The Thesis of the Ph.D. dissertation

Kinga K Mátyás Petrovicsné

Keszthely

2021

HUNGARIAN UNIVERSITY OF AGRICULTURE AND LIFE SCIENCES (MATE)

FESTETICS DOCTORAL SCHOOL

Comparative analysis of expressed genes in male and female inflorescences of the common ragweed (*Ambrosia artemisiifolia* L.)

THESIS OF Ph.D. DISSERTATION

DOI: 10.54598/001520

Kinga K Mátyás Petrovicsné

KESZTHELY

2021

A doktori iskola

megnevezése: Festetics Doktori Iskola (FDI)

tudományága: Növénytermesztés és Kertészeti Tudományok

Vezetője: Dr. habil Anda Angéla DSc.
egyetemi tanár, MTA doktora
Magyar Agrár- és Élettudományi Egyetem,
Környezettudományi Intézet

Témavezető: Dr. Virág Eszter
Debreceni Egyetem,
Természettudományi Kar,
Biotechnológiai és Mikrobiológiai Tanszék

....

Az iskolavezető jóváhagyása A témavezető jóváhagyása

TABLE OF CONTENT

Introduction

Common ragweed is native to North and Central America (Lorenzi and Jeffery 1987; Kovalev 1989), but now this species is present world-wide; Africa (CJB, 2016), Asia (Flora of China Editorial Committee, 2016), Australia (Council of Heads of Australasian Herbaria, 2016) and Europe (Euro + Med, 2016). The distribution of the common ragweed in Europe covers the area at medium latitude characterized by continental climate started its expansion from two centers southwestern France and southwestern Hungary and meaning a current problem from both agricultural and public health aspects (Makra et al. 2004). Since in the last 20-25 years the number of ragweed pollen has dramatically increased, it would be important to get global attention and find the optimal control against the rapid spread of the plant caused by extraordinary plant-adaptability, a large proportion of derelict land areas, the absence of specific pest insects and diseases, and not least the development of high-level resistance against herbicides. Considering the invasion migratory pathways, there is a high risk for the introduction of herbicide-resistant genotypes from Eastern to Western Europe.

Understanding the genetic control of sex determination during the plant life cycle may contribute to finding an ecologically safer strategy in the common ragweed control. Flower development is a complex and accurately coordinated biological and morphological process consisting of spatial regulation of a considerable number of organ-specific genes during the life cycle of higher plants (Taiz et al. 2015). Investigated genes discussed in this study were chosen based on the genetic knowledge of *Arabidopsis* flowering (Chandler 2011, Irish 2010). Based on *Arabidopsis* transcriptome analysis the majority of expressed transcripts were found in the reproductive part of flowers instead of the

perianth,5 reflecting the more complex anatomy of tissue and cell types within stamens and carpels and major developmental events such as ovule and pollen formation (Chandler 2011). Members of the *Asteraceae* family are characterized by a more complex and modified floral formula compared to *Arabidopsis thaliana. Asteraceae* are characterized by capitulum inflorescence and this flower head can be monogamous or heterogamous (Harris 1999). The inflorescences are protected by bracts, and their calyx reduced to pappus, scales, or coronula, have a corolla, stamens are connate and their cypsela fruit type is achene (Katinas et al. 2016). Therefore, we used the ABC(E) gene sequences of different species from the *Asteraceae* family (*Gerbera sp.; Helianthus sp.; Tagetes sp.; Chrysanthemum sp.*) to determine the flower organ identify genes (FOIGs). Due to the adverse effects on human health and weed control technique - mentioned above - it is of paramount importance to identify genes and pathways that regulate the development of flower organs and are responsible for pollen production and seed mass.

Next generation sequencing (NGS) is an advantaged technology to capture the diversity of differentially expressed transcripts in male and female flowers. In the present work, we report Illumina RNA-sequencing of two developmental stages of female flower compared with male and leaf transcriptomic data. Differences between expression levels of the above-described flowering pathways and two gene categories such as MIGs and FOIGs (ABC(E) ortholog - MADS box genes) identified in each type of flowers are also discussed. The genetic events of staminate and pistillate inflorescence formation were investigated through expression analysis in wild-growing and *in vitro* cultivated plants.

Objectives of the study

The aim of the present work is to identify genes that control the common ragweed flowering and their specificity to male and female inflorescence formation.

The aims can be summarized as:

1. Create transcriptome libraries from *A. artemisiifolia* L., early and late developmental stages of male and female inflorescences.

2. Identify the coding sequences and the expression level of flowering related genes of *A. artemisiifolia* L. inflorescences: Meristem Identity Genes (MIGs), Floral Organ Identity Genes (FOIGs), and Transcription factors (TF). Validation of tissue specific genes with RT-qPCR technique.

3. Determine the coding sequences and the expression level of the specific floral organ (*ABC*) genes during floral development.

4. Screening and identify of gender specific genes in male and female inflorescences and determine their expression level.

5. Determine the biochemical pathways that regulate male and female flower formation.

Materials and methods

Prior to genetic studies, the male and female inflorescences of the common ragweed were have dissected out, examined by digital microscopy using the stereomicroscopic focus stacking technique, and individual flower organs were identified. The RNA-seq experiments were based on common ragweed samples from four bulked individuals. Seeds were collected from Cserszegtomaj, Zala county, Hungary (GPS: 46.79528, 17.26005) and were surface sterilized for 1 min in 70% (v/v) ethanol and washed with tap water for 23-30 min, followed by dipping in a 7% solution of calcium hypochlorite with TWEEN 20 for 20 min, and rinsing three times with sterile distilled water. The sterilized seeds were then germinated on MS medium (Murashige and Skoog, 1962) with 0.3 mg/l metatopolin hormon supplemented with 2% sucrose/litre and solidified with 0.8% phytoagar. The pH was adjusted to 5.8 prior to autoclaving at 120 °C for 20 min. All the cultures were maintained at 25 °C under a 16/8 light/dark photoperiodic regime with a light intensity of about 2400 to 2800 lux. Node segments (1.5 -2 cm long) were excised from 21- day-old plants and transferred to the above described MS medium.

