ORIGINAL RESEARCH PAPER
Volume - 13 | Issue - 11 | November - 2024 | PRINT ISSN No. 2277 - 8179 | DOI : 10.36106/ijsr
INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH

EXPLORING THE GENOME OF TOMATO SPOTTED WILT VIRUS (TSWV) FOR IDENTIFICATION OF MIRROR REPEATS

Biological Science

Vidhi Yadav Postgraduate (M.Sc.) Student – School of Life Sciences Starex University Gurugram.

Priya Yadav Postgraduate (M.Sc.) Student – School of Life Sciences Starex University Gurugram.

Dr. Sandeep Yadav* Assistant Professor, School of Life Sciences Starex University Gurugram.
*Corresponding Author ORCID- https://orcid.org/0000-0001-6668-7974

ABSTRACT

Background: A greater part of the genomic DNA comprises repetitive patterns scattered throughout the genome elucidating its effects through modification in the molecular processes of gene expression, genomic stability, and building of triplex structures, etc. The least studied and highly anticipated repetitive sequences are mirror repeats which have a significant role in the development of H DNA. This particular research aims to study the identification and distribution of mirror repeat patterns in the genome of Tomato Spotted Wilt Virus (TSWV). TSWV genome incorporates three different segments with varying lengths of nucleotides (L, M, and S). TSWV infects a variety of crops leading to remarkable losses in the overall weight, length and quality of the fruits. **Methods:** In this particular research, a well-acknowledged approach called FASTA PARALLEL COMPLEMENT BLAST (FPCB) was utilized. **Results:** All the genomic segments were found to have varying mirror repeat patterns, lengths & numbers distributed across the genome with the L-segment consisting of the maximum number of MR (208) while the S segment has the minimum (64). Mirror repeat with a maximum length of nucleotide was found in the L Segment (MR51) and mirror repeat with a minimum length of nucleotides (7) was found in all the genome segments types. **Conclusion:** This brief study will be helpful in the future to connect the impact of mirror repeats of TSWV with the contingency of diseases in agronomic crops.

KEYWORDS

TSWV (Tomato Spotted Wilt Virus), Mirror Repeats (MR), DNA, FPCB, NCBI, BLAST

INTRODUCTION

The genomic constitution of living domains shows the major fraction of the repetitive DNA. Various types of repetitive patterns of bases were observed.[1] They are responsible for a variety of roles including their involvement in the duplication of the genome [2], formation of intermediate agents (various RNA types) as well as in repairing genomic constitute, etc.[3] Repeat patterns are also called Junk DNA. The most popular and well-researched patterns among the variety of repeats are Tandem, Palindrome, Direct & Indirect type, SSR, SINE & LINE sequences, etc. The role they play in the fundamental principles of central dogma in molecular biology has great significance. Mirror Repeats (MR) are a different sort of repeat that has been identified in the genome but is not widely investigated despite playing important roles. These repeats are distinguished by having the same centre of symmetry on the same helix strand.[4] These are thought to be involved in the formation of unusual structure of the genome (like triplex) by folding.[5] Some sequences are responsible for developing neurological diseases like Friedreich's ataxia, a disorder that occurs in human beings due to the expansion of GAA repeat sequences.[6] Mirror repeats are thought to play critical roles in replication as well as transcription.[7-8] This particular study focused on identification of mirror repeats in Tomato spotted wilt viral RNA. Higher genetic variability, shorter generation time, and larger population size are some of the properties of RNA viruses chiefly responsible for their adaptation to new hosts.[9] The interaction between viral genes and proteins of the host plant/s determines the extensiveness of viral infection.[10] This virus (TSWV) is an ideal subject for in-depth investigation due to its intriguing biology and economic significance. Previous research conducted on *Bunyaviridae* indicates that viruses belonging to this family evolve through mutations, antigenic shifts, and host switchin.[11] The disease caused by TSWV spreading through a small insect, *Frankliniella occidentalis* (Western flower thrips) which acting as a vector for this virus. These western thrips possess the virus from infected plants at the time of their larval development and transfer it to the next host.[12] It was also reported that enhanced production of offspring by western thrips was discovered when they lay on TSWV-infected plants.[13] This can be attributed to the fact that more numbers of females were attracted to infected plants and higher chances of fecundity.[14] Symptoms of infection include characteristic purpling of tomato leaves, leaf distortion & plant stunting, and yellowing. Greatly affected plants may not form fruits leading to a huge market loss.[15] Economic losses because of TSWV are also seen in other plant species, especially in peppers.[16] An immense amount of research and efforts must be done to control and manage the spread of this virus and this deadly disease. The genome of TSWV is a negative sense ssRNA virus composed of 19.2-19.5 kb genetic material with tripartite genome (RNA) that is composed of Large (L), Medium (M), and Small (S) segments types.[17-18] The virology of this pathogen is widely acknowledged and crucial for the cultivation of agronomic crops across the globe. The L type segment is a negative sense RNA that encodes an RNA-dependent RNA polymerase (RdRp), while M type and S type are involved in their genetic organization.[19] Two glycoproteins, Gn and Gc project out from the cell surface, and a non-structural protein (NSm) participating in a cell-cell movement of the virus & encoded by M type.

