

Comprehensive genome-wide analysis of transposable elements in 30 different cultivars and wild rice species; reveals LTR elements are major repeat constitute

Comprehensive genome-wide analysis of transposable elements for its association with disease resistance genes in rice (*Oryza sativa* L)

Ravindra Raut¹

@Ravindra_Raut

raviraut206@gmail.com

Subhankar Roy Barman¹ Malali Gowda²

¹ Department of Biotechnology, National Institute of Technology Durgapur, India

² Center for Functional Genomics and Bioinformatics, The University of Health Sciences and Technology Bengaluru, India

1 Introduction

- Transposable elements (TE) or transposons are mobile DNA elements which exist in almost all eukaryotic genomes [1].
- Firstly these elements are identified and described by geneticist Barbara McClintock.
- TEs are categorized based on the intermediate substrate producing insertions (DNA/RNA) [2]
- Class I (retrotransposons):** SINEs [ALUs, MIRs, Penelope], LINEs [L1, L2, L3/CR1, R1/LOA/Jockey], LTRs [ERVL, Ty1/Copia, Gypsy/DIRS1]
- Class II (DNA transposons):** hAT, En-Spm, MuDR, Tourist/Harbinger, Mirage, P-element, Transib.

2 Objectives

- To study the Transposable Elements (TEs) in different rice genomes: (*Oryza sativa* Japonica, *Oryza sativa* indica and wild rice species).
- To study association of TEs and Plant Resistance (R)-genes.

3 Methods

The de novo repeat prediction performed using **RepeatModeler 1.0.11**. The program runs with RM-BLAST (NCBI) database as input for repeat modelling. At the core of RepeatModeler, there are programs called RECON, RepeatScout, Tandem Repeat Finder. RepeatModeler follows the nomenclature code of Dfam to classify TEs. The library file generated by RepeatModeler was used to estimate the coverage of different classes of repeats by the **RepeatMasker 4.1.0**.

4 Results

The rice genome assemblies are investigated to detect and identify the presence of transposable elements include previously published genomes; we selected thirty rice assemblies based on genome coverage that includes *Oryza sativa* Japonica Group, Indica Group, African wild rice, Australian wild rice, South American rice wild species and African outgroup species *Leersia perrieri* (Launert or Cutgrass) Figure 4.1.

