Genetic diversity, variety identification and gene detection in some Egyptian grape varieties by SSR and SCoT markers

Shafik D. Ibrahim[*], S. S. Adawy, M. A. M Atia, A. M. Alsamman, M. M. Mokhtar

Agricultural Genetic Engineering Research Institute (AGERI), ARC, Giza, Egypt

*Corresponding author: shafikdarwish@ageri.sci.eg

Abstract

Leaves of seven Egyptian grape varieties represent different range of agronomic and morphological traits were genotyped using Start-codon Targeted (SCoT) and Simple Sequence Repeats (SSR). The 24 SCoT primers generated 362 total fragments with 77.10% of polymorphism and 0.04 of average PIC. Dice coefficient that shows the genetic similarity and relationship was also used between the seven varieties. On one hand, SCoT analysis successfully characterized 73 unique positive and negative markers differentiating between the rootstock varieties especially those with green and red fruits. On the other hand, the seven SSR primers produced 73 fragments with 86.30% total polymorphism and 0.14 of average PIC. It also provided 19 unique positive and negative markers differentiating between the rootstock varieties. SO4 variety was identified by the highest number of positive unique markers (5). Both Scot and SSR analysis had covered 0.02% (0.10 Mbp) of *V.vinifera* genome. The SCoT covered 0.96 Mbp and SSR covered 846 bp. Together, the covered region encompassed about 22 genes of grape genome. Further, eight SCoT polymorphic bands were purified, cloned and sequenced. Of which, four SCoT sequences ($SCoT3_{600}$, $SCoT4_{450}$, $SCoT6_{200}$ and $SCoT12_{550}$) showed high similarity to some potential genes.

Keywords: SCoT, SSR, *in silico* PCR, phylogenetic, grape, genes.
Abbreviations: PA_polymorphic amplicons, SCoT_start-codon targeted, %P_ percentage of polymorphism of calculated amplicons, PIC_polymorphism information content, UPGMA_unweighted pair group method with arithmetic mean, SSR_Simple sequences repeat, IAG_in silico amplimers with genes, NPC4_non-specific phospholipase C4-like, PEPC1_phosphoenolpyruvate carboxylase 1-like, Chr_Chromsosme, AG_AGAMOUS, PCR_polymerase chain reaction, SSR_simple sequence repeat, SNP_single nucleotide polymorphism, DGAT_diglyceride acyltransferase, PD_membrane-bound phytoene dehydrogenase, TSA_Transcriptome Shotgun Assembly, NGST_next generation sequencing technologies.

Introduction

The size of the global consumption market of grape (*Vitis vinifera)* commodity is 68 million tonnes. In this market, the average of Egyptian share is 1.37 million tonnes, making grapes the second most important fruit commodity in Egypt after citrus (FAOSTAT 2013). Some problems make the characterization between the grape varieties is important and essential for germplasm certification, improvement and conservation. Such problems in modern grape varieties are the high heterozygosity, hermaphroditic, long juvenile period (3-4 years) and variety synonym (Alleweldt and Possingham 1988;Moreno-Sanz et al., 2011; This et al., 2006;). Finding easy and economic methods to characterize grape varieties will facilitate the identification of genetic quality and agronomic attributes. The SSRs are tandomly repeated DNA sequences, typically of 2–4 bp. Repetition of such markers is evenly distributed throughout the eukaryotic genome showing high allelic diversity and inherited in a codominant fashion (Powell et al., 1996). About 41.40% of the grape genome is composed of repetitive elements (Jaillon et al., 2007).This enabled many successful application of SSR markers in grape diversity (Liu et al., 2012), germplasm management studies (Laucou et al., 2011) and could contain SNP variation between different varities (Mokhtar et al., 2016). SCoT is a new molecular tool targeting plant genes using DNA markers designed based on the short conserved

regions flanking the ATG start codon (Collard and Mackill 2009). Besides, it is simple, low-cost, and highly polymorphic that provides extensive genetic information using universal primers in plants. Moreover, it provides reproducible dominant and single primer amplification reaction like RAPD. These advantages were validated in many studies of genetic diversity in mango (Luo et al., 2010) and date palm (Adawy et al., 2014). The objectives of current study are: 1) to present the potential of SCoT and SSR markers in evaluating the genetic diversity and varietal discrimination between the Egyptian grape varieties; 2) to compare between both marker types according to primer selectivity, genome coverage and the ability to target genetic regions using *in silico* PCR analysis; and 3) to detect new genes through sequencing of potential and polymorphic PCR fragments. Identification of such new genes will depend on gene annotation and similarity with known genes through the application of bioinformatics analysis.

Results and Discussion

The polymorphism analysis of SCoT markers

The 24 SCoT primers generated a band profile of 362 fragments with average of 15 fragment/primer. The fragment

sizes were ranged between 140 and 2,150 bp (Table S1). The maximum number of amplicons (22) was produced by primer SCoT-01, while the minimum number was nine amplicons and achieved by SCoT-20 primer. However, the total number of PA was 279 with an average 11.6 polymorphic band/primer. SCoT-17 and SCoT-01 primers produced the highest number of PA (18).The %P was ranged between 94.7% (SCoT-17) and 46% (SCoT-19). Therefore, the total percentage of amplicons polymorphism was 77.10%. SCoT-27 produced the maximum PIC (0.065), while the SCoT-5 produced the minimum PIC (0.0167). This makes the average of PIC/primer is 0.04 (Table S1).

