

DNA Testing for Surname Genealogical Purposes: ySTR Marker Panels Compared 1

The purpose of this paper is to provide some guidance as to the relative merit of the most popular ySTR DNA testing panels for genetic genealogical purposes, and in particular the relative merits of the offerings of Family Tree DNA, and Ancestry.com [Ancestry pulled the plug on these tests many years ago, now. JBR]

In the absence of adequate information and guidance from FTDNA, the premier testing company for genetic genealogy, I here use the “Chandler rates”, supplemented by a set of anonymously published rates, said to be empirically derived^[1] to calculate the average mutation rate for all the marker panels considered in this paper: FTDNA 12, 25, 37, 67, and 111, and Ancestry 30, and 43.^[2]

The basis for the individual marker mutation rate estimates used in this paper is most unsatisfactory, and I continue to actively monitor the net for better sources and better data, but since as genetic genealogists we have in the meantime problems to solve, some estimates, however rough, are better than none.

¹ See John F. Chandler, in “[Estimating Per-Locus Mutation Rates](#)” *J Genet Geneal*, 2:27-33, but this covers only his rather well-founded estimates for markers 1-37, laying out the methodology and providing some general idea of the database (8,430 37-marker haplotypes downloaded from the YSearch database, and calibrated on the 30,000-odd cumulative father-son paternity test haplotypes tabulated in L. Gusmao, *et al.*, “[Mutation Rates at Y Chromosome Specific Microsatellites](#)” *Human Mutation* 26(2005):520-528. Chandler himself points out in his paper, that the calculated error rates for these estimates are rather large, amounting to roughly 15%. The remaining estimates all come from [an anonymously provided, but widely used spreadsheet](#) referenced in [this archived Rootsweb posting](#).

The estimates for the remaining markers, from the FTDNA 38-67, and 68-111 extended marker panels (which collectively cover all the markers tested by Ancestry) are considerably more dubious because no information is provided regarding their derivation. The spreadsheet claims that the estimated mutation rates for markers 38-67 are actually Chandler’s, and as I understand from other scattered references, these may have been derived from observed marker value variances, which may come from the [SMGF Y-Chromosome database](#), or they may too be based on [FTDNA’s YSearch database](#). The introductory note to this spreadsheet suggests that the remaining rates, for markers 68-111 are based on a set of 3565 haplotype, presumably taken from the YSearch database.

An alternative set of estimates for whole marker panels is to be found in Igor L. Rozhanskii, Anatole A. Klyosov, “[Mutation Rate Constants in DNA Genealogy \(Y Chromosome\)](#)”, 1(2011):26-34, and co-author Klyosov has been a tireless advocate of his estimated mutation rates, and his methodology on the Rootsweb DNA lists. This paper is focused on computing TMRCA estimates, but I recommend it for the thorough coverage of the issues involved in analyzing sets of haplotypes for genetic genealogical purposes. More to the purpose, the paper includes a set of “mutation rate constants” for the FTDNA yDNA marker panels, whose actual derivation is needlessly complicated by an attempt to incorporate (for TMRCA calculation convenience) a nominal years/generation number of 25, but these constants for the various panels do readily yield estimates of their relative mutational sensitivity. In particular, I have calculated that the 67-marker panel adds 33% more mutational sensitivity to the 37-marker panel, and the 111-marker panel adds 253% to the 37- and 65% to the 67-marker panel. A potential strength of these co-authors’ estimates is that they are said to have been derived empirically from a great deal of practical analytical work with actual haplotype data and corresponding genealogies, but in the complete absence of any accounting for this data, except for the presentation of examples, leaves me reluctant to rely on it, when there is at least a specific accounting for Chandler’s methodology and modeling for markers 1-37. And in order to make full and consistent use of this data to compare FTDNA’s panels with Ancestry’s or others, I need specific marker mutation values and these are not provided in the Rozhanskii & Klyosov paper, since their mutation rate constants are based on whole panel analyses.

There are several other published papers that present methodologies and estimates for these marker mutation rates, and these use modeling approaches similar to Chandler’s, though like the Rozhanskii & Klyosov paper, they fail to provide sufficient details, or convincing rationales for their calibration against existing data. Anatole Klyosov has been a contemptuous critic of a couple of these papers, for reasons that seem cogent to me, and I see no value even in citing them.

² Ancestry claims more markers for these panels, e.g. for its extended panel, 46, but this is to count rare additional values beyond DYS464a-d; since FTDNA doesn’t count these, I’ve subtracted them from the Ancestry panel to level the playing field.

Average Mutations Rates for the Major Panels

First, here are the average estimated mutations rates per marker per generation (transmission event from father to son) for each of the marker panels in question:

	Avg Mutation Rate	Error Range
FTDNA 12-marker panel:	.00187	±.00028
" 25- " " :	.00278	±.00042
" 37- " " :	.00492	±.00074
" 67- " " :	.00335	
" 111- " " :	.00290	
 Ancestry 30-marker panel:	 .00266	
" 43-marker panel:	.00253	

Comparing Panels

The value of yDNA testing for genetic genealogical purposes depends entirely on, and is proportional to, the number of mutations one can expect. Thus, we might compare test panels on their respective chances of producing one or more mutations in any given generation, applying the formula

$$1 - (1 - \langle \text{avg mutation rate per marker} \rangle)^{\langle \# \text{markers in panel} - 1 \rangle}$$

thus

$$\begin{aligned} & 1 - (1 - .00492)^{36} = .163 \quad \text{for FTDNA-37} \\ \& \quad 1 - (1 - .00253)^{42} = .101 \quad \text{for Ancestry-43} \end{aligned}$$

Thus, the FTDNA 37+ panel is $(.163 - .101) / .101 = 61.4\%$ more likely to produce one or more mutations per generation than the Ancestry-43+ panel.

