SSR mining of black cumin (Nigella sativa L) transcriptome for molecular markers development

Ibrahim Celik
Pamukkale University - Kınıklı Campus: Pamukkale Universitesi
Adnan AYDIN (aydinadnann@gmail.com)
Igdir University: Igdir Universitesi https://orcid.org/0000-0002-8284-3751

Research Article

Keywords: Black cumin, Unigenes, functional annotations, microsatellites, genic SSR markers

Posted Date: May 10th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1495064/v1

Abstract

Black cumin (*Nigella sativa* L.) is an important aromatic plant due to the production of many health-beneficial metabolites. Up to date- molecular characterizations of the germplasm were performed by using general marker systems. In the present study, the transcriptome of the black cumin was searched for SSRs. Dinucleotide and trinucleotide repeats were the major repeat types which comprised 52.4% and 44.7% of total SSRs, respectively. TC/GA was the most frequent repeat motif (11.8%). A total of 4,779 SSR markers were developed. Functional annotation of 2,456 SSR markers was performed. For validation of the SSR markers, randomly selected 20 markers were tested in a subset population containing nine *Nigella sativa* and one *Nigella damascena* genotypes. A total of 13 SSR makers produced clear amplification and generated 41 polymorphic fragments with mean gene diversity (0.25). Principal Coordinate Analyses (PCoA) and MCMC methods based on hierarchical clustering analysis revealed two clusters and moderate diversity with a mean value of 0.66 and 0.57 for interspecific and intraspecific populations, respectively. This is the first report of the development of SSR markers from black cumin transcriptome. These high-throughput crop-specific SSR markers are valuable genomic sources for comprehensive genome analysis in black cumin.

1. Introduction

Black cumin (*Nigella sativa* L) belongs to the Ranunculaceae family is one of the important aromatic plants due to health-promoting effect of its extracts on many diseases such as dermatological complications, cancers, and type 2 diabetes. Thus, the crop is a major candidate for the development of phytotherapy accents in complementary and alternative medicine (Yimer et al. 2019, Islam et al. 2019). In addition, black cumin seeds are edible and used in the bakery and cheese industry as a spice for flavor (Bourgou et al. 2010). The crop is mainly cultivated in Southern Europe, the Middle East, Northern, and East Africa.

Despite the importance of the crop in the food and pharmaceutical industry, crop-specific molecular tools are quite limited. Thus, molecular characterization of the black cumin germplasm was performed by using general marker systems such as RAPD (Random Amplified Polymorphic DNA), ISSR (Inter-Simple Sequence Repeat), SRAP (Sequence-Related Amplified Polymorphism), and SCoT (Start Codon Targeted Polymorphism) markers (Al-Huqail and Al-Saad 2010; Iqbal et al. 2011; Poyraz 2014; Birhanu et al. 2015; Neghab and Panahi 2017; Mirzaei and Mirzaghaderi 2017; KorehKhosravi et al. 2018; Golkar and Nourbakhsh 2019; Bourgou et al. 2020). Up to date, none of the black cumin-specific molecular markers are available for comprehensive genome analysis such as mapping or diversity analysis. SSRs (Simple Sequence Repeats) are highly polymorphic motif repeats (1 to 6 nucleotides) distributed in the genome due to unequal crossover and replication slippage. Thus, SSRs are used as a reliable marker system in plant genomics due to genome-wide distribution, high rate of polymorphism and locus specificity (Ellegren 2004; Morgante et al. 2002; Senan et al. 2014) SSR markers can be developed from transcriptomic or genomic sequences. Transcriptome sequencing using next-generation sequencing technologies (NGS) provides a great opportunity to develop genic SSR markers. The transcriptome of the black cumin was sequenced by using the Illumina NGS platform. Although the transcriptome was assembled and available in the PhytoMetaSyn database (https://bioinformatics.tugraz.at/phytometasyn/), it was not mined for SSR (Xiao et al. 2013). The present study intended to develop and validate transcriptome-based SSR markers for genome analysis in black cumin.

2. Materials And Methods

2.1. Plant material

A total of nine *Nigella sativa* genotypes (eight landraces and one cultivar) and one *Nigella damascena* genotype were used to test SSR markers developed in the present study. A total of eight *Nigella sativa* landraces were collected from

seven cities (Amasya, Burdur, Eskisehir, Konya, Tokat, Şanlıurfa, and Samsun) in Turkey, one landrace was collected from Syria. *Nigella sativa* cv. Çameli and *Nigella damascena* were purchased commercially.

The seeds of the genotypes were sown in pots filled with peat pellets and grown in the climate-controlled growth cabinet with the following settings: 50% humidity and 25 and 30 ^{o}C for 16h:8h light/dark cycle, respectively. Genomic DNA was isolated from leaf tissue bulked from five plants per genotype using the protocol described by Aydin et al. (2018). The quality and quantity of the extracted genomic DNA were determined with a nano-drop spectrophotometer (Maestrogen, MN-913) and 1% agarose gel.

2.2. SSR screening of black cumin transcriptome

Transcriptome assembly data of black cumin generated by Illumina HiSeq next-generation sequencing platform was downloaded from the phytometasyn database (https://bioinformatics.tugraz.at/phytometasyn/). De novo transcriptome assembly details can be found in Xiao et al. (2013). The transcriptome was screened for five SSRs (from dinucleotide to hexanucleotide repeats) by using GMATA software (Wang and Wang 2016).