In some grape varieties, 17 out 36 SCoT primers generated clear amplification products and polymorphisms (Guo et al., 2012). They found the average number of bands per primer was 7.2, the fragment sizes were ranged between 100 and 2,000 bp, and the PIC value between 0.75 and 0.91. In mango, the average number of PCR fragments was 8.27\primer and the sizes were ranged between 250 and 4000 bp (Luo et al., 2010). Similar results were reported by others (Guo et al., 2012; Zeng et al., 2014; Feng et al., 2015). In this study, to examine the genetic relationships among the grape varieties, the Dice coefficient and UPGMA methods were used to show the similarity matrices and generate the phylogeny tree (Fig. 1). The highest similarity (85%) was detected between the two varieties Thompson and Early Superior. This was followed by 84% of similarity between the two varieties Flame and Red Globe, while the lowest similarity (62%) was detected between the variety Early Superior and both SO4 and Ramsey (Table S3). In current study, SCoT analysis was effective in differentiating between the rootstock varieties (SO4, Ramsey and Dog Ridge), the green-fruit table varieties (Early superior and Thompson), and the red-fruit varieties (Fig. 1). This may reflect the remote relationships among these varieties. SCoT and cluster analysis provided similar results in wild *Vitis* species that were originated from America and China, however, four main groups were presented (Guo et al., 2012). Similarly in potato and peanut, SCoT markers were effective for fingerprinting, detecting DNA polymorphism, and studying genetic relationships among closely related genotypes (Gorji et al., 2011; Xiong et al., 2011). In current study, SCoT markers have successfully characterized all tested grape varieties through unique positive and/or negative markers (Table 1). The size of these markers was ranged between 170 and 2090 bp (Table 1). A total number of 73 unique markers were identified by all SCoT primers (Table 1). The highest number of positive and negative unique markers (7 and 20, respectively) was identified in Dog Ridge variety, while the solely negative unique markers (14 and 8) were identified in the varieties Ramsey and Thompson, respectively (Table 1).

The polymorphism analysis of SSR markers

The seven SSR primers have generated reproducible profile of total 73 fragments, average of 10.4 fragment/primer, and fragment sizes ranged between 110 and 860 bp (Table S2). The SSR-21 primer produced the maximum number of amplicons (14), while SSR-24 produced the minimum (8). However, the SSR-21 primer produced 63 total number of PA, 9 polymorphic fragments/primer, in average, and maximum number of PA (13). The percentage of polymorphism of calculated amplicons was ranged between 100% (SSR-07) and 62.5% (SSR-24), where the total percentage of amplicons polymorphism was 86.3%. The SSR-17 primer produced the maximum PIC (0.18), while

SCoT-07 primer produced the minimum PIC (0.07), therefore, the average of point/primer was 0.14 (Table S2).

In 38 Italian grapevine varieties, six microsatellite loci revealed PIC as low as 0.06-0.12 (Zulini et al., 2015). Another six SSR primers were used to fingerprint 25 grape varieties from Perú and Argentinaand produced heterozygosity between 71% and 89%, and the PIC ranged between 0.70 and 0.88 (Martinez et al., 2006). Also, 10 SSR primers genotyped 61 Tunsian grape varities and proved that the number of alleles per locus ranged from 6 to 11, and the PIC rangedfrom 0.86 to 0.91 (Zoghlami et al.,2009).The variation in reported PIC values could be due to the different number of varietes used in these studies.

Based on Dice coefficient and UPGMA analysis, the highest similarity (79%) was detected between the varieties Flame and Red Globe (Table S3 and Fig. 1). This was followed by 78% of similarity between the varieties Thompson and Flame, while the lowest genetic similarity (48%) was detected between the varieties Red Globe and Dog Ridge (Table S3). The SSR analysis in current study has enabled to differentiate rootstock varieties (SO4, Ramsey and Dog Ridge) from other studied varieties. This supports the efficacy of SSR technique to differentiate and study the genetic relationship in Tunisian, Spanish, and French grapes (Zoghlami et al., 2009; Moreno-Sanz et al., 2011; Pelsy et al., 2010). The power of SSR technique was most effective to characterize all tested grape varieties through the production of unique positive and/or negative markers (Table 2). These markers ranged in size from 150 to 560 bp. A total number of 19 unique markers were identified by using all the SSR primers. Dog Ridge was the only variety to be identified by two negative unique markers and with no positive markers. Other varieties were identified by only positive markers such as Flame and Red Globe. SO4 variety was identified by the highest number (5) of positive unique markers, while Flame variety was identified by the lowest number (1) of positive unique markers.

Combined molecular data analysis

The cluster analysis of the seven grape genotypes for SSR and SCoT techniques had shown unique topology. PCR scored data, of SSR and SCoT, were combined to raise the rate of accurate relationships (Table S3 and Fig. 1).The highest genetic similarity (83%) was detected between the cultivars Early Superior and Thompson and between Flame and Red Globe. This was followed by 80% of similarity between the cultivars SO4 and Ramsey. Meanwhile, the lowest genetic similarity (61%) was detected between the cultivars Early Superior and both of SO4 and Ramsey (Table S3).The combine phylogeny supports the structure revealed by SCoT analysis (Fig. 1).This could be due to the high number of PCR patterns revealed by SCoTs.

