

CELLULAR & MOLECULAR BIOLOGY LETTERS

Volume 9, (2004) pp 919 – 933

<http://www.cmbi.org.pl>

Received 30 August 2004

Accepted 15 November 2004

POLYMORPHOM OF SEXUALLY DIFFERENT CUCUMBER (*Cucumis sativus* L.) NIL LINES

ZBIGNIEW PRZYBECKI, MAGDALENA EWA KOWALCZYK, JUSTYNA WITKOWICZ, MARCIN FILIPECKI and EWA SIEDLECKA

Department of Plant Genetics, Breeding and Biotechnology, Faculty of Horticulture, Warsaw Agricultural University, Nowoursynowska 166, 02-787 Warsaw, Poland

Abstract: Isolations of polymorphic sequences of two pairs of the NIL lines of cucumber (*Cucumis sativus* L.), which differ with respect to sex, were carried out using the subtraction hybridization methods of DSC (Differential Subtraction Chain) and GDDSC (Genetically Directed DSC). 266 DSC tags were isolated from the entire genome region, and 38 GDDSC tags were isolated from the region containing the sex genes. Based on the obtained results, the methods used may be considered highly effective. The attained sequences, like 11 AFLP clones obtained earlier [Witkowicz, J. *et al.* Cell. Mol. Biol. Lett. **8** (2003) 375-381], were characterized by analyzing their hybridization with differential (dhaom) and subtractive cDNA libraries (cDNAsubtractom) from 1- to 2- mm floral buds of the same lines, and by the sequencing of 28 tags. A high average degree of homology was found to exist in the genpolom to dhom and cDNAsubtractom, particularly in the case of "dominant" (when the tester used was a line in which the sex of the plants was dependent upon the dominant allele). This indicates a significant share of coding sequences in the polymorphic genomic tags as well as their share in flower formation. Many of these sequences originate from the sex gene region. Analysis of the sequenced tags showed their interesting composition, including many organelle sequences which transferred into the nucleus, and coding sequences that may participate in flower development, including sex formation.

Key Words: Cucumber, *Cucumis sativus* L., Sex Expression, Sex Determination, Flower Development, Subtraction, DSC, GDDSC, Tags, Omes

INTRODUCTION

Closely related forms, including those that isogenically and homozygotically differ essentially by a single trait (NIL – near isogenic lines), actually differ in many regions of the genome, at least at the level of the DNA [1, 2]. Variations at many genome sites are also demonstrated by allelic forms, with respect to a morphological trait that is inherited through a single gene, such as sex of the plant, when one of these forms has originated from the other by way of chemical mutagenesis, followed by selection and roughly 30 years of self-pollination [2, 3]. If such lines differ morphologically by only a single trait, but they vary at many locations at the DNA level, a number of questions arise, such as: how big is that variation? how are these differences distributed in the genome? what is the function of the sequences in which the differences are? and is their function connected with the trait that differentiates the phenotypes of these NIL lines? The results presented here are elements of research studies aimed at providing answers to at least some of the above questions.

The aim of this study was the isolation and partial identification of genome-differentiating sequences in the closely related cucumber genotypes mentioned above.

MATERIALS AND METHODS

According to the ‘omes’ convention [4] we would like to define some of the terms used herein. ‘Polymorphom’ is the set of differences between compared objects, and we denote it with the abbreviation ‘polom’. Subsets of polomes are ‘genpolom’, which is the abbreviation for genomic polymorphom, and ‘cDNApolom’ means cDNA polymorphom. Subsets of differences can be also distinguished using their generation criteria, for example mutation (‘mutabolom’) or recombination (‘recombinome’) or using their method of detection, such as subtraction (‘subtractom’). From the latter using the genomic subtraction method gives the ‘gensubtractome’, the genomic subtractive DSC method gives ‘dsubtractom’, GDDSC gives ‘gdssubtractom’, and cDNA-DSC gives ‘cDNAsubtractom’. When we use the differential hybridization method to screen cDNA libraries, we get ‘dhome’. Other terms like ‘AFLPome’ are similarly logically derived.

Plant material

The plant material consisted of two pairs of near isogenic lines (NIL) that differ with regard to sex. One line pair has different alleles at loci *m*, *Gy3* (a dominant female line with an *MMFFGyGy* genotype), and *HGy3* (an isogenic hermaphroditic line with an *mmFFGyGy* genotype). The other pair of lines has differing alleles at the loci *Gy*, *B10* (a monoecious Borszczagowski line with an *MMffGyGy* genotype), and *2gg* (a recessive female with an *MMffgygy* genotype isogenic to *B10*). Their hybrid F_1 generations, and the segregating F_2 generations were also used.

Isolation of genomic DNA and AFLP

Isolation of DNA and AFLP in the above-mentioned population was carried out in accordance with methods presented earlier [2].

