Tandem repeat variation in partially selfing populations
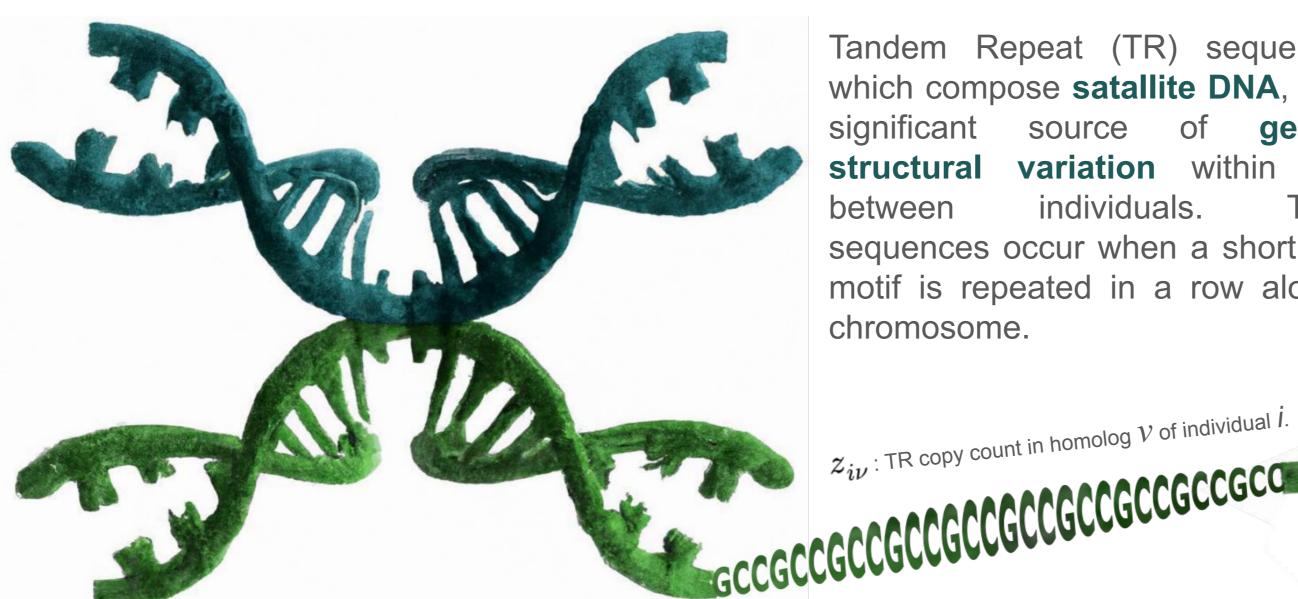
Vítor Sudbrack & Charles Mullon

Department of Ecology and Evolution (DEE)
University of Lausanne, Switzerland 🇨🇭

vitor.sudbrack@unil.ch   www.sudbrack.me   @vitorsudbrack
charles.mullon@unil.ch   www.lab-mullon.github.io   @MullonDlm

Tandem Repeat (TR) sequences, which compose **satellite DNA**, are a significant source of **genetic structural variation** within and between individuals. These sequences occur when a short DNA motif is repeated in a row along a chromosome.

$z_{i\nu}$: TR copy count in homolog ν of individual i.

GCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCC

Variation in TR sequences is thought to reflect a balance between purifying selection, genetic drift, **amplification**, and **unequal recombination** [1].

Adults carry two homologous sequences

Saltatory Amplification
with probability μ
per new chromatid [3]

z_{i1}
$z_{i1}^{(A)}$
$z_{i2}^{(A)}$
z_{i2}

Unequal recombination between homologs
with probability γ [2]

z_{i1}
$z_{i1}^{(R)}$
$z_{i2}^{(R)}$
z_{i2}

TR sequences are assumed to be costly, such that the **fecundity** of an individual is

$$f_i = 1 - s_{\mathrm{a}}\left(z_{i1} + z_{i2}\right)$$

Adults
Fixed population size, N

Number of gametes is proportional to fecundity

Life cycle

Syngamy, selfing with probability α

Gametes

Existing TR models assume panmixia and haploidy [1-3], but **inbreeding** affects selection, drift, and unequal recombination on TRs [4]. How do these effects combine to influence **TR variation**?

> We use **mathematical analyses** and **individual-based simulations** to study how inbreeding (due to partial selfing) shapes the polymorphism of TR sequences in diploids.

