

LTR retrotransposons are systematically younger than non-LTR retrotransposons in the *Drosophila melanogaster* genome.

Casey M. Bergman & Douda Bensasson

**Faculty of Life Sciences
University of Manchester**

**casey.bergman@manchester.ac.uk
douda.bensasson@manchester.ac.uk**

A brief introduction to transposable element (TE) evolution: the current paradigm

- **TEs are mobile DNA sequences, intra-genomic parasites**
- **Impact genome organization & evolution**
- **Transposition rates \gg excision rates**
- **Equilibrium maintained by natural selection**
- **Mode of natural selection is debated**
 - **deleterious effects of TE insertion**
 - **deleterious effects of transposition**
 - **deleterious effects of TE-mediated ectopic recombination**
- * **TE insertions observed at low frequency in nature**

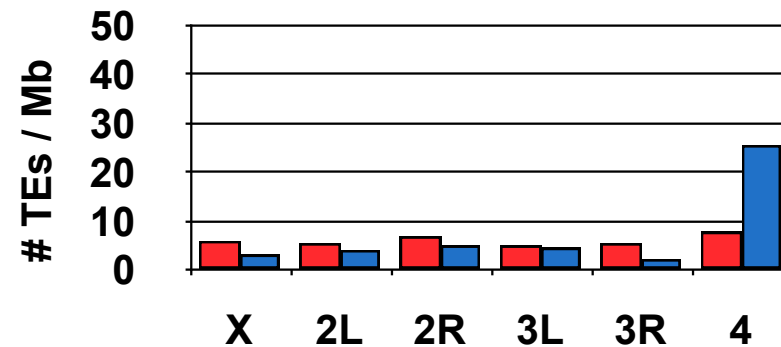
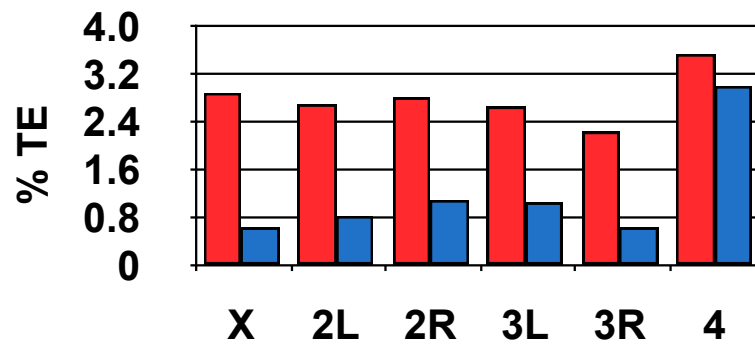
Subclasses of retrotransposon in the genus *Drosophila* differ in structure and abundance



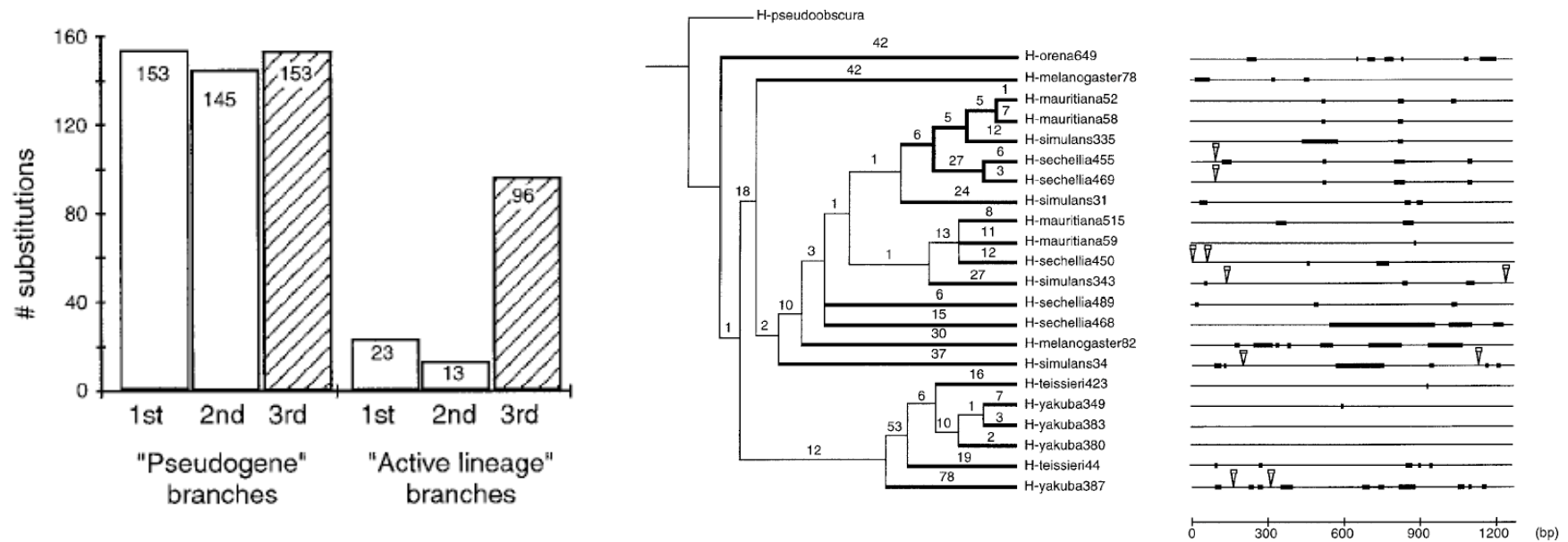
Long Terminal Repeat (LTR)



