a codominant marker system instead of dominant RAPDs, two fragments would be amplified, one from each allele. Hence, the dominant group would show both bands while the recessive group would only show one band. However, the coupling phase RAPD markers available are unable to discriminate between homozygote and heterozygote individuals of the ‘wild-type’ fraction. In addition, screening a population for the absence of a PCR fragment is less eligible since this absence may also be due to PCR reaction failures instead of a missing allele.
Both detected coupling phase RAPD markers were not applicable beyond the original screening population. Analysis of the fragment’s sequence did not reveal specific gene functions related to any flower trait or organ compartment. Instead, a Zn finger

domain and an NADH dehydrogenase I were identified, which are highly conserved and widely distributed proteins. NADH dehydrogrenases are mitochondrial transporter proteins for electrons, while Zn fingers are usually involved in DNA / RNA binding (transcription factors). Thus, assumptions regarding potential explanations for the occurrence of the markers in only one population are not possible based on this sequence information.

For this reason, efforts were made to improve the RAPD markers by increasing their specificity. This is frequently achieved by the transformation into Sequence Characterized Amplified Region (SCAR) markers (see, for example, Paran and Michelmore 1993). These markers theoretically offer an increased PCR-specificity and decrease the potential risk of identification of false-positives or false-negatives by means of an eventually existent low reproducibility of RAPD-PCR. In actual fact, the main application of improved SCAR markers and possibly a result of the economic significance of these factors are resistance traits, e.g. the selection for apple scrab resistance using RAPDs in *Malus floribunda* (Yang and Krueger 1994), the identification of molecular markers (AFLP, RFLP, SCAR) linked to the black spot resistance gene in *Rosa* sp. (Malek et al. 2000) or the selection for *Fusarium* resistance in *D. caryophyllus* using RAPD and RFLP techniques (Scovel et al. 2001). A recent work in *Fragaria vesca* (Albani et al. 2004) reported the successful development of ISSR-derived SCAR markers for *SFL* (*SEASONAL FLOWERING LOCUS*), and Scheef et al. (2003) described the improvement of species-specific SCAR markers in bentgrass (*Agrostis* spp.).

If SCAR primers are designed to include the initial RAPD binding sites, it may be possible to circumvent the loss of the polymorphism (see, for instance, Scheef et al. 2003, Zhang and Stommel 2001). This applies to cases where the RAPD binding site was identified by cloning instead of direct sequencing (Hernandez et al. 2003). If this approach is not possible, e.g. due to gaps within or missing ends of the sequence data, nested SCAR primers are designed from within the RADP fragment, thereby ignoring the specific site that originally introduced the polymorphism. Hence, the primers will (and frequently do) amplify monomorphic fragments, i.e. fragments that occur in both pools – which was true for all developed SCAR primer combinations in *C. vulgaris*, even if the one available RAPD binding sites was covered by the elongated primers. Thus, the development of functional SCAR markers failed for

C. vulgaris. In case of the 1.2 kb marker, the belatedly assumed cloning of multiple (at least) two PCR fragments may be an explanation. Regarding the 2.5 kb marker, the full length sequence of the fragment was absent. Hence, only one of the RAPD primer binding sites could be integrated into SCAR primers, which is not a positively supporting element for successful SCAR marker development. However, even SCAR primers designed from cloned RAPD fragments do not allow the preservation of polymorphisms that were initially based on simple RAPD primer mismatches.

Recovery of such an initial but lost polymorphism, e.g. by a co-dominant marker system with increased benefits, can be achieved using a variety of techniques, e.g. SSCP-analysis (Orita et al. 1989). SSCP studies include the denaturing of the DNA and the electrophoresis of these samples on a special non-denaturing polyacrylamide gel. Applying this technique, the single strand DNA (ssDNA) strands differentially migrate within a gel matrix due to the conformational differences that are formed. These differences are the result of one or several SNPs within the monomorphic fragment, which cannot be detected using standard electrophoresis conditions. By applying this technique, two different ssDNA fragments within individuals of the ‘wild-type’ pool of the initial segregating population were identified. In accordance to the initial RAPD marker, differing banding patterns between ‘normal-‘ and ‘bud-flowering’ individuals were identified, although the rate of false positives was higher for the SSCP fragments (6.9% to 9.3%) compared to the RAPD ‘A’ marker (2.3%).
With regard to the practical applicability of the marker system within horticultural and breeding companies, the SSCP technique could be considered unfavorable when compared to common agarose gel electrophoresis, due to its sophisticated technical requirements. Acquisition of the necessary equipment for this type of analysis is quite unrealistic for medium-sized breeding companies, which are in the majority in the *C. vulgaris* community.

The occurrences of deviating segregation ratios (Table 3, Chapter 2.2.1.) indicates the existence of additional factors influencing the feature characteristics of the ‘bud-flowering’ trait. On the one hand, these factors may be of environmental terms, e.g. temperature and light conditions on the day of pollination or during the germination phase. This supposition may fit to crossings #3 and #4, since in these cases the same parent individuals were used. However, both populations differed markedly in

their total number of seeds (#3: 52; #4: 132) and in their germination rate (#3: 0.94; #4: 0.18), respectively. On the other hand, genetic preconditions which may range from differing pollen acceptabilities of different crossing combinations to a diverse genetic fitness of seeds of, e.g. the 'bud-flowering' type may influence the segregation ratios. Furthermore, the assumption of simple monogenic inheritance by a major gene may be imprecise if one or several more background genes are taken into account that may regulate the anticipated recessive allele of the 'bud-flowering' trait.

Hence, in order to obtain adequate practicability, further improvements of the 'bud-flowering' trait marker are essential.
The development of CAPS markers (see e.g. Zhang and Stommel 2001) is a promising approach. These markers are either created by the digestion of monomorphic SCAR-fragments from both bulks with randomly chosen restriction enzymes, which may result in polymorphic banding patterns. Or, which is the more pragmatic procedure, SCAR fragments from several genotypes of each pool are cloned and sequenced. As a result, polymorphisms (e.g. SNPs, InDels) may become detectable that could enable the identification of trait-related recognition sites for restriction enzymes. Using this technique, SCAR fragments of both pools could be differentiated in agarose gels by simple restriction assays due to differing digestion patterns.
The discriminative effectiveness of a repulsion phase marker is much higher than its coupling phase pendant (Haley et al. 1994). Thus, re-screening *C. vulgaris* backcrossing populations by BSA in combination with AFLP markers is assumed to be more efficient for targeting monogenic traits (Wehling 1998), although co-dominant marker systems, e.g. micro-satellites, are generally preferred.

