# Intertwining of transposable elements and non-coding RNAs in plant genomes

Douglas Silva Domingues

w/ Daniel Longhi Fernandes Pedro,  Alexandre Rossi Paschoal

São Paulo State University, Institute of Biosciences at Rio Claro, Brazil
**Federal Technology University of Paraná, Graduation Program in Bioinformatics, Brazil**

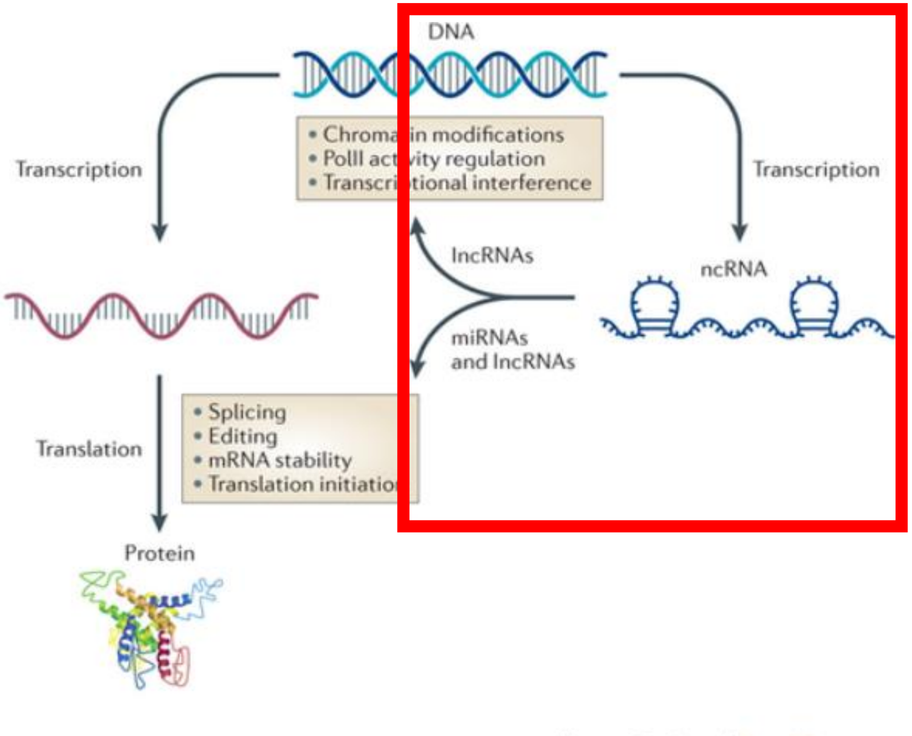
KEEP
CALM
AND
STUDY
PLANTS

# Topics

- ncRNAs – Non-coding RNAs
- TEs – Transposable Elements
- PlaNC-TE: a comprehensive knowledgebase of non-coding RNAs and transposable elements in plants
- Next steps
- Take home message

# What are ncRNAs?

# non-coding RNAs (ncRNAs)

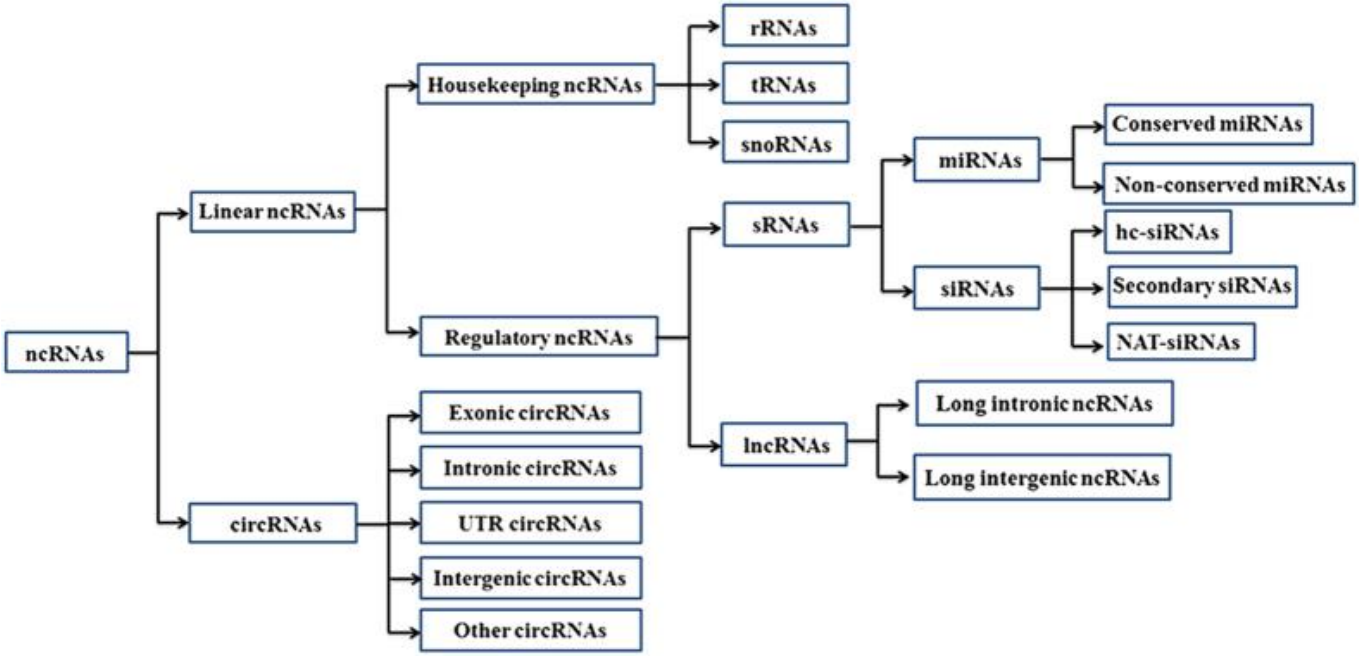


Nature Reviews | Drug Discovery

Wahlestedt, 2013

Sequences that are not translated into protein

# ncRNAs Classification

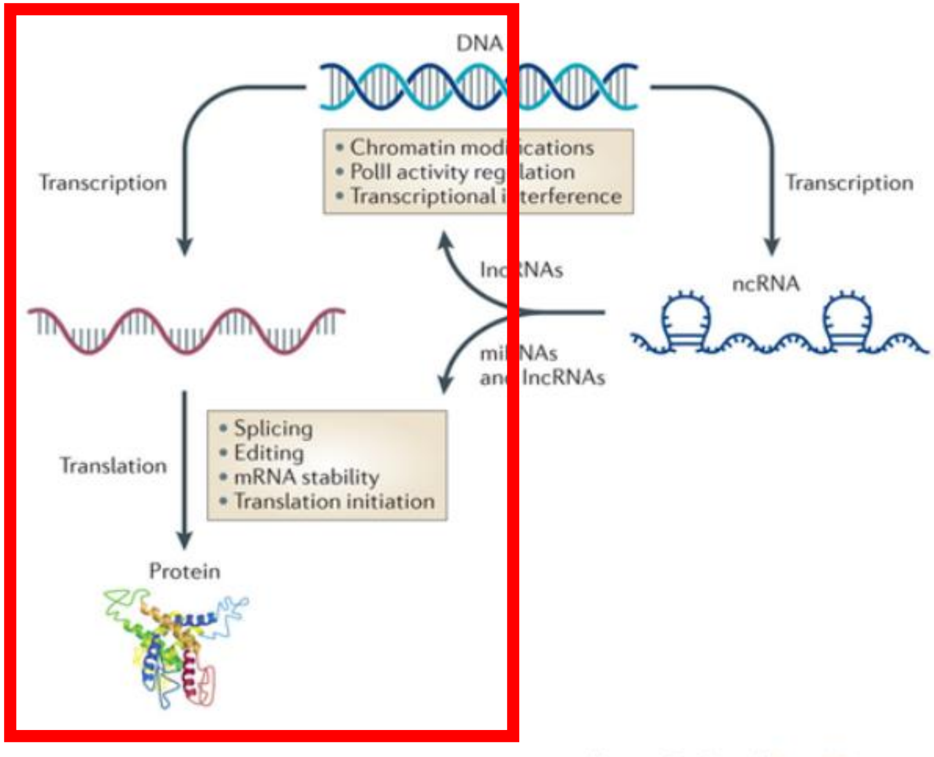


Liu et al. 2017

Organized by:
Location
Length
Function
Structure

# Central dogma of molecular biology.

DNA -> RNA -> Protein



Nature Reviews | Drug Discovery

Wahlestedt, 2013

KEEP CALM.

