

# Identification and characterization of genes connected with flower morphogenesis in cucumber

MAGDALENA PAWEŁKOWICZ<sup>1\*</sup>, PAWEŁ OSIPOWSKI<sup>1</sup>, MICHAŁ WOJCIESZEK<sup>1</sup>,  
 RAFAŁ WÓYCICKI<sup>1,4</sup>, JUSTYNA WITKOWICZ<sup>1,2,3</sup>, DIRK HINCH<sup>2</sup>, ZBIGNIEW PRZYBECKI<sup>1</sup>

<sup>1</sup>Department of Plant Genetics, Breeding & Biotechnology, Faculty of Horticulture and Landscape Architecture, Warsaw University of Life Sciences – SGGW, Warszawa, Poland

<sup>2</sup>Max-Planck-Institute of Molecular Plant Physiology, Potsdam, Germany

<sup>3</sup>Department of Plant Physiology, Ruhr-University Bochum, Bochum, Germany

<sup>4</sup>Department of Biology, Faculty of Science, McGill University, Montreal, Canada

\*Corresponding author: magdalena\_pawelkowicz@sggw.pl

## Abstract

Sex determination and flower morphogenesis are very broad and complex processes controlled at many levels. Four clones have been isolated from cucumber transcriptomes, mapped onto the cucumber genome and checked if the corresponding genes expression differed between the vegetative and generative tissues (leaf, shoot apex, and 1- to 2-mm flower buds) of monoecious and gynoeceious cucumber lines. To determine the role, and characteristics of identified genes in flower morphogenesis, as well as to understand the flower reproduction in cucumber, comprehensive computational studies using upstream regulatory elements and protein motifs were performed. A genome-wide overview of cucumber clones revealed that sequence of only one clone was mapped in the coding site. The gene was described as *CsPSTK1* encoding serine/threonine kinase. The results allow us to conclude that cucumber generative organs differ in responsiveness to plant hormones due to the distinct signal transductions that are mediated by protein kinases in male and female organs of the floral buds and shoot apices. Protein kinases may be an alternative way for hormonal signal transduction in flowers of the opposite sex, taking part in the inhibition of unwanted generative organs that cause the development of a unisex flower.

**Key words:** *Cucumis sativus*, flower morphogenesis, sex determination, tissue specification, cucumber genome, protein serine/threonine kinase

## Introduction

Rapid, whole genome sequencing tools, together with computational comparative analyses, provide a great opportunity to study genes and proteins that are involved in basic cellular processes of organogenesis. In this paper we identify and characterize genes that may be involved in sex determination and flower morphogenesis in cucumber. The data from the complete genome of *C. sativus* L. cv. Borszczagowski were used for our analyses (Wóycicki et al., 2011). During the past several years, a significant progress has been made in elucidating the mechanisms of plant sex determination through extensive studies on flowering plants (Dellaporta and Calderon-Urrea, 1993), and particularly the monoecious cucumber (Malepszy and Niemirowicz-Szczytt, 1991). Cucumber is a diploid species with seven pairs of chromosomes ( $2n = 14$ ), with the genome, estimated at 367 Mb, that has already been sequenced (Wóycicki et al., 2011;

Huang et al., 2009; Cavagnaro et al., 2010). Transcriptome sequencing in cucumber has been reported twice (Wu et al., 2010; Guo et al. 2010), but the processes by which sex expression is determined have not been fully elucidated.

At the early stages of development, cucumber floral primordia are bisexual and contain initial forms of both anthers and pistils. Sex determination follows selective arrest of development of either the staminate or the pistillate primordia (Bai et al., 2004). Cucumber sex expression is mainly determined by the genes *F*, *m*, *a*, *gy*, and *h*. The Female (*F*) gene regulates the number of female flowers and was cloned and shown to encode 1-aminocyclopropane-1-carboxylic acid synthase (*CsACSIG*), which is involved in ethylene biosynthesis (Trebish et al., 1997; Mibus et al., 2004). The andromonoecious gene (*M*) was also identified as a previously characterized putative 1-aminocyclopropane-1-carboxylic acid syn-

these gene *CsACS2*, while the *m* allele is mutated at a conserved site (Gly33Cys). These isoforms display reduced enzymatic activity (Boualem et al., 2009; Li et al., 2009). The *F* gene promotes femaleness, while the *m* gene regulates the appearance of hermaphroditic flowers on the plant. Little is known about *a*, *gy* and *h* genes. In a study it was shown that gene *a* may be linked to the gene of copper transporter (Responsive to ANtagonist1) designed as *CsRANI* (Terefe, 2005).

The role of plant hormone ethylene in cucumber sex determination is well established (Li et al., 2012). Cucumber gender can be influenced by many environmental factors, including photoperiod, temperature, and plant hormones such as ethylene, auxin, and gibberellic acid (Atsmon et al., 1968; Perl-Treves, 1999; Yamasaki et al., 2005). The expression of the *CsACS2* gene is correlated with ethylene production and which can also induce the expression of the *CsACS2* gene; therefore suggesting a positive feedback (Li et al., 2012). Identification of more mutations in the cucumber sex expression pathway and cloning of those genes may reveal more information for better understanding of the process (Li et al., 2012).

In this study, six methods were applied to screen the differences between four isogenic cucumber lines (two pairs of near-isogenic lines [NILs]) which differed by flower gender. Four methods (random amplification of polymorphic DNA [RAPD], amplified fragment length polymorphism [AFLP], differential subtraction chain [DSC], and genetically directed differential subtraction chain [GD-DSC]) were applied to screen the differences between genomes and two methods were applied for transcriptome profiling (DH differential hybridization or cDNA-DSC), for the 1 to 2-mm floral buds (male, female, and hermaphrodite) (Przybecki et al., 2003, 2004). We confirmed, by macroarray analysis, varied expression of 707 sequences. The aim of this study was to find genes which depict very stable expression at three developmental stages (leaves, shoot apex and floral buds) and differentiated in monoecious and gynoecious lines. As a result of our investigation, four genes that may be involved in flower development were selected. Those genes are as follows: *CsB10dhB2\_93*, *CsB10dhB2\_94*, *CsB10dhB2\_148*, and *CsB10dhB2\_317*.

To identify those genes, we screened the Borszczagowski's cucumber genome database (ACYN00000000 at <http://www.ncbi.nlm.nih.gov>), investigating the anno-

tation and structure of each gene as well as the possible protein motifs in the deduced amino acid sequences of the polypeptides coded by these genes. To gain an insight into the functional roles of the analyzed genes, we also compared their expression patterns in three different organs of the chosen cucumber lines.