Subtractive methods

Subtractions were carried out between the mentioned NIL line pairs using the DSC method in accordance with Luo *et al.* [5], as well as between bulks of 20 F₂ individuals of the above-mentioned lines using the GDDSC method. The GDDSC method is GDRDA based [6]: a genetically-directed, DSC subtraction method. The trait directing the selection and subtraction in GDDSC was the sex of the plants, while the bulks were created for the traits (sex) inherited in a dominant (dominant female sex and monoeciousness) and recessive (hermaphroditicity and recessive female sex) manner. The subtractions (DSC and GDDSC) were carried out in both directions; each of the forms used for the subtractions served as the tester and then as the driver. In order to simplify the description, the terms “recessive” or “dominant” library have been adopted, depending on whether the tester was the population carrying the recessive or the dominant allele. DSC (including for the GDDSC) was modified; the creation of representations, the conditions for the DNA probe hybridization, and the amplification algorithm were applied as in RDA [1], with the exception of the number of cycles for the subtraction products, which was 43, in accordance with Luo *et al.* [5]. Restriction enzymes (Tab. 1), adapters and primers for DSC were as per Lisitsyn *et al.* [1]. The GDDSC subtractions were carried out using only the *Bgl* *g*III enzyme. The ratio of the driver DNA to the tester in the case of the DSC was 100:1, and the number of subtraction rounds was 3-4, while in the case of GDDSC, the ratio of driver DNA to the tester was 1:1, and the number of subtraction rounds was 4-6. The subtraction products were separated by electrophoresis on agarose gels and PAA, and cloned.

Cloning and creation of tag libraries (subtractomes)*DSC (d) and GDDSC (gd) subtractomes*

The products of each subtraction were cloned into the pCRII-TOPO vector, using an Invitrogen TOPO Cloning Kit TA according to the manufacturer procedure. This resulted in subtractive libraries, here called subtractomes. These libraries are stored as glycerol bacteria stocks at -75°C [7]. As many subtraction libraries were created as there were DSC subtractions performed between the lines, and as there were GDDSC subtractions performed between the bulks of the F₂ generations.

Libraries of the clones (tags) of the DSC and GDDSC

In order to create the clone libraries, the subtraction libraries were plated as per Sambrook *et al.* [7]. Bacterial colonies were randomly isolated, and the inserts were identified by cutting them out of the plasmids using the *Eco*RI restriction enzyme. The lengths of the inserts were compared with the lengths of the

fragments following the last round of subtraction, and then stored as glycerol stocks at -75°C and as isolated plasmids, according to [7] at -20°C .

Libraries of the AFLP(af) clones – AFLPomes

The bands differentiating the above-mentioned lines were cut out of the PAA gel, placed into sterile water and stored at -20°C . For cloning, the isolated tags were amplified and then cloned in the manner presented above.

Dot-blot hybridization

The obtained subtractomes (d and gd) and the individual clones from AFLPome [2] were transferred onto membranes and hybridized using the dot-blot method as per Urbańczyk-Wochniak *et al.* [8]. cDNA subtractome and dhome were used as the probes. The cDNA subtractome were previously constructed by subtraction of cDNA cut with the *MboI* restriction enzyme (cDNA libraries) [9]. The dhome (libraries of differential cDNA clones – DH libraries) were obtained as a result of differential hybridization [9, 10] of the cDNA libraries of 1- to 2-mm floral buds of the lines used in this study.

RESULTS

Subtractive analysis

DSC

As a result of the three DSC subtractions (created using three different restriction enzymes) carried out between the utilized NIL cucumber lines, we obtained dsubtractom, consisting of a number of DSC differences (DSC tags), presented in Tab. 1. Differences in the effectiveness of subtraction may be observed, depending on the restriction enzyme used, and expressed as the number of the obtained tags (differences). The greatest number of differences was obtained using the *BglIII* restriction enzyme, with the exception of subtraction 2gg_{B10}. Not far behind (at about 12%) were the results obtained using *BamHI*. The *HindIII* restriction enzyme was definitely less effective, since as much as 50% less differences were obtained than with *BglIII*.

Tab. 1. Results of the DSC subtraction of pairs of NIL lines of the cucumber. The name of the line written in normal font denotes the tester, while the name of the line written as a subscript of the tester denotes the driver.

Subtraction *tester _{driver}	Number of tags			Σ
	HindIII	BglIII	BamHI	
B10 _{2gg}	11	28	25	64
2gg _{B10}	9	24	27	60
Gy3 _{HGy3}	26	30	27	83
HGy3 _{Gy3}	11	29	19	59
Σ	57	111	98	266

Some positive impact on the effectiveness of the subtraction may also be noted. It seems to depend on its direction (Tab. 1). More fragments were obtained when the wild form served as the tester rather than the recessive (mutant) form. While the difference is small (~6%) and its statistical significance may raise some doubts in the subtractions between B10 and 2gg, the ~31% difference in the case of lines Gy3 and HGy3 seems far more certain.

GDDSC

As a result of the GDDSC subtraction between the used NIL lines, we obtained gdssubtractome, which consists of a number of fragments (GDDSC tags) differentiating the adjacent region of the sex genes in the pairs of the NIL lines presented in Tab. 2. The DNA was digested with the *Bgl*III restriction enzyme; the remaining enzymes were not used. Similarly as for the subtraction between the lines, it appears that there are more isolated tags when the trait directing the tester is dominant. Comparing the numbers of tags obtained following DSC subtraction of the lines and following GDDSC using the *Bgl*III restriction enzyme, a significant reduction in their amount may be observed in the case of the GDDSC. The GDDSC tags amount to ~34% relative to the DSC of the lines. However, the GDDSC tags only originate from the region of the directing gene, namely the gene that determines the sex of the plants.