3.3. The Flower

3.3.1. Summary of Results

The investigation of disputable and unknown flower organ identities in *C. vulgaris* was accomplished by combining phenotypic and molecular analysis.
The morphological analyses comprised the macroscopic study of mature flowers, the imaging of the epidermal cell structure of the perianth by Scanning Electron Microscopy (SEM) and the histological analysis of flower development. All aspects were

comparatively investigated for the ‘wild-type’ and the economically important ‘bud-flowering’ type. As a prerequisite for any prospective molecular experiments, the first partial sequences of two *C. vulgaris* MADS-box transcription factors were identified and analyzed. Using 3’-RACE-PCR, C-terminal sequences of a B-like (*AP3*/*DEF*) and E-like (*SEP1/2*) gene were cloned. The identities of these genes were confirmed by BLAST and motif analysis: the *C. vulgaris AP3/DEF*-like gene *CvAP3* [GQ202026] contained the euAP3 and the PI-derived motif, the *C. vulgaris SEP1/2*-like gene *CvSEP1* [GQ202027] contained the SEP I and the SEP II motif (Chapter 2.3.1.). A statistical significant phylogenetic classification of these genes was not possible, presumably due to the common high variability of C-terminal sequences (Chapter 2.3.2.).

The identified genes were targeted for expression analysis by qRT-PCR in different flower organs. Expression levels of *CvAP3* gradually decreased from petals over sepals to bracts in both flower types. However, an almost fourfold higher expression was observed in petals of the ‘wild-type’ genotypes compared to its ‘bud-flowering’ counterpart. *CvSEP1* expression did not differ markedly between whorl I and whorl II organs of both flower types, but was clearly lower in the outermost whorl of green leaves. Both genes were expressed in vegetative tissue (foliage) which is quite unusual; however, the accordance of the PCR product with the predicted amplicons was confirmed by cloning and sequencing.

Macroscopical analysis revealed, that the ‘wild-type’ may be described by the floral formula $Ca^{4}Co^{(4)}A^{8}G^{(4)}$, the most common ‘diplocalyx’ subgroup of the ‘bud-flowering’ type by $Ca^{4+4}Co^{0}A^{0}G^{(4)}$, respectively, whereas the ‘bud-flowering’ variety ‘David Eason’, which presumably represents a characteristic class B gene deficiency mutant, is represented by $Ca^{4+4}Co^{0}A^{0}G^{(8+4)}$ (Ca: Calyx; Co: Corolla; A: Androecium ; G: Gynoecium). These formulas imply for the ‘bud-flowering’ phenotype it is deficient both of petals (Co^{0}) and stamens (A^{0}). On the one hand, this hypothesis is morphologically supported by the absence of ‘dome-shaped’ epidermis surface cells in whorl II organs; however, the overall number of colored leaves and their positioning has not changed in comparison to the ‘wild-type’. On the other hand, staining and cell structure appearance of stamens, petals and sepals were comparable by histological analysis: the bases of stamens and petals looked similar in the ‘wild-type’ phenotype, whereas in the ‘bud-flowering’ type, both whorl I and whorl II organs resembled each other. Moreover, we confirmed the lack of stamens during the floral developmental

process in the ‘bud-flowering’ phenotype since neither evidence for organ primordia nor belated abortion of preceding initiated stamens were observed.
The combined evaluation of molecular (expression levels of *CvAP3* and *CvSEP1*) and macroscopical (occurrence of stomata, coloring) data for the outermost whorl of green leaves allowed to classify these organs as bracts.
As another prerequisite for prospective genomic application, the genome size of *C. vulgaris* (~1.18 pg/2C, ~1,154 Mbp) was determined by flowcytometry.

3.3.2. Discussion & Conclusions

A flower is concertedly, although not comprehensively defined as a ‘determinate, compressed, bisexual reproductive axis composed of megasporangia (carpels), microsporangia (stamens) and a sterile perianth composed of at least one sterile laminar organ’ (see, e.g. Bateman et al. 2006, Baum and Hileman 2006, Theissen and Melzer 2007). More generally, the flower is one of the major breeding targets in ornamental crops since it has a significant impact on the overall attractiveness of the plant. In addition, the event of flowering, i.e. the transition from the vegetative to the generative stage, is an essential component of the developmental process in angiosperms that assures reproduction and the continuance of the species (Tan and Swain 2006).
The identification of floral homeotic genes whose regulative transcriptional functions are known to play an important role in flower organ identities and development led to the genetic ‘ABC’-model (Coen and Meyerowitz 1991) and its extension regarding class D (Angenent et al. 1995, Colombo et al. 1995) and class E genes (Pelaz et al. 2000, 2001a, 2001b, Ditta et al. 2004). A molecular-based support results from the ‘floral quartet’ model (Theissen 2001). The ‘ABCDE’-model is applicable to an extensive array of angiosperms (Kim et al. 2005), but not to all (Bowman 1997).
The essentials of the ‘ABCDE’-model are the MADS-box transcription factors that control other genes whose products are directly or indirectly involved in flower organ identity (Theissen and Melzer 2007). Most MADS-box proteins are of the MIKC-type in which the MADS-domain is (i) the most conserved domain and (ii) the key determinant of DNA binding within the ‘CArG-Box’ (‘CC-A rich-GG’, 5’-CC(A/T)$_6$GG-3’, Theissen and Melzer 2007). MADS-box genes have been identified in eudicots, monocots and even in humans (Krizek and Fletcher 2005). The amount of genes per flowering organism may exceed 100 (*A. thaliana*: 107, Teeri et al. 2006). Known regula-

tive mechanisms of MADS-box genes are several genes of the flowering initiation phase, e.g. *LFY* (*LEAFY*) or *UFO* (*UNUSAL FLOWER ORGANS*). In addition, microRNAs (miRNAs) and epigenetic controls (e.g. methylation or chromatin remodeling) are currently being discussed and investigated as additional regulative factors (Zik and Irish 2003).

Since only minor variations of gene activities are responsible for the wide range of flower morphology observations (Goto et al. 2001, Smyth 2001), investigations of homologs are applicable even in genera lacking sequence information, as is the case for *C. vulgaris*. Elucidation of the homeotic flower organ identity genes and the flower development is therefore a feasible research field that may also have a high practical impact on both agricultural and ornamental crops.

Nevertheless, alterations of the flower shape in some non-eudicot species (which include approx. 25% of angiosperm species (Buzgo et al. 2005)) necessitated modifications of the 'ABCDE'-model. In *Tulipa* and other *Liliaceae* (monocots), for example, the two outmost whorls consist of colored (pigmented) so-called tepals. The *Viridiflora* mutants of *Tulipa* (van Tunen et al. 1993) transform their first two whorls of tepaloid organs into sepals and the stamens of whorl III into carpels, respectively. Thus, it was suggested that class B transcription factors are expressed in these three whorls. Hence, the outer boundary of the B genes has moved towards the outer rim of the flower whorls and therefore, it is termed as 'shifting boundary' (Bowman 1997) or 'sliding boundary' (Kramer et al. 2003) model. *Viridiflora* mutants have also been described for roses (green rose, *Rosa chinensis,* Judd 2002). In *Amborella trichopoda* (*Amborellaceae*, phylogenetic classification unknown, Stevens 2001) gradual morphological transitions between the different organ types have been reported (Buzgo et al. 2004), e.g. the tips of the inner tepals resemble the connective tips of the stamens. Supported by expression analysis of B-function homologs (Kim et al. 2005), these finding led to the formulation of the 'fading borders' model, in which these continuing morphological changes are explained by different levels of expression of the floral organ identity regulator factors. Figure 13 (next page) comparatively outlines the three described models.