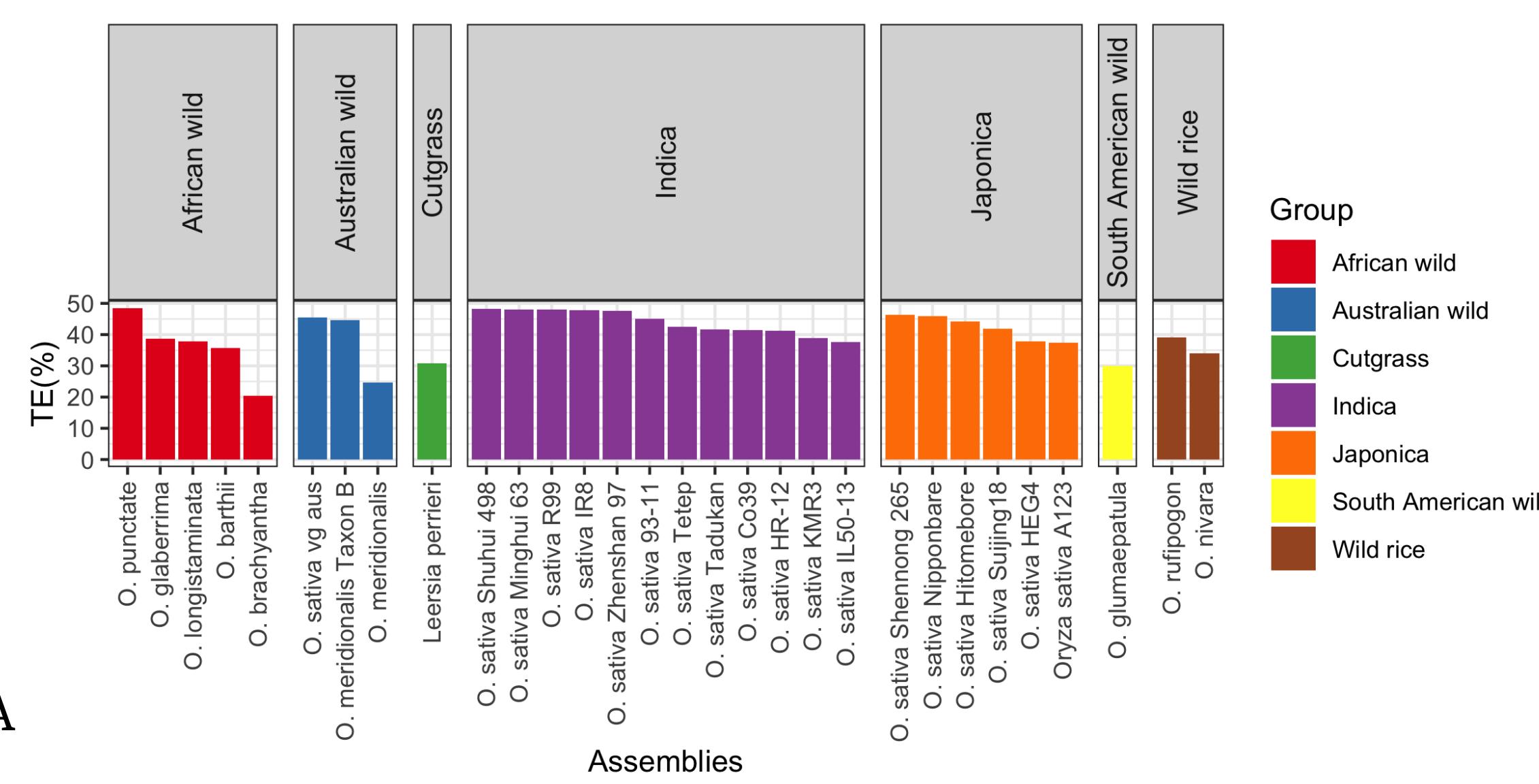


Figure 4.1: The percentage of transposable elements in *oryza* genome and classification of assemblies based on the genomic groups.

4.1 Transposable Elements Classes in rice genome assemblies

This study revealed unique pattern of TE Classes in rice genome assemblies (Figure 4.2).