2.3. Marker development for SSR polymorphism

For amplification of SSRs identified in the present study, SSR markers fulfilled primer design parameters such as 120–400 bp and 60°C for PCR product size and optimum annealing temperature, respectively by using Primer3 software in GMATA pipeline (Wang and Wang 2016).

The present study aimed to validate SSR markers developed in the present study for amplification and polymorphism. To achieve this aim, randomly selected 20 SSR markers were tested in a subset of black cumin germplasm (Table 1). PCR mix contained 1X PCR buffer, 2 mM $MgCl_2$, 0.28 mM dNTPs, 1 U Taq Polymerase, 0.5 µM forward and reverse primers and 50 ng genomic DNA. The PCR protocol was one cycle of 5 min at 94 °C, followed by 35 cycles of 30 s at 94°C, 30 s at 55 °C (annealing) and 45 s at 72 °C, with a final extension step of 10 min at 72 °C. PCR fragments were separated using 4% agarose gel electrophoresis and DNA High-Resolution Kit of QIAxcel Advanced System Capillary electrophoresis (Qiagen). PCR fragments were scored dominantly (1 for presence and 0 for absence). Gene diversity of all tested SSR markers was calculated based on the formula of Roldan-Ruiz et al. (2000) implemented in GDdom software (Abuzayed et al. 2017).

Table 1. SSR primer pairs used for validation.

Primer ID	Sequence	PCR fragment	SSR Motif	Repetitions
nsSSR157-F	TAGTTCAGGGCTGAGCAGGT	208	TG/CA	5
nsSSR157-R	GCACACACTCCCTAGCACTTC			
nsSSR446-F	GACCATGGACACAACCATGA	282	TCACAG/CTGTGA	8
nsSSR446-R	ACCACCAATCATCACCCCTA			
nsSSR551-F	CCGGCATGAGACAAGAAACT	400	TC/GA	5
nsSSR551-R	CTGAAAATGCACGTCACCAC			
nsSSR701-F	TCTTTCCCTGCTCTTTCCAA	336	GAA/TTC	5
nsSSR701-R	TGGTGATGGTGGTGGTAATG			
nsSSR1482-F	TTGGCCCAGAAAGTGTAACC	161	CA/TG	6
nsSSR1482-R	GGGAACAGCCATATCGAAGA			
nsSSR1716-F	TGGTCATGGAGTTCAAGGTG	400	GAT/ATC & AG/CT	6 & 6
nsSSR1716-R	TCGCTCTTCTTCCTCCTTCTC			
nsSSR1895-F	GAGAGAGGGTGAAGGGTTGC	334	GAAG/CTTC	6
nsSSR1895-R	CCCCAAGCACCATAGAGTGT			
nsSSR1912-F	AGACTGTAGCACGGTGGATG	121	AG/CT	7
nsSSR1912-R	GAACCAGACAACCTTTTCTTTC			
nsSSR1941-F	TTGACGCGTACTCATTGACAG	161	GAA/TTC	5
nsSSR1941-R	CCTTCCGCTTCTCATAACCA			
nsSSR2136-F	TCAACACCTCAATCGCTGTC	280	CTT/AAG	5
nsSSR2136-R	TCCACATTGGAGGTCACAAA			
nsSSR2139-F	TCCACAACCACCACTTCAAA	400	CTC/GAG	5
nsSSR2139-R	TGCGAGGAGGAGAAGAAGAA			
nsSSR2221-F	TCAGGAAACAGTTGTTGAAACC	121	TTC/GAA	6
nsSSR2221-R	CAAGGCCTGAAGATGATTGC			
nsSSR2239-F	TCTCTTCTTTTGCCGCCTCT	120	CCA/TGG & TC/GA	5 & 5
nsSSR2239-R	GAGAAGAATTGGGTGGTGGA			
nsSSR2404-F	TCAGATCGTTCCTTTCCTTGA	207	GAA/TTC	6
nsSSR2404-R	AATCTCCGGCCCCTAAATAA			
nsSSR2419-F	TGTGTGGTGCAGAGTGAAGA	121	GAA/TTC	5
nsSSR2419-R	CACCTCCATCACCTCCTTGT			
nsSSR2543-F	TCCCAGTGCAGAAGACAAAA	340	TA/TA & CTG/CAG	5 & 7
nsSSR2543-R	CAGCAGCAGTAGCAGCAAAC			

nsSSR2567-F	AGAAGGGGTTGTGGTTGGTT	120	AG/CT	5
nsSSR2567-R	TTTCCTCCCTTCCACTACTCG			
nsSSR3047-F	CATCTGGGTTGTGGGTTTCT	161	CCT/AGG	5
nsSSR3047-R	AACCAGTAGGAGCTGCAGGA			
nsSSR3342-F	CTAAGCTGCCCTCCTTCCTT	279	TC/GA	5
nsSSR3342-R	GGAGAATCGACGGATCAGAG			
nsSSR3533-F	AGGGGCGAATAGTCCTTTCT	206	CT/AG	5
nsSSR3533-R	TTCGGTCCATATACAGAATTGC			

2.4. Diversity analysis

Firstly, Principal Coordinate Analyses (PCoA) was performedusing the Multi Variety Statistical Package (MVSP 3.13O, Kovach Computing Services, Pentraeth, UK) software based on the *Jaccard* similarity coefficient of marker data.

