Research Article

Indian Journal of Science and Research

Kumar V & Maurya S.K., Ind. J. Sci. Res. 2026, 6(2), 124-134

ISSN 2583 – 2913

COMPARATIVE ANALYSIS OF CONFORMATIONAL, ARCHITECTURAL AND THERMODYNAMIC SIGNALS IN UPSTREAM FLANKS OF MPING ELEMENTS INSERTIONS IN TWO SUBSP. OF CULTIVATED RICE

Virendra Kumar[1] and Sanjeev Kumar Maurya[1,2]*

Abstract: Transposable elements rewire gene regulation in eukaryotes. The proliferation and elimination of mobile genetic elements are well known to play a significant role in genome size variation within a single genus. The insertion of Transposable elements occurred millions of years ago, and active element copies are very low; therefore, it is exceedingly difficult to obtain enough data to test this hypothesis rigorously. Though a class II (Transposons) rice endogenous transposable element, mPING, is still active, its transposition is found to be highly significant under stress conditions, which would be key to answering this question. The bio-computational approach was used to analyse the dispersion of mPING in two cultivated rice subspecies, indica and japonica, which showed 14 and 52 insertions, respectively. The 50 base pairs 5' strand upstream flanking sequences were analysed for structural and physical properties. The comparative analysis revealed a significant variation, which further indicated hidden physical/structural genomic sequence specificity.

Keywords: Transposable elements, Retrotransposon, Bio-computational approach, Structural, Physical.

Introduction: Rice is one of the most important food crops cultivated worldwide. As a model organism, it also promotes our understanding of cereal genetics and functional genomics. The genome size variation in the genus *Oryza* is more than 4-fold and ranges from 295 Mbp in *Oryza brachyantha* to 1283 Mbp in the polyploid *Oryza ridleyi*. Transposable elements (TEs) are the hidden genomic segments named after their characteristics that they transpose, change locations in a genome and increase copy numbers[1]. TEs are divided into two groups according to the intermediate transposition mechanism. Class I elements transpose with the help of an RNA intermediate so-called copy and paste and Class II elements, DNA transposons, whose transposition process is called the cut-and-paste mechanism. Both classes have autonomous and non-autonomous elements. The transposition mechanism and sequence characteristics of elements are further used to group them into transposon superfamilies. The natural mobility of transposable elements is vital for numerous aspects, including evolution; they contribute to the mutational repertoire of the host genome, provide crucial mechanisms in certain circumstances of adoption, and influence the functional activity of the host genome[2]. The role of mPING in stress is found to be very significant and inserted roughly 20 new transpositions per plant per generation in the *japonica* rice cultivar 'Gimbozu' under ordinary cultivation conditions[3-4]. However, under various stress conditions, small populations may obtain thousands of new insertions[5] in the rice genome, and the mPing elements were distributed in the 1-bp to 500-bp upstream regions of genes[6]. The

***Corresponding author**

[1]Department of Plant Science, M. J. P. Rohilkhand University, Bareilly, India – 243006

[2] Department of Botany, Constituent Government College, Puranpur, Pilibhit, India – 262122

E-mail: mauryasanjeev15[at]gmail.com

DOI: https://doi.org/10.5281/zenodo.19511040

Article recived on: 22 March 2026

Published on web: 10 April 2026, www.ijsronline.org

ubiquitous nature and abundance of mPING in the rice genome make it an important candidate responsible for the diverse set of genomic changes in the organism [5].

The mPing (miniature Ping) is a member of the miniature inverted-repeat transposable element (MITE) superfamily of class II elements indigenous to rice and still active in its transposition[7]. The mPING a 430-bp long non-autonomous TE (Unable to encode TPase) derived from the deletions of PING elements, was identified in 2003. These mPING elements are dependent for their transposition on a member of the PIF/Harbinger superfamily of class 2 elements (PING)[8]. Generally, transposable elements are silenced by the host surveillance mechanism[9], however, in the case of mPING elements, the surveillance mechanism of the host is unable to silence them, therefore they remain active and increase their copy number in the rice genome without killing them. The previous shreds of evidence suggest the transposition of mPING elements massively catalysed by TPase encoded by the Ping element. The identical structure of the subterminal regions of the Ping element and structural divergence of TPase of other potentially active TEs have remained mPING active[1]. Most transposase proteins (TPase) of class II elements are made up of a conserved catalytic domain (DDE motif) and a DNA-binding domain, whereas in the case of the PING element, these domains are encoded separately by two ORFs, ORF1 and ORF2, respectively a Myb-like DNA-binding protein, and a TPase without a DNA-binding domain. Myb-like DNA-binding protein directly binds to the subterminal regions of the transposon to recruit the transposase[10].

In this research work, the bio-computational approach was used for the comprehensive study of the mPING element distribution in the rice genome. The upstream sequences of the insertion sites were also used to analyse the conformational, architectural, and thermodynamic signals.