A Mutational Sensitivity Index for Comparing Panels

An even simpler calculation yields what might be called a “mutational sensitivity index”,

$$\begin{aligned} .00492 \times 37 &= .1820 \\ .00253 \times 43 &= .1088 \end{aligned}$$

and in this comparison, the FTDNA 37-marker test has 67.4% more mutational sensitivity than the best Ancestry test. Here are some mutational sensitivity (and price) comparisons between various currently available tests:

	Average Mutation Rate	\$Price as of Sep2014	Mutational Sensitivity Index	Value/ Price Index
FTDNA-25	.00278	124	.0695	5.6
FTDNA-37	.00492	149	.1820	12.2
FTDNA-67	.00335	238	.2245	9.4
FTDNA-111	.00290	339	.3219	9.5
 Ancestry-30+	 .00266	 149	 .0798	 5.4
Ancestry-43+	.00253	179	.1088	6.1

Clearly, the FTDNA 37-marker panel produces the best “bang for the buck”, and it happens also to be the lowest resolution test capable of sorting most testees definitively into their particular [genealogical patrilineage](#). When considering the price of these various tests, one should be aware also that FTDNA has had a Holiday sale for the month of December in each of the last four years, in which the price of many tests and upgrades are dropped. Typically, the \$149 price of the 37-marker test (assuming that it’s ordered through one of the surname projects—it cost \$169 if you order it directly from the company) is dropped to \$119.

Should I Upgrade from the Ancestry tests to FTDNA-37?

Yes, especially if you’ve only done the Ancestry-30+ test. If you can prove that you’ve tested at Ancestry, you can upgrade to FTDNA-37 for \$119 by printing out and sending in [this form](#) along with the proof. Or, if you have commissioned the better Ancestry-43+ test, you can get the upgrade for just \$58, by going to [the FTDNA order page](#) and scrolling down to and clicking the link “Transfer Y-DNA46 + Y-DNA37”; you will be asked to submit the results of your Ancestry testing with your order. This is well worth doing, both because the FTDNA-37 test has 40% more mutational sensitivity than even the best Ancestry test, and because you will then become an FTDNA customer, and be exposed to a much larger data base of possible matches.^[3]

Should I Extend 37-Markers to 67, or to 111?

FTDNA also offers upgrades to its higher resolution tests for not much more than the differential cost of ordering these tests in the first place. The current prices for these upgrades (and sale prices I’ve seen) are:

		\$Price	(Sale)
FTDNA-12	-> 37	99	69
FTDNA 12	-> 67	189	148
FTDNA 12	-> 111	---	---
FTDNA 37	-> 67	99	79
FTDNA 37	-> 111	220	188
FTDNA 67	-> 111	129	109

By my calculations, the extension of FTDNA-37 to 67 markers yields only about a 23% boost in mutational sensitivity, at a cost 60% higher, and yields only about a 52% chance of producing one or more additional mutations over say 7 generations, and these estimates are consistent with the data I’ve seen. The markers in the 38-67 panel are relative duds and rarely mutate. The markers in the 68-111 panel are somewhat better, but the catch is that to test these markers you first have to test the 38-67 markers.

In general, I think it’s better to spend the money testing additional people at 37-markers than going for 67, at least initially; a 37-marker test can always be extended later to 67, for \$99, versus \$89 if you order the 67-marker test in the first place.

However, the value of all testing is directly proportional to the number of people you match to, who belong to your DNA-confirmed patrilineage, and where your patrilineage is large and others have already tested to 67 or 111 markers, the calculus can work out differently. By the same token, where the patrilineage is small, testing has little or no value. Thus, if you turn out initially to be a patrilineage of one (because you have no matches), the value of your testing is nil—except that, of course, you will have got the ball rolling for your patrilineage and perhaps encouraged others to test

³ There are many other advantages to being an FTDNA customer, which you can read about in [this comparison chart](#).

who were sitting on the fence. If you happen to have a rare surname that hasn't appeared yet in the YSearch database or elsewhere, you should definitely test, just to get your name out there.

In considering whether to extend your test, the same factors apply. The goal, always, is to try to find matching mutations with people who have already tested, and the chances of this depend both on the mutational sensitivity of the test or extension you are contemplating, and on the number of people who have already tested to that level. There are times when extending is very desirable indeed.

Extending Your Haplotype by Testing Individual ySTR Markers

Alternatively, where enough other people have already tested in your patrilineage, that it has become possible to identify certain shared mutations that mark particular family sub-branches, you can determine whether you belong to those sub-branches by testing just those additional markers. The cost for testing individual markers is just \$20 each, and it is rare that you would ever want to test more than three or four of these. To my knowledge, only Family Tree DNA offers the ability to test individual STR markers, though they don't as yet have full coverage for their extended 67- and 111-marker panels.