



# TRANSPOSABLE ELEMENT ANNOTATION AND POLYMORPHISM IDENTIFICATION IN WILD STRAWBERRY (*FRAGARIA VESCA*)

Rocio Tolley<sup>1</sup>, David Posé-Padilla<sup>3</sup>, Pablo Manavella<sup>2</sup>, Ileana Tossolini<sup>1,2</sup>

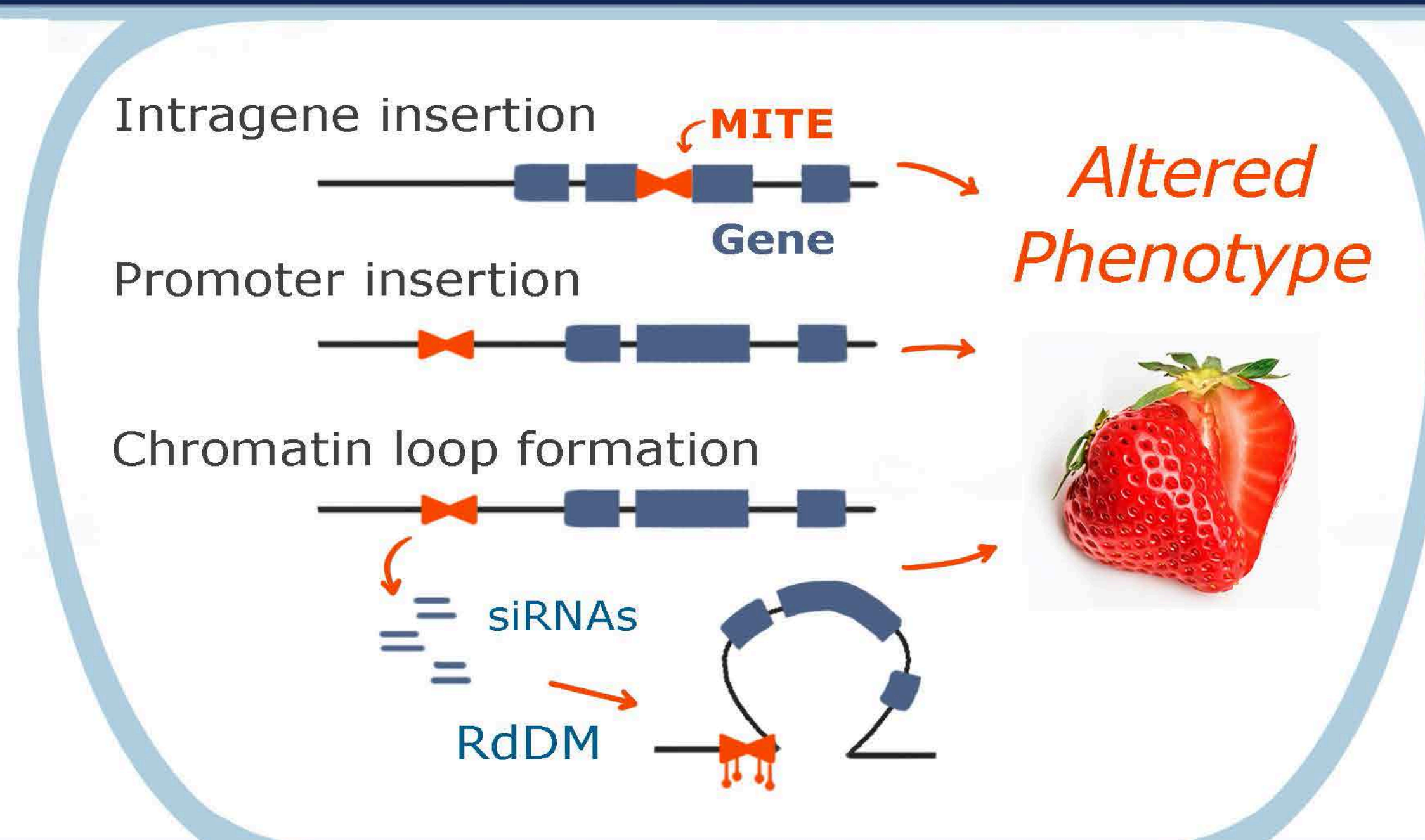
<sup>1</sup>Facultad de Ingeniería, Universidad Nacional de Entre Ríos, Oro Verde, Argentina. <sup>2</sup>Instituto de Agrobiotecnología del Litoral (IAL, CONICET-UNL), Santa Fe, Argentina. <sup>3</sup>Instituto de Hortofruticultura Subtropical y Mediterránea, Málaga, España