Tab. 2. Results of the GDDSC subtraction of the F₂ hybrid pairs of NIL lines. Crossed lines are denoted by parenthesis, the remaining are as in Tab. 1.

subtraction tester _{driver} (hybrids)	Number of tags
J _{rZ} (B10x2gg)	11
rZ _J (B10x2gg)	8
dZ _H (Gy3xHGy3)	10
H _{dZ} (Gy3xHGy3)	9
Σ	38

J-bulk of the monoecious plants F₂, rZ-bulk of the recessive female plants F₂, dZ-bulk of the dominant female plants F₂, H-bulk of the hermaphroditic plants F₂.

Characteristics of the isolated sequences

The characteristic feature of the polom (DSC, GDDSC and AFLP) that differentiates the genomes of the NIL lines is that some of the clones share homology to the coding sequences of cDNA subtractome and dhome of the floral buds at the 1- to 2-mm stage obtained earlier [9, 10]. Some of the polom's tags were sequenced, including the AFLP products, which demonstrated total co-segregation with the *M* and *Gy* genes using the method of low resolution BSA [2].

Hybridization

The results of hybridization, presented in Tab. 3, point to certain regularities in its overall display. First and foremost, it should be stated that hybridization

between cDNApolom and genpolom exists and in many cases is very intensive. As indicated by the results of the hybridization of single AFLP clones, about 45% of them had their counterparts in expression libraries. Clones 27afB10 (RuBisCO large subunit) and 28afB10 (Apple developing fruit cDNA) hybridized with the biggest t intensity.

Tab. 3. Results of the hybridization of genpolome and 1- to 2-mm floral buds cDNApolom.

Genpoloms		cDNApoloms (cDNA probes)						
		cDNAsubtractom (MboI)				dhome		
	Name	Enzyme	B10 _{2gg}	2gg _{B10}	Gy3 _{Hgy3}	Hgy3 _{Gy3}	B10 _{2gg}	Gy3 _{Hgy3}
dsubtractoms and gdsbtractoms	Gy3 _{Hgy3}	BamHI	+	-	++	-	-	-
	Hgy3 _{Gy3}		++	-	+++	-	-	+
	B10 _{2gg}		-	-	-	-	-	-
	2gg _{B10}		+++	-	++++	+	-	+++
	Gy3 _{Hgy3}		+++++	+	+++++	++	-	+++++
	Hgy3 _{Gy3}	+++	++	++++	+	-	++	
	dZ _H	BgIII	+++++	+	+++++	-	-	+++++
	B10 _{2gg}		+++++	-	++++	-	-	+++++
	2gg _{B10}		+	-	+++	-	-	+
	J _{rZ}		++++	-	+++++	-	-	+++
	rZ _J		-	-	-	++	-	-
	Gy3 _{Hgy3}	HindIII	+++	-	+++++	+	-	+++
	Hgy3 _{Gy3}		-	-	-	-	-	-
	B10 _{2gg}		+	-	+	+	-	+
	2gg _{B10}		++	-	++	-	-	+
26afB10	+		-	-	-	-	-	
AFLPome (Single AFLP clones)	27afB10	++++	-	++	-	-	++	
	28afB10	-	-	+++	-	+	+++	
	29afB10	-	-	-	-	-	-	
	31afB10	-	-	-	-	-	-	
	32afGy3	-	-	-	-	-	-	
	33afGy3	-	-	-	-	-	-	
	34afGy3	-	-	-	-	-	-	
	36afGy3	+	-	-	-	-	-	
	37afGy3	-	-	-	-	-	-	
	39afGy3	-	-	-	+	-	-	

Scale of signal intensity: (+++++) very strong; (++++) strong; (+++) moderate; (++) weak; (+) very weak; As in Tabs. 1, 2 and 4.

We noticed that hybridization is far more intensive when, as in the case with expression libraries, the tester was either a line or a bulk with a dominant allele at the sex-determining loci. The only exception was the DH B10_{2gg} because there was no hybridization with genpolom. The hybridization was the strongest in the case of the subtractom d and gd in the case of BgIII. In addition, only the

above libraries hybridized with the "recessive" cDNA subtractome, i.e. when the plants with a recessively-conditioned sex type served as the tester.

The results of a hybridization of pairs of antagonistic gensubtractomes (libraries Gy3_{Hgy3}^H and the Hgy3_{Gy3}^H as well as B10_{2gg}^B and 2gg_{B10}^B) are also quite interesting, because one of the pair hybridizes very intensively, while the other not at all (Tab. 3). This situation is very similar to the J_{rZ}^{Bl} and rZ_J^{Bl} pair, the only difference being that the "recessive" gensubtractome hybridizes poorly with the one cDNA subtractome (Hgy3_{Gy3}).

Sequencing

Seventeen clones of gensubtractome (genome subtractome) and eleven clones of AFLPome (demonstrating coupling with dominant alleles of the sex genes – [2]) were sequenced (Tab. 4). The obtained sequences were aligned with BLASTNnr (designated N in Tab. 4), BLASTest (designated N:EST in Tab. 4) and BLASTX algorithm (designated X in Tab. 4). The last two databases were used with the aim to find coding sequences. The homology was considered significant when the e-value was < 0.01.