Materials and methods: The sequences of genus *Oryza* were collected from the National Centre for Biotechnology Information (http://www.ncbi.nlm.nih.gov/) and Ensembl (http://www.ensemblgenomes.org). The mPING element is obtained from Repbase (http://www.girinst.org/repbase). The dispersion of the previously characterised mPING repetitive elements in the genome was searched by Stand-alone BLAT v472 (BLAST-Like Alignment Tool) using parameters -stepSize=5 -repMatch=2253 -minIdentity=95[11]. The 100% query coverage from BLAT output was filtered by the pslCDnaFilter tool using the parameter -minCover=1.00. The results of BLAT were processed and assembled for significant information with the help of bash one-liners (awk 'NR > 5 {print $10, $14, $11, $12, $13, $9, $15, $16, $17}' OFS='\t' output.psl > filtered_coords.tsv). The output data fed into the TBtools-II to visualise the distribution of mPING elements on the genomes (Circos diagram)[12]. The In-house bioperl script was used for the extraction of the 50 bp 5’ upstream flanking region of mPING insertion. The upstream sequences were used to develop the web logo by using version 3 of the WebLogo tool[13]. The biophysical properties of these sequences were also analysed and compared by using the stand-alone tools package DnaFVP (version 0.9.5) (http://dnafvp.sourceforge.net/). The mean values of biophysical properties of both species mPING upstream sequence used to prepare the line charts to show comparative variation in values across DNA sequences length by LibreOffice Calc. To identify divergence between the biophysical properties, paired t-tests across 20 DNA structural parameters were performed. To account for the inflation of Type I errors associated with multiple hypothesis testing, p-values were transformed into q-values using the Benjamini-Hochberg False Discovery Rate (FDR) procedure. The p and q-values were used to develop the heatmap by using software TBtools-II.

Results and Discussion: In the present study, ten species of the genus *Oryza* were analysed for the genome-wide dispersion of mPING elements (Table 1). The analysis revealed that out of ten, only three species (*Oryza rufipogon*, *Oryza nivara*, and *Oryza sativa*) show the presence of the mPING elements. The species *Oryza rufipogon* and *Oryza nivara*, respectively, show 6 and 4 insertions of the mPING elements. However, cultivated rice subspecies *japonica and indica* show 52 and 14 insertions, respectively (Table 1). The number of mPING-inserted copies in *Oryza rufipogon* and *Oryza nivara* is negligible. Therefore, the analysis was focused on the two subspecies of *Oryza sativa*. The genome-wide dispersion of mPING elements in both *Oryza sativa subsp. japonica* and *Oryza sativa subsp. indica* is shown with the help of the circos tool (Figure 1). The results of genome-wide dispersion analysis in *Oryza sativa subsp. japonica* given in figure 1 depict ununiform distribution of 52 mPING elements throughout the genome. The *Oryza sativa subsp. japonica* chromosome number 1 to 12, respectively, have 5, 7, 10, 5, 3, 4, 1, 6, 2, 2, 2 and 5 insertions. However, in *Oryza sativa subsp. indica* shows 3, 1, 1, 2, 2, 1, 1, 2 and 1 insertion copies of mPING elements respectively on chromosomes 1, 2, 3, 4, 5, 6, 9, 11 and 12. The analysis revealed that the insertion of mPING element copies is highly variable in two subspecies of *Oryza sativa.*

Table 1 List of Rice Species along with their genome size and insertions of mPING elements.

S. No	Species	Genome Size	mPING element
1	*Oryza barthii*	~355	_
2	*Oryza brachyantha*	~295	_
3	*Oryza glaberrima*	~330	_
4	*Oryza glumaepatula*	~420	_
5	*Oryza longistaminata*	~400	_
6	*Oryza meridionalis*	~381	_
7	*Oryza nivara*	~384	4
8	*Oryza punctata*	~441	_
9	*Oryza rufipogon*	~384	6
10	*Oryza sativa japonica*	~422	52
11	*Oryza sativa indica*	~420	14

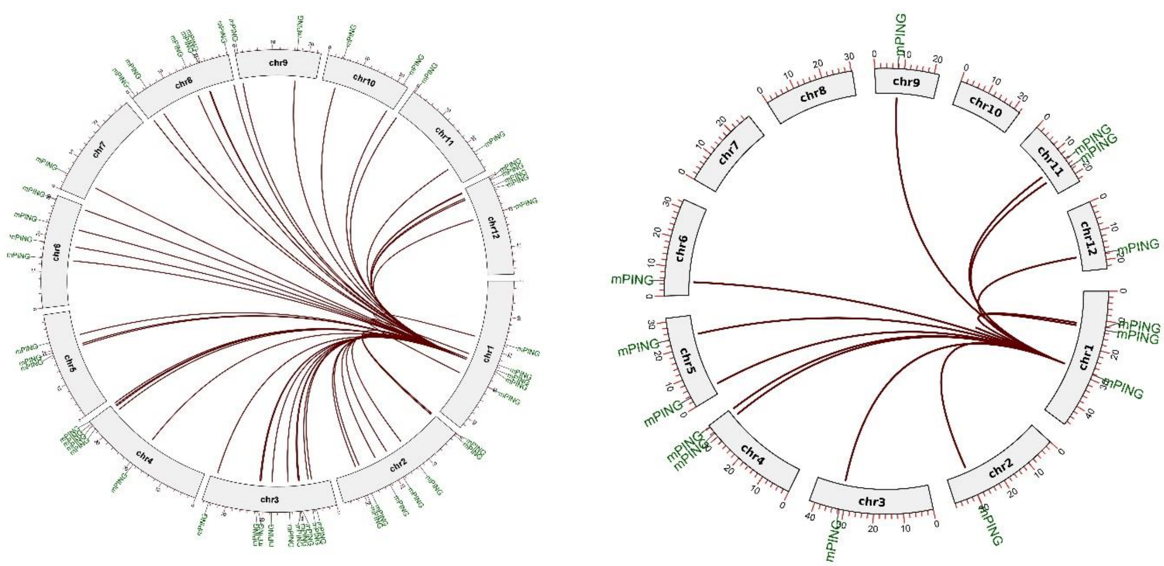

Figure 1 Circos diagram showing the insertion of mPING element copies in two subspecies of *Oryza sativa.* (A) *Oryza sativa japonica* (B) *Oryza sativa indica.*

Besides the distribution of mPING elements in the rice genome, it is not well-known whether the target-specific loci or their preferences to integrate within the genome in each species of rice are unique or different. Since transposable elements require precise conformational, architectural, and thermodynamic signals to integrate themselves at a specific position in the genome and the upstream flanking sequence of TEs insertions shows suitable nucleotide biophysical properties to favour transposition[2], we set out to document the comparative visualisation of repeats upstream sequences conformational, architectural, and thermodynamic signals. The combined view of nucleotide sequence structural features indicates the existence of a hidden physical/structural code for better understanding and interpretation of the underlying genomic sequences' functionality. The Weblogo was developed from 5' upstream of mPING elements from both *Oryza sativa subsp. indica and Oryza sativa subsp. japonica* shown in figure 2.