Our mathematical model assumes a regime of weak selection, rare amplification and large population size, i.e. $s_{\mathrm{a}} \sim \delta, \mu \sim \delta, N^{-1} \sim \delta$ for a small δ, and arbitrary rate of unequal crossover. We track the evolution of 3 quantities:

MEAN

$$\bar{z} = \frac{1}{N}\sum_{i=1}^{N}\frac{z_{i1}+z_{i2}}{2}$$

VARIANCE WITHIN individuals

$$\sigma_{\mathrm{W}}^2 = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{z_{i1}-z_{i2}}{2}\right)^2$$

VARIANCE BETWEEN individuals

$$\sigma_{\mathrm{B}}^2 = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{z_{i1}+z_{i2}}{2}-\bar{z}\right)^2$$

REFS.: [1] Charlesworth, B., Sniegowski, P., & Stephan, W. (1994). The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature*. [2] Stephan, W. (1986). Recombination and the evolution of satellite DNA. *Genetics Research*. [3] Stephan, W. (1987). Quantitative variation and chromosomal location of satellite DNAs. *Genetics Research*. [4] Buschiazzo, E., & Gemmell, N. J. (2006). The rise, fall and renaissance of microsatellites in eukaryotic genomes. *Bioessays*.

Key findings

Decomposition of variance

The total variance in TR copy count is

$$\sigma^2 = \sigma_{\mathrm{B}}^2 + \sigma_{\mathrm{W}}^2$$

and as selfing becomes more frequent, the proportion of variation between individuals relative to within individuals increases,

$$\sigma_{\mathrm{B}}^2 = \left(\frac{1+F_{\mathrm{IS}}}{2}\right)\sigma^2$$

where F_{IS} is the **coefficient of inbreeding** and depends on both selfing and unequal recombination rates:

$$F_{\mathrm{IS}} = \frac{\alpha\left(1-\frac{\gamma}{2}\right)^2}{2-\alpha\left(1-\frac{\gamma}{2}\right)^2} + \mathcal{O}(\delta)$$
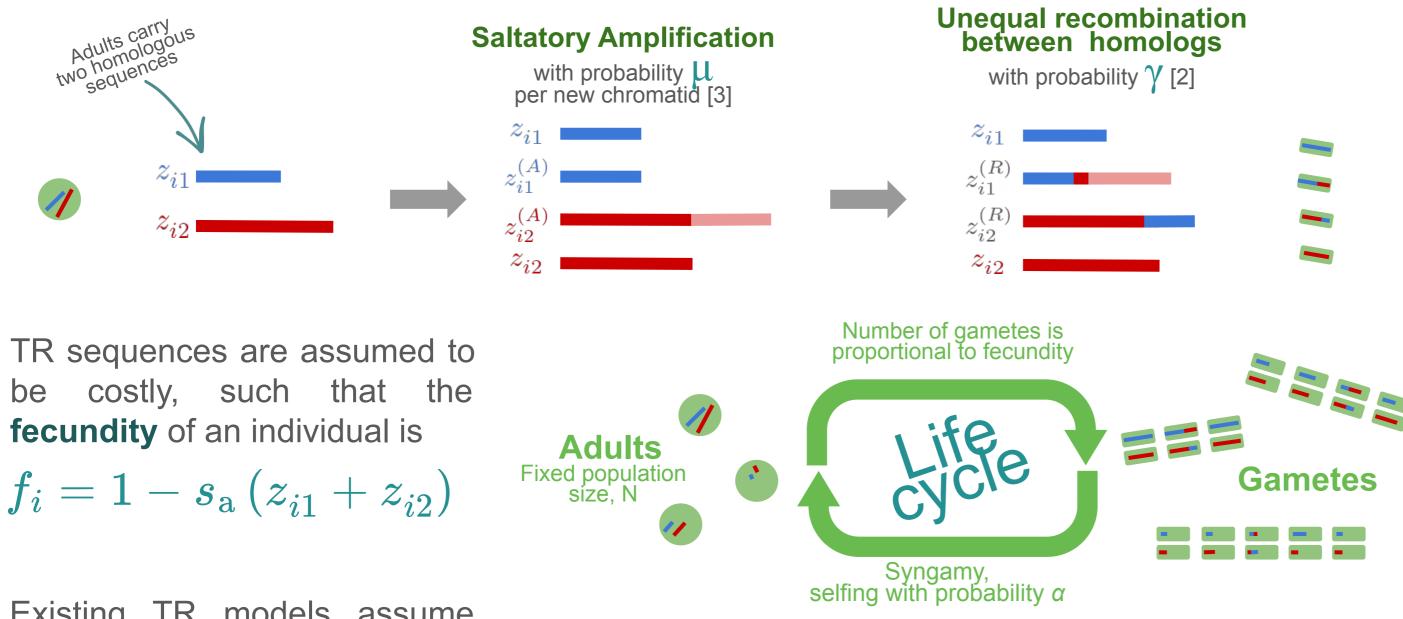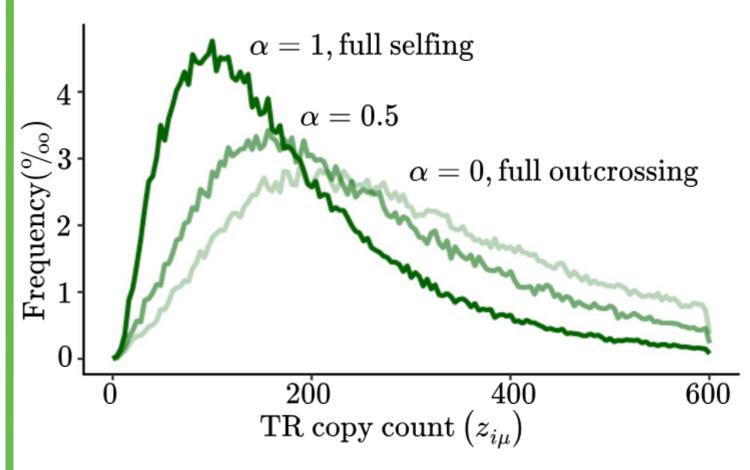
Change in variance