In silico PCR result analysis revealed by SCoT and SSR markers

Out of the 24 SCoT and the seven SSR primers, eight SCoT and three SSR primers have revealed *in silico* amplicons. By then, the total number of fragments proposed was 229 with an average of 20.8 fragment/primer, giving a product size ranged between120 and 1,473 bp (Table 3 and Fig. 2). Whereas, the virtual SCoT analysis produced 226 *in silico* amplicons (IA) while SSR proposed three IA. SCoT revealed higher average of IA/primer (28.3) than SSR (1) (Table 3).The two different marker techniques collectively covered

Table 1. Grape varieties characterized by unique positively and/or negatively SCoT markers identifying each variety.

Variety name	Unique negative markers			Unique positive markers			
	PN	MBS	TMC	PN	MBS	TMC	GT
SO4	SCoT-16	1453	1			8	9
				SCoT-06	585		
				SCoT-29	340		
				SCoT-21	170		
				SCoT-13	460		
				SCoT-05	1183&600		
				SCoT-20	920		
				SCoT-03	408		
Red Globe	SCoT-29	815	3			3	6
	SCoT-04	567					
	SCoT-10	730					
				SCoT-01	170		
				SCoT-16	1161		
				SCoT-02	970		
Dog Ridge	SCoT-08	596	7			13	20
	SCoT-17	828					
	SCoT-04	650&445					
	SCoT-27	858					
	SCoT-10	1270					
	SCoT-14	958					
				SCoT-24	765&555		
				SCoT-01	485&175		
				SCoT-25	275		
				SCoT-14	730&1695&300		
				SCoT-03	250		
				SCoT-23	765&555		
				SCoT-05	375		
				SCoT-22	360		
Flame	SCoT-27	750	1			4	5
				SCoT-17	751		
				SCoT-01	736		
				SCoT-27	2090		
				SCoT-13	405		
Early Superior	SCoT-24	800	2			9	11
	SCoT-23	800					
				SCoT-28	310		
				SCoT-12	1030		
				SCoT-21	330		
				SCoT-17	1520&1010		
				SCoT-19	564		
				SCoT-22	408&140		
				SCoT-09	1520		
Ramsey			0			14	14
				SCoT-28	1400		
				SCoT-24	1020&700&520		
				SCoT-06	530		
				SCoT-25	470		
				SCoT-13	630		
				SCoT-03	890		
				SCoT-08	730		
				SCoT-21	440		
				SCoT-23	1020&700&510		
				SCoT-16	1254		
Thompson			0			8	8
				SCoT-28	351		
				SCoT-24	421		
				SCoT-29	631		
				SCoT-17	590&375		
				SCoT-23	420		
				SCoT-22	898		
				SCoT-09	1164		
Total			14			59	73

Primer name (PN), size of marker Band (MBS), total number of markers (TMC), grand total markers (GT).
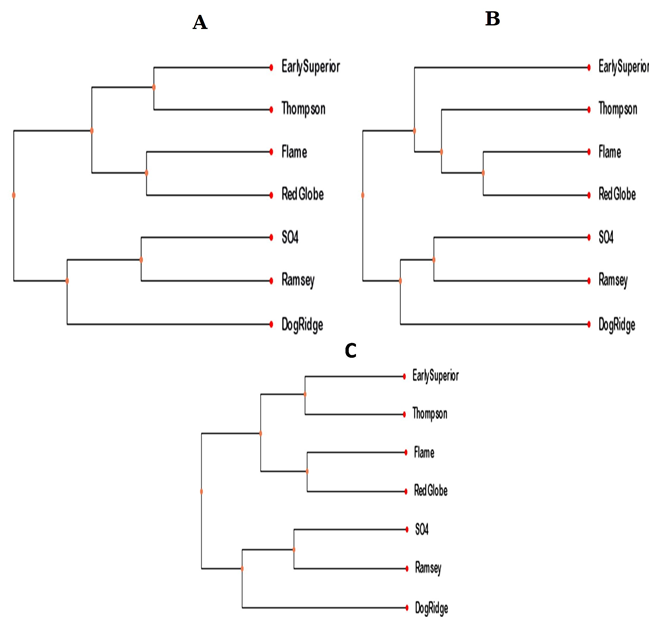
Fig 1. Dendrogram for the 7 grape varieties constructed from the SCoTs (A), SSR (B) and combined (C) data using UPGMA and similarity matrices computed according to Dice coefficient.

about 0.025% (0.10Mbp) of the grape genome; where SCoT and SSR amplimers cover 0.96 Mbp and 846 bp, respectively (Table 4).The chromosomal coverage revealed by SSR and SCoT amplimers across grape chromosomes were shown in Table 4 as well. Given that, the distribution of the *virtual* amplimers revealed by SCoT and SSR amplimers across grape genome is shown in figure 4. SCoT-20 marker covered the largest genomic area, while SSR-07 gave only one amplimer located on grape chromosome Vv7 (red font).Out of the proposed 229 amplimers, only 52 were IAG (that are close or adjoin genes). The SCoT and SSR have given 51 and one amplification, respectively. Having 22 genes as the actual number, the SCoT was the highest (21 genes) technique of targeting and SSR was the lowest (1 gene) (Table 3).