## INTRODUCTION

Transposable elements (TEs) are repetitive DNA sequences capable of translocating and multiplying within a host genome.

One of the mechanisms plants have developed to silence TEs is RNA-directed DNA methylation (RdDM).

A subclass of TEs, Miniature Inverted-repeat TEs (MITEs) has recently been shown to alter nearby gene expression via siRNA (small interfering RNA) production, triggering RdDM and chromatin reorganization<sup>[1]</sup>.



Strawberries are a crop with significant economic and nutritional value that has a wide geographic distribution, suggesting a high degree of adaptability. Studying TEs in strawberry could help identify regions with biotechnological potential which, upon editing, may result in favorable phenotypic traits, such as quicker growth rate or improved organoleptic properties. However, TEs have not been annotated in the reference genome *Fragaria vesca*, nor has MITEs variability across populations been analyzed.

## MATERIALS AND METHODS

### TE Annotation

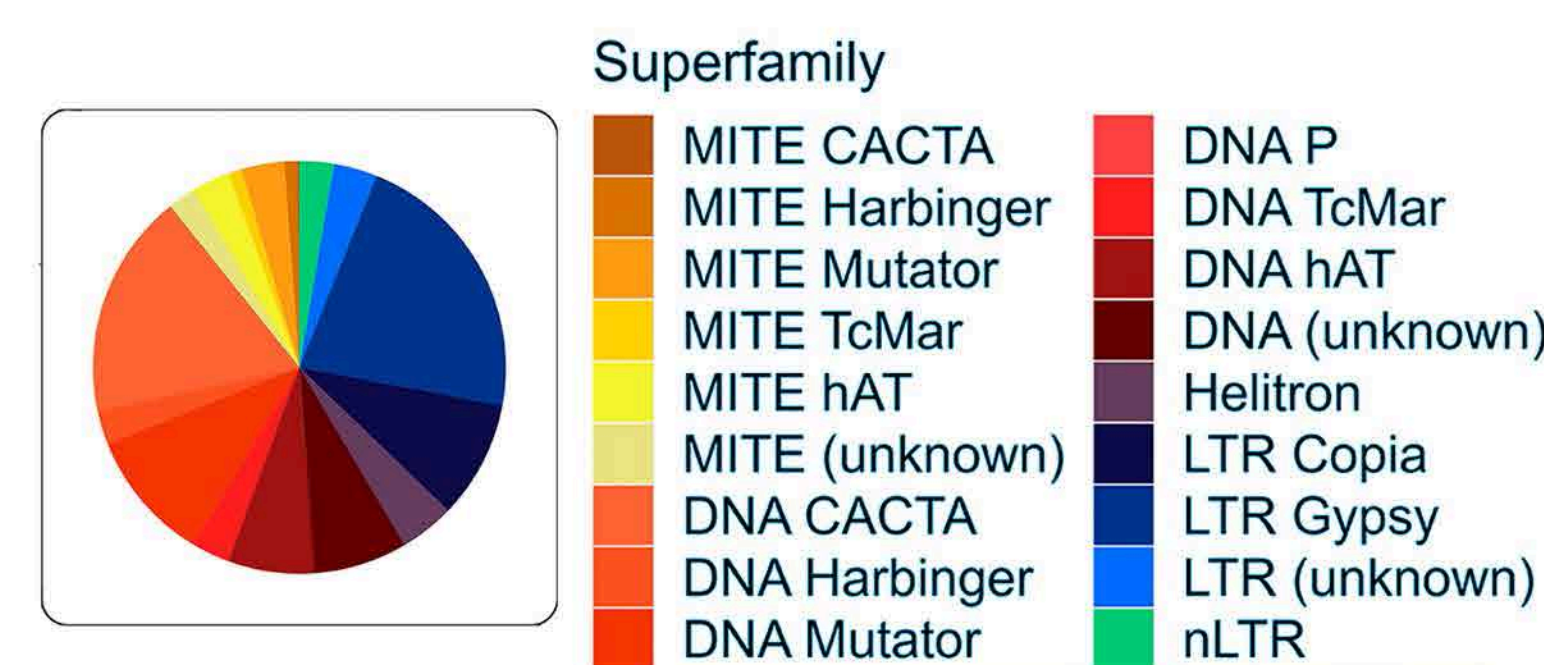
EDTA<sup>[2]</sup> + DeepTE<sup>[3]</sup>