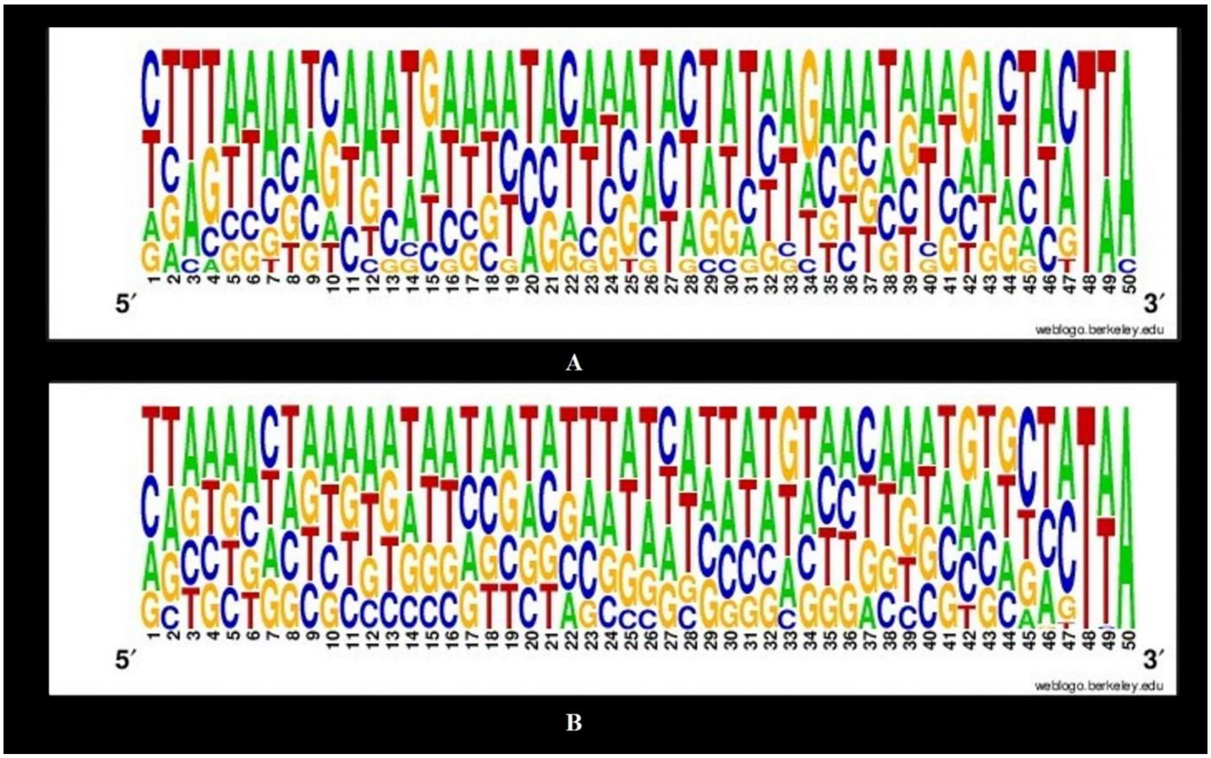

Figure 2 Weblogo developed from 5' upstream of mPING elements insertions in two subspecies of *Oryza sativa.* (A)*Oryza sativa subsp. japonica.* (B) *Oryza sativa subsp. indica.*

In the case of *Oryza sativa subsp. indica*, we found 5' AT-rich flanking upstream sequences of mPING elements. Similarly, in *Oryza sativa subsp. japonica* 5' flanking regions of mPING insertion are AT-rich. Thus, we can say that AT-rich regions can serve as signals for the insertion of mPING elements into the host genome (Figures 1 and 2). The 5' upstream flanking regions of mPING elements from both subsp. of *Oryza sativa* were analysed. The comparative and/or contrasting visualisation of conformational, architectural, thermodynamic signals, i.e. A-phylicity[14], Z-DNA stabilising energy[15], duplex stability disrupt energy[16], and duplex stability free energy shown in figure 3. The physical and chemical landscape of DNA is governed by a sophisticated array of thermodynamic, architectural, and conformational signals that extend far beyond simple sequence information. At the thermodynamic level, signals such as A-phylicity and Z-DNA stabilising energy dictate the energetic cost of transitioning from the standard B-form to the more compact A-form or the left-handed Z-form, respectively; these transitions are often biologically triggered by hydration levels or by torsional stress induced by RNA polymerase. Similarly, duplex stability free energy and disrupt energy quantify the "melting" potential of the helix, where lower stability often correlates with genomic regions that must be frequently opened, such as promoters. Base stacking energy[17], DNA denaturation[18], B-DNA twist[19], and protein-DNA twist[20] are given in figure 4. The Propeller twist[21], protein-induced deformability, DNA bending stiffness[22], Bendability[19], nucleosome preference[23], and GC content are given in figure 5. Base stacking energy, the primary stabiliser of the double helix, alongside DNA denaturation profiles, provides a map of the internal cohesion of the strands.