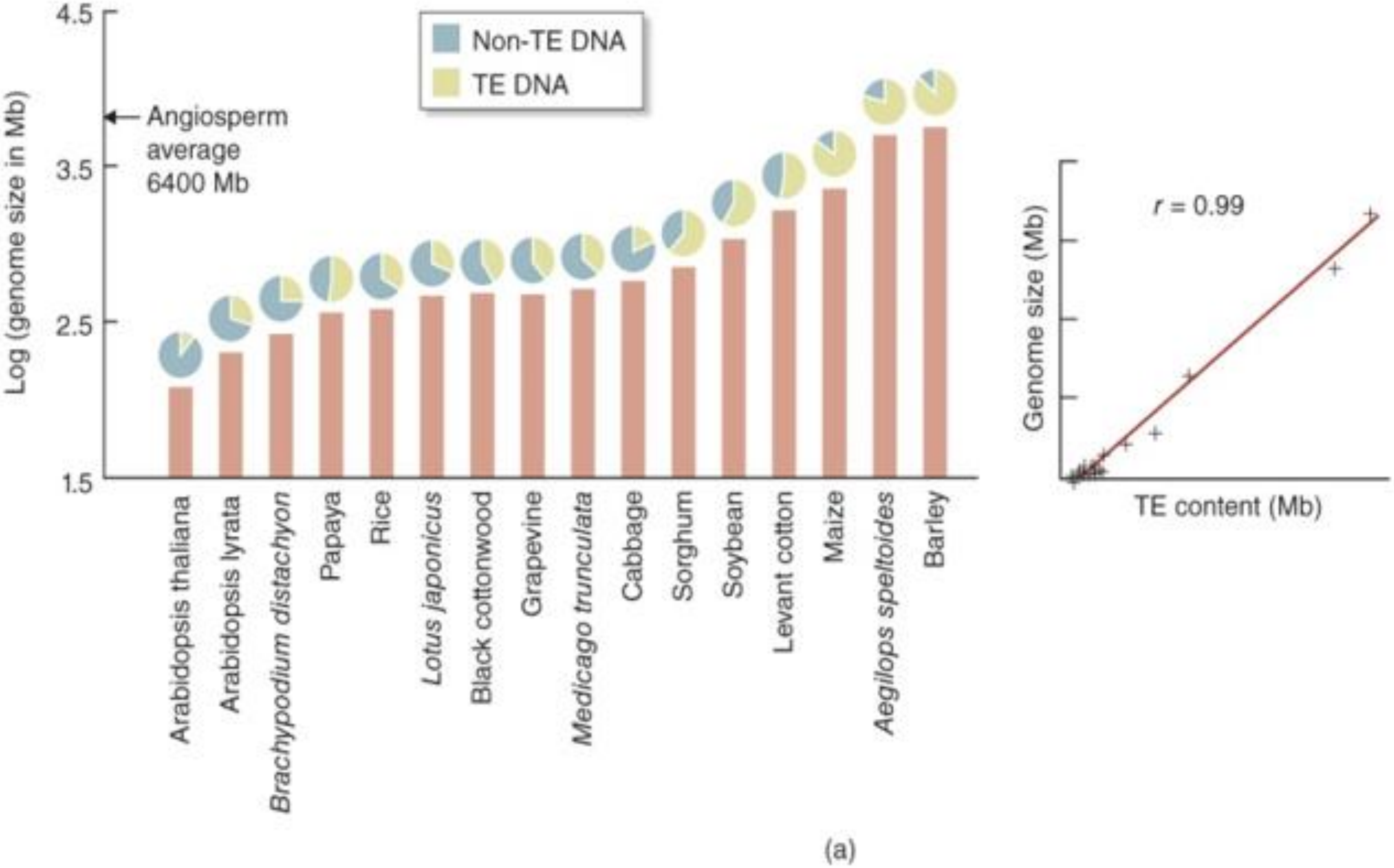
PLANTS

HAVE

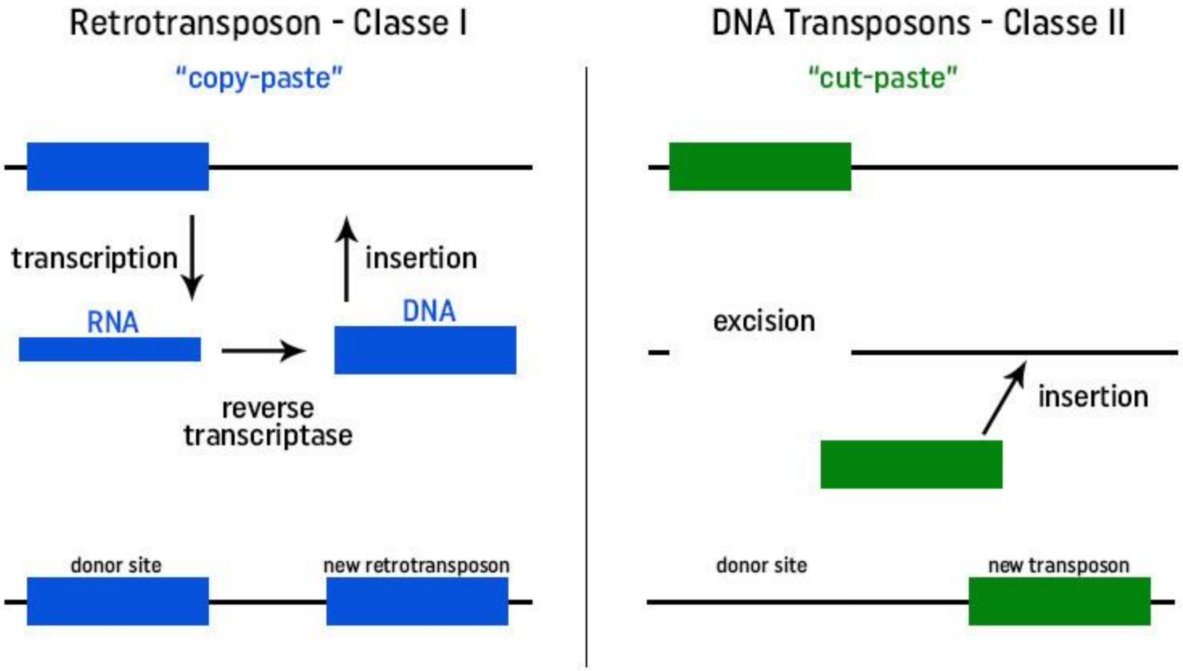
PROTEIN, TOO.

# What are TEs?

# Major componentes in plant genomes and relevant to genome size!



# Transposable Elements – Classes



Own

Sequences that can change their position within a genome.

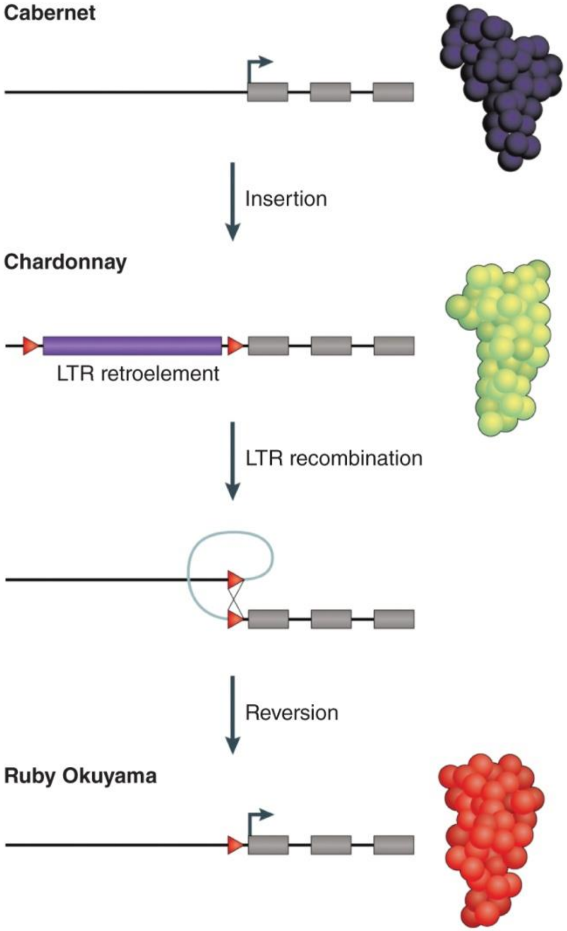
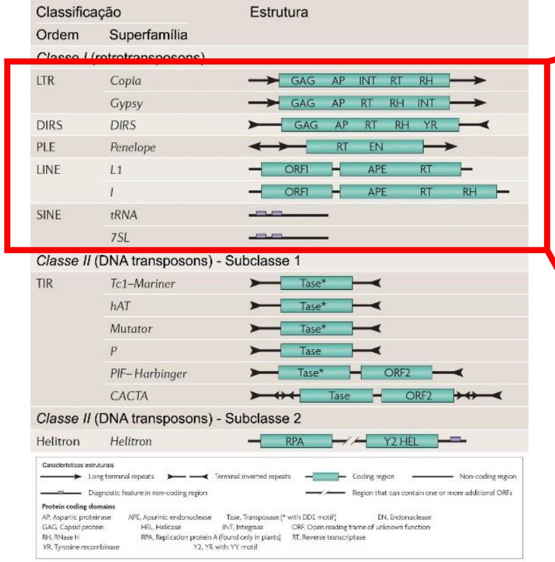


**Figure 3.10** Control of fruit color in grapes by a retrotransposon. Cabernet grapes have a fully functional pigment gene (exons indicated by gray boxes). Insertion of a retrotransposon just upstream of the gene blocks pigment production and leads to green Chardonnay grapes. The LTRs of the element can recombine and remove most of the transposon, but one LTR is left, causing reduced transcription of the locus in Ruby Okoyama grapes. Lisch (2013). Reproduced with permission of Macmillan Publishing Ltd.

*Plant Genes, Genomes and Genetics*, First Edition. Erich Grotewold, Joseph Chappell and Elizabeth A. Kellogg.
© 2015 John Wiley & Sons, Ltd. Published 2015 by John Wiley & Sons, Ltd.
Companion Website: www.wiley.com/go/grotewold/plantgenes.

# Transposable Elements – Hierarchy



| Classificação | | Estrutura |
|---|---|---|
| Ordem | Superfamília | |
| *Classe I* (retrotransposons) | | |
| LTR | *Copia* | GAG AP INT RT RH |
| | *Gypsy* | GAG AP RT RH INT |
| DIRS | *DIRS* | GAG AP RT RH YR |
| PLE | *Penelope* | RT EN |
| LINE | *L1* | ORFI APE RT |
| | *I* | ORFI APE RT RH |
| SINE | *tRNA* | |
| | *7SL* | |
| *Classe II* (DNA transposons) - Subclasse 1 | | |
| TIR | *Tc1–Mariner* | Tase* |
| | *hAT* | Tase* |
| | *Mutator* | Tase* |
| | *P* | Tase |
| | *PIF–Harbinger* | Tase* ORF2 |
| | *CACTA* | Tase ORF2 |
| *Classe II* (DNA transposons) - Subclasse 2 | | |
| Helitron | *Helitron* | RPA Y2 HEL |

Características estruturais

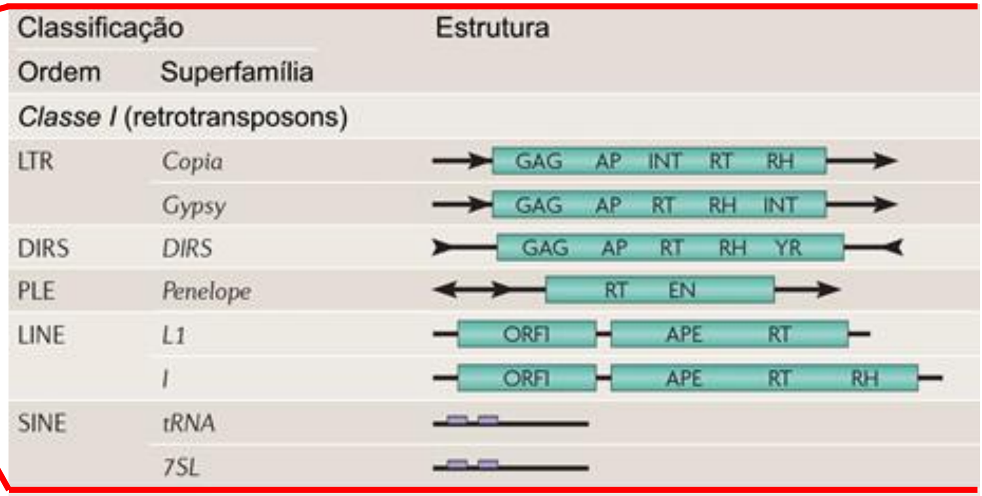
Long terminal repeats
Terminal inverted repeats
Coding region
Non-coding region
Diagnostic feature in non-coding region
Region that can contain one or more additional ORFs

Protein coding domains

AP, Aspartic proteinase
GAG, Capsid protein
RH, RNase H
YR, Tyrosine recombinase
APE, Apurinic endonuclease
HEL, Helicase
RPA, Replication protein A (found only in plants)
Y2, YR with YY motif
Tase, Transposase (* with DDE motif)
INT, Integrase
ORF, Open reading frame of unknown function
RT, Reverse transcriptase
EN, Endonuclease

