the plant journal

SEB Society for Experimental Biology

The Plant Journal (2014) **80,** 582–591

doi: 10.1111/tpj.12656

Young but not relatively old retrotransposons are preferentially located in gene-rich euchromatic regions in tomato (*Solanum lycopersicum*) plants

Yingxiu Xu[1,2] and Jianchang Du[3,*]

[1]Bioinformatics Group, Institute of Industrial Crops, Jiangsu Academy of Agricultural Sciences, Nanjing 210014, China,
[2]Key laboratory of Cotton and Rapeseed (Nanjing), Ministry of Agriculture, Nanjing 210014, China, and
[3]State Key Laboratory of Cotton Biology, Institute of Cotton Research of Chinese Academy of Agriculture Sciences, Anyang 455000, China

SUMMARY

Long terminal repeat (LTR) retrotransposons are the major DNA components of flowering plants. They are generally enriched in pericentromeric heterochromatin regions of their host genomes, which could result from the preferential insertion of LTR retrotransposons and the low effectiveness of purifying selection in these regions. To estimate the relative importance of the actions of these two factors on their distribution pattern, the LTR retrotransposons in *Solanum lycopersicum* (tomato) plants were characterized at the genome level, and then the distribution of young elements was compared with that of relatively old elements. The current data show that old elements are mainly located in recombination-suppressed heterochromatin regions, and that young elements are preferentially located in the gene-rich euchromatic regions. Further analysis showed a negative correlation between the insertion time of LTR retrotransposons and the recombination rate. The data also showed there to be more solo LTRs in genic regions than in intergenic regions or in regions close to genes. These observations indicate that, unlike in many other plant genomes, the current LTR retrotransposons in tomatoes have a tendency to be preferentially located into euchromatic regions, probably caused by their severe suppression of activities in heterochromatic regions. These elements are apt to be maintained in heterochromatin regions, probably as a consequence of the pericentromeric effect in tomatoes. These results also indicate that local recombination rates and intensities of purifying selection in different genomic regions are largely responsible for structural variation and non-random distribution of LTR retrotransposons in tomato plants.

- Can the authors convince you that selection is at play?

- Are the right that LTR retros are “preferentially”located in euchromatin

Brief summary

- Comparison of the distribution and age estimates of LTR retrotransposons in the genome of tomato (*Solanum lycopersicum*).

- Young elements tend to reside in euchromatin.

- Older elements tend to accumulate in heterochromatin.

- Age is negatively correlated with local recombination rate.

Brief summary

- Tomato is a good model for investigating these dynamics because..
  - (i) high-quality genome, which facilitates the annotation of LTR retrotransposons;
  - (ii) large proportion of the genome is heterochromatic (~80%).
  - (iii) A fine-scale genetic map is available.