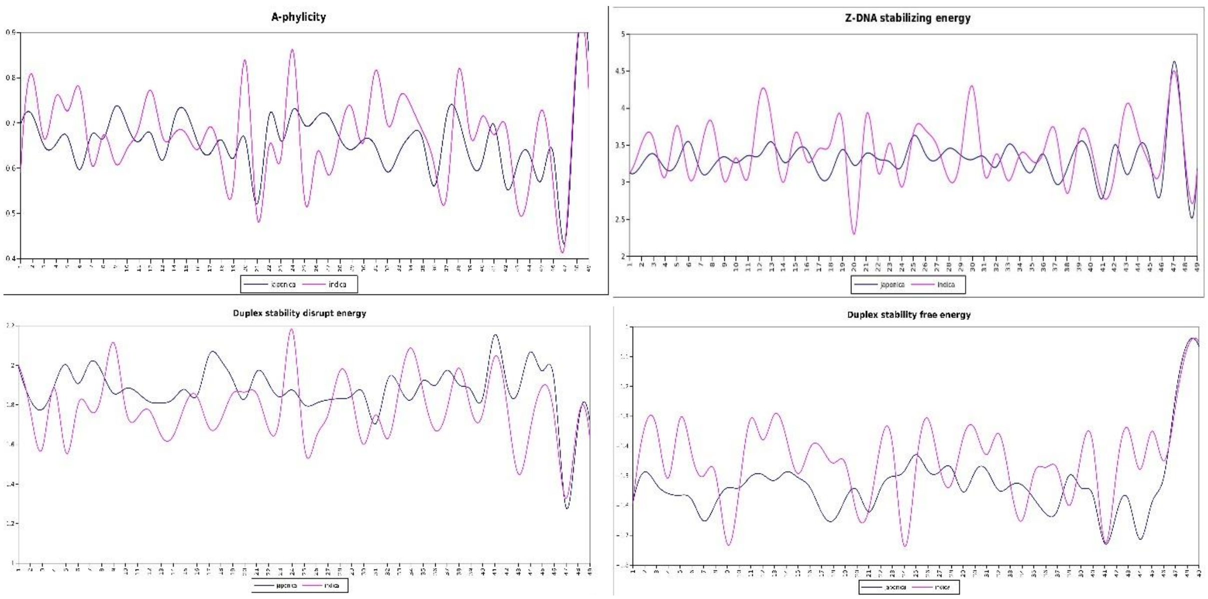

Figure 3 The comparative visualisation of A-phylicity, Z-DNA stabilising energy, duplex stability disrupts energy and duplex stability free energy shown in 5' upstream flanking regions of mPING elements of both subsp. of *Oryza sativa.*

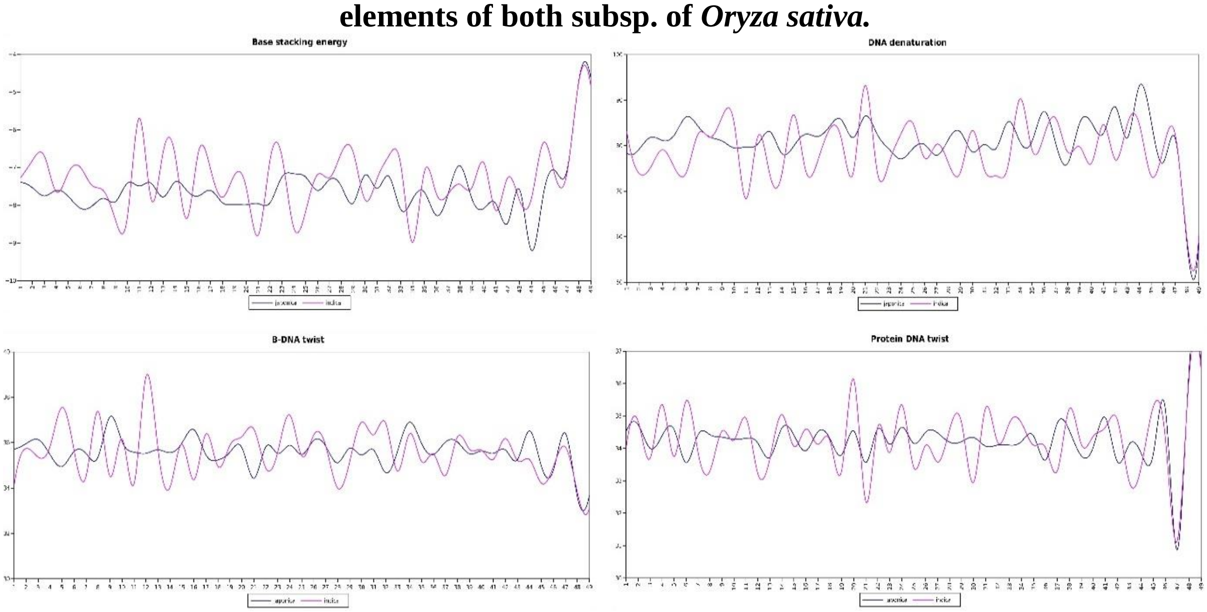

Figure 4 The comparative visualisation of Base stacking energy, DNA denaturation, B-DNA twist and Protein DNA twist shown in 5' upstream flanking regions of mPING elements of both subsp. of *Oryza sativa.*