Gene expression patterns and morphological analyses of the remarkable diversity of floral structures in angiosperms suggest the 'ABCDE'-model to be a 'derived condition' (Soltis et al 2006) i.e. a state derived from an 'ancestral program'. Furthermore,

the ‘silding boundary’ model has presumably developed independently for monocots (e.g. *Tulipa*) and basal eudicots (e.g. *Ranunculus*, Soltis et al. 2006). Hypothetically, the ‘fading border’ may reflect the ancestral program from which the other models were derived.

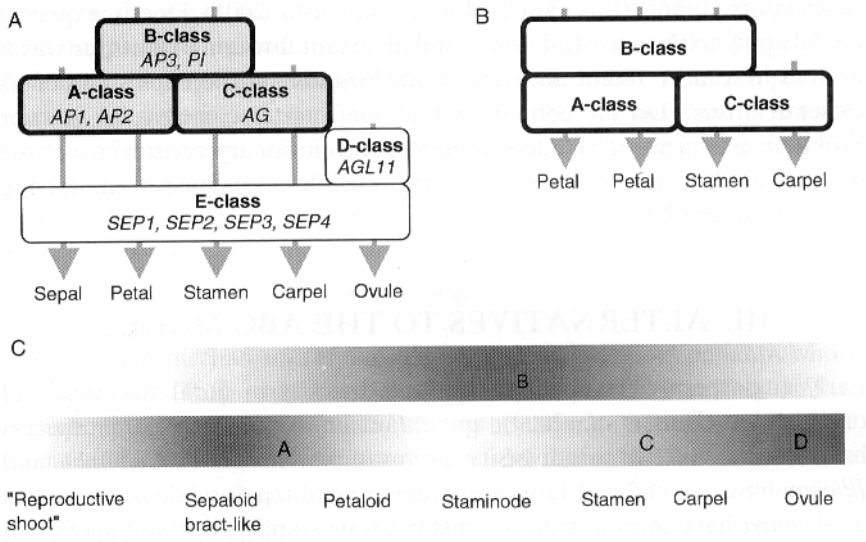


Figure 13: Comparison of classic ‘ABCDE’-model of floral organ identity (A) with sliding boundaries (B) and fading borders (C) models (reproduced from Soltis et al. 2006, Figure 1).

As a conclusion, the analysis of gene expression of two identified MADS-box genes in *C. vulgaris* showed once more that the strict application of these models is not feasible in every situation i.e. for every species. Therefore, this molecular approach to clarify flower organ identity requires completion by cloning and analysis of the remaining transcription factors. For example, it is known that both *AP3* plus *PI* (Geuten et al. 2006) and *SEP3* genes can concertedly regulate perianth morphology, e.g. in *A. thaliana* (Urbanus et al. 2009).

The created phylograms for *C. vulgaris* (Figure 10, Figure 11) may provide access to the phylogenetic progress of a species or, more obvious, of the gene itself. However, they provide a deviant view on the phylogeny more often than those trees based on data usually used for phylogenetic analysis i.e. chloroplast DNA. This is a consequence of the origin and quality of the data. On the one hand, MADS-box genes contain the conserved MADS-box and therefore, the resulting tree does not precisely correspond to the phylogeny i.e. similarities may be overestimated. On the other hand, C-terminal ends comprise conserved motifs, too (Figure 8, Figure 9) but additionally, they are known to be highly variable in sequence (Zahn et al. 2005). Thus, homology may even be underestimated. Presumably due to the short length of the *C.*

vulgaris sequences *CVAP3* and *CVSEP1*, the presented phylograms do not satisfactorily resolve the phylogenetic relationships. Full length cloning of these genes may improve the situation.

The flower shape is a key trait related to the overall attractiveness of ornamental crops: the regulation of the expression of floral regulators allows the direct modification of the flower shape. By altering expression profiles of the class ABCDE genes, any floral organ type can be generated in any whorl (Krizek and Fletcher 2005).
For example, filled / double flowers are valuable to the horticultural industry. According to the 'ABCDE'-model, these flowers are usually a result of a class C deficiency, since antagonistic class A genes are expressed as a substitute in the inner whorls, and thereby induce (i) the loss of flower determinacy and (ii) the formation of sepals and petals. Partially filled flowers have already been achieved by genetic modifications in *Gerbera* (*Asteraceae*). Teeri et al. (2006) describe the 'Regina' variety, which was transformed with an antisense C-function construct (*GAGA2*); thus additional petals and indeterminate meristems were formed (Yu et al. 1999). By way of transformation with a B-function gene (35S-*GGLO1*), the cosuppression of *GGLO1* resulted in the development of pappus hairs instead of petals (Yu et al. 1999). In *Chrysanthemum morifolium* (*Asteraceae*), Aida et al (2008) were able to partially induce a secondary corolla within the ray florets by an *Agrobacterium*-mediated transformation with an antisense *Chrysanthemum-AGAMOUS*-homolog.
In lily (*Lilium* spp.), the 'Aphrodite' variety converts its stamens into petals but maintains its carpels and floral determinacy in 'wild-type' style (Akita et al. 2008). An investigation into this unusual C-mutant revealed a reduced C-class gene expression in whorl 3 and the morphological similarity of whorl 3 petals with the tepals in the outer whorls. The further reduction of C-gene expression in lily flowers by genetic engineering assays would therefore allow the induction of double-flowered lilies (Akita et al. 2008).
With respect to the flower shape of *C. vulgaris*, which is a major feature of attraction in this ornamental crop, the investigation and initial clarification of the underlying molecular and morphological features were necessary. Isolation of the first MADS-box sequences in *C. vulgaris* (*CvAP3*, *CvSEP1*) and analysis of their expression represented the first step towards the molecular-based understanding of flower organ

identity. The morphological characterization of organ homology may be a supplementary and supporting advancement.
Naturally occurring *C. vulgaris* phenotypes with changed flower morphologies are already known, albeit uncharacterized. For example, the filled 'Radnor' variety (Figure 1CD of Chapter 2.3.) is assumed to be a potential candidate for C-function deletion. Furthermore, the 'David Eason' variety (Figure 4C of Chapter 2.3.) develops carpels instead of stamens; in addition to the morphological appearance of its petaloid sepals, which are comparable to the corresponding organs of the standard 'bud-flowering' type, this variety may be a candidate genotype for B-gene deletion in whorls II and III.

