



HAL
open science

Mutation rate estimates for 110 Y-chromosome STRs combining population and father-son pair data

Miguel Navascués, Concetta Burgarella

► **To cite this version:**

Miguel Navascués, Concetta Burgarella. Mutation rate estimates for 110 Y-chromosome STRs combining population and father-son pair data. Annual Meeting of the SMBE 2010, 2010, Lyon, France. 2010. hal-02751488

HAL Id: hal-02751488

<https://hal.inrae.fr/hal-02751488>

Submitted on 26 Feb 2021

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.

Mutation rate estimates for 110 Y-STRs combining population and father-son pair data

Concetta Burgarella^{1,2} & Miguel Navascués^{1,3}

1 CNRS UMR 7625 Écologie et Évolution, Paris (France)

2 CIFOR-INIA, Madrid (Spain)

3 INRA, UMR CBGP (INRA/IRD/Cirad/Montpellier SupAgro) Montpellier (France) navascues@supagro.inra.fr



INTRO

Y-chromosome microsatellites (Y-STRs) are typically used for kinship analysis and forensic identification as well as for inferences on population history and evolution. All applications would greatly benefit from reliable locus-specific mutation rates, to improve forensic probability calculations and interpretations of diversity data. However, estimates of mutation rate from father-son transmissions are available for few loci and have large confidence intervals, due to the small number of meioses usually observed. By contrast, population genetic diversity data exist for many more Y-STRs, holding unused information about their mutation rates.

METHODS

'Homozygosity' ($H = \sum p_i^2$) can be used to estimate the parameter $\theta = 2N\mu = ((1/H) - 1)/2$ for each locus. The ratio between estimates of θ from H for two loci (R_H) provides an estimate of the relative mutation rates between loci.

A logistic regression was performed between the proportion of observed mutations in father-son transmissions (i.e. meiosis) and the predictive variables R_H (for a reference locus), size of repeat motif (tri-, tetra-, penta- or hexanucleotide) and structure of repeat motif (simple or complex). Data were obtained from several publications. This regression was used to predict mutation rates on 110 Y-STRs loci. The error of this method was evaluated by simulations.

RESULTS

- Simple Y-STRs mutate faster than those with a complex structure (fig. 1)
- Repeat size influences mutation rate: $\mu_{tetra} > \mu_{hexa} > \mu_{penta} \approx \mu_{tri}$ (fig. 1)
- Locus specific mutation rate estimates include estimates for 30 loci lacking meiosis observations and for 41 with a previous estimate of zero (not shown, see table for a summary)
- Regression based estimates have a lower error if the number of observed meioses is low. However, with enough data, direct estimation of mutation rates from meiosis is preferable (fig. 2)