Bioinformatics analysis for sequenced PCR bands

Out of the eight SCoT-cloned bands, only four were successfully sequenced ($SCoT3_{600}$, $SCoT4_{450}$, $SCoT6_{200}$, and $SCoT12_{550}$). The BLASTs analysis (NCBI, website) was used to find local similarities and compare inquired nucleotide sequences against public databases. The $SCoT3_{600}$, $SCoT4_{450}$ and $SCoT6_{200}$ sequences gave high similarities and coverage with grape genome regions. About 55 grape genes similar to $SCoT3_{600}$ sequence were distributed across grape genome, except for chromosomes 1, 6, 8 and 16. This suggests a conserved domain shared among all these genes. In a window of 1kb, the closest genes to $SCoT3_{600}$ were located in Chr 4, 18, 17 with no known functions. In a window of 5 kb, some genes have been detected such as *NPC4* (Chr15), *PEPC1* (Chr19) and *AG* (Chr10) (Fig. 3). The *NPC4* is a plasma membrane-bound isoform of NPC protein, which plays a role in membrane rearrangement during phosphate starvation and abiotic stresses (Pejchar et al., 2015). The *PEPC1* is another peptide that initially helps catalyse and fix the atmospheric CO_2in C4 plants (Ku et al., 1999). The *AG* protein is important for carpel identity and ovule and flower development (Ray et al., 1994). In a window of 12 kb, matching sequences to $SCoT3_{600}$ and $SCoT4_{450}$ were detected such as cytochrome-like protein (Chr18), lysine-tRNA ligase-

like protein, transcript variant X_2 (Chr15) ,mediator of RNA polymerase II, transcription factor (Chr1), O-acyltransferase WSD1 (Chr15), phytoene dehydrogenase, (Chr12), probable serine/threonine-protein kinase (Chr17), probable WRKY transcription factor 72, (Chr17),putative disease resistance RPP13-like protein (Chr17),tRNA-Glu (Chr17) and tRNA-Val (Chr3) (Fig. 3-C). In Arabidopsis, the DGAT enzyme has a *WSD1* type, which is capable of using C18, C24 and C28 alcohols and C16 fatty acid to produce wax esters (Li et al.,, 2008). The PD is an enzyme in carotenoid and pigment biosynthesis (Schmidt and Sandman 1990). The serine/threonine-protein kinase family has been functionally characterized. The functions include signal ling auxin, brassinosteroid, phototropism; regulating rapamycin pathway, and cell stress responses (Uhrig et al., 2013). The WRKY is a transcriptional regulator and integral signalling molecule in plants (Rushton et al., 2010). Plants having avirulence proteins are guarded through the indirect interaction with the pathogen's virulence proteins. This triggers the defence systems including the hypersensitive responses, which restricts the pathogen growth (Bittner-Eddy et al., 2000). Although, $SCoT4_{450}$ sequence has matched hits with eight regions on grape genome, all were having unknown functions. However, $SCoT6_{200}$ sequence matched hits for the transcription factor *UPBEAT1* and another unknown gene in a window of 5kb and 12kb, respectively.*UPBEAT1* is modulating the balance between cellular proliferation and differentiation in root growth (Tsukagoshi et al., 2010). The alignment of $SCoT12_{550}$ sequence against flowering plants database retrieved maximum coverage of 65% for different putative genes. Such putative genes include ribonuclease H protein and reteroelements (Fig. 3). The retrotransposons are retrovirus-related mobile sequences that have the potential to replicate via RNA intermediates and increase the genome size by insertion into new sites. Such elements have been reported in cultivated grapes as a key role in fruit colour variation and abiotic stresses (Cadle-Davidson and Owens 2008; Tillett et al., 2011). TSA is an archive of computationally assembled

314

Table 2. Grape varieties characterized by unique positively and/or negatively SSR markers.

Variety name	Unique negative markers			Unique positive markers			
	PN	MBS	TMC	PN	MBS	TMC	GT
SO4			0	SSR-27 SSR-21 SSR-24	270&206&225 190 260	5	5
Flame			0	SSR-27	256	1	1
Red Globe			0	SSR-27 SSR-21	264 390	2	2
Dog Ridge	SSR-06 SSR-07	560 180	2			0	2
Early Superior			0	SSR-17 SSR-08	410&390 190	3	3
Ramsey			0	SSR-27 SSR-24	244 299	2	2
Thompson			0	SSR-07 SSR-21	470&440&150 280	4	4
Total			2			17	19

Primer name (PN), size of marker Band (MBS), total number of markers (TMC) and grand total markers (GT) identifying each variety.
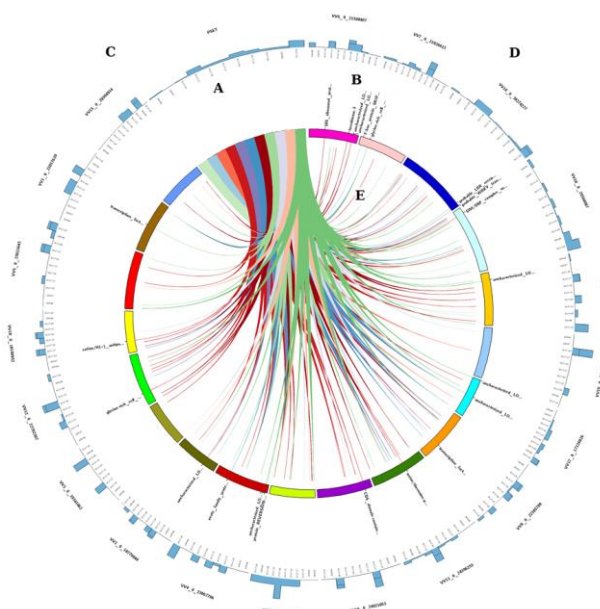