TE identification and initial classification was carried out on *F. vesca* v4 and v6 genome assemblies with EDTA, which combines homology- and structural-based methods. DeepTE was then used to reclassify EDTA's TE library.

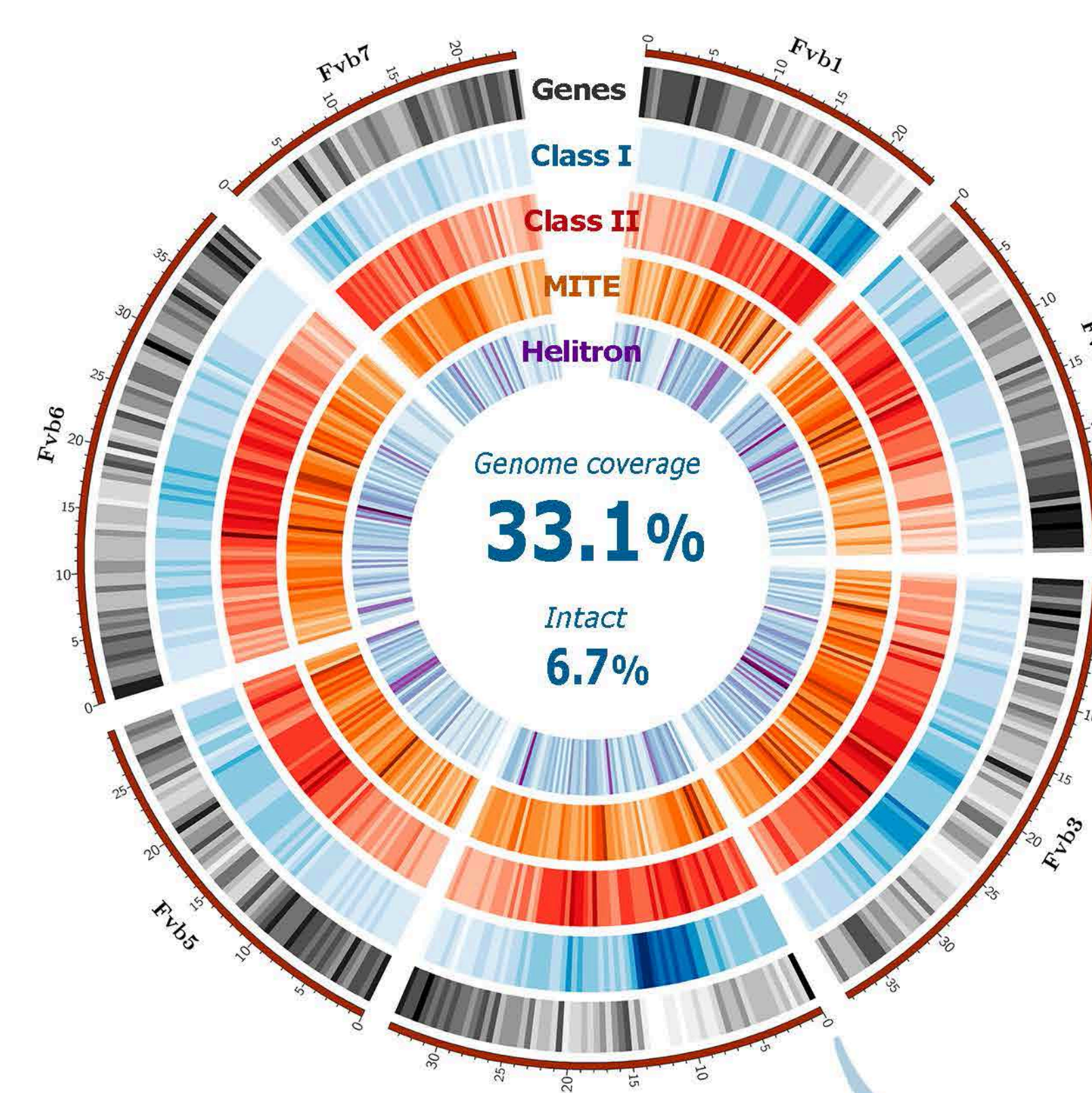
## RESULTS

**Table 1.** TE class counts and total genomic coverage in *F. vesca* v4 and v6.

	Class I	Class II	MITE	Helitron
v4				
Copies	52 039	66 869	15 040	5 987
Coverage (%)	17.02	16.63	1.94	1.35
v6				
Copies	55 862	69 994	14 101	4 609
Coverage (%)	16.57	18.09	1.78	1.08

