BIOLOGIA PLANTARUM 55 (3): 577-580, 2011

BRIEF COMMUNICATION

Comparison of genomic SSR and EST-SSR markers for estimating genetic diversity in cucumber

J. HU, L. WANG and J. LI*

Open Key Laboratory of Horticultural Plant Physiology and Genetic Improvement, High School of Henan Province; Key Laboratory of Fruit Plant and Cucurbit, Henan Province; College of Horticulture, Henan Agricultural University, Zhengzhou 450002, P.R. China

Abstract

Thirteen genomic microsatellite (gSSR) and sixteen expressed sequence tag (EST)-SSR (eSSR) markers were compared to estimate genetic diversity among 29 cucumber (*Cucumis sativus* L.) accessions. gSSR markers detected mean 4.46 alleles with a mean polymorphic information content (PIC) of 0.664, against eSSR markers with mean 3.38 alleles and a mean PIC of 0.397. gSSRs amplified more null alleles than eSSRs. Genetic diversity within the accession set was estimated by construction of dendrograms using gSSR or eSSR data. There was a clear consistency between gSSR and eSSR trees in terms of positioning of most cucumber germplasms. gSSR markers could separate various types of cucumber germplasms on the whole, although clustering of some accessions was not based on their geographical origins in eSSR tree. eSSR markers identified an independent sub-cluster containing five accessions resistant to downy mildew, suggesting a probable relationship between eSSRs and disease-resistance trait in cucumber. The Mantel test between gSSR and eSSR matrices revealed a good fit correlation (r = 0.836). The general dendrogram constructed using the combined data of gSSRs and eSSRs was similar to those obtained separately with each marker.

Additional key words: *Cucumis sativus*, dendrograms, polymorphic information content.

Cucumber (*Cucumis sativus* L.), being one of the economically important vegetables in the world, occupies a large cultivated area, ranking fourth after tomato, onion and cabbage (Pitrat *et al*. 1999). Despite its large morphological variability, cucumber has a narrow genetic base (Staub *et al*. 2005), which limits development of new cucumber cultivars by cross-breeding. Thus, genetically diverse cucumber materials are in demand as they allow improvement of certain agronomic traits and reduction of vulnerability to environmental stresses. To date, the degree of genetic diversity in cucumber has been assessed with a number of DNA markers (Dijkhuizen *et al*. 1996, Danin-Poleg *et al*. 2001, Staub *et al*. 2005, Zhuang *et al*. 2008, Sikdar *et al*. 2010), which provide useful genetic information for cucumber cross-breeding. However, most of the markers are anonymous ones which are likely to have no close linkage to transcribed sequences and thus no known genic function. This limits their applications in some extent. Express sequence tag (EST)-derived microsatellites are such type of markers which are easily obtained by electronic search of EST databases. EST-microsatellite (eSSR) polymorphism is associated with transcribed regions of the genome and reflects the genetic diversity inside or adjacent to the genes (Varshney *et al*. 2005), which may be functionally more informative than genomic SSR (gSSR) that are more widely used. In this study, we compared gSSR and eSSR markers for estimating genetic diversity of cucumber accessions from a wide collection. The information offered by the present research would favour

Received 10 March 2010, *accepted* 10 April 2010.

Abbreviations: EG, EO, NC, OP, SC - Europe greenhouse, Europe open-field; north China, occident processing and south China accessions, respectively; eSSR - expressed sequence tag microsatellite; EST - expressed sequence tag; gSSR - genomic microsatellite; PIC - polymorphic information content; UPGMA - unweighted pair group method using arithmetic averages; W - wild species.

Acknowledgements: This work is supported by National Key Technology R & D Program (No. 2009BADB8B02). We are grateful to Prof. Jiancan Feng, College of Horticulture, Henan Agricultural University, for his assistance in the laboratory work on this project.