Fig 2. The *in silico* PCR results for the SCoT and SSR primers set (A) against *V. vinifera* genome, revealing the possible adjoined genes (B), primer total genome coverage percentages statics (C) , primer total chromosome coverage percentages statics (D) and the position of possible PCR amplimers with genes (extended lines) or without (short lines) (E).

sequences of primary data of ESTs and NGST. Alignment using Blastn and BlastX against the TSA database, results revealed that $SCoT3_{600}$, $SCoT4_{450}$ and $SCoT12_{550}$ sequences have high similarity coverage of expressed genes/proteins in grapes, except for $SCoT6_{200}$ sequence that has low similarity coverage (Fig.3).

Materials and methods

Plant material and DNA extraction

Leaves of fourteen Egyptian grapes belongs to seven varieties (two replicates for each variety) with different morphological traits (Early Superior, Thompson, Flame, Red Globe, SO4, Ramsey and Dog Ridge) were kindly provided by the Horticultural Research Institute - ARC. Early Superior, Thompson, Flame and Red Globe are table grapes, which are used for fruit productions. Early Superior and Thompson produces fruits with red color, while Flame and Red Globe produces fruits with red colors. Both Thompson and Flame varieties used in this study are seedless varieties. SO4, Ramsey and Dog Ridge varieties are used as a rootstock in grape grafting due to their resistance to nematode and phylloxera insect (Gaprial et al. 2004).

The total DNA was isolated from collected varieties using DNAeasy Plant Mini Kit (Qiagen, Santa Clarita, CA) (Cat no. 69104). Agarose gel electrophoresis confirmed that the DNA was of high molecular weight with no degradation.

The PCR procedure for SCoT analysis

A twenty four SCoT primers employed in the present study (Table S1) were designed by Collard and Mackill (2009)

Table 3. PCR statics revealed using SCOT, SSR and total primers through *in silico* PCR analysis against *V.vinifera* genome.

Technique	NP	TIA	TCGA	GCP%	TIAP	IAG	TG
SSR	3	3	846	0.0002%	1.0	1	1
SCoT	8	226	96077	0.0225%	28.3	51	21
Primer Set	11	229	106452	0.0250%	20.8	52	22

Number of primers (NP), total number of the *in silico amplimers*(TIA), total covered genomic area (TCGA) , total covered genome area percentage compared to all genome sequence length (GCP%) , the average of TIA per primer (TIAP),number of *in silico* amplimers that are close or adjoin genes (IAG) and the number of actual genes that are targeted (TG)
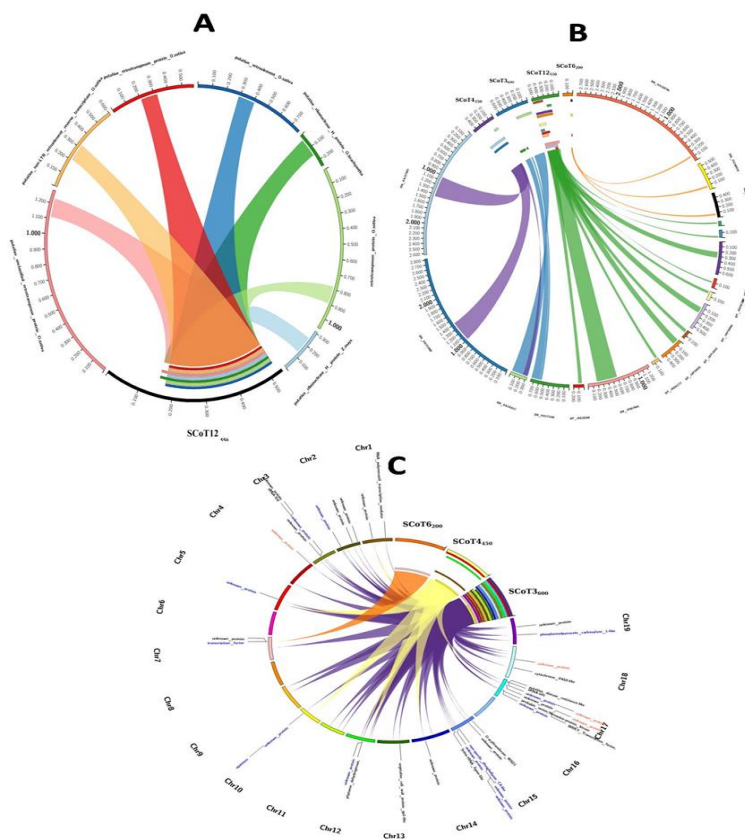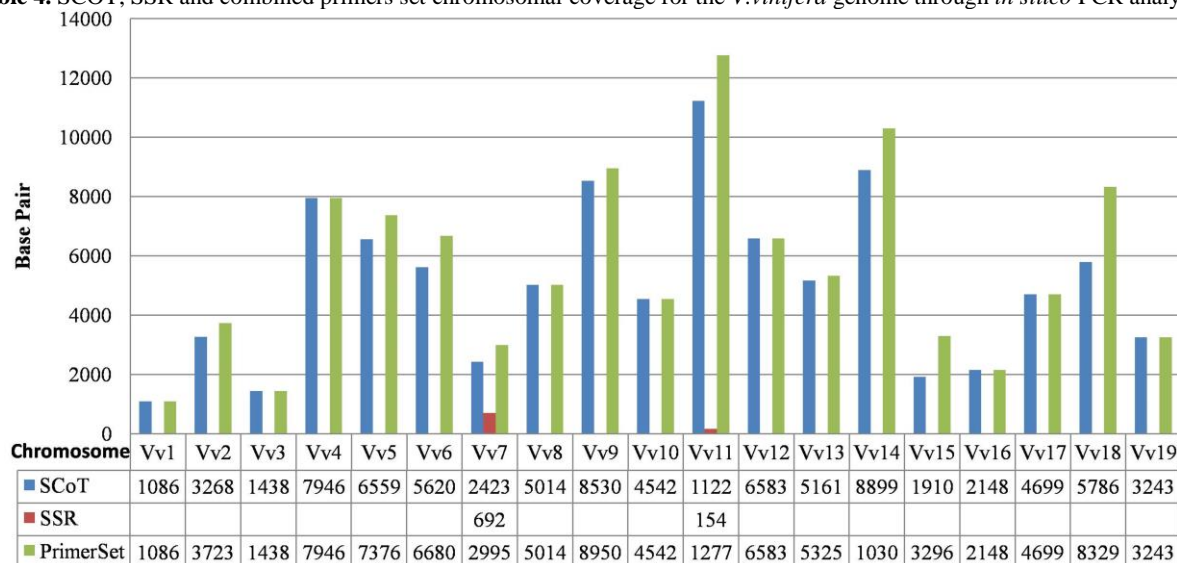