non-LTR (LINE-like)



Terminal branches in non-LTR retrotransposon genealogies exhibit a 'pseudogene-like' mode of sequence evolution



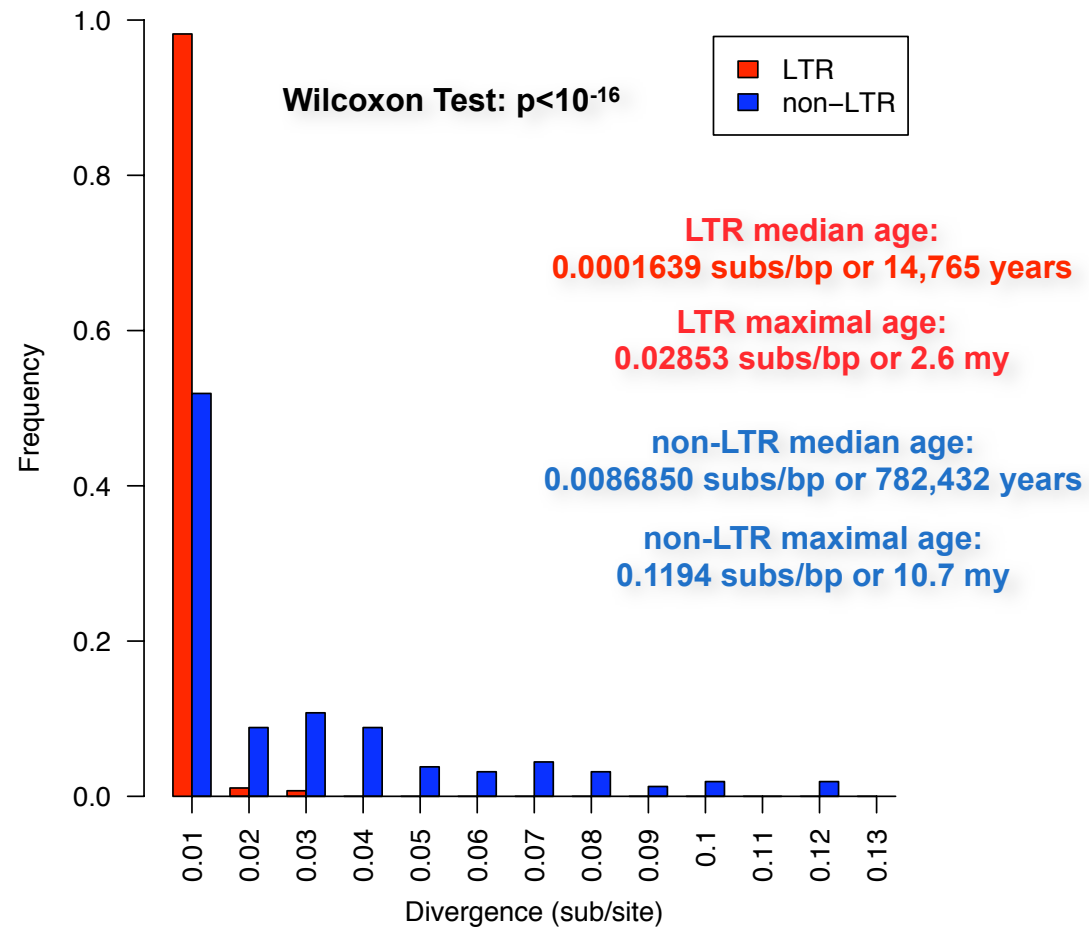
Petrov and Hartl (1998) *Mol. Biol. Evol.* 15:293-302
 Blumenstiel and Hartl (2002) *Mol. Biol. Evol.* 19:2211-2225

