



**HAL**  
open science

## **Structural variation involving transposable elements associated with grain width in cultivated rice.**

Marie-Christine Carpentier, Christel Llauro, Wan-Yi Qiu, Yue-Ie Hsing,  
Olivier Panaud

► **To cite this version:**

Marie-Christine Carpentier, Christel Llauro, Wan-Yi Qiu, Yue-Ie Hsing, Olivier Panaud. Structural variation involving transposable elements associated with grain width in cultivated rice.. JOBIM 2022: JOURNÉES OUVERTES DE BIOLOGIE, INFORMATIQUE ET MATHÉMATIQUES 2022, Jul 2022, RENNES, France. hal-04919402

**HAL Id: hal-04919402**

**<https://hal.science/hal-04919402v1>**

Submitted on 31 Jan 2025

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Structural variation involving transposable elements associated with grain width in cultivated rice



Marie-Christine Carpentier<sup>1,2</sup>, Christel Llauro<sup>1,2</sup>, Wan-Yi Qiu<sup>3</sup>, Yue-le Hsing<sup>3</sup> and Olivier Panaud<sup>1,2</sup>



<sup>1</sup> Centre National de la Recherche Scientifique (CNRS), Laboratory of Plant Genome and Development. UMR 5096, 66860 Perpignan, France

<sup>2</sup> University of Perpignan, Laboratory of Plant Genome and Development. UMR 5096, 66860 Perpignan, France

<sup>3</sup> Institute of Plant and Microbial Biology, Academia Sinica, Nankang 115 Taipei, Taiwan

Corresponding author: marie-christine.carpentier@univ-perp.fr



## Context

**Rice**, *Oryza sativa*, is the staple food for half the world population. Rice varieties are very well structured into 7 different varietal groups of which the 2 main ones are *Indica* and *Japonica*.

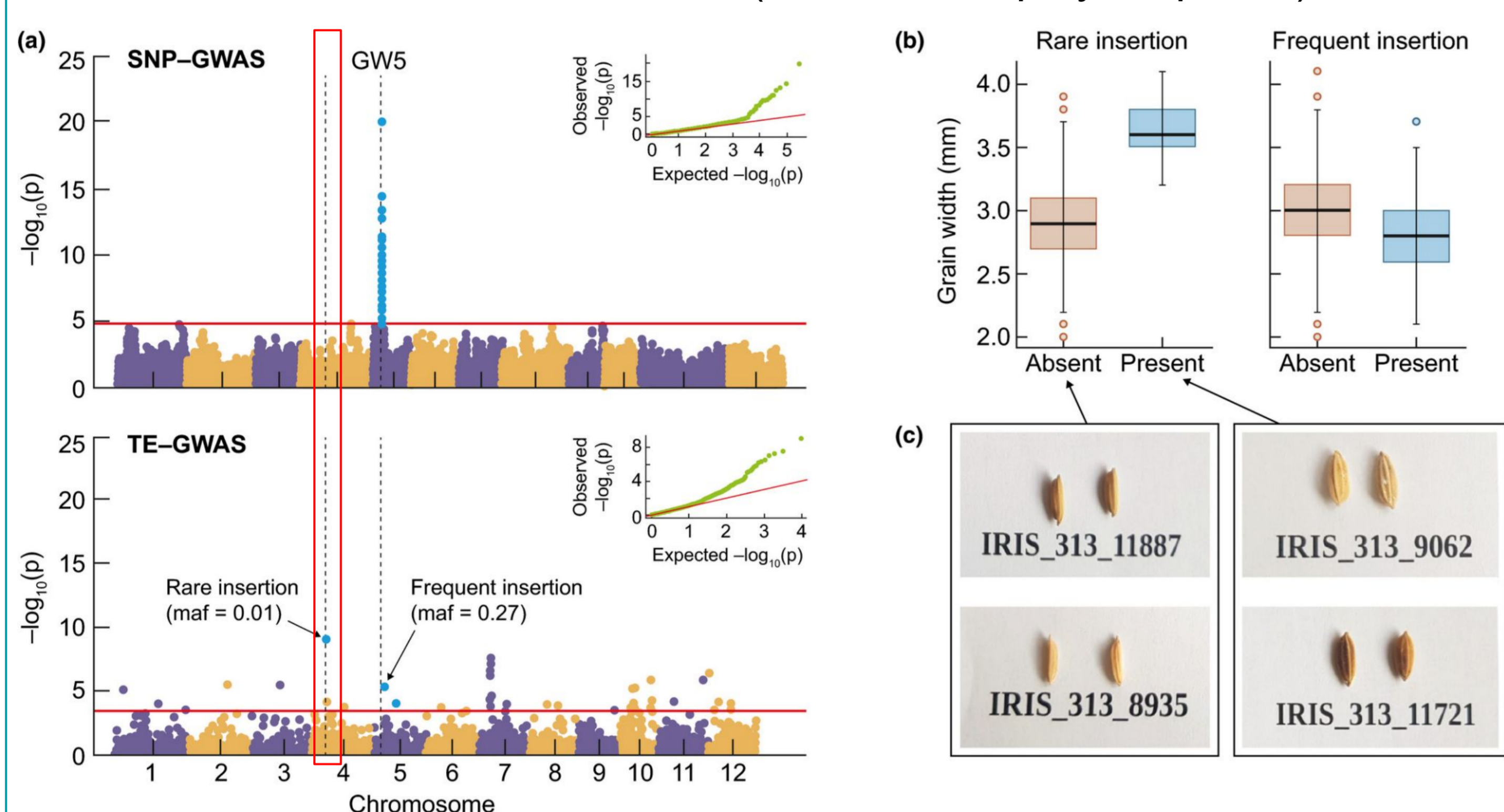
This model species benefits from various genomic resources among which genome sequences of 3000 rice varieties that were publicly released in 2014 (2x75pb Illumina)<sup>1</sup>.