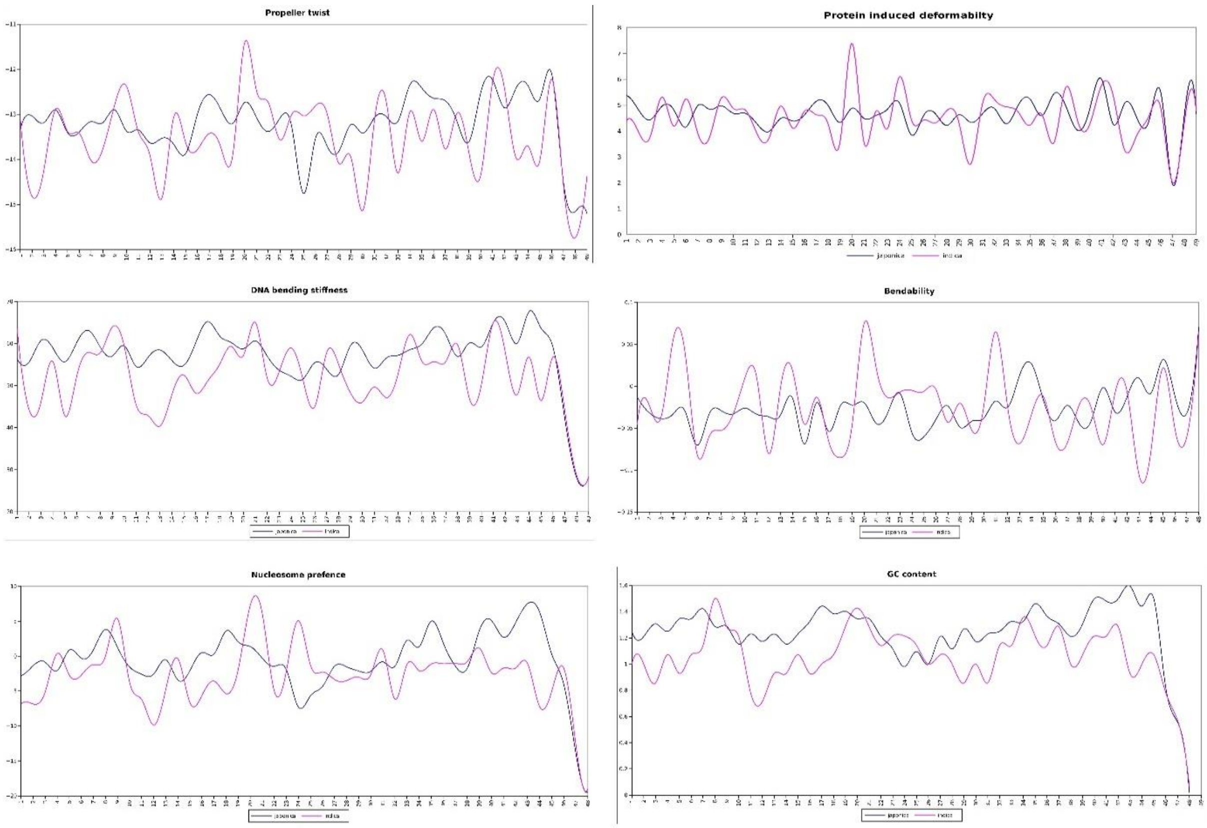

Figure 5 The comparative visualisation of Propeller twist, Protein induced deformability, DNA bending stiffness, Bendability, nucleosome preference, and GC content shown in 5' upstream flanking regions of mPING elements of both subsp. of *Oryza sativa.*

Architecturally, DNA exhibits sequence-dependent geometry characterized by B-DNA twist, Protein-DNA twist, and Propeller twist, which describe the angular orientation and rigidity of base pairs. These geometric signals, combined with protein-induced deformability, bending stiffness, and bendability, determine how easily a DNA segment can be wrapped or distorted by binding proteins. For instance, high nucleosome preference is often found in sequences with periodic bendability that allows the DNA to coil tightly around histone octamers. The Major-groove Hydrogen bonding Pattern (Hbondw1, Hbondw2, Hbondw3, Hbondw3p, Hbondw2p, Hbondw1p), of the mPING element upstream flanks in both subsp. respectively given in figure 6. Crucially, the Major-groove Hydrogen bonding patterns (Hbondw1 through Hbondw3p) represent the "direct readout" mechanism for proteins. By sensing the specific spatial arrangement of hydrogen bond donors and acceptors across the major groove's width, transcription factors can distinguish between base pairs without unzipping the helix, allowing for high-affinity, sequence-specific recognition that is further stabilised by the local GC content and structural flexibility.

The heatmap of p-value and q-value of these physiochemical properties is also given in figure 7 and 8. The p and q-values talk about the compatibility of data with the null hypothesis. The smaller p-value indicates that the data do not develop randomly. However, the concept of q-value given by John Storey is the false discovery rate (FDR) analogue of the p-value. The q-value for an observed p-value is the minimum FDR at which the feature can be considered significant. The 20 distinct DNA structural features across 50-nucleotide-long upstream flanking mPING retrotransposon sequences were evaluated by dual-

threshold ($p < 0.05$ and 0.25) statistical analysis to identify regions of divergence between the two subspecies. The $p < 0.05$ to identify highly significant structural differences and $p < 0.25$ to capture broader trends of divergence.
Analysis at 95% confidence level ($p < 0.05$) shows the structural context of mPing insertion sites remains largely conserved. While the tri-nucleotide GC content at positions 12 and 43, DNA bendability at position 43 and stacking energy at 45 is significant divergent. However, the analysis at 75% confidence level ($p < 0.25$) reveals broader structural trends and regional "hotspots" that may signify subtle differences in how mPing interacts with the genome in *indica* and *japonica* subspecies. To distinguish primary sequence-driven differences and broader structural trends between japonica and indica subspecies, a comparative t-test significance for 20 DNA structural parameters of the 50 bp upstream of mPing insertion sites at 95% ($p < 0.05$) and 75% ($p < 0.25$) confidence intervals, as given in table 2.