Hierarchical clustering analysis was performed using MCMC methods in MrBayes V3.2.1 x64 software program. One cold and 3 heated chains were run starting from a random tree for 10 million Markov chain Monte Carlo (MCMC) generations, with chains sampled every 100th cycle (Karaca et al, 2015). A consensus dendrogram with a 50% majority rule was constructed with the posterior probabilities indicated at the nodes (Karaca et al., 2013).

2.5. Functional annotation of the transcriptome

Sequences containing SSRs were converted in FASTA format and functional annotation was performed for revealing molecular functions of SSRs using the Go Feat functional annotation tool (Araujo et al. 2018).

3. Results

3.1. SSR screening of black cumin transcriptome

Black cumin transcriptome containing 67,591 unigenes 51,740,837 bp in size were screened for SSRs. As result, 8,038 SSRs were identified in 6,709 unigenes (9.93% of total unigenes). The size of the unigenes ranged from 200 nt to 14,144 nt with a mean value of 2,578.5 ± 1,289.1 nt (mean ± SE). SSR length ranged from 10 nt to 96 nt with a mean value of 14.66 ± 0.03 nt. (Online Resource 1, Table S1). Most of the Unigenes contained single SSRs (86.3% of total unigenes containing SSRs). Dinucleotide and trinucleotide repeats were the major repeat types which comprised 52.4% and 44.7% of total SSRs, respectively. Rest of the repeat types (tetranucleotide, pentanucleotide, and hexanucleotide repeats) had low frequency which comprised 2.9% of the total SSRs (Table 2). TC/GA was the most frequent repeat motif which comprised 11.8% of the total motifs. The rest of the SSR motifs had a frequency of less than 10% (Table 3). TCT/AGA was the most abundant trinucleotide SSR motif (4.7% of total SSRs).

Table 2. SSR types identified in the black cumin transcriptome.

Motif length	Number of SSRs	Frequency (%)
Dinucleotide	4,215	52.4
Trinucleotide	3,590	44.7
Tetranucleotide	174	2.2
Pentanucleotide	27	0.3
Hexanucleotide	32	0.4
Total	8,038	

Table 3. The most frequent SSR motifs

SSR motif	Number of SSRs	% of SSR motif
TC/GA	946	11.8
CT/AG	736	9.2
AG/CT	720	9.0
GA/TC	637	7.9
TCT/AGA	380	4.7
TTC/GAA	349	4.3
CTT/GAA	295	3.7
GAA/TTC	289	3.6
AGA/TCT	243	3.0
AAG/CTT	239	3.0

3.2. Marker development for SSR polymorphism

For amplification, a total of 4,779 primer pairs targeting 5,297 SSRs (65.9% of total SSRs) were developed. Detailed primer data are given in Online Resource 2. Randomly selected 20 SSR primers were tested in a subset of the *Nigella sativa* population and a *Nigella damascena* genotype as an out-group. A total of 13 SSR markers (65%) had clear amplification and one of them (nsSSR2136) was polymorphic in agarose gel electrophoresis (Figure 1).

Thus, non-polymorphic SSR markers were separated by using capillary electrophoresis for better resolution. A total of nine SSR markers (69.2%) were transferable to *Nigella damascena*. PCR fragment number arranged from 1 to 6 with a mean value of 3.31 ± 1.26. All SSR markers except for nsSSR17 were polymorphic in all genotypes. All markers produced 41 polymorphic PCR fragments (97.7% of total PCR fragments). Interspecific Gene diversity of markers ranged from 0.18 to 0.37 with a mean value of 0.25. nsSSR1912 had the highest gene diversity (Table 4). All markers produced 30 polymorphic PCR fragments (90.9% of the total fragments) in *Nigella sativa* genotypes. The average fragment number ranged from 1 to 5 with a mean value of 2.54 ± 1.28. All SSR markers except for nsSSR2221 and nsSSR2567

were polymorphic. Intraspecific gene diversity of markers ranged from 0.10 to 0.31 with a mean value of 0.25. nsSSR1912, nsSSR1941, and nsSSR3047 had the highest gene diversity (Table 4).

Table 4. Gene diversity values of Nigella sativa specific SSR markers.

			Nigella sativa genotypes & Nigella domesticane		Nigella sativa genotypes	
SSR marker	SSR motif	Repetitions	No. of polymorphic fragments//total no. fragments (%)	Gene diversity ± SE	No. of polymorphic fragments//total no. fragments (%)	Gene diversity ± SE
nsSSR446	TCACAG/CTGTGA	8	6/6 (100)	0.28 ± 0.03	5/5 (100)	0.28 ± 0.06
nsSSR701	GAA/TTC	5	5/5 (100)	0.28 ± 0.05	5/5 (100)	0.28 ± 0.05
nsSSR1716	GAT/ATC & AG/CT	6 & 6	4/4 (100)	0.32 ± 0.07	3/3 (100)	0.28 ± 0.01
nsSSR1912	AG/CT	7	3/3 (100)	0.37 ±0.08	2/2 (100)	0.31 ± 0.13
nsSSR1941	GAA/TTC	5	4/4 (100)	0.31 ± 0.06	4/4 (100)	0.31 ± 0.06
nsSSR2136	CTT/AAG	5	3/3 (100)	0.23 ± 0.04	1/2 (50)	0.11 ± 0.09
nsSSR2139	CTC/GAG	5	3/3 (100)	0.30 ± 0.06	2/2 (100)	0.23 ± 0.09
nsSSR2221	TTC/GAA	6	2/2 (100)	0.18	0/1 (0)	0
nsSSR2239	CCA/TGG & TC/GA	5 & 5	2/2 (100)	0.2	2/2 (100)	0.2
nsSSR2419	GAA/TTC	5	3/3 (100)	0.27 ± 0.04	2/2 (100)	0.13 ± 0.05
nsSSR2543	TA/TA & CTG/CAG	5 & 7	4/4 (100)	0.22 ± 0.03	2/2 (100)	0.10 ± 0.05
nsSSR2567	AG/CT	5	0/1 (0)	0	0/1 (0)	0
nsSSR3047	CCT/AGG	5	3/3 (100)	0.36 ± 0.07	2/2 (100)	0.31 ± 0.13

