IDENTIFICATION AND EXPRESSION OF POLLEN ALLERGEN TRANSCRIPTS IN DIFFERENT ORGANS OF THE COMMON RAGWEED (*AMBROSIA ARTEMISIIFOLIA* L.)

*Eszter Farkas[1], Zoltán Farkas[1], Géza Hegedűs[2], Barbara Kutassy[1], Balázs Kolics[1], Kinga Klára Mátyás[1], Tamás Parrag[1], Izabella Solti[1], Eszter Virág[1], János Taller[1]**

1. Department of Plant Science and Biotechnology, Georgikon Faculty, University of Pannonia, H-8360 Keszthely, Festetics str. 7, Hungary

2. Department of Economic Methodology, Georgikon Faculty, University of Pannonia, H-8360 Keszthely, Festetics str. 7, Hungary

**e-mail address: taller@georgikon.hu*

Abstract

An increasing number of individuals are affected in the temperate climate zone by allergy caused by the pollen of the common ragweed (*Ambrosia artemisiifolia* L.). For 11 allergen families of ragweed the polypeptide sequences are available, and for most the nucleotide sequence is also known. However, little is known about the expression of these allergen genes. In the present study transcriptomes of the male and female flowers, as well as of leaves were produced and transcripts with high similarity or identity to the allergen genes were

identified. All the analyzed 11 allergen genes could be detected in the male flowers and with the exception of Amb a 6 also in leaves. Amb a 1, Amb a 3, Amb a 8 and Amb a 12 were expressed also in female flowers. The strongest expression was found for Amb a 5 and then for Amb a 3. Our results indicate that expression of ragweed pollen allergens is not restricted to the pollen, but they function also in other organs of the plant. The results presented here contribute to our understanding on the function and role of the genes of pollen allergens in common ragweed.

Keywords

Ambrosia artemisiifolia, common ragweed, pollen allergen, Amb a, transcript, gene expression

Összefoglalás

Az ürömlevelű parlagfű (*Ambrosia artemisiifolia* L.) pollenje által kiváltott allergia egyre több embert érint az e gyomnövénnyel fertőzött területeken. A parlagfű 11 allergén géncsaládja fehérje-szinten ismert, és legtöbbjük nukleotid szekvenciája is feltárásra került. Ugyanakkor, az allergén gének expressziójáról kevés ismeret áll rendelkezésre. A világ különböző régióiból származó parlagfű növényekben azonosított allergén géneknek a hazai populációból való kimutathatósága, illetve működésének tanulmányozása céljából transzkriptomokat készítettünk a parlagfű hím és nő virágjából, valamint leveléből, és azonosítottuk az ismert allergén génekkel azonos vagy nagyfokú hasonlóságot mutató transzkripteket. A hím virágokból mind a 11 allergén gént ki tudtuk mutatni, és az Amb a 6 kivételével a levelekből is. Az Amb a 1, Amb a 3, Amb a 8 and Amb a 12 gének a nővirágokban is mutattak expressziót. A génkifejeződés az Amb a 5, majd az Amb a 3

esetében volt a legerősebb. Eredményeink azt mutatják, hogy az ismert parlagfű allergének, függetlenül attól, hogy fő vagy allergológiai szempontból kisebb jelentőségű allergének, a hazai növényekben is jelen vannak. Továbbá, eredményeink egyértelműen igazolják, hogy a parlagfű pollen allergén génjei nemcsak a hímvirágokban működnek, hanem a növény egyéb részeiben is van szerepük. Itt bemutatott eredményeink hozzájárulnak a parlagfű allergéneket kódoló gének működésének és szerepének jobb megértéséhez.

Introduction

The common ragweed (*Ambrosia artemisiifolia* L.) is a noxious, invasive weed in large areas of the temperate climate zone (Kazinczi et al. 2008). From 1995 it is the most widespread weed in Hungary (Novák et al. 2009). Besides endangering yield safety of crops the pollen of ragweed is the clinically most important seasonal aeroallergen (Rafnar et al. 1991). However, the pollen is the main source of ragweed allergy, it is known that also the green parts of the plant may cause dermatitis in sensitive patients.