The structural analysis confirms that the primary integration site for mPing is highly conserved, while the flanking upstream sequences contain significant structural "fingerprints" that distinguish the two subspecies. The notable divergence in DNA bendability and nucleosome preference suggests that the subspecies-specific genomic environment might influence mPing activity by altering the local physical accessibility of the DNA to the transposition machinery. These findings provide a structural basis for understanding the variations in mPing distribution and mobilization observed between the *indica* and *japonica* subspecies of rice.

Table 2 Summary of Statistically Significant DNA Structural Divergences in the 50 bp Upstream Region of mPing Insertion Sites. Comparison between indica and japonica rice subspecies using t-tests. Significant positions are identified at both 95% ($p < 0.05$) and 75% ($p < 0.25$) confidence levels, with nucleotide positions indexed from 1 to 50.

Feature	Significant Positions 95% confidence	Count 95% confidence	Significant Positions 75% confidence	Count 75% confidence
Hbondw1	0	0	43	1
Hbondw2	0	0	3, 45	2
Hbondw3	0	0	7, 43	2
Hbondw3p	0	0	3, 45	2
Hbondw2p	0	0	7, 43	2
Hbondw1p	0	0	43	1
A-phylicity	0	0	6, 12, 20, 26	4
ZDNA stab energy	0	0	12, 20, 30	3
Disrupt energy	0	0	25, 26	2
Free energy	0	0	12, 13, 26, 43, 45	5
Stacking energy	45	1	11, 32, 45	3
DNA denature	0	0	6, 20, 32, 45	4

B-DNA twist Ohler	0	0	5, 9, 12	3
Protein DNA twist	0	0	6, 20, 45	3
Propeller twist	0	0	20, 25, 43	3
P induced deform	0	0	20, 30, 37, 44	4
DNA bend stiff	0	0	12, 13, 26, 45	4
Bendability	43	1	4, 8, 10, 11, 12, 20, 31, 33, 43	9
Nuclpref	0	0	11, 12, 20, 24, 29, 42, 43	7
Tri GC content	12, 43	2	11, 12, 29, 43	4

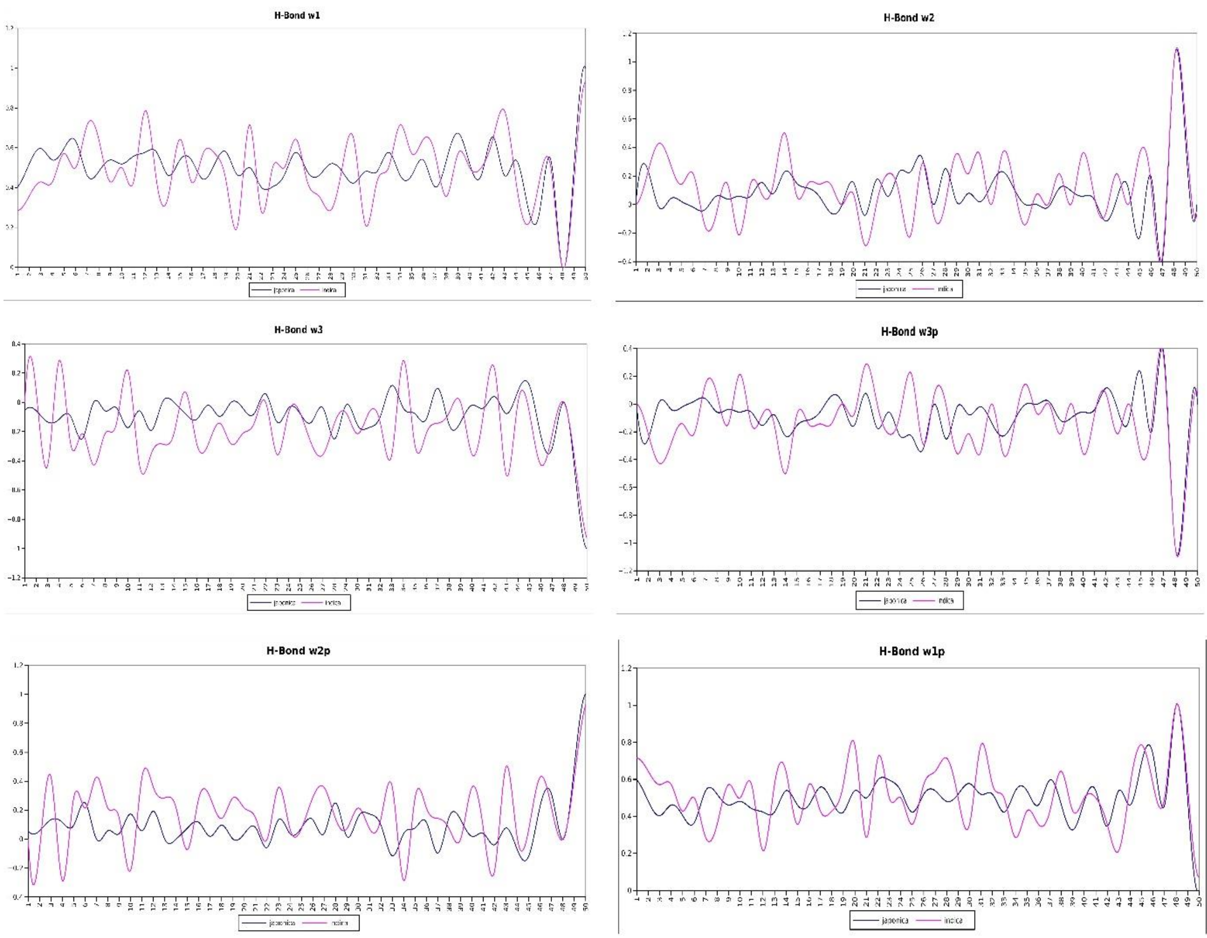

Figure 6 The comparative visualisation of Major-groove Hydrogen bonding Pattern (Hbondw1, Hbondw2, Hbondw3, Hbondw3p, Hbondw2p, Hbondw1p) shown in 5' upstream flanking regions of mPING elements of both subsp. of *Oryza sativa.*