3.3. Diversity analysis

Performance of newly developed SSR markers in terms of diversity was tested using Principal Coordinate Analyses (PCoA). As result, the PCoA plot contained two clusters (cluster A and B). Cluster A (green colored) contained two landraces collected from Amasya and Eskişehir and one cultivar (Çameli). Cluster B (red-colored) contained four landraces collected from Şanlıurfa, Burdur, Tokat, Samsun. Although the landrace collected from Konya is located close to Turkish landraces, it was not clustered. While landrace from Syria was located in the upper right quadrant of the PCA plot, *Nigella damascena was* located in the lower right quadrant of the PCA plot apart from the rest of the material

(Figure 2). Average interspecific diversity (including *Nigella damascena)* ranged from 0.29 (71% similarity) to 0.99 (1% similarity) with a mean value of 0.66. Intraspecific diversity (excluding *Nigella damascena)* ranged from 0.29 (between landraces collected from Tokat and Burdur) to 0.82 (between landraces collected from Konya and Syria) with a mean value of 0.57.

MCMC methods-based hierarchical clustering analysis was performed using distance matrix calculated from binary data of the SSR markers developed in the present study. As expected, *Nigella damascena* was not clustered with *Nigella sativa* genotypes. The dendrogram contained two clusters (blue and green colored clusters). While blue colored cluster contained landraces from Burdur and Tokat, red-colored clusters contained six landraces collected from Syria, Şanlıurfa, Samsun, Konya Eskişehir and Amasya and one cultivar (Çameli) (Figure 3). Landraces collected from Syria and Konya had the highest branch length.

3.4. Functional annotation of the transcriptome

Functional annotations of unigenes containing SSR markers were performed. As result, SSR markers had 2,972 hits for 539 molecular functions. SSR markers had ATP, metal ion, DNA, nucleic acid, RNA, and zinc ion binding functions had the highest frequency. A total of 292, 160, 154, 145, 134, and 105 SSR markers had these functions, respectively. For biological processes, SSR markers had 1,432 hits in 530 processes. Regulation of transcription and translation was the most prevalent biological processes with 83 and 52 SSR marker hits, respectively. Also, cellular components of SSR markers were predicated. As result, SSR markers had 2002 hits for 217 cellular components. Integral membrane and nucleus have the most frequent cellular components with 1,063 and 243 hits. Overall, 2,141 markers had hits for uncharacterized proteins and 6 for unplaced genic scaffolds. A total of 167 SSR markers did not have any hits, while 2,456 SSR (51% of total SSRs) markers were annotated. Detailed annotation results of SSR markers are given in Online Resource 3.

4. Discussion

4.1. SSR screening of black cumin transcriptome

The transcriptome of black cumin was mined for SSRs for marker development. Although less than 10% of the unigenes contained SSRs, a sufficient number of SSRs were identified. The frequency of SSR markers was similar to the frequency of SSRs identified in *Coriandrum sativum* (10%) and a range from 2% to 17% reported in dicotyledonous species (Tulsani et al. 2020). The frequency of dinucleotide repeats was slightly higher than trinucleotide repeats. Although trinucleotide repeats had a higher frequency than dinucleotide repeats in transcriptomes due to positive selection in *Oryza sativa*L. *Arabidopsis thaliana*, *Glycine max* L., *Zea mays* L., *Triticum aestivum* L. and *Arachis hypogaea*, recent studies performed in *Coriandrum sativum* L. (45.5% of Dinucleotide and 34.6% of trinucleotide repeats), *Camellia sinensis* (52.54% of dinucleotide and 22.64% of trinucleotide repeats), *Sesamum indicum* L. (67.07% of dinucleotide and 24.89% of trinucleotide repeats), *Cajanus cajan* L. Millspaugh (60.41% of dinucleotide and 34.52% of trinucleotide repeats) reported that dinucleotide repeats were the most commonly similar to present study (Morgante, et al. 2002; Wei et al. 2011; Dutta at al. 2011, Zhang et al. 2012; Tan et al. 2013; Tulsani et al. 2020). The predominance of TC/GA SSR motif in black cumin was conflicted to previous studies due to highest prevalence of AG/CT in transcribed regions of *Oryza sativa*, *Arabidopsis thaliana*, *Glycine max*, *Zea mays*, *Triticum aestivum* genomes (Morgante et al. 2002; Wei et al. 2011). The reason for black cumin-specific high frequency of TC/GA repeats is unclear due to complex evolution and selection mechanisms of SSR motifs (Sonah et al. 2011).