Each of the sequenced subtractive clones possessed a significant similarity to the sequence in the nucleotide database. Whereas among the AFLP clones, only two (12%) contained fragments with a high degree of homology to the organelle DNA of a cucumber, the remainder did not possess fragments with a significant resemblance to the sequences in the data bases. Among the subtractive clones, 7 possessed two or three non-overlapping fragments that demonstrated a similarity to various sequences. Among all of the found homologies, as many as 8 were of organelle origin, including two from a cucumber – one of a chloroplast and one mitochondrial. The mitochondrial sequence was highly homologous to a mutation site in mitochondrial DNA that is responsible for the mosaic type (MSC mutation [11, 12] demonstrating a pleiotrophic effect) of leaves, and constituted a portion of clone 41 d2gg_{B10}^B. A portion of the sequence of clone AFLP 27afB10 was a sequence with a high degree of similarity to the gene of the large subunit of RuBisCO (rbcL), located within the chloroplasts. The remainder of the clones contained 6 ribosomal sequences (21%), 4 satellites (14%) including 3 from a cucumber, 1 snRNA, and 4 fragments of nuclear genes, including genes for cell cycle control, the dopamine receptor, neomycine resistance, and TGF-beta Type I receptor.

A comparison of these same clones to an EST database demonstrated that 8 clones (29%), including two DSC clones, did not possess any significant similarity, and the remainder bore a resemblance to many distinctly different coding sequences. Among them, 5 sequences were found that resemble sequences of the cDNA of flowers at various stages, including 0-3 mm flower

Tab. 4. Results of sequencing of some of the subtractive and AFLP clones (gd and AFLP coupled to the sex genes).

Clone (length bp)	Compared sequence	Identity	Description	Accession number	e-value
43 d2gg _{B10} ^B (337)	57	<i>Cucumis sativus</i>	N: Cucumber satellite DNA type III	X03770.1	3e-06
41 d2gg _{B10} ^B (876)	496	<i>Cucumis sativus</i>	N: MSC16 mit. tRNA-Thr (trnT) gene	AY258272.1	0.0
	88	<i>Crassostrea gigas</i>	N: TGF-beta Type I receptor	AJ544074	2e-40
	90	<i>Tall Fescue</i>	N:EST: 44 deg C Heat Stress cDNA	CK801672.1	1e-41
33 d2gg _{B10} ^B (178)	72	<i>Cucumis sativus</i>	N: ITS1; 5.8S rRNA gene	AJ488213.1	2e-27
	72	<i>Royal Gala</i>	N:EST: Royal Gala cDNA clone	CN943070.1	5e-25
32 d2gg _{B10} ^B (357)	185	<i>Cucumis sativus</i>	N: ITS1, and ITS2, hardwickii strain	CSA488213	4e-58
	88	<i>L. esculentum</i>	N:EST: tomato flower bud	AW928996	1e-21
20 d2gg _{B10} ^B (259)	180	<i>Cucumis sativus</i>	N: 5S rRNA gene and inter-spacer region	AY424363.1	5e-69
	92	<i>Zea mays</i>	N:EST: Zea mays sperm cell cDNA.	CF675057.1	4e-21
21 d2gg _{B10} ^B (147)	39	<i>A. thaliana</i>	N: clone bl-2 satellite seq.	AF494859.1	3e-10
	40	<i>A. thaliana</i>	N:EST: cDNA clone tk63h09.g7	CO048708.1	5e-12
18 d2gg _{B10} ^B (394)	298	<i>A. thaliana</i>	N: <i>Arabidopsis thaliana</i> chloroplast	AP000423.1	e-154
	384	<i>Malus x domestii.</i>	N:EST: Royal Gala cDNA clone	CN923517	e-160
	83	<i>Pinus thunbergii</i>	X: chloropl. ORF42f	NP042478.1	5e-10
16 d2gg _{B10} ^B (451)	364	<i>Cucurbita pepo</i>	N: 26S ribosomal RNA gene,	AF479108.1	e-153
	323	<i>L esculentum</i>	N:EST: flower, 8 mm to preanth buds	BI930533	e-132
	197	<i>Candida albica</i>	X: microspore hypot. protein CaO19.6835	EAK91016.1	7e-21
14d2gg _{B10} ^B (452)	366	<i>Cucurbita pepo</i>	N: 26S rRNA gene,	AF479108.1	e-141
	363	<i>L. esculentum</i>	N:EST: tomato flower, buds 0-3 mm	BI925834	e-131
	134	<i>Plasmod. yoelii</i>	X: hypothetical protein	EAA18799.1	0.002
12d2gg _{B10} ^B (91)	33	<i>Polyosoma</i>	N: chl. part. ndhF gene for NADH dehyd	AJ4197021.1	1e-05
	26	<i>cunninghami</i>	N: cell cycle checkpoint mRNA	U83666.1	2e-04
	26	<i>Rattus norvegicus</i>	N:EST: cDNA for abiotic stressed leaves	CD720928.1	1e-05
		<i>Vitis vinifera</i>			
2 d2gg _{B10} ^B (331)	167	<i>A. thaliana</i>	N: <i>A. thaliana</i> chloroplast DNA	AP000423.1	4e-06
	26	<i>Vitis vinifera</i>	N:EST: cDNA for abiotic stressed leaves	CD720928.1	8e-04
	54	<i>A. thaliana</i>	X: ycf1 [<i>Arabidopsis thaliana</i>]	NP_051117.1	9e-06
1 d2gg _{B10} ^B (355)	194	<i>Cucumis sativus</i>	N: ITS1, 5.8S rRNA gene and ITS2, strain hardwickii	AJ488220.1	2e-66
1gdrZ _J ^B (439)	289	<i>Lotus corniculatus</i>	N: chloroplast DNA, complete genome	AP002983.1	3e-84
	278	<i>Vitis vinifera</i>	N:EST: abiotic stressed leaves	CD719377.1	2e-73
	150	<i>Citrus jambhiri</i>	X: ACR toxin-sensit inducing protein	BAB85481.1	6e-11
17 dGy3 _{Hgy3} ^H (419)	243	<i>Cucumis sativus</i>	N: Cucumber satellite DNA type I	X03768.1	8e-67
	243		N: Cucumber satellite DNA type II	X03769.1	1e-55
	59	<i>Festuca ruginosa</i>	N:EST: heat stress cDNA	CK802895.1	8e-23
15 dGy3 _{Hgy3} ^H (1135)	210	<i>pPGKneo-II</i>	N: neomycine transferase	AF335420.3	1e-76
	109	<i>Branchiost.</i>	N: dopamine D1/beta receptor	AJ005433.1	4e-51
	177	<i>Lanceolata</i>	N: Cucumber satellite DNA type III	X03770.1	2e-45
	224	<i>Cucumis sativus</i>	N:EST: fem. gonad MZ cDNA 5'(catalase)	BM733797.1	1e-81
	65	<i>Ascaris suum</i>	X: photosystem II protein D1	CAD24436.1	2e-05
		<i>Palmaria descipiens</i>			
13 dGy3 _{Hgy3} ^H (419)	174	<i>A. thaliana</i>	N: U2.6 snRNA gene	X52312.1	1e-48
	105	<i>Crassostrea gigas</i>	N: thetaRI gene for TGF-beta Type I receptor	AJ544074	1e-37
	196	<i>Ipomea nil</i>	N:EST: flowers and flower bud cDNA	BJ555563	8-54
	83	<i>Tall fescue</i>	N:EST: 39 degC Heat stress	Ck802584	9e-39
	70	<i>A. thaliana</i>	X: vacuolar ATP synthase catalytic subunit-rel./ vacuolar proton pump-related	NF_173126.1	4e-06
1 gddZ _H ^{BI} (579)	160	<i>Cucurbita pepo</i>	N: trnR-rnr5 integrat. chloroplast	AY396262.1	6e-46
	160	<i>Oryza sativa</i>	N: chromosome 3 BAC clone	AL731605.3	1e-43
	240	<i>L. esculentum</i>	N:EST: breaker fruit cDNA 5' end.	BM412815.1	1e-50