Kumar V & Maurya S.K., Ind. J. Sci. Res. 2026, 6(2), 124-134

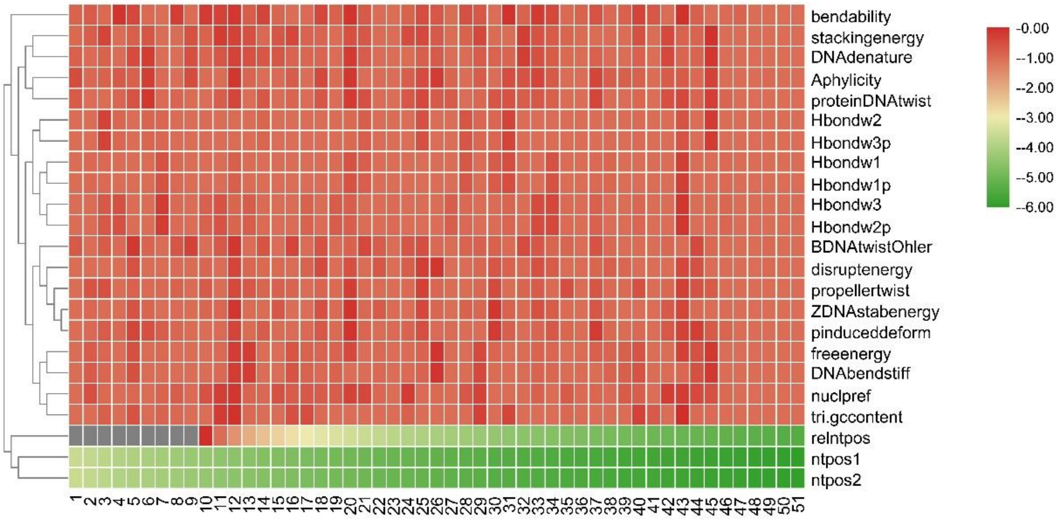

Figure 7 The heatmap based on p-value for all above discussed physiochemical properties of 5' upstream flanking regions of mPING elements of both subsp. of *Oryza sativa*.

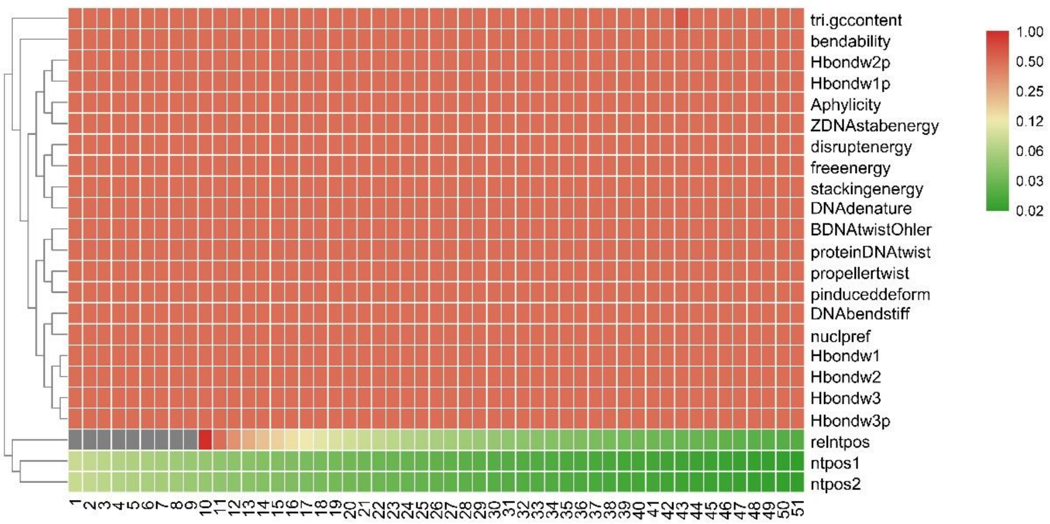

Figure 8 The heatmap based on q-value for all above discussed physiochemical properties of 5' upstream flanking regions of mPING elements of both subsp. of *Oryza sativa*.

Conclusion: In the present study, analysis revealed the invariable distribution of mPING repetitive elements in the genus *Oryza*. As we know, a variety of structural/physical features are associated with the functionality of genome regulatory regions. Two subspecies of the traditionally cultivated rice species, that is *Oryza sativa* subsp. *indica* and *Oryza sativa* subsp. *japonica* respectively, shows 14 and 52 insertions. Previously, we showed that upstream flanking sequences of TEs exhibit certain biophysical features that facilitate transposition[2]. However, the comparative analysis of the mPING retrotransposon upstream flanking sequence from both subsp. of *Oryza sativa* showed a significant variation between the structural and physical properties. Thus, this research work revealed that each TEs in each crop species uses its own particular nucleotide biophysical properties to transpose under different conditions. Therefore, a wide work on these lines is required rather than formulating a model for all organisms. Thus, the big challenge is to predict which TEs use what properties to transpose in which crop species, and under which conditions.

Author Contributions: Virendra Kumar: Conceptualisation, Perform Software Analysis, and Writing – original draft. Sanjeev Kumar Maurya: Writing – review, editing and Data interpretation.