Identified LTR retros

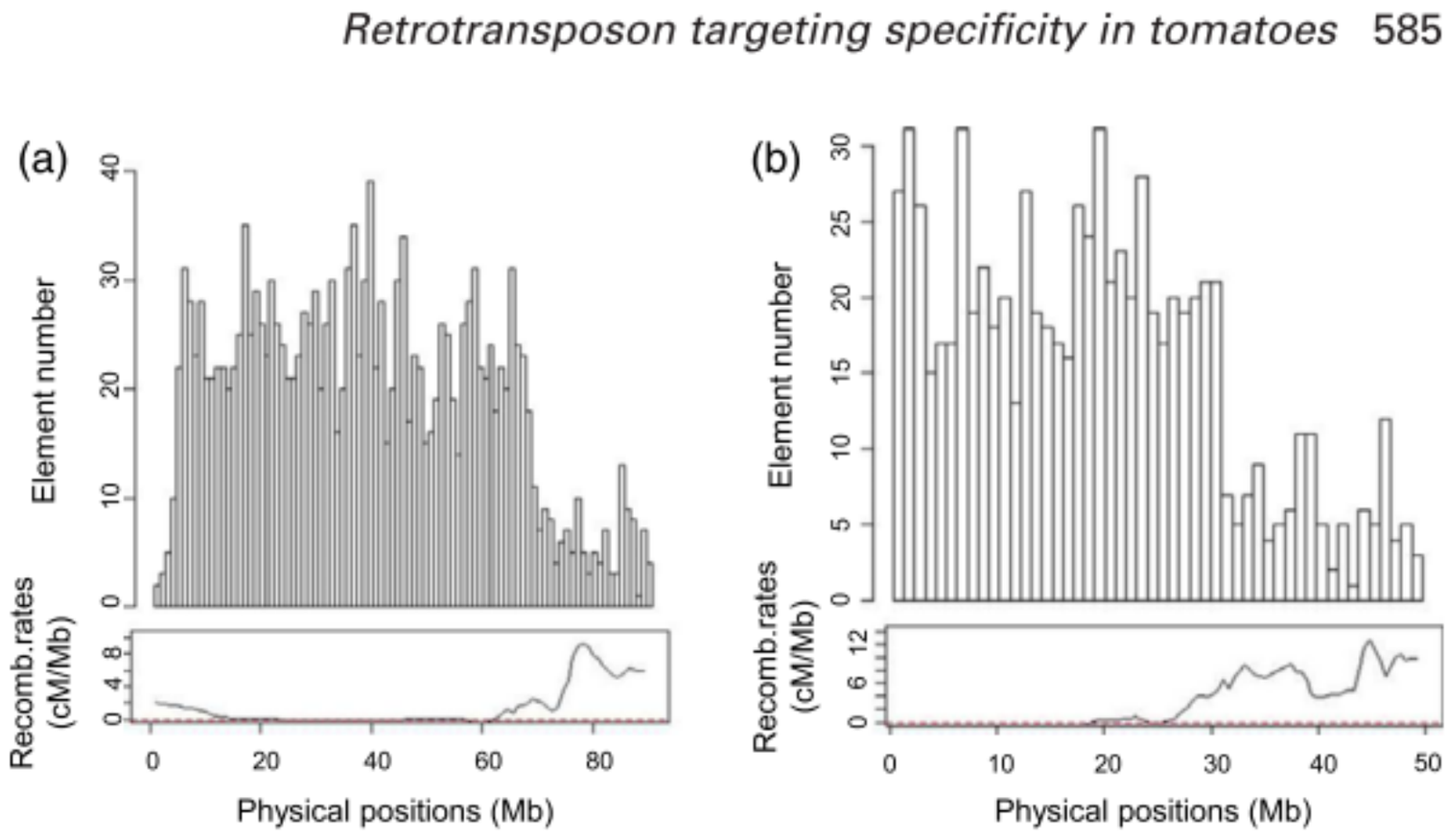
- Used LTR_STRUC and LTR sequences were used a queries in a CROSS_MATCH analysis.

- Identified >15,000 elements: “Each element was manually inspected to avoid incorrect annotation from the program”.

- Recombination rate was estimated by comparing physical and genetic based maps (>2000 markers).

Distribution of LTR retros vs recombination rate

- Used 1 Mbp windows
- Estimated LTR copy number
- Copy-number lower in regions of high recombination.



LTR retros vs recombination rate

Table 2 Distribution, insertion time, and genetic recombination rate of long terminal repeat (LRT) retrotransposons in the tomato genome

