

Why the Y Chromosome? – A Look at Male Lineage and Ancestry

Nancy L. Elwess*, Felecia Edwards, and Sandra M. Latourelle

Department of Biological Sciences

Plattsburgh State University, USA.

Email: nancy.elwess@plattsburgh.edu.

*The author for correspondence.

Abstract: Up until a short time ago the Y chromosome played the role of the juvenile delinquent within human chromosomes. It was considered to be rich in junk, short on genes, and rapidly degenerating. Now the Y chromosome is growing up by providing a means for investigating human migration. Through the use of genetic markers on the Y chromosomes, students from a college bioinformatics course were able to determine the migration routes over tens of thousands of years from the 100+ male DNA samples they collected.

Keywords: Y Chromosome, Genetic Markers, Lineage, Migration

Introduction

The year was 1990. It was the time when geneticists published findings on the part of the Y chromosome that conferred maleness, the SRY (Sex Reversal on Y chromosome) gene. Once again the Y chromosome is in the journals. Now, it has been elevated from genomic junkyard to evolutionary revelation [Lewis, 2003].

This tiny chromosome is quickly becoming mighty in how it reveals evidence of ancient demographics histories. In the palindromes (regions of DNA in which the sequence of nucleotides are identical with an inverted sequence in the complementary strand) found within the DNA of the Y chromosome are the clues to the past. To determine the migration route of ancient man a time line was needed; DNA provided the clock. The longer the DNA lineage has been in existence, the more mistakes/mutations the DNA sequence is likely to contain. These mutations on the Y chromosome provide genetic markers (Table 1). By tracking these genetic markers a timeline for man's migration routes can be determined (Figure 1). This approach provides a unique *double helix*: the combining of history and genetics.

Initially, the focus in determining time lines was on mitochondrial DNA that is inherited maternally. It descends from generation to generation from mother to daughter. Mitochondrial DNA evidence shows that females were more mobile than males [Seilstad et al., 1994; Underhill et al., 1996; Ruis-Linares et al., 1996]. This suggested that in most African tribes it was the women who did the traveling probably to find mates [Shnayerson,2005; Cavalli-Sforza, 1997]. The men on the other hand stayed put, so the story that the mitochondrial DNA told

only provided half of the answer. The Y chromosome was needed to provide further insight into the "Out of Africa" theory and man's migration. Recently, genetic markers on the non-recombinant region of the Y chromosome have been used as the male complement to mitochondrial DNA. These markers allow for the reconstruction and tracing of ancient human migration routes (Figure 1).

The Y chromosome has become such a research interest that the National Geographic Society is undertaking its most ambitious project, the Genographic project. Under the direction of population geneticist, Dr. Spencer Wells, and at a cost of 40 million dollars over five years, the Genographic project is establishing eleven DNA-sampling centers with the goal of collecting over 100,000 DNA samples worldwide [Shnayerson,2005].

Marker	Years ago	Base pair (bp) size
M168	50,000 years	473 bp
M130	50,000 years	205 bp
M96	40,000 years	440bp
M9	40,000 years	340 bp
M45	35,000 years	352 bp
M173	30,000 years	220 bp
M242	20,000 years	365 bp
M3	10,000 years	241 bp
M2	10,000 years	209 bp
M122	10,000 years	393 bp
M19	<10,000 years	333 bp

Table 1. Genetic Markers for specific time periods and their expected PCR product size

RETRACING THE MOVEMENT OF GENETIC LINEAGES AROUND THE WORLD

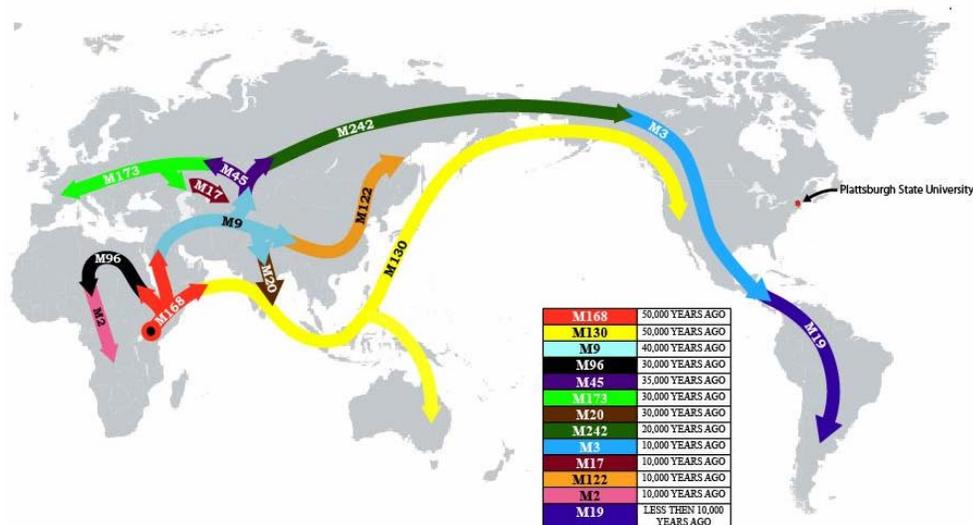


Figure 1. Proposed migration routes of man using genetic markers as suggested by Kunzig (2004). Map created by Kyle Martin, a student in the bioinformatics course.