* Author for correspondence; fax: (+86) 371 63558096, e-mail: lijw555@sohu.com

selection of the cucumber genotypes having "true genetic diversity" inside or adjacent to the genes for cross-breeding or molecular mapping. Plant materials consisted of 29 cucumber (*Cucumis sativus* L.) accessions, representing a majority of germplasms currently cultivated in China. These accessions included 12 North China (NC) accessions (NC1, NC2, NC3 - NC12), 9 South China (SC) accessions (SC1, SC2, SC3 - SC9), 3 Europe greenhouse (EG) accessions (EG1, EG2 and EG3), 2 Europe open-field (EO) accessions (EO1 and EO2), 2 occident processing (OP) accessions (OP1 and OP2), and one wild species (W) (*Cucumis hystrix*). The NC and SC accessions and wild species are all Chinese native germplasms, and the others were introduced from Europe and America in the last century and domesticated in China. All the accessions were provided by Group of Plant Genetics and Breeding, Henan Agricultural University, China, and subjected to DNA extraction with a CTAB method (Murray and Thompson 1980).

Thirteen gSSR and 16 eSSR primers were obtained from the published primer sequences (Watcharawong-paiboon and Chunwongse 2008) and our previous research (Hu *et al*. 2010), respectively. The reaction system and program of all PCR amplifications were the same as in previous report (Hu *et al*. in press). Amplified products were electrophoresed in 6 % non-denaturing polyacrylamide gels (acrylamide : bisacrylamide; 19:1) and the gels were silver-stained according to the method proposed by Creste *et al*. (2001) and photo-documented.

Amplified bands were visually scored as present (1) or absent (0) for each gSSR or eSSR marker. With the marker data, binary matrices were established and then used to calculate Jaccard's similarity coefficients (Jaccard 1908). Cluster analysis was performed to construct dendrogram based on the similarity matrix data using the unweighted pair group method using arithmetic averages (UPGMA) and the *SHAN* clustering program. To investigate the goodness of fit of the two marker data, a Mantel test was conducted using the *COPH* and *MXCOMP* programs. All analyses were performed with the *NTSYS-pc 2.10* software (Rohlf 2000). Polymorphism information content (PIC) values were calculated to evaluate diverse level of each SSR marker according to Anderson *et al*. (1993) using the formula: $PIC = 1 - \sum P_{ij}^2$, where P_{ij} is the frequency of the j_{th} allele (marker) for the i_{th} EST-SSR locus.

Both gSSR and eSSR markers amplified distinct band patterns among the 29 cucumber accessions and each marker revealed polymorphism. The 13 gSSR markers detected 58 alleles with a mean of 4.46 and the 16 eSSR markers detected 54 alleles with a mean of 3.38. Obviously, gSSR marker detected more allelic number than eSSR marker, showing a higher discrimination power. Also, PIC values presented a similar result (Table 1). gSSR markers had high PIC values that ranged from 0.295 to 0.835 with a mean of 0.664, while comparable low PIC values were observed with eSSR markers and ranged from 0.185 to 0.578 with a mean of 0.397. The present results are in agreement with the previous reports in a variety of plant species (Cho *et al*. 2000, Danin-Poleg *et al*. 2001, Chabane *et al*. 2005). This fact may be due to possible selection against alterations in the conserved coding sequences (Scott *et al*. 2000), which limits SSR variation in these regions.

Table 1. Comparison of gSSRs and eSSRs in cucumber. The number of alleles, null alleles, multiple alleles and polymorphic information content (PIC) were investigated in 29 cucumber accessions (Y means detection of null/multiple alleles, N means no detection of null/multiple alleles).