Most retrotransposon families exhibit a ‘pseudogene-like’ mode of sequence evolution

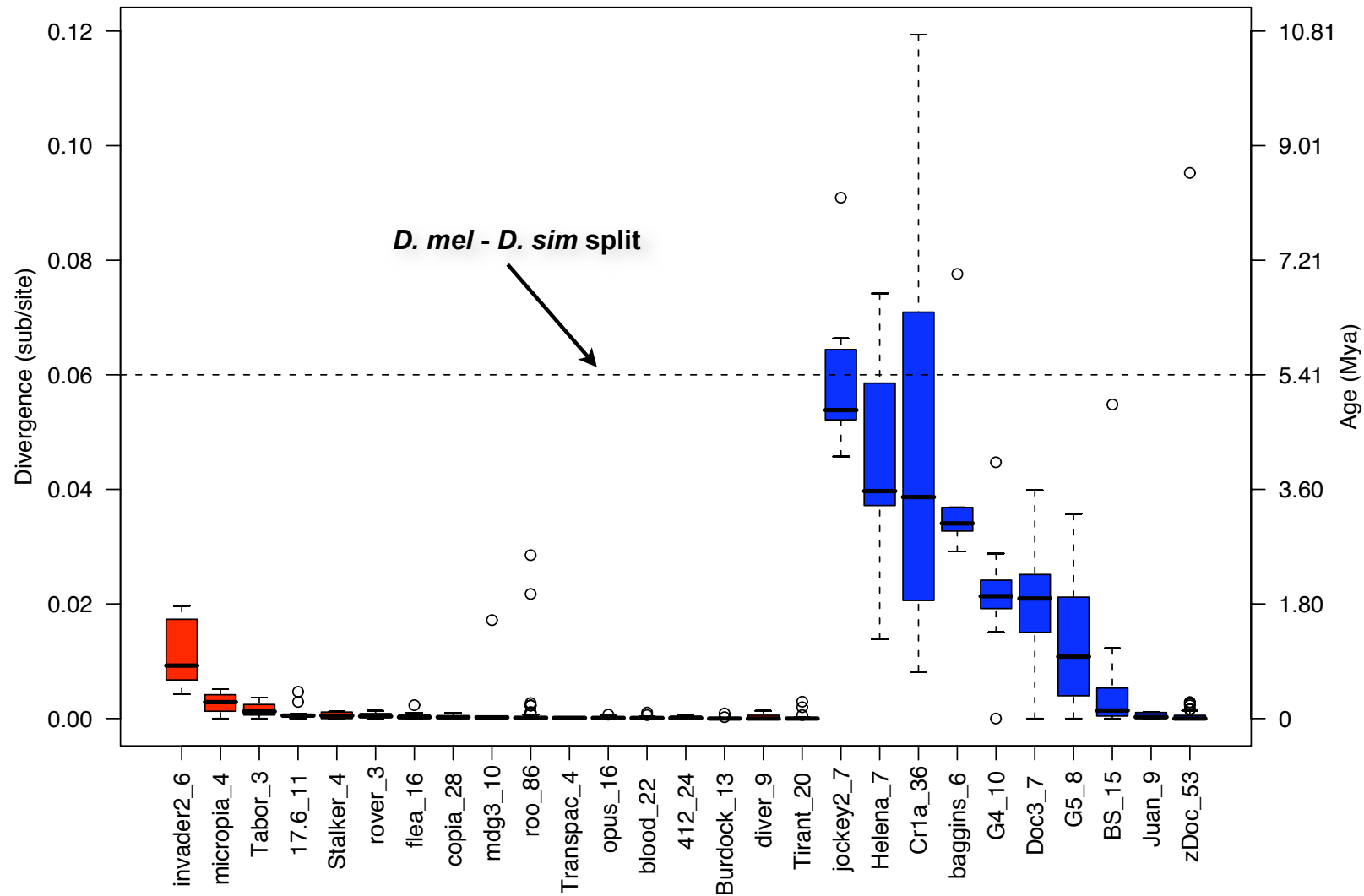
	# families	# elements	Total bp Surveyed	1st	2nd	3rd	Total Point Sub.	P (Ho)	ψ
All non-LTR	19	377	836,819	1,515	1,424	1,917	4,884	3.56E-24	N
ψ non-LTR	10	158	336,748	791	746	781	2,341	0.192	Y
All LTR	27	385	1,973,013	677	603	1,120	2,420	2.18E-44	N
ψ LTR	17	279	1,491,867	272	267	307	851	0.159	Y
Grand Total	46	762	2,809,832	2,192	2,027	3,037	7,304	5.18E-61	N
Total ψ	27	437	1,828,615	1,063	1,013	1,088	3,192	0.06	Y