The main objective for our project was to provide students an opportunity to take part in a complete research experience. By complete, we mean experimental design, laboratory techniques, data collection, and data analysis. The data published here is the result of a research project undertaken by students as part of their junior/senior level college bioinformatics course. The six students working on this project collected DNA samples from over 100 male college students. Through the use of genetic markers, which can date the Y chromosome back 10,000 years, 20,000 years, 30,000 years, 40,000 years, and 50,000 years, the students were able to determine the male lineages, migration routes and the ancestry of the Y chromosome from the tested DNA samples. Their results along with the protocols used are presented in this paper.

Methods

I. DNA isolation by Saline Mouthwash

Prepare a 0.9% saline solution (0.9 g NaCl per 100 mL distilled water); aliquot 10 mL 0.9% saline solution into 50 mL polypropylene tubes. One tube will be needed for each sample collected. Make sure to include at least one female sample to provide a negative control. Also make a 10% Chelex® solution (Chelix 100 Resin from Bio-Rad), 100 µl of 10% chelex will be needed per sample; 10 mL is enough for 100 samples. Chelex does not dissolve in solution, so it will need to be shaken prior to each use. Also it is suggested that a truncated tip be used when pipetting the Chelex solution to freely allow the uptake of the beads within the solution.

1. Have each participant pour 10 mL of the saline solution (0.9% NaCl) into his or her mouth and vigorously swish for 30 seconds.
2. Expel saline solution back into the polypropylene tube. Number the tube.
3. Store samples on ice or in refrigerator until the steps below can be done.
4. Swirl to mix cells in the tube and transfer 1 mL (1000 µl) of the liquid to a 1.5 mL tube, number the tube.
5. Place the sample tubes in a balanced configuration in a microcentrifuge, and spin for 1 minute.
6. Carefully pour off supernatant into paper cup or sink. Be careful not to disturb the cell pellet at the bottom of the test tube. A small amount of saline will remain in the tube.
7. Resuspend cells in remaining saline (~30 µl) by pipetting in and out. (If needed, saline solution may be added to facilitate resuspension.)
8. Withdraw 30 µl of cell suspension, placing it into a new screw cap 1.5 mL tube (make sure to number the tubes) then add 100 µl of 10% Chelex (Shake this solution prior to use). Shake the tube well to mix.

9. Boil cell sample for 10 minutes. Use boiling water bath, heat block, or program thermal cycler for 10 minutes at 99°C. Then, cool tube briefly on ice (optional).
 10. After boiling, shake tube. Place in a balanced configuration in a microcentrifuge, and spin for 1 minute.
 11. Transfer 30 µl of supernatant (containing the DNA) to a clean 1.5 mL tube. Avoid cell debris and Chelex beads. This DNA sample will be used for setting up the testing of the DNA samples with the genetic markers. Make sure to mark the tubes with the correct sample numbers.
 12. Store your sample on ice or in the refrigerator.
6. Following amplification, load each sample into a 1% agarose gel along with a 100 bp marker. If using ethidium bromide, 10 µl of 10mg/mL ethidium bromide can be added to the agarose mixture (50mL) prior to pouring the gel.
 7. The gels can be run for 30-40 min at 120 volts.
 8. Visualize each gel using UV light (for ethidium bromide gels). If ethidium bromide was not used, stained the gel in methylene blue (or a similar DNA stain) for 30 min after that destain in water for 10 minutes then view with a white light. Record the results whether or not they revealed the presence the genetic marker of interest.

II. DNA Amplification Procedure using Polymerase Chain Reaction

1. Using a micropipet with fresh tips add 17.5 µl dH₂O, 2.5 µl of the 20 µM forward primer and 2.5 µl of the 20 µM reverse primer to a PCR tube containing a Ready-To-Go PCR Bead (Amersham). Tap tube with finger to dissolve bead. Table 2 contains all the respective primer sequences for the genetic markers.
2. Use fresh tip to add 2.5 µl of human DNA (from Part I) to reaction tube, and tap to mix. Pool reagents by pulsing in a microcentrifuge or by sharply tapping tube bottom on lab bench.
3. Label the cap of the tube with the sample number.
4. Add one drop of mineral oil on top of reactants in the PCR tube. Note: Thermal cyclers with heated lids do not require use of mineral oil.
5. Store all samples on ice until ready to amplify according to the following: Program thermal cycler for 35 cycles according to the following cycle profile. Each program may be linked to 4°C to hold samples after completion of amplification. However, amplified DNA also holds well at room temperature.
Denaturing time and temperature 30 sec - 94°C
Annealing time and temperature 30 sec - 54°C
Extending time and temperature 45 sec - 72°C