26afB10 (169)					
27afB10 (261)	101	<i>Cucumis sativus</i>	N: RuBisCO (rbcL) gene, partial cds; cp gene for cp product	AF206755	1e-48
	88	<i>Malus x domestica</i>	N: EST: Royal Gala 24 DAFB cDNA	CN872400	9e-37
	34	<i>Buxbaumia aphylla</i>	X: RuBisCO large subunit	AAG44273	7e-13
28afB10 (125)	24	<i>Malus x domestica</i> (cultivated apple)	N: EST: Apple developing fruit cDNA similar to essential related to bicoid interacting protein (2M394)	CN544885	0.004
29afB10 (77)	71	<i>Cucumis sativus</i>	N: clone F16 mit. genomic sequence	AF282393	6e-28
31afB10 (173)					
32afGy3 (96)					
33afGy3 (90)					
34afGy3 (79)	22	<i>Plasmod falciparu 3D7</i>	N: EST: asexual cDNA sequence	BI815313	0.002
36afGy (87)	27	<i>Mus musculus</i>	N: EST: 7-day old embryo cDNA clone	BB652274	0.003
37afGy3 (86)					
39afGy3 (52)					

As names of clones or libraries: af-denotes a marker (tag) or an AFLP library; d-denotes a marker (tag) or a DSC library; gd-denotes a marker (tag) or a GDDSC library. The symbols located on the right side indicate the name of the line where it occurred. Information was not included in the table if a significant similarity was not found. The rest of the descriptions same as in other tables.

buds, mainly of *Lycopersicon esculentum*, and there were 2 clones from fruits, 5 clones of plants genes for abiotic stress, and 5 cDNA of the various portions of vegetative plants (stems, buds, leaves, and whole plants). Three clones resembled animal EST. Clone 15dGy3_{HGy3}^H demonstrated 95% similarity over a length of 224 bp, to a sequence (catalase) isolated from the female organs of *Ascaris suum* (nematode). Among the 19 homologous EST sequences, only 4 were not of plant origin. In the protein data base, as many as 20 (71%) of the clones did not possess counterparts.

The length of all the subtractive clones as well as a part of their sequence homologous to database sequences was much longer than the lengths of the AFLP clones.