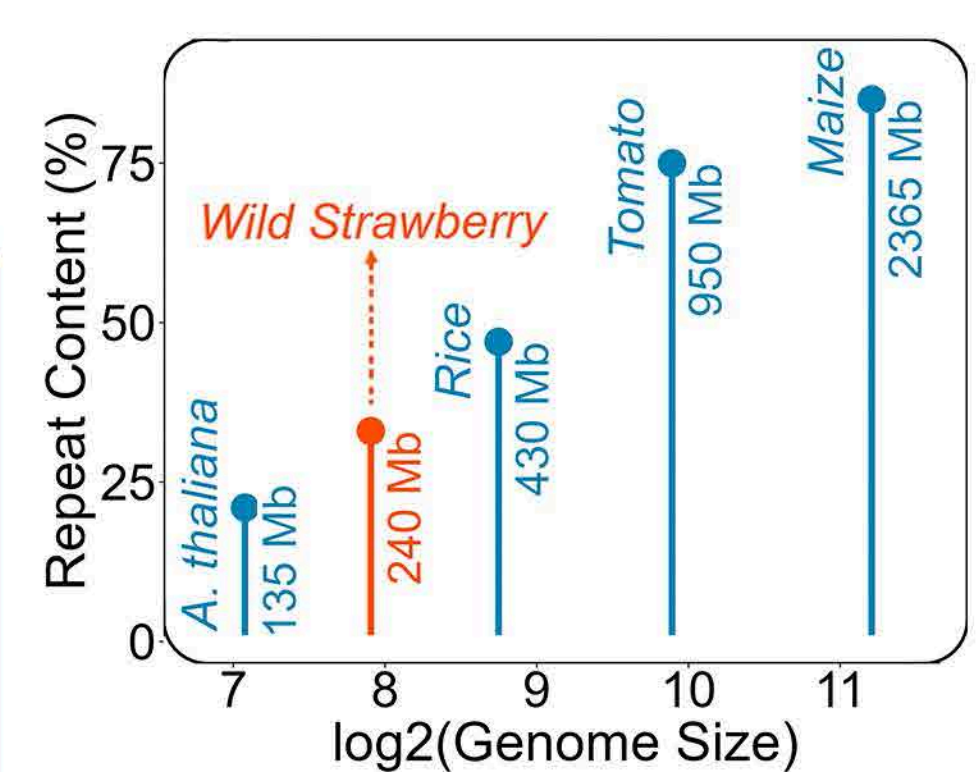


**Fig. 1.** TE superfamily distribution in *F. vesca* v4.



**Fig. 2.** TE density by class.

Class I and Helitron density is higher