Nine developmental stages of female flowers were collected from 12 *in vitro* cultivated individuals of common ragweed. For RNA-seq two categories of nine phenological phases (1F and 2F) were established and combined such as early 1F and late 2F developmental stages. Male flower buds were collected from genotypes used in F, M, L transcriptomes (Virág et al. 2016). For gene expression comparisons we used 1M and 2M libraries. Along with RT-qPCR analysis, we used four additionally wild-growing genotypes (collected in Keszthely, Zala county, Hungary, GPS: 46.7654716, 17.2479554). Samples were frozen in liquid nitrogen immediately after collection. RNA extraction and on-column DNase digestion were performed from each category using TaKaRa

Plant RNA Extraction Kit (Takara Bio Inc; Japan) following the manufacturer's instruction. RNA integrity and quantity were assessed on an Agilent 2100 Bioanalyzer (Agilent Technologies; USA).

Enrichment of mRNA, cDNA synthesis, and library preparation for Illumina NextSeq paired-end sequencing were carried out using TruSeq™ RNA sample preparation kit (Low-Throughput protocol) using an oligo$(dT)_{18}$. The RNA-seq was performed using the Illumina NextSeq5000 system. Sequence reads of male inflorescence (M), leaves (L), and female (F) inflorescence containing all nine developmental stages were used from earlier reported and deposited data in National Centre for Biotechnology Information's (NCBI) Sequence Read Archive (SRA) under the accession SRP08007 (Virág et al. 2016). During preprocessing, raw reads quality was examined using the FastQC quality control software (Andrews 2010). Based on FastQC report, sequences found to be represented more than 0.1% of the total and low-quality bases (corresponding to a 0.1% sequencing error rate) were removed and trimmed using a self-developed application 'GenoUtils' written in Visual Studio integrated development environment in C#. Conversion of .fastq to .fasta files was also performed using this in-house application. A sample-specific multiple assemblies of cleaned reads from seven libraries (M, L, F, 1F, 2F, 1M, and 2M) were performed. *De novo* assembly of each sample type was performed by using Trinity (Haas et al. 2013). For reference, guided alignments Bowtie2 short read aligner (Langmead et al. 2009) was applied. First, a combined read set was assembled from the three sample libraries (M, F, and L) to generate a 'reference' *de novo* transcriptome assembly that was deposited in the NCBI TSA database under the accession GEZL00000000 and reported in our previous study (Virág et al. 2016). During this process, reads with a certain length of overlap were combined to form contigs (kmer size, K, = 25). This combined dataset was used as a reference shotgun assembly in the further alignments. In order to perform

sample-specific expression analysis, we screened the differentially expressed sequences by aligning the original sample reads to the reference followed by abundance estimation using RSEM (Li and Dewey 2011).

The resulting differentially expressed transcripts were further clustered according to their expression patterns by applying Microsoft SQL Server Management Studio. Protein coding regions were extracted from the reference assembly using TransDecoder and further characterized according to likely functions based on sequence homology or domain content using BLAST+ (Altschul et al. 1990). Separately, *de novo* assemblies of 1F, 2F, 1M, and 2M libraries were performed using Trinity and deposited in the NCBI TSA database under the accession GFWB00000000 (1F) and GFWS00000000 (2F). Using these reference contig sets 1F and 2F unique transcripts were realigned applying BLASTn with E value less than 10^{-5}. The resulting narrowed and specified sequences were annotated using NCBI nr protein database. Normalized expression of selected genes was performed digitally and expression values were compared. The obtained reads from each library were mapped to the selected gene CDSs by using Bowtie2. The mapped reads were used to estimate the transcriptome level by the reads per kilobase per million mapped reads (RPKM) method (Mortazavi et al. 2008). For automatic calculation of RPKM values of genes, a self-developed pipeline as module of GenoUtils were applied.

Reverse transcription was performed starting from 1 µg total RNA using Maxima H Minus First Strand cDNA Synthesis Kit with dsDNase (Thermo Scientific) according to the manufacturer's protocol, using an oligo$(dT)_{18}$ and random hexamer primers (Thermo Scientific), the final volume was 20 µl. cDNA (1 µl) was used for real-time PCR amplification on a Bio-Rad CFX96 System. qPCR analysis (and efficiency) was performed with 1µl of cDNA on a Bio-Rad CFX96 System using Xceed qPCR SG Mix (Institute of Applied Biotechnologies). The relative gene expression was calculated with ΔΔCt

method using Bio-Rad CFX Manager™ Software v3.1. Primer efficiency was analyzed with CFX Manager™ Software v3.1 (Bio-Rad). PCR was performed as follows: an initial activation of the polymerase enzyme at 95 °C for 2 min was followed by 46 cycles at 95 °C 5 sec, 59 °C for 30 sec, ended with a 5 sec melt analysis ranging between 65-95 °C with 0.5 increments. Primers were designed with Primer3 (Koressaar and Remm 2007, Untergrasser et al. 2012) based on *in silico* sequence prediction of CDS. Gene expression analysis was established based on three technical and biological replicatesand normalized with the reference gene glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) (Hodgins et al. 2013).

Potential coding regions within reconstructed transcripts that were insufficiently represented by detectable homologies to known proteins were predicted based on metrics tied to sequence composition by applying TransDecoder include with Trinity. Running this application on the Trinity-reconstructed transcripts the candidate protein-coding regions could be identified based on nucleotide composition, open reading frame (ORF) length, and (optional) Pfam domain content. To predict their functions the latest non-redundant (Nr) protein database 10/03/2016 (ftp://ftp.ncbi.nlm.nih.gov/blast/db/FASTA/nr.gz) were used. For further annotation of unigenes using various bioinformatics approaches, the unigenes were firstly searched against the non-redundant database and the Swiss-Prot protein database using local BLASTx (Gish and States 1993) with an E value cutoff of 10-5. With Nr annotation, Blast2GO (Götz et al. 2008) was used to get GO annotation according to molecular function, biological process, and cellular component ontologies http://www.geneontology.org.