	Primers	Number of alleles	Null alleles	Multiple alleles	PIC
Genomic SSR	CSJCT14	4	Y	N	0.754
	CSJCT42	3	N	Y	0.457
	CSJCT71	4	Y	N	0.735
	CSJCT97	2	N	N	0.295
	CSJCT252	6	Y	Y	0.828
	CSJCT266	3	N	N	0.437
	CSJCT323	4	Y	N	0.782
	CSJCT358	5	N	N	0.771
	CSJCT390	7	Y	N	0.806
	CSJCT632	4	N	Y	0.511
	CSJCT641	3	Y	N	0.718
	CSJCT662	7	N	Y	0.835
	CSJCT933	6	N	Y	0.702
	mean	4.46			0.664
EST-SSR	EC11	5	Y	Y	0.578
	EC12	2	Y	N	0.279
	EC13	3	N	N	0.426
	EC15	4	N	Y	0.354
	EC18	2	N	N	0.285
	EC19	3	N	N	0.347
	EC20	3	N	Y	0.247
	EC22	3	N	N	0.328
	EC24	4	N	N	0.438
	EC27	3	N	N	0.461
	EC28	4	Y	Y	0.441
	EC34	5	N	Y	0.642
	EC35	3	N	N	0.392
	EC39	4	N	Y	0.404
	EC41	2	N	N	0.185
	EC49	4	N	N	0.540
	mean	3.38			0.397

To better understand SSR variation in cucumber, null alleles and multiple alleles detected by gSSR or eSSR markers were investigated (Table 1). Six out of the 13 gSSRs (*ca*. 50 %) detected null alleles, against only 3 out of 16 eSSRs (*ca*. 20 %). eSSR null alleles existed in only one or two accessions, while gSSR null alleles appeared in 4 - 12 accessions. Since null alleles are the possible result of mutation events (*e.g*. single nucleotide polymorphisms) occurring at the priming sites, it is reasonable to believe that the transcribed portions of

cucumber genome bear less mutations under selective pressures than non-coding regions, and hence, the rate at which SSR flanking regions are free to evolve is lower in EST sequences. This may explain a lower proportion of null alleles detected by eSSRs, although a larger data would be needed to verify this possibility. In addition, the amplification of multiple alleles was frequently observed with the two marker systems (5 of the 13 gSSRs versus 6 of the 16 eSSRs). Presumably, in this case, each of the markers targeted more than one homoeolocus in cucumber genome. Amplification of multiple loci by a single microsatellite primer pair was also reported in bread wheat (Gupta *et al*. 2003).

Pairwise comparison was performed between the 29 accessions and Jaccard's similarity coefficients were separately calculated using the gSSR and eSSR data. The former generated a mean similarity coefficient of 0.497 and the latter generated a higher value of 0.664, again reflecting polymorphic differences between the two markers. Combination of the two data gave a moderate value of 0.551.

Dendrograms based on the gSSR or eSSR data were separately constructed and the two markers revealed a high similarity in dendrogram topologies (Fig. 1*A*,*B*). The 29 accessions were generally divided into two broad clusters, which was in line with the different geographical origins of the accessions. The Chinese native germplasms (NC and SC accessions and a wild species) formed an independent group that fell well outside of the group including all non-Chinese native germplasms (EG, EO and OP accessions). Within the group containing Chinese native germplasms, most NC and SC accessions were clearly separated, particularly in the gSSR tree where only two accessions (SC1 and NC6) failed to divide according to their geographical origins. Moreover, some accessions with known lineage relationship were clustered together well in both trees, such as NC8 and NC9, both of which were self progenies of a known cultivar, so did 'SC2' and 'SC3'. The wild species 'W' was closely clustered with some SC accessions in all trees, indicating that it could be an ancestry of Chinese cucumber, especially the SC accessions. This opinion was supported by our previous study with chloroplast SSRs (Hu *et al*. 2009), where the wild species shared a haplotype with most SC accessions. From these, it is clear that there is a strong consistency on positioning of most cucumber accessions between the gSSR and eSSR trees. On the contrary, some differences in the distribution of certain accessions were observed in the phylogenetic trees constructed by different markers. For example, although most NC and SC accessions were well separated in gSSR tree, several NC and SC accessions ('NC4', 'NC6', 'NC10', 'SC8' and 'SC9') were closely clustered and distinctly diverged from the other groups in eSSR tree. Interestingly, the five accessions had a common trait: resistance to downy mildew under artificially inoculated condition (personal observation). This fact suggested that these eSSR markers could be related to some important agronomic traits like disease resistance in cucumber, and could be an alternative marker system to functional diversity studies. Besides, non-Chinese native accessions were clustered based on their geographical origins in gSSR tree, however, positioning of some accessions (EG and OP accessions) appeared confused in eSSR tree. These differences between the two trees could be due to the differences in targeted DNA regions or