## Material and methods

### *Plant material and libraries*

To examine cucumber gene expression patterns among different lines and organs, and test their correlation with sex determination, we derived two pairs of near isogenic lines from the Borszczagowski cultivar: the monoecious B10 (*AAffMMGyGy*) line, with male and female flowers and the gynoecious 2gg (*AAffMMgygy*) line, with only female flowers. The plants were cultivated in plastic pots in the greenhouse with a 16h/8h day/night photoperiod (25-27°C during the day/18-20°C during the night). The light intensity in the greenhouse was 1500  $\mu\text{mol}$  (quantum)/ $\text{m}^2/\text{s}$  of photosynthetically active radiation. For RNA isolation, leaves, shoot apex, and floral buds (1-2 mm) were frozen in liquid nitrogen, and stored at  $-80^\circ\text{C}$ . Collection of plant material was divided into two phases: phase 1 (ph1), organs collected up to the sixth node; and phase 2 (ph2), organs collected between the sixth and fifteenth node. For the B10 line, ph1 consisted of male buds, and ph2 consisted of female and male buds. In our previous investigations we combined four methods (RAPD, AFLP, DSC, GD-DSC) to screen for differences between the genomes of four isogenic lines (two pairs of NILs) which differed by flower gender. We also performed transcriptome profiling (DH or cDNA-DSC) for 1 to 2 mm cucumber floral buds (male, female, and hermaphrodite). As a result, differential genomic and cDNA libraries of cucumber NILs were constructed and subsequently used for macroarray analysis (Przybecki et al., 2003, 2004).

### *Macroarray validation*

In order to validate the above-mentioned techniques and to identify genes that displayed stable, but differential, expression patterns, not only in 1 to 2 mm floral buds, but also in leaves and the differentiating shoot apex in the two phases of growth (ph1 and ph2), the macroarray analysis was performed. Macroarray experiments are described in detail by Witkowicz (unpublished data). Altogether 1395 clones were spotted onto nylon

membranes. To minimize variations, each clone was spotted twice on the nylon and each membrane had two technical replicates. Additionally, three biological replicates in hybridization procedure were performed. The values were therein normalized by background subtraction and reference hybridization. The radioactive intensity of each spot was quantified as a raw value. The local background value was subtracted from the raw value; and normalization of all membranes was performed first, by dividing each spot value by a mean of all values from the membrane, and then, dividing every spot value by a reference hybridization value relevant to each spot. To verify the quality of the macroarray experiments, the reproducibility of macroarray hybridizations was assessed by calculating the Pearson's correlation coefficient between every pair of biological replicates, for each of the three tissues.

### **Identification of genes in the *Cucumis sativus* genome**

To identify the sequences of differentially expressed genes in the cucumber genome, a blast analysis, using blastn and tblastx alignments v2.2.22 (<http://www.ncbi.nlm.nih.gov/>), with an e-value =  $1e^{-10}$  was applied. The blastn analysis was performed to select the genomic contigs from the cucumber genome sequence of the B10 line, Borszczagowski cultivar ACYN00000000 (Wóycicki et al., 2011). To compare cDNA sequences, we used tblastn on cDNA from selected clones and a library of cucumber transcript csb10a\_v1\_genes (<http://csgenome.sggw.pl>) that were obtained after *in silico* structural annotation by GM HMM software (Borodovsky and Lomsadze, 2011a). After the two-level searching databases, the sequence alignments resulted in the identification of genes that were mapped into a *C. sativus*-sequenced contigs dataset and were also present in the predicted transcriptome dataset. To investigate the genetic structure, we analyzed intron and exon locations with GeneMark.hmm ES software (<http://www.exon.gatech.edu>), on the whole contigs assembly training set (Wóycicki et al., 2011). For promoter analyses, we used *Cucumis sativus* promoters, upstream – 1000 nt from the transcription start site (TSS).

### **Bioinformatics analysis**

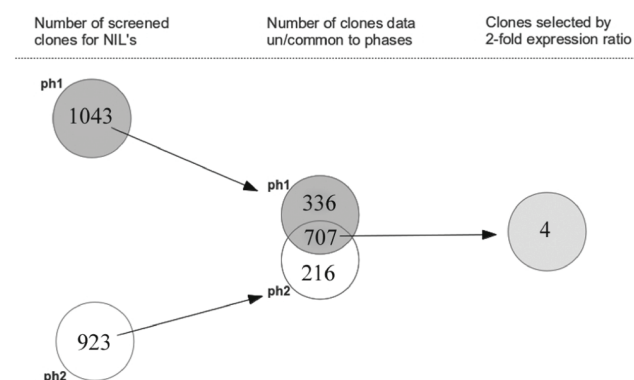
The functional analysis was performed using the following methods: 1) to find homologous sequences in the

database and to classify the gene ontology – Blast and GoRetriever (<http://www.geneontology.org>); 2) to analyze the presence of protein motifs – Prosite Proteomic Server (<http://expasy.org/prosite/>), InterProScan (<http://www.ebi.ac.uk/Tools/pfa/iprscan/>) Pfam tool (<http://pfam.sanger.ac.uk/>); 3) to predict the transmembrane helix – the TMHMM 2.0 server (<http://www.cbs.dtu.dk/services/TMHMM-2.0/>); 4) to identify the presence and location of signal peptide: SignalP (Petersen, 2011; <http://www.cbs.dtu.dk/services/SignalP/>); 5) to analyze the cis-regulatory elements along the promoter regions: PlantCare (Lescot et al., 2002; <http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>); 6) to predict secondary structure of the determined *C. sativus* orthologs: GOR V (Kloczkowski et al., 2002; <http://gor.bb.iastate.edu/>) and Jpred3 (Cole et al., 2008; <http://www.compbio.dundee.ac.uk/www.jpred/>); and 7) for cellular localization of protein: TargetP (Emanuelsson et al., 2000; <http://www.cbs.dtu.dk/services/TargetP/>) and WolfPsort (<http://wolfsort.org/>).

## **Results**

### **Macroarray**

In our investigation, we combined four methods (RAPD, AFLP, DSC, GD-DSC) to screen the differences between the genomes of four isogenic lines (two pairs of NILs) that differed with regard to flower gender, and two methods for transcriptome profiling (DH or cDNA-DSC) for the 1-2 mm floral buds (male, female, and hermaphrodite) (Przybecki et al., 2003, 2004). In the macroarray experiments, sscDNA from leaf, shoot apices and 1-2 mm floral buds of lines B10 and 2gg during two phases of growth (ph1 and ph2; see materials and methods) was used. When the quality of the macroarray experiments was verified, Pearson's correlation coefficients ranged from 0.7 to 0.99 for normalized data. These results indicate that the variation among biological replicates was rather small. When we investigated the level of hybridization for B10 and 2gg to estimate the liability of transcriptome profiles, transcripts specific for B10 resulted in 94% of the B10 clones hybridized with sscDNA B10, and 90% of the 2gg clones hybridized with 2gg sscDNA. Altogether, after normalization, the expression profiles of 1043 (ph1) and 923 (ph2) genes were chosen for further analysis, according to the general scheme presented in Figure 1. To identify genes that might be involved in differential floral morphogenesis, gene expression



**Fig. 1.** The scheme of macroarray analysis. The number of total and selected clones in phase 1 (ph1) and phase 2 (ph2)

ratio of more than 2-fold was regarded as significant. Seven hundred and seven sequences (common to ph1 and ph2) whose transcripts were differentially accumulated between the lines were obtained. As a result of the investigation performed under very strict conditions, four genes with differential expressions between NILs but a stable expression profiles between the two phases were identified. Those genes: *CsB10dhB2\_93*, *CsB10dhB2\_94*, *CsB10dhB2\_148*, and *CsB10dhB2\_317* may be involved in flower development.