in pericentromeric regions, while MITE density displays greater overlap with gene-rich regions.



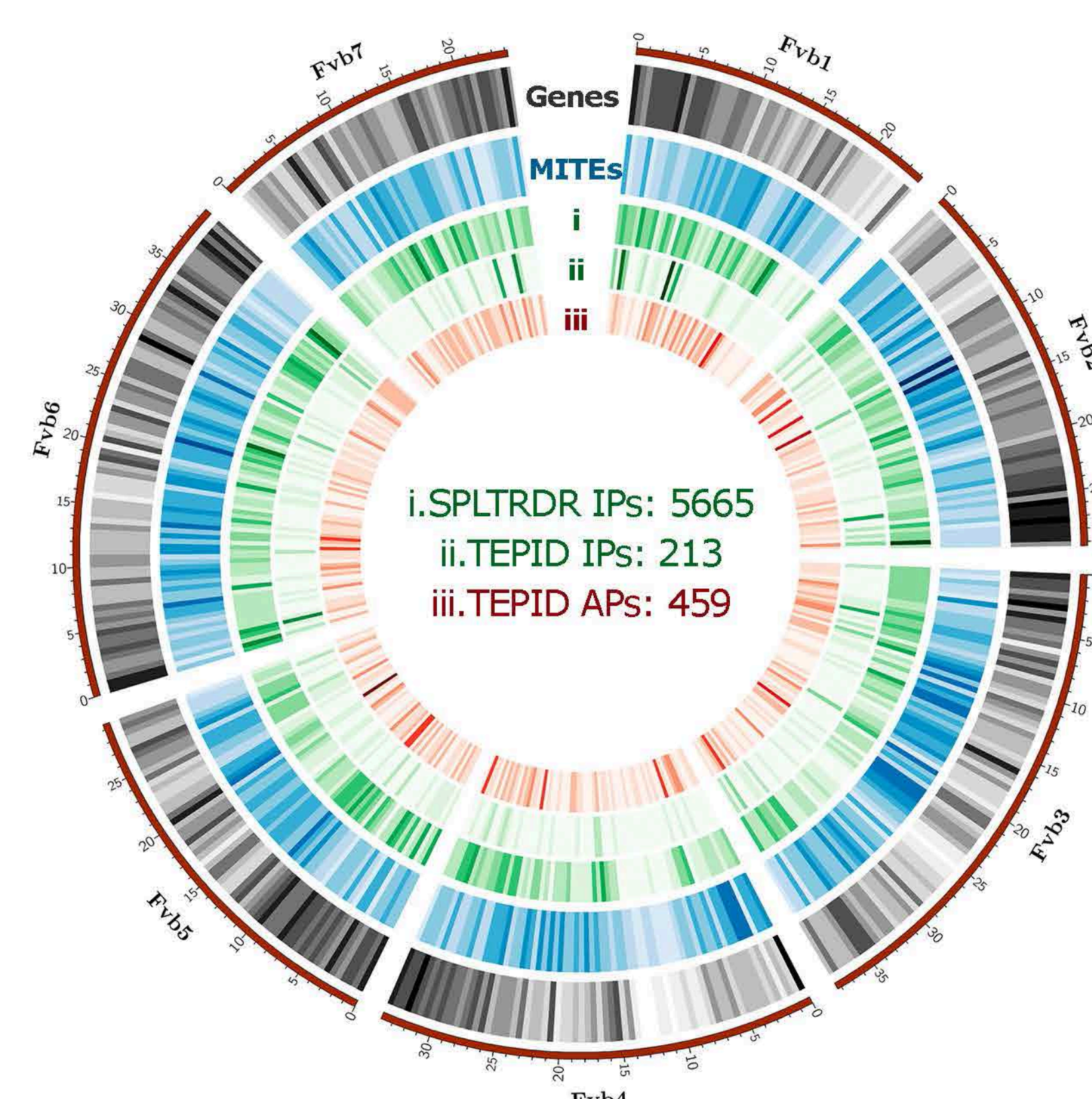
**Fig. 3.** TE content by genome size.