4.2. Marker development for SSR polymorphism

SSR markers were validated in a subset of the *Nigella sativa* population. Seven SSR markers (35%) of the tested primers did not produce PCR fragments, similar frequency of nonamplified primers were also reported in genic SSR in *Cajanus cajan* (28.8%), *Sesamum indicum* (20%), and *Camellia sinensis* (12.2%) (Dutta at al. 2011; Wei et al. 2011; Tan at al. 2013). The reason for this limitation might be due to large introns in genes. (Wei et al. 2011). Non-transferability of some of the SSR markers to *Nigella damascena* was expected due to species specific genes in *Nigella sativa*. The average fragment number of SSR markers (3.31 ± 1.26 and 2.54 ± 1.28 for interspecific and intraspecific populations, respectively) developed in the present study was lower than *Coriandrum sativum* (4.47), *Sesamum indicum* (6.55), and *Cajanus cajan* (6.25) specific SSR markers (Tulsani et al. 2020; Dutta et al. 2011; Wei et al. 2011). This difference might be due to different genome organizations of plants such as duplications. Also, scoring of alleles of SSR primers produced a high number of fragments was complicated. Although the number of targets of the primers was not clear in the black cumin due to the absence of reference genome, a lower number of fragments showed that current SSR markers targeted single. Validation of the markers showed that SSR markers have moderate gene diversity (0.25). Gene diversity value was similar to intraspecific gene diversity of opium poppy specific genic SSR (0.27) and lower than interspecific gene diversity of *Origanum* genus species (0.36). The differences of gene diversity of SSR markers due to different population types (interspecific or intraspecific populations) (Taşcıoğlu, et al. 2018; Şelale et al. 2013).
Mean gene diversity of SSR markers was lower than ISSR (0.35) and RAPD (0.42) markers tested in Iranian populations of *Nigella sativa* contained 18 genotypes(KorehKhosravi et al. 2018*)* Although this showed that SSR markers had lower discrimination power than ISSR and RAPD markers, SSR markers were reliable marker system due to higher reproducibility than ISSR and RAPD markers.

4.3. Diversity analysis

Principal Coordinate Analyses (PCoA) and MCMC methods based on hierarchical clustering analysis were used to test the performance of SSR markers developed in the present study. Although a small set of *Nigella sativa* populations was used, clusters were determined successfully based on tested SSR markers. Although PCoA and MCMC plots had different topology, they had similar results based on diversity such as branch lengths. The average diversity of the population analyzed in the present study was higher than the average diversity of the core set population containing 30 *Nigella sativa* genotypes analyzed by sequence-related amplified polymorphisms (SRAP) and start codon targeted (SCoT) markers (0.31) (Golkar and Nourbakhsh. 2019) and lower than 39 Iranian *Nigella sativa* genotypes analyzed by ScoT markers (Mirzaei and Mirzaghaderi 2017). This difference might be due to the different populations analyzed in the studies. PCoA and hierarchical clustering analysis together have clearly demonstrated both interspecific and intraspecific diversity revealed by SSR markers.

4.4. Functional annotation of the transcriptome

Functional annotation of SSR markers was determined based on homology analysis. The proportion of annotated unigenes (51%) in the present study was lower than the average proportion of 20 nonmodel plants transcriptome (88.15%) reported by Xiao et al. (2013). This difference might be due to the different methods used in the two studies and the shorter SSR sequence length used in the annotation in the present study. In the present study, SSR markers had the function in biosynthetic processes were identified which is useful for targeted analysis of loci control metabolite synthesis due to natural source of the crops for various health-related metabolites (Yimer et al. 2019; Islam et al. 2019).

5. Conclusion

In the present study, high-throughput black cumin-specific genic SSR markers were developed. The number of SSR markers was sufficient for comprehensive genome analyses such as mapping and diversity. Validation of SSR markers showed that these markers can be used in diversity analysis due to sufficient gene diversity value and accurate

dendrogram. Also, the function of the SSR markers was predicted based on homology analysis. The present study is the first report of SSR markers in *Nigella sativa* and the markers developed in the present study can be used in black cumin genome analysis.

Declarations

Acknowledgments

We would like to thank Dr. Muhittin Kulak (Igdir University, Turkey) for providing plant seed materials.

Funding:

Not applicable

Conflicts of interest/Competing interests:

The authors declare that they have no conflict of interest.

Availability of data and material:

Not applicable

Code availability:

Not applicable

Ethics approval:

This research did not involve animal or human participants.

Consent to participate:

Not applicable

Consent for publication:

Not applicable

Authors' contributions:

IC designed SSR markers, analyzed data and drafted manuscript. AA performed laboratory experiments, analyzed data and wrote the manuscript.