Thus, our results may be a useful starting point for prospective approaches to alter flower shapes in *C. vulgaris* by genetic engineering techniques. Until now, however, in vitro propagation of *C. vulgaris* was only shown to be functional for axillary shoot producing cultures for one variety (Gebhardt and Friedrich 1987), and an efficient transformation protocol for this woody species is absent. Hence, the specific production of 'bud-flowering' genotypes is a task that lies in the distant future. Nonetheless, *C. vulgaris* usually flowers each year and, provided optimal propagation and cultivation conditions are given, is able to flower only a few months after cutting multiplication or sowing. Thus, further experiments targeting the homeotic MADS-box genes and their upstream regulators are conceivable in principle due to the positive preconditions.
Genetic engineering has developed into a breeding method of interest for ornamentals (see, e.g. Tanaka et al. 2005 and Chandler and Tanaka 2007 and references therein) as it may support traditional breeding techniques. However, negative public acceptance regarding the application of genetic engineering and biotechnological techniques in food and plant science both in the USA and Europe (Klingemann et al. 2006) will presumably hinder any practical implementation in both agricultural and ornamental crops (see, e.g. Debener and Winkelmann 2009 and references therein).
Modification of the flower shape is only one of the targets of genetic engineering; until now, results in ornamental crops have been published for, e.g. flower color modifications of gerbera (Lu et al. 2003), carnation (Florigene Moon® series, Chandler and Tanaka 2007 and references therein) and rose (Katsumoto et al. 2007) or for modifications of post-harvest features (vase life) in carnation (Chandler 2007). Variation of

the flower color is usually targeted by influencing the metabolic pathway of flavonoids, which are considered most important for flower colors from red to blue. In rose, (Katsumoto et al. 2007) transgenic plants expressed a flavonoid 3´,5´hydroxylase introduced from viola, that led to the accumulation of delphinidine and relatively blue flowers from these experiments. Although such 'blue roses' have never been achieved by traditional breeding efforts, the transgenic plants have not yet been marketed. However, Florigene (http://www.florigene.com, as at May 2009) 'considers merchandising after production, distribution, and sales structures have been properly established'.

A blue *Calluna* would be an interesting breeding target, too (Heidepflanzen de Winkel, pers. comm.), since the common flower colors only range from white to lilac. Nonetheless, in-depth knowledge is even lacking for the well understood metabolic pathways of anthocyanins (Debener and Winkelmann 2009 and references therein) in this species. Such an approach towards the improvement of breeding in *C. vulgaris* is therefore rather unrealistic at present.

Furthermore, data from molecular analysis of floral regulator genes is applied to study evolutionary pathways in and of model organisms, for example in orchids (see, e.g. Mondragón-Palomino and Theissen 2009). As with lilies and tulips, the flowers of the *Orchidaceae* (monocot) consist of two whorls of petaloid organs (tepals) surrounding the inner reproductive organs. But even the tremendous morphological diversity of orchid flowers can be explained by an underlying genetic code (the 'orchid code', Mondragón-Palomino and Theissen 2008), which is once again based on known floral homeotic MADS-box genes. In this adapted model, the function of *DEF*-like B genes is distributed to four different clades, but the functions of class A and class C genes, as well as of *GLO*-like class B genes, remain untouched (Mondragón-Palomino and Theissen 2008). However, this functional divergence in class B gene expression patterns is able to explain the development of the different tepal organs of the orchid perianth: expression of clade 1 and 2 leads to the three outer tepals, expression of clades 1, 2 and 3 specifies the lateral tepals and the simultaneous expression of all four clades of *DEF*-like genes leads to the development of the lip.

Thus, identification of additional MADS-box genes in *C. vulgaris* and related genera may eventually account for the phylogenetic clarification of the disputable relationship

of the *Ericales,* since this order is still a polytomic group within the asterids (Stevens 2001).

The main drawback for any comparative genetic study of the flower morphology in perennial plants such as *C. vulgaris* is the separate and distinguishable juvenile and adult phase. The transition between these phases only affects some meristems, since continued growth is maintained by other meristems resting in their vegetative state (Tan and Swain 2006). The duration of the juvenile phase is species-dependant and can last, e.g. for the *Citrus* tree (Davis and Albrigo 1994), from two to more than ten years. Although sequences of some woody plants have already been isolated, e.g. from apple (*Malus sylvestris,* e.g. Kotoda et al. 2000), eucalyptus (*Eucalyptus grandis*, e.g. Dornelas et al. 2004), citrus (e.g. Endo et al. 2005) and wild rose (*Rosa rugosa*, Matsumoto and Fukui 2007), the confirmation of the gene function within mutants of woody plants is complex. On the one hand, the need for intact and mature plants is a simple task in annual plants but complicated and time-consuming in perennials. Thus, long generation times usually hinder their practical application and investigation. On the other hand, limited genetic knowledge of the respective crop is a common and supplementary disadvantage of woody perennials within these surroundings (Tan and Swain 2006). However, perennial flowering genes function like their *A. thaliana* counterparts (Wilkie et al. 2008).
Instead of 'horticultural trees' (Wilkie et al. 2008), the rose may become the model organism for studying the flowering process in perennial plants (Foucher et al. 2008) since it is capable of recurrent blooming and moreover, the juvenile phase is reduced to six to eight weeks (Foucher et al. 2008).

The control of flowering also includes understanding floral initiation. From the breeder's point of view, the critical parameters for initiation of flowering in *C. vulgaris* were already outlined by Dixon and Dutton (1987) and Kwolek and Woolhouse (1982). Flowering in *C. vulgaris* is induced by long-day treatment and inhibited by short-day conditions. Temperatures above 14°C promote flowering, too. Breeding *C. vulgaris* could be fundamentally improved, i.e. accelerated, if additional flowering periods were available for crossings, since each additional flowering season per year would allow a doubling of the amount of crossing combinations. This intention necessitates a reproducible system for flowering control under practical circumstances, i.e.

light and temperature conditions should be easily deployable in small breeding companies using typical horticultural equipment.
However, light and temperature are not the sole factors controlling flower initiation in either annual or perennial plants, although the mechanisms and effective environmental factors are deemed to be similar (Wilkie et al. 2008). Apart from photoperiodic events, floral induction is known to be induced by gibberellins and vernalization mechanisms. The underlying genetic factors involved (e.g. *FT* ('florigen', *FLOWERING LOCUS T*), the floral meristem identity gene *LFY* and *FRI* (*FRIGIDA*)) are well-known in *A. thaliana* and have therefore already been partially identified in woody species (Wilkie et al. 2008 and references therein).
Thus, comprehensive additional efforts at both the molecular level and at the environmental stage are crucial to achieving a sustainable advance in breeding the most valuable trait of *C. vulgaris* – the flower.

