ISSN 2394 - 7837

ARTICLE

INSILICO METHOD FOR THE IDENTIFICATION OF SNPs AND SSRs IN AMORPHOPHALLUS.

Nisha. T. Nair, Prabhakumari. C, Nair Sreecha Chandran

Received 05/10/2018 Accepted 01/01/2019

Abstract

Amorphophallus has a place with the sort of perpetual plants dissipated in the tropics of West Africa and South Asia. Corms of this class have abnormal state of water-dissolvable glucomannan and thusly it has for some time been utilized as a restorative herb and sustenance source. Dasheen mosaic infection and numerous other unreported putative infections are associated with various viral mosaic contaminations of Amorphophallus. Due to the absence of genetic markers, genetic investigations of Amorphophallus have been held up. Molecular genetic markers are a standout amongst the most incredible assets for genome examination and gathering of heritable qualities with fundamental hereditary variety. Single nucleotide polymorphisms (SNPs) and Simple Sequence Repeats (SSRs) created through high-throughput techniques prompted the uprising in their utilization as atomic markers. Bioinformatics techniques are useful for creating molecular markers subsequently making their improvement quicker and less expensive. By now, a number of computer programs have been implemented with the aim of identifying molecular markers from sequence data. Here w a different pipeline was developed for the identification of SNPs and SSRs.

Keywords: Single Nucleotide Polymorphism, Simple Sequence Repeats

Introduction

Amorphophallus paeoniifolius is one of the important aroid tuber vegetable crops cultivated for corms in tropical and the subtropical region of Asia (Ravi et al 2009).The Amorphophallus tubers are a delicacy in food and rich in nutrients. It is also used as a vegetable in delicious cuisines. The tuberous roots of the plant posses blood purifier properties and have been used traditionally for the treatment of piles, abdominal disorders, tumours, enlargement of spleen, asthma and rheumatism. They are traditionally used in arthralgia, elephantiasis, tumors, inflammations, hemorrhoids, hemorrhages and various other diseases. The stalk roots of the plant have conjointly been according to possess tonic, internal organ and course properties (Anuradha Singh, 2013).

Among the various diseases affecting *A. paeoniifolius*, the collar rot and the mosaic disease are the main ones (Misra 1997; Misra et al. 2003; Misra and Nedunchezhiyan 2008; Babu et al. 2011a). In India, the mosaic disease has rised as a devastating disease in recent years (Hegde et al. 2006; Khan et al. 2006). The symptoms of mosaic disease include severe leaf deformation, yellowing; puckering to shoe string. DsMV is a single stranded positive sense potyvirus. It is transmitted via vegetative propagation. Use of disease free planting material can be used as a cultural measure of immediate success to prevent the spread of this important disease.

Due to the importance of *A. paeoniifolius* and the disease affecting the crop causes serious constraints to its production. Producing virus resistance improves breeding goal. The utilization of virus resulting resistance by transgenic expression of sequences from DsMV in the plant presents as a promising approach. This approach proved very powerful to combat plant diseases caused by several viruses including potyviruses (Sonoda et al. 1999; Scorza et al 2001; Jan et al. 1999; Mundembe et al. 2009).

Genetic marker is a gene with a known location on a chromosome and associated with a particular gene or traits. It can be described as a variation, which may be arising due to mutation or alteration in the genomic loci. Genetic markers are closely linked with the target gene. Genetic markers are broadly classified in to two categories, which are classical marker and molecular marker. Morphological, cytological and biochemical markers are type of classical markers. Examples for classical markers are Restriction Fragment Length Polymorphism (RFLP), Amplified Fragment Length Polymorphism (AFLP), Simple Sequence Repeats (SSR), Single Nucleotide etc.

Materials and Methods

Sample Collection

The DsMV affected Amorohophallus paeoniifolius samples

CEPCI Laboratory and Research Institute, Cashew Bhavan, Mundakkal West, Kollam-691001.
Corresponding author, e mail: sreecha_nair123@rediffmail.com

