Annals of Arid Zone 61(1): 71-74, 2022

Short Communication

Deciphering Chloroplast Genome of Indian Cultivar Gujarat Cumin 4: First Report

Rukam S. Tomar[1], Ramesh K. Solanki[2]*, M.A. Iquabal[3], S.M. Padhiyar[1] and J.K. Kheni[1]

[1]*Junagarh Agricultural University, Junagarh 362 001, India*

[2]*ICAR-Central Arid Zone Research Institute, Jodhpur 342 003, India*

[3]*ICAR-Indian Agricultural Statistic Research Institute, New Delhi 110 012, India*

Received: February 2022

Cumin (*Cuminum cyminum*), one of the most important spice crop of hot arid regions of western India and of Mediterranean countries mainly Syria, Turkey, UAE and Iran is grown for its aromatic and nutraceutical values since time immemorial (Kafi *et al.*, 2006). The crop is well adapted to cool and dry weather, and is sensitive to both abiotic and biotic stresses causing productivity loss. Looking to the genotypic or ecotype variation existing in the Mediterranean regions it has been suggested as its place of origin (Kafi *et al.*, 2006). India is leader in cumin cultivation, in year 2020-21 more than 8.56 lakh tons seed was produced from an area of around 12.4 lakh ha. Around 2.98 lakh tons was exported of worth INR 4251 corers (Anonymous, 2021). China was the leading importing country for year 2020-21 of Indian cumin by purchasing 1.02 lakh tons of seed worth INR 1397 corers (Anonymous, 2021). Indian western hot arid and semi-arid zone represented by states of Rajasthan and Gujarat is the ‘Bowl of Cumin cultivation’, over the years the area and production is increasing gradually; for the year 2020-21 the Rajasthan and Gujarat state acreage was 7.7 lakh ha and 4.69 lakh ha respectively with production of around 4.2 lakh tons from both the states (Anonymous, 2021). In Rajasthan, more than 80% area in cumin is under arid zone covering districts of Jodhpur, Barmer, Jalore and Bikaner (Anonymous, 2019). National productivity has increased over decades by adoption of high yielding cultivars, cv. Gujarat Cumin 4 (GC 4) is most popular and predominantly grown in all cumin growing area for the last one and half decade for its high yield potential and resistance/tolerance to *Fusarium* wilt disease (Singh and Solanki, 2015).

The genotypic constitution of GC 4 has prominently dominated for its adaptability over location over years, hence deciphering the genomic information resting in the nuclear part and it’s organelles i.e., chloroplast (cp) and mitochondria (mt) can provide valuable nucleotide bio-informatics database for knowing the genetic strength of the phenotype. It also gives opportunity for the molecular breeder to understand the genomic diversity which may be helpful in planning genetic improvement program following advance biotechnological tools.

Whole genome sequencing of GC 4 was carried out using Next Generation Sequencer Ion S5 at the Division of Biotechnology and Biochemistry, Junagarh Agricultural University, Junagarh, Gujarat (India); total 56.57 GB nucleotide data was generated for analysis. The chloroplast reads were filtered from the generated data by comparing them with reference genomes using bowtie2 (version 2.3.4.3). The mapped chloroplast reads were assembled into a circular contig using CLC Genomics Workbench v. 3.6.1. (Fig. 1). The complete chloroplast genome of cumin was preliminarily annotated using the online program DOGMA (Dual Organellar Genome Annotator; http://dogma.ccbb.utexas.edu) followed by manual correction.

The analysis of GC 4 genomic sequences deciphered a circular chloroplast genomic map of 155,443 bp in length (Fig. 1) with an overall GC content of 37.29%, it is submitted in the NBCI Genbank database having submission number OM731709. A total of 187 genes were annotated, of which 35 were exons, 18 were introns, 147 were CDS and 62 were responsible for tRNA. Genes related to Photosynthesis (rubisco, Photosystem I, Photosystem II, ATP synthase, Cytochrome and NADPH

