

ISSN 0002-3078



AZƏRBAYCAN MİLLİ ELMLƏR AKADEMİYASI
NATIONAL ACADEMY OF SCIENCES OF AZERBAIJAN
НАЦИОНАЛЬНАЯ АКАДЕМИЯ НАУК АЗЕРБАЙДЖАНА

MƏRUZƏLƏR

REPORTS

ДОКЛАДЫ

www.science.az

1
2016

MÜNDƏRİCAT

Fizika-riyaziyyat və texnika elmləri

<i>Hüseynov R.V., Axundov Ə.Y.</i> "Zəif" parabolik tənliklər sistemi üçün bir tərs məsələ haqqında.....	6
<i>Tenirbekov N.M., Baygereyev D.R.</i> İzotermik olmayan üçfazlı süzülmə məsələsi üçün global qoyuluşda adaptiv şəbəkə üsulu.....	11
<i>Əliyev T.A., Musayeva N.F., Süleymanova M.T., Qazızadə B.İ.</i> Küyün diferensial normal paylanma funksiyasının hesablanması üçün ədədi alqoritmləri.....	18
<i>Qasımova R.E., Hüseynov V.A., Hacıyeva B.T.</i> İfratgüclü maqnitləşmiş ulduzlarda neytrinolar tərəfindən elektron-pozitron cütlərinin yaranması proseslərində spin asimetriyaları.....	23
<i>Nəsirov V.İ., Bayramov R.B.</i> $K_{1-x}Cs_xNO_3$ ($x=0,015; 0,035; 0,045$) monokristallarında polimorf çevrilmələrin kinetikasi.....	29

Kimya

<i>Mübasov V.M., Əliyeva L.İ., Nəcəfova G.Ə., Hacıyeva S.Y., Mövsümova P.Ə., Əliyev B.M., Əlizadə A.A.</i> Müalicəvi naftalan neftinin quruluş xüsusiyyətləri və onun təmizlənməsinin perspektiv üsulları.....	34
<i>Əhmədova R.H., İbrahimov H.C., Rüstəmov M.İ., Kondratenko Y.V.</i> Mikroməsaməli silikoalüminofosfat katalizatorunun iştirakı ilə metanoldan olefinlərin sintezi prosesinin xüsusiyyətləri.....	40
<i>Leynalov E.B., Nəsibova G.Q.</i> Nanokarbon birləşmələri ilə kataliz olunan neft karbohidrogenlərinin aerobik oksidləşmə xüsusiyyətləri.....	44
<i>Əliyev Ö.M., Əjdərova D.S., Maqşudova T.F., Rəhimova V.M., Səlimova T.A.</i> $Cu_2S - PbCuSbS_3$ və $CuSbS_2 - Pb_5Sb_4S_{11}$ sistemlərində faza tarazlığı.....	49
<i>Məmmədov A.N., Tağıyev E.R., Babanlı M.B., Tağıyev D.B.</i> $YbTe - Sb_2Te_3 - Bi_2Te_3$ sisteminin Gibbs enerjisi və likvidus səthlərinin hesablanması və 3D modelləşdirilməsi.....	53

Yer elmləri

<i>Quliyev H.H.</i> Materiyanın azad halının səciyyələri haqqında.....	57
<i>Səhbazov E.Q., Yusifov R.Ə., Cabbarova K.Ş., Hüseynli X.X., Əliyev C.A., Hüseynov M.Ə.</i> İşlənmənin son mərhələsində olan neft yataqlarının istismarı zamanı yaranan duz çöküntülərinin qarşısının alınması üçün nanosistem.....	61

Biologiya və tibb elmləri

<i>Şarifova S.S., Əliyev R.T., Şahmuradov İ.Ə.</i> Müxtəlif pomidor növlərinin ESTs (Expressed Sequence Tags) ardıcılıqlarında sadə təkrarlanan motivlərin rast gəlmə tezliyi və tiplərinin <i>in silico</i> analizi.....	66
<i>Bayramov Ş.M.</i> <i>Brachypodium distachyon</i> bitkisinin cücərən və yetişməkdə olan dənələrində fosfoenolpiruvat karboksilaza fermentinin gen ekspressiyasının və zülal miqdarının tənzimlənməsi.....	71