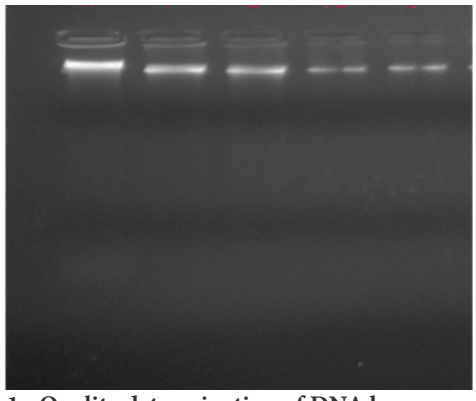

Figure 1 - Quality determination of DNA by agarose gel

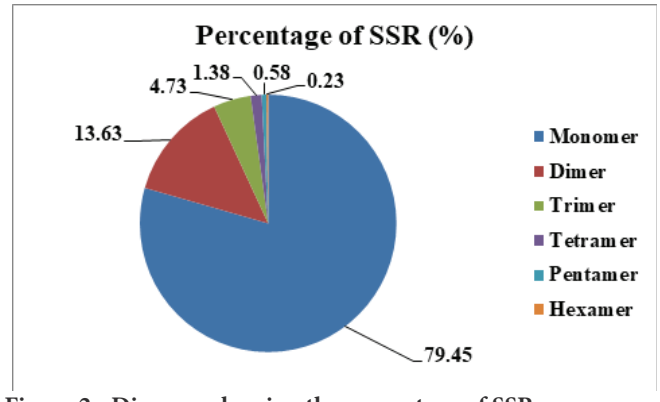

Figure 2 - Diagram showing the percentage of SSR

Table 1. Distribution of different repeat type classes of SSR

Serial No.	Type of SSR	Number of SSR	Percentage of SSR (%)
1	Monomer	688	79.45
2	Dimer	118	13.63
3	Trimer	41	4.73
4	Tetramer	12	1.38
5	Pentamer	5	0.58
6	Hexamer	2	0.23
Total Number of SSR		866	100

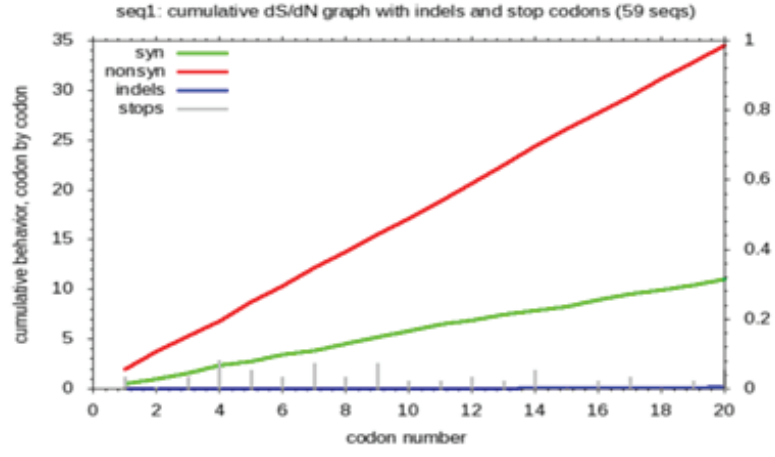

Figure 2 - Graph showing the distribution of SNP.

were collected
Plant part used – leaves

DNA Isolation

Leaf samples of 100mg of Amorphophallus were crushed with CTAB buffer using Morten and Pestle. After that add 1ml of extraction buffer (prewarmed at 65°C). 750µl of the crude extract was transferred to fresh (2ml) tube. The sample was incubated at 65°c for 1hour. Then centrifugation was carried out at 1000rpm for 10 minute at 27°C. Then

Table 2. Distribution of SNP in sample 1

Per Codon				
codon	indel	syn	nonsyn	Aa
1	0.00	0.46	1.97	E
2	0.00	0.51	1.72	I
3	0.00	0.65	1.58	L
4	0.00	0.74	1.52	L
5	0.00	0.41	1.97	Y
6	0.00	0.73	1.53	T
7	0.00	0.34	1.88	S
8	0.00	0.69	1.62	S
9	0.00	0.72	1.74	R
10	0.00	0.61	1.60	T
11	0.00	0.62	1.70	R
12	0.02	0.42	1.78	F
13	0.02	0.56	1.91	L
14	0.02	0.42	186	N
15	0.02	0.43	1.80	R
16	0.02	0.68	1.54	T
17	0.02	0.56	1.61	S
18	0.02	0.41	1.78	F
19	0.02	0.51	1.68	S
20	0.02	0.57	1.68	T

the supernatant was transferred to fresh 2ml tube. 10µl of RNase was added and mixed properly by tapping and incubated at 37°C for 1 hour. Then equal volume of chloroform: isoamyl alcohol (24:1) was added to supernatant and centrifuged at 15000rpm for 10 minute at 4°C. Transfer the aqueous layer or supernatant into fresh tube. Add 0.8 volume of ice cold isopropanol to the supernatant (800µl) and slowly mix. Incubate at -20°C for 1 hour or overnight. Following centrifugation at 12000rpm or 15000rpm for 10 minutes at 4°C the pellet was washed with 0.5 ml of 70% ethanol of centrifuging at 12000rpm for 5 minutes at 4°C. The washing was repeated with 0.5 ml of 75% ethanol. DNA pellet obtained was air dried at 37°C for 30 minutes and then dissolved in 15µl sterile distilled water or 1X TE buffer. After incubating at 37°C for 1 hour, the DNA was stored at -20°C.