A nucleocapsid protein (N) along with an additional non-structural protein [20], NSs having RNA silencing suppressor activity is encoded by S RNA.[21] The replication of virions in plants and thrips insects is a major part played by the G1 and G2 glycoproteins.[22] For efficient cell-to-cell transportation of this virus, an accurate and appropriate interaction between these glycoproteins is required.[23] All of these structures help in the successful entry and infection of the TSWV virus to the host cell. A comparative study has indicated that this virus has a proportional conserved genome.[24] A complex interaction takes place inside the Western thrips between the virus and host protein products for infection and replication of the virus. The first interaction takes place in the midgut lumen where through the utilization of glycoproteins interplays with the midgut receptors to enter the cells via an endocytic pathway.[25] From this point, migration to visceral muscle cells and salivary glands takes place for the sole purpose of multiplication and replication of the virus. During this process, the GP1 glycoprotein is thought to cover another GP2 glycoproteinwhich serves as a viral attachment protein aiding in the translocation of viral particles to other body parts and vertebrate cells.[26] The virus changes the behavior of its vectors for efficient transmission since vectors are mobile leading to viruses spreading within and between standing crops.[27] This particular bioinformatics based research on TSWV helps in many ways at the molecular level to study the viral species in many prospective.

MATERIAL & METHODS

The identification of the mirror repeats sequences in the Tomato spotted wilt virus (TSWV) genome utilized a manual FPCB bioinformatics based approach.[28-29] It includes a four-step process, started from using public domain databases (for gene/genome sequences), Sequence conversion tool and BLAST analysis (for aligning the sequences).

1) Step1: Downloading sequences

The complete sequence of the genome (L, M, and S types) of Tomato spotted wilt virus (TSWV) searched on the NCBI website and downloaded in FASTA format.[30] Each type's downloaded genome sequence file was divided into regions to make the process very simple and more accurate.