WICKER et al. 2007

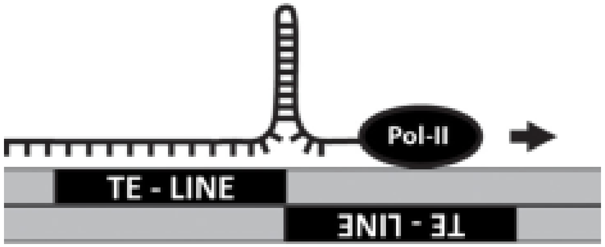
Class | Order | Superfamily



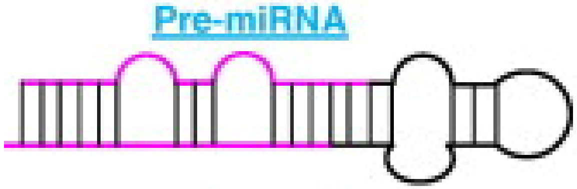
| Classificação | | Estrutura |
|---|---|---|
| Ordem | Superfamília | |
| *Classe I* (retrotransposons) | | |
| LTR | *Copia* | GAG AP INT RT RH |
| | *Gypsy* | GAG AP RT RH INT |
| DIRS | *DIRS* | GAG AP RT RH YR |
| PLE | *Penelope* | RT EN |
| LINE | *L1* | ORFI APE RT |
| | *I* | ORFI APE RT RH |
| SINE | *tRNA* | |
| | *7SL* | |

**LTR**: Long Terminal Repeat
**non-LTR**: non-Long Terminal Repeat

# Why ncRNA:TEs?

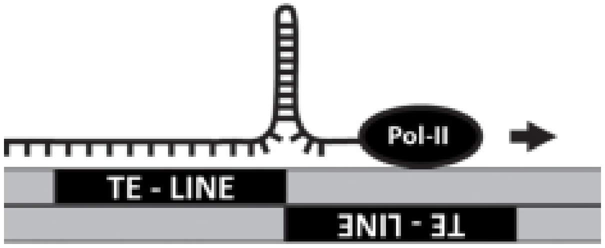


*Roberts et al. 2014*

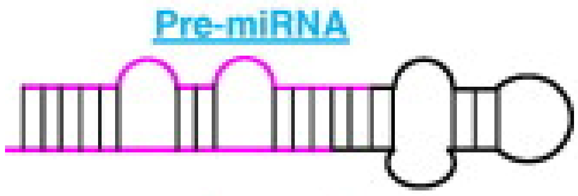


*Adapted from: Maiti et al. 2012*

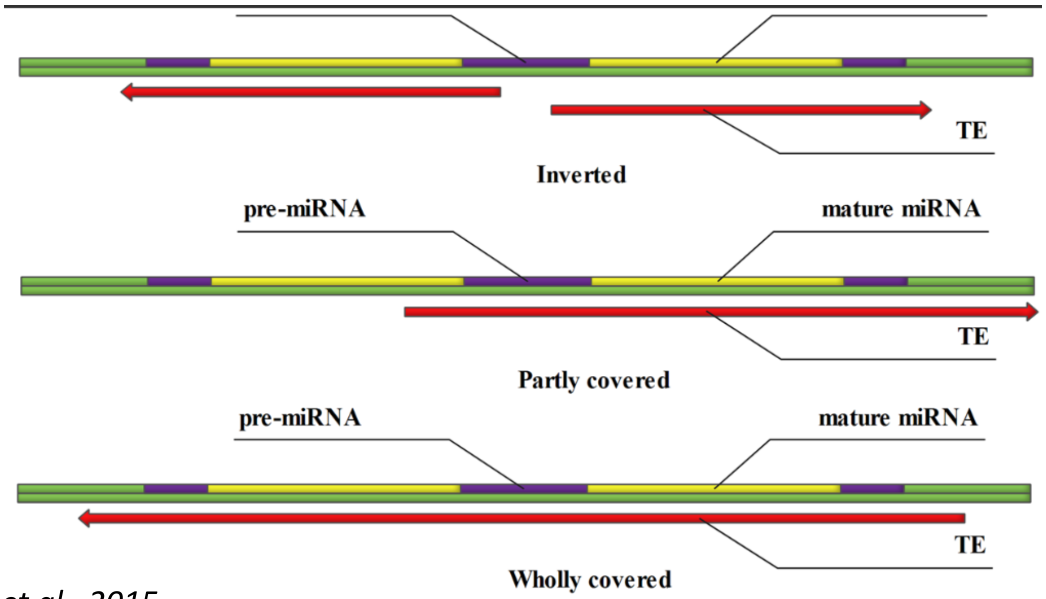
# Why ncRNA:TEs?



*Roberts et al. 2014*

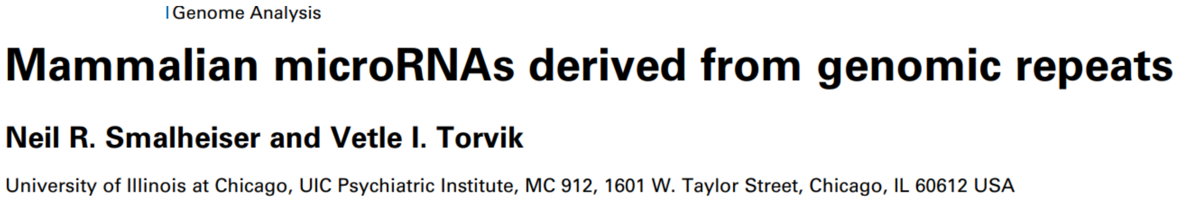


*Adapted from: Maiti et al. 2012*



*Qin et al., 2015*

# - ncRNA and TEs: known but ignored at in large-scale analyses

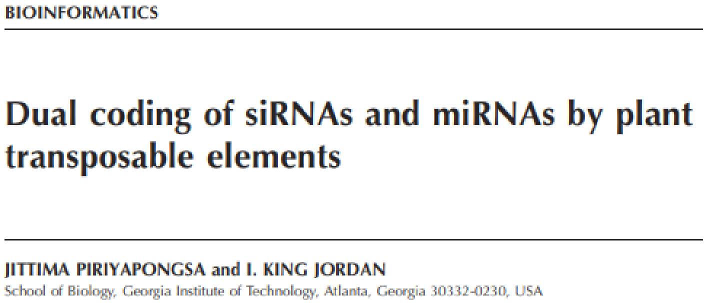


Genome Analysis

**Mammalian microRNAs derived from genomic repeats**

**Neil R. Smalheiser and Vetle I. Torvik**

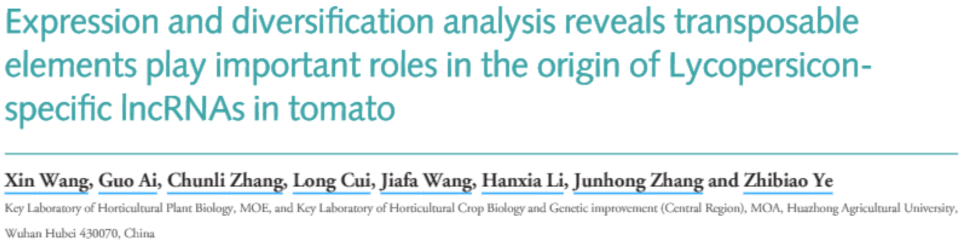
University of Illinois at Chicago, UIC Psychiatric Institute, MC 912, 1601 W. Taylor Street, Chicago, IL 60612 USA



BIOINFORMATICS

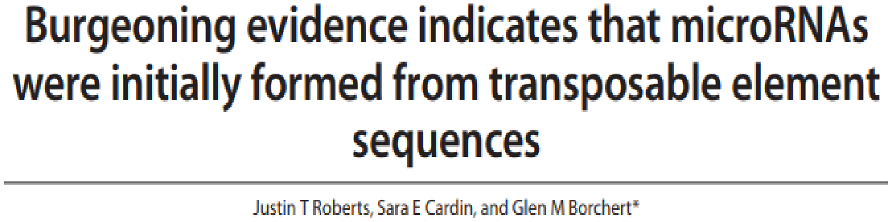
**Dual coding of siRNAs and miRNAs by plant transposable elements**

JITTIMA PIRIYAPONGSA and I. KING JORDAN
School of Biology, Georgia Institute of Technology, Atlanta, Georgia 30332-0230, USA



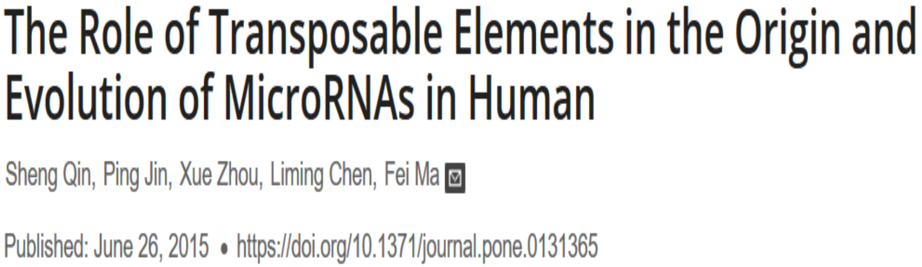
**Expression and diversification analysis reveals transposable elements play important roles in the origin of Lycopersicon-specific lncRNAs in tomato**

Xin Wang, Guo Ai, Chunli Zhang, Long Cui, Jiafa Wang, Hanxia Li, Junhong Zhang and Zhibiao Ye
Key Laboratory of Horticultural Plant Biology, MOE, and Key Laboratory of Horticultural Crop Biology and Genetic improvement (Central Region), MOA, Huazhong Agricultural University, Wuhan Hubei 430070, China



**Burgeoning evidence indicates that microRNAs were initially formed from transposable element sequences**

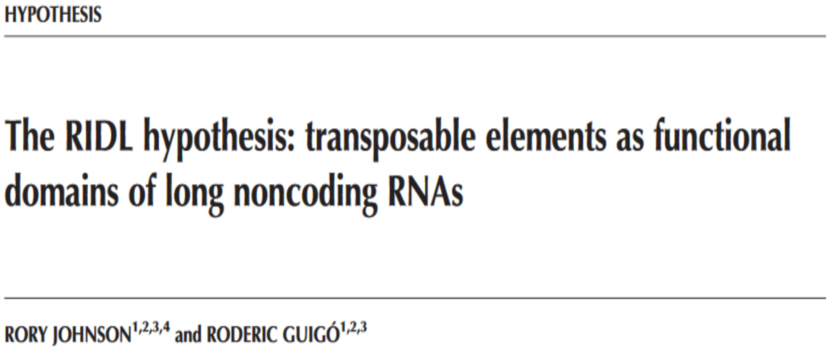
Justin T Roberts, Sara E Cardin, and Glen M Borchert*



**The Role of Transposable Elements in the Origin and Evolution of MicroRNAs in Human**

Sheng Qin, Ping Jin, Xue Zhou, Liming Chen, Fei Ma

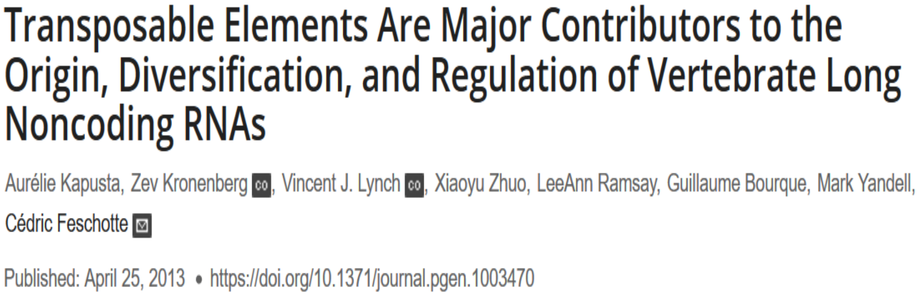
Published: June 26, 2015 • https://doi.org/10.1371/journal.pone.0131365