### The differential expression of genes

After normalization and comparison of two sets of plant collection phases (ph1 and ph2), four transcriptomic sequences were designated for further characterization: *CsB10dhB2\_93*, *CsB10dhB2\_94*, *CsB10dhB2\_148*, and *CsB10dhB2\_317* (Fig. 1). Each of them represented a gene that displayed the stable expressional pattern in both sets of plant material for ph1 and ph2 (Fig. 2). *CsB10dhB2\_93* was highly expressed in B10 lines, especially in the shoot apex, when compared to the 2gg shoot apex. In leaves, the expression pattern was slightly higher in the B10 line, but in 1-2 mm floral buds, the expression dramatically increased in the 2gg as compared to the B10 line. *CsB10dhB2\_94* displayed an opposite expression pattern to *CsB10dhB2\_93*. Its leaves had a higher expression in the B10 than in the 2gg line; however, in the shoot apex, 2gg expression was very high and remained high into the next developmental stage (1-2-mm floral bud). The next two, *CsB10dhB2\_148* and *CsB10dhB2\_317*, had very similar expression patterns in their leaves and shoot apices: their expression was at the same level, but the 1-2 mm floral buds had a higher ex-

pression in the B10 line. The expression patterns of *CsB10dhB2\_93* and *CsB10dhB2\_94* were opposite, which also occurred in the pair *CsB10dhB2\_148* and *CsB10dhB2\_317*.

### Identification and annotation of genes

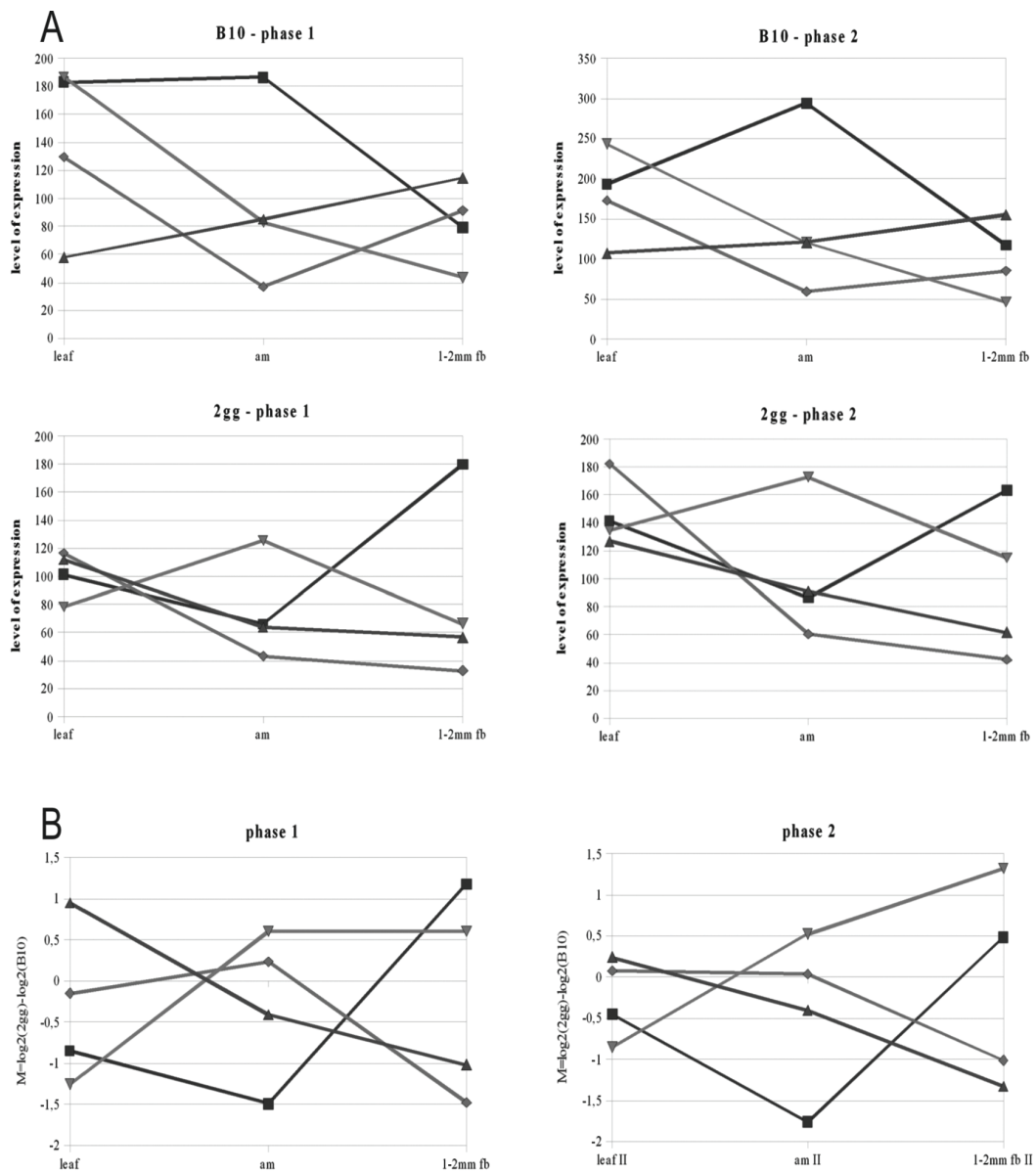
To identify selected sequences in the cucumber genome, a blastn/tblastx on the cucumber genome database was performed. As a result of a macroarray analysis in which the ph1 and ph2 experiment using a rigorous normalization method were compared, four clones i.e., *CsB10dhB2\_93*, *CsB10dhB2\_94*, *CsB10dhB2\_148*, and *CsB10dhB2\_317* were identified for further analysis. The final candidates for the cucumber genome were then compared on the basis of high similarity, e-values, and bit scores. *CsB10dhB2\_93* appeared homologous to gene no. 1, which is on the 4116 contig, *CsPSTK1* (putative kinase serine/threonine). The gene was mapped on chromosome 5 at 68,065 cM (~6.5 cM from the *M/m* gene), on the 1042 scaffold. The gene was 2707-bp long and was located on the sense DNA strand. The open reading frame was 1104-bp long, and encoded a 367-aa-long protein with a predicted molecular mass of 40.78 kDa (Fig. 3).