59% of retrotransposon families exhibit a pseudogene-like mode of evolution on terminal branches

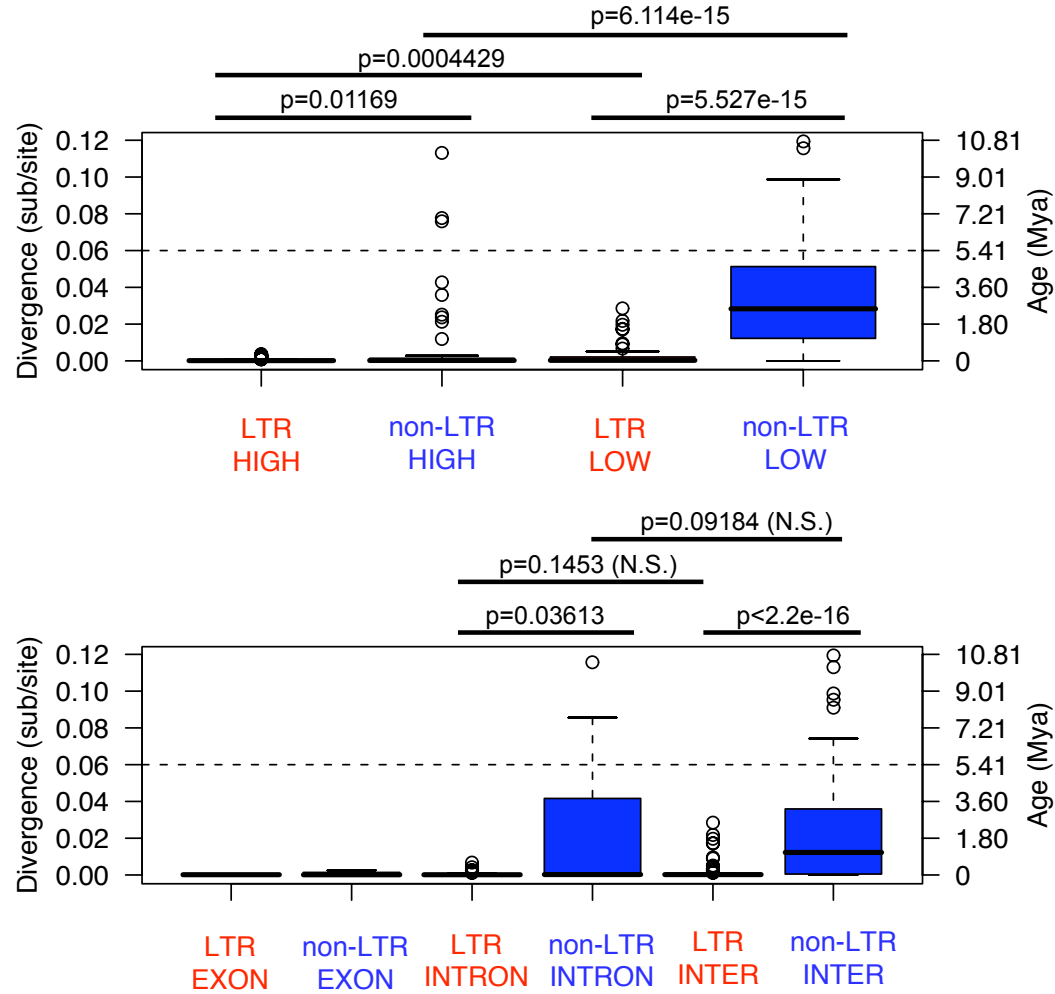
LTR elements are systematically younger than non-LTR elements in the *D. melanogaster* genome



