Annals of Agrarian Science

Journal homepage: http://journals.org.ge/index.php

Draft annotation of stilbene synthase genes of Georgian grape varieties

I. Pipia*, V. Tabidze

Institute of Molecular Genetics, Agricultural University of Georgia, Kakha Bendukidze University Campus, # 240, David Aghmashenebeli Alley, Tbilisi, 0159, Georgia

Received: 6 September 2020; accepted: 21 September 2020

A B S T R A C T

In the times of genome-sequencing the bioinformatics is becoming the vital importance for the field of genomics. Here we present draft annotation of stilbene synthase (STS) candidate genes of four Georgian grape varieties - Chkhaveri, Saperavi, Meskhuri Mtsvane and Rkatsiteli by web-based system MEGANTE. It has shown the existence of STSs in the chromosomes 10 and 16. Abundance of STSs candidates is found on chromosome 16. The obtained data is caring significant information for the future gene-expression studies of STSs of Georgian grape varieties.

Keywords: Stilbene synthase genes, MEGANTE, Chkhaveri, Saperavi, Meskhuri Mtsvane, Rkatsiteli.

*Corresponding author: Dr. Ia Pipia: E-mail address: i.pipia@agruni.edu.ge

Introduction

In the last few decades, genome-based research has become a very popular and accessible approach for studying of the structures and functions of genomes. Identification of the primary genome structure, which involves the detection and analysis of genes and intergenic regions, is crucial for studying the molecular basis of that unique features that causes individuality of each species and also, for the assessment of phylogenetic links between different taxa of living organisms. Much of the data generated by genome sequencing is processed by bioinformatical programs and softweres, that provides scientists with more or less organized information about the nucleotide sequence organizations of the certain genomes. However, it should be noted that bioinformatical data mostly have to be checked and corrected by so called "manual curation" as mistakes/mismatches in machinery annotation is not a rare case.

To date the complete genomes of the numbers of plant species are sequenced and annotated. The amounts of those species are increasing day by day. Among fully sequenced and annotated genomes is the genome of grapevine (*Vitis vinifera* ssp. *vinifera*) – the World's most important agricultural crop [1-5]. According to the data received from the sequence-analyses of firstly fully sequenced grape variety - Pinot noir, *Vitis vinifera* ssp. *vinifera*, is a diploid (2n=38) taxa with 19 chromosomes in karyotype. Genome size in *V.vinifera* is relatively smaller than in other plants - 475-500 Mb. The length of mitochondrial and chloroplast genomes are 750,000 and 160,000 bp respectively [6, 7, 5, 8]. It is shown that the genome carries up to 29.000 genes [1,2].

Tabidze et al., [5] performed whole genome sequencing of well known Georgian grape varieties: Chkhaveri, Saperavi, Meskhuri mtsvane and Rkatsiteli by using Illumina HiSeq. Pinot Noir nuclear, mitochondrial and chloroplast DNA were used as reference. In the frame of the mentioned research among other results, length of chromosomes of Georgian grape cultivars (19 chromosomes for the each grape variety), the amounts of genes and SNPs were detected. Interestingly that in the same work the successful annotation of terpene synthase genes was conducted by using of web-based program MEGANTE – the method which we used in the presented research.

Stilbens represent a small group of phenylpropanoids, which are found in at least 72 non-relat-

ed plant species. Stilbenes are accumulated in cells in response to biotic and abiotic factors (i.e. stress infection, mechanical injury, ultraviolet radiation, chemicals). Well-known resveratrol – stilbene compound, who has the antioxidant activity, can increase resistance to pathogens (e.g., *Botrytis cinerea*, *Eutypa lata*, *Plasformin viticola*). Besides, stilbens are participating in plant-killer relationships, and also have the ability to slow down the progression of some diseases [9].

The major enzymes in the biochemical cascade of phenylpropanoids are stilbene synthases (STS). The main goal of the presented research was detection and annotation of STS genes of four Georgian grape varieties: Chkhaveri, Saperavi, Meskhuri Mtsvane and Rkatsiteli, by using web-annotation system – MEGANTE [10].

Materials and Methods

Web-annotation system MEGANTE (https://megante.dna.affrc.go.jp) was used for annotation of the stilbene synthases (STS) genes of Chkhaveri, Saperavi, Meskhuri mtsvane and Rkatsiteli. The system automatically run different programs and integrates the results to select the consensus exon-intron sites and to predict open reading frame (ORFs) regions. MAGANTE makes possible functional annotation, which includes searches of identities against known proteins and functional domains. The final data generated by the system can be downloaded what makes easy to use it.