Results and Discussion

In this study, we present an RNA-seq approach to get closer to the mechanism of floral development in *A. artemisiifolia* using Illumina sequencing. For 1F, 2F, 1M, 2M libraries early and late categories of phenophases were collected and sequenced *de novo* from *in vitro* and wild-growing plants. This work and related NGS project are deposited in the NCBI Bioproject under the ID PRJNA335689. A total of 39,664,366 (1F) and 37,127,852 (2F) 2*80 bp reads were generated by the Illumina NextSeq500 system. In total of 92% were high quality sequences from raw data: numerically 36,491,216 (1F) and 34,157,623 (2F). These high quality, processed paired-end reads were used to assemble into contigs and transcripts. Their *de novo* assembly resulted in 109,452 (1F) and 97,239 (2F) contigs.

The identification of differences between the male and female floral transcriptomes was based on *in silico* analysis. We performed a local blast alignment by using the determined 80 CDS fasta sequences (MK098047-MK088126) against male (M) and female (F) transcriptomes. No qualitative differences were found among the sequences of commonly expressed genes in two floral organs. The observed expression differences were validated by investigating the transcript expression levels of homolog genes responsible for pollen and embryo formation in model organism *A. thaliana*. The reliability of the male transcriptome was characterized by overexpression of the homologs of *AMS, LAP3, LAP5, LAP6, MSE1, PME, PG1, PG2*. The female transcriptome validity was justified by overexpression of *YAB1, YAB4,* and *ANT* transcription factors. Additionally, we introduced *SUP, MYB5, MYB61, TCP12* as female candidate genes based on our filtering by Microsoft SQL Server Management Studio. Based on our results, it can be identified that our libraries are tissue-specific.

Functional and structural analysis of *A. thaliana* gene homologs were investigated to elucidate floral regulation in the common ragweed. No interaction was found between *LFY, FRI,* and *FLC* transcription factors that determine the initiating steps during vegetative growth. Overexpression of photoperiodic pathway related transcripts of *PHYB* in 1F, 2F, 1M, and 2M libraries and *COP1, CO, and FT* in the L library were found. In Arabidopsis, the expression of *PHYB* affects indirectly *FT* and is limited to the leaf, so for common ragweed, the expression of *PHYB* in flowers is interesting and unusual data. This may suggest that *PHYB* could also regulate flower development steps that occur after the FT switch. The vernalization pathway gene homolog *VRN1*, showed high values in all developmental stages (1F, 2F, 1M, 2M), which function is to inhibit the flowering repressor *FLC* (Sung and Amasino 2004). We found *SOC1* homolog only in 1F and 2F libraries. *SOC1* can integrate flowering signals from different pathways and activat the *APETALA1 (AP1)* and *LEAFY (LFY)* flower meristem identity genes. The lack of *SOC1, LFY* (RPKM=0), and overexpression of *FT* (RPKM=1487) in wild-growing samples indicated that flowering transition was regulated more intensively by the photoperiodic pathway in these samples. In contrast, hormonal pathway related genes *(SOC1, GID1A, GAI, Ga2ox8, CAL)* were more expressed in *in vitro* plants. Therefore we concluded, that both cultivation conditions reached the *AP1* expression that is the "first pass" of floral architecture formation, however, the vegetative pathway was altered under *in vitro* conditions. We observed that morphogenesis of each flower type occurs by two different regulatory pathways, the development of male flowers through the photoperiodic pathway by the FT / FD complex, while that of female flowers under the hormonal pathway is regulated by the AGL24 / SOC1 complex.

The ABC(E) genes responsible for the formation of flower organs were determined based on *Compositae* species (*Gerbera sp, Helianthus sp, Tagetes*

sp, *Chrysanthemum sp*) found in public databases (NCBI, UniProt). The coding sequences and their expression patterns of the following thirteen gene homologs were determined: *AP1, AP2, AP3 1, AP3 2, AP3 3, PI, AG, SEP1, SEP2 1, SEP2 2, SEP3 1, SEP3 2 and SEP4*. The *AP2, AP3 2,* and *PI* genes were not or very low expressed in the female samples, therefore we considered them as male specific genes of the common ragweed. *PI* with *AG* determine the stamen and *AP2* with *AP3* determine the petal formation that are male flower elements in the common ragweed.

Transcription factors (TFs) play important roles in plant development and flower morphogenesis responding to the environment, based on homolog sequences, the coding sequences of 11 TFs genes in common ragweed were determined. Based on RPKM values, homologs of the *CONSTANS-LIKE* zink-finger TF family were found to be characteristic in both female and male samples (*COL4*, *COL5*). Upregulation of *COL4* was found in 1F and 2F libraries indicating a stronger interaction of female tissues with the circadian clock and light signals than in male flowers. Downregulation of *COL9* was observed in all wild-growing samples. We observed the higher *COL5* expression in female flowers, mainly in the case of *in vitro* plants where pistillate flowering showed intensive growth under short day conditions. The next most characteristic transcript was an *ILR3* homolog of *IAA-LEUCINE RESISTANT 3* protein representing the helix-loop-helix protein family. Since this transcript was found to be expressed roughly equally in all investigated samples (1F, 2F, 1M, 2M, M, and L), we assumed that to keep at appropriate level the iron homeostasis is essential to all flower types, in fact, to maintain hormonal and redox balance for fertility (Sudre et al. 2013).

The auxin efflux carrier homolog *PIN1* was also found to be upregulated in early female flowers (1F) indicating the female tissue determination might be under strong control of auxin. In the ethylene response pathway, the protein

EIN3 regulates ethylene-responsive genes as a positive regulator and results in flowering delay via repression of the floral meristem-identity genes *LFY* and *SOC1* (Achard et al. 2007). In our flowering system, it was upregulated in M, 1M, 2M libraries resulting no expression of *SOC1* and *LFY* in male tissues. NAC-domain containing transcription factors are essential for normal plant morphogenesis representatives of which are the *CUP-SHAPED COTYLEDON* proteins encoded by *CUC1* and *CUC2* genes less expressed in our flower samples. The homeodomain like *MYB* transcription factors are involved in the control of the cell cycle of plants. The gibberellin pathway related *MYB33* homolog was expressed only in male flower in wild-growing *Ambrosia* samples in agreement with Rocheta (2014), and Millar (2005) in which studies the *GAMYB-like MYB33* was proved to facilitate anther development redundantly (Millar and Gubler 2005, Rocheta et al. 2014). In *in vitro* originated 1F and 2F libraries homolog transcripts of *MYB33* were also found indicating repeatedly altered hormonal regulation of this system.