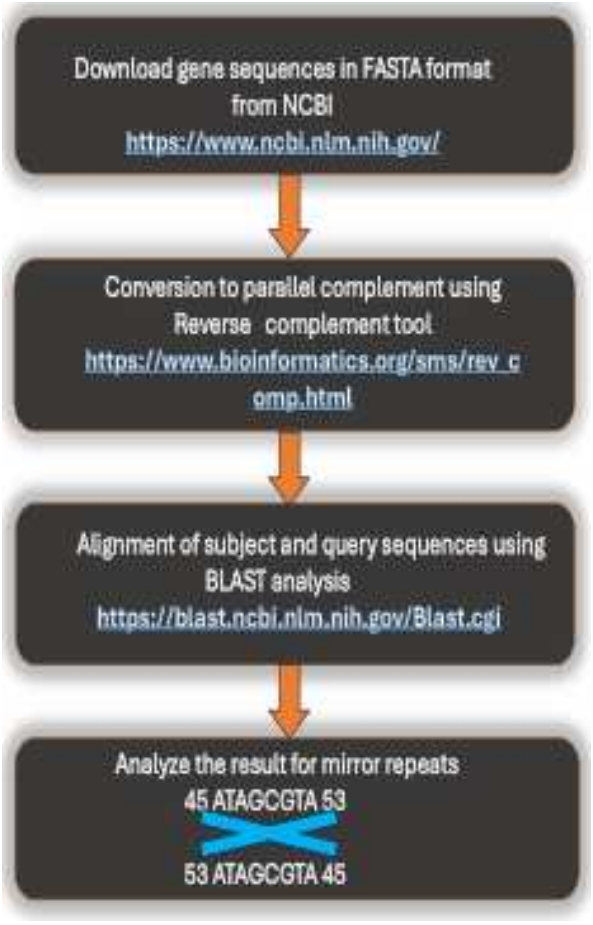

Figure 1: A flowchart of the methodology utilized in the identification of the mirror repeats in the TSWV's genome

2) Step 2: Conversion To Parallel Complement

After dividing the complete sequence file into smaller regions, each split region of the genome sequence is called the query sequence. Each query sequence is then paste into a bioinformatics tool called the Reverse complement tool to get its parallel complement. The parallel complement counterpart represents your subject sequence.

3) Step 3: Mirror repeats identification

Both types of sequences were aligned using BLAST as a homology searching tool to obtain somewhat similar sequences with specific algorithmic parameters like word size, and threshold frequency which are adjusted, accordingly using BLAST. Then, the obtained alignment results are analyzed to trace the mirror sequences in the genome or query sequence. It is found that, if the position number of the both aligned sequence are exactly reversed then a mirror repeat will be appear in the alignments.

4) Step 4: Result and analysis

RESULT & DISCUSSION

By FBCB analysis, the alignment results achieved in the Tomato spotted wilt virus genome (TSWV) contain multiple homolog alignment sequences in the tripartite segments that infect the agronomic and horticultural crops. This particular bioinformatics study performed on the genome of TSWV laid its foundation from previous studies performed on the Human insulin gene [31], bacteria (*E. coli* genes) [29], and fungi (*Candida albicans*).[32] The succeeding graphs present the distribution of number of mirror repeat sequences in the respective segments of the virus. A Minimum length of 7 nucleotides was almost found in different regions of the tripartite segments of the viral genome while a maximum length of 44 bps was seen in the L segment. A maximum of 37 and 31 were seen in type M and S segments respectively.

Table 1: An Outline Of The Region-wide Distribution Of No Of Mirror Repeat In Different Segments Of The TSWV's Genome

Genome segment/s	Regions	No. of Mirror repeats
Genome Segment L	Region 1	30
	Region 2	24
	Region 3	24
	Region 4	24
	Region 5	26
	Region 6	20
	Region 7	29
	Region 8	31
Genome Segment M	Region 1	24
	Region 2	26
	Region 3	14
	Region 4	16
Genome Segment S	Region 1	24
	Region 2	26
	Region 3	14

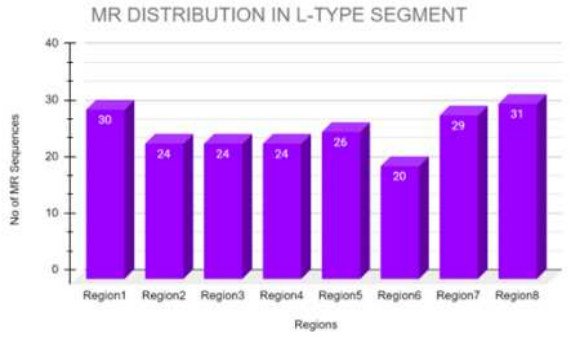

Figure 2: Depicted the occurrence of total MR sequences (region wise) in Segment L

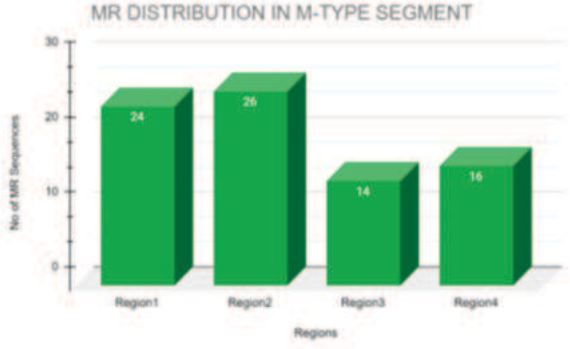

Figure 3: Shows the occurrence of total no (region wise) of MR sequences in Segment M