Agarose gel electrophoresis:

Prepare the agarose gel (50 ml 1x TBE buffer add 0.5 g agarose) after boiling and cooling add 1 µl ethidium bromide solution mix well without making bubbles. Choose an suitable comb for forming the samples slots in the gel. The warm agarose solution is poured into the mold. Then allow the gel to set completely (20-30 minutes at room temperature) carefully remove the cello tape and place the casting tray on the buffer tank. Combine the sample loading buffer and the DNA samples (2 µl sample loading buffer+4 µl DNA sample). Gradually load the sample mixture in to the slots of the submerged gel using a micropipette. Run the gel until the bromophenol blue migrate to an appropriate distance through the gel. After switching off the power gel tray may be removed and placed directly on a Trans illuminator.

Sequencing of Obtained Data

The isolated DNA samples were sequenced using rbcL and matK in Rajiv Gandhi Centre for Biotechnology, Thiruvanadhapuram.

SSR Prediction Usng MISA (MIcroSAtellite identification tool)

MISA is a SSR prediction tool. Pasted sample sequence in FASTA file given in input area. The output file formats of SSR data are MISA or GFF. The monomer ((A) n), dimer ((AT) n), trimer ((ATT) n), tetramer ((CCGG) n), pentamer ((CGATA)n), hexamer ((AGCTAA)n) are types of SSR. Types of SSR can also be detected by using MISA.

SNP Prediction Using Bioinformatics ToolSNAP

SNAP is a SNP identification tool found in HIV Database which is accessible in https://www.hiv.lanl.gov/content/sequence/SNAP/SNAP.html. SNAP calculates synonymous (they may leave the sequence unchanged) and non-synonymous substitution (amino acid sequence changes) rates based on a set of codon-aligned nucleotide sequences. The sample sequences are inserted in the input and the options are selected. Then date was submitted.

Primer Designing Using PRIMER3PLUS

Primer3plus is a user friendly tool which is available in http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi. Primers are selected according to GC content and melting temperature. The primer sequence should have GC content between 40 and 60%, with the 3' of a primer ending in C or G to promote binding. The melting temperature, (Tm) of the hairpin must be less than the annealing temperature for the reaction; on average it should range between 55°C and 65°C.