In order to identify exclusively expressed transcripts in the sex-specific floral tissues, we used Bowtie2 alignments of F, M, L, 1F, and 2F libraries for the *Ambrosia* reference transcriptome GEZL00000000. The transcript datasets were extracted from Trinotate heat map and post-processed using Microsoft SQL Server Management Studio. The queries for unique organ-specific sequences resulted in 5659 (M), 1691 (F), and 4267 (L). The number of exclusive and shared transcripts in wild-growing male and female flowers and leaf are visualized in Venn diagram. For visualiszation we used FunRich (V3) software (Pathan et al. 2015). The number of annotable unique transcripts was 10507 in M, F and L libraries. Functional annotation of the transcriptome sequences and analysis of the annotation of three-specific RNA pools (M, F, L) were performed using BLASTx search (Gish and States 1993) and Blast2GO (Götz et al. 2008, Conesa et al. 2005).

In order to identify exclusively expressed genes in female flowers, the combined assembly of five (M, L, F, 1F, and 2F) non-normalized libraries were analysed. In the analysis, 1691 unique transcripts were found in the wild growing F library, representing the nine developmental stages. Transcripts longer than 700 bp were further investigated. After this filtering, we found 60 transcripts showing evaluable ORF, among which nine genes were well annotable using NCBI nr and Swiss-Prot databases. Based on this, the following protein coding and transcription factors were identified: *Protein kinase 1A (PBL9), SUP, Alpha carbonic anhydrase-7 (ACA7), Laccase-2 (LAC2), TCP12, MYB5, MIB61, Zinc finger protein WIP2,* and *Protein Brassinosteroid enhanced expression 1 (BEE1).* The uniquely expressed transcript number in male tissues was 4549 of which sequences longer than 700 bp were filtered out. In this way, total coding sequences of 41 genes were identified based on NCBI nt Blast and NCBI ORF finder databases. The 12 most characteristic male specific genes, *PCC13-62, CYP450 86B1, GDSL2, TET8, OAS, MYB80, MYB26, MYB35, MYB44, PMADS2, NIP, SYN,* their RPKM values, and functional properties are described in this thesis.

For the RT-qPCR validation, we selected the genes showing the largest differences and most interesting from functional aspects. *TCP12, SUP,* and *LAP6* were selected to justify their exclusive expression in female and male flowers. *STIG1* was chosen according to its opposite expression in male flowers. Additionally, we selected also *PI* as ABC(E) transcript exclusively present in staminate tissues. The relative expression value of *PI* gene in male flowers was 8170.66, in leaf 1.59, and in female flowers 3077.59. For the *LAP6* gene, the highest expression value was 11599.13 in male flowers, 0.16 value was in leaf, and 486.28 value was in female flowers. The expression values of the *STIG1* gene were in male 1000.38, in leaf 0.38, and in female tissues 40.9. For the *SUP* gene, a relative expression value of 0.83 was detected in male samples, 0.63 in

leaf samples, and 540.53 in female flowers. The *TCP12* gene was expressed with a relative value of 12.97 in male samples, 3.95 in leaf samples, and 146.44 in female samples. To evaluate the results, genes were grouped according to their specificity. According to this, the female transcriptome was validated based on *TCP12* and *SUP* expressional patterns. Female specificity of *SUP* and *TCP12* was confirmed. The expression of *TCP12* was three times higher than of the *SUP* in female flowers confirming the *in silico* RPKM results. The male transcriptome was validated according to the expression pattern of *LAP6*, *STIG1*, and *PI*. The non-exclusive expression of *PI* was observed indicating the ABC(E) function and B class activity in both flower types. Expression of *STIG1* homolog in male flowers was confirmed by RT-qPCR, indicating the modified pistil origin in the pistillodium structure.

Conclusions

The genetic program of the formation of separated male and female flower development in monoecious plants is a less understood mechanism in flowering biology. Available genomic data of the majority of these species are insufficient or inexistent to date. It makes it more difficult, that the reproduction biology of *A. artemisiifolia* is less studied and the majority of information has been revealed exclusively in the pollen formation as part of male flowering biology. Today, large DNA or RNA data sets of the non-model organisms may be attainable using NGS technology; therefore, we sequenced and compared transcriptome libraries of wild-growing male, female, leaf, *in vitro* female flowers and male flowers representing early and late developmental phases. Comparative studies revealed a subset of transcripts that were differentially expressed in the different libraries known in flower or plant development in *A.*

thaliana. However, genes showing differential expression previously were not characterized in *A. artemisiifolia* during flowering.

Genome-wide transcriptional profiling in seven libraries revealed a high number of transcripts that were differentially expressed playing role - non-exclusively - in plant or flower development, but also in signal transduction, redox, and abiotic stress mechanisms. Induction of floral meristem initiation is preceded by a complex regulation network just in the vegetative phase which regulatory elements are the so-called flowering pathway genes. The expression pattern of components of these pathways is consistent with the short day-induced flowering of common ragweed and suggests that the initiation of flowering depends on PHYB-related light signals.

Because of the small amount of collectable pistillate flowers, plants were also *in vitro* cultivated on meta-topolin supplemented media. In this way, influences of hormonal and photoperiodic pathways were modelled on the flowering pathway genes. Based on gene expression differences of flowering pathways in male and female samples and the shifts to intensive female morphogenesis under *in vitro* cultivation conditions, we concluded that before the generative transition, the determination of floral gender takes place just in the vegetative phase during the vegetative pathway dominancy that defines the subsequent floral organ morphogenesis. Transition to *AP1* expression and FOIGs induction is led by hormonal and photoperiodic pathways, and depending on the relative dominancy of these two routes, initiation of the female or the male flower development is facilitated. Thus, flower identity is decided just in vegetative growth. Investigation of male flower indicated that photoperiodic pathway may induce the generative transition through FT / FD complex; however, gibberellin-related *SOC1* expression was not observed indicating it has no role in male gender formation.