### Cis-acting elements in the promoter region of *CsPSTK1*

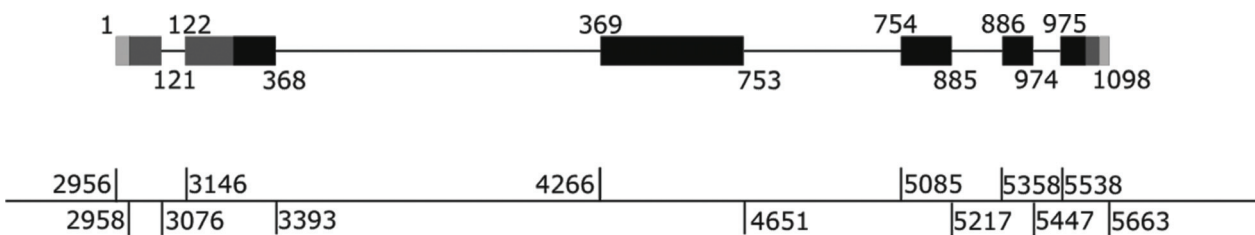
An analysis of cis-regulatory elements along the 1000-bp upstream promoter region of *CsB10dhB2\_93* using PlantCare software, revealed many elements with functions such as light and heat stress responsiveness, regulation of transcription levels, auxin and salicylic acid responsiveness, expression in endosperm and circadian control. Common promoter elements and enhancer regions as well as protein binding site were also identified. The detail cis-elements are presented in Table1.

### Bioinformatic analysis

The Blast analysis was performed for the following four clones: *CsB10dhB2\_93*, *CsB10dhB2\_94*, *CsB10dhB2\_148*, and *CsB10dhB2\_317*. The e-values, presented in paranthesis, indicate the evidence for homology. *CsB10dhB2\_93* (0.0) is homologous to the receptor-like serine/threonine (dual)-protein kinase ALE2 from *Vitis vinifera* (GI:100259692), and Jpred predicted homology ( $5e^{-57}$ ) to the BRASSINOSTEROID INSENSITIVE 1-associated receptor kinase 1 (AtBAK1; BRI1-associated recep-



**Fig. 2.** The expression profiles of genes with sequences homologous to sequences of four clones selected by normalization on macroarrays and a comparative analysis between phase 1 (ph1) and phase 1 (ph2). Analysis was performed for three organs: leaves, shoot apex and floral buds at a 1-2 mm long stage, of two cucumber lines B10 and 2gg. A) expression within one line B10 and 2gg, respectively;  $\blacklozenge$  DHB10\_093,  $\blacksquare$  DHB10\_094,  $\blacktriangle$  DHB10\_148,  $\times$  DHB10\_317, B) the differential expression between two lines;  $\blacksquare$  CsB10DHB2\_93,  $\blacklozenge$  CsB10DHB2\_94,  $\blacksquare$  CsDHB2\_148,  $\blacktriangle$  CsDHB2\_317



**Fig. 3.** The structure of *CsPSTK1* gene. The dark gray squares represent exons with a light gray part as start and stop codons, the black region is Pkinase domain and the straight lines are introns. The lower graph shows the alignment of the gene on contig 4116

**Table 1.** Results of the upstream cis-elements search in Plant CARE software for CsB10dhB2\_93 promoter

Motive name (No.*)	Function
TGA-element (1)	auxin-responsive element
5UTR Py-rich Stretch (2)	cis-acting element conferring high transcription levels
HSE (2)	cis-acting element involved in heat stress responsiveness
TCA-element (1)	cis-acting element involved in salicylic acid responsiveness
Skn-1_motif (1)	cis-acting regulatory element required for endosperm expression
Circadian (3)	cis-acting regulatory element involved in circadian control
CAAT-box (22)	common cis-acting element in promoter and enhancer regions
TATA-box (45)	core promoter element around – 30 of transcription start
AACA_motif (1)	involved in endosperm-specific negative expression
as-2-box (1)	involved in shoot-specific expression and light responsiveness
Box III (1)	protein binding site
GA-motif (1)/ GATA-motif (1)/ I-box (1)/ 3-AF1 binding Site (1)/Box I (1)/ GT1-motif (3)/ Sp1 (1)/ ACE (1)/ ATCC-motif (1)/ ATCT-motif (2)	light responsive element
AAGAA-motif (2)/ AC-II (1)/ Unnamed_4 (8)	unknown

\*the number of the repetitions of a given motive is in paranthesis

tor kinase) from *A. thaliana* (GI:29427920). According to gene ontology classification, the gene was assigned to the following categories: 1) cellular component: endosome and plasma membrane; 2) molecular function: protein kinase activity and 3) biological process: defense response to microorganisms, protein phosphorylation; cell death and growth, brassinosteroid-mediated signaling pathway.

For the remaining clones, the e-values indicate homology is as follows: *CsB10dhB2\_94* ( $4e^{-04}$ ) to unknown protein from *Zea mays* (ACR38164.1); *CsB10dhB2\_148* ( $7e^{-20}$ ) – *Fagus sylvatica* mRNA for ethylene receptor 1, ers1 (AJ420194.1); and *CsB10dhB2\_317* ( $2e^{-27}$ ) to chloride channel form *Oryza sativa* OsCLC (AB069968.1). For further protein analysis, the *CsB10dhB2\_93* was selected as its full cDNA sequence was found on genome map and amino acid sequence was predicted from it. The other sequences, due to the poor predictability obtained by blast analysis as well as the lack of cucumber genes and protein sequences, were not included for protein investigation.

### Protein secondary structure predictions

We used the JPred and GOR V algorithms to predict *CsB10dhB2\_93* secondary structure (Fig. 4). Predictions from both methods had a moderate match with the ex-