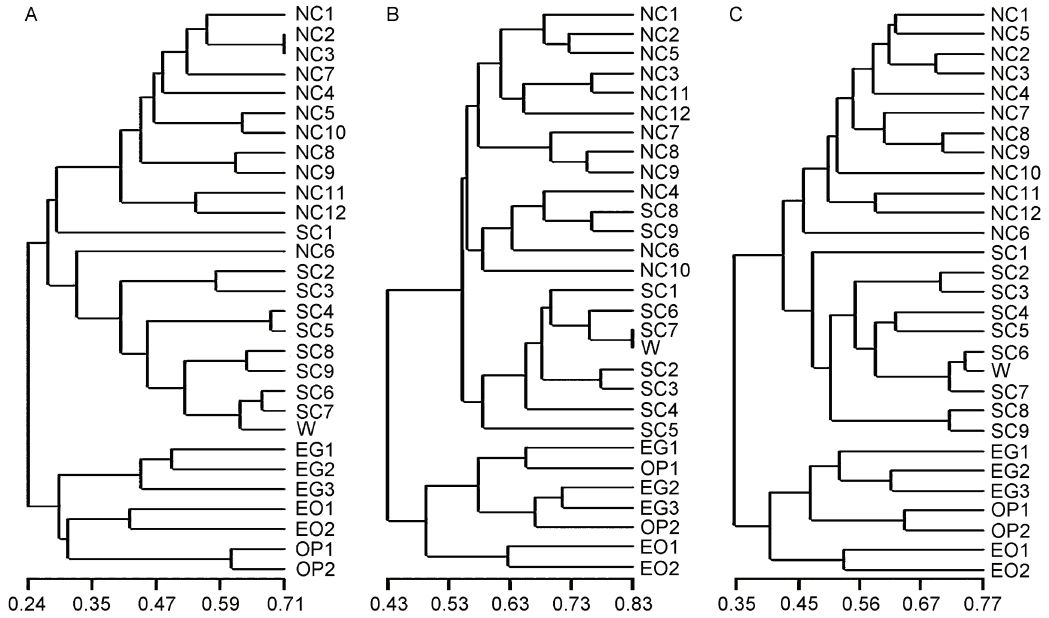

Fig. 1. Dendrograms of cluster analysis of 29 cucumber accessions using unweighted pair-group method with arithmetic averages (UPGMA) based on gSSR (*A*), eSSR (*B*) and gSSR + eSSR (*C*) data.

genomic coverage rate of the different marker systems (Varshney *et al*. 2005).

The correlation between gSSR and eSSR similarity matrices was examined by the Mantel test and revealed a good fit between the two data (r = 0.836). As expected, the general dendrogram (Fig. 1*C*) that was constructed using the combined data of the two markers was very similar to those obtained separately with each marker. This combined dendrogram clearly separated the accessions according to their geographical origins, demonstrating the importance of gSSR and eSSR dataset in the phylogenetic studies. The dendrogram from gSSR data was the most congruent with the general dendrogram as shown by the Mantel test (r = 0.944) between gSSR data and the combined data of the two markers.

To conclude, in this study, we compared gSSR and eSSR variations among 29 cucumber accessions and found that compared to eSSR markers, gSSR markers revealed higher polymorphism but detected more null alleles. Genetic diversity within the accession set was well resolved by each of the two markers, largely revealing a consistency between gSSR and eSSR trees. In comparison to gSSR or eSSR dendrogram, the combined dendrogram more clearly showed the genetic relationship among the accessions. So, estimation of genetic diversity in plant species with a narrow genetic base could be more efficient if different marker systems are used like gSSRs and eSSRs.

References

Anderson, J.A., Churchill, G.A., Autrique, J.E., Tanksley, S.D., Sorrells, M.E.: Optimizing parental selection for genetic linkage maps. - Genome **36**: 181-186, 1993.