On the other hand, expression of FT / FD was lacking in *in vitro* female samples that was probably due to the action of *SVP* or *AGL24*. Transcripts of *LFY* homolog were not found in any of the libraries, therefore it appears likely that floral organ development is induced through *LMI* - *CAL* regulation in female flowers. However, *FUL* and directly *(AP1)* - *AP2* regulation in male flowers in only induced when the *AP1* repression by *TFL1* was also in effect. Expression analysis of transcription factors indicated that similarly to other species investigated so far, the circadian clock takes a significant part in timing flowering initiation.

Thesis Points

1. Identification of male and female flower organs of *Ambrosia artemisiifolia* L. and determination of their developmental stages. Creation of transcriptome libraries of female and male inflorescences from early and late developmental stages and their biological validation.

2. Determination of the coding sequence of 80 flowering regulatory genes in *Ambrosia artemisiifolia* L. that were published in the NCBI (international) database. Identification their expression level with *in silico* and experimental method.

3. Uniquely expressed gender-specific genes were determined comparing all of the libraries. 9 female and 14 male specific genes were identified and validated.

4. Thirteen sp-called ABC(E) genes, such as floral architecture formation genes, in *A. artemisiifolia* L. monoecious flowering system were identified. These results showed that the *APETALA 2*, *APETALA3 2*, and *PISTILLATA* genes were no or slightly expressed in the female inflorescence, oppositely to male inflorescence, therefore so they can be determined male-specific in this species.

5. The development of male flowers occurs through the photoperiodic pathway which is regulated by the FT / FD complex; however, regulation of female flowers depends on the hormonal pathway affecting AGL 24 / SOC1 complex regulation.

Publication List

Publications related to the PhD thesis

Refereed International Journals

Mátyás KK, Hegedűs G, Taller J, Farkas E, Decsi K Kutasy B, Kálmán N, Nagy E, Kolics B, Virág E *Different expression pattern of flowering pathway genes contribute to male or female organ development during floral transition in the monoecious weed Ambrosia artemisiifolia L. (Asteraceae).* PEERJ 7 (2019) doi.org/10.7717/peerj.7421. **IF.2.38**

Virág E, Hegedűs G, Barta E, Nagy E, **Mátyás K**, Kolics B, Taller J *Illumina sequencing of common (short) ragweed (Ambrosia artemisiifolia L.) reproductive organs and leaves*. FRONTIERS IN PLANT SCIENCES data report (2016) doi:10.3389/fpls.2016.01506. **IF: 4.298**

National Publications (In Hungarian)

Mátyás KK, Bódis J, Virág E, Taller J, Pintér Cs *Az ürömlevelű parlagfű (Ambrosia artemisiifolia L.) virágzatának részletes leírása sztereomikroszkópos rétegfotózás használatával.* BOTANIKAI KÖZLEMÉNYEK 107: 1 pp. 103-109., 7 p. (2020) **IF: 0.52**

International Conferences

Mátyás KK, Taller J, Hegedűs G, Kolics B, Farkas E, Nagy E, Parrag T, Farkas Z, Virág E *Identification of floral organ identity genes in the common ragweed (Ambrosia artemisiifolia L)*. Fiatal Biotechnológusok III. Országos Konferenciája (2018), Budapest, Magyarország, 2018. március 28-29., 31p.

Virág E, Nagy E, **Mátyás K**, Kolics B, Kutasy B, Decsi K, Taller J . *Analysis of molecular background of flowering in the common ragweed (Ambrosia artemisiifolia L.)*. In: Pannonian Plant Biotechnology Workshop "Integration fundamental research into the practical agriculture".(2015) Ljubljana, Szlovénia, pp. 33-34.

National Conferences (In Hungarian)

Mátyás KK, Pintér Cs, Bódis J, Virág E, Taller J *Az ürömlevelű parlagfű (Ambrosia artemisiifolia L.) hím és nő virágzatának morfológiai jellemzése rétegfotózással.* MAGYAR GYOMKUTATÁS ÉS TECHNOLÓGIA (2019) 20. évf. 1. sz, 84 p.

Virág E, Hegedűs G, **Mátyás K**, Nagy E, Taller J *Az ürömlevelű parlagfű váltivarú virágzatának kialakulásért felelős gének NGS alapú vizsgálata*. „GENETIKAI MŰHELYEK MAGYARORSZÁGON" XVI. Minikonferencia (2017), Szeged, Magyarország, 2017. szeptember 8. Elektronikus kiadvány, 1 p.

Conference articles (In Hungarian)

Mátyás KK, Taller J, Hegedűs G, Farkas E, Nagy E, Parrag T, Farkas Z, Virág E *Az ürömlevelű parlagfű (Ambrosia artemisiifolia L.) hím és nő virágzatában kifejeződő gének összehasonlító elemzése.* XXIV. Ifjúsági Tudományos Fórum (2018), 2018. május 24, Keszthely, Elektronikus kiadvány, 10 p.

Virág E**, Mátyás KK**, Taller J, Farkas Z, Parrag T, Hegedűs G *Az ürömlevelű parlagfű (Ambrosia artemisiifolia L.) váltivarú virágzatáért felelős genetikai háttér felderítése bioinformatikai elemzésekkel*. XXIV. Ifjúsági Tudományos Fórum (2018), 2018. május 24, Keszthely, Elektronikus kiadvány, 12 p.