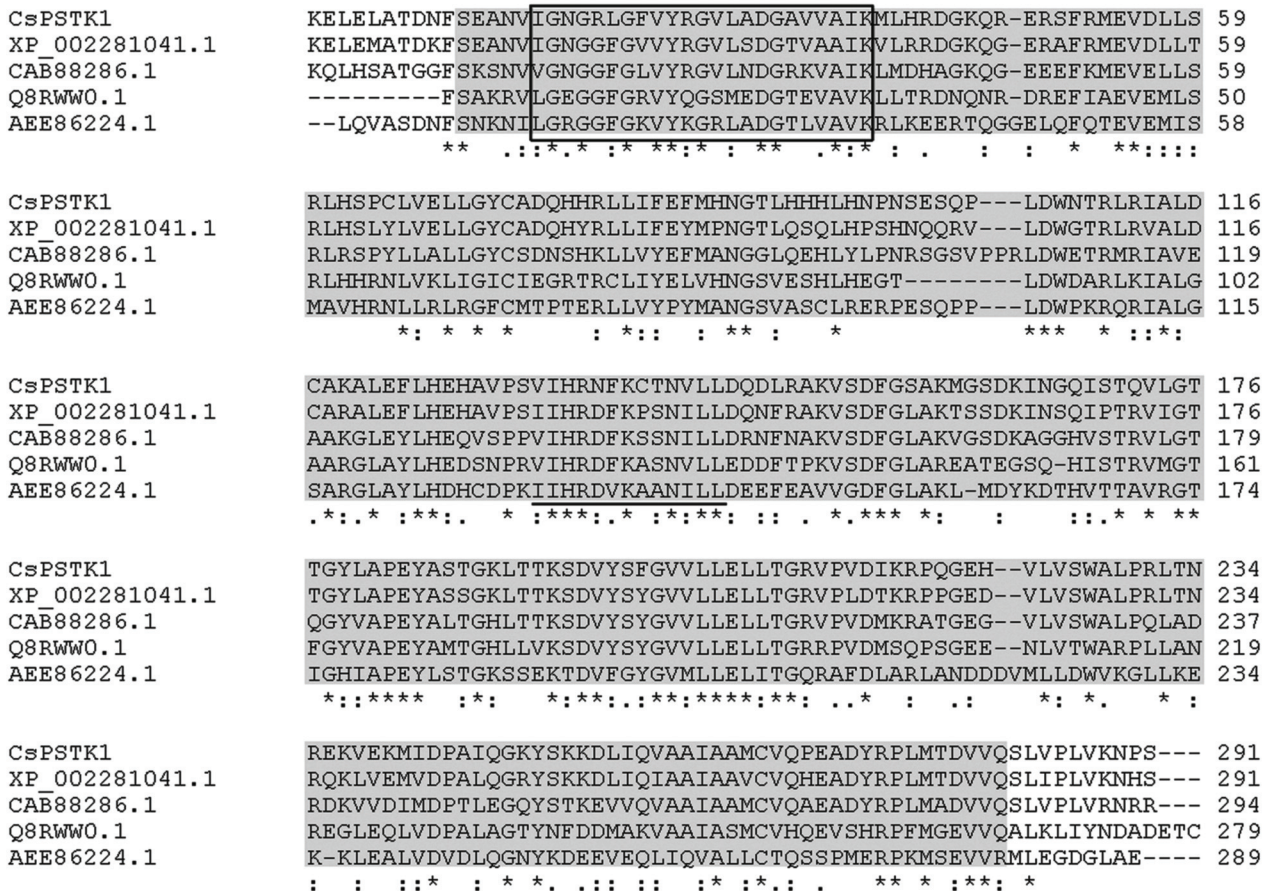
perimental data. In some parts of the sequence, the following differences in prediction were observed. Helical/extended structure differences occurred five times. Some of the structures were present in one prediction, but not in the others (four times). Ambiguities were also seen in the domain regions and at the serine/threonine binding site. This might have been due to different algorithms and prediction methods. Despite this, we observed that some predictions overlapped, indicating that further analysis is needed to comprehend the protein structure.

A bioinformatic analysis showed the presence of a serine/threonine protein kinase domain, which accounted for more than 75% of the protein and was located between the 79th and 356th amino acid residues. The presence of that domain was confirmed by three separate analyses (Prosite, Pfam, InterProScan). CsPSTK1 lacked a cleavage site and a signaling peptide (SignalP). The TmHMM algorithm revealed that CsPSTK1 is not a trans-membran protein, next TargetP and WolfPsort showed its putative location to be in the mitochondria. CsPSTK1 and four sequences were selected with the best hit in BLAST for multiple, sequence alignments, which were performed by CLUSTALW2 (Larkin, 2007) (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>). Large con-





**Fig. 4.** The secondary structure: CsPSTK1 secondary structure prediction ranging from M1 to N367. The first row figures represent the consensus JPred structure prediction while the second row figures represent GOR V prediction. “E” stands for the extended and “H” for the helical structure, “-” stands for the others



**Fig. 5.** The ClustalW alignment of five proteins. XP\_002281041.1 from *V. vinifera*, CAB88286.1, Q8RWW0.1, AEE86224.1 from *A. thaliana*. The gray region stands for Pkinase domain, the underlined sequence is the enzyme active site predicted by InterProScan and the black box shows the ATP binding site

served region were identified, of 278 amino acid residues long, of Pkinase domain. Since N-termini and C-termini are less conserved, they were excluded from the selected sequences. A comparison of domains is shown in Figure 5.

## Discussion

Identification of the mechanism of plant sex determination and flower morphogenesis that leads to genetic diversity is of great importance and interest to many research groups. Flower evolution includes a variety of processes carried out in accordance to the genetic code and regulated at multiple levels of organization by both endogenous and exogenous factors. This study focused on molecular identification of cucumber sequences that are involved in flower sex morphogenesis. Rapid, whole genome sequencing tools, together with computational software, allowed us to study the proteins involved in the basic cellular responses and processes of organogenesis (flower morphogenesis). The goal of the study was to identify the genes that are differentially expressed between lines. We wanted to identify genes which are not influenced by exogenous stimuli but are correlated with sex determination genes in cucumber. Such genes should have constant and at least 2-fold change in expression between the cucumber lines during the whole plant growth (in ph1 and ph2) and in three organs (leaves, shoot apex, floral buds). Four genes that differentiated expression patterns within the two investigated lines, B10 and 2gg were selected.

For *CsB10dhB2\_93* sequence, *CsPSTK1* gene was identified (putative serine/threonine kinase). The remaining clones did not show significant similarity to predicted *in silico* genes in cucumber genomes. This may have been caused by the absence of those sequences in the genome draft (Wóycicki, 2011). The identified clones need further investigation.

A detailed description of *CsPSTK1* revealed that this protein is a putative kinase serine/threonine kinase. In the literature they are correlated with female flower formation and sterility in *Arabidopsis* (Tanaka, 2007). This protein was also connected to the floral regulation in cucumber through ethylene biosynthesis, via brassinosteroids (BR) regulation. Additionally, in cucumber, there are several processes connected with sex determination and flower morphogenesis and some of them are

regulated by serine/threonine kinase such as cell proliferation, programmed cell death (apoptosis) or cell differentiation (Hao, 2003; Delorme, 2004, Champion, 2004). *CsPSTK1* gene shows high similarity to *ALE2* protein kinase. The investigation with *Arabidopsis* showed that the mutant-designed, abnormal leaf shape 2 (*ale2*) had epidermal effects resulting in irregular leaf morphology. At the reproductive stage, *ale2* plants generated malformed and fused floral organs. In addition, the *ale2* plants were sterile, mainly due to defects in the female reproductive tissue, although male fertility was also reduced (Tanaka et al., 2007). This study clearly shows that this gene may be correlated with flower morphogenesis due to its differential expression between the male and female organs of cucumber B10 and 2gg lines. In addition, the 2gg line is sterile, with only female flowers on the plant. For the B10 line, in ph1, male flowers occurred with undeveloped carpels, and in ph2, female flowers occurred with reduced stamens. The *CsPSTK1* was strongly expressed in the leaves and shoot apex of B10, in both male ph1 phase and in female ph2 phase. Instead in the floral buds the expression was dramatically reduced. In case of the 2gg line, expression of *CsPSTK1* was opposite to this in 2gg line: in leaves and shoot apex it was rather low but in floral bud it appeared high. Interestingly in later stages, after floral buds had developed, *CsPSTK1* transcripts reappeared in shoot apex of the 2gg line. In addition, the ovule of *Arabidopsis* mutant, *ale2*, was different from its wild type plants. In this case, ovule surfaces were disorganized and possessed irregularly orientated cross walls and deformed cells. The mature ovules contained degenerated embryo sacs, suggesting that the sterility of *ale2* plants was a result of the failure of ovules to develop (Tanaka, 2007). A further histological analysis is needed in cucumber NILs to determine the detailed organization of the developing flower buds.