---

*E-mail: rksolanki.ars@gmail.com

Table 1. Gene annotation of the Cumin var. GC 4 Chloroplast genome

Category	Group	Genes
Photosynthesis related genes	Rubisco	rbcL
	Photosystem I	psaA, psaB, psaC, psaI, psaJ
	Photosystem II	psbA, psbB, psbC, psbD, psbE, psbF, psbH, psbI, psbJ, psbK, psbL, psbM, psbT, psbZ
	ATP synthase	atpA, atpB, atpE, atpF, atpH, atpI
	Cytochrome complex	petA, petB, petD, petN, petL, petG
	Cytochrome synthesis	ccsA
	NADPH dehydrogenase	ndhA, ndhB, ndhC, ndhD, ndhE, ndhF, ndhH, ndhG, ndhJ, ndhK, ndhI
Transcription and translation related genes	Transcription	rpoA, rpoC2, rpoC1, rpoB
	Ribosomal proteins	rps2, rps3, rps4, rps8, rps11, rps12, rps14, rps15, rps16, rps19, rpl2 (×2), rpl14, rpl16, rpl20, rpl22, rpl23 (×2), rpl32, rpl33, rpl36
	Translation initiation factor	infA
RNA genes	Transfer RNA	trnH-GUG, trnK-UUU (×2), trnQ-UUG, trnS-GCU, trnG-UCC (×2), trnR-UCU, trnC-GCA, trnD-GUC, trnY-GUA, trnE-UUC, trnT-GGU, trnS-UGA, trnG-UCC, trnfM-CAU, trnS-GGA, trnT-UGU, trnL-UAA (×2), trnF-GAA, trnV-UAC (×2), trnM-CAU, trnW-CCA, trnP-UGG, trnI-CAU, trnL-CAA (×2), trnV-GAC, trnI-GAU (×3), trnA-UGC (×2), trnR-ACG (×2), trnN-GUU (×2), trnL-UAG, trnN-GUU, trnR-ACG, trnA-UGC (×2), trnV-GAC, trnI-CAU
Other genes	RNA processing	matK
	Carbon metabolism	cemA
	Fatty acid synthesis	accD
	Proteolysis	clpP
Unknown function	Conserved ORFs	ycf1, ycf2, ycf4

dehydrogenase), transcription, translation, ribosomal protein, translational factors, tRNA, RNA processing, carbon metabolism, fatty acid synthesis were observed (Table 1 and Fig. 1).

Genomic information on cumin crop is meager. In the era of molecular breeding, basic information of the nucleotide sequences in the crop genome is essential. Initially, in cumin arbitrary PCR based markers were used to know the molecular diversity few examples are use of RAPD markers (Bahraminejed *et al*., 2012) and SRAP markers (Bhatt *et al.,* 2017). Carrot is the model crop of Apiaceae family, hence looking to its closeness to cumin its SSR markers were screened for knowing genetic diversity (Kumar *et al.,* 2014). Cumin specific markers were also developed by obtaining next generation sequencing data to design 25 SSR primers and 30 genotypes were characterized; of the 25 primers 21 showed cross transferability in other Apiaceae species also (Bharti *et al*., 2018). Flow cytometric analysis revealed high genome size variation in cumin genotypes ranging from 1.74 pg to 2.66 pg Indian cultivars, in which GC 4 genome size was found to be of 2.51 pg (Hasan *et al.,* 2016) and in other studies GC 4 genome was reported to be of 2.66 pg (Tomar *et al.,* 2022). Cumin complete genome sequence information is yet to come, in year 2020 a local cultivar of cumin from Kuerle, Xinjiang, China was sequenced to obtain chloroplast genome of 157,389 bp (GenBank accession number: MN 901636; Zhou and Liu, 2020) which had showed closeness to *Dacus carota* cp genome.

Cumin is predominantly cultivated in India and more than 80% of the world demand is met by Indian cumin export. Though information on cp genome of cumin is available on a local genotype from uncultivated areas of China, but the present study on widely cultivated Indian variety GC 4 can provide better insight for strategically planning of molecular breeding programs in the country. In the present investigation the chloroplast