Table 3. List of deigned primers

Sl No	Left primer	Length	Tm	GC	Right primer	Length	Tm	GC
1	GTCCCCCTTTCTCGCGTAAA	20bp	60.0 C	55.0 %	GGATCTCCGATCGGGAAACC	20 bp	60.0 C	60.0 %
2	TGTTCGCATCGGATCTCAGG	20 bp	59.9 C	55.0 %	GAAATCAACCGAGACCCCGT	20 bp	60.0 C	55.0 %
3	TTCCCATTTCCTCGCTGACC	20 bp	59.9 C	55.0 %	TGGAAAAACCAGCAAGCAGC	20 bp	60.0 C	55.0 %
4	TTCCCATTTCCTCGCTGACC	20 bp	60.0 C	55.0 %	TGGAAAAACCAGCAAGCAGC	20 bp	59.9 C	50.0 %
5	GCTGCTTGCTGGTTTTTCCA	20 bp	59.9 C	50.0 %	CTGGGGTTGAAGAAGGTCCC	20 bp	60.0 C	60.0 %
6	GAGTCTGCTTGGCTGGTAGG	20bp	60.1 C	60.0 %	AAGAAGGGTATTGCGGGCAA	20 bp	60.0 C	50.0 %
7	TGGCCTGCACCTTTTCAAGA	20 bp	60.1 C	50.0 %	CAGAGTGGCTCGAGGAGTTG	20 bp	60.1 C	60.0 %
8	TTGCCCGCAATACCCTTCTT	20 bp	60.0 C	50.0 %	AAAGGCACACTGCAAGGGAT	20 bp	60.2 C	50.0 %
9	GGGTCAGTGACGTACAAGCA	20 bp	60.0 C	55.0 %	TTGTTTGGGCTGCCCTTAGT	20 bp	59.8 C	50.0 %
10	CAACTCCTCGAGCCACTCTG	20 bp	60.1 C	60.0 %	TTTCAGGGCCCCAAAGAAGG	20 bp	60.2 C	55.0 %
11	TGGCCAAACGTTTCAAACCC	20bp	59.8 C	50.0 %	GGCTGCGTGTATACTGTCGA	20 bp	59.9 C	55.0 %
12	CCTCCAGACATGCACAACGA	20 bp	60.3 C	55.0 %	GGAAGGGCCCCTAATGCTTT	20 bp	60.0 C	55.0 %
13	CAAGGCACATCCTCAACCGA	20 bp	60.3 C	55.0 %	GGATCCATTGGAACCCGGTT	20 bp	60.0 C	55.0 %
14	TTGCCCTGATTGGATGAGCA	20 bp	59.7 C	50.0 %	CCATTGGAACCCGGTTCTCA	20 bp	60.0 C	55.0 %
15	CTCAAGCACAGTCTCACGGA	20 bp	59.7 C	55.0 %	TTCTTTGCCTTTGGTCGGGT	20 bp	60.1C	50.0 %
16	ATCCCTGCCACGAAAGACTG	20bp	60.0 C	55.0 %	AACGTCCTTGAGAGCCCAAG	20 bp	60.0 C	55.0 %
17	GGGGTAGCAATTGGGGAGAC	20 bp	60.1 C	60.0 %	CTTCCCGGTTAGCTGTGTGT	20 bp	60.0 C	55.0 %
18	AGCAATTGGGGAGACGATGG	20 bp	60.1 C	55.0 %	TGTGTGTATGCTGCCTCCAG	20 bp	60.0 C	55.0 %
19	CTTGGGCTCTCAAGGACGTT	20 bp	60.0 C	55.0 %	TCCAACTGCCAGAACAGCAT	20 bp	59.9 C	50.0 %
20	AGAGAGGACGGCATAGTGGT	20 bp	60.0 C	55.0 %	AGATCCACAGGACCCTTGGT	20 bp	60.2C	55.0 %
21	ATCCCTGCCACGAAAGACTG	20bp	60.0 C	55.0 %	AACGTCCTTGAGAGCCCAAG	20 bp	60.0 C	55.0 %
22	GGGGTAGCAATTGGGGAGAC	20 bp	60.1 C	60.0 %	CTTCCCGGTTAGCTGTGTGT	20 bp	60.0 C	55.0 %
23	AGCAATTGGGGAGACGATGG	20 bp	60.1 C	55.0 %	TGTGTGTATGCTGCCTCCAG	20 bp	60.0 C	55.0 %
24	CTTGGGCTCTCAAGGACGTT	20 bp	60.0 C	55.0 %	TCCAACTGCCAGAACAGCAT	20 bp	59.9 C	50.0 %
25	AGAGAGGACGGCATAGTGGT	20 bp	60.0 C	55.0 %	AGATCCACAGGACCCTTGGT	20 bp	60.2C	55.0 %

Results and Discussion

DNA Quality Determination

By using CTAB method DNA were isolated. Quality of isolated DNA checked by agarose gel electrophoresis (AGE). The DNA samples were electrophoresed in agarose gel and orange bands were observed under uv transluminator. (Fig-

ure: 1)

Sequencing of Obtained DNA

Sequence data of selected samples were obtained from Rajiv Gandhi Centre for Biotechnology, Thiruvananthapuram. The sequence date was in the .rtf format.

SSR PREDICTION USING BIOINFORMATICS TOOLS

SSR Prediction using MISA

From 3514bp sequence, about 866 SSR were predicted. 688 monomer type SSR, 118 dimer type SSR, 41 trimer type SSR, 12 tetramer type SSR, 5 pentamer type SSR and 2 hexamer type SSR were predicted by using MISA tool. In MISA tool monomer type of SSR found to have highest percentage and hexamer type have lowest percentage of SSR.

SNP Prediction Using Bioinformatics SNAB

A total of 5 contigs were used for identifying SNPs using SNAP. Based on the type of SNPs these were further classified into synonymous SNP and non-synonymous SNPs. About more SNPs are non-synonymous SNPs. This means that all these SNPs will effect a change in the translated protein. Some SNPs were synonymous means that the mutations will not cause any change in the system. Based on the type of polymorphism these SNPs can be classified into InDel. In all samples non-synonymous SNPs shows higher values compared to synonymous SNPs and InDel.