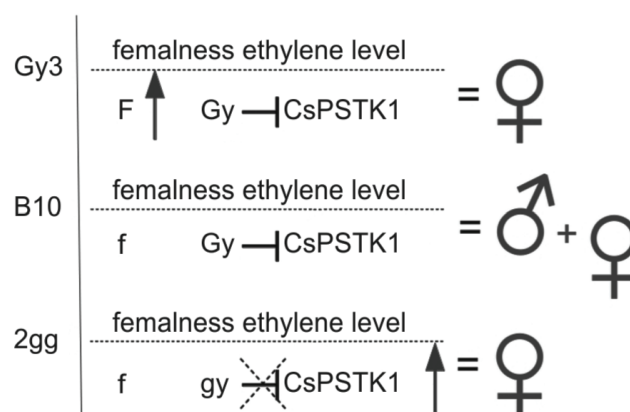
The *CsB10dhB2\_93* gene showed similarity to gene which encodes BRASSINOSTEROID INSENSITIVE 1-associated receptor kinase 1. Brassinosteroids, including BRI1 and BAK1, are perceived by the cell surface receptor, kinase complex (Tang, 2010). Brassinosteroids are plant hormones that increase ethylene production and also interact with each other (Hansen et al., 2010). Ethylene is regarded as essential in sex determination in cucumber, (Li et al., 2012). The *F* and *M* genes responsible for sex determination are the *CsACSIG* and



*CsACS2* genes, respectively. *ACS* genes are transcriptionally regulated during development as well as in response to environmental stimuli (Wang et al., 2005). Recent evidence suggests that an important component of the regulation of ACS and the biosynthesis of ethylene occurs post-transcriptionally, by stabilization of ACS proteins (Argueso et al., 2007). Several hormones, including auxins, cytokinins, and BR, are known to increase the biosynthesis of ethylene (Arteca, 2008). BR is perceived by a BRI1 receptor, a leucine-rich receptor-like kinase that can dimerize with and phosphorylate BAK1 (Li and Jin, 2007). There may be a correlation between *CsPSTK1*, and BAK1 receptor in BR signaling pathway that influences ethylene by regulating *ACS* genes in cucumber. Our results confirm the above reported findings. We found that *CsPSTK1* expression in female floral buds of 2gg line was higher than in male buds of B10 line. BR may regulate ethylene production which influences female flowers in the 2gg line. The expression of this gene, however, is stable within the two phases of the B10 line (phase 1 of only male flowers and phase 2 with male and female flowers but with an excess of males). We believe that in B10 line, the expression in floral buds in ph2 should be altered to display a pattern similar to that in 2gg, but remained at the same level. This is possible because the monoecious line in ph2 possesses more male flowers than female flowers. The percentage of male and female buds during the collection of plant materials could not be verified; however, male floral buds were predominant. B10 and 2gg lines have no *F* gene; therefore, it is speculated that ethylene production in those lines is regulated differently than in the lines that possess this gene. An increase in gene expression of *CsPSTK1* in the female 2gg line (recessive *ff*) increased the synthesis of ethylene in 1-mm floral buds, possibly initiating female flower formation. Increasing the expression of *CsPSTK1* would thereby be the same as having *F* gene (which encodes ACC synthase) present; however, in buds of the monoecious B10 line with a large predominance of male flowers, no increase in *CsPSTK1* expression was observed.

*CsPSTK1* activity may be associated with the dominant allele of *Gy* gene, since *CsPSTK1* in line 2gg is linked to the recessive allele *gy* (Kubicki, 1974). We believe this regulation is negative (Fig. 6). Brassinosteroids that take part in the above process are likely influenced directly by the female flower formation. The *Gy*

may inhibit the *CsPSTK1*, but in a mutant 2gg, the inhibition is removed and *CsPSTK1* has a positive effect as it increases the level of ethylene and promotes female flower production. Cucumber species, however, have different types of females in Gy3 line (*AAFFMMGyGy*) (Przybecki et al., 2003); therefore, ethylene production may be mainly regulated by *F* gene and the presence of *Gy* may inhibit *CsPSTK1* (Fig. 6).



**Fig. 6.** A hypothetical scheme of the correlation between the *F*, *Gy* and *CsPSTK1* and their effect on ethylene production and sex expression in three different genotypes of cucumber. The “ $\bar{}$ ” - the inhibition of *CsPSTK1* by dominant *Gy* gene; while - “ $\uparrow$ ” represents the regulation of ethylene production

BR appears to be closely linked to the flowering time regulation (Li et al., 2010) by repressing the expression of the Flowering Locus C (Domagalska et al., 2007). Another well-known defect in BR biosynthetic and signaling mutants is male infertility, but the BR mechanism that regulates male fertility is unknown. A recent study combined microscopy analysis with transcript profiling to show that many of the well-studied genes that regulate anther and pollen development have altered the expression in BR mutants, leading to observable defects in anther and pollen morphology (Clouse, 2011). A novel role for BRs in gynoecium and ovule development has been recently identified using genetic screens that identified a BR biosynthetic enzyme as a suppressor of mutations in the *Arabidopsis* SEUSS gene, encoding a transcriptional adaptor protein that is involved in floral organ identity and the development of the carpel margin meristem (Nole-Wilson et al., 2010). The promotion of cell expansion through BR-regulated expression of cell wall modification genes, the ion and water transport, and

cytoskeleton rearrangements are well documented in individual genetic and global microarray analyses (Kim and Wang, 2010). Like auxin and cytokinin, BRs are not merely involved in simple cell elongation, but also modulate the timing of cell division and differentiation within complex processes such as organ formation (Clouse, 2011). Furthermore, brassinosteroids and photomorphogenesis are connected by an antagonistic interaction between BR and light that targets common downstream genes (Sun et al., 2010). BR can also directly regulate important light signaling components (Luo et al., 2010).

The *CsPSTK1* gene possesses many cis-acting elements in the promoter region that are involved in light response, auxin, and hormone signaling. It is known that sex determination in plants may be modulated by light, or that auxin may trigger ethylene accumulation to cause the formation of the female flower (Liu et al., 2012).