Humanitar və ictimai elmlər

<i>Bağırov A.M.</i> Qloballaşma və ədəbiyyatların dialoqu.....	77
<i>Cabbarov F.R.</i> Qafqazın Baş hakimi A.M.Dondukov-Korsakovun dövründə (1882-1890-cı illər) Qafqaz administrasiyasının ermənilərə münasibətdə siyasəti.....	83
<i>Quliyeva X.N.</i> Multikulturalizm etikası - tərəqqi konsepsiyası.....	87

DOI: 575.1/2+577.21

IN SILICO ANALYSIS ON FREQUENCY AND TYPES OF SIMPLE SEQUENCE REPEATS IN EXPRESSED SEQUENCE TAGS (ESTs) OF DIFFERENT TOMATO SPECIES

S.S.Sharifova¹, R.T.Aliyev¹, Corresponding Member of ANAS I.A.Shahmuradov^{2,3}

The present study focuses on computational mining and survey of the frequency and types of simple sequence repeats in Expressed Sequence Tags (ESTs) of three different tomato species. A thousand ESTs sequences from each of *S.lycopersicum* L., *S.pennellii* Corell. and *S.habrochaites* S.Knapp & D.M.Spooner were retrieved from the National Center for Biotechnology Information (NCBI) database and searched for the presence of simple sequence repeat (SSRs) motifs. A total of 35 EST-derived SSRs within 32 unigens of *S.lycopersicum*, 51 SSRs within 46 unigens of *S.pennellii* and 64 SSRs in 55 unigens of *S. habrochaites* were identified. Trinucleotide repeats were the most abundant class of microsatellites in all the examined species (54%, 39% and 45% respectively). The hexanucleotide repeats were the second most frequent EST-SSRs in cultivated tomato (26%) and *S. pennellii* (27%), whereas trinucleotide motifs were followed by pentanucleotides in *S. habrochaites* (28%). The AGC/CTG motif was the most frequent among all types of repeats identified in *S. lycopersicum* ESTs, whereas the AAT/ATT and AT/AT were the most abundant in *S. pennellii* and AAATT/AATTT, ACC/GGT, AAT/ATT, ATC/ATG and AT/AT in *S. habrochaites* respectively.

Keywords: SSRs, tomato, repeats, unigene, ESTs

Introduction

The Solanum section of *Solanaceae* family contains the domesticated tomato (*Solanum lycopersicum* L.) and its 12 closest wild relatives [9]. Mating systems in tomato species varies from allogamous self-incompatible to facultative allogamous and self-compatible to autogamous and self-compatible, and has been proved to play an important role in their evolution. The self-incompatibility system in wild tomatoes has a strong relationship with the extent of outcrossing and allelic polymorphisms [8]. *Solanum pennellii* and *Solanum habrochaites* species have both self-incompatible and self-compatible populations and they are considered important species for improvement of cultivated tomato germplasms. *S. pennellii* is an important donor for its extreme stress tolerance and unusual morphology [1], whereas *S. habrochaites* possess many important traits for disease resistance, cold tolerance and several agronomic traits in tomato breeding [4]. All of the 13 species of Solanum have been proposed for sequencing by the SOL-100 project and are being sequenced (<http://solgenomics.net/organism/sol100/view>).

S.lycopersicum cv. Heinz 1706 and *S. pimpinellifolium* LA1589 were already sequenced and assembled by the International Tomato Genome Sequencing Consortium [11]. Such kinds of projects generate large amounts of sequence information which are being used extensively for analyzing genome structures and complexities and to mine useful genes and molecular markers [10].

The expressed sequence tags (ESTs) databases are important resources to develop functional SSR markers for genotyping, diversity and phylogenetic studies, genetic/association mapping and tagging of useful genes. In a last decade, *in silico* approaches to screen for simple sequence repeats from EST sequences become a practicable, less time consuming and inexpensive alternative. The present study focuses on computational mining and examining of the abundance and types of SSRs in ESTs sequences of three different tomato species.

