Gene Conversions in Palindromic Regions of the Y chromosome

A. Turner[1], N.M. Myres[2], J.E. Ekins[2], N. Angerhofer[2], S.R. Woodward[2]
1) Menlo Park, CA, USA, DNACousins@aol.com; 2) Sorenson Molecular Genealogy Foundation (SMGF), Salt Lake City, UT, USA

Abstract

Several commonly tested microsatellite markers on the Y chromosome are located in palindromic regions. The Sorenson Molecular Genealogy Foundation (SMGF) database includes results for the multi-copy markers DYS385 a/b, DYS459 a/b, DYS464 a/b/c/d and YCAII a/b. These markers typically have multiple alleles within an individual. However, gene conversion may result in a single allele for a marker ("homozygosity") and violate the single-step mutation model for microsatellites. The YCAII marker in particular demonstrates this phenomenon with an excess of homozygosity, compared to expectations based on the mutation rate for microsatellites.

Introduction

Complete mapping of the Y chromosome was complicated by long regions of duplicated segments that differed only slightly. Publication of the complete sequence in 2003[1] illustrated the palindromic nature of these duplications, with homologous segments arranged in inverted order. The palindromes can be visualized as arms on a hairpin turn formed by Watson-Crick base pairing.[2] The symmetry of the arms exceeds 99.9%, more than typical for segmental duplications on the autosomal chromosomes. This symmetry may be driven by gene conversion, where sequences on one arm overwrite the homologous sequence on the other arm and restore uniformity to the two arms.[3]

Two Forms of Recombination

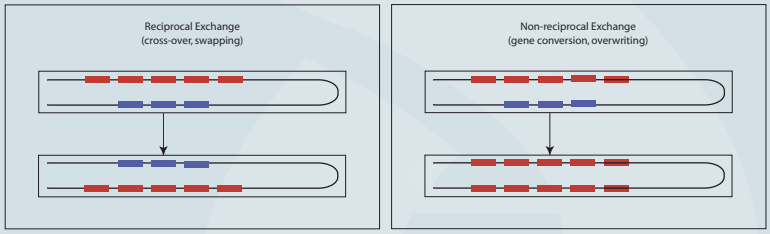

Figure 1. Two forms of recombination

The 99.9% symmetry figure is based on single nucleotide polymorphisms with a low mutation rate, on the order of 2.0×10^{-8} per base per generation. Microsatellite markers (Short Tandem Repeats), with a higher mutation rate, on the order of 2.0×10^{-3} per marker per generation, may provide a more dynamic window into gene conversion events. Microsatellite markers on the Y chromosome are used extensively for genealogical purposes, and large databases are available for data mining.

The model for microsatellite mutations involves stepwise changes in the number of short tandem repeats. For instance, the dinucleotide repeat pattern CA can increase or decrease by one repeat by a process of replication slippage, as shown in Figure 2. Father/son studies of the microsatellite mutation rate show that 96% of changes involve a single step, with most of the remaining 4% being two-step mutations.[4]

Methods

The SMGF database contains over 15,000 records with Y chromosome data for 37 microsatellite markers, including 4 multi-copy markers (DYS385 a/b, DYS459 a/b, DYS464 a/b/c/d, and YCAII a/b). DNA samples are collected from volunteers, who also contribute ancestral pedigree data for at least four generations. Microsatellite markers are analyzed with conventional techniques.

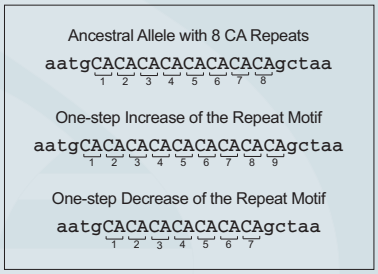

Figure 2. Stepwise microsatellite mutational model

Records were filtered from the database according to the following criteria:

a) Individuals who were estimated to belong to the most common European haplogroup, R1b, were included. The basis for selection was closeness of fit to the Atlantic Modal Haplotype. As originally described by Wilson,[5] this haplotype of six single-copy markers plus its one-step neighbors is uncommon in other haplogroups. This dataset is thus most likely derived from a single distant common ancestor who lived thousands of years ago.

DYS388	DYS390	DYS391	DYS392	DYS393	DYS394
12	24	11	13	13	14

Table 1. Atlantic Modal Haplotype

b) Individuals who were estimated to be closely related on the basis of their full 37-marker haplotype were then removed, in order to avoid artificial inflation of subgroups by large familial clusters in the database. Matches on 34 or more of the 37 markers imply a Most Recent Common Ancestor within a timeframe of a few hundred years.

The final dataset consists of 3326 records, representing many independent lines of descent from the founder of haplogroup R1b. The multi-copy marker with the strongest modal value was YCAII, with 2538 (76.3%) records showing a haplotype of 19 and 23 repeats for YCAIIa and YCAIIb. (It is not possible to tell which physical location on the Y chromosome has which value; repeats are always listed in ascending order. See Figure 3 for the location of the YCAII markers.)

This strong mode may be a consequence of a mutation rate that is somewhat lower than the average of other microsatellite markers.[6] YCAII also has the largest spread of modal values, with a difference of four repeats between 19 and 23. This makes the distinction between several one-step mutations and one gene conversion event more striking. By comparison, the modal value for DYS459 a/b is 9-10.

