

## Intraclade Age Uncertainty Analysis

---

Consider a table of  $N$  haplotypes each of which has  $Z$  makers. Let  $V_j$  denote the variance of column  $j$  in that table. The intraclade or coalescence age of that set of haplotypes is given by Nordtvedt (2012) as:

$$A = \frac{1}{\mu_s} \sum_1^Z V_j \quad (1)$$

The term  $\mu_s$  is the sum of the mutation rates, i.e.

$$\mu_s = \sum_1^Z \mu_j \quad (2)$$

The uncertainty in the age  $\Delta A$  is assumed to be given by the square root of the sum of the square on the individual uncertainties in accordance with equation (3):

$$(\Delta A)^2 = \left(\Delta\mu_s \frac{\delta A}{\partial\mu_s}\right)^2 + \sum_1^Z \left(\Delta V_j \frac{\delta A}{\delta V_j}\right)^2 \quad (3)$$

There are two derivatives to be evaluated:

$$\frac{\delta A}{\partial\mu_s} = \frac{-1}{\mu_s^2} \sum_1^Z V_j \quad (4)$$

$$\frac{\partial A}{\partial V_j} = \frac{V_j}{\mu_s} \quad (5)$$

Introduce the error of the average mutation rate as:

$$\varepsilon = \frac{\Delta\mu_s}{\mu_s} \quad (6)$$

$$\left(\Delta\mu_s \frac{\delta A}{\partial\mu_s}\right)^2 = \varepsilon^2 \left(\sum_1^Z V_j\right)^2 \quad (7)$$

Equation (7) is the first term in equation (3) and represents the uncertainty in the computed age due to uncertainty in the sum of the mutation rates. McDonald (2017)

## Intracade Age Uncertainty Analysis

---

surveyed the available data on STR mutation rates. In the data of Heinila (2012), Burgarella et al. (2011) and Willems et al. (2016) there are 54 markers in common. That data is shown in table (2) and summarized in table (1). From that data, it is concluded that as an order of magnitude approximation:

$$\varepsilon = 0.10 \tag{8}$$

Evaluation of the second term in Equation (3) involves computing  $\Delta V_j$  for each of the Z STR markers in the computation. According to Wonnapijit, Chinnery & Samuels (2010) these are:

$$\Delta V_j = \sqrt{\left(\frac{2}{N-1}\right)} V_j \tag{9}$$

Hence equation (3) becomes:

$$\Delta A = \left(\frac{1}{\mu_s}\right) \sqrt{\left(\varepsilon^2 \left(\sum_1^Z V_j\right)^2 + \left(\frac{2}{N-1}\right) \sum_1^Z V_j^2\right)} \tag{10}$$

Table 1

Number of Markers	54
Sum Heinila	0.104
Sum Burgarella	0.117
Sum Willems	0.103
Mean Sum	0.108
Std Dev	0.008
Std Dev/Sqrt(2)	0.011
Std Error*1.96	0.099

## Intracade Age Uncertainty Analysis

---

Table 2 – A set of 54 STR mutation rates as reported by three different authors  
[from summary by McDonald (2017)]

Marker	Heinila	Burgarella	Willems
DYF406S1	0.00161	0.00473	0.00214
DYS19	0.00168	0.00284	0.00228
DYS388	0.00058	0.00046	0.00058
DYS389i	0.00234	0.00220	0.00255
DYS391	0.00276	0.00202	0.00165
DYS392	0.00060	0.00048	0.00047
DYS426	0.00011	0.00046	0.00008
DYS434	0.00028	0.00258	0.00031
DYS435	0.00023	0.00228	0.00022
DYS436	0.00007	0.00044	0.00010
DYS437	0.00083	0.00233	0.00078
DYS438	0.00049	0.00075	0.00059
DYS439	0.00471	0.00101	0.00508
DYS441	0.00167	0.00371	0.00189
DYS442	0.00329	0.00193	0.00265
DYS445	0.00092	0.00247	0.00072
DYS450	0.00011	0.00047	0.00022
DYS454	0.00020	0.00218	0.00037
DYS455	0.00027	0.00214	0.00019
DYS456	0.00539	0.00327	0.00377
DYS458	0.00717	0.00478	0.00920
DYS460	0.00331	0.00249	0.00208
DYS461	0.00203	0.00297	0.00260
DYS462	0.00056	0.00277	0.00069
DYS481	0.00438	0.00694	0.00467
DYS485	0.00158	0.00056	0.00105
DYS487	0.00079	0.00046	0.00121

Marker	Heinila	Burgarella	Willems
DYS492	0.00023	0.00044	0.00023
DYS494	0.00022	0.00042	0.00014
DYS495	0.00115	0.00056	0.00077
DYS505	0.00166	0.00299	0.00190
DYS510	0.00317	0.00241	0.00247
DYS511	0.00129	0.00239	0.00162
DYS522	0.00199	0.00277	0.00209
DYS525	0.00154	0.00236	0.00145
DYS533	0.00371	0.00257	0.00191
DYS537	0.00131	0.00228	0.00141
DYS540	0.00131	0.00231	0.00117
DYS549	0.00499	0.00247	0.00454
DYS556	0.00120	0.00251	0.00120
DYS561	0.00165	0.00183	0.00181
DYS565	0.00072	0.00242	0.00058
DYS568	0.00047	0.00230	0.00066
DYS570	0.00893	0.00420	0.00780
DYS575	0.00018	0.00216	0.00009
DYS576	0.01109	0.00418	0.01373
DYS578	0.00023	0.00255	0.00020
DYS590	0.00019	0.00043	0.00014
DYS593	0.00023	0.00044	0.00016
DYS594	0.00043	0.00051	0.00047
DYS607	0.00248	0.00373	0.00177
DYS641	0.00037	0.00218	0.00020
DYS643	0.00135	0.00073	0.00192
Y-GATA-A10	0.00410	0.00290	0.00420

## Intraclade Age Uncertainty Analysis

---

### References

Nordtvedt (2014) Generations111T.xlsx, a spreadsheet that is no longer online, but as of 28 Jan 2018, it can be found in the Internet Archive:

<https://web.archive.org/web/20120616154221/http://knordtvedt.home.bresnan.net/>).

McDonald (2017) variance\_calculator\_3b.ods, a spreadsheet, The University of Manchester

Wonnapijit, Chinnery & Samuels ( 2010), The American Journal of Human Genetics, Volume 86, Issue 4, p540–550, Equation 6.