3.4. Perspectives for *C. vulgaris*

3.4.1. Breeding

Improving the breeding conditions of *C. vulgaris* is closely linked to the enhancement of the genetic diversity available for crossing. For this reason, exploitation of the naturally available gene pool is an essential task in order to consequently breed for interesting traits. In addition, a global gathering of wild type species may be profitable. Based on such a reference collection, systematic crossings that exploit important traits, e.g. an extended winter hardiness from wild samples or additional flower colors, should be performed. Further breeding targets of economical relevance are the foliage colors, flower size, date of flowering or elongation of the flowering time (Heidepflanzen de Winkel, pers. comm.). New growth habits, e.g. a more compact foliage structure, would enable easier handling during the pruning stages (Heidepflanzen de Winkel, pers. comm.). These quantitative traits require concerted breeding, which has not yet been performed in *C. vulgaris*. Compared to other ornamental crops, breeding in *C. vulgaris* is still assumed to be at the early stage.
In parallel, it is necessary to optimize the selection process concerning its expenditure of time and general effectiveness. This implies the definition of the most important traits as applicable selection criteria, since these economically relevant criteria

do not necessarily match the DUS criteria for variety discrimination. Additionally, breeding could be improved by optimizing the crossing strategy itself since, up to now, explanations for the observed diverse crossing results, i.e. fertilization successes, remain unclear. Thus, perfecting the pollination technique regarding the factors time, manual labour and crossing combinations could be useful. Furthermore, the capability of reproducibly controlling the flower induction mechanism in *C. vulgaris* would be helpful, since repeated belated crossings of the biological reproduction cycle in autumn would become feasible.
In addition, the circumvention of crossing barriers with other genera, e.g. by embryo rescue techniques, is a continuative assay to further advance the genetic diversity of ornamental crop *C. vulgaris*.

3.4.2. Molecular research

With regard to the economically most significant 'bud-flowering' trait, two key tasks are available for prospective research. First, clarification of the flower organ mutation(s) underlying this phenotype would be a helpful prerequisite for prospective breeding strategies. Second, a diagnostic molecular marker is presumably able to improve the breeding process. These two pathways should be pursued in parallel.

However, the establishment of a diagnostic marker may require the generation of new and preferably larger segregating populations. These could be used for further predictions regarding the mode of inheritance of the 'bud-flowering' trait as well as for mapping strategies, since markers linked more closely to the trait of interest are identifiable in larger populations. For such mapping strategies, a methodological switch to molecular techniques that allow superior throughput and performance, e.g. AFLPs, should be considered. The newly generated populations could be (i) screened for additional marker loci and (ii) used for fine-mapping of the trait of interest. In parallel, these populations would assist in the identification of co-segregating traits. Nevertheless, supplementary efforts should be initiated to improve the available RAPD markers by, e.g. re-sequencing the fragments or design of functional SCAR or CAPS markers.

Future research efforts in the area of flower organ identity in *C. vulgaris* may target the molecular basics of the 'bud-flowering' trait. The identification of additional flower

organ-related homologs is therefore a requirement, which could be accomplished by additional RACE experiments targeting the remaining MADS-box transcription factors and their improvement by full-length cloning. Another promising but technically more sophisticated approach towards differential expression patterns of 'wild-type' and 'bud-flowering' genotypes would be the construction of SAGE (serial analysis of gene expression) databases. Although SAGE analysis does not require existent sequence information, its value for a species lacking molecular data, such as *C. vulgaris,* remains to be seen. However, the approach would be independent of any knowledge regarding candidate genes. Recognized gene sequences could be exploited towards primer design for expression studies in flower tissues.

Due to the small genome size of *C. vulgaris*, the development of a BAC library may be a feasible advancement for *C. vulgaris* genomics, since even new genes may be discovered by (partial) sequencing of the library.

Finally, the specific production of 'bud-flowering' genotypes by genetic engineering, i.e. the specific induction of the 'bud-flowering' trait while maintaining all other desirable features of an individual, may be assumed to be the ultimate target of a *C. vulgaris* breeder. However, the nonexistent public acceptance of genetically modified plants is presumably the major obstacle to the successful commercial launch of such plants.

3.4.3. Final conclusion

Besides the establishment of publicly available molecular data concerning genotyping issues and flower organ identity, morphological observations of valuable characteristics of *C. vulgaris* were published. Furthermore, the generation of segregating populations led to the formulation of a hypothesis regarding the inheritance of the economically most significant 'bud-flowering' trait and the establishment of a related basic MAS system initially showed the applicability of molecular techniques with relevance to practically important traits within this species.

The results presented within this thesis and the derived conclusions and indications may be useful to the future improvement of breeding strategies in *C. vulgaris*. Although implementation of molecular methods is still at the initial stage in this species, this thesis has demonstrated the basic feasibility of a set of methods that could offer quite different perspectives for prospective usages.

3.5. References

Aida R, Komano M, Saito M, Nakase K, Murai K (2008) **Chrysanthemum flower shape modification by suppression of chrysanthemum-AGAMOUS gene.** Plant Biotechnology 25:55-59

Akita Y, Horikawa Y, Kanno A (2008) **Comparative analysis of floral MADS-box genes between wild-type and a putative homeotic mutant in lily.** Journal of Horticultural Science & Biotechnology 83(4):453-461

Albani MC, Battey NH, Wilkinson MJ (2004) **The development of ISSR-derived SCAR markers around the SEASONAL FLOWERING LOCUS (SFL) in *Fragaria vesca***. Theoretical and Applied Genetics 109:571-579

Angenent GC, Franken J , Busscher M, van Dijken A, van Went JL, Dons HJM, van Tunen AJ (1995) **A novel class of MASD box genes is involved in ovule development in petunia.** Plant Cell 7, 1569-1582

Bateman RM, Hilton J, Rudall PJ (2006) **Morphological and molecularphylogenetic context of the angiosperms: constrasting the 'top-down' and 'bottom-up' approaches used to infer the likely characteristics of the first flowers.** Journal of Experimental Botany 57:3471-3503

Baum DA, Hileman LC (2006) **A developmental genetic model for the origin of the flower.** In: Ainsworth C: Flowering and its manipulation. Sheffield, UK: Blackwell Publishing

Benecke M (1997) **DNA typing in forensic medicine and in criminal investigations: a current survey.** Naturwissenschaften 84:181-188

Bowman JL (1997) **Evolutionary conservation of angiosperm flower development at the molecular and genetic levels.** Journal of Biosciences 22(4):515-527

Buzgo M, Soltis DE, Soltis PS (2004) **Floral developmental morphology of *Amborella trichopoda* (*Amborellaceae*).** International Journal of Plant Sciences 165:925-947

Buzgo M, Soltis PS, Kim S, Soltis DE (2005) **The making of the flower.** Biologist 52:149-154

Chandler S, Tanaka Y (2007) **Genetic modification in floriculture.** Critical reviews in Plant Science 26:169-197

Chandler SF (2007) **Practical lessons in the commercialization of genetically modified plants - long vase-life carnation.** In: Proceedings of the International Symposium on Plant Biotechnology: from bench to commercialization. 764:71-81

Coen ES, Meyerowitz EM (1991) **The war of the whorls: genetic interactions controlling flower development.** Nature 353:31-37

Colombo L, Franken J, Koetje E, van Went J, Dons HJ (1995) **The *Petunia* MADS-box gene *FBP11* determines ovule identity.** Plant Cell 7:1859-1868

CPV.4937 (2007) **A study on the possible use of isoenzyme characteristics for the assessment of distinctness, uniformity and stability in *Calluna vulgaris* L..** Final Report, Agreement number CPV.4937

Davis FS, Albrigo IG (1994) **Citrus**. CAB International, Wallingford, UK

de Riek J, Dendauw J, Leus L, de Loose M, van Bockstaele E (2000) **Variety protection by use of molecular markers: some case studies.** Plant Biosystems 135(1):107-113

Debener T, Matthiesch L (1999) **Construction of a genetic linkage map for roses using RAPD and AFLP markers.** Theoretical and Applied Genetics 99:891-899

Debener T, Winkelmann T (2009) **Ornamentals**. In: Kempken F, Jung C: Transgenic Plants in Breeding and Crop Production. Springer-Verlag, Berlin, in press.