Materials and Methods

A thousand EST sequences from each of *S.lycopersicum* L., *S.pennellii* Corell. and *S. habrochaites* S.Knapp & D.M.Spooner were re-

retrieved from the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>) and assembled using CAP3 programme (<http://doua.prabi.fr/software/cap3>) for identification of non-redundancy. The non-redundant unigene sequences were screened for the presence of EST-SSR motifs using the Micro Satellite identification tool (MISA) (<http://pgrc.ipk-gatersleben.de/misa>). SSR detection criteria was fixed at 7, 5, 4, 3, and 3 repeat units for di-, tri-, tetra-, penta-, and hexanucleotide motifs respectively. Mononucleotide repeats were not included in the SSR search criteria. For compound repeats the maximum default interruption length was set at 100 bp. The Primer 3 (<http://bioinfo.ut.ee/primer3-0.4.0/>) tool was used for checking a possibility to design primer pairs for identified SSR repeats.

Results and Discussion

A thousand EST sequences from each of *S. lycopersicum*, *S. pennellii* and *S. habrochaites* were used to computational analysis on frequency and types of EST-SSRs. All of the downloaded EST sequences were assembled with the CAP3 sequence assembly program resulting in production of 674 unigens including 57 contigs and 617 singletons in *S. lycopersicum*, 536 unigens consisted of 104 contigs and 432 singletons in *S. pennellii*, and 533 unigens comprising 261 contigs and 272 singletons in *S. habrochaites* respectively.

A total of 35 EST-derived simple sequence repeats within 32 sequences were identified after mining of 674 non-redundant unigens of cultivated tomato (*S. lycopersicum* L.). Three sequences contained more than one SSR and two motifs were found in the compound formation (table).

A total of 51 SSRs within 46 unigens were identified in 536 non-redundant sequences of *S. pennellii*, while 64 SSR repeats were found in 55 different unigens among of 533 sequences of *S. habrochaites*. Five and eight sequences were carrying more than one SSR and three and five motifs were found in the compound formation in *S. pennellii* and *S. habrochaites* respectively.

The frequency of simple sequence repeats in ESTs reflects the density of SSRs in the transcribed region of the genome. The frequency of the SSR repeats within all retrieved *S. lycopersicum* EST sequences was 3.5%, while this number was corresponded to 5.1% in *S. pennellii* and 6.4% in *S. habrochaites*. Since, random sequencing within cDNA libraries usually results in a high proportion of redundant ESTs, we performed SSR search after elimination of redundancy. Thus, only small fractions of screened non-redundant ESTs (5.2% in *S. lycopersicum*, 9.5% in *S. pennellii* and 12% in *S. habrochaites*) from the respective species contained SSR repeats, giving an average density of one SSR per 14.8 kb in *S. lycopersicum*, per 6.5 kb in *S. pennellii* and per 7.1 kb in *S. habrochaites* species. This indicates that SSR repeats frequency and density in wild species were much more than that of cultivated tomato. In other reports, an EST-SSRs have been observed to be correspond to one every 11.1 kb in tomato, 3.8 kb in pepper, 14.7 kb in lettuce, 13.8 kb in *Arabidopsis thaliana*, 3.4 kb in rice, 8.1 kb in maize, 7.4 kb in soybean, 20.0 kb in cotton and 14.0 kb in poplar [2, 14, 15]. The variations of frequencies among different studies were considered mainly due to the criteria used to identify SSRs, size of data set, database mining tools and EST sequence redundancy [12].

We classified all repeat types taking sequence complementarity into account and examined their occurrence. Trinucleotide repeats in our survey were the most abundant class of microsatellites in all the examined species (54%, 39% and 45% respectively). The hexanucleotide repeats were found to be the second most frequent EST-SSRs in cultivated tomato (26%) and *S. pennellii* (27%), whereas the most abundant trinucleotide motifs were followed by pentanucleotides in *S. habrochaites* (28%).