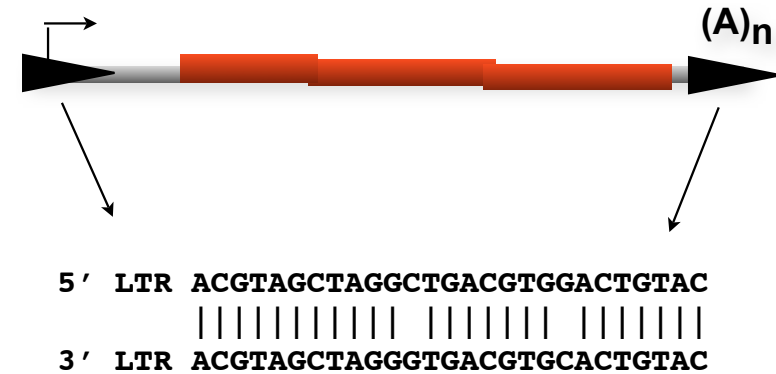
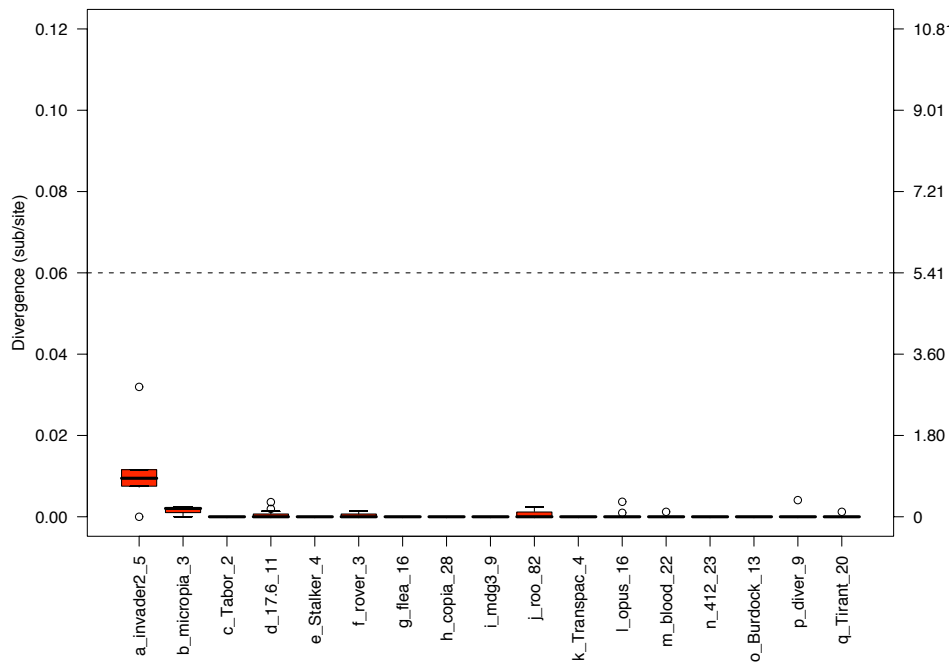
LTR and non-LTR families exhibit drastically different age structures



Differences in age between LTR and non-LTR elements are not cause by recombination or transcription



Intra-element LTR-LTR comparisons support age estimates based on terminal branch length