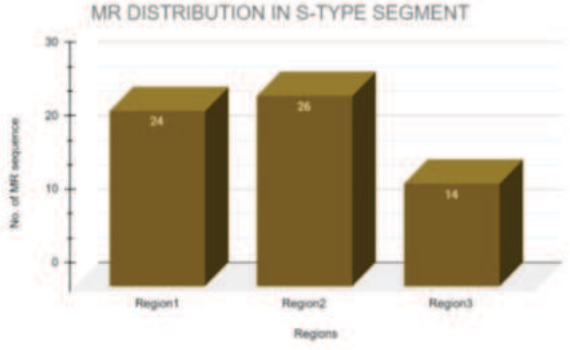

Figure 4: Depicted the graphical representation of total no of MR sequences (region wise) in Segment S

The following table contains the details of the distribution of some selected mirror sequences in the respective genome segment type of TSWV, along with other information regarding their length, position, etc.

Table 2 - Outlines The Presence Of Some Selected Mirror Repeat

Sequences In The TSWV genome

	GENOME TYPE L		
MR Number	**Sequence**	**Length**	**Position**
MR9	GAGATGGTCTCTCTGTTTGAG	21	289
MR18	AACTATGTGATATATAAAGAATCAA	25	532
Mr23	TCGTT-TCTATAGTTATATTGAAACCTATTGCT	32	713
MR47	CAAACAA-ATAAAGCAAGAGATATAACAAAC	30	579
MR51	AAGGTTCGTGTTTAACCAGAAACAAAAATGAATTT-TGCATGAAA	44	745
MR54	AAGTATAGACATGTCTTCTCTG-ACTCATATGAA	33	959
MR55	TttttCTACTTCATAATTTT	20	51
MR92	TATCTCAAAACAAAACTATAT	21	631
MR115	CAATGAAAGCTTATCACAACACT-TTGGAATGTTAC	35	430
MR119	TTTCAAAA-GTTTTTGCATAACTTT	24	611
MR144	TGCTAAAATTTGATTTAGATTCATATTGT	29	732
MR180	CAAAGAGATTGTAGATGATGATATCAGAGAAAC	33	272
MR182	AGAAGGTTTGGAAAACTGCAAAAG-TGTGGAAGA	33	344
MR185	TTCAGAGATTGT-TGACAAAGCAAAACAGTATACTAGAGATTT	42	482
	GENOME TYPE M		
MR25	TCTTTAAATATC--ACTTATTTAAGCT-TAAATTTCT	34	11
MR28	CTAATATAGAAGATTGAATC	20	286
MR40	CATCATTTTGATCTATTATATTCTCTGGTTCTTCTAC	37	659
	GENOME TYPE S		
MR14	AAAGTGAATGTTCTATCTCCTAACAGAAA	29	632
MR27	GAAAGTTAGATTTAG-GTGAAAG	22	200
MR46	AAACACACTTATTTAAAATTTA-ACACACTAA	31	881
MR57	TCTTAGATTT—GATAGTATTGAGATTCT	27	403

CONCLUSION

The current study focused on the identification of MR sequences in the genome of the Tomato Spotted Wilt Virus (TSWV) virus. The virus is mainly responsible for infecting tomato cultivation throughout the world. The total number of mirrors repeat patterns were identified in the genome of TSWV is 352 in which the L segment consists of most MR sequences (208), M Segment has 80 MR sequences and lastly S Segment hasthe least number of MR sequences (64). In the genome of TSWV, the MR sequences with the maximum number of nucleotides/length in the L-type segment, M- type M & S- type segments are MR51 (44), MR40 (37), and MR46 (31) respectively. Mirror repeats and their significance in the genome of organisms have been previously studied through various pieces of research. The influence of TSWV on the growth and development of a wide variety of crops is highly anticipated and we have tried to find out the existence of mirror repeat sequences (MR) in the genome of TSWV that will gives a hint about their prominent roles in the genomic biology of this virus during infection.

Author Contribution Statement

Ms. Vidhi Yadav designed the experimental setup & analyzed the *in silico* data. Ms. Priya Yadav helps in data analysis as well as drafting of the manuscript. Dr. Sandeep Yadav (corresponding author) revised & confirm the final version of the manuscript.