Compared to the cultivated tomato, *S. pennellii* and *S. habrochaites* EST sequences had the highest density of dinucleotide repeats with frequencies of 16% and 11% respectively. The result of high trinucleotide repeat frequency is in conformity with a number of studies.

trieved from the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>) and assembled using CAP3 programme (<http://doua.prabi.fr/software/cap3>) for identification of non-redundancy. The non-redundant unigene sequences were screened for the presence of EST-SSR-motifs using the Micro Satellite identification tool (MISA) (<http://pgrc.ipkgatersleben.de/misa>). SSR detection criteria was fixed at 7, 5, 4, 3, and 3 repeat units for di-, tri-, tetra-, penta-, and hexanucleotide motifs respectively. Mononucleotide repeats were not included in the SSR search criteria. For compound repeats the maximum default interruption length was set at 100 bp. The Primer 3 (<http://bioinfo.ut.ee/primer3-0.4.0/>) tool was used for checking a possibility to design primer pairs for identified SSR repeats.

Results and Discussion

A thousand EST sequences from each of *S. lycopersicum*, *S. pennellii* and *S. habrochaites* were used to computational analysis on frequency and types of EST-SSRs. All of the downloaded EST sequences were assembled with the CAP3 sequence assembly program resulting in production of 674 unigens including 57 contigs and 617 singletons in *S. lycopersicum*, 536 unigens consisted of 104 contigs and 432 singletons in *S. pennellii*, and 533 unigens comprising 261 contigs and 272 singletons in *S. habrochaites* respectively.

A total of 35 EST-derived simple sequence repeats within 32 sequences were identified after mining of 674 non-redundant unigens of cultivated tomato (*S. lycopersicum* L.). Three sequences contained more than one SSR and two motifs were found in the compound formation (table).

A total of 51 SSRs within 46 unigens were identified in 536 non-redundant sequences of *S. pennellii*, while 64 SSR repeats were found in 55 different unigens among of 533 sequences of *S. habrochaites*. Five and eight sequences were carrying more than one SSR and three and five motifs were found in the compound formation in *S. pennellii* and *S. habrochaites* respectively.

The frequency of simple sequence repeats in ESTs reflects the density of SSRs in the transcribed region of the genome. The frequency of the SSR repeats within all retrieved *S. lycopersicum* EST sequences was 3.5%, while this number was corresponded to 5.1% in *S. pennellii* and 6.4% in *S. habrochaites*. Since, random sequencing within cDNA libraries usually results in a high proportion of redundant ESTs, we performed SSR search after elimination of redundancy. Thus, only small fractions of screened non-redundant ESTs (5.2% in *S. lycopersicum*, 9.5% in *S. pennellii* and 12% in *S. habrochaites*) from the respective species contained SSR repeats, giving an average density of one SSR per 14.8 kb in *S. lycopersicum*, per 6.5 kb in *S. pennellii* and per 7.1 kb in *S. habrochaites* species. This indicates that SSR repeats frequency and density in wild species were much more than that of cultivated tomato. In other reports, an EST-SSRs have been observed to be correspond to one every 11.1 kb in tomato, 3.8 kb in pepper, 14.7 kb in lettuce, 13.8 kb in *Arabidopsis thaliana*, 3.4 kb in rice, 8.1 kb in maize, 7.4 kb in soybean, 20.0 kb in cotton and 14.0 kb in poplar [2, 14, 15]. The variations of frequencies among different studies were considered mainly due to the criteria used to identify SSRs, size of data set, database mining tools and EST sequence redundancy [12].

We classified all repeat types taking sequence complementarity into account and examined their occurrence. Trinucleotide repeats in our survey were the most abundant class of microsatellites in all the examined species (54%, 39% and 45% respectively). The hexanucleotide repeats were found to be the second most frequent EST-SSRs in cultivated tomato (26%) and *S. pennellii* (27%), whereas the most abundant trinucleotide motifs were followed by pentanucleotides in *S. habrochaites* (28%).