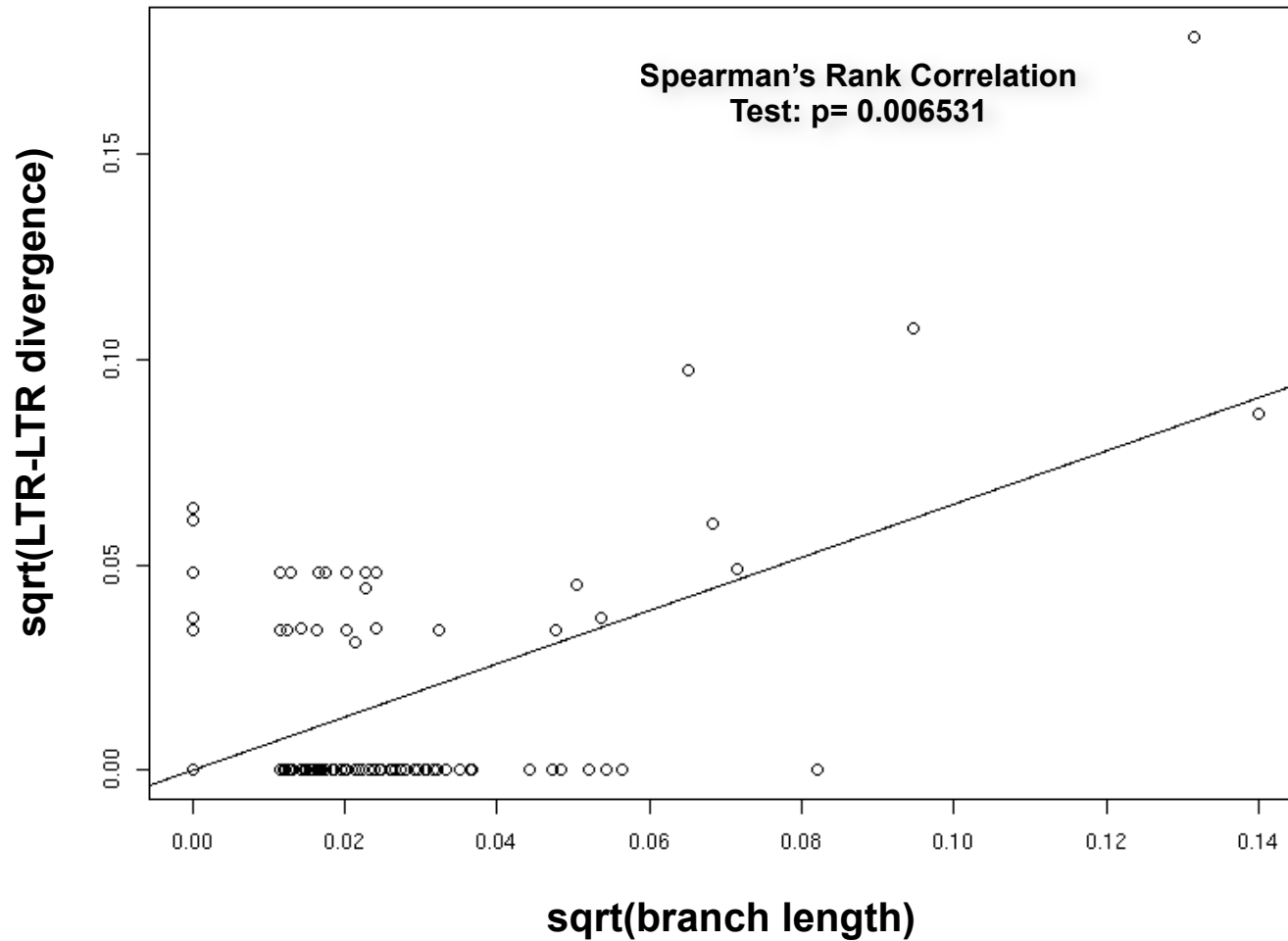


$$T = D / (2 * r)$$

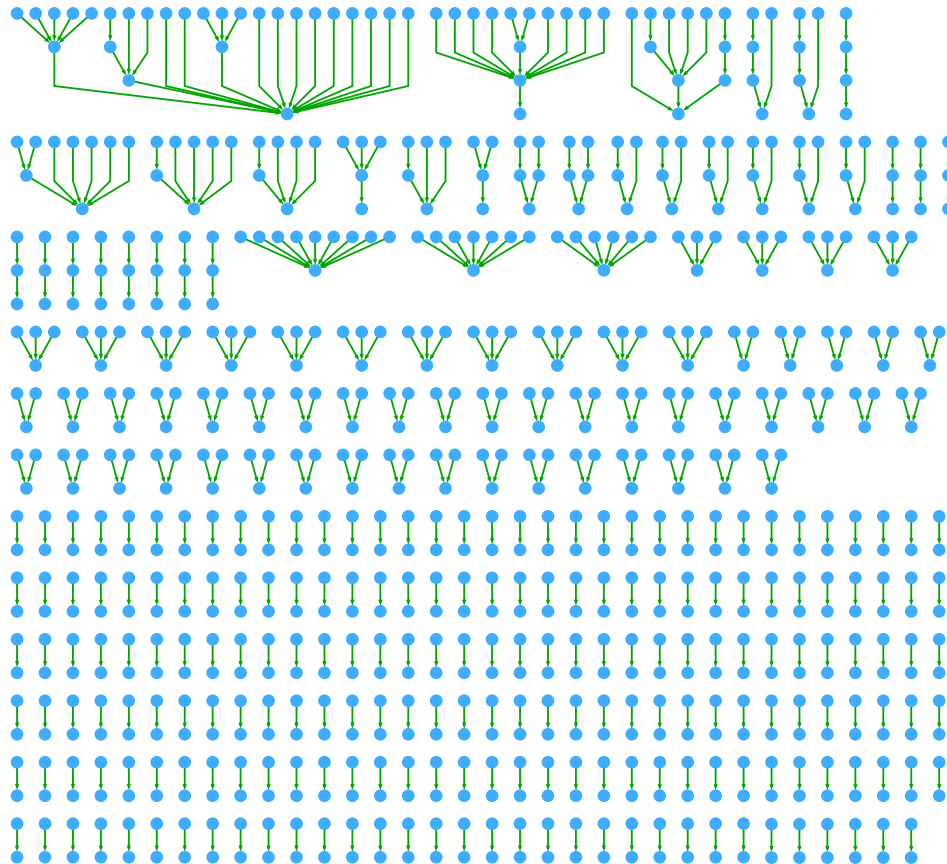
T - absolute time
D - 5' vs. 3' LTR divergence
r - neutral substitution rate (0.0111/my)