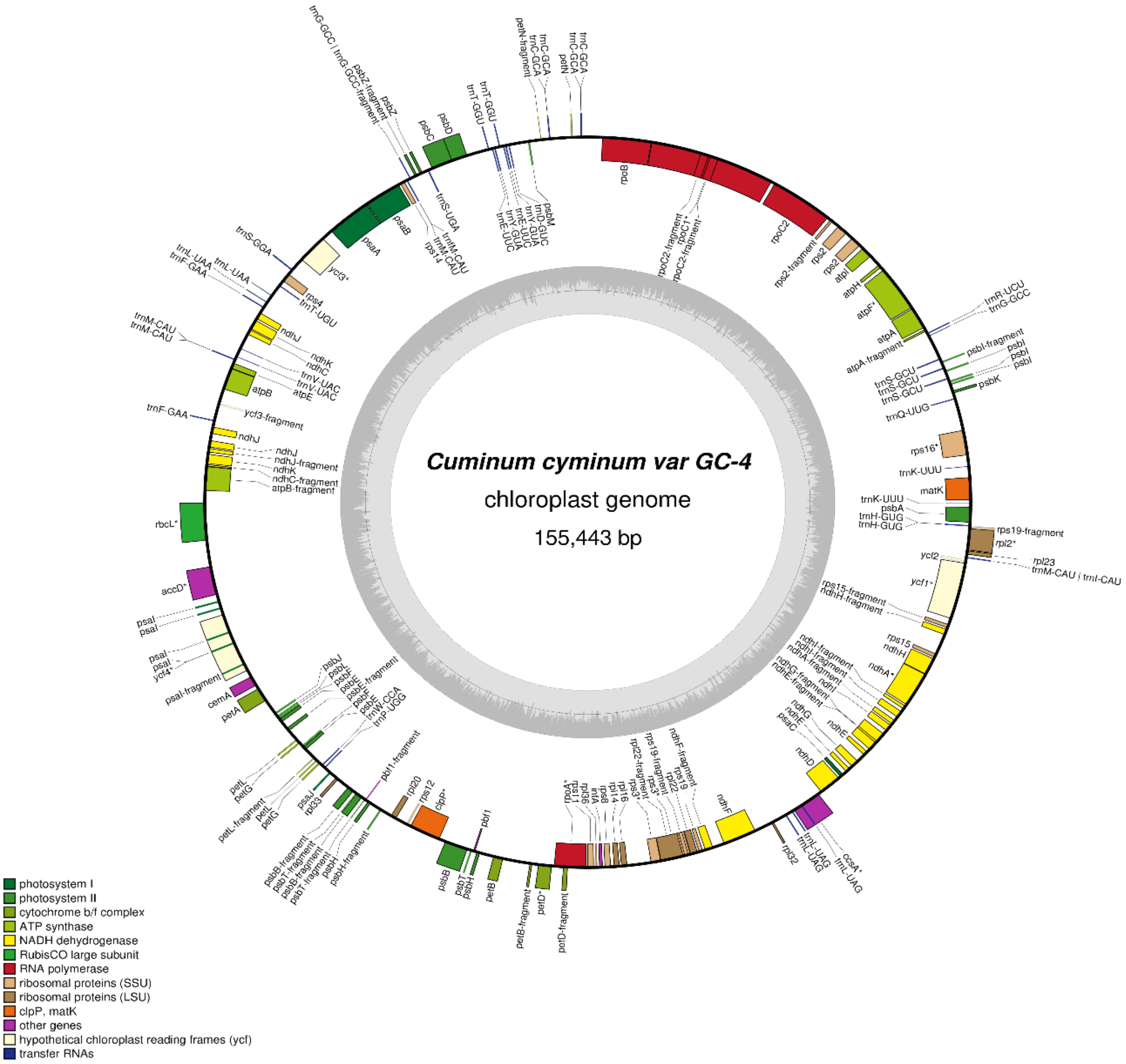

Fig. 1. Circular map of cumin var. GC 4 chloroplast (cp) genome. Gene map showing 187 annotated genes with different functional groups that are colour-coded on outer circle as transcribed clock-wise (outside) and transcribed counter clock-wise (inside). The inner circle indicates the GC content as dark grey plot.

genome information deciphered for GC 4 and is first report on a cultivated variety. Further comparison of the genome with that of reported cp genomes of cumin and other Apiaceae crop will help in drawing a robust phylogenetic tree of the genus. It can also help in development of organelle specific markers and in functional biology studies in the crop.

Acknowledgment

The authors are highly thankful to ICAR for project Grant under ICAR-Extra Mural Fund and are also thankful to Director ICAR-NRCSS, Ajmer, Rajasthan (India) and The Head, Division of Biotechnology and Biochemistry, Junagarh Agricultural University, Junagarh, Gujarat (India) for providing needful facilities and support.

References

Anonymous 2019. *Rajasthan State Agriculture Statistics*. https://agriculture.rajasthan. gov.in

Anonymous 2021. *Spices Board, India*. www. indianspices.com

Bahraminejad, A., Ghasem, M.N., Mihdzar, A.K. and Mohd, R.B.Y. 2012. Molecular diversity of cumin

(*Cuminum cyminum* L.) using RAPD markers. *Australian Journal of Crop Science* 6(2): 194-199.

Bharti, R., Kumar, S. and Parekh, M.J. 2018. Development of genomic simple sequence repeat (SSR) markers in cumin and their application in diversity analyses and cross-transferability. *Industrial Crop and Products* 111: 158-164.

Bhatt, J., Kumar, S., Patel, S. and Solanki, R. 2017. Sequence-related amplified polymorphism (SRAP) markers based genetic diversity analysis of cumin genotypes. *Annals of Agrarian Science* 15(4): 434-438.

Hasan, A.M., Sakure, A.A., Fougat, R.S. and Kumar, S. 2016. Estimation of nuclear DNA content in cumin. *International Journal of Seed Spices* 6(1): 96-99.

Kafi, M., Mohaseel, M.H.R., Koocheki, A. and Nassiri, A. 2006. *Cumin* (*Cuminum cyminum*) *Production and Processing*. Science Publishers, Enfield, NH, USA ISBN-57808-50407.

Kumar, S., Asamadi, M.H. and Fougat, R.S. 2014. Transferability of carrot (*Daucus carota*) microsatellite markers to cumin (*Cuminum cyminum*). *International Journal of Seed Spices* 4(1): 88-90.

Singh, Balraj and Solanki, R.K. 2015. Status of seed spices research and development in India. *Indian Journal of Agricultural Sciences* 85(2): 151-156.

Tomar, R.S., Solanki, R.K., Padhiyar, S.M., Kheni, J.V. and Kakani, R.K. 2022. Flow cytometric genome size estimation of major seed spice grown in India. *The Pharma Innovation* 11(6): 507-510.

Zhou, J. and Liu, Z. 2020. The complete chloroplast genome and phylogenetic analysis of *Cuminum cyminum*. *Mitocondrial DNA Part B*, 5(1): 1079-1080.

Printed in March 2022