Ditta G, Pinyopich A, Robles P, Pelaz S, Yanofsky MF (2004) **The *SEP4* gene of *Arabidopsis thaliana* functions in floral organ and meristem identity.** Current Biology 14:1935-1940

Dixon GR, Dutton KJ (1987) **Manipulation of flowering in white heather.** Acta Horticulturae 205:219-224

Dornelas MC, Neves Do Amaral WA, Rodriguez APM (2004) ***egLFY,* the *Eucalyptus grandis* homolog of the Arabidopsis gene *LEAFY* is expressed in reproductive and vegetative tissues**. Brasilian Journal of Plant Physiology 16:105-114

Dunemann F, Kahnau R, Stange I (1999) **Analysis of complex leaf and flower characters in *Rhododendron* using a molecular linkage map**. Theoretical and Applied Genetics 98:1146-1155

Eeuwijk FA, Law JR (2004) **Statistical aspects of essential derivation, with illustrations based on lettuce and barley.** Euphytica 137:129-137.

Endo T, Shimada T, Fujii H, Kobayashi Y, Araki T, Omura M (2005) **Ectopic expression of an FT homolog from Citrus confers an early flowering phenotype on trifoliage orange (*Poncirus trifoliata* L. Raf.).** Transgenic Research 14:703-712

Foucher F, Chevalier M, Corre C, Soufflet-Freslon V, Legeai F, Hibrand-Seint Oyant L (2008) **New resources for studying the rose flowering process**. Genome 51:827-837

Gebhardt K, Friedrich M (1987) **Micropropagation of *Calluna vulgaris* cv. Beale, H.E.** Plant Cell Tissue and Organ Culture 9(2):137-145

GenDG (2009): law draft of the German Federal Government of October 13, 2008; enacted April 24, 2009; published as **Gendiagnostikgesetz** - GenDG, printing matter 16/10532

Geuten K, Becker A, Kaufmann K, Caris P, Janssens S, Viaene T, Theissen G, Smets E (2006) **Petaloidy and petal identity MADS-box genes in the balsaminoid *Impatiens* and *Marcgravia*.** The Plant Journal 47:501-518

Goto K, Kyozuka J, Bowman JL (2001) **Turning floral organs into leaves, leaves into floral organs.** Current Opinion in Genetics & Development 11:449-456

Haley SD, Afanador L, Kelly (1994) **Selection for monogenic pest resistance traits with coupling- and repulsion-phase RAPD markers**. Crop Science 34:1061-1066

Handa T, Kita K, Wongsawad P, Kurashige Y, Yukawa T (2006) **Molecular phylogeny-assisted breeding of ornamentals.** Journal of Crop Improvement 17:51-68

Heckenberger M (2004) **Identification of essentially derived varieties in maize (*Zea mays* L.) using molecular markers, morphological traits, and heterosis.** Dissertation, University of Hohenheim.

Heckenberger M (2004). **Identification of essentially derived varieties in maize (*Zea mays* L.) using molecular markers, morphological traits, and heterosis.** Ph.D. thesis, University of Hohenheim

Hernandez P, Dorado G, Ramirez MC, Laurie DA, Snape JW, Martin A (2003) **Development of cost-effective *Hordeum chilense* DNA markers: molecular aids for marker–assisted cereal breeding.** Hereditas 138:54-58

Judd WS, Campbell CS, Kellog EA, Donoghue MJ (2002) **Plant Systematics: a phylogenetic approach**. Sinauer, Sunderland, Mass.

Katsumoto Y, Fukuchi-Mizutani M, Fukui Y, Brugliera F, Holton TA, Karan M, Nakamura N, Yonekura-Sakakibara K, Togami J, Pigeaire A, Tao GQ, Nehra NS, Lu CY, Dyson BK, Tsuda S, Ashikari T, Kusumi T, Mason JG, Tanaka Y (2007) **Engineering of the Rose Flavonoid Biosynthetic Pathway**

Successfully Generated Blue-Hued Flowers Accumulating Delphinidin. Plant Cell Physiol. 48(11):1589-1600

Kim S, Koh J, Yoo MJ, Kong H, Hu Y, Ma H, Soltis PS, Soltis DE (2005) **Expression of floral MADS-box genes in basal angiosperms: implications for the evolution of floral regulators**. Plant Journal 43:724-744

Klekowski EJ Jr. (1988) **Mutation, developmental selection, and plant evolution**. Columbia University Press, New York.

Klingeman W, Babbit B, Hall C (2006) **Master gardener perception of genetically modified ornamental plants provides strategies for promoting research products through outreach and marketing**. HortScience 39:1327-1331

Kotoda N, Wada M, Komori S, Kidou S, Abe K, Masuda T, Soejima J (2000) **Expression pattern of homologues of floral meristem identity genes *LFY* and *AP1* during flower development in apple**. Journal of the American Society for Horticultural Science 125:398-403

Kramer EM, Di Stilio VS, Schluter PM (2003) **Complex patterns of gene duplication in the APETALA3 and PISTILLATA lineages of the *Ranunculaceae*.** International Journal of Plant Sciences 164:1-11

Kramer EM, Dorit RL, Irish VF (1998) **Molecular Evolution of Genes Controlling Petal and Stamen Development: Duplication and Divergence Within the APETALA3 and PISTILLATA MADS-Box Gene Lineages.** Genetics 149:765-783

Krizek BA, Fletcher JA (2005) **Molecular mechanisms of flower development: an armchair guide.** Nature Genetic Reviews 6:688-698

Kumar LS (1999) **DNA markers in plant improvement: an overview**. Biotechnology Advances 17:143-182

Kwolek AVA, Woolhouse HW (1982) **Studies on the Dormancy of *Calluna vulgaris* (L.) Hull, during Winter: The Effect of Photoperiod and Temperature on the Induction of Dormancy and the Annual Cycle of Development.** Annals of Botany 49:367-376

LeBuanec B (2004) **Protection of Plant-Related Innovations: Evolution and Current Discussion.** IP Strategy Today 9:1-18