Interestingly, the *CsB10dhB2\_148* sequence shows similarity to the sequence of mRNA of ethylene receptor 1-ERS1 in *Fagus sylvatica*. In *Arabidopsis* research, overexpression of gene *ERS1* resulted in elevated phenotypic ethylene responses including growth inhibition and early senescence. It also represses ethylene responses and has inhibitory effects on ETR1-specific ethylene signaling (Liu 2010). The expression of this gene was very high in the gynoeocious 2gg line (especially in the shoot apex and floral buds) in comparison to the monoecious B10 line. In Yamasaki (2003), *ERS* gene expression was up-regulated by ethylene in the shoot apices of monoecious and gynoeocious, but not andromonoecious, cucumber. This suggests that ethylene induction of *CS-ERS* expression is mediated at the *M* locus. The expression of *CS-ERS* in both monoecious and andromonoecious cucumber thereby reflects the level of ethylene that is produced by *CS-ACS2* through mediation from a product of the *M* locus (Yamasaki 2003).

*CsB10dhB2\_94* showed similarity to the chloride channel, displaying a variety of important physiological and cellular roles including regulation of pH, volume homeostasis, cell transport, cell migration, cell proliferation, and cell differentiation (Nakamura, 2006). High expression of this gene occurred in B10 buds, but we could not connect its expression to a cell differentiation function within flower organ morphogenesis. The connection between the above *CsPSTK1* gene (clone *CsB10dhB2\_93*), which codes putative protein kinases, may interact with ion channels.

The least is understood about the *CsB10dhB2\_317* clone, whose similarity to three unknown protein form *Zea mays* is low. It appears to be a novel protein involved in flower morphogenesis and sex determination in cucumber. We suppose that *CsB10dhB2\_93*, *CsB10dhB2\_94*, *CsB10dhB2\_148*, and *CsB10dhB2\_317* are all correlated with flower morphogenesis in cucumber plants. To prove this, future studies expanding our work, the whole genomic sequence in the 2gg line are needed. This would help identify the suspected mutation site in either the coding sequence or the promoter sequence that causes the differential expression detected in this study.

A genome-wide overview of cucumber clones revealed *CsPSTK1* as the gene that encodes protein serine/threonine kinase. We conclude that the identified protein kinases may be an alternative way for hormonal signal transduction in flowers of opposite sex, taking part in the inhibition of unwanted generative organs that cause development of a unisex flower.

### Acknowledgements

This research was supported by grants from the Ministry of Science and Higher Education (MNiSW) - N302 3633 33 and the National Science Center - 2011/01/B/NZ2/01631.

### References

- Argueso C.T., Hansen M., Kieber J.J. (2007) *Regulation of ethylene biosynthesis*. J. Plant Growth Regul. 26: 92-105.
- Arteca R.N., Arteca J.M. (2008) *Effects of brassinosteroid, auxin, and cytokinin on ethylene production in Arabidopsis thaliana plants*. J. Exp. Bot. 59(11): 3019-3026.
- Atsmon D., Lang A., Light E.N. (1968) *Contents and recovery of gibberellins in monoecious and gynoeocious cucumber plants*. Plant Physiol. 43(5): 806-810.
- Bai S.L., Peng Y.B., Cui J.X., Gu H.T., Xu L.Y., Li Y.Q., Xu Z.H., Bai S.N. (2004) *Developmental analyses reveal early arrests of the spore-bearing parts of reproductive organs in unisexual flowers of cucumber (Cucumis sativus L.)*. Planta 220(2): 230-240.
- Borodovsky M., Lomsadze A. (2011a) *Gene identification in prokaryotic genomes, phages, metagenomes, and EST sequences with GeneMarkS suite*. Curr. Protoc. Bioinformatics Chapter 4: Unit 4.5.1 – 17.
- Borodovsky M., Lomsadze A. (2011b) *Eukaryotic gene prediction using GeneMark.hmm-E and GeneMark-ES*. Curr. Protoc. Bioinformatics Chapter 4: Unit 4.6.1 – 10.
- Boualem A., Troadec C., Kovalski I., Sari M.A., Perl-Treves R., Bendahmane A. (2009) *A conserved ethylene biosynthesis enzyme leads to andromonoecy in two cucumis species*. PLoS One 4(7): e6144.
- Cavagnaro P.F., Senalik D.A., Yang L., Simon P.W., Harkins T.T., Kodira C.D., Huang S., Weng Y. (2010) *Genome-*