References

1. Abuzayed M, El-Dabba N, Frary A, Doganlar S (2017) GDdom: an online tool for calculation of dominant marker gene diversity. Biochem Genet 55(2):155–157
2. Al-Huqail A, Al-Saad F (2010) DNA fingerprinting and genotyping of four black seed (*Nigella sativa* L.) Taxa. J King Abdulaziz University: Metrol Environ Arid Land Agricultural Sci 142(574):1–31
3. Araujo FA, Barh D, Silva A et al (2018) GO FEAT: a rapid web-based functional annotation tool for genomic and transcriptomic data. Sci Rep-UK. https://doi:10.1038/s41598-018-20211-9

4. Aydin A, Ince AG, Uygur Gocer E, Karaca M (2018) Single cotton seed DNA extraction without the use of enzymes and liquid nitrogen. Fresenius Environ Bull 27(10):6722–6726)
5. Birhanu K, Tileye F, Yohannes P, Said M (2015) Molecular diversity study of black cumin (*Nigella sativa* L.) from Ethiopia as revealed by inter simple sequence repeat (ISSR) markers. Afr J Biotechnol 14(18):1543–1551. DOI:10.5897/AJB2015.14567
6. Bourgou S, Pichette A, Marzouk B, Legault J (2010) Bioactivities of black cumin essential oil and its main terpenes from Tunisia. S Afr J Bot 76(2):210–216. https://doi.org/10.1016/j.sajb.2009.10.009
7. Dutta S, Kumawat G, Singh BP et al (2011) Development of genic-SSR markers by deep transcriptome sequencing in pigeonpea [*Cajanus cajan* (L.) Millspaugh]. BMC Plant Biol 11(1):17. https://doi.org/10.1186/1471-2229-11-17
8. Ellegren H (2004) Microsatellites: simple sequences with complex evolution. Nat Rev Genet 5(6):435–445. https://doi.org/10.1038/nrg1348
9. Golkar P, Nourbakhsh V (2019) Analysis of genetic diversity and population structure in *Nigella sativa* L. using agronomic traits and molecular markers (SRAP and SCoT). Ind Crop Prod 130:170–178. https://doi.org/10.1016/j.indcrop.2018.12.074
10. Iqbal MS, Nadeem S, Mehboob S et al (2011) Exploration of genotype specific fingerprinting of *Nigella sativa* L. using RAPD markers. Turk J Agric For 35(6):569–578. doi:10.3906/tar-1001-622
11. Islam MT, Khan MR, Mishra SK (2019) **() An updated literature-based review**: **phytochemistry**, **pharmacology and therapeutic promises of** Nigella sativa **L**. **Oriental Pharm Exp Med 1–15**. https://doi.org/10.1007/s13596-019-00363-3
12. Karaca M, Ince AG, Aydin A, Ay ST (2013) Cross-genera transferable e-microsatellite markers for 12 genera of the *Lamiaceae* family. J Sci Food Agric. https://doi.org/10.1002/jsfa.5982
13. Karaca M, Ince AG, Aydin A, Elmasulu SY, Turgut K (2015) Microsatellites for Genetic and Taxonomic Research on Thyme (*Thymus* L.). Turk J Biol. https://doi.org/10.3906/biy-1406-20
14. KorehKhosravi SH, Masoumiasl A, Dehdari M (2018) A comparative analysis of RAPD and ISSR markers for assessing genetic diversity in Iranian populations of *Nigella sativa* L. Cel Mol Biol 64(1):52–59. https://doi.org/10.14715/cmb/2018.64.1.10
15. Mirzaei K, Mirzaghaderi G (2017) Genetic diversity analysis of Iranian *Nigella sativa* L. landraces using SCoT markers and evaluation of adjusted polymorphism information content. Plant Genet Resour 15(1):64–71. https://doi.org/10.1017/S1479262115000386
16. Morgante M, Hanafey M, Powell W (2002) Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. Nat Genet 30(2):194–200. https://doi.org/10.1038/ng822
17. Neghab MG, Panahi B (2017) Molecular characterization of Iranian black cumin (*Nigella sativa* L.) accessions using RAPD marker. BioTechnologia *98*(2). https://doi.org/10.5114/bta.2017.68308
18. Poyraz I (2014) An efficient DNA isolation method from *Nigella sativa* L.(*Ranunculaceae*) seeds for RAPD and ISSR analysis. Bilecik Seyh Edebali university journal of science 1(1):22–27. https://degipark.org.tr/tr/pub/bseufbd/issue/22290/239070
19. Roldan-Ruiz I, Dendauw J, Bockstaele EV et al (2000) AFLP markers reveal high polymorphic rates in ryegrasses (*Lolium* spp.). Mol Breed 6:125–134. https://doi.org/10.1023/A. : 1009680614564
20. Sonah H, Deshmukh RK, Sharma A et al (2011) Genome-wide distribution and organization of microsatellites in plants: an insight into marker development in Brachypodium. PLoS ONE. https://doi.org/10.1371/journal.pone.0021298
21. Senan S, Kizhakayil D, Sasikumar B, Sheeja TE (2014) Methods for development of microsatellite markers: an overview. Not Sci Biol 6(1):1–13. https://doi.org/10.15835/nsb619199