Compared to the cultivated tomato, *S. pennellii* and *S. habrochaites* EST sequences had the highest density of dinucleotide repeats with frequencies of 16% and 11% respectively. The result of high trinucleotide repeat frequency is in conformity with a number of studies.

Table

Results of microsatellite search in three tomato species

Parameters	Plant species		
	<i>Solanum lycopersicum</i>	<i>Solanum pennellii</i>	<i>Solanum habrochaites</i>
Total number of sequences examined:	1 000	1 000	1 000
Total size of examined sequences (bp):	767 424	598 857	598 857
Total number of non-redundant sequences:	674	536	533
Total size of non-redundant sequences (bp):	516 277	332 868	453 530
Total number of identified SSRs:	35	51	64
Total number of SSR containing sequences:	32	46	55
Total number of sequences containing more than 1 SSR:	3	5	8
Total number of SSRs present in compound formation:	2	3	5
Total number of dinucleotide motifs:	2	8	7
Total number of trinucleotide motifs:	19	20	29
Total number of tetranucleotide motifs:	1	4	3
Total number of pentanucleotide motifs:	4	5	18
Total number of hexanucleotide motifs:	9	14	7
Frequency of EST-SSRs	3.5%	5.1%	6.4%
Density of EST-SSRs (SSRs/kb)	14.8 kb	6.5 kb	7.1 kb

Varshney et al. reported that trinucleotide repeats in plants are the most common, followed by either dinucleotide or tetranucleotide repeats [13]. It is considered that expansion or deletions of trinucleotide and hexanucleotide repeats in coding region do not perturb reading frames and therefore less frequency of other repeat types in coding regions is a result of negative selection against frameshift mutations [7].

Among all types of repeats identified in *S. lycopersicum* ESTs, AGC/CTG, AAC/GTT, AAG/CTT and AAT/ATT units, considering sequence complementary, were the most dominant repeat types, while ACG/CGT was the less frequent. The most abundant repeat classes in *S. pennellii* were AAT/ATT, AAG/CTT and AAC/GTT among the trinucleotides with frequencies of 35%, 20% and 20% respectively, followed by ATC/ATG (15%). The AT/AT (75%) was the most frequent among dinucleotides, whereas TAAAA was the most dominant (80%) among pentanucleotides. The AAAAAT/ATTTTT motif was the most common among hexanucleotide repeats with a frequency of 21%.

In *S. habrochaites* AT/AT (71%) was found to be the most frequent repeat types among dinucleotides. In term of trinucleotide motifs, the most common types were ACC/GGT, AAT/ATT, AAG/CTT and ATC/ATG representing an average frequency of 31%, 21%, 17% and 17% re-

spectively. The AATAT/ATATT repeat motif had the highest frequency (67%) among identified types of pentanucleotides followed by ACCCG/CGGGT (22%).

The high frequency of AAG/CTT motif type was observed in all three species examined in our study that is in agreements with other researches, since several authors had been identified that AAG/CTT were the most frequent trinucleotide motif in majority of plants [6]. Besides, AAT/ATT motif was also among the most frequent motif types in all three species.

Among all types of identified trinucleotide motifs the AAC/GTT, AAG/CTT, AAT/ATT, AGC/CTG, ATC/ATG were observed in all three species, whereas ACG/CGT motif type was found only in EST sequences of cultivated tomato, CCG/CGG only in *S. habrochaites* and ACT/AGT in both of wild species. It was also observed that all types of identified motifs were AT-rich in all three species, which is in close agreement with a number of previous studies [3].

The location of SSRs and length of SSR flanking sequences determines the possibility of designing specific primers. About 1 to 5% of ESTs from different plant species have been found to contain SSRs suitable for marker development [5]. It proved to be difficult or impossible to design primers for many of the EST-SSRs in our study, but a number of them

can be used for this purpose and to develop of EST-SSR markers.