Lu CY, Chandler SF, Madon JG Brugleria F (2003) **Florigene flowers: from laboratory to market.** In: Plant Biotechnology and beyond. Pp.333-336. Ed.: Vasil IK, Kluwer Academic Publishers, Dordrecht

Mahy G, Ennos RA, Jacquemart A-L (1999) **Allozyme variation and genetic structure of *Calluna vulgaris* (heather) populations in Scotland: the effect of postglacial recolonization**. Heredity 82:654-660

Mahy G, Jacquemart A-L (1998) **Mating systems of *Calluna vulgaris*: self-sterility and outcrossing estimations.** Canadian Journal of Botany 76:37-42

Mahy G, Vekemans X, Jacquemart A-L, de Sloover JR (1997) **Allozyme diversity and genetic structure in South-Western population of heather, *Calluna vulgaris*.** New Phytologist 137:325-334

Malek B von, Weber WE, Debener T (2000) **Identification of molecular markers linked to *Rdr1*, a gene conferring resistance to blackspot in roses.** Theoretical and Applied Genetics 101:977-983

Matsumoto S, Fukui H (2007) **ABCDE model for wild rose (*Rosa rugosa* Thunb.Ex Murray) floral development.** Acta Horticulturae 751:369-373

Meikle A, Paterson S, Finch RP, Marshall G, Waterhouse A (1999) **Genetic characterization of heather (*Calluna vulgaris* (L.) Hull) subject to different managment regimes across Great Britain**. Molecular Ecology 8:2037-2047

Michelmore RW, Paran I, Kesseli RV (1991) **Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations.** Proceedings of the National Academy of Sciences USA 88:9828-9832

Mondragón-Palomino M, Theissen G (2008) **MADS about the evolution of orchid flowers**. Trends in Plant Science 13:51-59

Mondragón-Palomino M, Theissen G (2009) **Why are orchid flowers so diverse?** Reduction of evolutionary constraints by paralogues of class B floral homeotic genes. Annals of Botany 1-12

Orita M, Iwahana H, Kanazawa H, Hayashi K, Sekiya T (1989) **Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphism**. Proceedings of the National Academy of Sciences 86:2766-2770

Paran I, Michelmore RW (1993) **Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce.** Theoretical and Applied Genetics 85:985-993

Pelaz S, Ditta GS, Baumann E, Wisman E, Yanofsky MF (2000) **B and C floral organ identity functions require SEPALLATA MADS-box genes**. Nature 405:200-203

Pelaz S, Gustafson-Brown C, Kohalmi SE, Crosby WL, Yanofsky MF (2001a) ***APETALA1* and *SEPALLATA3* interact to promote flower development.** Plant Journal 26:385-394

Pelaz S, Tapia-Lopez R, Alvarez-Buylla ER, Yanofsky MF (2001b) **Conversion of leaves into petals in *Arabidopsis***. Current Biology 11:182-184

Rendell S, Ennos RA (2002) **Chloroplast DNA diversity in *Calluna vulgaris* (heather) populations in Europe.** Molecular Ecology 11:69-78

Rout GR, Mohapatra A (2006) **Use of molecular markers in ornamental plants: a critical reappraisal.** European Journal of Horticultural Science 71(2):53-68

Scheef EA, Casler MD, Jung G (2003) **Development of species-specific SCAR markers in bentgrass.** Crop Science 43:345-349
Schum A (2003) **Mutation breeding in ornamentals: an efficient breeding method**. Acta Horticulturae 612:47-60

Schum A (2003) **Mutation breeding in ornamentals: an efficient breeding method?** Acta Horticulturae: 612:47-60

Scovel G, Ovadis M, Vainstein A, Reuven M, Ben-Yephet Y (2001) **Marker assisted selection for resistance to *Fusarium oxysporum* in the greenhouse carnation.** Acta Horticulturae 552:151-156

Scovel G, Ben-Meir H, Ovadis M, Itzhaki H, Vainstein A (1998) **RAPD and RFLP markers tightly linked to the locus controlling carnation (*Dianthus caryophyllus*) flower type**. Theoretical and Applied Genetics 6(1): 117-122

Smyth DR (2001) **Flower development**. Current Biology 11:82-84

Soltis PS, Soltis DE, Kim S, Chanderbali A, Buzgo M (2006) **Expression of floral regulators in basal angiosperms and the origin and evolution of ABC-function**. In: Advances in Botanical Research - Developmental genetics of the flower. Eds.: Soltis DE, Leebens-Mack JH, Soltis PS. Elsevier Ltd.

Staub JE, Gabert A, Wehner TC (1996) **Plant Variety Protection: a consideration of genetic relationship.** HortScience 31:1086-1091

Staub JE, Meglic V (1993) **Molecular genetic markers and their legal relevance for cultivar discrimination: a case study in cucumber**. HortTechnology 3(3):291-300

Stevens PF (2001 onwards) **Angiosperm Phylogeny Website**. Version 9, June 2008 http://www.mobot.org/MOBOT/research/APweb/

Tan CH, Swain SM (2006) **Genetics of flower initiation and development in annual and perennial plants.** Physiologia Plantarum 128:8-17

Tanaka Y, Katsumoto Y, Brugliera F, Mason J (2005) **Genetic engineering in floriculture**. Plant Cell, Tissue and Organ Culture 80:1-24

Teeri TH, Kotilainen M, Uimari A, Ruokolainen S, Peng NGY, Malm U, Pöllänen E, Broholm S, Laitinen R, Elomaa P, Albert VA (2006) **Flower developmental genetics of *Gerbera (Asteraceae).*** Advances in Botanical Research 44:323-351

Theissen G (2001) **Development of floral organ identity: stories from the MADS house**. Current Opinion in Plant Biology 4:75-85

Theissen G, Kim J, Saedler H (1996) **Classification and phylogeny of the MADS-box multigene family suggest defined roles of MADS-box gene subfamilies in the morphological evolution of eukaryotes**. Journal of Molecular Evolution 43:484-516

Theissen G, Melzer R (2007) **Molecular mechanisms underlying origin and diversification of the Angiosperm flower.** Annals of Botany 1-17

Urbanus SL, de Folter S, Shchennikova AV, Kaufmann K, Immink RGH, Angenent GC (2009) **In planta localisation patterns of MADS domain proteins during floral development in *Arabidopsis thaliana*.** BMC Plant Biology 9:5

van Tunen AJ, Eikelboom W, Angenent GC (1993) **Floral organogenesis in *Tulipa***. Flowering Newsletter 16:33-37

Vosman B, Visser D, Voort JR van der, Smulders MJM, van Eeuwijk F (2004) **The establishment of 'essential derivation' among rose varieties, using AFLP.** Theoretical and Applied Genetics 109:1718-1725

Wehling P (1998) **Nutzbarmachung biochemischer und molekularer Markermethoden für die praktische Züchtung.** In: Schriften zu Genetischen Ressourcen, Band 08: Züchterische Nutzung pflanzengenetischer Ressourcen - Ergebnisse und Forschungsbedarf.