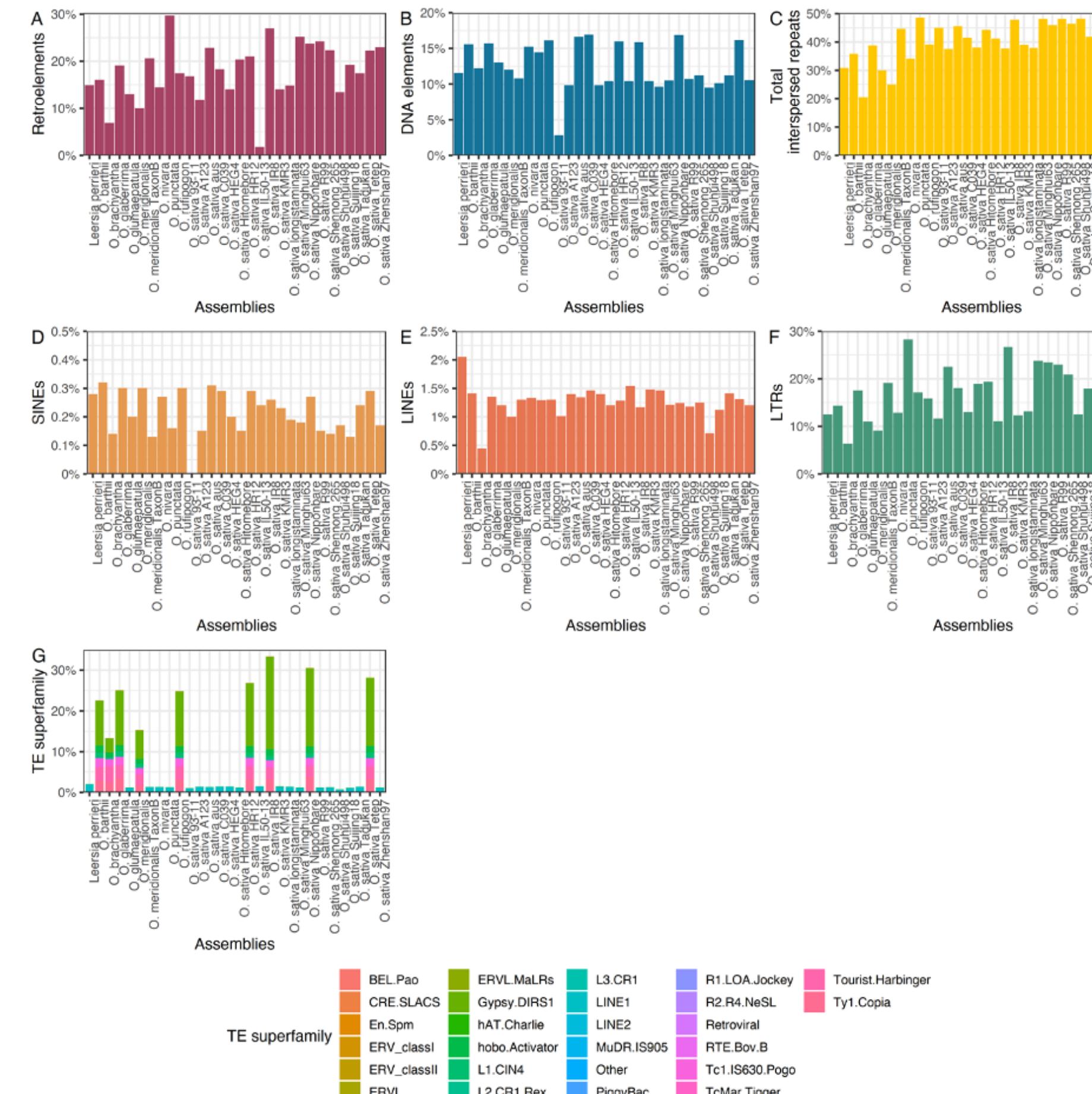


Figure 4.2: Analysis of rice genome assemblies for (A) Retroelements, (B) DNA elements, (C) Total interspersed repeats, (D) SINES, (E) LINEs, and (F) LTR elements and (G) TE superfamily.

4.2 Association of TEs and Plant Resistance (R)-genes

Plant Resistance (R)-genes in rice encode products that recognize pathogens and triggers disease resistance, but its' expression can be influenced by TEs. For eg. RPM1-Disease Resistance Gene have TE footprint sequences in its exonic regions, protein product and in protein domain (Figure 4.3), (Figure 4.4), (Figure 4.5).