Because MEGANTE has some limitations in the sizes of uploaded sequence-files, at the initial stage of the study, the sequences of chromosome 10 and 16 of Chkhaveri, Saperavi, Meskhuri mtsvane, Rkatsiteli and reference Pinot noir were fragmented (cutted) by the program ARTEMIS. For each grape genome ARTEMIS generated 5.000.000 bp long sequences what was corresponded to the MEGANTE's allowed size - 10 Mb. According to the chromosome lengths presented in the Table 1, all five genomes were cut into three fragments of 5.000.000 and one fragment of 3.000.000 bp lengths.

Chromosomal sequence containing nucleotide sequences of stilben synthases - chromosomes 10 and 16 of Georgian grape varieties were obtained from the link provided by Agricultural University of Georgia:

https://drive.google.com/open?id=0B4mIJZ9E-7ht7Q1dDTm9HbGFIR1U.

As reference genome sequences of Pinot noir were used (Genoscope:

http://www.genoscope.cns.fr/externe/GenomeBrowser/Vitis).

The annotation process included the numbers of steps. More concretely: Finding the appropriate consensus exon-intron structures for predicting open reading frames (ORFs) at each locus and making functional annotation, including a similarity search against known proteins and a functional domain search for the predicted ORFs.

Results and Discussion

In our previous work sequence of all 19 chromosomes of four Georgian grape cultivars was determined by resequencing using Pinot noir as a reference genome [5]. Chromosomal lengths of chromosome 10 and 16 where existence of STS candidate genes were found are representing in Tab.1.

49 candidate STS genes were predicted on chromosomes 10 and 16 of all studied Georgian cultivars by web-annotation program MEGANT. Among them, on chromosome 10, only one STS gene was observed in Chkhaveri, Rkatsiteli and Meskhuri mtsvane. At the same time, the presence of two STS genes was detected in Pinot noir, and none was found in Saperavi. Differences in gene number were observed on chromosome 16 of all Georgian grape cultivars. 14 STS genes were found in Meskhuri Mtsvane, 13 in Saperavi, 10 in Chkhaveri and 9 in Rkatsiteli (Table 2).

Table 1. *Chromosomal lengths of Georgian grape varieties according to Tabidze et al., [5].*

Chromosome	Rkatsiteli (bp)	Saperavi (bp)	Meskhuri Mts. (bp)	Chkhaveri (bp)	Pinot noir (bp)
16	22,021,800	22,050,608	22,056,729	22,057,692	22,053,297
10	18,122,339	18,139,422	18,146,577	18,143,602	18,140,952

Table 2. *Numbers of annotated stilbene synthase candidate genes in chromosomes 10 and 16.*

Cultivar	Chromosome 10	Chromosome 16
Chkhaveri	1	10
Meskhuri Mtsvane	1	14
Rkatsiteli	1	9
Saperavi	0	13
Pinot noire	2	12

STS-predictions of Chromosome 10

Single STS gene was detected in chromosome 10 of the tree Georgian grape varieties: Meskhuri Mtsvane, Chkhaveri, Rkatsiteli by system ME-GANTE. Predicted STSs genes were found in the range of 14.203.661 – 14.285.672 bp of chromosome 10. At the same time, the presence of two STS genes was detected in Pinot noir, and none was found in Saperavi (Table 3).

By the maximum-likelihood method phylogenetic tree of stilbene synthase candidate genes on chromosome 10 was built. As seen on Fig. 1, there are two main branches on the tree: one of them (on the bottom of Fig) is represented just with P2 or Pinot noir's STS candidate gene found in the sequence region 14.284.187-14.285.672 bp (Table 3). The second branch (on the top of Fig.) contains all other STS candidate genes which are more closely related to each other.