Fig 3. (A) SCoT12$_{550}$ BLAST results against the NCBI non-redundant database showing coverage comparison with different genes belong to different plant species. (B) SCoT3$_{600}$, SCoT4$_{450}$, SCoT6$_{200}$ and SCoT12$_{550}$ BLAST results against the NCBI TSA database showing coverage comparison with different transcriptome expressed sequences belong to *Vitis vinifera*. (C) SCoT3$_{600}$, SCoT4$_{450}$ and SCoT6$_{200}$ BLAST results against the *Vitis vinifera* genome showing coverage comparison and the closest genes to these sequences in 1kbp (red), 5kbp (blue) and 12kbp (black) window.

Table 4. SCOT, SSR and combined primers set chromosomal coverage for the *V.vinifera* genome through *in silico* PCR analysis.

Chromosome	Vv1	Vv2	Vv3	Vv4	Vv5	Vv6	Vv7	Vv8	Vv9	Vv10	Vv11	Vv12	Vv13	Vv14	Vv15	Vv16	Vv17	Vv18	Vv19
SCoT	1086	3268	1438	7946	6559	5620	2423	5014	8530	4542	1122	6583	5161	8899	1910	2148	4699	5786	3243
SSR							692				154								
PrimerSet	1086	3723	1438	7946	7376	6680	2995	5014	8950	4542	1277	6583	5325	1030	3296	2148	4699	8329	3243

based on the consensus sequences of translation initiation codon region in higher plants with ATG codon at positions +1, +2, +3 ; 'G' at position + 4; and 'A,' 'C,' and 'C' at positions + 7, + 8, and + 9, respectively. The PCR reaction was performed as mention by (Luo et al., 2010) in a total volume of 25 μl containing 1X reaction buffer, 1.5 mMMgCl2, 0.2 mM dNTP, 0.4 μM of a single primer; 50 ng genomic DNA and 2U of Taq DNA polymerase. The PCR amplification conditions were carried out as follows: an initial denaturation step at 94 ^{O}C for 3 min. followed by 36 cycles of 94 ^{O}C for 50 s. 50 ^{O}C for 1 min. and 72 ^{O}C for 2 min; the final extension at 72 ^{O}C was held for 5 min. A total of 5 μl PCR products were separated on 1.5% agarose gel in 1 X TBE buffer using electrophoresis technique. The agarose gel was visualized UV light after staining with ethidium bromide.

The PCR procedure for SSR analysis

Seven SSR primers were used in this study (Table S2). The PCR amplification was performed in a 25 μl reaction volume containing about 3μl (10 ng/ μl) genomic DNA, 2 μl primer / each (F and R), 1.5 mMMgCl2, 0.2 mM of each dNTP, 0.4 μM of a single primer; 50 ng genomic DNA and 2U of Taq DNA polymerase . The thermal cycler was programmed with an initial step of 5 min at 94° C; the amplification reaction was carried out using 40 cycles of 40 s at94° C, an annealing step of 1 min at specific annealing temperature for each primer and an elongation step of 1min at 72°C; and finally a 7 min extension at 72° C.

The PCR scoring Data analysis

PCR fragments on gels were recorded as presence (1) or absence (0) for all samples and final data sets included both polymorphic and monomorphic bands. Only patterns that are existed in both sample replicate were chosen for analysis. Then a binary statistic matrix was constructed. Dice's similarity matrix coefficients were calculated between accessions using the unweighted pair group method with arithmetic averages (UPGMA) and this matrix was used to construct a phylogentic dendrogram using the online dendrogram construction utility DendroUPGMA (http://genomes.urv.es/UPGMA/) (Garcia-Vallvé et al., 1999). The PIC was calculated using PowerMarker software (Liu and Muse 2005).