Ragweed is wide source of different type pollen allergens which are on different importance regarding their allergenicity. Nomenclature of allergens is approved by the WHO/IUIS (World Health Organization and International Union of Immunological Societies). In the allergen nomenclature database of WHO/IUIS (http://www.allergen.org/) 11 different types of common ragweed allergens are identified. Each represent a different family, which involve an unknown number of isoforms. Two proteins, a pectate lyase (Amb a 1) (Rafnar et al., 1991; Nandy et al., 2011; Augustin et al., 2012; Augustin et al., 2013) and a cysteine protease (Amb a 11) (Bouley et al., 2015) are considered as major allergens with a prevalence of >90% and 66% in sensitization, respectively. Others are minor proteins, like plastocyanins (Amb a 3

(Klapper et al., 1980; Taller et al., 2016) and Amb a 7 (Roebber et al., 1991), defensin (Amb a 4) (Léonard et al., 2010), lipid transfer protein (Amb a 6) (Roebber et al., 1983; Hiller et al., 1998), profilin (Amb a 8), and polcalcins (Amb a 9 and Amb a 10) (Wopfner et al., 2008). For Amb a 7 no protein or DNA sequence is available. Amb a 5 is also a minor ragweed allergen with unknown function, although the 3D structure of the protein has been resolved (Ghosh et al., 1993; Metzler et al., 1992). A recently identified new allergen, Amb a 12 is an enolase (Bordas-Le Floch et al., 2015b). Besides these 11 allergens of the IUIS database a further minor ragweed allergen is known, which is a cysteine protease inhibitor (Amb a CPI) (Rogers et al., 1993).

Comprehensive knowledge of the allergen repertoire of ragweed pollen is prerequisite for accurate diagnosis and efficient immunotherapy. To this end we performed a transcriptome analysis on different developmental stage male and female flowers, as well as on leaves, using pooled samples of six phenotypically different plants. One of the aims of the present study was to explore whether the known allergen families, which were identified in plants from different geographic regions of the world, can be detected in the randomly chosen Hungarian ragweed plants. Further aim of the study was to explore whether the allergens are expressed only in the male flowers or also in female flowers and/or leaves, and to estimate their expression level.

Materials and Methods

Plant material

Six phenotypically different ragweed plants growing under natural conditions in Keszthely, Hungary were used. During the flowering season young shoots were covered with transparent

paper bags to protect from pollen contamination. Male racemes were covered separately. Male and female flowers were collected in seven and nine developmental stages, respectively, as follows:

Male flowers: stage 1.: 1 mm raceme; stage 2.: 4-5 mm long raceme; stage 3.: 10 mm long raceme; stage 4.: not yet opened nest from a 10 cm long raceme; stage 5.: opened nest without opened flowers of that 10 cm long raceme; stage 6.: a nest with partially opened flowers; stage 7.: a nest in stage of maximal pollen release.

Female flowers: stage 1.: young flowers without visible pistils; stage 2.: flowers with 1-2 mm pistils; stage 3.: flowers with 3-4 mm pistils; stage 4.: flowers with 5-6 mm pistils; stage 5.: flowers with full length pistils; stage 6.: flowers just after fertilization; stage 7.: fertilized flowers with living pistil; stage 8.: fertilized flowers with dying pistils; stage 9: flowers with clearly developing seeds in them.

Leaf samples were collected in five developmental stages: stage 1. leaves were collected from the youngest shoots, while stage 5. leaves from old branches of the same plant, and the other three stages were collected in between these two.

Collected samples were immediately frozen in liquid nitrogen and stored at -80^{o}C until analysis.