22. Şelale H, Çelik I, Gültekin V, Allmer J, Doğanlar S, Frary A (2013) Development of EST-SSR markers for diversity and breeding studies in opium poppy. Plant Breed 132(3):344–351. https://doi.org/10.1111/pbr.12059
23. Taşcıoğlu T, Sadıkoğlu N, Doğanlar S, Frary A (2018) Molecular genetic diversity in the *Origanum* genus: EST-SSR and SRAP marker analyses of the 22 species in eight sections that naturally occur in Turkey. Ind Crops Prod 123:746–761. https://doi.org/10.1016/j.indcrop.2018.07.027
24. Tan LQ, Wang LY, Wei K et al (2013) Floral transcriptome sequencing for SSR marker development and linkage map construction in the tea plant (*Camellia sinensis*). PLoS ONE 8(11):e81611. https://doi.org/10.1371/journal.pone.0081611
25. Tulsani NJ, Hamid R, Jacob F et al (2020) Transcriptome landscaping for gene mining and SSR marker development in Coriander (*Coriandrum sativum* L.). Genomics 112(2):1545–1553. https://doi.org/10.1016/j.ygeno.2019.09.004
26. Xiao M, Zhang Y, Chenc X, Lee EJ et al (2013) Transcriptome analysis based on next-generation sequencing of non-model plants producing specialized metabolites of biotechnological interest. J Biotechnol 166(3):122–134. http://dx.doi.org/10.1016/j.jbiotec.2013.04.004
27. Wang X, Wang L (2016) GMATA: an integrated software package for genome-scale SSR mining, marker development and viewing. Front Plant Sci. https://doi.org/10.3389/fpls.2016.01350 *7*, 1350
28. Wei W, Qi X, Wang L et al (2011) Characterization of the sesame (*Sesamum indicum* L.) global transcriptome using Illumina paired-end sequencing and development of EST-SSR markers. BMC genom 12(1):451. https://doi.org/10.1186/1471-2164-12-451
29. Yimer EM, Tuem KB, Karim A et al (2019) *Nigella sativa* L.(black cumin): a promising natural remedy for wide range of illnesses. Evid-based Complement Altern Med. https://doi.org/10.1155/2019/1528635
30. Zhang J, Liang S, Duan J et al (2012) De novo assembly and characterisation of the transcriptome during seed development, and generation of genic-SSR markers in peanut (*Arachis hypogaea* L.). BMC genom 13(1):90. https://doi.org/10.1186/1471-2164-13-90

Figures

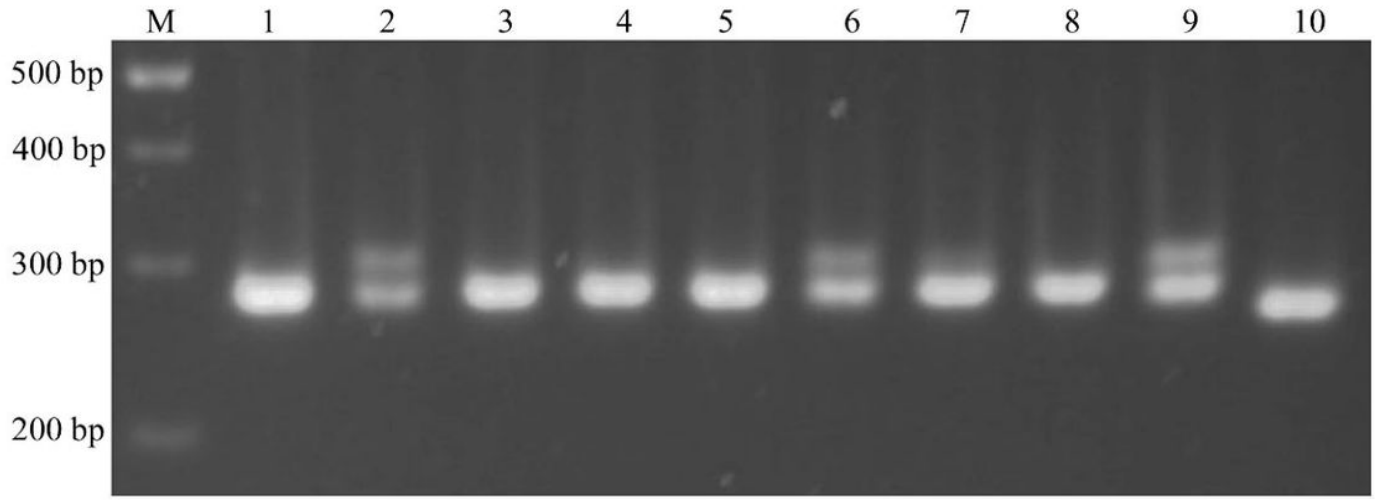

Figure 1

Agarose gel electrophoresis image of NS-2136 EST-SSR primer.