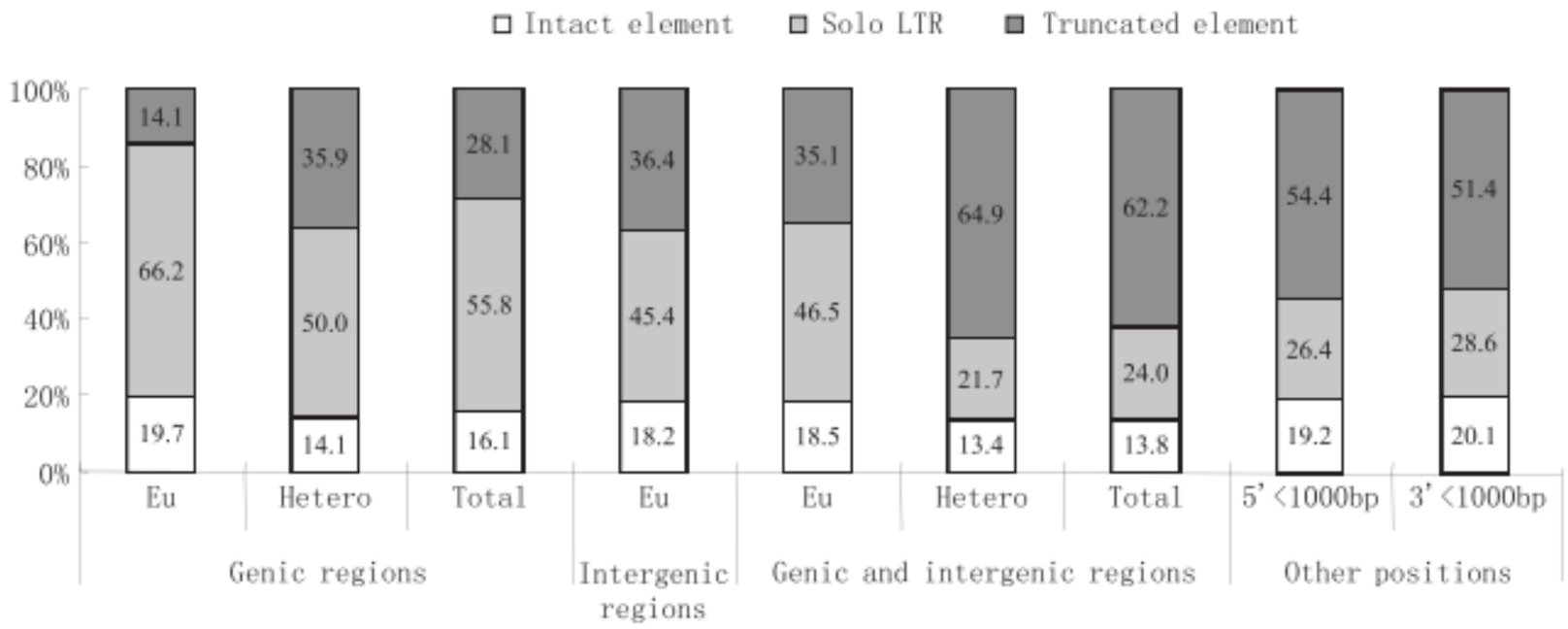
	Intact element			Solo LTR			Truncated element			Total		
	Eu.	Hetero.	Eu./ hetero.	Eu.	Hetero.	Eu./ hetero.	Eu.	Hetero.	Eu./ hetero.	Eu.	Hetero.	Eu./hetero.
No. of elements	252	1808	0.14	642	2936	0.21	487	8784	0.05	1381	13 528	0.10
Element density	1.45	3.09	0.47	3.69	5.01	0.72	2.80	15.00	0.18	7.93	23.10	0.34
Insertion time	1.18	2.86	0.41	/	/	/	/	/	/	1.18	2.86	0.41
Recombination rate	5.58	0.46	12.13	6.19	0.44	14.17	5.26	0.31	17.16	5.75	0.36	16.14

eu., euchromatin; hetero., heterochromatin.

- Important to note that the age estimates of intact elements are different in eu- vs heterochromatin.

- Is this number a mean? why no SD/SE?

LTR retros: genomic distribution



- Authors claim that truncated retros ~ removal via IR, whereas solo LTR ~ removal by UR.