DISCUSSION

Quality of subtraction

Before now, the DSC method was used only rarely for the purpose of searching for differences within genomes or pools of mRNA. In the field of plant genetics, two publications are available on the Internet [13, 14], respectively pertaining to rice and to *Tripsacum dactyloides*, and two publications are available in printed form [15, 16], respectively pertaining to rice and to *A. thaliana*. In the case of our research, satisfactory results were obtained following a number of

modifications mentioned in the materials and methods section. Aside from searching for differences in the entire genome, we also attempted to isolate those differences from the region of the sex genes' loci (the only trait morphologically differentiating the analyzed pairs of lines). For that purpose, it was decided to guide the DSC subtraction in a manner similar to how it was done with the RDA [6], i.e. to perform GDDSC. So far, there have been no publications on the use of such a modification of the DSC method. The sex of the plant was the control trait. The subtracted populations consisted of the sex-segregating F₂ plants DNA pools. This should permit for a good segregation of tags coupled strongly to the sex genes from the remainder, because of the uniform distribution within the subtracted bulks of allelic sequences either not coupled or only weakly coupled with the sex genes. The size of the area from which the tags linked with the guiding gene will be "fished out" depends on the amount of the bulks used for the subtraction; the greater the size, the smaller the area from which they originate. The number of tags should also decrease, but the remainder are more and more coupled to the guiding gene. The precise size of the area from which the subtracted fragments originate was not calculated under the conditions of this experiment. However, it may be estimated that with the number sizes used, it will not be greater than 2 cM. It may be concluded that in our case, a ~1/3 variability is possible to find using the DSC method with the *Bgl*III restriction enzyme, and is located very close to the gene determining trait differences in the NIL lines. Similar results were obtained using the AFLP method, considering the fact, that only the markers linked to the dominant alleles were presented [2]. The obtained subtraction results, as may be gathered from the amount of the differences attained using the given restriction enzymes, seem to testify to the high effectiveness of the methods used herein. The restriction enzymes used do not give a uniform number of differences, which may be associated with the average length of the fragments generated by them as well as with the extent of genome simplification [1, 19,20]. The *Hind*III is an exception, since it gives twice as few tags than each of the remaining two. This may be the consequence of the lesser genome simplification given by the *Hind*III [1]. For that reason, in the case of this restriction enzyme, more rounds of subtraction were performed [18]. *Bgl*III, in accordance with suggestions previously published [20], because it gives a short average fragment length (2699 bp) and generates a large representation, is suitable for searching for differences when the subtracted populations are closely related to each other.

The question arises as to whether the identified tags are really polymorphic? With reference to genome tags, such testers (RFLP) were not done. Yet the results of earlier testing of the cDNA subtractome [9,17] indicate the great precision of this method.

In this study, even though there were somewhat less differences in nearly all the "recessive" gensubtractomes than in the "dominant" ones, the similar amounts of tags within a majority of the "isogenic" gensubtractome pairs (i.e. obtained as a consequence of reversing the subtraction direction, in the area of the isogenic line pair for ex. Gy3_{HGy3} and Hgy3_{Gy3}) seem to confirm the precision of the

subtraction (Tabs. 1 and 2). This kind of a result may be justified by stating that regardless of the direction of subtraction, the sites at which the genomes differentiate do not vary and their number does not change. However, in the case of two gensubtractomes (Hgy3_{Gy3}^H and Hgy3_{Gy3}^B) an inexplicable significant reduction in the size of the "recessive" dsubtractoms occurred (when a line of plants in which sex is dependent on the recessive alleles serves as the tester – this also pertains to bulks of GDDSC).

Analysis of the isolated sequences

The number of differences within the genomes seems quite large, considering that these are near isogenic lines. In the context of a large morphological similarity and a large number of differences in the genomes of these lines, the information-bearing content of the changed fragments seems very interesting. The presented analysis was aimed at their initial characterization in that regard. Of particular interest is the determination of whether these are coding sequences, what they are coding and what these products are for, and also whether these sequences are expressed in flowers, and if they are expressed differentially. For that reason, a hybridization was carried out with genpoloms (subtractome d and gd, AFLPome) versus cDNApolomes (cDNAsubtractomes and dhomes) representing the differences between 1- to 2-mm floral buds [9,10] cDNA sets. A portion of the genomic clones was sequenced.

Hybridization

The obtained hybridization results conclude that a large portion of genpolom contains sequences that are homologous to the products of the genes which are differentially expressed in 1- to 2-mm floral buds of the analyzed lines (to cDNApoloms). Analysis of the hybridization of the individual clones of AFLPome indicates that it may even be as much as ~50% of the clones.

Other aspects of the obtained results may be as follows: (i) the relationship of the hybridization intensity to the types of cDNApoloms with which the genpolom hybridize; (ii) the restriction enzyme that has been utilized for the formation of the genome representation.

A hybridization with the "dominant" cDNApoloms rather than with the "recessive" (Tab. 3).

The existence of a differential expression on the one hand may indicate that these genes (differentiating line genomes) play important functions in the formation of the various types of flowers, while on the other hand, a lack of the transcripts of these genes in the "recessive" cDNAlibraries may point to either a lack of or a reduction in transcription due to their mutation. It may be for that reason that the "recessive" cDNAsubtractomes contain 2-3 times fewer clones than the "dominant" ones [9]. There is one more question: why do the "dominant" genpolom libraries hybridize in a similar fashion to the "recessive" ones? The answer may be that despite the mutation, there remains sufficiently many homologies between the genomic clones and the cDNA; in addition, the

latter, for reasons presented earlier, mainly occur in the “dominant” cDNApoloms (cDNAsubtractom and dhome).