Cloning and sequencing potential PCR bands

Eight SCoT PCR bands which show variations among grape varieties were separated by electrophoresis in agarose gel and recovered by mechanical excision of small gel plugs. DNA was eluted and purified. DNA fragments were inserted into an appropriate plasmid vector according to the manufacturer's instructions (pGEM-T vector, Promega Corporation, Cat. No. A1360). DNA sequencing was carried out by Sanger sequencing technique (Sanger et al., 1977) using M13 primers.

Bioinformatics analysis procedures

To estimate the quality of the generated sequence traces, the original forward and reverse sequences were assembled using CodonCode Aligner 3.0 (CodonCode Co., USA). The NCBI online BLAST tools were used with its default parameters to align the generated sequences using BLAST algorithm (Altschul et al., 1994) against the NCBI database. PERL

scripts were used to visualize BLAST results using Circos software package (Circos 0.66) (Krzywinski et al., 2009). The whole genome sequence analysis (PERL scripts) was used to identify genes in a window of 1, 5 and 12 kb in both upstream and downstream through the published grape genome (Jaillon et al., 2007) (Fig. 3).

In silico PCR analysis parameters

The full genome of *Vitis vinifera* (~487 Mbp, n=19) (Jaillon et al., 2007) were used as template for the *in silico* PCR analysis against all studied PCR primers to reveal the possible PCR amplimers. PERL computer language scripts were used for the *in silico* PCR analysis according to the following criteria: the maximum length of produced amplimer ≤ 1500 bp and the minimum length ≥ 50.The maximum acceptable sequence mismatch between the primer and the DNA template for SCoT ≤ 4 bp and SSR ≤ 1bp. Circos software package (Circos 0.66) was used for visualizing output results in a circular layouts (Krzywinski et al., 2009).

Conclusion

Our results demonstrated that high genetic variability exists between the Egyptian grape varieties under study. Besides, the potential use for SCoT and SSR techniques to generate high number of polymorphic markers is assigned. Plausibly, *virtual* and actual amplified PCR fragments proved that SCoT technique was successful to target generic regions across grape genome. More over; our result provided sufficient primed marker design. The SCoT3$_{600}$, SCoT4$_{450}$, SCoT6$_{200}$ and SCoT12$_{550}$ sequences gave good similarities with genes of high potential functions, and suggesting future studies are needed to decipher its structure. Our study is putting the basis for cultivar selection and association mapping of Egyptian grape varieties.

Acknowledgments

References

Adawy SS, Jiang J, Atia MAM (2014) Identification of novel sex-specific PCR-based markers to distinguish the genders in Egyptian date palm trees. Int J Agric Sci Res. 4:45–54.

Alleweldt G, Possingham JV (1988). Progress in grapevine breeding. Theor Appl Genet. 75:669–673.

Altschul SF, Boguski MS, Gish W, Wootten JC (1994) Issues in searching molecular sequence databases. Nat Genet. 6:119–129.

Collard BCY, Mackill DJ (2009) Start codon targeted (SCoT) polymorphism: a simple, novel DNA marker technique for generating gene-targeted markers in plants. Plant Mol Biol Report. 27:86–93.

Bittner-Eddy PD, Crute IR, Holub EB, Beynon JL (2000) RPP13 is a simple locus in *Arabidopsis thaliana* for alleles that specify downy mildew resistance to different

avirulence determinants in *Peronospora parasitica*. Plant J. 21:177–188.

Cadle-Davidson MM, Owens CL (2008) Genomic amplification of the Gret1 retroelement in white-fruited accessions of wild *Vitis* and interspecific hybrids. Theor Appl Genet. 116:1079–1094.

FAOSTAT, F FAO ''Statistical Database'' Retrieved March 25, 2016. http://faostat.fao.org. Rome.

Feng S, He R, Yang S, Chen Z, Jiang M, Lu J, Wang H (2015) Start codon targeted (SCoT) and target region amplification polymorphism (TRAP) for evaluating the genetic relationship of *Dendrobium* species. Gene. 567 (2): 182-188.

Gaprial FG , Razek EA, Hassan FF (2004) cultivation and production of grape. Central administration for agricultural extended services. Technical bulletin. No. 1.

Garcia-Vallvé S, Palau J, Romeu A (1999) Horizontal gene transfer in glycosyl hydrolases inferred from codon usage in *Escherichia coli* and *Bacillus subtilis*. Mol Biol Evol. 16:1125–1134.

Gorji AM, Poczai P, Polgar Z, Taller J (2011). Efficiency of arbitrarily amplified dominant markers (SCoT, ISSR and RAPD) for diagnostic fingerprinting in tetraploid potato. Am J potato Res. 88:226–237.

Guo DL, Zhang JY, Liu CH (2012) Genetic diversity in some grape varieties revealed by SCoT analyses. Mol Biol Rep. 39:5307–5313.

Jaillon O, Aury JM, Noel B, Policriti A, Clepet C, Casagrande A, Choisne N, Aubourg S, Vitulo N, Jubin C , Vezzi A (2007) The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature. 449:463–467.

Krzywinski M,Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ , Marra MA(2009) Circos: an information aesthetic for comparative genomics. Genome Res. 19:1639–1645.