Conflict Of Interest: -All the authors declare no any conflict of Interest

Funding Acknowledgement: - NA

REFERENCES

1. Kashi, Y., & King, D. G. (2006). Simple sequence repeats as advantageous mutators in evolution. *TRENDS in Genetics*, *22*(5), 253-259.
2. Shapiro, J. A., & von Sternberg, R. (2005). Why repetitive DNA is essential to genome function. *Biological Reviews*, *80*(2), 227-250.
3. Rao, S. R., Trivedi, S., Emmanuel, D., Merita, K., & Hynniewta, M. (2010). DNA repetitive sequences-types, distribution and function: a review. *J Cell Mol Biol*, *7*(2), 1-11.
4. Mirkin, S. M., Lyamichev, V. I., Drushlyak, K. N., Dobrynin, V. N., Filippov, S. A., & Frank-Kamenetskii, M. D. (1987). DNA H form requires a homopurine–homopyrimidine mirror repeat. *Nature*, *330*(6147), 495-497.
5. Belotserkovskii, B. P., Veselkov, A. G., Filippov, S. A., Dobrynin, V. N., Mirkin, S. M., & Frank-Kamenetskii, M. D. (1990). Formation of intramolecular triplex in homopurine-homopyrimidine mirror repeats with point substitutions. *Nucleic acids research*, *18*(22), 6621-6624.
6. Krasilnikova, M. M., & Mirkin, S. M. (2004). Replication stalling at Friedreich's ataxia (GAA) n repeats in vivo. *Molecular and cellular biology*, *24*(6), 2286-2295.
7. Wang, G., & Vasquez, K. M. (2017). Effects of replication and transcription on DNA structure-related genetic instability. *Genes*, *8*(1), 17.
8. Belotserkovskii, B. P., Mirkin, S. M., & Hanawalt, P. C. (2013). DNA sequences that interfere with transcription: implications for genome function and stability. *Chemical reviews*, *113*(11), 8620-8637.
9. Tsompana, M., Abad, J., Purugganan, M., & Moyer, J. W. (2005). The molecular population genetics of the Tomato spotted wilt virus (TSWV) genome. *Molecular Ecology*, *14*(1), 53-66.
10. Schneider, W. L., & Roossinck, M. J. (2001). Genetic diversity in RNA virus quasispecies is controlled by host-virus interactions. *Journal of virology*, *75*(14), 6566-6571.
11. Bohlman, M. C., Morzunov, S. P., Meissner, J., Taylor, M. B., Ishibashi, K., Rowe, J., ... & St. Jeor, S. C. (2002). Analysis of hantavirus genetic diversity in Argentina: S segment-derived phylogeny. *Journal of Virology*, *76*(8), 3765-3773.
12. Gupta, R., Kwon, S. Y., & Kim, S. T. (2018). An insight into the tomato spotted wilt virus (TSWV), tomato and thrips interaction. *Plant Biotechnology Reports*, *12*, 157-163.
13. Maris, P. C., Joosten, N. N., Goldbach, R. W., & Peters, D. (2004). Tomato spotted wilt virus infection improves host suitability for its vector Frankliniella occidentalis. *Phytopathology*, *94*(7), 706-711.
14. Bautista, R. C., Mau, R. F., Cho, J. J., & Custer, D. M. (1995). Potential of tomato spotted wilt tospovirus plant hosts in Hawaii as virus reservoirs for transmission by Franklinieila occidentalis (Thysanoptera: Thripidae). *Phytopathology*, *15*, 32.
15. Sevik, M. A., & Arli-Sokmen, M. (2012). Estimation of the effect of Tomato spotted wilt virus (TSWV) infection on some yield components of tomato. *Phytoparasitica*, *40*, 87-93.
16. Roggero, P., Masenga, V., & Tavella, L. (2002). Field isolates of Tomato spotted wilt virus overcoming resistance in pepper and their spread to other hosts in Italy. *Plant Disease*, *86*(9), 950-954.
17. Adkins, S. (2000). Tomato spotted wilt virus—positive steps towards negative success. *Molecular plant pathology*, *1*(3), 151-157.