Similar genomic distribution was found in *F. vesca* v6, with the exception of Helitron TEs.

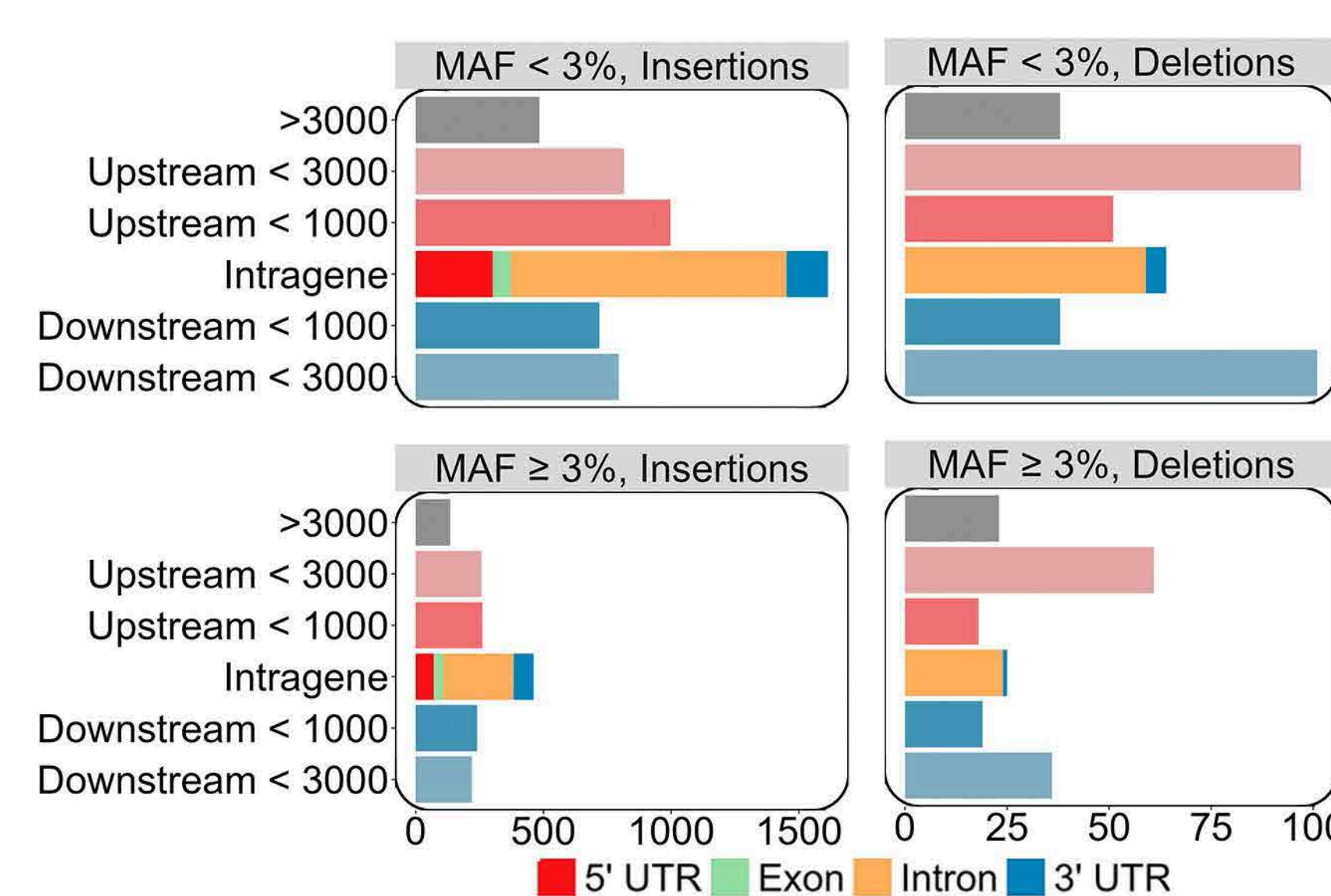
### TE Polymorphism Identification

TEPID<sup>[4]</sup> + SPLITREADER<sup>[5]</sup>

The analysis was carried out on 210 *F. vesca* resequenced accessions, provided by the Instituto de Hortofruticultura Subtropical y Mediterránea (IHSM-CSIC).



**Fig. 5.** MITE insertion polymorphism (IP) and absence polymorphism (AP) distribution.



**Fig. 6.** Polymorphic MITEs close to genes in *Fragaria vesca* v4.

MITE insertions occur at a higher rate in gene-rich regions, but do not spread as readily throughout the population, while there is a higher proportion of frequent (minor allele frequency, MAF ≥ 3%) deletions.