HYPOTHESIS

**The RIDL hypothesis: transposable elements as functional domains of long noncoding RNAs**

RORY JOHNSON[1,2,3,4] and RODERIC GUIGÓ[1,2,3]



**Transposable Elements Are Major Contributors to the Origin, Diversification, and Regulation of Vertebrate Long Noncoding RNAs**

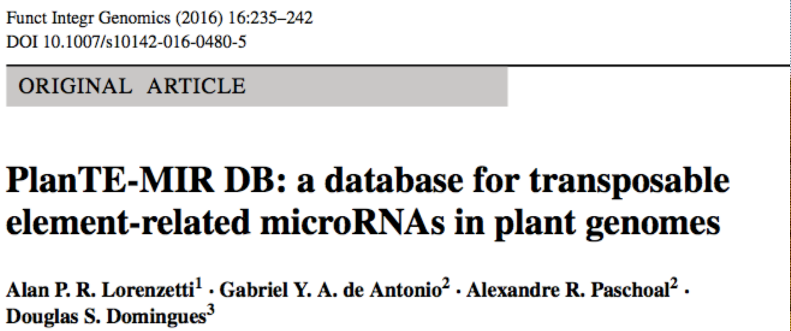
Aurélie Kapusta, Zev Kronenberg, Vincent J. Lynch, Xiaoyu Zhuo, LeeAnn Ramsay, Guillaume Bourque, Mark Yandell, Cédric Feschotte

Published: April 25, 2013 • https://doi.org/10.1371/journal.pgen.1003470

# Initial efforts

- First database to organize such information in plants

- PlanTE-MIR DB {10 *ssp}* – *v.1 - 2016*
    - miRNA:TE {152 evidences in 10 genomes}

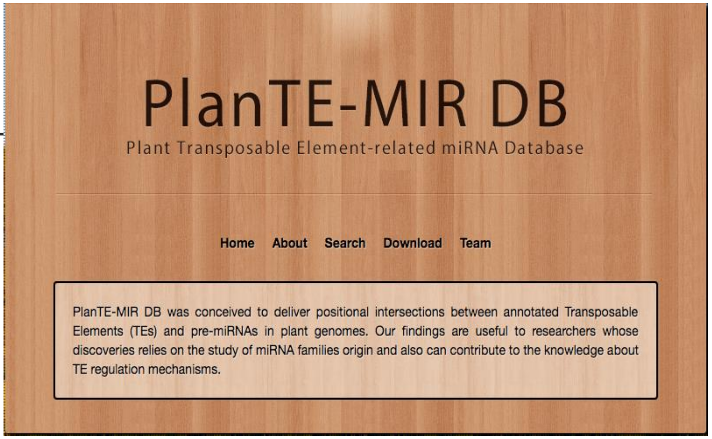


Funct Integr Genomics (2016) 16:235–242
DOI 10.1007/s10142-016-0480-5

ORIGINAL ARTICLE

**PlanTE-MIR DB: a database for transposable element-related microRNAs in plant genomes**

**Alan P. R. Lorenzetti[1] · Gabriel Y. A. de Antonio[2] · Alexandre R. Paschoal[2] · Douglas S. Domingues[3]**



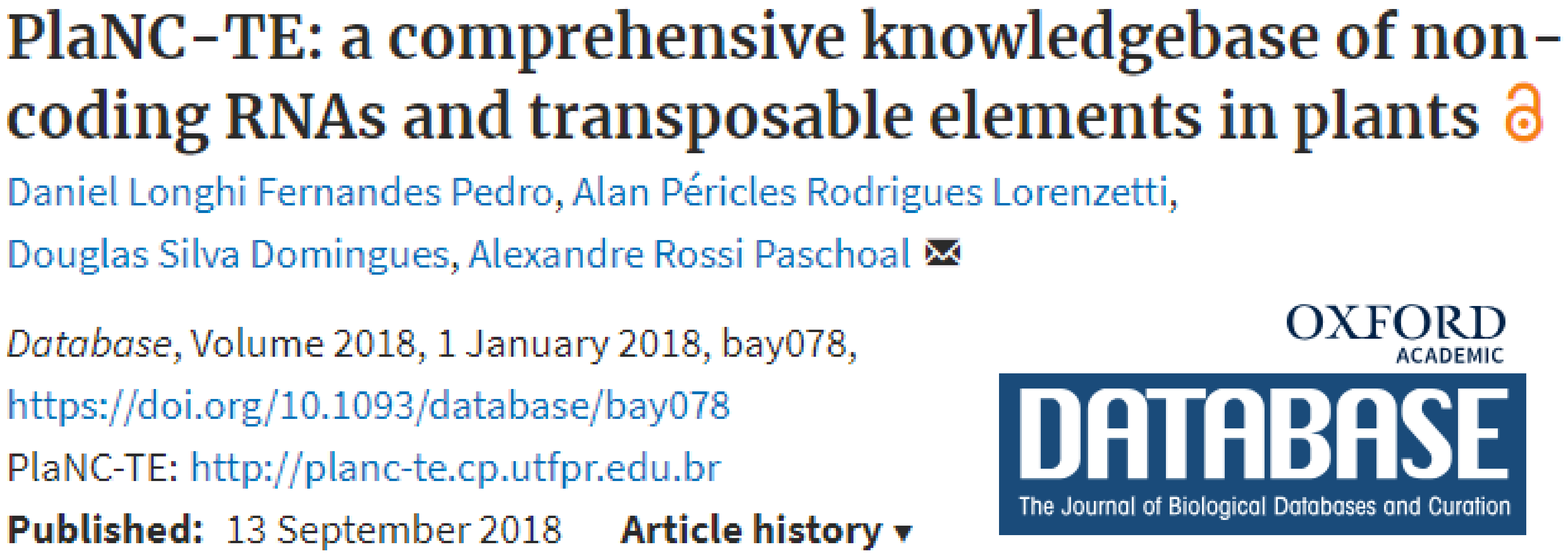
PlanTE-MIR DB

Plant Transposable Element-related miRNA Database

PlanTE-MIR DB was conceived to deliver positional intersections between annotated Transposable Elements (TEs) and pre-miRNAs in plant genomes. Our findings are useful to researchers whose discoveries relies on the study of miRNA families origin and also can contribute to the knowledge about TE regulation mechanisms.

What if we expand this to all public plant genomes with TE and ncRNA annotation data?

# http://planc-te.cp.utfpr.edu.br



## PlaNC-TE: a comprehensive knowledgebase of non-coding RNAs and transposable elements in plants

Daniel Longhi Fernandes Pedro, Alan Péricles Rodrigues Lorenzetti, Douglas Silva Domingues, Alexandre Rossi Paschoal ✉

*Database*, Volume 2018, 1 January 2018, bay078,
https://doi.org/10.1093/database/bay078
PlaNC-TE: http://planc-te.cp.utfpr.edu.br

**Published:** 13 September 2018 **Article history** ▾

OXFORD ACADEMIC

DATABASE

The Journal of Biological Databases and Curation

# Objectives

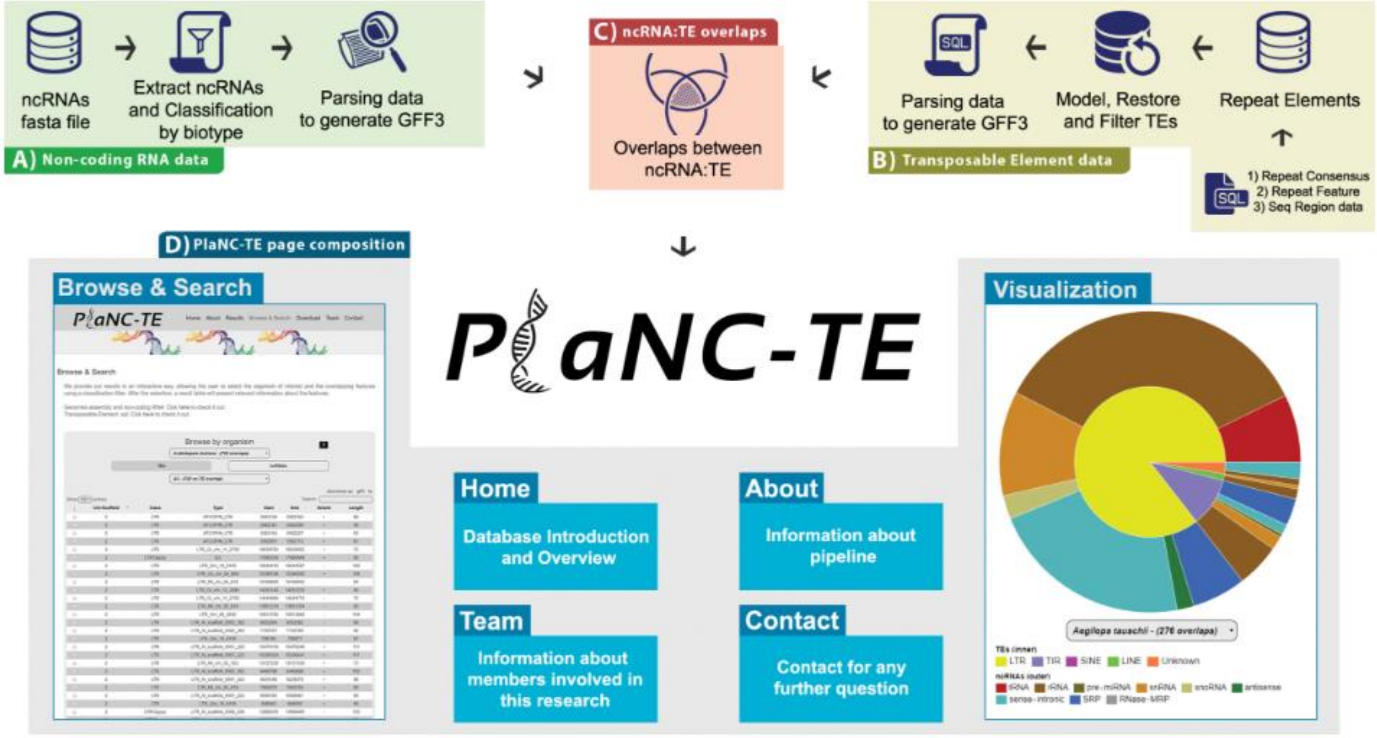
- Extend PlanTE-MIR to all plant genomes available in Ensembl (53 species)

- Extend to all ncRNA classes available

- Make available a well-organized data
  - Lack of an organized repository of ncRNA:TEs for complete genomes in plants
  - Standardize outputs

- Stimulate studies in TEs and ncRNAs in plant genomes

# PlaNC-TE - *Workflow*


ncRNAs fasta file
Extract ncRNAs and Classification by biotype
Parsing data to generate GFF3
A) Non-coding RNA data
C) ncRNA:TE overlaps
Overlaps between ncRNA:TE
SQL
Parsing data to generate GFF3
Model, Restore and Filter TEs
Repeat Elements
B) Transposable Element data
1) Repeat Consensus
2) Repeat Feature
3) Seq Region data
D) PlaNC-TE page composition
Browse & Search
PlaNC-TE
Browse by organism
PlaNC-TE
Home
Database Introduction and Overview
About
Information about pipeline
Team
Information about members involved in this research
Contact
Contact for any further question
Visualization
Aegilops tauschii - (276 overlaps)
TEs (inner)
LTR
TIR
SINE
LINE
Unknown
ncRNAs (outer)
tRNA
rRNA
pre-miRNA
snRNA
snoRNA
antisense
sense-intronic
SRP
RNase-MRP

A) Selecting ncRNAs.