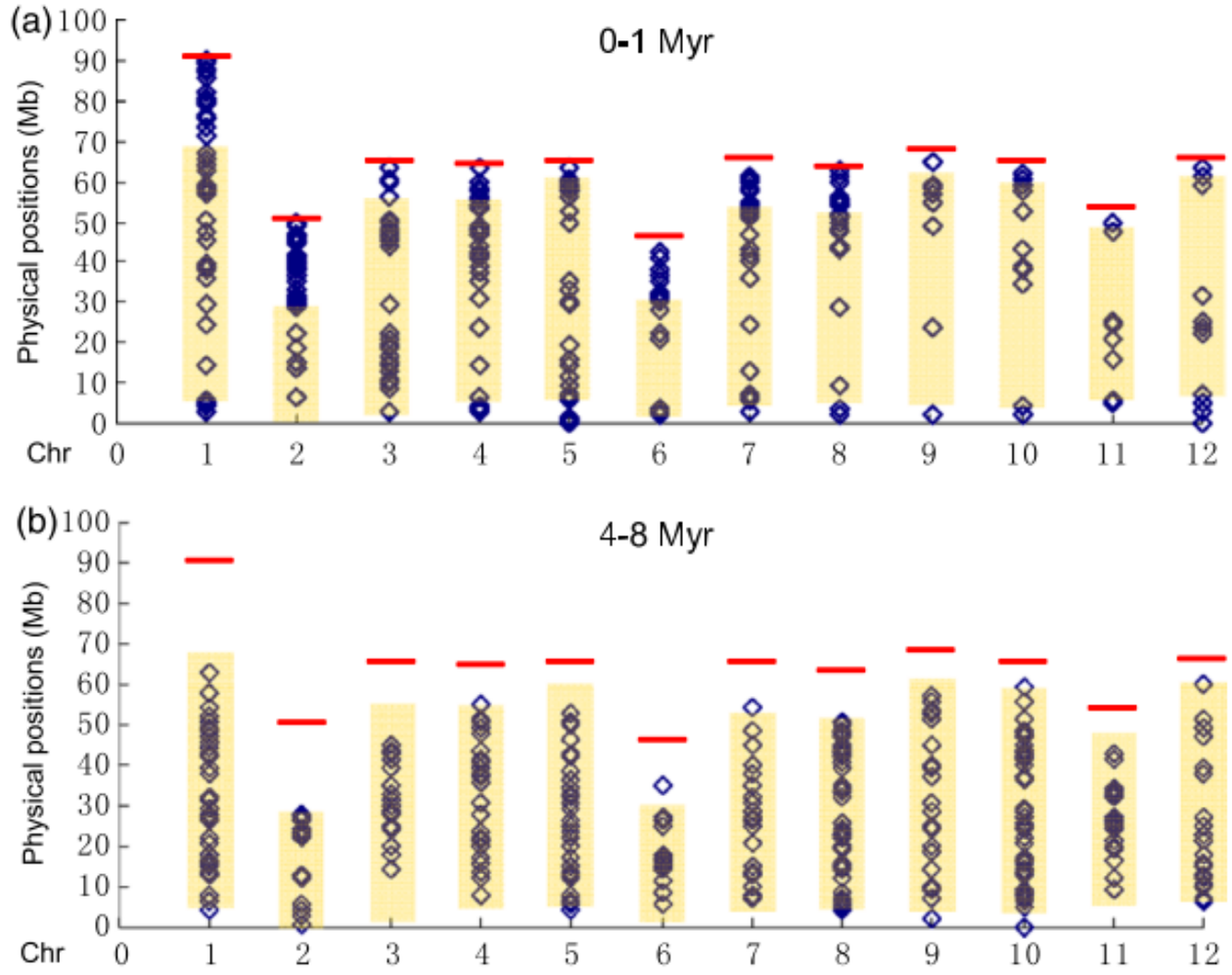
- Yet later they claim that most removal is by UR?? So solo LTRs must dominate.