The change in total variance of TRs depends on the **interaction between unequal recombination and selfing**:

$$\mathbb{E}[\Delta\sigma^2] = -\frac{1}{2}\gamma\sigma_{\mathrm{W}}^2$$
$$+\frac{1}{12}\gamma\left[\bar{z}^2 + \sigma_{\mathrm{B}}^2 - 1\right] + \mathcal{O}(\delta)$$

such that unequal recombination tends to make different sequences more similar, and similar sequences more different. The first case occurs more frequently in outcrossed offspring, while the second occurs more frequently in selfed offspring.

Change in mean

The mean number of TR copies per chromosome reflects a **balance between amplification and purifying selection**. The latter is stronger when selfing is present because selfing leads to greater differentiation between adults:

$$\mathbb{E}[\Delta\bar{z}] = \frac{1}{4}\mu(\bar{z}+1) - 2s_{\mathrm{a}}\sigma_{\mathrm{B}}^2 + \mathcal{O}(\delta^2)$$
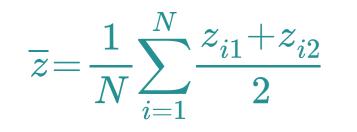
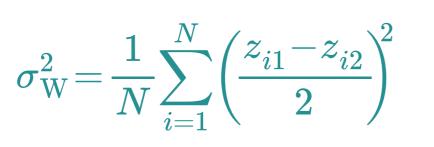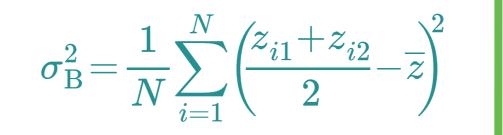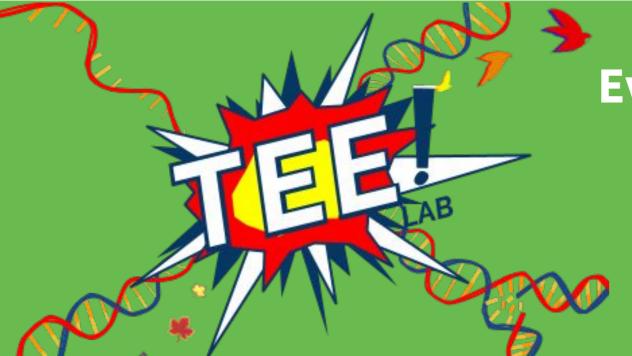> Selfing leads to **shorter** and **less diverse** TR sequences

Together, our findings suggest that **selfing and inbreeding are important factors in shaping TR sequences** and, more broadly, genetic structural variation.

Theoretical Evolutionary Ecology lab at

TEE! LAB

UNIL | Université de Lausanne

← Read it later at
sudbrack.me/TR

Swiss National Science Foundation