**Transposable elements (TE)** are mobile genetic elements abundant in plant genomes. These mobile entities can provide a more precise picture of rice genome dynamics on a shorter evolutionary scale.

**TRACKPOSON** : pipeline developed to detect all retrotransposons insertions in the 3000 genomes dataset, to discover landscape TE polymorphism in cultivated rice<sup>2</sup>.

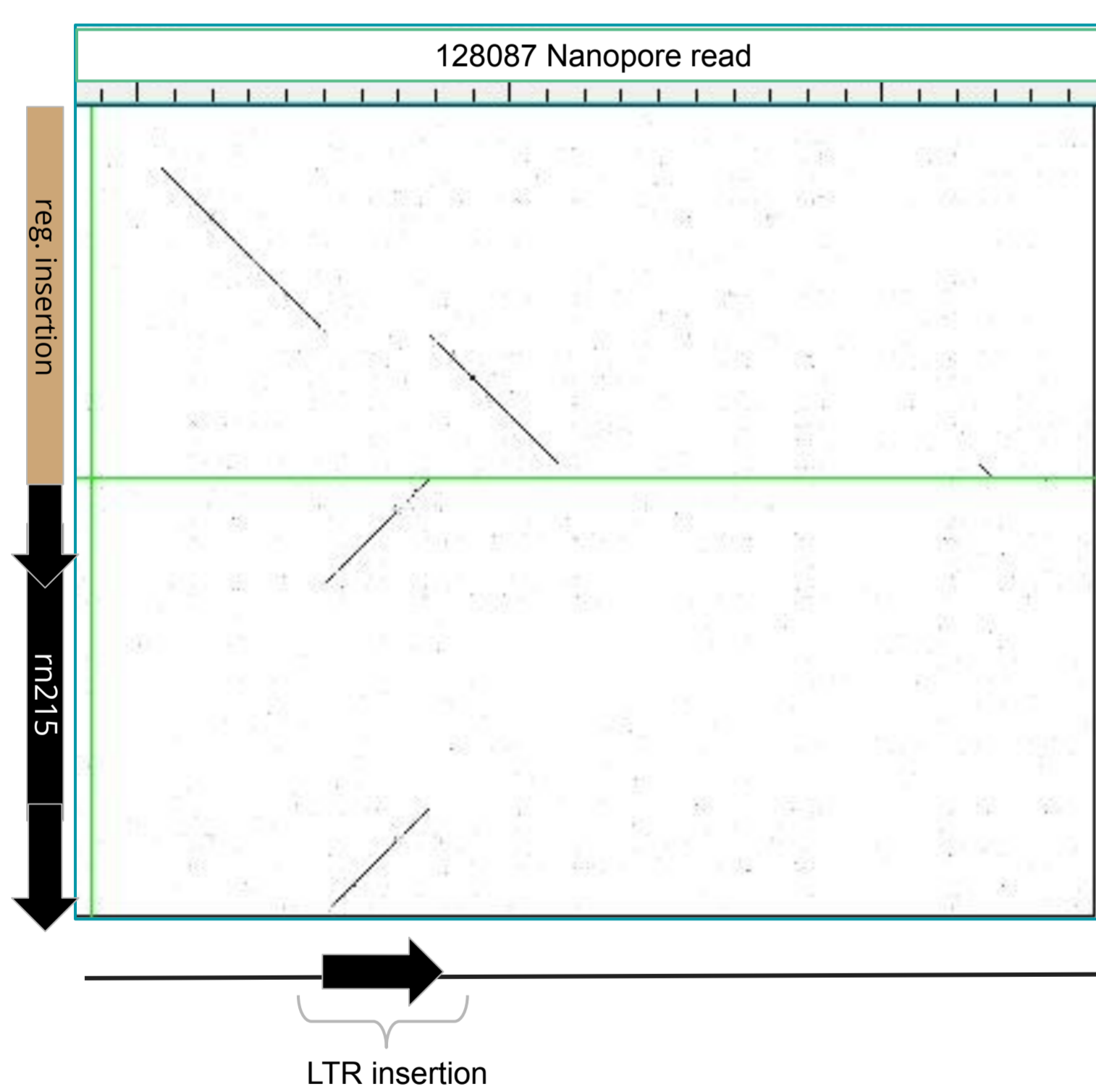
## TE-GWAS : TE genome wide association study<sup>3</sup>