However, all of that does not explain the nearly absolute lack of hybridization with the B10_{2gg} dhome, which indicates the existence of sequences different than those within the remaining cDNApoloms (at least the “dominant” ones). However, a majority of the B10_{2gg} dhome clones hybridize, in a better or worse fashion, with a majority of the clones of the cDNAsubtractome B10_{2gg} [9] (that hybridize with the genpolom), which might mean, that the cDNAsubtractome contains a portion of clones homologous to dhome B10_{2gg} and a portion homologous to the genpolom. In support of this presumption, it may be helpful to consider the hybridization results of three RAPD markers entirely coupled with the *Gy* loci (in mapping the average resolution) [9]. In this case, the opposite result was obtained, i.e. these markers hybridized with the dhome B10_{2gg}, but did not hybridize with the cDNAsubtractom B10_{2gg}.

With regard to the effectiveness (measured by the intensity of hybridization) of the restriction enzymes in generating tags from the coding regions, then similarly to the number of tags generated generally, *Bgl*III appears to be the best. This may stem from the fact that the number of the tags was highest, although *Bam*HI generated only slightly fewer fragments and the level of hybridization may be assessed as similar to that for *Hind*III. Therefore, it appears to be a greater preference of the coding regions. Serving to support that statement may be the fact that there is a complete lack of hybridization of the dsubtractoms B10_{2gg}^B and Hgy3_{Gy3}^H, while the B10_{2gg}^{BI} and Hgy3_{Gy3}^{BI} dsubtractoms hybridized with great strength.

Another significant aspect would seem to be the clearly visible presence of the signals of hybridization of the GDDSC libraries (dgsubtractoms), particularly the strong hybridization of the “dominant” libraries. This may signify that fragments of active genes which reside in the vicinity of the sex genes and become expressed in the 1- to 2-mm floral buds have been isolated from the genomes. Some fragments of the sex genes may also be present. However, to prove this hypothesis, further studies are required

Sequencing

The analysis of the sequences presented here (Tab. 4) points to their extensive variety. It also confirms the hybridization results; in other words, the presence in the gensubtractom of multiple coding fragments (nearly ~70% of the sequenced clones possess them). Among these, clones similar (fulfilling the herein assumed significance criteria) to those found earlier in reproduction-associated organs (5 floral and 2 fruit). Particularly interesting from the perspective of development and determination of flower sex may be the genes that are expressed within floral buds and flowers, fruits, stress genes, cell cycle control genes, and genes associated with the proton pump. The differential expression of the genes responsible for the flower type-dependent stress response has been found in the flowers of plants [21, 22, 23] including the 1- to 2-mm floral buds of a cucumber [9, 10]. It may also be interesting to consider the occurrence of

the sequences of the dopamine receptor, or catalase from the female organs of *Ascaris suum*. In addition, snRNA was found, which is significant due to the fact that short RNAs may play a role as specific regulators at the level of the transcript.

The most numerous group was that of the sequences, the role of which, as of present, in the development of flowers and sex, might be difficult to identify. These consist of organelle sequences (nearly 30% of sequences contain them), rDNA sequences (found in over 20% of the sequences) and satellites (found in about 14% of the sequences).

The question arises as to whether such a large number of subtracted organelle sequences is an artifact, or if they were really inside the nucleus. It seems that the answer may lie in the analysis of the two clones that contain sequences similar to the organelle DNA of a cucumber. Tags 41d2gg_{B10}^B and 27afB10 possess organelle-coding fragments, the first fragment of the large subunit of RuBisCO (rbcL), and a second fragment of a sequence in which a deletion, as already mentioned, gives a mosaic phenotype (in this case it was the MSC16). In both cases, the flanking sequences were not of organelle origin. Therefore, it may be concluded that both sequences were inside the nucleus, since otherwise, all the clones would have to be organelle sequences. In the case of these sequences, it may even be possible to estimate the time of their nuclear transfer. It must have happened after 1972 since it was in that year when the 2gg line was obtained from the B10 line in a process of chemical mutagenesis (ethylenimine) [3]. If it were otherwise, both of the lines would have had these sequences and they would have been removed during the subtraction. DNA transfer from the organelles to the nucleus is a known phenomenon and it occurs very frequently – this may take place in the cells of *Nicotiana tabacum* (from the chloroplasts to the nucleus) in *in vitro* cultures with a frequency of 1/5000000 [24, 25]. Such a large number of the herein obtained organelle-related tags seem to confirm the significant frequency of such a transfer.

Acknowledgements. This work was financed from the following grants of the State Scientific Committee: 029/ PO6/0.7 (which was a part of bigger project PBZ/KBN/029/P06/2000) and 5PO6A 001 17.

REFERENCES

1. Lisitsyn, N., Lisitsyn, N. and Wigler, M.H. Cloning the differences between two complex genomes. **Science** 259 (1993) 946-951.
2. Witkowicz, J., Urbańczyk-Wochniak, E. and Przybecki, Z. AFLP marker polymorphism in cucumber (*Cucumis sativus* L.) Near isogenic lines differing in sex expression. **Cell. Mol. Biol. Lett.** 8 (2003) 375-381.
3. Kubicki, B. New sex types in cucumber and their uses in breeding work. XIXth International Horticultural Congress Warszawa 11-18 September (1974), 475-485.