Publications not related to the PhD thesis

Refereed International Journals

Kolics É, Specziár A, Taller J, **Mátyás KK,** Kolics B *Lithium chloride outperformed oxalic acid sublimation in a preliminary experiment for Varroa mite control in pre-wintering honey bee colonies,* ACTA VETERINARIA HUNGARICA 68 (4). 370-373. (2021) **IF: 1,019** (2019)

Kolics É, Parrag T, Házi F, Szepesi K, Heltai B, **Mátyás K** Kutasy B; Virág E; Taller J; Orbán L, Kolics B *An Alternative, High Throughput Method to Identify Csd Alleles of the Honey Bee*. INSECTS 11: 8 Paper: 483 (2020) **IF: 2,220** (2019)

Kolics É, **Mátyás K**, Taller J, Specziár A, Kolics B *Contact Effect Contribution to the High Efficiency of Lithium Chloride Against the Mite Parasite of the Honey Bee* INSECTS 11: 6 Paper: 333 (2020) **IF: 2,220** (2019)

Farkas E, Farkas Z, Hegedűs G, Kutasy B, Kolics B, **Mátyás KK**, Parrag T, Solti I, Virág E, Taller J *Identification and expression of pollen allergen transcripts in different organs of the common ragweed (Ambrosia artemisiifolia L.)*. GEORGOKIN FOR AGRICULTURE 23 (2) (2019) 116-130.

Taller J, Decsi K, Farkas E, Nagy E**, Mátyás KK,** Kolics B, Kutasy B, Virág E *De novo Transcriptome Sequencing Based Identification of Amb a 3-like Pollen Allergen in Common Ragweed (Ambrosia artemisiifolia L.).* JOURNAL OF BOTANICAL SCIENCES (2015) 5: 2 pp. 12-16.

Taksonyi P, Kocsis L, **Mátyás K**, Taller J *Investigation into effective control of powdery mildew with strobilurins in viticulture – a case study.* MITTELUNGEN KLOSTERNEUBURG (2013) 63: 1 pp. 30-37. **IF: 0,031**

Taksonyi P, Kocsis L, **Mátyás KK**, Taller J *The effect of quinone outside inhibitors fungicides on powdery mildew in a grape vineyard in Hungary.* SCIENTIA HORTICULTURAE (2013) 161 pp. 233-238.) **IF:1,504**

Gorji MA, **Mátyás KK,** Dublecz Zs, Decsi K, Cernák I, Hoffmann B, Taller J, Polgár Zs *Investigation of in vitro osmotic stress tolerance in tetraploid potato and identification of major QTLs.* AMERICAN JOURNAL OF POTATO RESEARCH 89 (2012): 6 pp. 453-464. **IF:1,250**

Mátyás KK, Taller J, Cseh A, Poczai P, Cernák I *Development of a simple PCR-based assay for the identification of triazine resistance in the noxious plant common ragweed (Ambrosia artemisiifolia) and its applicability in higher plants*. BIOTECHNOLOGY LETTERS 33 (2011): 12 pp. 2509-2515. **IF:1,683**

Poczai P, **Mátyás KK**, Szabó I, Varga I, Hyvönen J, Cernák I, Mosapour Gorji A, Decsi K, Taller J *Genetic Variability of Thermal Nymphaea (Nymphaeaceae) Populations Based on ISSR Markers: Implications on Relationships, Hybridization, and Conservation*. PLANT MOLECULAR REPORTER 29 (2011): 4 pp. 906-918.) **IF: 2,453**

Poczai P, **Mátyás K**, Taller J, Szabó I *Study of the origin of the rarely cultivated edible Solanum species: morphological and molecular data.* BIOLOGIA PLANTARUM 54 (2010) : 3, pp. 543-546. **IF: 1,582**

National Publications (In Hungarian)

Taller J, Nagy E, Decsi K, Kutasy B, **Mátyás KK**, Farkas E, Kolics B, Barta E, Virág E *A transzkriptomika hasznosítása a gyomkutatásban. Esettanulmány a legelterjedtebb gyomnövényünkkel, az ürömlevelű parlagfűvel (Ambrosia artemisiifolia* L.*).* MAGYAR GYOMKUTATÁS ÉS TECHNOLÓGIA 17 (2016) 5-15.

Mátyás KK, Vignesh M, Taller J *Előzetes eredmények az ürömlevelű parlagfű (Ambrosia artemisiifolia L.) Kárpát-medencei populációgenetikai vizsgálatáról.* MAGYAR GYOMKUTATÁS ÉS TECHNOLÓGIA 13 (2012) 21-36.

Conference articles (In English)

Mátyás KK, Kolics B, Csép A, Nagy E, Taller J *Isolation and preliminary analysis of the variability of the ALS (Acetolactate synthase) gene from common ragweed (Ambrosia artemisiifolia L.).* In: Bene Sz (szerk.) 20th Youth Scientific Forum (2014): University of Pannonia Georgikon Faculty Keszthely, Magyarország, Elektronikus kiadvány *pp. 386-396.*

International Conferences

Kolics B, Sajtos Zs, **Mátyás K**, Kolics É, Taller J, Baranyai E. *"Lithium Chloride - Hazard or Possibility?"* In 46th Apimondia International Apicultural Congress Montréal Québec - Canada, 8-12 September 2019, Elektronikus kiadvány p. 285.

Kolics B, **Mátyás K**, Kutasy B, Parrag T, Házi F, Kolics É, Orbán L, Taller J *Novel method to identify the sex alleles of the honey bee,* In 46th Apimondia International Apicultural Congress Montréal Québec Canada, 8-12 September 2019 Elektronikus kiadvány p58.

Mátyás KK, Poczai P, Cernák I, Benoit DL, Taller J *Identification of linuron resistance in the common ragweed (Ambrosia artemisiifolia L.)* The 10th International Ph.D. Student Conference on Experimental Plant Biology (2012) Brno, Czech Republic, September 3-5 2012 Bulletin-Book of Abstracts 117 p.

Mátyás KK, Murthy V, Gorji AM *Population Genetic analysis of Common ragweed (Ambrosia artemisiifolia L.) in Europe using DNA – Based Molecular Markers* In: Bohren C, Bertossa M, Schoenenberger N, Rossinelli M, Conedera M (szerk.) 3rd International Symposium Environmental Weeds and Invasive Plants (2011): European Weed Research Society (EWRS) p127.

National Conferences (In Hungarian)

Solti I, Farkas E, **Mátyás K**, Székvári K, Kolics B, Kutasy B, Nagy E, Kálmán N, Hegedűs G, Taller J, Virág E *Az ürömlevelű parlagfű (Ambrosia artemisiifolia L.) főbb allergénjeinek génkifejeződés vizsgálata*. MAGYAR GYOMKUTATÁS ÉS TECHNOLÓGIA 20 (2019): 1p. 87.