B) Filtering TEs.

C) ncRNA:TE analysis

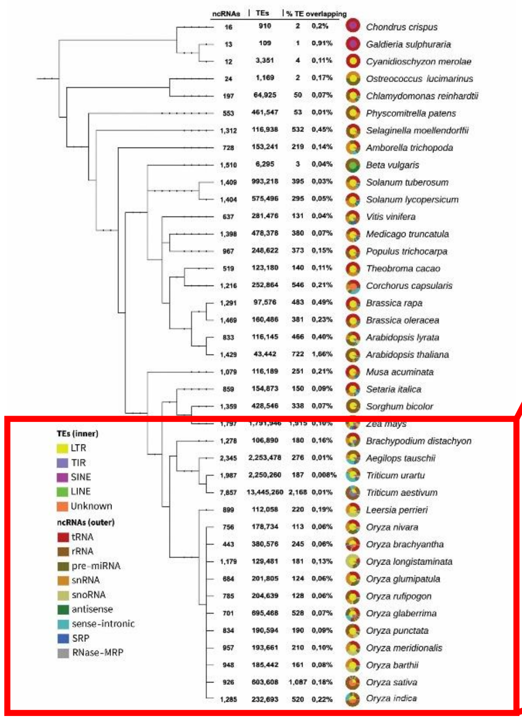
D) PlaNC-TE webpages

# PlaNC-TE: A comprehensive knowledgebase of non-coding RNAs and transposable elements in plants.

- Genomic sequence source
  - Ensembl Plants
  - 53 genomes

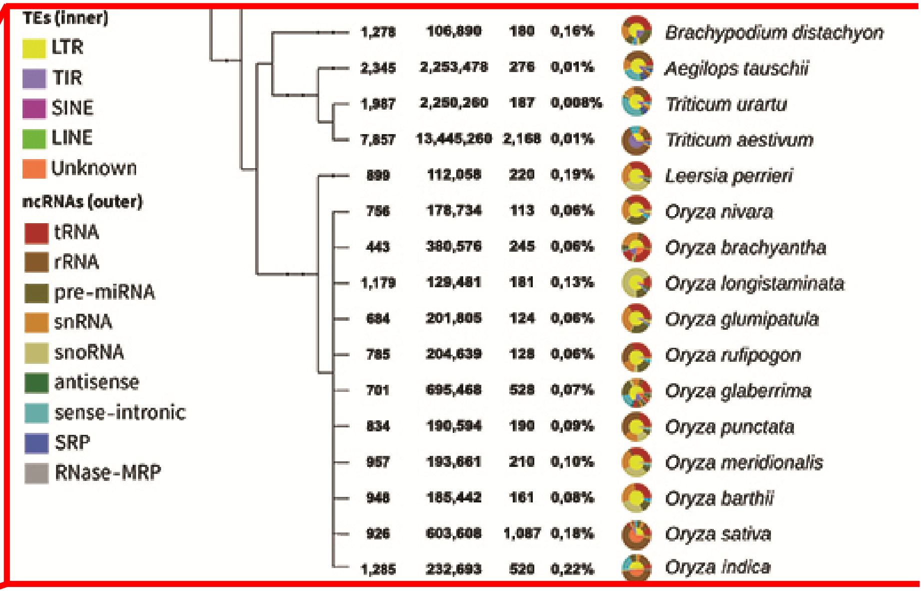
- Retrieved sequences in ncRNAs and TEs
  - ncRNAs – 58,390 records (53 genomes)
  - TEs – 31,217,630 records (45 genomes)

# PlaNC-TE – *Phylogenetic tree*



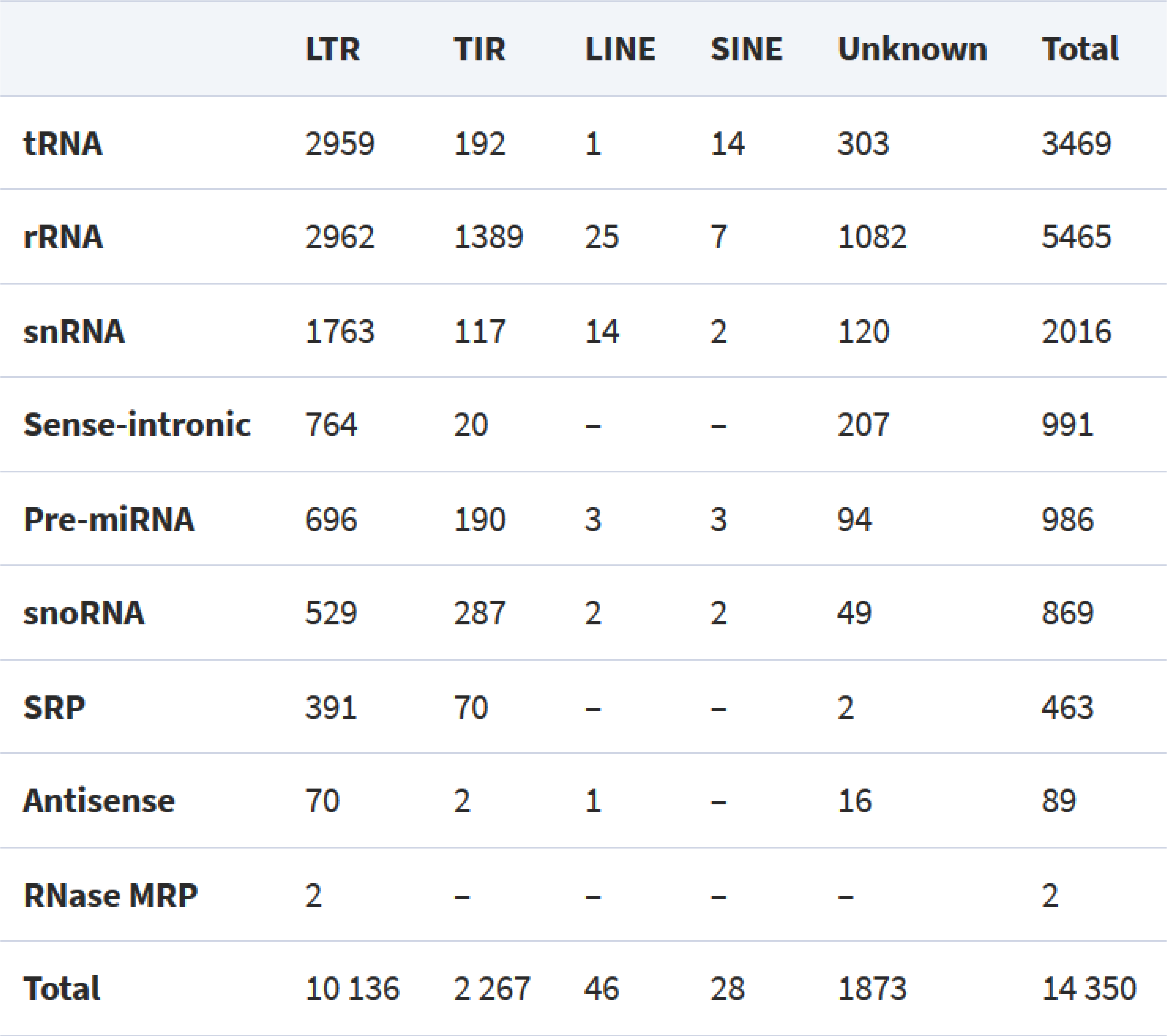
| ncRNAs | TEs | % TE overlapping | | |
|---|---|---|---|---|
| 16 | 910 | 2 | 0,2% | *Chondrus crispus* |
| 13 | 109 | 1 | 0,91% | *Galdieria sulphuraria* |
| 12 | 3,351 | 4 | 0,11% | *Cyanidioschyzon merolae* |
| 24 | 1,169 | 2 | 0,17% | *Ostreococcus lucimarinus* |
| 197 | 64,925 | 50 | 0,07% | *Chlamydomonas reinhardtii* |
| 553 | 461,547 | 53 | 0,01% | *Physcomitrella patens* |
| 1,312 | 116,938 | 532 | 0,45% | *Selaginella moellendorffii* |
| 728 | 153,241 | 219 | 0,14% | *Amborella trichopoda* |
| 1,510 | 6,295 | 3 | 0,04% | *Beta vulgaris* |
| 1,409 | 993,218 | 395 | 0,03% | *Solanum tuberosum* |
| 1,404 | 575,496 | 295 | 0,05% | *Solanum lycopersicum* |
| 637 | 281,476 | 131 | 0,04% | *Vitis vinifera* |
| 1,398 | 478,378 | 380 | 0,07% | *Medicago truncatula* |
| 967 | 248,622 | 373 | 0,15% | *Populus trichocarpa* |
| 519 | 123,180 | 140 | 0,11% | *Theobroma cacao* |
| 1,216 | 252,864 | 546 | 0,21% | *Corchorus capsularis* |
| 1,291 | 97,576 | 483 | 0,49% | *Brassica rapa* |
| 1,469 | 160,486 | 381 | 0,23% | *Brassica oleracea* |
| 833 | 116,145 | 466 | 0,40% | *Arabidopsis lyrata* |
| 1,429 | 43,442 | 722 | 1,66% | *Arabidopsis thaliana* |
| 1,079 | 116,189 | 251 | 0,21% | *Musa acuminata* |
| 859 | 154,873 | 150 | 0,09% | *Setaria italica* |
| 1,359 | 428,546 | 338 | 0,07% | *Sorghum bicolor* |
| 1,797 | 1,791,946 | 1,915 | 0,10% | *Zea mays* |
| 1,278 | 106,890 | 180 | 0,16% | *Brachypodium distachyon* |
| 2,345 | 2,253,478 | 276 | 0,01% | *Aegilops tauschii* |
| 1,987 | 2,250,260 | 187 | 0,008% | *Triticum urartu* |
| 7,857 | 13,445,260 | 2,168 | 0,01% | *Triticum aestivum* |
| 899 | 112,058 | 220 | 0,19% | *Leersia perrieri* |
| 756 | 178,734 | 113 | 0,06% | *Oryza nivara* |
| 443 | 380,576 | 245 | 0,06% | *Oryza brachyantha* |
| 1,179 | 129,481 | 181 | 0,13% | *Oryza longistaminata* |
| 684 | 201,805 | 124 | 0,06% | *Oryza glumipatula* |
| 785 | 204,639 | 128 | 0,06% | *Oryza rufipogon* |
| 701 | 695,468 | 528 | 0,07% | *Oryza glaberrima* |
| 834 | 190,594 | 190 | 0,09% | *Oryza punctata* |
| 957 | 193,661 | 210 | 0,10% | *Oryza meridionalis* |
| 948 | 185,442 | 161 | 0,08% | *Oryza barthii* |
| 926 | 603,608 | 1,087 | 0,18% | *Oryza sativa* |
| 1,285 | 232,693 | 520 | 0,22% | *Oryza indica* |