Used TRACKPOSON to define TIPs (TE insertion polymorphism) matrix.



Significant association between an insertion of TE and rice grain width. If the TE is present, the grain is larger.

## TE-GWAS validation with Nanopore read



Validation of TRACKPOSON results with Nanopore read.

Just a part of TE are inserted (LTR) in variety 127087.

Next step : compare contigs including the region of insertion of the both varieties 128087 and 127244.

## Indica seeds & Nanopore sequencing

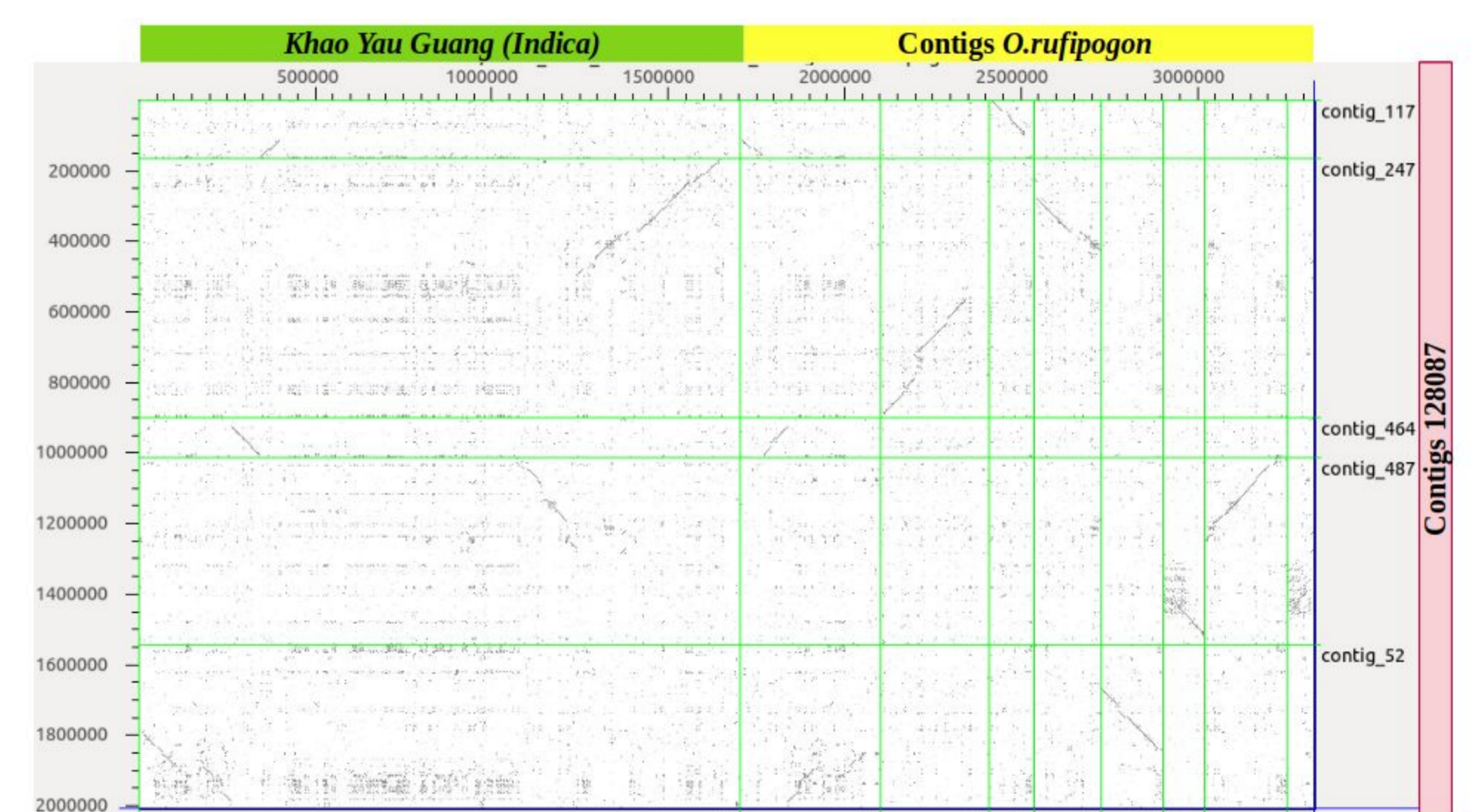


Long-reads sequencing with Oxford Nanopore technology of 2 cultivated rice varieties.

De novo assembly with Flye v2.8.3 software.

Statistics of assembly	128087 (large grain)	127244 (thin grain)
Total assembly size (pb)	382,904,989	389,051,348
Number of contigs	2,152	873
N50	752,396	2,816,480

## Origine of insertion region : wild rice introgression?



Contigs around TE insertion (ie contig 487) seem to have more similarity with ancestral wild rice *O. rufipogon* compare to *Indica* reference genome varieties. Maybe this interest region is an introgression from wild rice and its size is estimated around 800kb.

## Conclusions & Perspectives

With analysis of short and long reads sequencing, we highlighted the presence of a TE insertion and its correlation on the width of the grain in *Indica* varieties.

This insertion is located in a complex region of the *Indica* genome. It seems that this region comes from an introgression of an estimated size of 800kb. The origin of the introgression could be the ancestral wild rice, *O. rufipogon*.

In the future, genomic analysis of contigs with this insertion of variety 128087 is underway : with annotation of gene, TE and further comparative analysis with contig of variety 127244.

Genetic analysis is ongoing by crossing the 2 varieties together and creating a rice population to analysis the segregation of the insertion. In parallel, wet lab validation (PCR) are also in progress to validate the presence of TE insertion in rice plants.

## References

1 - The 3,000 rice genomes project. The 3,000 rice genomes project. GigaScience, December 2014.

2 - Carpentier *et al.* Retrotranspositional landscape of Asian rice revealed by 3000 genomes. Nature Communications, December 2019.

3 - Roland Akakpo, Marie-Christine Carpentier, Yue le Hsing, and Olivier Panaud. The impact of transposable elements on the structure, evolution and function of the rice genome. New Phytologist, April 2020.

## Acknowledgments

This work was supported by a CNRS/Région Languedoc Roussillon research grant and by the University of Perpignan via Domitia.

This work was funded by an Academia Sinica Thematic Project AS-TP-107-L02 in Taiwan.