4. Greenbaum, D., Luscombe, N.M., Jansen, R., Qian, J. and Gerstein, M. Interrelating different types of genomics data, from proteome to secretome: 'Oming in on function' 2002. <http://bioinfo.mbb.yale.edu/e-print/omes-genomeres/text.pdf>
5. Luo, J.H., Puc, J.A., Slosberg, E.D., Yao, Y., Bruce, J.N., Wright, T.C., Becich, M.J. and Parsons, M. Differential subtraction chain, a method for identifying differences in genomic DNA and mRNA. **Nucleic Acids Res.** 27e (1999) 24e, i-viii
6. Lisitsyn, N.A., Segre, J., Kusumi, K., Lisitsyn, N.M., Nadeau, J.H., Frankel, W.N., Wigler, M.H. and Lander, E.S. Direct isolation of polymorphic markers linked to a trait by genetically directed representational difference analysis. **Nat. Genet.** 6 (1994) 57-63.
7. Sambrook, J., Fritsch, E.F. and Maniatis, T. Molecular cloning. A laboratory manual. Ed.II Ford N., Nolan C., Ferguson M. (Cold Spring Harbor Laboratory Press N.Y.), 1989.
8. Urbańczyk-Wochniak, E., Filipecki, M. and Przybecki, Z. A useful protocol for *in situ* RT-PCR on plant tissues. 7 **Cell. Mol. Biol. Lett.** (2002) 7-18.
9. Kowalczyk, M.E. Analiza porównawcza produktów ekspresji genów w pąkach kwiatowych ogórka (*Cucumis sativus* L.) Doctoral thesis 2004 Warsaw University of Agriculture, Department of Plant Genetics, Breeding and Biotechnology.
10. Przybecki, Z., Kowalczyk, M.E., Siedlecka, E., Urbańczyk-Wochniak, E. and Malepszy, S. The isolation of cDNA clones from cucumber (*Cucumis sativus* L.) floral buds coming from plants differing in sex. **Cell. Mol. Biol. Lett.** 8 (2003) 421-438.
11. Malepszy, S., Burza, W. and Śmiech, M. Characterization of a cucumber (*Cucumis sativus* L.) somaclonal variant with paternal inheritance. **J. Appl. Genet.** 37 (1996) 65-78.
12. Lilly, J.W., Bartoszewski, G., Malepszy, S. and Havey M.J. A major deletion in the cucumber mitochondrial genome sorts with the MSC phenotype. **Curr. Genet.** 40 (2001) 144-151.
13. Zhao, X. and Chen, F. Enhancing the technology base for rice improvement. <http://farrer.riv.csu.edu.au/rice2001/research/res-03.htm>.
14. Hoisington, D. and Kanapiu, F. Key results from Applied Biotechnology Center (ABC) – Mexico. www.africancrops.net/
15. Jaccoud, D., Peng, K., Feinstein, D. and Kilian A. Diversity arrays: a solid state technology for sequence information independent genotyping. **Nucleic Acids Res.** 29e (2001) 25e.
16. Gong, Z., Koiwa, H., Cushman, M.A., Ray, A., Bufford, D., Kore-eda, S., Matsumoto, T.K., Zhu, J., Cushman, J.C., Bressan, R.A. and Hasegawa, P.M. Gene that are uniquely stress regulated in salt overly sensitive (sos) mutants. **Plant Physiol.** 126 (2001) 363-375.
17. Kowalczyk, M.E. and Przybecki, Z. Construction of differential expression profiles by cDNA – DSC method in cucumber floral buds (*Cucumis sativus* L.). Pol. Kong Genet. Gdańsk. 7-10 IX (2004).

18. Fujisawa, M., Hayashi, K., Nishio, T., Bando, T., Okada, S., Yamato, K.T., Fukuzawa, H. and Ohyama, K. Isolation of X and Y chromosome-specific DNA markers from a liverwort, *Marchantia polymorpha*, by representational difference analysis. **Genetics** 159 (2001) 981-985.
19. Bishop, D.T., Wiliamson, J.A., and Skolnik, M.H. A model for restriction fragment length distributions. **Am. J. Hum. Genet.** 35 (1983) 795-815.
20. Lisitsyn, N.A. Representational difference analysis. Genome analysis. in: **A Laboratory Manual** (Green, E.D., Birren, B., Klapholz, S., Myers, R.M. and Hieter, P. Eds) Cold Spring Harb, NY, Cold Spring Harb. Lab. Press, 1997-1999. Volume 4: Mapping Genomes 217-258.
21. Zabaleta, E., Oropeza, A., Jimenes, B., Salerno, G., Crespo, M. and Herrera-Estrella, L. Isolation and characterization of genes encoding chaperonin 60 β from *Arabidopsis thaliana*. **Gene** 112 (1992) 175-181.
22. Tsukaya, H., Takahashi, T., Naito, S. and Komeda, Y. Floral organ-specific and constitutive expression of an *Arabidopsis thaliana* heat-shock HSP 18.2:GUS fusion gene is retained even after homeotic conversion of flowers by mutation. **Mol. Gen. Genet.** 237 (1993) 26-32.
23. Moris, K., Holmes, L., Cordai, L., Huttley, A., Carre, I. and Jackson, S.D. U don't RING me flowers anymore! **Comp. Biochem Physiol. Part A** 134 (2003) S1-S237.
24. Martin, W. Gene transfer from organelles to the nucleus: frequent and in big chunks. **Proc. Natl. Acad. Sci. U.S.A.** 100 (2003) 8612-8614.
25. Stegemann, S., Hartmann, S., Ruf, S. and Bock, R. High-frequency gene transfer from the chloroplast genome to the nucleus. **Proc. Natl. Acad. Sci. U.S.A.** 100 (2003) 8828-8833.