Nagy E, Virág E, **Mátyás K**, Kutasy B, Solti I, Kolics B, Taller J *Mikroszatellitek az ürömlevelű parlagfű (Ambrosia artemisiifolia L.) kloroplasztisz genomjában* MAGYAR GYOMKUTATÁS ÉS TECHNOLÓGIA 20 (2019:) 1 p. 85.

Nagy E, **Mátyás K**, Kolics B, Virág E, Kutasy B, Taller J *Molekuláris genetikai vizsgálatok a virágzás szabályozottságának feltárására az ürömlevelű parlagfűben (Ambrosia artemisiifolia L.)* MAGYAR GYOMKUTATÁS ÉS TECHNOLÓGIA 16 (2015): 1 p. 67

Varga I, Nagy V, Baltazár T, **Mátyás KK** *A fehér fagyöngy (Viscum album L.) elleni herbicides védekezés hatékonyságának vizsgálata,* In: XXII. Keszthelyi Növényvédelmi Fórum (2012) Keszthely, 2012.január 25-27., Magyarország: Pannon Egyetem Georgikon Kar p. 124

Taller J, Cernák I, **Mátyás KK**, Poczai P, Cseh A *Az ürömlevelű parlagfű (Ambrosia artemisiifolia L.) atrazin és ALS-gátló hatóanyagú herbicidekkel szembeni rezisztenciájának molekuláris genetikai vizsgálata*. In: Lehoczky É, Budai P, Kadlicskó S, Marczali Zs, Nádasy M, Nádasyné Iharosi E, Takács A (szerk.) XX. Keszthelyi Növényvédelmi Fórum (2010): 20

éves a Keszthelyi Növényvédelmi Fórum 1990-2010, Keszthely, 2010. január 27-29, Magyarország p. 151

Conference articles (In Hungarian)

Solti I, Farkas E, **Mátyás K**, Székvári K, Kolics B, Kutasy B, Nagy E, Kálmán N, Hegedűs G, Taller J, Virág E *Az ürömlevelű parlagfű (Ambrosia artemisiifolia L.) allergénjeinek génkifejeződés vizsgálata.* In: Dr. Bene Szabolcs (szerk.) XXV. Ifjúsági Tudományos Fórum (2019) Pannon Egyetem Georgikon Kar, Keszthely, 2019. május 23. CD kiadvány ISBN száma: 978-963-9639-98-0.

Kutasy B, **Mátyás K**, Virág E, Decsi K, Nagy E, Farkas E, Hegedűs G, Kolics B, Taller J *Herbicid célgének in silico azonosítása az ürömlevelű parlagfűben (Ambrosia artemisiifolia L.)*. In: Bene, Szabolcs (szerk.) XXIII. Ifjúsági Tudományos Fórum (2017), Pannon Egyetem Georgikon Kar, Keszthely, 2017.május26., CD kiadvány, ISBN 978-963-9639-87-4.

Kutasy B, Virág E, Decsi K, Nagy E, Farkas E, **Mátyás K**, Kolics B, Taller J *Herbicid célgének vizsgálata az ürömlevelű parlagfűben (Ambrosia artemisiifolia L.).* In: Nagy Zita Barbara (szerk.) LVIII. Georgikon Napok: Felmelegedés, ökolábnyom, élelmiszerbiztonság. 462 p., Keszthely, Magyarország, 2016.09.29-2016.09.30. pp. 486-496. ISBN:978-963-9639-85-0

Farkas E, Decsi K, Nagy E, **Mátyás K**, Kolics B, Kutasy B, Virág E, Taller J *Pollen allergének transzkriptomikai vizsgálata az ürömlevelű parlagfűben (Ambrosia artemisiifolia L.)* In: Nagy Zita Barbara (szerk.) LVIII. Georgikon Napok: Felmelegedés (2016), ökolábnyom, élelmiszerbiztonság. 462 p., Keszthely, Magyarország, 2016.09.29-2016.09.30., pp. 92-99. ISBN:978-963-9639-85-0

Nagy E, Virág E, **Mátyás K**, Kolics B, Kutasy B, Taller J *Az ürömlevelű parlagfű (Ambrosia artemisiifolia L.) kloroplasztisz genomjának vizsgálata* In: Nagy Zita Barbara (szerk.) LVIII. Georgikon Napok (2016): Felmelegedés, ökolábnyom, élelmiszerbiztonság. 462 p., Keszthely, Magyarország, 2016.09.29-2016.09.30., pp. 479-485. ISBN:978-963-9639-85-0

Nagy E, **Mátyás K**, Kolics B, Virág E, Kutasy B, Taller J *Molekuláris genetikai vizsgálatok a virágzás szabályozottságának feltárására az ürömlevelű parlagfűben (Ambrosia artemisiifolia L.)* In: Pannon Egyetem Georgikon Kar Állattudományi és Állattenyésztéstani Tanszék (szerk.) XXI. Ifjúsági Tudományos Fórum (2015). Konferencia helye, ideje: Keszthely, Magyarország, 2015.05., 6 p. ISBN:978-963-9639-78-2

Tóth G, **Mátyás KK**, Buda B *A szombathelyi Isis-szentélyben feltárt emberi csontmaradványok vizsgálatának előzetes eredményei* In: Újlaki Pongrácz, Zsuzsánna (szerk.) Hadak útján: Népességek és iparok a népvándorlás korában: A népvándorláskor fiatal kutatóinak XVI. Konferenciáján (Nagykovácsi, 2005. szeptember 26-28.), Nagykovácsi, Magyarország: Pars Kft., (2006) pp. 33-36., 4p

References

ACHARD, P., BAGHOUR, M., CHAPPLE, A., HEDDEN, P., VAN DER STRAETEN, D., GENSCHIK, P., MORITZ, T. & HARBERD, N. P. 2007. The plant stress hormone ethylene controls floral transition via DELLA-dependent regulation of floral meristem-identity genes. *Proceedings of the National Academy of Sciences,* 104**,** 6484-6489.