Wilkie JD, Sedgley M, Olesen T (2008) **Regulation of floral initiation in horticultural trees.** Journal of Experimental Botany 59(12):3215-3228

Wolfe AD, Liston A (1998) **Contributions of PCR-based methods to plant systematics and evolutionary biology.** In: Soltis DE, Soltis PS, Doyle JJ: Molecular systematics of plants II – DNA sequencing. Kluwer Academic Publishers, Boston-London

Yang HY, Kruger J (1994) **Identification of a RAPD marker linked to the *Vf* gene for scab resistance in apples.** Euphytica 77:83-87

Yu DY, Kotilainen M, Pollanen E, Mehto M, Elomaa P, Helariutta Y, Albert VA, Teeri TH (1999) **Organ identity genes and modified patterns of flower development in *Gerbera hybrida* (*Asteraceae*).** Plant Journal 17:51-62

Zahn LM, Leebens-Mack J, dePamphilis CW, Ma H, Theissen G (2005) **To B or Not to B a Flower: The Role of DEFICIENS and GLOBOSA Orthologs in the Evolution of the Angiosperms**. Journal of Heredity 96(3):225–240

Zhang Y, Stommel JR (2001) **Development of SCAR and CAPS markers linked to the Beta gene in tomato.** Crop Science 41:1602-1608

Zik M, Irish VF (2003) **Flower development: initiation, differentiation, and diversification**. Annual Review of Cell and Developmental Biology 19:119-140

4. Appendix

4.1. List of Figures

The cover picture is reproduced by courtesy of Peter de Winkel (Heidepflanzen de Winkel, Goch, http://www.dewinkel.de).

4.2. List of Tables

4.3. Acknowledgments

First and primarily, I wish to thank Dr. Annette Hohe for leaving this fascinating project to me, her comprehensive guidance and excellent scientific support throughout this investigation and her endless patience regarding manifold manuscript corrections. My thanks are extended to my doctoral advisor, Prof. Dr. Thomas Debener for his scientific backup and support in respect of my external Ph.D. study. Additionally, I would like to thank Prof. Dr. Traud Winkelmann for the preceding and critical review of this thesis as well as for her subsidiary efforts during the project execution.

Primarily, I am very much obliged to Katja and Jörg Krüger (IGZ) for their skilled design and realization of the daily laboratory work and any organizational matters regarding my everyday life in Erfurt. Additional thanks go to Claudia Hönemann, Anke Müller, Yvonne Klopotek and Luisa Hiese for valuable scientific discussions, proofreading of this thesis, their laboratory assistance and any non-scientific courtesy.
Furthermore, I would like to thank any unmentioned workers and employees of the IGZ who participated in or supported this investigation as well as all students and trainees who contributed to both projects during the last years.

I additionally wish to thank Dr. Marcus Linde, Dr. Anja Biber and Inka Gawenda of the Leibniz-University of Hannover for their competent advices. I owe gratitude to Prof. Dr. Günter Theissen, Dr. Dajana Lobbes, Dr. Mariana Mondragon, Dr. Hans Pohl and Dr. Pia Nutt of the Friedrich-Schiller-University of Jena for their practical and academic support, especially during the methodological setup and the investigations concerning the flower molecular biology.
Special thanks go to the owner of the cooperating company, Peter de Winkel, and his head of breeding Friederike Vogt regarding their encouragement of the project and hence of my work.

Last but foremost, I wish to thank my whole family, but especially Anja who supported me in various and indispensable ways and who knew how to encourage and motivate me during this period.

4.4. Curriculum Vitae

Personal Data	
Name	Thomas Borchert (né Schmidt)
Address	In der Görtzbach 43 35041 Marburg
Birthday, -place	09.01.1980 Friedberg (Hessen), Germany

Education	
1986 – 1990	Primary School Geschwister-Scholl-Schule, Niddatal
1990 – 1999	Secondary School Augustinergymnasium, Friedberg
1999 – 2000	Basic military service Rotenburg an der Fulda
2001 – 2005	University of Applied Sciences, Gießen Specialization: Biotechnology, Dipl.-Ing. (FH)
2005	Diploma thesis: „Vergleichende Untersuchung der Zellteilungsaktivität und Morphologie zygotischer und somatischer Embryonen von *Cyclamen persicum* MILL.“

Work Experience	
2004 – 2005	Leibniz-Institute of Vegetable and Ornamental Crops, Erfurt Trainee
2005 – 2007	Heidepflanzen de Winkel, Goch Ph.D. student (AiF project)
since 2007	Leibniz-Institute of Vegetable and Ornamental Crops, Erfurt Ph.D. student (BMELV project)

Awards	
2005	Prize of ADIVK (Arbeitskreis Deutsche In Vitro Kulturen) for the diploma thesis

Manuscripts (submitted)	
2009 (incl.[4])	**Borchert, T.**, Eckardt K, Fuchs J, Krueger K, Hohe, A.: 'Who's who' in different flower types of *Calluna vulgaris* (*Ericaceae*): morphological and molecular analyses of flower organ identity. BMC Plant Biology, submitted: May 26, 2009[5]

Publications (peer-reviewed)	
2009 (incl.)	**Borchert, T.**; Hohe, A.: Identification of molecular markers for the flower type in the ornamental crop *Calluna vulgaris*. Euphytica, in press
2008 (incl.)	**Borchert, T.**; Krüger, J.; Hohe, A.: Implementation of a model for identification of EDVs in vegetatively propagated *Calluna vulgaris* varieties. BMC Genetics 9:56
2007	**Borchert, T.**; Fuchs, J.; Winkelmann, T.; Hohe, A.: Variable DNA content of *Cyclamen persicum* regenerated via somatic embryogenesis: rethinking the concept of long-term callus and suspension cultures. Plant Cell, Tissue and Organ Culture 90 (3), 255-263
2006	**Schmidt, T.**; Ewald, A.; Seyring, M.; Hohe A.: Comparative analysis of cell cycle events in zygotic and somatic embryos of *Cyclamen persicum* indicates strong resemblance of somatic embryos to recalcitrant seeds. Plant Cell Reports 25 (7), 643-650

[4] included within this thesis
[5] accepted and revision requested as at July 10, 2009

Declaration of Authorship / Erklärung zur Dissertation

Hierdurch erkläre ich, dass die vorliegende Dissertation „Grundlegende Untersuchungen zur Genetik, Züchtung und Blütenorganidentität von *Calluna vulgaris* (L.) Hull“ selbstständig verfasst und alle benutzten Hilfsmittel sowie evtl. zur Hilfeleistung herangezogene Institutionen vollständig angegeben wurden.

Die Dissertation wurde nicht schon als Diplom- oder ähnliche Prüfungsarbeit verwendet.

Hannover, den 29.05.2009

[Borchert]