see also Bowen and McDonald (2001) *Genome Res* 11:1527-1540

Intra-element LTR-LTR age estimates correlate with terminal branch age estimates



Genome-wide patterns of TE nesting support the recent LTR insertion hypothesis



	LTR	non-LTR	outer total
LTR	151	39	190
non-LTR	46	46	92
inner Total	197	85	282

Using length of TEs for outer & number of TEs for inner to calculate expected number of nests.

G-test for all 4 types of nest
 $p=0.000188$

G-test for different subclass nests
 $p=0.017$

Implications of recent LTR insertion - I

Current paradigm is built on a non-random sample of TE families, which is biased towards LTR elements

Table 1. *Transposition rates in the mutation accumulation lines*

Element family	Element copy no. ^a						No. of lines analysed	No. of new insertions	Transposition rate per element ^c	Upper 95% confidence limit
	X	2-Bal	2-MA	3	4	Total ^b				
* <i>mdg-1</i> (2181)	4.0	7	6	9.3	0.0	17.8	22	3	1.77×10^{-4}	
* <i>opus</i> (2217)	2.0	4	7	11.9	0.0	18.4	21	1	5.98×10^{-5}	
* <i>copia</i>	4.0	9	12	14.8	0.0	27.3	21	10	4.03×10^{-4}	
* 1731 (2158)	0.0	0	0	1.0	0.0	1.0	20	0		3.44×10^{-4}
* 297	3.5	16	8	19.1	1.0	33.9	21	4	1.30×10^{-4}	
* 412	1.0	12	12	8.8	0.0	21.3	23	0		1.41×10^{-4}
<i>jockey</i> (2161)	4.0	8	24	16.2	1.0	35.2	21	1	3.13×10^{-5}	
? 2244	2.8	9	5	12.2	1.0	21.6	20	1	5.35×10^{-5}	
<i>pogo</i>	2.8	3	5	6.1	0.0	11.5	20	0		3.01×10^{-4}
* <i>roo</i> (B104)	16.7	19	24	30.2	0.0	60.1	20	7	1.35×10^{-4}	
* 17.6 (2210)	1.6	3	1	8.5	1.0	12.3	21	0		2.68×10^{-4}
Total	42.4	90	104	138.1	4.0	260.4	20.9	27	1.15×10^{-4}	

Implications of recent LTR insertion - II

Current paradigm assumes transposition-selection equilibrium, which may be violated for LTR elements

Table 4 Models of selection that most closely fit published X chromosomal proportions of occupied sites