ncRNAs | TEs | % TE overlapping



| | | | | |
|---|---|---|---|---|
| 1,278 | 106,890 | 180 | 0,16% | *Brachypodium distachyon* |
| 2,345 | 2,253,478 | 276 | 0,01% | *Aegilops tauschii* |
| 1,987 | 2,250,260 | 187 | 0,008% | *Triticum urartu* |
| 7,857 | 13,445,260 | 2,168 | 0,01% | *Triticum aestivum* |
| 899 | 112,058 | 220 | 0,19% | *Leersia perrieri* |
| 756 | 178,734 | 113 | 0,06% | *Oryza nivara* |
| 443 | 380,576 | 245 | 0,06% | *Oryza brachyantha* |
| 1,179 | 129,481 | 181 | 0,13% | *Oryza longistaminata* |
| 684 | 201,805 | 124 | 0,06% | *Oryza glumipatula* |
| 785 | 204,639 | 128 | 0,06% | *Oryza rufipogon* |
| 701 | 695,468 | 528 | 0,07% | *Oryza glaberrima* |
| 834 | 190,594 | 190 | 0,09% | *Oryza punctata* |
| 957 | 193,661 | 210 | 0,10% | *Oryza meridionalis* |
| 948 | 185,442 | 161 | 0,08% | *Oryza barthii* |
| 926 | 603,608 | 1,087 | 0,18% | *Oryza sativa* |
| 1,285 | 232,693 | 520 | 0,22% | *Oryza indica* |

https://doi.org/10.1093/database/bay078 - Figure 1



| | LTR | TIR | LINE | SINE | Unknown | Total |
|---|---|---|---|---|---|---|
| **tRNA** | 2959 | 192 | 1 | 14 | 303 | 3469 |
| **rRNA** | 2962 | 1389 | 25 | 7 | 1082 | 5465 |
| **snRNA** | 1763 | 117 | 14 | 2 | 120 | 2016 |
| **Sense-intronic** | 764 | 20 | – | – | 207 | 991 |
| **Pre-miRNA** | 696 | 190 | 3 | 3 | 94 | 986 |
| **snoRNA** | 529 | 287 | 2 | 2 | 49 | 869 |
| **SRP** | 391 | 70 | – | – | 2 | 463 |
| **Antisense** | 70 | 2 | 1 | – | 16 | 89 |
| **RNase MRP** | 2 | – | – | – | – | 2 |
| **Total** | 10 136 | 2 267 | 46 | 28 | 1873 | 14 350 |

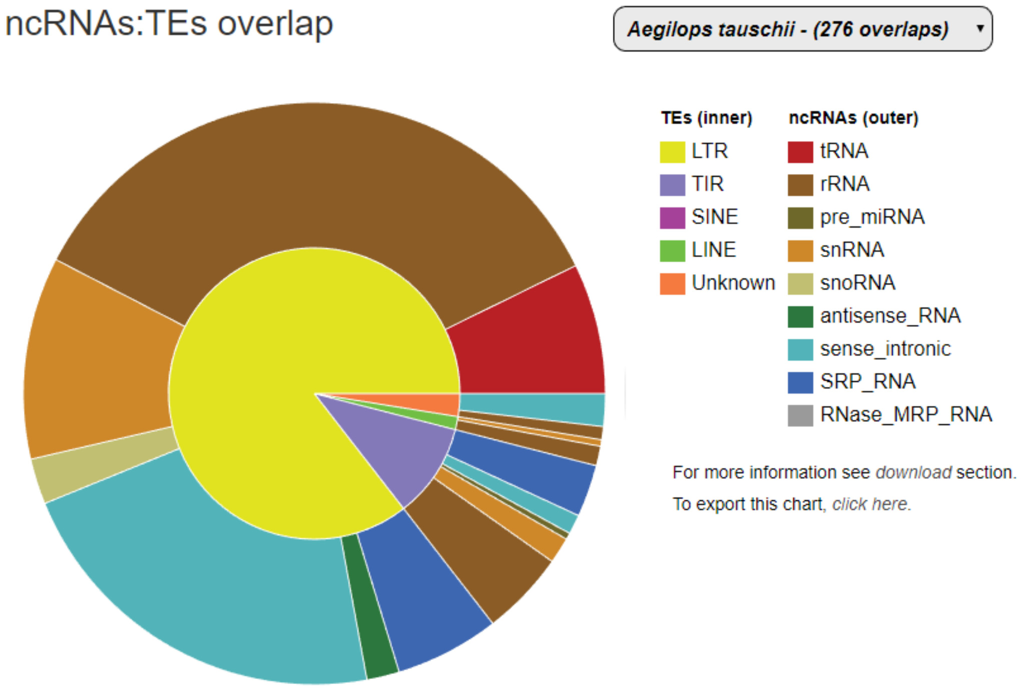
# Overall features

- Overlap features:
  - ~41% of the overlaps are among 4 genomes:
    - *Triticum aestivum; Zea mays;* ***Oryza sativa; Arabidopsis thaliana***

- Overlap records between ncRNA:TE **| Public data available**
  - Visualization tools (Charts by genome and jBrowse)
  - **14.350 overlaps in 40 genomes**

- Scripts developed in Perl + Bash
  - Automatic updates
  - ZendFramework2, Php7, MySQL, CSS3, HTML5, JavaScript and Debian9.

# PlaNC-TE – *Detailed info*



ncRNAs:TEs overlap

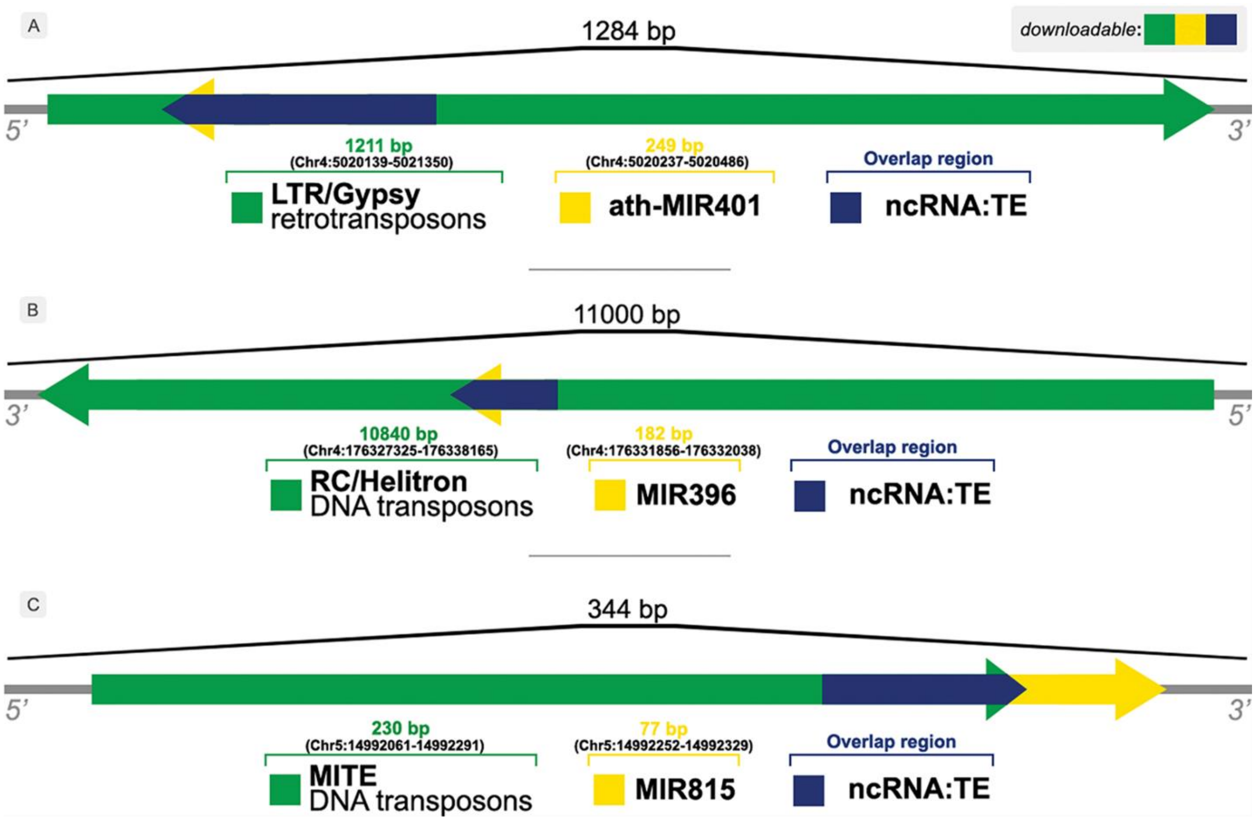
*Aegilops tauschii - (276 overlaps)* ▾

For more information see *download* section.

To export this chart, *click here*.

- First page
- Select genome
- View ncRNA:TE overlaps

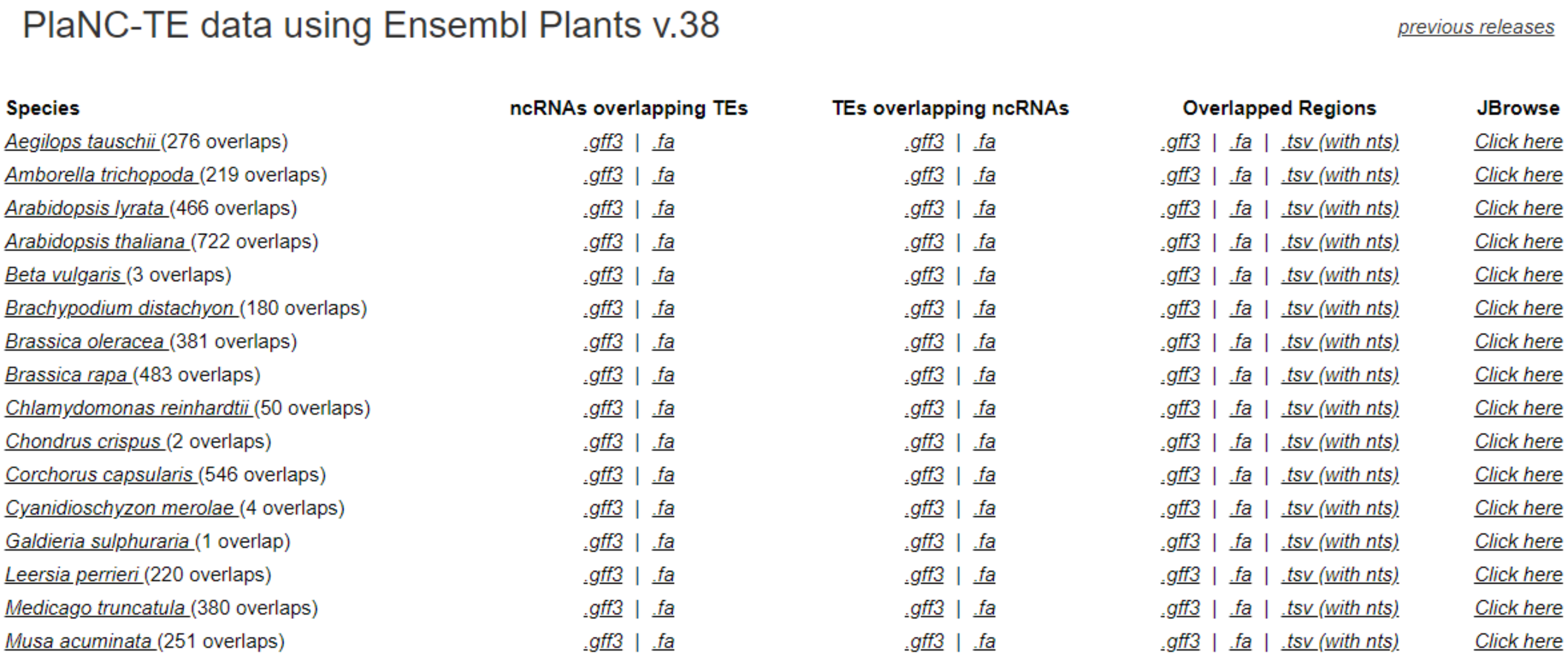
# PlaNC-TE – *Examples of nc:TEs*



From: PlaNC-TE: a comprehensive knowledgebase of non-coding RNAs and transposable elements in plants