Recent LTR insertions localize to euchromatin

- Younger retros appear to be located in euchromatin

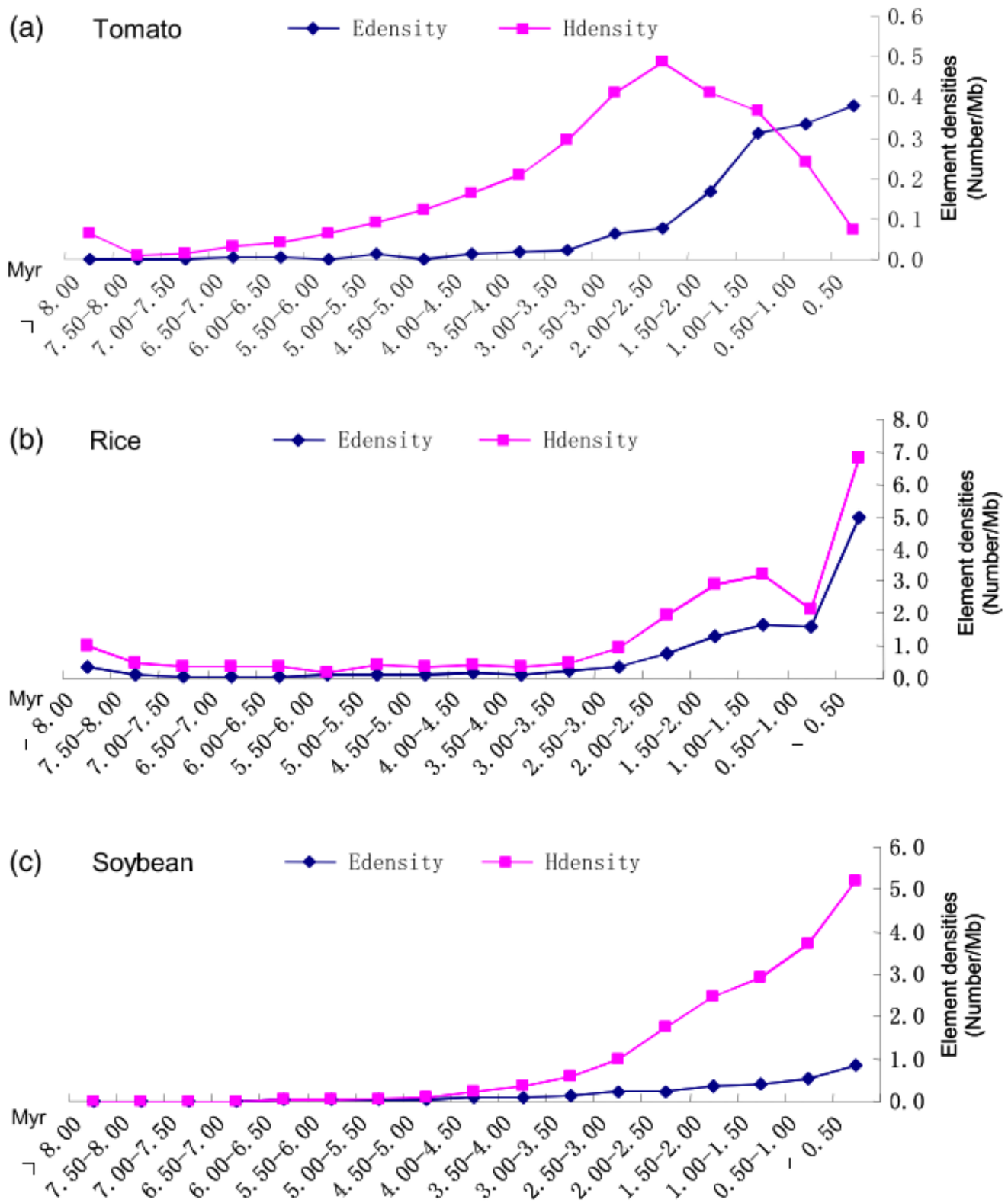
- Older insertions are more or less absent in euchromatin

- Authors tested this using a T-test??



- Highlight the difference between tomato and other genomes

- clearly shows that young LTR retros are enriched in euchromatin in tomato but NOT rice or soybean.



Recombination rate vs insertion time

- Negative correlation between insertion time and recombination rate
- Only for younger insertion times.

Table 3 Correlation analysis of insertion time and genetic recombination rates

Categories	r^a	P^b
All intact elements	−0.269	<0.0005
Insertions 0–1.0 Mya	−0.306	<0.0005
Insertions 1.0–2.0 Mya	−0.210	<0.0005
Insertions 2.0–3.0 Mya	0.010	>0.50
Insertions 3.0–4.0 Mya	0.036	>0.50
Insertions 4.0–5.0 Mya	0.046	>0.50
Insertions 5.0–8.0 Mya	0.034	>0.50
Copia elements inserted within 3.0 Mya	−0.385	<0.0005
Gypsy elements inserted within 3.0 Mya	−0.471	<0.0005
Euchromatic elements inserted within 3.0 Mya	−0.201	<0.001
Heterochromatic elements inserted within 3.0 Mya	−0.194	<0.0005

Mya, million years ago.
[a]Pearson correlation coefficients.
[b]*P* values.

Conclusions

- Why is there a negative correlation between recombination rate and insertion age?
- Why is tomato different from soy and rice?
- Is this result entirely surprising?