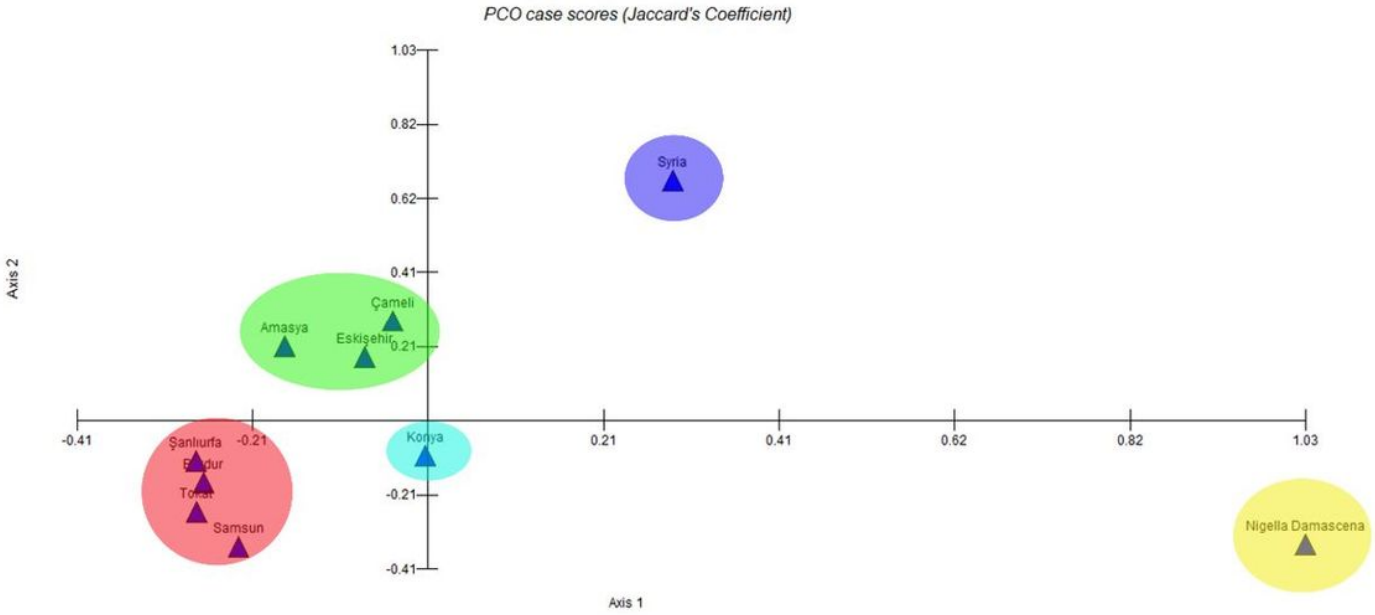

Figure 2

PCoA analysis plot of the first 2 axes.

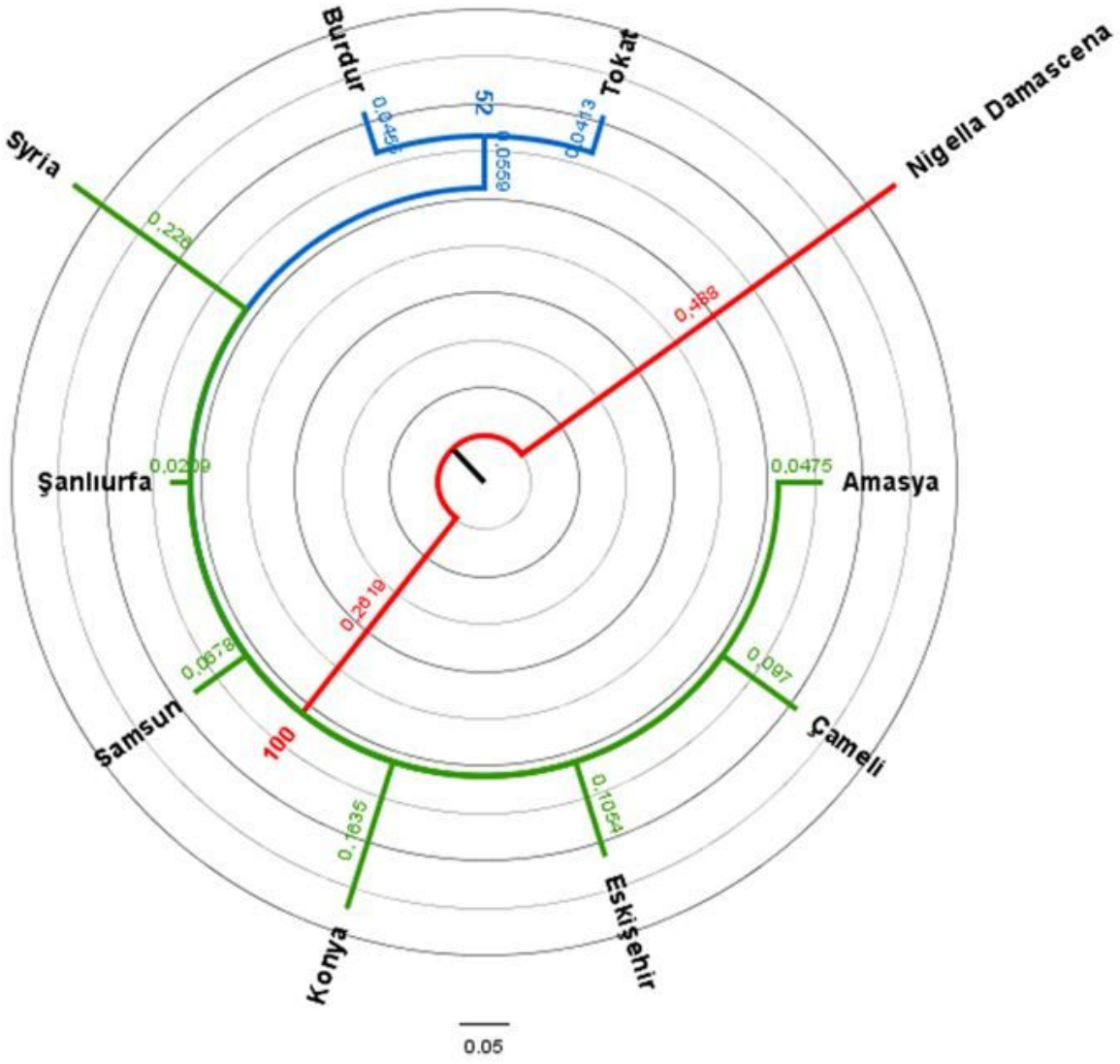

Figure 3

The 50% majority-rule consensus dendrogram based on SSR markers develped in the presnt study.