- wide characterization of simple sequence repeats in cucumber (*Cucumis sativus* L.). *BMC Genomics*. 11: 569.
- Champion A., Kreis M., Mockaitis K., Picaud A., Henry Y. (2004) *Arabidopsis* kinome: after the casting. *Funct. Integr. Genom.* 4(3): 163-187.
- Clouse S.D. (2011) *Brassinosteroid signal transduction: from receptor kinase activation to transcriptional networks regulating plant development*. *Plant Cell* 23(4): 1219-1230.
- Cole C., Barber J.D., Barton G.J. (2008) *The Jpred 3 secondary structure prediction server*. *Nucl. Acids Res.* 36(Web Server issue): W197-201.
- Dellaporta S.L., Calderon-Urrea A. (1993) *Sex determination in flowering plants*. *Plant Cell* 5(10): 1241-1251.
- Delorme V.G., McCabe P.F., Kim D.J., Leaver C.J. (2000) *A matrix metalloproteinase gene is expressed at the boundary of senescence and programmed cell death in cucumber*. *Plant Physiol.* 123(3): 917-927.
- Domagalska M.A., Schomburg F.M., Amasino R.M., Vierstra R.D., Nagy F., Davis S.J. (2007) *Attenuation of brassinosteroid signaling enhances FLC expression and delays flowering*. *Development*. 134(15): 2841-2850.
- Emanuelsson O., Nielsen H., Brunak S., von Heijne G. (2000) *Predicting subcellular localization of proteins based on their N-terminal amino acid sequence*. *J. Mol. Biol.* 300: 1005-1016.
- Guo S., Zheng Y., Joung J.G., Liu S., Zhang Z., Crasta O.R., Sobral B.W., Xu Y., Huang S., Fei Z. (2010) *Transcriptome sequencing and comparative analysis of cucumber flowers with different sex types*. *BMC Genom.* 11: 384.
- Hansen M., Chae H.S., Kieber J.J. (2009) *Regulation of ACS protein stability by cytokinin and brassinosteroid*. *Plant J.* 57(4): 606-614.
- Hao Y.J., Wang D.H., Peng Y.B., Bai S.L., Xu L.Y., Li Y.Q., Xu Z.H., Bai S.N. (2003) *DNA damage in the early primordial anther is closely correlated with stamen arrest in the female flower of cucumber (Cucumis sativus L.)*. *Planta* 217(6): 888-895.
- Huang S., Li R., Zhang Z., Li L., Gu X., Fan W., Lucas W.J., Wang X., Xie B., Ni P. et al. (2009) *The genome of the cucumber, Cucumis sativus L.* *Nat Genet.* 41(12): 1275-1281.
- Kim T.W., Wang Z.Y. (2010) *Brassinosteroid signal transduction from receptor kinases to transcription factors*. *Ann. Rev. Plant Biol.* 61: 681-704.
- Kloczkowski A., Ting K.L., Jernigan R.L., Garnier J. (2002) *Combining the GOR V algorithm with evolutionary information for protein secondary structure prediction from amino acid sequence*. *Proteins* 49(2): 154-166.
- Kubicki B. (1974) *New sex types in cucumber and their uses in breeding work*. *Proc. XIX Intl. Hort. Congr.* 3: 475-485.
- Larkin M.A., Blackshields G., Brown N.P., Chenna R., McGettigan P.A., McWilliam H., Valentin F., Wallace I.M., Wilm A., Lopez R. et al. (2007) *Clustal W and Clustal X version 2.0*. *Bioinformatics* 23(21): 2947-2948.
- Lescot M., Déhais P., Thijs G., Marchal K., Moreau Y., Van de Peer Y., Rouzé P., Rombauts S. (2002) *PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences*. *Nucl. Acids Res.* 30(1): 325-327.
- Li J., Jin H. (2007) *Regulation of brassinosteroid signaling*. *Trends Plant Sci.* 12: 37-41.
- Li J., Li Y., Chen S., An L. (2010) *Involvement of brassinosteroid signals in the floral-induction network of Arabidopsis*. *J. Exp. Bot.* 61(15): 4221-4230.
- Li Z., Huang S., Liu S., Pan J., Zhang Z., Tao Q., Shi Q., Jia Z., Zhang W., Chen H. et al. (2009) *Molecular isolation of the M gene suggests that a conserved-residue conversion induces the formation of bisexual flowers in cucumber plants*. *Genetics* 182(4): 1381-1385.
- Li Z., Wang S., Tao Q., Pan J., Si L., Gong Z., Cai R. (2012) *A putative positive feedback regulation mechanism in CsACS2 expression suggests a modified model for sex determination in cucumber (Cucumis sativus L.)*. *J. Exp. Bot.* 63(9): doi: 10.1093/jxb/ers123.
- Liu Q., Xu C., Wen C.K. (2010) *Genetic and transformation studies reveal negative regulation of ERS1 ethylene receptor signaling in Arabidopsis*. *BMC Plant Biol.* 10: 60 doi:10.1186/1471-2229-10-60
- Luo X.M. et al. (2010) *Integration of light- and brassinosteroid – signaling pathways by a GATA transcription factor in Arabidopsis*. *Dev. Cell* 19: 872-883.
- Malepszy S., Niemirowicz-Szczytt K. (1991) *Sex determination in cucumber (Cucumis sativus) as a model system for molecular biology*. *Plant Sci.* 80: 39-47.
- Mibus H., Tatlioglu T. (2004) *Molecular characterization and isolation of the F/f gene for femaleness in cucumber (Cucumis sativus L.)*. *Theor. Appl. Genet.* 109(8): 1669-1676.
- Nakamura A., Fukuda A., Sakai S., Tanaka Y. (2006) *Molecular cloning, functional expression and subcellular localization of two putative vacuolar voltage-gated chloride channels in rice (Oryza sativa L.)*. *Plant Cell Physiol.* 47(1): 32-42.
- Nole-Wilson S., Rueschhoff E.E., Bhatti H., Franks R.G. (2010) *Synergistic disruptions in seuss cyp85A2 double mutants reveal a role for brassinolide synthesis during gynoecium and ovule development*. *BMC Plant Biol.* 10: 198 doi: 10.1186/1471-2229-10-198.
- Perl-Treves R. (1999) *Male to female conversion along the cucumber shoot: approaches to studying sex genes and floral development in Cucumis sativus*. *Sex determination in plants*. BIOS Scientific Publishers Ltd., pp. 189-216.
- Petersen T.N., Brunak S., von Heijne G., Nielsen H. (2011) *SignalP 4.0: discriminating signal peptides from transmembrane regions*. *Nat. Meth.* 8(10): 785-786.
- Przybecki Z., Kowalczyk M.E., Siedlecka E., Urbańczyk-Wochniak E., Malepszy S. (2003) *The isolation of cDNA clones from cucumber (Cucumis sativus L.) floral buds coming from plants differing in sex*. *Cell Mol. Biol. Lett.* 8(2): 421-438.
- Przybecki Z., Kowalczyk M.E., Witkiewicz J., Filipecki M., Siedlecka E. (2004) *Polymorphism of sexually different cucumber (Cucumis sativus L.) NIL lines*. *Cell Mol. Biol. Lett.* 9(4B): 919-933.

- Sun Y., Fan X.Y., Cao D.M., Tang W., He K., Zhu J.Y., He J.X., Bai M.Y., Zhu S., Oh E. et al. (2010) *Integration of brassinosteroid signal transduction with the transcription network for plant growth regulation in Arabidopsis*. *Dev. Cell* 19: 765-777.
- Tanaka H., Watanabe M., Sasabe M., Hiroe T., Tanaka T., Tsukaya H., Ikezaki M., Machida C., Machida Y. (2007) *Novel receptor-like kinase ALE2 controls shoot development by specifying epidermis in Arabidopsis*. *Development* 134: 1643-1652.
- Tang W., Deng Z., Wang Z.Y. (2010) *Proteomics shed light on the brassinosteroid signaling mechanisms*. *Curr. Opin. Plant Biol.* 13(1): 27-33.
- Terefe D., Tatlioglu T. (2005) *Isolation of a partial sequence of a putative nucleotide sugar epimerase, which may involve in stamen development in cucumber (Cucumis sativus L.)*. *Theor. Appl. Genet.* 111(7): 1300-1307.
- Trebitsh T., Staub J.E., O'Neill S.D. (1997) *Identification of a 1-aminocyclopropane-1-carboxylic acid synthase gene linked to the female (F) locus that enhances female sex expression in cucumber*. *Plant. Physiol.* 113(3): 987-995.
- Wang N., Shih M., Li N. (2005) *The GUS reporter-aided analysis of the promoter activities of Arabidopsis ACC synthase genes AtACS4, AtACS5, and AtACS7 induced by hormones and stresses*. *J. Exp. Bot.* 56: 909-920.
- Wóycicki R., Witkowicz J., Gawroński P., Dąbrowska J., Lomsadze A., Pawelkowicz M., Siedlecka E., Yagi K., Pląder W., Seroczyńska A. et al. (2011) *The genome sequence of the North-European cucumber (Cucumis sativus L.) unravels evolutionary adaptation mechanisms in plants*. *PLoS One* 6(7): e22728; doi:10.1371/journal.pone.0022728
- Wu T., Qin Z., Zhou X., Feng Z., Du Y. (2010) *Transcriptome profile analysis of floral sex determination in cucumber*. *J. Plant. Physiol.* 167(11): 905-913.
- Yamasaki S., Fujii N., Takahashi H. (2003) *Photoperiodic regulation of CS-ACS2, CS-ACS4 and CS-ERS gene expression contributes to the femaleness of cucumber flowers through diurnal ethylene production under short-day conditions*. *Plant, Cell Environ.* 26: 537-546.
- Yamasaki S., Fujii N., Takahashi H. (2005) *Hormonal regulation of sex expression in plants*. *Vitam Horm.* 72: 79-110.