## CONCLUSIONS

TE annotation and MITE polymorphism detection was carried out for the diploid strawberry *Fragaria vesca*, which serves as a model plant for cultivated strawberry as well as the *Rosaceae* family.

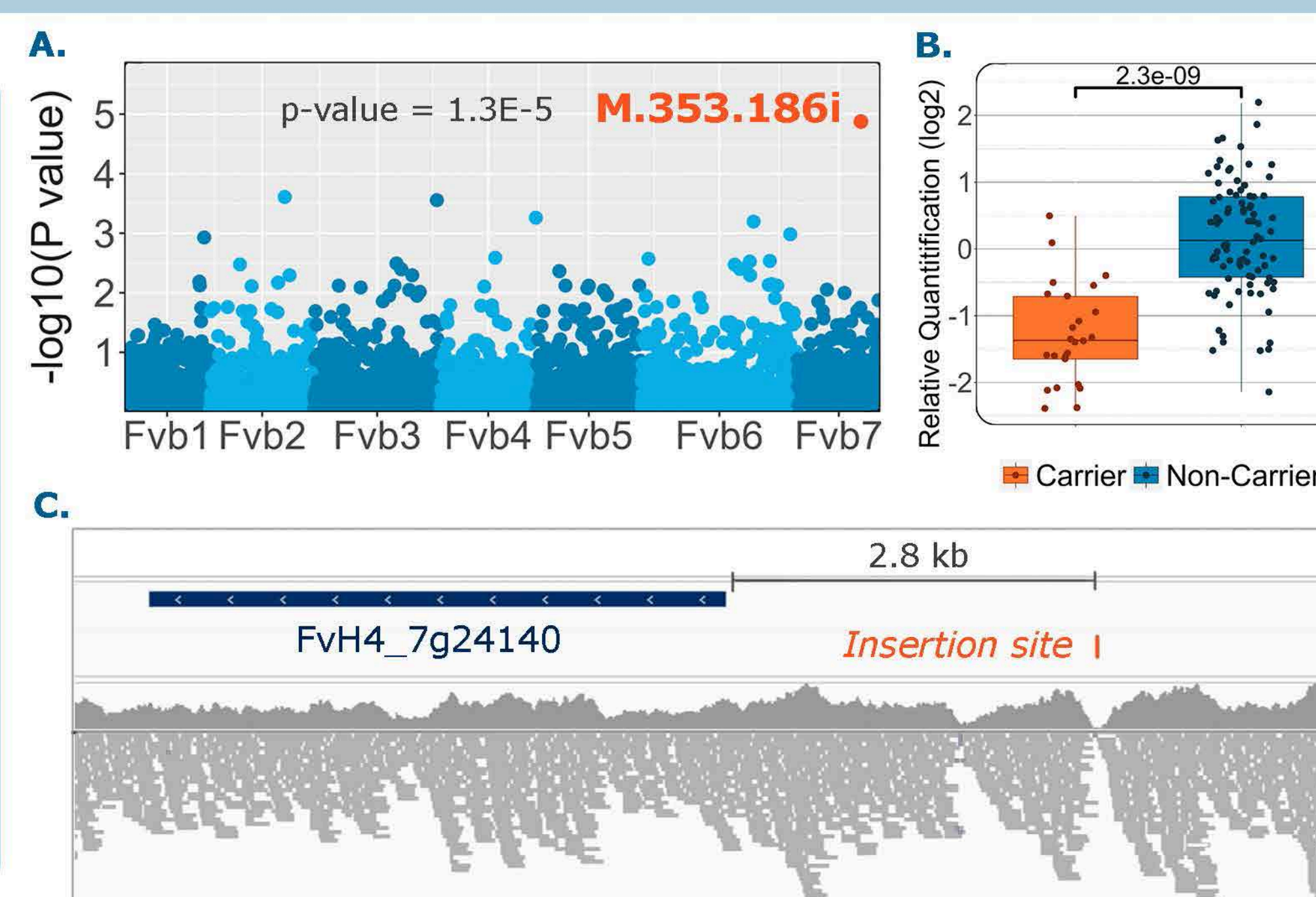
Future analyses will enable the determination of the effects of these variable MITEs on *F. vesca* phenotypic traits, potentially leading to the identification of non-coding sequences that may be used as biotechnological targets by CRISPR-Cas9 seeking to generate plants with altered fruit characteristics.