Database (Oxford). 2018;2018. doi:10.1093/database/bay078

# PlaNC-TE – *Download*

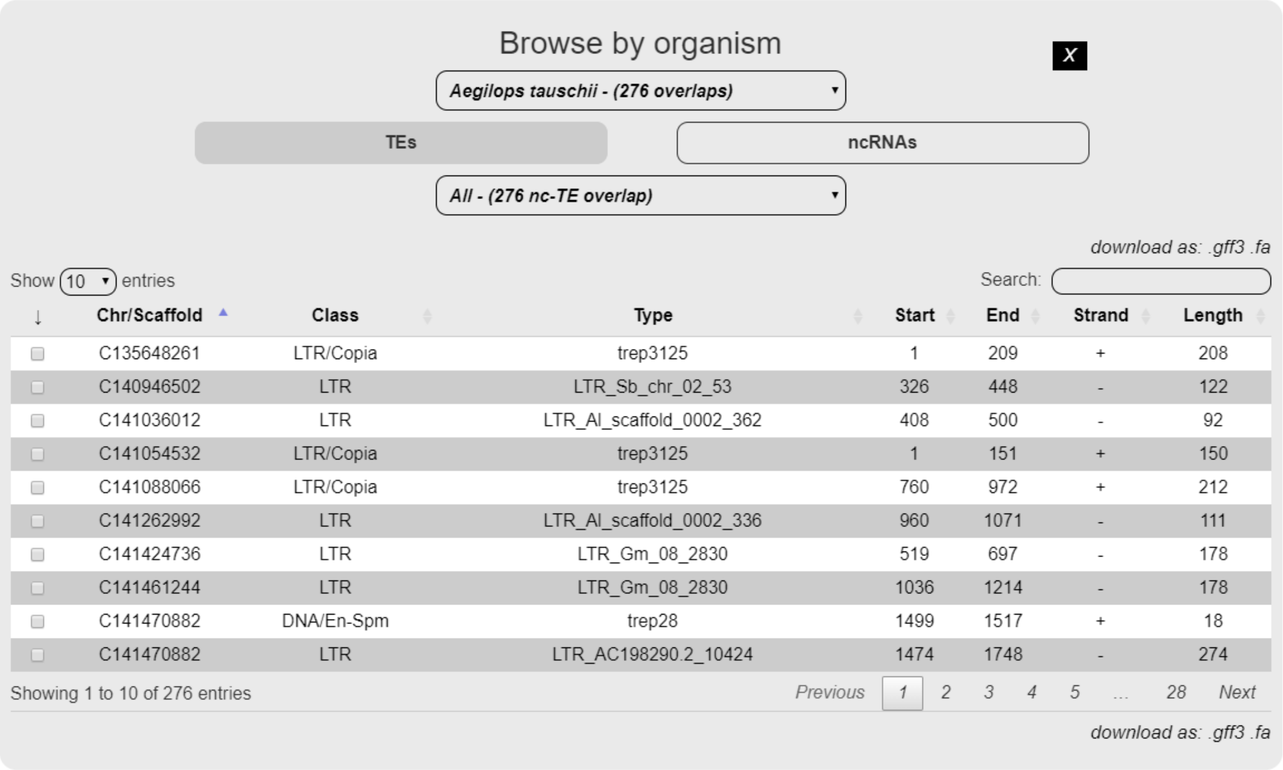


## PlaNC-TE data using Ensembl Plants v.38

*previous releases*

| Species | ncRNAs overlapping TEs | TEs overlapping ncRNAs | Overlapped Regions | JBrowse |
|---|---|---|---|---|
| *Aegilops tauschii* (276 overlaps) | *.gff3* \| *.fa* | *.gff3* \| *.fa* | *.gff3* \| *.fa* \| *.tsv (with nts)* | *Click here* |
| *Amborella trichopoda* (219 overlaps) | *.gff3* \| *.fa* | *.gff3* \| *.fa* | *.gff3* \| *.fa* \| *.tsv (with nts)* | *Click here* |
| *Arabidopsis lyrata* (466 overlaps) | *.gff3* \| *.fa* | *.gff3* \| *.fa* | *.gff3* \| *.fa* \| *.tsv (with nts)* | *Click here* |
| *Arabidopsis thaliana* (722 overlaps) | *.gff3* \| *.fa* | *.gff3* \| *.fa* | *.gff3* \| *.fa* \| *.tsv (with nts)* | *Click here* |
| *Beta vulgaris* (3 overlaps) | *.gff3* \| *.fa* | *.gff3* \| *.fa* | *.gff3* \| *.fa* \| *.tsv (with nts)* | *Click here* |
| *Brachypodium distachyon* (180 overlaps) | *.gff3* \| *.fa* | *.gff3* \| *.fa* | *.gff3* \| *.fa* \| *.tsv (with nts)* | *Click here* |
| *Brassica oleracea* (381 overlaps) | *.gff3* \| *.fa* | *.gff3* \| *.fa* | *.gff3* \| *.fa* \| *.tsv (with nts)* | *Click here* |
| *Brassica rapa* (483 overlaps) | *.gff3* \| *.fa* | *.gff3* \| *.fa* | *.gff3* \| *.fa* \| *.tsv (with nts)* | *Click here* |
| *Chlamydomonas reinhardtii* (50 overlaps) | *.gff3* \| *.fa* | *.gff3* \| *.fa* | *.gff3* \| *.fa* \| *.tsv (with nts)* | *Click here* |
| *Chondrus crispus* (2 overlaps) | *.gff3* \| *.fa* | *.gff3* \| *.fa* | *.gff3* \| *.fa* \| *.tsv (with nts)* | *Click here* |
| *Corchorus capsularis* (546 overlaps) | *.gff3* \| *.fa* | *.gff3* \| *.fa* | *.gff3* \| *.fa* \| *.tsv (with nts)* | *Click here* |
| *Cyanidioschyzon merolae* (4 overlaps) | *.gff3* \| *.fa* | *.gff3* \| *.fa* | *.gff3* \| *.fa* \| *.tsv (with nts)* | *Click here* |
| *Galdieria sulphuraria* (1 overlap) | *.gff3* \| *.fa* | *.gff3* \| *.fa* | *.gff3* \| *.fa* \| *.tsv (with nts)* | *Click here* |
| *Leersia perrieri* (220 overlaps) | *.gff3* \| *.fa* | *.gff3* \| *.fa* | *.gff3* \| *.fa* \| *.tsv (with nts)* | *Click here* |
| *Medicago truncatula* (380 overlaps) | *.gff3* \| *.fa* | *.gff3* \| *.fa* | *.gff3* \| *.fa* \| *.tsv (with nts)* | *Click here* |
| *Musa acuminata* (251 overlaps) | *.gff3* \| *.fa* | *.gff3* \| *.fa* | *.gff3* \| *.fa* \| *.tsv (with nts)* | *Click here* |

# PlaNC-TE – *Search & Browse*



Browse by organism

*Aegilops tauschii - (276 overlaps)*

*All - (276 nc-TE overlap)*

*download as: .gff3 .fa*

Show 10 entries

Search:

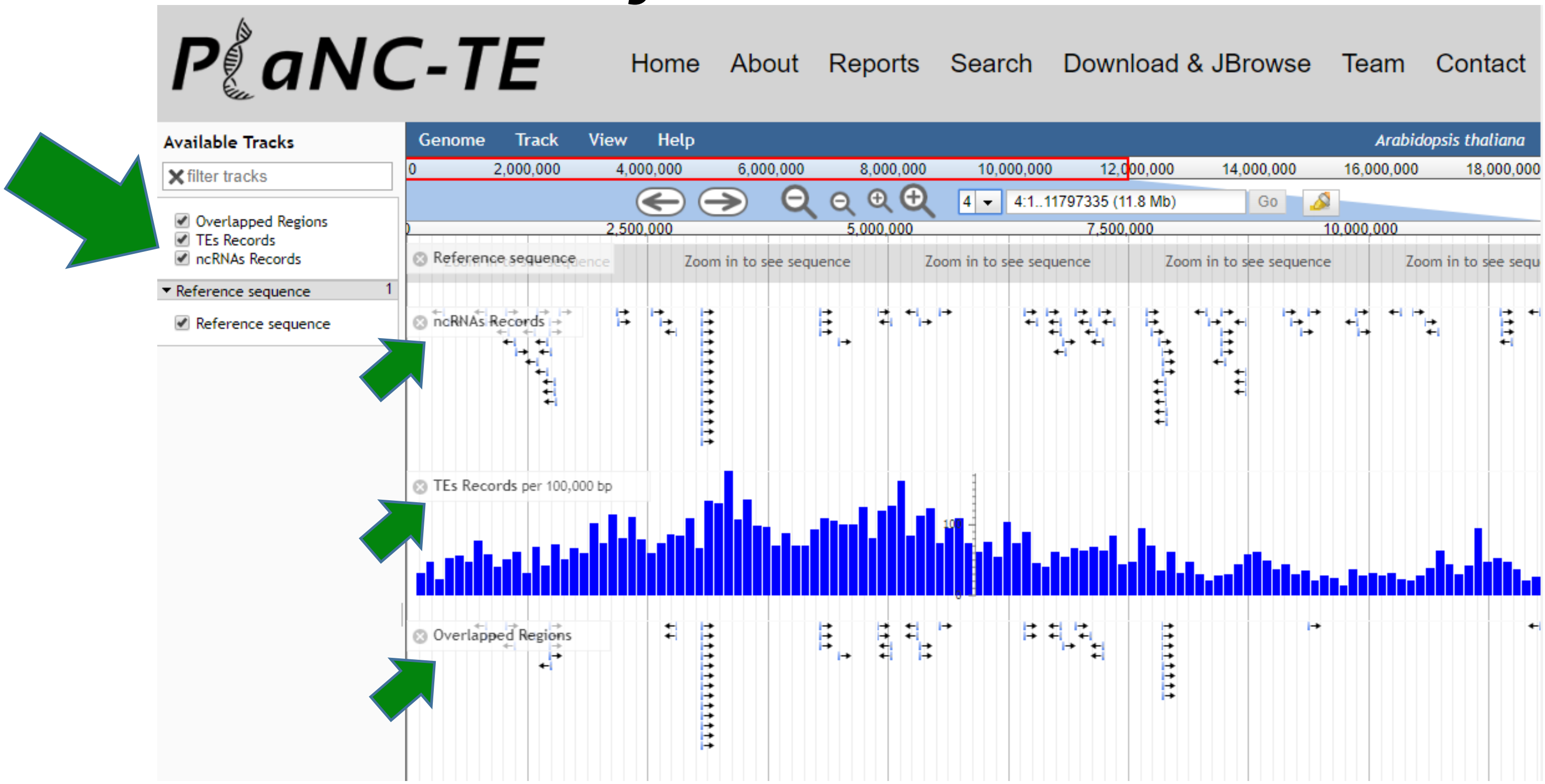
| ↓ | Chr/Scaffold ▲ | Class | Type | Start | End | Strand | Length |
|---|---|---|---|---|---|---|---|
| □ | C135648261 | LTR/Copia | trep3125 | 1 | 209 | + | 208 |
| □ | C140946502 | LTR | LTR_Sb_chr_02_53 | 326 | 448 | - | 122 |
| □ | C141036012 | LTR | LTR_Al_scaffold_0002_362 | 408 | 500 | - | 92 |
| □ | C141054532 | LTR/Copia | trep3125 | 1 | 151 | + | 150 |
| □ | C141088066 | LTR/Copia | trep3125 | 760 | 972 | + | 212 |
| □ | C141262992 | LTR | LTR_Al_scaffold_0002_336 | 960 | 1071 | - | 111 |
| □ | C141424736 | LTR | LTR_Gm_08_2830 | 519 | 697 | - | 178 |
| □ | C141461244 | LTR | LTR_Gm_08_2830 | 1036 | 1214 | - | 178 |
| □ | C141470882 | DNA/En-Spm | trep28 | 1499 | 1517 | + | 18 |
| □ | C141470882 | LTR | LTR_AC198290.2_10424 | 1474 | 1748 | - | 274 |