Molecular analyses

A male and a female flower, as well as a leaf sample was created by bulking the subsamples of the different developmental stages, in a way that each stage should be represented equally. Molecular analyses were performed as described in Virág et al. (2016). Briefly:

RNA of these three samples was extracted by TaKaRa Plant Extraction kit (Takara Bio, Japan). For cDNA synthesis the Illumina TruSeq RNA sample preparation kit was used.

RNA-sequencing was performed on an Illumina HiSeq2000 system (Illumina, USA). Each fragment was pair-end sequenced 100 nucleotide deep.

Bionformatics analyses

The http://allergen.org/ and Uniprot databases were used for allergen annotation. For the primary reconstruction of allergen gene expression analysis we used the Trinity *de novo* assembler, followed by transcript abundance calculation using Bowtie short read aligner and Geneious®. Normalized expression level values (RPKM – reads per kilobase per million mapped reads) were assigned with GenoUtils that was developed in Visual Studio and C# programming language (Virág and Hegedűs, 2018). For screening for highly similar sequences in the transcriptome datasets the available DNA sequences of nine Amb a gene families of the IUIS database and the Amb a CPI (Bordas-Le Floch et al., 2015), as well as the Amb a 5 sequence (Gosh et al., 1993) was used.

Results

The transcriptome datasets of male and female flowers, as well as of leaves have been screened with all published DNA sequences of 11 Amb a gene families for highly similar transcripts. Results are summarized in Table 1.

Table 1. Expression of Ambrosia artemisiifolia allergens

Allergen	Identification code[1]	Protein length (aa)[2]	Coverage (%)[3]	Identity (aa)[4]	Identity (%)[5]	Expression in male flowers (RPKM)	Expression in female flowers (RPKM)	Expression in leaves (RPKM)
Amb a 1	TR45018|c2_g2_i13	396	100	396/396	100	2138.04	1.40	27.73
Amb a 3	TR20404|c0_g3_i1	101	97	87/101	90	7958.35	2.39	71.24
Amb a 4	TR30326|c0_g2_i1	129	82	129/135	96	5517.13	0.00	1.33
Amb a 5	TR27249|c0_g1_i1	45	100	45/45	100	17671.01	0.00	6.31
Amb a 6	TR44274|c4_g7_i2	118	100	114/118	97	4673.33	0.00	0.00
Amb a 8	TR37128|c0_g1_i2	133	100	133/133	100	2533.45	2.67	41.92
Amb a 9	TR28466|c1_g1_i1	83	100	81/83	98	662.77	0.00	18.63
Amb a 10	TR38368|c0_g1_i2	160	100	160/160	100	349.50	0.00	20.42
Amb a 11	TR43337|c0_g2_i1	386	100	379/386	98	819.88	0.00	3.01
Amb a 12	TR43144|c0_g1_i1	445	100	445/445	100	671.89	137.50	1358.49
Amb a CPI	TR36272|c0_g1_i1	92	100	90/92	97	131.97	0.00	9.79

[1]The identification code of predicted allergen transcript in Ambrosia artemisiifolia transcriptome database

[2]The size of published allergen proteins (aa: number of amino acids)

[3]The percentage of query covered by alignment to the published allergen sequence

[4]The identity of query covered by alignment to the published allergen sequence (aa: number of amino acids)

[5]The percentage of identity by alignment to the published allergen sequence

Colors refer to the level of gene expression. Non-expressed genes are indicated with green color, while deep red indicates the highest expression level. Yellow color show transition expression value. The numbers show the RPKM value of each expressed genes. High RPKM value show higher gene expression level.

Sequence similarity

Transcripts which exactly matched sequences of allergens were obtained for the Amb a 1, Amb a 5, Amb a 8, Amb a 10 and Amb a 12. These five transcripts had identical sequence with the corresponding Amb a gene and covered them 100%.

Except for Amb a 3 and Amb a 4 we could identify transcripts in all three transcriptome which covered 100% the corresponding Amb a gene. Except for Amb a 3 the sequence similarity between the transcripts and allergen genes was very high, >96%.