Table 3. *Genome regions of stilbene synthase candidate genes in chromosome10. Numbers represent bp range of the certain genome region, numbers in brackets indicate the first letter of cultivar and the number of candidate gene, ND - means "not detected".*

Genome Regions on Chromosome 10				
Pinot noir	Meskhuri Mtsvane	Chkhaveri	Saperavi	Rkatsiteli
			ND	14.203.661-14.205.226 (R1)
	14.221.262-14.222.827 (M1)		ND	
14,263,951-14,306,350 (P1)			ND	
		14.266.493-14.312.372 (C1)	ND	
14.284.187-14.285.672 (P2)			ND	

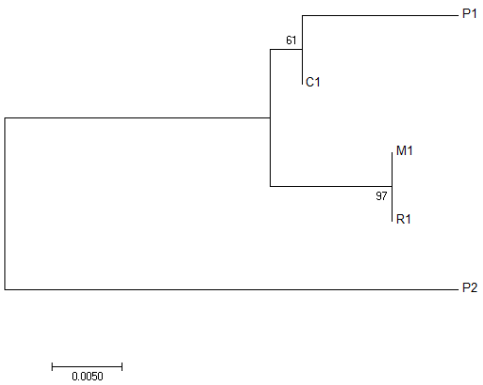

Fig. 1. *Molecular phylogenetic analysis of stilbene synthase candidate genes on chromosome 10 by maximum-likelihood method. Analyses were conducted in MEGA7 [11]*

STS-predictions of Chromosome 16

Table 4 represents positions of predicted Stilbene synthase candidate genes in chromosome16. Differences in gene number were observed on chromosome 16 of all Georgian grape cultivars. 14 STS genes was found in Meskhetian green, 10 in Chkhaveri and 9 in Rkatsiteli and 13 in Saperavi. Position in chromosome also represents the genomic organization of the STS gene cluster, where several additional genes, which are not detected in the Pinot noir reference genome, were identified. Comparative genomics reveals the presence of gene collinearity in a wide range of organisms from different cultivars within a species to genetically distinct genomes. Populations within a species and closely related organisms have extensive regions of gene collinearity. Comparative analysis of Georgian grape samples reflects a high level of similarity among those genes, which are located in similar positions on the chromosome, and confirmed the fact that all these genes are orthologous. The difference in gene number between the analyzed grape cultivars, especially the detection of additional genes that are absent in reference DNA, may help to understand the different volatile profiles in different Georgian grape cultivars.

Conclusion

The presented work represents the first large-scale attempt of STS gene-annotation of Georgian grape varieties. The results offer the new information concerning predicted molecular structures of mentioned genomic regions via providing their STS candidate gene sequences. In the future obtained results can be facilitating for gene-expression studies.

References

[1] Jaillon O, Aury J-M, Noel B, Policriti A, ClepetCh Casagrande A, Choisne N, Aubourg S, Vitulo N, Jubin C, Vezzi A, Legeai F, HugueneyPh Dasilva C, Horner D, Mica E, Jublot D, Poulain J, Bruyère C, Billault A, Segurens B, Gouyvenoux M, Ugarte E, Cattonaro F, Anthouard V, Vico V, Cr Del Fabbro, Alaux M, Di Gaspero G, Dumas V, Felice N, Paillard S, Juman I, Moroldo M, Scalabrin S, Canaguier A, Le Clainche I, Malacrida G, Durand E, Pesole G, Laucou V, ChateletPh Merdinoglu D, Delledonne M, Pezzotti M, Lecharny A, Scarpelli C, Artiguenave F, Pè ME, Valle G, Morgante M, Caboche M, Adam-Blondon A-B, Weissenbach J, Quétier F,Wincker P.,The

Table 4. *Genome regions of predicted Stilbene synthase candidate genes in chromosome16: Numbers represent bp range of the certain genome region.*

Genome Regions on Chromosome 16				
Pinot noir	Meskhuri mtsvane	Chkhaveri	Saperavi	Rkatsiteli
				12.622.981-12.624.590
13.636.905-13.638.529				
		16.242.833-16.244.650		
16.268.745-16.270.647		16.272.686-16.290.362	16.267.430-16.285.104	
16.276.509-16.278.378	16.279.101-16.280.970			
16.335.626-16.337.351	16.338.201-16.339.926	16.339.530-16.341.255	16.334.268-16.335.993	16.316.372-16.318.097
				16.378.850-16.380.652
16.385.834-16.387.530	16.388.483-16.390.180		16.384.530-16.386.226	
16.398.235-16.400.037	16.407.841-16.415.954	16.402.185-16.402.362	16.396.933-16.398.735	
				16.447.349-16.459.201
16.466.725-16.478.614	16.469.648-16.471.182	16.470.641-16.482.501	16.465.424-16.467.236	
	16.479.686-16.481.202		16.475.749-16.477.265	16.471.859-16.489.947
16.491.316-16.493.132	16.493.885-16.495.700	16.495.151-16.513.265	16.489.897-16.491.712	
16.503.637-16.509.480	16.506.205-16.512.041		16.502.237-16.508.080	
16.511.156-16.512.830	16.513.717-16.515.391	16.514.941-16.516,615	16.509.756-16.511.430	
	16.528.433-16.530.348	16.529.638-16.531.553	16.524.500-16.526.415	
		16.559.284-16.561.126		16.535.728-16.537,570
16.555.671-16.557.513	16.558.138-16.559.980		16.554.143-16.555.985	
	16.576.448-16.585.497			
	16.589.900-16.591.437	16.590.963-16.592.500	16.585.902-16.587.425	16.587.795-16.604.156
16.607.177-16.608.743	16.609.684-16.611.250		16.605.617-16.607.183	16.607.012-16.608.647
				16.653.181-16.655.099

grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla, Nature, 449 (2007) 463–468.

[2] Velasco, R., Zharkikh, A., Troggio, M., Cartwright, D.A., Cestaro, A., Pruss, D., Pindo, M., Fitzgerald, L.M., Vezzulli, S.,Reid, J., et al., A high quality draft consensus sequence of the genome of a heterozygous grapevine variety, 2007, PLoS ONE 2 (12), e1326 https://doi.org/10.1371/journal.pone.0001326. PubMed.

[3] Di Genova, A., Almeida, A.M., Muñoz-Espinoza, C., Vizoso, P., Travisany, D., Moraga, C., Pinto, M., Hinrichsen, P.,Orellana, A., and Maass, A., Whole genome comparison between table and wine grapes reveals a comprehensive catalog of structural variants, BMC Plant Biol. 14 (1), 7 (2014) https://doi.org/10.1186/1471-2229-14-7.

[4] Gambino, G., Dal Molin, A., Boccacci, P., Minio, A., Chitarra, W., Avanzato, C.G., Tononi, P., Perrone, I., Raimondi, S., Schneider, A., et al., Whole-genome sequencing and SNV genotyping of 'Nebbiolo' (Vitis vinifera L.) clones. Sci. Rep. 7 (1), 17294 (2017). https://doi.org/10.1038/s41598-017-17405-y. PubMed.

[5] Tabidze, V., Pipia, I., Gogniashvili, M., Kunelauri, N., Ujmajuridze, L., Pirtskhalava, M., Vishnepolsky, B., Hernandez, A.G., Fields, C.J., and Beridze, T., Whole genome comparative analysis of four Georgian grape cultivars. Mol. Genet. Genomics 292 (6), (2017) 1377–1389 https://doi.org/10.1007/s00438-017-1353-x. PubMed.

[6] Pipia I., Roma L., Tabidze V. et al.,Annotation of whole-plastid genomes of wild grapes (Vitis vinifera subsp. sylvestris) from the south Caucasus, Europe and the Mediterranean Basin. Acta Hortic. 1248, 2019. ISHS 2019. DOI 10.17660/ActaHortic.2019.1248.34.

[7] Jansen R.K., Kaittanis C., SaskiC., Lee S.-B., TomkinsJ., Alverson A. and Daniell H., Phylogenetic analyses of Vitis (Vitaceae) based on complete chloroplast genome sequences: effects of taxon sampling and phylogenetic methods on resolving relatishipships among rosids. BMC Evolutionary Biology 6 (1) (2006) 32.

[8] Vazha Tabidze, Grigol Baramidze, Ia Pipia, Mari Gogniashvili, Levan Ujmajuridze, Tengiz Beridze, Alvaro G Hernandez, Barbara Schaal, The complete chloroplast DNA sequence of eleven grape cultivars. Simultaneous resequencing methodology. OENO One, Vol. 48 No. 2 (2014) DOI: https://doi.org/10.20870/oeno-one.2014.48.2.1568.

[9] Vannozzi, A., Dry I.B., Fasoli M., Zenoni S., Lucchin M., Genome-wide analysis of the grapevine stilbene synthase multigenic family: genomic organization and expression profiles upon biotic and abiotic stresses, BMC Plant Biol., 12, 130, 2012. https://doi.org/10.1186/1471-2229-12-130

[10] Hisataka Numa, Takeshi Itoh, (2013). MEGANTE: A web-based system for integrated plant genome annotation. Plant and Cell Physiology, Volume 55, Issue 1, January 2014, Page e2. https://doi.org/10.1093/pcp/pct157

[11] Sudhir Kumar, Glen Stecher, Koichiro Tamura, MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol 2016 Jul;33(7),1870-4 (2016), doi: 10.1093/molbev/msw054.