Simple sequence repeats derived from ESTs essentially represent expressed genic sequences and are candidates in the development of functional markers for comparative genomic studies. Obtained results on survey of simple sequence repeats demonstrate their abundance in expressed parts of genome and potential of ESTs for development of microsatellite markers by mining of available tomato databases.

This work was supported by the Science Development Foundation under the President of the Republic of Azerbaijan-Grant № EIF-2012-2(6)-39/03/3

REFERENCES

1. Bolger A., Scossa F., Bolger M.E. et al. The genome of the stress-tolerant wild tomato species *Solanum pennellii*. *Nature genetics*. 2014.
2. Cardle L., Ramsay L., Milbourne D. et al. Computational and experimental characterization of physically clustered simple sequence repeats in plants// *Genetics*, 2000, 156(2), 847-854.
3. Cavagnaro P.F., Senalik D.A., Yang L. et al. Genome-wide characterization of simple sequence repeats in cucumber (*Cucumis sativus* L.)// *BMC genomics*, 2010, 11(1), 569.
4. Ercolano M.R., Sebastiano A., Monti L. et al. Molecular characterization of *Solanum habrochaites* accessions// *J. of Genetics and Breeding*, 2005, 59(1), 15.
5. Kantety R.V., La Rota M., Matthews D.E., & Sorrells M.E. Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat// *Plant molecular biology*, 2002, 48(5-6), 501-510.
6. Kumpatla S.P., & Mukhopadhyay S. Mining and survey of simple sequence repeats in expressed sequence tags of dicotyledonous species// *Genome*, 2005, 48(6), 985-998.
7. Metzgar D., Bytof J., & Wills C. Selection against frameshift mutations limits microsatellite expansion in coding DNA// *Genome research*, 2000, 10(1), 72-80.
8. Peralta I.E., & Spooner D.M. Morphological characterization and relationships of wild tomatoes (*Solanum* L. Sect. *Lycopersicon*). *Monographs In Systematic Botany*, 2005, 104, 227.
9. Peralta I.E., Spooner D.M., & Knapp S. Taxonomy of wild tomatoes and their relatives (*Solanum* sect. *Lycopersicoideae*, sect. *Juglandifolia*, sect. *Lycopersicon*; *Solanaceae*). 2008
10. Sonah H., Deshmukh R.K., Singh V.P. et al. Genomic resources in horticultural crops: status, utility and challenges// *Biotechnology advances*, 2011, 29(2), 199-209.
11. Tomato Genome Consortium. The tomato genome sequence provides insights into fleshy fruit evolution// *Nature*, 2012, 485(7400), 635-641.
12. Varshney R.K., Graner A., & Sorrells M.E. Genic microsatellite markers in plants: features and applications// *TRENDS in Biotechnology*, 2005, 23(1), 48-55.
13. Varshney R.K., Thiel T., Stein N. et al. In silico analysis on frequency and distribution of microsatellites in ESTs of some cereal species// *Cellular and Molecular Biology Letters*, 2002, 7(2A), 537-546.
14. Yi G., Lee, J.M., Lee S., Choi D., & Kim B.D. Exploitation of pepper EST-SSRs and an SSR-based linkage map// *Theoretical and Applied Genetics*, 2006, 114 (1), 113-130.
15. Yu J.K., La Rota M., Kantety R.V., & Sorrells M.E. EST derived SSR markers for comparative mapping in wheat and rice// *Molecular Genetics and Genomics*, 2004, 271(6), 742-751.

¹Genetic Resources Institute of ANAS

²Institute of Botany of ANAS

³Azerbaijan Medical University
saidasharifzade@yahoo.com

MÜXTƏLİF POMİDOR NÖVLƏRİNİN ESTs (EXPRESSED SEQUENCE TAGS) ARDİCİLLİQLARINDA SADƏ TƏKRARLANAN MOTİVLƏRİN RAST GƏLMƏ TEZLİYİ VƏ TİPLƏRİNİN *IN SILICO* ANALİZİ