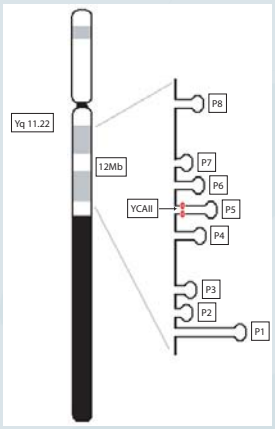

Figure 3. Panlindromic region of the Y chromosome

Results

Thirty different haplotypes (combinations of alleles for YCAII a/b) were observed. The frequency of one-step mutations from the 19-23 modal haplotype is markedly decreased, ranging from 44 to 284 samples, as shown in Table 2 and Figure 4. There is a tendency for microsatellite markers with longer repeat stretches to be more mutable, and this is observed for YCAII, with the two larger clusters being derived from the 23 allele. There is also a tendency for long repeats to become shorter,[7] and the largest cluster 19-22 is in accord with this trend.

Two-step mutations from the modal show another sharp decrease, and two theoretically possible haplotypes (18-24 and 20-24) have no exemplars at all in this dataset.

Three-step haplotypes continue to show decreasing numbers, as do most four-step haplotypes (four with one examplar and ten with no exemplars). However, two of the sixteen possible four-step haplotypes sharply counter the trend for decreasing frequency, with 80 samples homozygous for 19-19 and 43 samples homozygous for 23-23. This suggests that another mechanism for generating haplotypes is in play. Gene conversion, where the palindromic arm with 19 repeats overwrites the arm with 23 repeats (or vice versa), is a plausible mechanism for this trend reversal. In fact, the few examples of five and six-step mutations could be more easily derived as one or two step mutations from the homozygous states of 19-19 and 23-23.

Haplotype	N	Steps
19-23	2538	0
19-22	284	1
19-24	130	1
20-23	50	1
18-23	44	1
19-21	79	2
21-23	21	2
17-23	11	2
19-25	7	2
20-22	3	2
18-22	2	2
19-20	6	3
20-21	4	3
22-23	4	3
16-23	3	3
21-22	2	3
17-22	1	3
18-21	1	3
18-25	1	3
19-26	1	3
19-19	80	4
23-23	43	4
17-21	1	4
20-20	1	4
21-21	1	4
22-22	1	4
23-24	4	5
18-19	1	5
17-19	1	6
23-25	1	6

Table 2. Frequency of YCAII haplotypes, with number steps from modal value 19-23

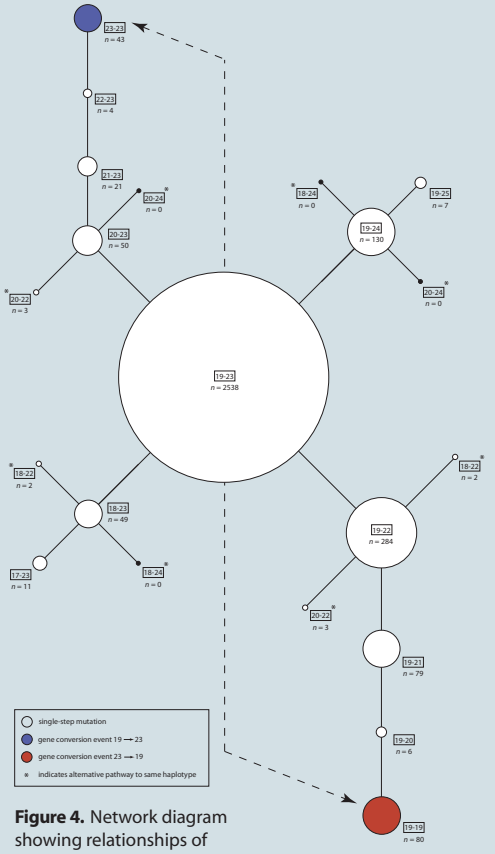

Figure 4. Network diagram showing relationships of haplotypes according to the single-step model

Conclusion

The process of gene conversion is of more than theoretical interest. Many of the multi-copy genes in the palindromes are involved in spermatogenesis. Mutations in one copy may be repaired by replacing the segment with an intact version of another copy, or conversely, mutations may be propagated to all copies.[8]

Microsatellite markers, with their relatively high mutation rate, may provide a way of calibrating the frequency of gene conversion. The SMGF database is large enough to include representative samples of many of the theoretically possible YCAII haplotypes. For the shorter allele of YCAII, the rate of gene conversion from 19-23 to 19-19 appears to be similar to the rate of a one-step mutation from 19-23 to 20-23. For the longer allele of YCAII, the rate of gene conversion from 19-23 to 23-23 appears to be about one-third the rate of a one-step mutation from 19-23 to 19-22.

References

[1] Skaletsky H et al "The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes" (2003) Nature 423(6942):825-37
[2] Graves J "The degenerate Y chromosome – can conversion save it?" (2004) Reproduction, Fertility and Development 16:527-534
[3] Hurles M et al "A singular chromosome" (2003) Nature Genetics 34(3):246-247
[4] Gusmao L et al "Mutation Rates at Y Chromosome Specific Microsatellites" (2005) Human Mutation 26(6):520-8
[5] Wilson J et al "Genetic Evidence for different male and female roles during cultural transitions in the British Isles" (2001) PNAS 98(9):5078-5083
[6] Hutchinson L et al "Direct determination of mutation characteristics of Y chromosome STR loci", ASHG Annual Meeting, Toronto, Canada 2004; http://www.smgf.org/resources/papers/ASHG2004-4.pdf
[7] Gusmao, ibid
[8] Graves, ibid

SORENSON MOLECULAR GENEALOGY FOUNDATION
www.smgf.org