In the IUIS database just the amino acid sequence of the Amb a 3 allergen is registered. There was no nucleic acid sequence available. From the transcriptomes used in the present study we recently identified a transcript (Taller et al., 2016) that encodes a protein with 90% similarity to the Amb a 3 amino acid sequence published by Klapper et al. (1980). In Table 1. the Amb a 3 isoform identified by Taller et al. (2016) is compared to the Amb a 3 protein of Klapper et al. (1980).

Expression levels

Expression of the Amb a transcripts of the three transcriptomes are shown in Table 1.

All identified transcripts were expressed in the male flowers, and except Amb a 6 they were expressed also in leaves. In female flowers expression could be detected just for the Amb a 1, Amb a 3, Amb a 8 and Amb a 12 allergen transcripts.

In male flowers the strongest expression was observed for Amb a 5 and then for the Amb a 3 isoform. Strong expression was detected for Amb a 4 and Amb a 6 and then for Amb a 8 and Amb a 1.

In female flowers expression of only of Amb a 12 is notably, since it is relatively high, being about one-fifth of that found in male flowers.

Also in leaves the highest expression was detected for Amb a 12 with an RPKM value about double of that found in male leaves. Except Amb a 12, where expressed, the expression of the Amb a genes is relatively low in leaves.

Discussion

Identification of new allergens from non-model plant species such as *A. artemisiifolia* are classically hindered by the paucity of protein or genomic information available in public databases (Bordas-Le Floch et al., 2015). For Amb a 7 neither a protein, nor a nucleotide sequence is known, and for Amb a 3 just a polypeptide sequence (Klapper et al., 1980) was available. For this later, a transcript coding for a protein with 90% similarity to that polypeptide was identified recently (Taller et al., 2016). However, this Amb a 3 isoform showed the second highest expression in male flowers in the present study, to prove the allergenic nature of it requires further functional and allergological analyses.

Amb a 1, Amb a 3, Amb a 8 and Amb a 12 were expressed in all three organs, indicating a more general functional role for these proteins. Amb a 1 is the major ragweed pollen allergen, since 95% of ragweed-sensitive individuals react to it. Further, the Amb a 1 protein is highly abundant, comprising about 6% of the total protein of pollen (Rafnar et al., 1991). In our study we found strong, but not outstanding expression for Amb a 1 in male flowers. Amb a 1 is a pectate lyase, which act in cell wall softening and have an important role in various plant developmental processes (Marin-Rodrigez et al., 2002). In the style pectate lyases facilitate pollen tube emergence and penetration of the pollen (Taniguchi et al., 1995), which can be a possible reason for abundancy of Amb a 1 in the ragweed pollen.

The function of Amb a 5 is not cleared yet. In male flowers we observed outstandingly high expression for this minor allergen. It needs to be studied what is the role of Amb a 5 in the plant and why is it expressed at such a high level.

Amb a 12 is a newly identified minor allergen belong to the enolases. In this study we found a twice so high expression for Amb a 12 in leaves than in male flowers. In this aspect it is the only one among the studied allergens, where expression in male flowers is lower than in leaves. Functional analysis of enolases in the plant and in the pollen would be necessary to understand the reasons of this phenomena.

Except Amb a 6 all examined allergens were expressed also in leaves, indicating that function of the majority of Amb a allergens is not restricted to the pollen. Many of the pollen allergens are evolutionarily conserved and are involved in stress responses, as the pathogenesis related (PR) proteins, and in metabolic processes, e.g. in cell wall metabolism, while other allergens emerged during evolution (Chen et al., 2016). Expression in leaves and female flowers and the molecular type of the different Amb a allergens indicate that they have a more general role in plants. However, whether allergens in the pollen are produced exclusively in the anther or they are accumulated to the pollen from other parts of the plant needs further investigations.