18. Best, R. J. (1968). Tomato spotted wilt virus. *Advances in virus research*, *13*, 65-146.
19. de Oliveira Resende, R., de Haan, P., de Avila, A. C., Kitajima, E. W., Kormelink, R., Goldbach, R., & Peters, D. (1991). Generation of envelope and defective interfering RNA mutants of tomato spotted wilt virus by mechanical passage. *Journal of General Virology*, *72*(10), 2375-2383.
20. Kormelink, R., Kitajima, E. W., De Haan, P., Zuidema, D., Peters, D., & Goldbach, R. (1991). The nonstructural protein (NSs) encoded by the ambisense S RNA segment of tomao spotted wilt virus is associated with fibrous structures in infected plant cells. *Virology*, *181*(2), 459-468.
21. Qiu, W. P., Geske, S. M., Hickey, C. M., & Moyer, J. W. (1998). Tomato Spotted WiltTospovirusGenome Reassortment and Genome Segment-Specific Adaptation. *Virology*, *244*(1), 186-194.
22. Bandla, M. D., Campbell, L. R., Ullman, D. E., & Sherwood, J. L. (1998). Interaction of tomato spotted wilt tospovirus (TSWV) glycoproteins with a thrips midgut protein, a potential cellular receptor for TSWV. *Phytopathology*, *88*(2), 98-104.
23. Whitfield, A. E., Kumar, N. K. K., Rotenberg, D., Ullman, D. E., Wyman, E. A., Zietlow, C., ... & German, T. L. (2008). A soluble form of the Tomato spotted wilt virus (TSWV) glycoprotein GN (GN-S) inhibits transmission of TSWV by Frankliniella occidentalis. *Phytopathology*, *98*(1), 45-50.
24. Kormelink, R., Storms, M., Van Lent, J., Peters, D., & Goldbach, R. (1994). Expression and subcellular location of the NSM protein of tomato spotted wilt virus (TSWV), a putative viral movement protein. *Virology*, *200*(1), 56-65.
25. Whitfield, A. E., Ullman, D. E., & German, T. L. (2004). Expression and characterization of a soluble form of tomato spotted wilt virus glycoprotein GN. *Journal of virology*, *78*(23), 13197-13206.
26. Margaria, P., Bosco, L., Vallino, M., Ciuffo, M., Mautino, G. C., Tavella, L., & Turina, M. (2014). The NSs protein of tomato spotted wilt virus is required for persistent infection and transmission by Frankliniella occidentalis. *Journal of Virology*, *88*(10), 5788-5802.
27. Shalileh, S., Ogada, P. A., Moualeu, D. P., & Poehling, H. M. (2016). Manipulation of Frankliniella occidentalis (Thysanoptera: Thripidae) by Tomato Spotted Wilt Virus (Tospovirus) via the host plant nutrients to enhance its transmission and spread. *Environmental Entomology*, *45*(5), 1235-1242.
28. Yadav, P., Kumari, J., Yadav, P., Yadav, R., Yadav, S., Sharma, D., ... & Yadav, S. (2023). Identification of Mirror Repeats in Viral Genomes using FPCB Analysis. *bioRxiv*, 2023-04.
29. Yadav, S., Yadav, U., & Sharma, D. C. (2022). In silico approach for the identification of mirror repeats in selected operon genes of Escherichia coli strain K-12 substrain MG1655. *Biomedical and Biotechnology Research Journal (BBRJ)*, *6*(1), 93-97.
30. Yadav, S., Yadav, U., & Sharma, D. C. (2022). In-silico evaluation of 'mirror repeats' in HIV genome.(2021). *Int. J. Life Sci*, *11*(5), 81-87.
31. Yadav, M., Sehrawat, B., Yadav, P., Kumari, J., Yadav, P., Yadav, S., ... & Yadav, S. (2023). In-Silico analysis of Human Insulin Gene for Mirror repeat sequences. *bioRxiv*, 2023-07.
32. Sehrawat, B., Yadav, P., Sarjeet, M., Yadav, S., Yadav, V., Goyal, N., ... & Yadav, S. (2024). In-silico Evaluation of Mirror Repeats in Some Selected Genes of Candida albicans. *Journal of Biology and Nature*, *16*(1), 39-52.