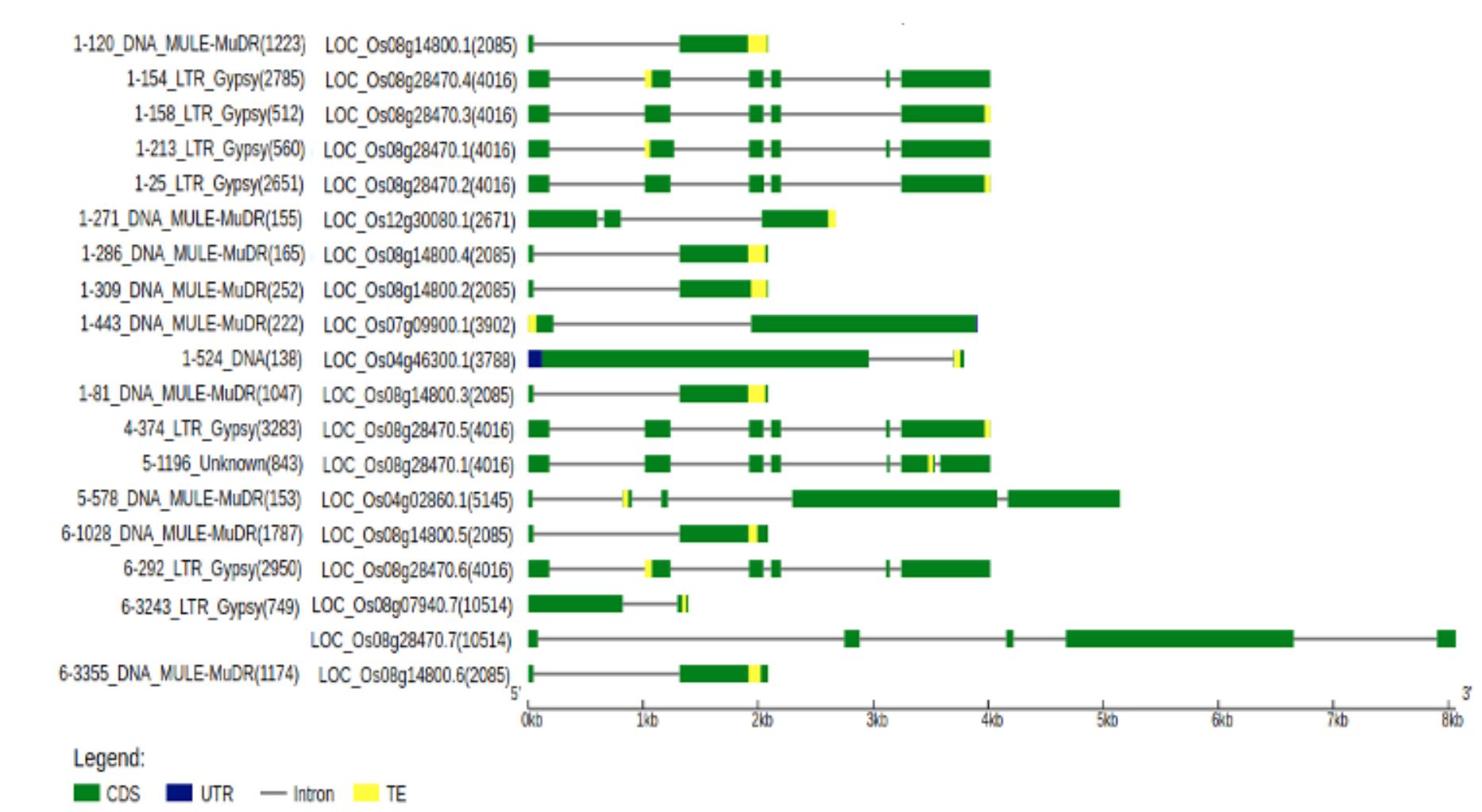


Figure 4.3: TE footprints sequences in *RPM1*-Disease Resistance Gene

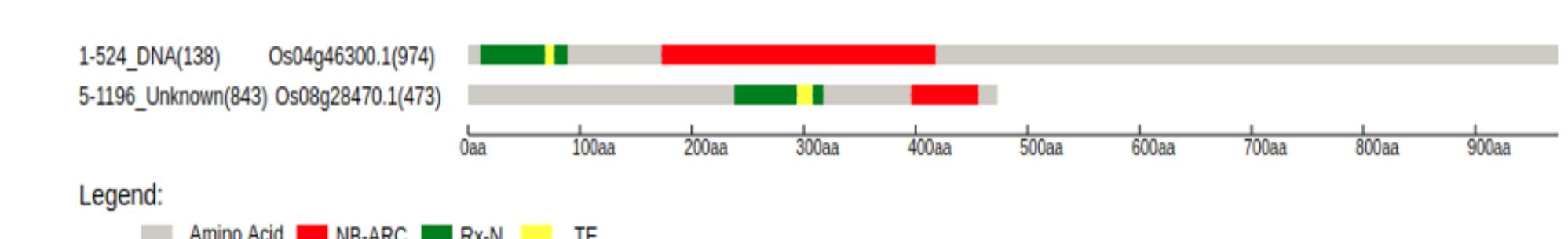


Figure 4.4: TE footprints in Rx-N domain of *RPM1*-Disease Resistance Protein

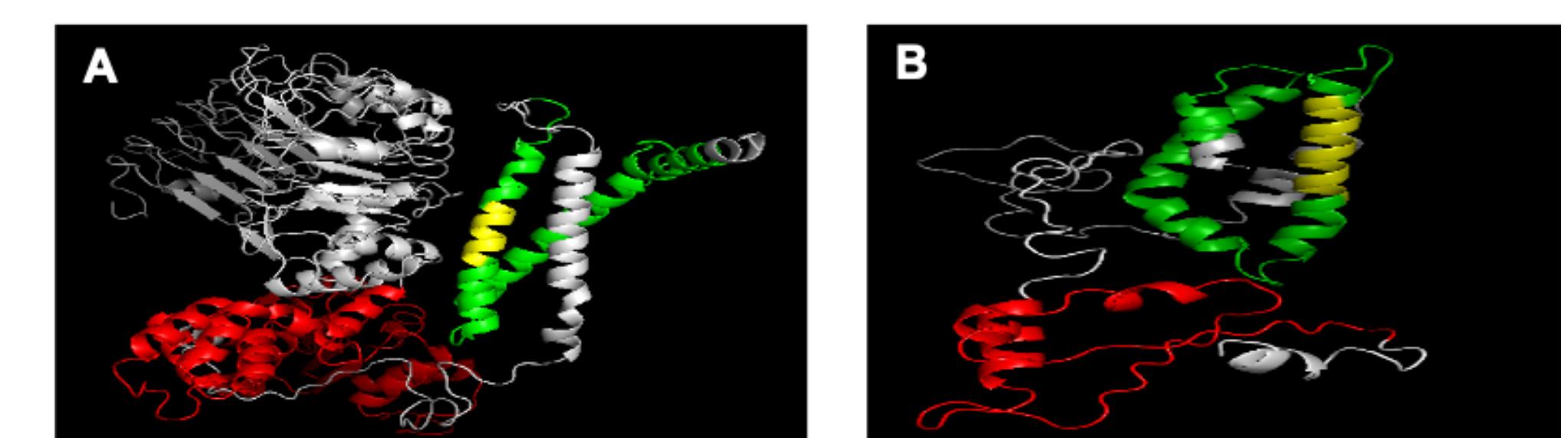


Figure 4.5: The 3D structure of *RPM1*-Disease Resistance Proteins (A. *Os04g46300.1* and B. *Os08g28470.1*) showing TE footprints inside Rx-N domain

5 Conclusions

- This study revealed the repeats, constitute 20.47–48.55% of the rice genome assemblies, with average 39.01%.
- The LTR class **Ty1/Copia** and **Gypsy/DIRS1** account for a significant portion of the genome.
- Some TEs have association with Rice disease resistance (R) genes.

6 References

- Xiong (1990). <https://doi.org/10.1002/j.1460-2075.1990.tb07536.x>
- Wicker (2007). <https://doi.org/10.1038/nrg2165>