In this study we could identify each analyzed Amb a gene family in the randomly chosen Hungarian ragweed plants. Originally, the Amb a genes were identified from plants on different geographic origin, like North-America or Western-Europe. This indicates, that the analyzed Amb a allergens can be widespread in ragweed populations and possibly present in every ragweed plant. In ongoing studies we are examining the expression and function of the different allergens in different organs and their role during development of the ragweed plant.

Acknowledgments

This work was supported by the K100919, Postdoc-2014-44 and UNKP-16-4 grants of the Hungarian Scientific Research Fund (OTKA), Hungarian Academy of Sciences (MTA) and New National Excellence Programme, Ministry of Human Resources of the Hungarian Government.

References

Alvarez-Buylla E.R., Benítez M., Corvera-Poiré A., Cador Á.C., de Folter S., de Buen A.G., Garay-Arroyo A., García-Ponce B., JaimesMiranda F., Pérez-Ruiz R.V., Piñeyro-Nelson A., Sánchez-Corralesa Y.E. Flower Development. In: The Arabidopsis Book. The American Society of Plant Biologists; Rockville, MD.: 2010. doi:10.1199/tab.0127, http://www.aspb.org/publications/arabidopsis/.

Augustin, S., Stock. M., Cromwell, O., Nandy, A., Reese, G. 2012. Proteomic and immunological characterization of ragweed allergens. *World Allergy Organ Journal*. 5 (2). 23–24.

Augustin, S., Wald, M., Asero, R., Reese, G., Klysner, S., Nandy, A. 2013. Assessment of Amb a 1 isoallergens as basis for development of a recombinant ragweed immunotherapeutic vaccine. *Allergy*. 68 (97). 111.

Bouley, J., Groeme, R., Le Mignon, M., Jain, K., Chabre, H., Bordas-Le Floch, V., Couret, M., Bussières, L., Lautrette, A., Naveau, M., Baron-Bodo, V., Lombardi, V., Mascarell, L., Batard, T., Nony, E.,Moingeon, P. 2015. Identification of the cysteine protease Amb a 11 as a novel major allergen from short ragweed. *Journal of Allergy and Clinical Immunology*. 136 (4). 1055–1064.

Bordas-Le Floch, V., Groeme, R.,Chabre, H., Baron-Bodo, V., Nony E., Mascarell, L., Moingeon, P. 2015a. New insights into ragweed pollen allergens. *Current Allergy and Asthma Reports*. 15 (11). 63.

Bordas-Le Floch, V., Le Mignon, M., Bouley, J., Groeme, R., Jain, K., Baron-Bodo, V., Nony, E., Mascarell, E., Moingeon, E. 2015b. Identification of Novel Short Ragweed Pollen Allergens Using Combined Transcriptomic and Immunoproteomic Approaches. *PLoS ONE* 10 (8). 1-18.

Chen, M., Xu, J., Devis, D., Shi, J., Ren, K., Searle, I., Zhang, D. 2016. Origin and Functional Prediction of Pollen Allergens in Plants. *Plant Physiology*. *172* (1). 341–357.

Ghosh, B., Perry, M.P., Rafnar, T., Marsh, D.G. 1993 Cloning and expression of immunologically active recombinant Amb a V allergen of short ragweed (Ambrosia artemisiifolia) pollen. *Journal of Immunology*. 150 (12). 5391–5392.

Hiller, K. M., Lubahn, B.C., Klapper, D.G. 1998. Cloning and expression of ragweed allergen Amb a 6. *Scandinavian Journal of Immunology*. 48. 26-36.

Kazinczi G., Béres I., Novák R., Bíro K., Pathy Z. 2008. Common ragweed (Ambrosia artemisiifolia): A review with special regards to the result in Hungary I. Taxonomy, origin and distribution, morphology, life cycle and reproduction strategy. *Herbologia.* 9. 55-91.