### MITE insertion - Volatilome Whole-Genome Association Study (GWAS)

GWAS was performed using a generalized linear model (GLM), based on volatile compound data associated to 131 of the resequenced *F. vesca* accessions<sup>[6]</sup>.

A significant association was found with 2-pentylfuran, which impacts strawberry aroma and was recently found to be variable between geographically distinct wild strawberry populations.

The significantly associated insertion is located within 3 kb upstream of the gene *FvH4\_7g24140*, known as *HSPRO2*, or ortholog of sugar beet *HS1 PRO-1 2*.



**Fig. 7.** GWAS analysis.

**A.** Manhattan plot of MITE IP-GWAS for 2-pentylfuran quantification. **B.** Distribution of 2-pentylfuran quantification between carriers and non-carriers of the M.353.186i insertion. **C.** Visualization of the insertion site on the reference genome, and the read alignment of one of the carrier accessions.

## REFERENCES

- Arce, A. et al. (2023), Polymorphic inverted repeats near coding genes impact chromatin topology and phenotypic traits in *Arabidopsis thaliana*. *Cell reports*, 42(1), 112029.
- Ou, S. et al. (2019), Benchmarking transposable element annotation methods for creation of a streamlined, comprehensive pipeline. *Genome Biol* 20, 275.
- Yan, H. et al. (2020), DeepTE: A computational method for de novo classification of transposons with convolutional neural network. *Bioinformatics*, 36(15), 4269–4275.
- Quadrana, L. et al (2016) The *Arabidopsis thaliana* mobilome and its impact at the species level. *eLife* 5:e15716.
- Stuart, T. et al. (2016), Population scale mapping of transposable element diversity reveals links to gene regulation and epigenomic variation. *eLife* 5:e20777.
- Urrutia, M. et al. (2023), Diversity of the volatilome and the fruit size and shape in European woodland strawberry (*Fragaria vesca*). *Plant J*.