Showing 1 to 10 of 276 entries

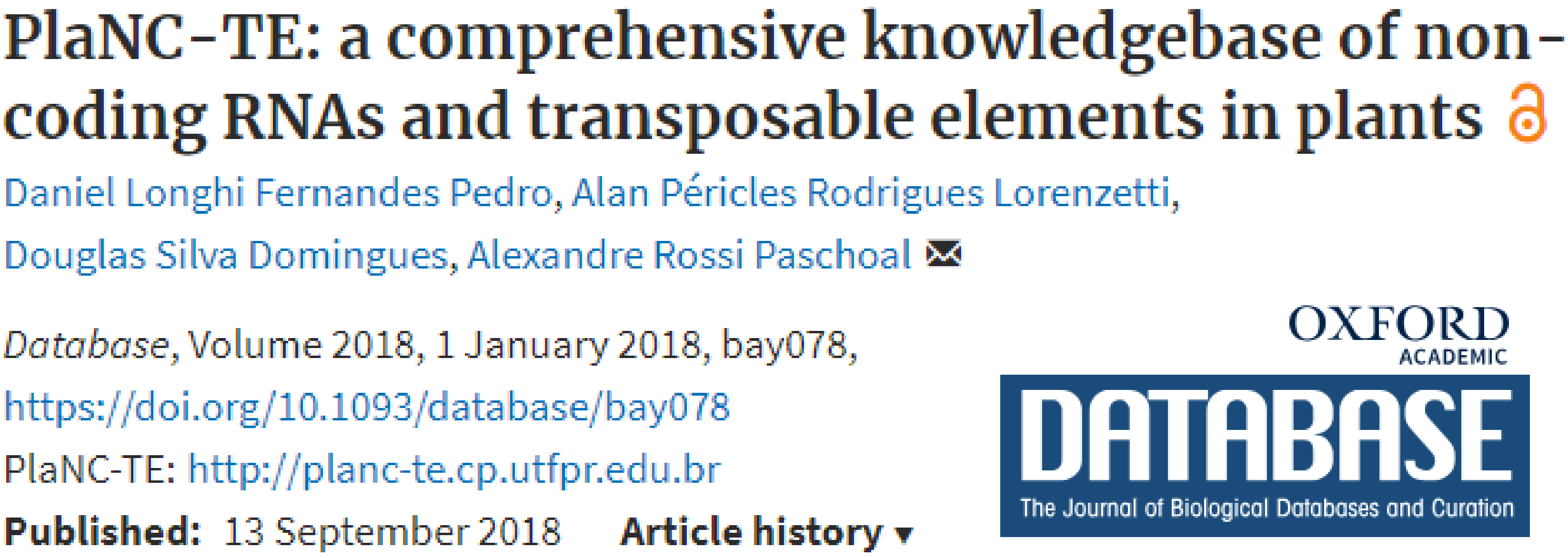
*Previous 1 2 3 4 5 … 28 Next*

*download as: .gff3 .fa*

# PlaNC-TE – *jBrowser*



# http://planc-te.cp.utfpr.edu.br



## PlaNC-TE: a comprehensive knowledgebase of non-coding RNAs and transposable elements in plants

Daniel Longhi Fernandes Pedro, Alan Péricles Rodrigues Lorenzetti, Douglas Silva Domingues, Alexandre Rossi Paschoal ✉

*Database*, Volume 2018, 1 January 2018, bay078, https://doi.org/10.1093/database/bay078

PlaNC-TE: http://planc-te.cp.utfpr.edu.br

**Published:** 13 September 2018 **Article history ▾**

OXFORD ACADEMIC

DATABASE

The Journal of Biological Databases and Curation

# So…

- Extend PlanTE-MIR to all plant genomes available in Ensembl (53 species)

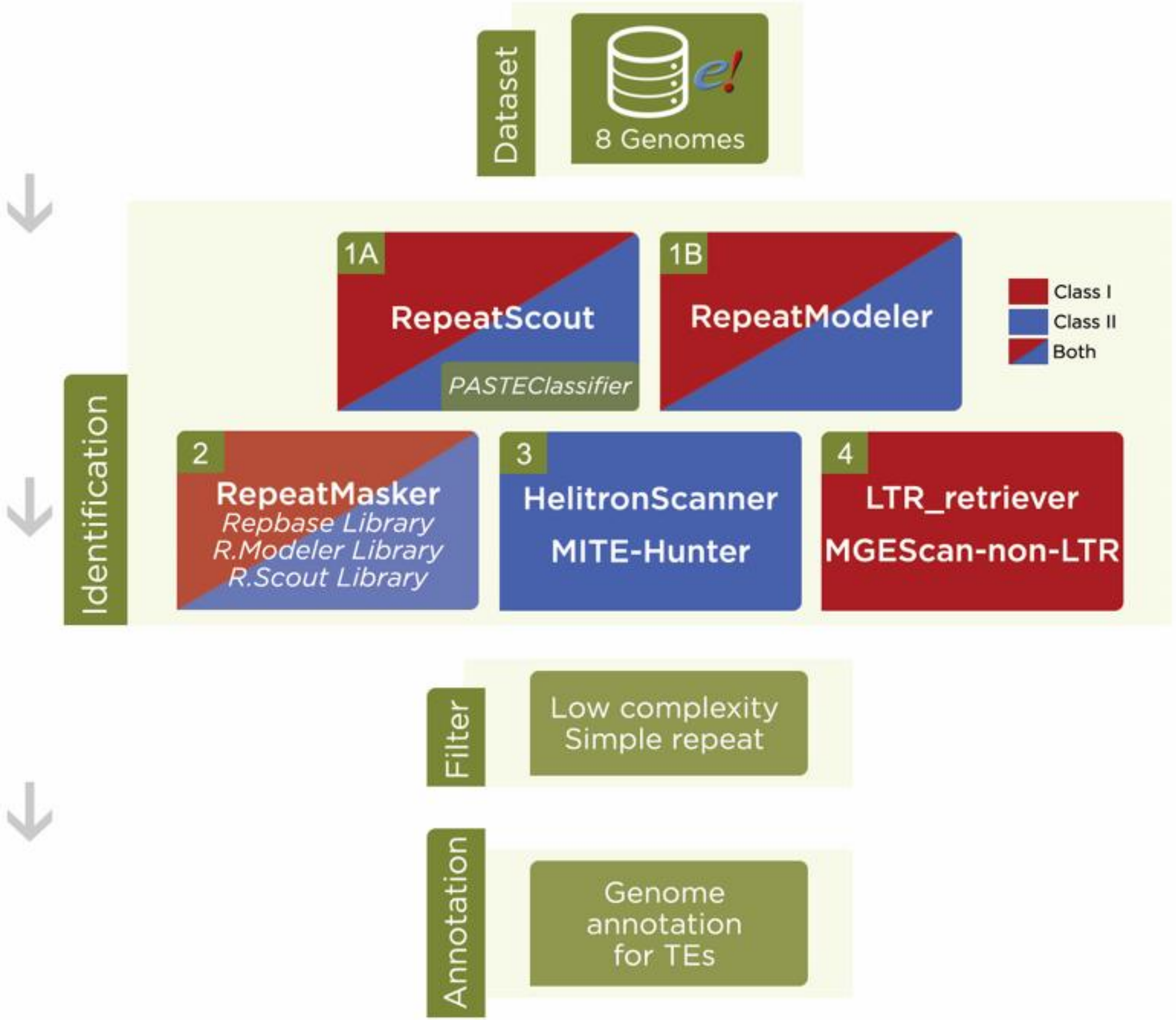
# So…

- Extend PlanTE-MIR to all plant genomes available in Ensembl (53 species)
    - But TEs data are available only for 40 genomes!

# So…

- Extend PlanTE-MIR to all plant genomes available in Ensembl (53 species)
  - But TEs data are available only for 40 genomes!
  - If TEs are a major component of genomes, something is wrong!

# Phase 2: Re-annotation of TEs in complete plant genomes


Dataset
e!
8 Genomes
Identification
1A
RepeatScout
PASTEClassifier
1B
RepeatModeler
Class I
Class II
Both
2
RepeatMasker
Repbase Library
R.Modeler Library
R.Scout Library
3
HelitronScanner
MITE-Hunter
4
LTR_retriever
MGEScan-non-LTR
Filter
Low complexity
Simple repeat
Annotation
Genome
annotation
for TEs

# Initial analyses raised the number of TE entries in genomes

| Plant genomes | Ensembl Plants | Our approach |
|---|---|---|
| *A. lyrata* | 116,145 | 391,425 |
| *A. thaliana* | 43,442 | 63,879 |
| *B. vulgaris* | 6,295 | 984,280 |
| *B. rapa* | 97,576 | 434,231 |
| *C. sativus* | - | 176,333 |
| *M. acuminata* | 116,189 | 637,112 |
| *P. trichocarpa* | 248,622 | 864,831 |
| *V. vinifera* | 281,476 | 834,298 |

# Take home message

- We still need standardization and better annotation (at least of TEs) in plant genomes
- Up to now, TE annotation is heavily based in alignment: curated datasets can be an starting point for other computational approaches
- Long-term goal: Are any specific characteristic (*feature*) of TE and/or ncRNA that distinguish ncRNA:TE association?

# Team/Acknowledgements

Daniel L. F. Pedro

Dr. Alexandre Rossi Paschoal

Alan P. Rodrigues Lorenzetti

Tharcísio Amorim

Bioinformática UTFPR

UTFPR
UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
CAMPUS CORNÉLIO PROCÓPIO

unesp

Funding

CNPq
Conselho Nacional de Desenvolvimento Científico e Tecnológico

CAPES

# Thank you!

e-mail: douglas.domingues@unesp.br