Klapper, D. G., Goodfriend, L., Capra, J. D. 1980. Aminoacid sequence of ragweed allergen Ra3. *Biochemistry.* 19 (25). 5729-5734.

Léonard, R., Wopfner, N., Pabst, M., Stadlmann, J., Petersen, B. O., Duus, J., Himly, M., Radauer, C., Gadermaier, G., Razzazi-Fazeli, E., Ferreira, F., Altmann, F. 2010. A new allergen from ragweed (Ambrosia artemisiifolia) with homology to Art v 1 from mugwort. *The Journal of Biological Chemistry*. 285 (35). 27192-27200.

Marín-Rodriguez, M.C., Orchard, J., Seymour, G. B. 2002. Pectate lyases, cell wall degradation and fruit softening. *Journal of Experimental Botany.* 53 (377). 2115-2119.

Metzler, W.J., Valentine, K., Roebber, M., Marsh, D.G., Mueller, L. 1992. Proton resonance assignments and three-dimensional solution structure of the ragweed allergen Amb a V by nuclear magnetic resonance spectroscopy. *Biochemistry*. 31 (37). 8697–705.

Nandy, A., Augustin, S., Mitulski, L., Cromwell, O. 2011. Isoallergen analysis of pectate lyases (Amb a 1 and Amb a 2) from commercial short ragweed pollen. *Journal of Allergy and Clinical Immunology*. 127 (2). AB168.

Novák, R., Dancza, I., Szentey, L., Karamán J. (2009). Magyarország szántóföldjeinek gyomnövényzete. (in Hungarian). Ötödik Országos Szántóföldi Gyomfelvételezés. FVM, Budapest.

Rafnar, T., Griffith, I. J., Kuo, M., Bond, J. F., Rogers, B. L., Klapper, D. G. 1991.Cloning of Amb a I (Antigen E), the major allergen family of short ragweed pollen. *The Journal of Biological Chemistry.* 266 (2). 1229-1236.

Roebber, M., Hussain, R., Klapper, D. G., Marsh, D. G., 1983. Isolation and properties of a new short ragweed pollen allergen, Ra6. *Journal of Immunology.* 131 (2). 706-711.

Roebber, M., Marsh, D. G. 1991. Isolation and characterization of allergen Amb a VII from short ragweed pollen. *Journal of Allergy and Clinical Immunology*. 87. 324.

Rogers, B. L., Pollock, J., Klapper, D.G., Griffith, I. J. 1993. Sequence of the proteinase-inhibitor cystatin homologue from the pollen of Ambrosia artemisiifolia (short ragweed). *Gene*. 133. 219–221.

Smyth, D. R., Bowman J. L., Meyerowitz E. M. 1990. Early flower development in Arabidopsis. *Plant Cell*. 2 (8). 755-767.

Taller J., Decsi K., Farkas E., Nagy E., Mátyás K. K., Kolics B., Kutasy B., Virág E. 2016. De novo transcriptome sequencing based identification of Amb a 3-like pollen allergen in common ragweed (Ambrosia artemisiifolia). *Journal of Botanical Science* 5 (2). 36-40.

Taniguchi, Y., Ono, A., Sawatani, M., Nanba, M., Kohno, K., Usui, M., Kurimoto, M., Matuhasi T. 1995. *Cry j* I, a major allergen of Japanese cedar pollen, has a pectate lyase enzyme activity. *Allergy* 50**.** 90–93.

Virág, E. and Hegedűs, G. 2018. Digitális gén expressziós elemzések szoftveres támogatása. 48. Membrán transzport konferencia, Sümeg, Május 15-18. 2018. (poszter) (in Hungarian)

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

Wopfner, N., Gruber, P., Wallner, M., Briza, P., Ebner, C., Mari, A., Richter, K., Vogel, L., Ferreira, F. 2008. Molecular and immunological characterization of novel weed pollen pan-allergens. *Allergy*. 63 (7). 872–881.