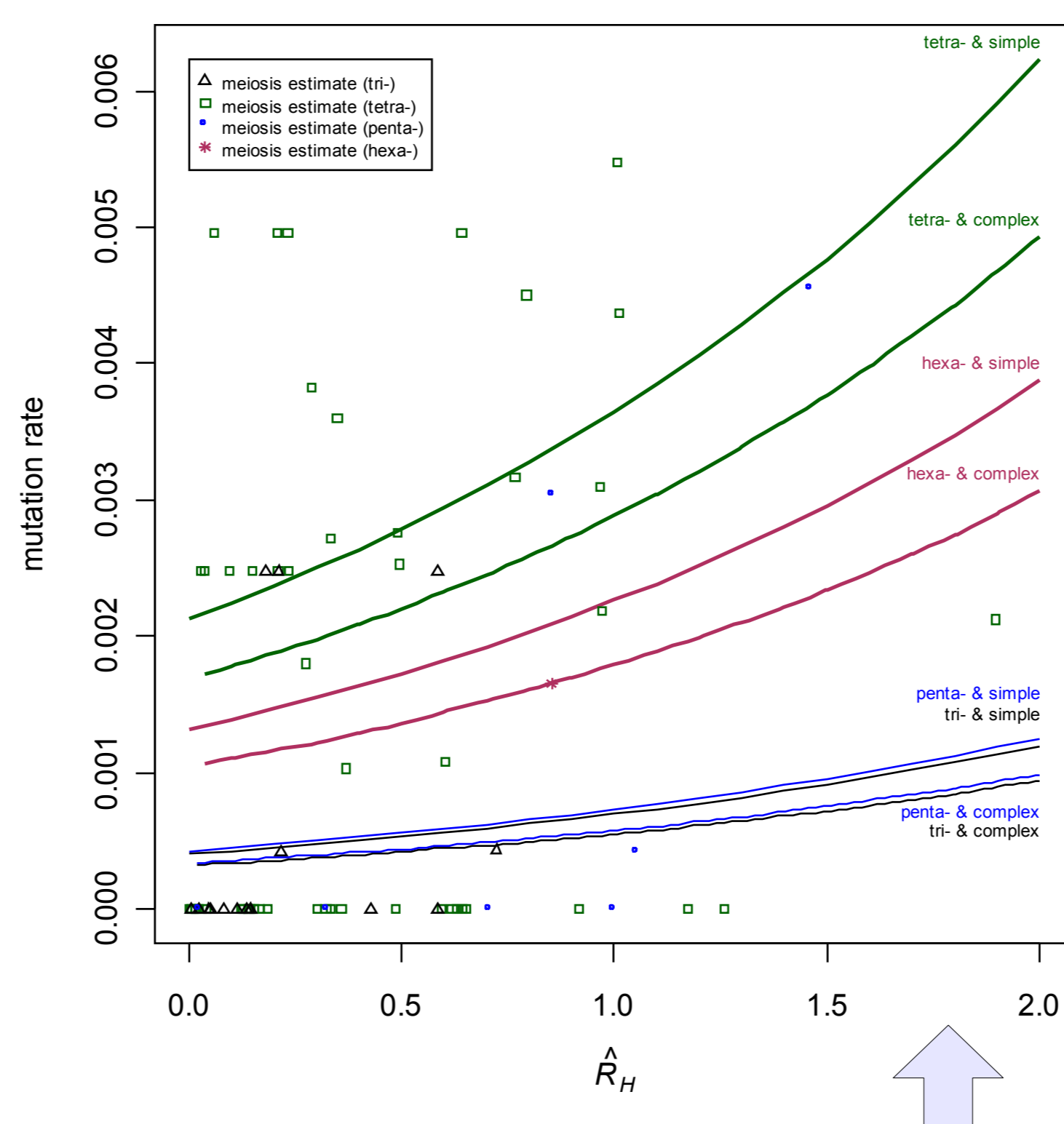


Figure 1. Logistic regression curves in function of relative genetic diversity to a reference locus (R_H) and size and structure of repeat motif.

Figure 2. Relative mean squared error (ReIMSE) for mutation rate estimates in simulated datasets of meiosis and population data.

Table 1. Average mutation rate estimates for each category of Y-STRs. Locus specific estimates for 110 loci will be presented elsewhere (article submitted to *European Journal of Human Genetics*)

repeat type	# of loci (with meiosis)	mean μ (meiosis)	mean μ (regression)
tri+simple	18 (16)	9.57×10^{-4}	8.15×10^{-4}
tri+complex	1 (1)	4.30×10^{-4}	4.75×10^{-4}
tetra+simple	44 (37)	2.30×10^{-3}	2.64×10^{-3}
tetra+complex	29 (17)	2.69×10^{-3}	3.11×10^{-3}
penta+simple	8 (6)	5.78×10^{-4}	5.95×10^{-4}
penta+complex	8 (1)	4.56×10^{-3}	5.11×10^{-4}
hexa+complex	2 (1)	1.65×10^{-3}	1.56×10^{-3}