Ku MS, Agarie S, Nomura M, Fukayama H, Tsuchida H, Ono K, Hirose S, Toki S, Miyao M , Matsuoka M (1999) High-level expression of maize phosphoenolpyruvate carboxylase in transgenic rice plants. Nat Biotechnol. 17:76–80.

Laucou V, Lacombe T, Dechesne F, Siret R, Bruno JP, Dessup M, Dessup T, Ortigosa P, Parra P, Roux C, Santoni S (2011). High throughput analysis of grape genetic diversity as a tool for germplasm collection management. Theor Appl Genet. 122:1233–1245.

Liu C, Fan X, Jiang J, Guo D, Sun H, Zhang Y, Feng J (2012) Genetic diversity of Chinese wild grape species by SSR and SRAP markers. Biotechnol Biotechnol Equip. 26:2899–2903.

Liu K , Muse SV (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics. 21:2128–2129.

Luo C, He XH, Chen H, Ou SJ, Gao MP (2010) Analysis of diversity and relationships among mango cultivars using Start Codon Targeted (SCoT). markers. Biochem Syst Ecol. 38:1176–1184.

Martinez LE, Cavagnaro PF, Masuelli RW, Zuniga M (2006) SSR-based assessment of genetic diversity in South American *Vitis vinifera* varieties. Plant Sci. 170:1036–1044.

Mokhtar MM, Adawy SS, El-Assal SE-DS, Hussein EHA (2016) Genic , Intergenic SSR Database Generation, SNPs Determination and Pathway Annotations, in Date Palm (Phoenix dactylifera L.). PLoS one .11(7):e0159268.

Moreno-Sanz P, Loureiro MD, Suárez B (2011) Microsatellite characterization of grapevine (*Vitis vinifera*

L.) genetic diversity in Asturias (Northern Spain). Sci Hortic (Amsterdam). 129:433–440.

Pejchar P, Potocký M, Krčková Z, Brouzdová J, Daněk M, Martinec J(2015). Non-specific phospholipase C4 mediates response to aluminum toxicity in *Arabidopsis thaliana*. Front Plant Sci. 6:66.

Pelsy F, Hocquigny S, Moncada X, Barbeau G, Forget D, Hinrichsen P, Merdinoglu D (2010) An extensive study of the genetic diversity within seven French wine grape variety collections. Theor Appl Genet. 120:1219–1231.

Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S, Rafalski A (1996) The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Mol Breed. 2:225–238.

Ray A, Robinson-Beers K, Ray S, Baker SC, Lang JD, Preuss D, Milligan SB, Gasser CS (1994) Arabidopsis floral homeotic gene BELL (BEL1) controls ovule development through negative regulation of AGAMOUS gene (AG). Proc Natl Acad Sci. 91:5761–5765.

Rushton PJ, Somssich IE, Ringler P, Shen QJ (2010) WRKY transcription factors. Trends Plant Sci. 15:247–258.

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci. 74:5463–5467.

Schmidt A, Sandman G (1990) Cloning and nucleotide sequence of the crtI gene encoding phytoene dehydrogenase from the cyanobacterium Aphanocapsa PCC6714. Gene 91:113–117.

Li F., Wu X, Lam P, Bird D, Zheng H, Samuels L, Jetter R , Kunst L (2008) Identification of the wax ester synthase/acyl-coenzyme A: diacylglycerol acyltransferase WSD1 required for stem wax ester biosynthesis in Arabidopsis. Plant Physiol. 148(1):97-107.

This P, Lacombe T, Thomas MR (2006) Historical origins and genetic diversity of wine grapes. Trends Genet 22:511–519.

Tillett RL, Ergül A, Albion RL, Schlauch KA, Cramer GR, Cushman JC (2011) Identification of tissue-specific, abiotic stress-responsive gene expression patterns in wine grape (*Vitis vinifera* L.) based on curation and mining of large-scale EST data sets. BMC Plant Biol. 11:1.

Tsukagoshi H, Busch W, Benfey PN (2010) Transcriptional regulation of ROS controls transition from proliferation to differentiation in the root. Cell. 143:606–616.

Uhrig RG, Labandera AM, Moorhead GB (2013). Arabidopsis PPP family of serine/threonine protein phosphatases: many targets but few engines. Trends Plant. Sci 18:505–513.

Xiong F, Zhong R, Han Z, Jiang J, He L, Zhuang W, Tang R (2011) Start codon targeted polymorphism for evaluation of functional genetic variation and relationships in cultivated peanut (*Arachis hypogaea* L.) genotypes. Mol Biol Rep 38:3487–3494.

Zeng B, Zhang Y, Huang LK, Jiang XM, Luo D, Yin G (2014) Genetic diversity of orchardgrass (*Dactylis glomerata* L.) germplasms with resistance to rust diseases revealed by Start Codon Targeted (SCoT) markers. Biochem Syst Ecol. 54:96–102.

Zoghlami N, Riahi L, Laucou V, Lacombe T, Mliki A, Ghorbel A, This P (2009) Origin and genetic diversity of Tunisian grapes as revealed by microsatellite markers. Sci Hortic (Amsterdam). 120:479–486.

Zulini L, Russo M , Peterlunger E (2015) Genotyping wine and table grape cultivars from Apulia (Southern Italy) using microsatellite markers. VITIS-Journal Grapevine Res 41:183.