ALTSCHUL, S. F., GISH, W., MILLER, W., MYERS, E. W. & LIPMAN, D. J. 1990. Basic local alignment search tool. *Journal of molecular biology,* 215**,** 403-410.

ANDREWS, S. 2010. FastQC: a quality control tool for high throughput sequence data. Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom.

CHANDLER, J. (2011) The hormonal regulation of flower development. Journal of Plant Growth Regulation 30:242-254

CJB (2016). African Plant Database. Conservatoire et Jardin Botaniques de la Ville de Geneve, Geneva, Switzerland, and South African National Biodiversity Institute, Pretoria, South Africa. Geneva, Switzerland: CJB/SANBI. http://www.ville-ge.ch/musinfo/bd/cjb/africa/

CONESA, A., GÖTZ, S., GARCÍA-GÓMEZ, J. M., TEROL, J., TALÓN, M. & ROBLES, M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics,* 21**,** 3674-3676.

COUNCIL OF HEADS OF AUSTRALASIAN HERBARIA 2016. Australia's Virtual Herbarium., Australia: Council of Heads of Australasian Herbaria. http://avh.ala.org.au

EURO+MED 2016. Euro+Med PlantBase - the information resource for Euro-Mediterranean plant diversity. http://www.emplantbase.org/home.html

FLORA OF CHINA EDITORIAL COMMITTEE 2016 Flora of China. St. Louis, Missouri and Cambridge, Massachusetts, USA: Missouri Botanical Garden and Harvard University Herbaria. http://www.efloras.org/flora_page.aspx?flora_id=2

GÖTZ, S., GARCÍA-GÓMEZ, J. M., TEROL, J., WILLIAMS, T. D., NAGARAJ, S. H., NUEDA, M. J., ROBLES, M., TALÓN, M., DOPAZO, J. & CONESA, A. 2008. High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic acids research,* 36**,** 3420-3435.

HAAS, B. J., PAPANICOLAOU, A., YASSOUR, M., GRABHERR, M., BLOOD, P. D., BOWDEN, J., COUGER, M. B., ECCLES, D., LI, B. & LIEBER, M. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature protocols,* 8**,** 1494-1512.

HARRIS, E. M. 1999. Capitula in the Asteridae: a widespread and varied phenomenon. *The Botanical Review,* 65**,** 348.

HODGINS, K. A., LAI, Z., NURKOWSKI, K., HUANG, J. & RIESEBERG, L. H. 2013. The molecular basis of invasiveness: differences in gene expression of native and introduced common ragweed (A mbrosia artemisiifolia) in stressful and benign environments. *Molecular ecology,* 22**,** 2496-2510.

IRISH, V. F. 2010. The flowering of Arabidopsis flower development. *The Plant Journal,* 61**,** 1014-1028.

KORESSAAR, T. & REMM, M. 2007. Enhancements and modifications of primer design program Primer3. *Bioinformatics,* 23**,** 1289-1291.

KOVALEV, O. 1989. Spread of adventitious plants of the tribe Ambrosia in Eurasia and methods of biological control of weeds of the genus Ambrosia L.(Ambrosieae, Asteraceae). *Trudy Zoologicheskii, Institut Akademii Nauk SSSR,* 189**,** 7-23.

LANGMEAD, B., TRAPNELL, C., POP, M. & SALZBERG, S. L. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology,* 10**,** R25.

LI, B. & DEWEY, C. N. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC bioinformatics,* 12**,** 323.

LORENZI, H. J. & JEFFERY, L. S. 1987. Weeds of the United States and their control. AVI Book, New York, N.Y. 355 pp.

MAKRA, L., JUHÁSZ, M., BORSOS, E. & BÉCZI, R. 2004. Meteorological variables connected with airborne ragweed pollen in Southern Hungary. *International Journal of Biometeorology,* 49**,** 37-47.

MILLAR, A. A. & GUBLER, F. 2005. The Arabidopsis GAMYB-like genes, MYB33 and MYB65, are microRNA-regulated genes that redundantly facilitate anther development. *The plant cell,* 17**,** 705-721.

MORTAZAVI, A., WILLIAMS, B. A., MCCUE, K., SCHAEFFER, L. & WOLD, B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature methods,* 5**,** 621-628.

MURASHIGE, T. & SKOOG, F. 1962. A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiologia plantarum,* 15**,** 473-497.

ROCHETA, M., SOBRAL, R., MAGALHÃES, J., AMORIM, M. I., RIBEIRO, T., PINHEIRO, M., EGAS, C., MORAIS-CECÍLIO, L. & COSTA, M. M. 2014. Comparative transcriptomic analysis of male and female flowers of monoecious Quercus suber. *Frontiers in plant science,* 5**,** 599.

SUDRE, D., GUTIERREZ-CARBONELL, E., LATTANZIO, G., RELLÁN-ÁLVAREZ, R., GAYMARD, F., WOHLGEMUTH, G., FIEHN, O., ÁLVAREZ-FERNÁNDEZ, A., ZAMARREÑO, A. M. & BACAICOA, E. 2013. Iron-dependent modifications of the flower transcriptome, proteome, metabolome, and hormonal content in an Arabidopsis ferritin mutant. *Journal of experimental botany,* 64**,** 2665-2688.

SUNG, S. & AMASINO, R. M. 2004. Vernalization in Arabidopsis thaliana is mediated by the PHD finger protein VIN3. *Nature,* 427**,** 159-164.

TAIZ, L., ZEIGER, E., MØLLER, I. & MURPHY, A. 2015. Plant physiology and development. Sinauer Associates. *Inc., Sunderland*.

UNTERGRASSER, A., CUTCUTACHE, I., KORESSAAR, T., YE, J., FAIRCLOTH, B., REMM, M. & ROZEN, S. 2012. Primer3–new capabilities and interfaces. *Nucleic Acids Research,* 40(15):e115

VIRÁG, E., HEGEDŰS, G., BARTA, E., NAGY, E., MÁTYÁS, K., KOLICS, B. & TALLER, J. 2016. Illumina sequencing of common (short) ragweed (Ambrosia artemisiifolia L.) reproductive organs and leaves. *Frontiers in plant science,* 7**,** 1506.