S.S.Şərifova, R.T.Əliyev, İ.Ə.Şahmuradov

Tədqiqat işi pomidorun üç müxtəlif növünün EST ardıcılıqlarında (ESTs- ekspressiya olunan ardıcılıq hissələri) mikrosatellit motivlərin rast gəlmə tezliyinin və tipinin kompüter proqramları vasitəsilə araşdırılması və təhlilinin aparılmasına həsr edilmişdir. Biotexnoloji Məlumatlar Milli Mərkəzinin (NCBI) məlumat bazasından *S. lycopersicum* L., *S. pennellii* Corell. və *S. habrochaites* S.Knapp & D.M.Spooner növlərinin hər birinin min EST ardıcılığı yüklənmiş və SSR motivlərinin tapılması üçün işlənmişdir. *S. lycopersicum* növünün 32 unigenində 35, *S. pennellii* növünün 46 unigenində 51 və *S. habrochaites* növünün 55 unigenində 64 mikrosatellit təkrarı aşkar edilmişdir.

Trinukleotid təkrarları bütün öyrənilən növlərdə ən zəngin mikrosatellit tipi olmuşdur (müvafiq olaraq 54%, 39% və 45%). Heksanukleotidlər, mədəni pomidor (26%) və *S. pennellii* növündə (27%) növbəti ən çox rast gəlinən təkrarlar olmuş, *S. habrochaites* növündə isə trinukleotidlərdən sonra ən çox müşahidə olunanlar pentanukleotid təkrarları (28%) olmuşdur. Ardıcılıqların komplementarlığını nəzərə aldıqda, AGC/CTG motivi *S. lycopersicum* EST ardıcılıqlarında ən çox müşahidə olunan təkrar tipi olmuşdur. *S. pennellii* növündə AAT/ATT və AT/AT, *S. habrochaites* növündə isə AAATT/AATTT, ACC/GGT, AAT/ATT, ATC/ATG və AT/AT ən çox rast gəlinən təkrar tipləri olmuşdur.

Açar sözlər: SSRs, pomidor, təkrarlar, unigen, ESTs

IN SILICO ANALIZ ЧАСТОТЫ И ТИПОВ ПРОСТЫХ ПОВТОРЯЮЩИХСЯ ПОСЛЕДОВАТЕЛЬНОСТЕЙ В ESTs (EXPRESSED SEQUENCE TAGS) ИЗ РАЗЛИЧНЫХ ВИДОВ ТОМАТА

С.С.Шарифова, Р.Т.Алиев, И.А.Шахмуратов

Данная статья посвящена компьютерной обработке и обследованию частоты и типов простых повторяющихся последовательностей в ESTs (Expressed Sequence Tags) из трех различных видов томата. В общем 1000 EST последовательностей от каждого из трёх видов томата (*S. lycopersicum* L., *S. pennellii* Corell. и *S. Habrochaites* S.Knapp & D.M.Spooner) из базы данных Национального Центра Биотехнологической Информации (NCBI) были взяты и исследованы на наличие SSR (single sequence repeats) мотивов. Всего было идентифицировано 35 EST производных SSRs в пределах 32 унигенов *S. lycopersicum*, 51 SSRs в пределах 46 унигенов *S. pennellii* и 64 SSRs в пределах 55 унигенов *S. habrochaites*. Тринуклеотидные повторы были наиболее распространенным классом микросателлитов для всех исследованных трёх видов (54%, 39% и 45% соответственно). Гексануклеотидные повторы оказались вторым по встречаемости EST-SSRs в культурном томате (26%) и *S. pennellii* (27%), далее следовали тринуклеотидные, а затем пентануклеотидные мотивы (28%) в *S. habrochaites*. AGC/CTG мотив, учитывая комплементарную последовательность, имел наибольшую частоту встречаемости среди всех типов повторов, выявленных в EST *S. lycopersicum*, в то время как AAT/ATT и AT/AT были наиболее распространены в *S. pennellii*, а AAATT/AATTT, ACC/GGT, AAT/ATT, ATC/ATG и AT/AT в *S. habrochaites*, соответственно.

Ключевые слова: SSRs, томат, повторы, униген, ESTs