Based on the cumulative behavior values of synonymous SNPs were found to be between 0.34-0.74, non synonymous SNPs were found to be between 1.52-1.97 and InDels were found to be 0-0.02. The graph below shows the ratio of synonymous, non-synonymous SNPS and InDels. From the graph Non-synonymous SNPs shows the higher value.

PRIMER DESIGNING USING PRIMER3PLUS

By using the primer3plus software 25 primers were generated. The primers were selected on the bases of GC content and melting temperature. The primer should have GC content between 40-60% and the average melting temperature between 55-65% was selected. On the bases of these criteria 8 primers were selected from 25 primers.

Conclusion

The present study is the development of SSR and SNP marker in Amorphophallus paeoniifolius against Dasheen mosaic virus using bioinformatics tools and primer designing using primer3plus. Amorphophallus paeoniifolius is the widely cultivated plant in south India. It is useful in elephantiasis, arthralgia, haemorrhoids, haemorrhages, vomiting, cough, bronchitis, asthma, anaemia etc. The Dasheen mosaic virus widely effects Amorphophallus paeoniifolius. The molecular marker studies will help to develop the plant which is tolerated to the Dasheen mosaic virus infection. SSR and SNP marker was developed with the help of bioinformatics tools. MISA is software which helps to predict the SSR and SNAP is software for the prediction of SNP. From these developed molecular markers, the primers were designed by using primer3plus. Then the primers are selected according to the GC content and the melting temperature. Further studies of the work include primers synthesis and validation of primers against different strains Amorphophallus paeoniifolius. Further studies are also needed for genetic diversity analysis and mutational studies of different samples of Amorphophallus paeoniifolius collected from various location of Kerala.

References

1. Abd-Elsalam, K. A. (2003). Bioinformatic tools and guideline for PCR primer design. african Journal of biotechnology, 2(5), 91-95.
2. Aggarwal, P. (2011). Bistro-Primer-Tool to design and validate specific PCR primer pairs for phylogenetic analysis.
3. Agung SEDAYU, Marcel C. M. EURLINGS, Barbara GRAVENDEEL and Wilbert L. A. HETTERSCHEID (2010). Morphological character evolution of Amorphophallus (Araceae) based on a combined phylogenetic analysis of trnL, rbcL and LEAFY second intron sequences, Botanical Studies, 51, 473-490.
4. Anil, S. R., Siril, E. A., & Beevy, S. S. (2011). Morphological variability in 17 wild elephant foot yam (*Amorphophallus paeoniifolius*) collections from southwest India. Genetic resources and crop evolution, 58(8), 1263-1274.
5. Anuradha Singh, Neeraj Wadhwa (2014). A Review on Multiple Potential of Aroid: *Amorphophallus paeoniifolius*, Int. J. Pharm. Sci. Rev. Res., 24(1), 55-60.
6. Baldwin, S., Pither-Joyce, M., Wright, K., Chen, L., & McCallum, J. (2012). Developments of robust genomic simple sequence repeat markers for estimation of genetic diversity within and among bulb onion (Allium cepa L.) populations. Molecular Breeding, 30(3), 1401-1411.
7. Beier, S., Thiel, T., Münch, T., Scholz, U., & Mascher, M. (2017). MISAweb: a web server for microsatellite prediction. Bioinformatics, 33(16), 2583-2585.
8. Binoy Babu, Vinayaka Hegde, T. Makeshkumar &M. L. Jeeva (2011). Molecular detection and identification of Dasheen mosaic virus infecting *Amorphophallus paeoniifolius*, Archives of Phytopathology and Plant Protection,Vol. 44, No. 13, 1248–1260.
9. C. Pan, Y.N. You, Y. Diao, Z.L. Hu1 and J.M. Chen (2012). Isolation and characterization of microsatellite loci for the herbaceous tuber crop, *Amorphophallus konjac* (Araceae), Genet. Mol. Res. 11 (4): 4617-4621.
10. Calef, C., Kuiken, C., Szinger, J., Gaschen, B., Abfalterer, W., Zhang, M., ... & Dalwani, A. (2005). Gateway to tools of HIV and HCV databases. HIV Mol. Immunol, 33, 56.
11. Chakravarthi, B. K., & Naravaneni, R. (2006). SSR marker based DNA fingerprinting and diversity study in rice (*Oryza sativa*. L). African Journal of Biotechnology, 5(9).
12. Chang, H. W., Chuang, L. Y., Cheng, Y. H., Hung, Y. C., Wen, C. H., Gu, D. L., & Yang, C. H. (2009). Prim-SNPing: a primer designer for cost-effective SNP genotyping. Biotechniques, 46(6), 421-431.