Acknowledgments: Authors are thankul to Prof. A. K. Jaitley for his guidance and the Bio-Linux community for the computational environment.

Funding: This research received no funding and was conducted using personal computational resources.

Conflict of Interest: The authors declare that there is no conflict of interest.

References

1. Naito. K., Monden. Y., Yasuda. K., Saito. H. and Okumoto. Y. 2014. *MPing: The bursting transposon*. Breed Sci. 64(2):109-114.
2. Khan. M. F., Shrivastava. K., Sinha. R., Kumar. V. and Jaitly. A.K. 2016. *Genome wide occurrence and insertion preferences of INGI/RIME and SLACS CRE transposable elements in Trypanosoma brucei*. Mobile Genetic Elements. 6(6).
3. Naito. K., Cho. E. Y., Yang. G., Campbell. M. A., Yano. K., Okumoto. Y., Tanisaka. T. and Wessler. S. R. 2006. *Dramatic amplification of a rice transposable element during recent domestication.* Proc. Natl. Acad. Sci. USA. 03:17620-17625.
4. Naito. K., Zhang. F., Tsukiyama. T., Saito. H., Hancock. C. N., Richardson. A. O., Okumoto. Y., Tanisaka. T. and Wessler. S. R. 2009. *Unexpected consequences of a sudden and massive transposon amplification on rice gene expression*. Nature. 461:1130-1134.
5. Teramoto. S., Tsukiyama. T., Okumoto. Y. and Tanisaka. T. 2014. *Early Embryogenesis Specific Expression of the Rice Transposon Ping Enhances Amplification of the MITE mPing.* LoS Genet. 10(6).
6. Yasuda. K., Ito. M., Sugita. T., Tsukiyama. T., Saito. H., Naito. K., Teraishi. M., Tanisaka. T. and Okumoto. Y. 2013. *Utilization of transposable element mPing as a novel genetic tool for modification of the stress response in rice.* Mol Breed. 32: 505-516.
7. Horibata. K. Y. and Kato. T. 2015. *Distorted genetic segregation of the transposon mPing at the long arm of chromosome 12 in rice.* Breeding Science. 65: 340-344.
8. Kum. R., Tsukiyama. T., Inagaki. H. et.al. 2015. *The active miniature inverted-repeat transposable element mPing post-transcriptionally produces new transcriptional variants in the rice genome.* Mol. Breeding. 35:159.
9. Kumar. V., Chandra. S. and Jaitly. A. K. 2019. *DNA Methylation: A Stabilizing and Regulatory Mechanism of Plant Genome.* G- J Environ Sci Technol. 7(3): 24-31.
10. Lin. X., Long. L., Shan. X., Zhang. S., Shen. S. and Liu. B. 2006. *In planta mobilization of mPing and its putative autonomous element Pong in rice by hydrostatic pressurization.* J Exp Bot. 57(10): 2313-2323.
11. Kent. W. J. 2002. BLAT the BLAST-like alignment tool. Genome Res. 12: 656-664.
12. Chen. C., Wu. Y., Li. J., Wang. X., Zeng. Z., Xu. J., Liu. Y., Feng. J., Chen. H., He. Y. and Xia. R. 2023. *TBtools-II: A "one for all, all for one" bioinformatics platform for biological big-data mining*. Molecular Plant, *16*(11), 1733–1742.
13. Crooks. G. E., Hon. G., Chandonia. J. M. And Brenner. S. E. 2004. *WebLogo: a sequence logo generator*. Genome Res. 14(6): 1188-1190.
14. Ivanov. V. and Minchenkova. L. 1994. *The A-form of DNA: In search of the biological role*. Mol Biol. 28(6): 1258.
15. Ornstein. R. L., Rein. R., Breen. D. and Macelroy. R. 2004. *An optimized potential function for the calculation of nucleic acid interaction energies I Base stacking.* Biopolymers, 17(10): 2341-2360.
16. Breslauer. K., Frank. R., Blocker. H. and Marky. L. 1986. *Predicting DNA duplex stability from the base sequence*. Proc Natl Acad Sci USA. 83(11): 3746.

17. Blake. R. D. and Delcourt. S. G. 1998. *Thermal stability of DNA.* Nucleic Acids Res. 26(14): 3323-3332.
18. Ho. P. S., Zhou. G. W. and Clark. L. B. 1990. *Polarized electronic spectra of Z-DNA single crystals.* Biopolymers. 30(1-2): 151–163.
19. Brukner. R., S´anchez. D., Suck. and Pongor. S. 1995. *Trinucleotide models for DNA bending propensity: Comparison of models based on DNase I digestion and nucleosome packaging data.* Journal of Biomolecular Structure & Dynamics. 13(2): 309-317.
20. Gorin. A. A., Zhurkin. V.B. and Olson. W. K. 1995. *B-DNA twisting correlates with base-pair morphology*. J Mol Biol. 247(1): 34-48.
21. Olson. W. K., Gorin. A. A., Lu. X. J., Hock. L. M. and Zhurkin. V. B. 1998. *DNA sequence-dependent deformability deduced from protein-DNA crystal complexes.* Proc Natl Acad Sci USA. 95(19): 11163-11168.
22. Sivolob. A. V. and Khrapunov. S. N. 1995. *Translational positioning of nucleosomes on DNA: The role of sequence-dependent isotropic DNA bending stiffness.* J Mol Biol. 247(5): 918-931.
23. Satchwell. S. C., Drew. H. R. and Travers. A. A. 1986. *Sequence periodicities in chicken nucleosome core DNA.* J Mol Biol. 191(4): 659-675.