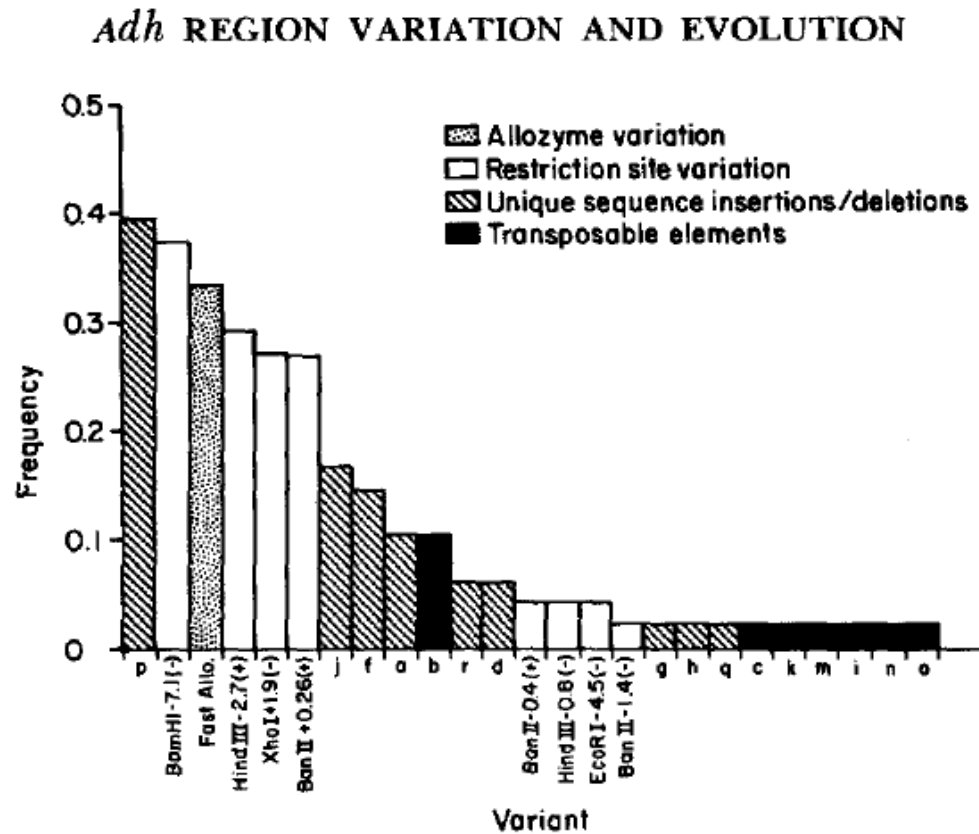
Retrotransposon	Most closely fitting model					
<i>297</i>	Random ^a	Random ^b	Random ^e	Random ⁱ		
<i>412</i>	Insertion ^a	Ectopic Exchange 1 or 2 ^b	Insertion ^e	Insertion ⁱ	Ectopic Exchange 1 ^k	Insertion ^l
<i>copia</i>	Insertion ^b	Insertion ^d	Insertion ^f	Insertion ^g	Insertion ^h	
<i>mdg-1</i>	Insertion ^a	Insertion ^c	Ectopic Exchange 1 ^d	Insertion ^j	Insertion ^l	
<i>mdg-3</i>	Random ^a	Random ^c	Random ^j			
<i>opus</i>	Random ^a	Insertion ⁱ				
<i>roo</i>	Random ^a	Ectopic Exchange 2 or Random ^e	Ectopic Exchange 2 ⁱ			

^aCarr et al. this study); ^bStrobel et al. (1979); ^cBelyaeva et al. (1984); ^dBiémont (1986); ^eMontgomery et al. (1987); ^fYamaguchi et al. (1987); ^gBiémont and Gautier (1988); ^hBiémont (1992); ⁱCharlesworth et al. (1992); ^jBiémont et al. (1994); ^kAulard et al. (1995); ^lVieira and Biémont (1996)

Carr et al. (2002) *Chromosoma* 110:511-518

Implications of recent LTR insertion - III

Current paradigm interprets low TE frequency as evidence for selection, but low frequency may reflect age since insertion



Aquadro *et al.* (1986) *Genetics* 114:1165-1190

Conclusions

- Both LTR and non-LTR elements can be used as pseudogene-like systems, although many retrotransposon families show evidence of selective constraint on terminal branches.
- LTR elements are systematically younger than non-LTR elements in *D. melanogaster*.
- Age of major LTR activity is on same time scale as colonization of Europe, whereas non-LTR elements have accumulated in waves since since speciation with *D. simulans*.
- LTR and non-LTR elements are older in regions of low recombination, but no differences in age are observed between intronic and intergenic regions.
- Terminal branch & intra-element LTR ages are correlated.
- Patterns of TE nesting support recent LTR insertion.
- “Sure facts” of TE evolution (equilibrium & low frequency) need to be re-evaluated & population genetic models need to include TE age

Acknowledgements

Dmitri Petrov, Minal Mehta
NERC, Royal Society