Chabane, K., Ablett, G.A., Cordeiro, G.M., Valkoun, J., Henry, R.J.: EST versus genomic derived microsatellite markers for genotyping wild and cultivated barley. - Genet. Resour. Crop Evol. **52**: 903-909, 2005.

Cho, Y.G., Ishii, T., Temnykh, S., Chen, X., Lipovich, L., McCouch, S.R., Park, W.D., Cartinhour, N.A.S.: Diversity of microsatellites derived from genomic libraries and GenBank sequences in rice (*Oryza sativa* L.). - Theor. appl. Genet. **100**: 713-722, 2000.

Creste, S., Neto, A.T., Figueira, A.: Detection of single sequence repeat polymorphisms in denaturing polyacrylamide sequencing gels by silver staining. - Plant mol. Biol. Rep. **19**: 229-306, 2001.

Danin-Poleg, Y., Reis, N., Tzuri, G., Katzir, N.: Development and characterization of microsatellite markers in *Cucumis*. - Theor. appl. Genet. **102**: 61-72, 2001.

Dijkhuizen, A., Kennard, W., Havey, M.J., Staub J.E.: RFLP variation and genetic relationships in cultivated cucumber. - Euphytica **90**: 79-87, 1996.

Gupta, P.K., Rustgi, S., Sharma, S., Singh, R., Kumar, N., Balyan, H.S.: Transferable EST-SSR markers for the study of polymorphism and genetic diversity in bread wheat. - Mol. Genet. Genomics **270**: 315-323, 2003.

Hu, J.B., Li, J.W., Zhou, X.Y.: Analysis of cytoplasmic variation in a cucumber germplasm collection using chloroplast microsatellite markers. - Acta Physiol. Plant. **31**: 1085-1089, 2009.

Hu, J.B., Zhou, X.Y., Li, J.W.: Development of novel EST-SSR markers for cucumber (*Cucumis sativus*) and their transferability to related species. - Sci. Hort. **125**: 534-538, 2010.

Jaccard, P.: Nouvelles recherches sur la distribution florale. - Bul. Soc. Vaudoise Sci. Nat. **44**: 223-270, 1908.

Murray, M.G., Thompson, W.F.: Rapid isolation of high molecular weight plant DNA. - Nucl. Acids Res. **8**: 4321-4325, 1980.

Pitrat, M., Chauvet, M., Foury, C.: Diversity, history, and production of cultivated cucurbits. - Acta Hort. **492**: 21-28, 1999.

Rohlf, F.J.: *NTSYS-pc*: Numerical Taxonomy and Multivariate Analysis System, *Version 2.1*. - Exeter Publications, New York 2000.

Scott, K.D., Eggler, P., Seaton, G., Rossetto, M., Ablett, E.M., Lee, L.S., Henry, R.J.: Analysis of SSRs derived from grape ESTs. - Theor. appl. Genet. **100**: 723-726, 2000.

Sikdar, B., Bhattacharya, M., Mukherjee, A., Banerjee, A., Ghosh, E., Ghosh, B., Roy, S.C.: Genetic diversity in important members of *Cucurbitaceae* using isozyme, RAPD and ISSR markers. - Biol. Plant. **54**: 135-140, 2010.

Staub, J.E., Chung, S.M., Fazio, G.: Conformity and genetic relatedness estimation in crop species having a narrow genetic base: the case of cucumber (*Cucumis sativus* L.). - Plant Breed. **124**: 44-53, 2005.

Varshney, R.K., Graner, A., Sorrells, M.W.: Genic microsatellite markers in plants: features and applications. - Trends Biotechnol. **23**: 48-55, 2005.

Watcharawongpaiboon, N., Chunwongse, J.: Development and characterization of microsatellite markers from an enriched genomic library of cucumber (*Cucumis sativus*). - Plant Breed. **127**: 74-81, 2008.

Zhuang, F.Y., Chen, J.F., Staub, J.E., Qian, C.T.: Assessment of genetic relationships among *Cucumis* spp. by SSR and RAPD marker analysis. - Plant Breed. **123**: 167-172, 2008.