Safety Issues

As expected in any laboratory where chemical reagents are used, students were required to wear gloves, goggles, and cautioned not to wear contact lenses during the laboratory experience. All reagents were labeled with appropriate hazard warnings to include health hazard, fire hazard, reactivity and any specific hazard. Used buffer from electrophoresis events was placed in a specified waste container under the chemical hood. Gels containing ethidium bromide residues were placed in a biohazard bag, located under a chemical hood. III.

Systematic Approach for Sample Analysis

Feeling somewhat like ‘ancestry detectives’, the research team chose a plan of attack to begin answering the questions: *Where did we come from?* and *How did we get here?* The students reviewed the literature to determine how the Y chromosome can be used as an evolutionary marker and what genetic markers they should use for their project (Jobling & Tyler-Smith, 2003; Underhill et al., 1996; Underhill et al., 1997; Underhill et al., 2001; Yuehai et al., 2001; Seielstad et al., 1994; Seielstad et al., 2003; Wells et al., 2001). Once the genetic markers were selected a research plan was designed (Figure 2). Testing was initiated using the 50,000-year genetic marker, M168. Basically each sample was tested using a specific genetic marker, as long as it tested positive for that particular marker (presence of a band on a gel, Figure 3) the experimenter would move on to the next genetic marker. All the samples started with the 50,000-year genetic marker and worked their way towards the 10,000- year markers. Once a sample failed to test positive for a particular marker, it had reached its endpoint in this study. For example, if a sample tested positive for M242 (20,000 years) but not for the M3 marker (10,000

years), then its endpoint would be M242 (20,000 years). There are additional genetic markers for the Y chromosome that could have been used, but it was

decided to test the main markers that were suggested in the literature.

Genetic Marker	Primer Sequences	Reference
M168	Forward primer-AGTTTGAGGTAGAATACTGTTTGCT Reverse primer- AATCTCATAGGTCTCTGACTGTTC	Underhill et al., 2001
M130	Forward primer-TATCTCCTCTTCTATTGCAG Reverse primer- CCACAAGGGGGAAAAACAC	Yuehai et al., 2001
M96	Forward primer- GTTGCCCTCTCACAGAGCAC Reverse primer- AAGGTCACCTGGAAGGATTGC	Underhill et al., 2001
M9	Forward primer- GCAGCATATAAACTTTTCAGG Reverse primer- AAAACCTAACTTTGCTCAAGC	Underhill et al., 1997
M45	Forward primer- GCTGGCAAGACACTTCTGAG Reverse primer- AATATGTTCTGACACCTTCC	Underhill et al., 2001
M173	Forward primer- AATTC AAGGGCATTTAGAACC Reverse primer-TATCTGGCATCCGTTAGAAAAGA	Wells et al., 2001
M242	Forward primer-AACTCTTGATAAACCGTGCTG Reverse primer- TCCAATCTCAATTCATGCCTC	Seielstad et al., 2003
M3	Forward primer- TAATCAGTCTCCTCCCAGCA Reverse primer-AAAATTGTGAATCTGAAATTTAAGG	Underhill et al., 1996
M2	Forward primer- AGGCACTGGTCAGAATGAAG Reverse primer- AATGGAAAATACAGCTCCCC	Seielstad et al., 1994
M122	Forward primer-AAGCAATTGAGATACTAATTCAC Reverse primer-CAACTTCTTCCCTCAACATAG	Wells et al., 2001
M19	Forward primer- CTGGTCATAACACTGGAAATC Reverse primer- TGAACCTACAAATGTGAAACTC	Underhill et al., 1997
M17	Forward primer-CTGGTCATAACACTGGAAATC Reverse primer-TGAACCTACAAATGTGAAACT	Underhill et al., 1997

Table 2. Genetic Marker Primer Sequences for each of the Genetic Markers

Results

From the time it took to read about Spencer Wells' work in a popular, science magazine [Kunzig, 2004] prior to the start of the semester to the incorporation of it into the curriculum to finally the project conclusion, it has been six months. Results have indicated that college students are quite unaware of their genealogy. In a questionnaire designed to accompany each donated sample of DNA, the donors

had very little knowledge of their paternal background beyond their fathers' points of origins. Any future testing done by us will simply include an assigned number for each donor. But for now, at the molecular level, let's focus on the lineages of one hundred young men from around the world and attending a small university in up-state New York, USA

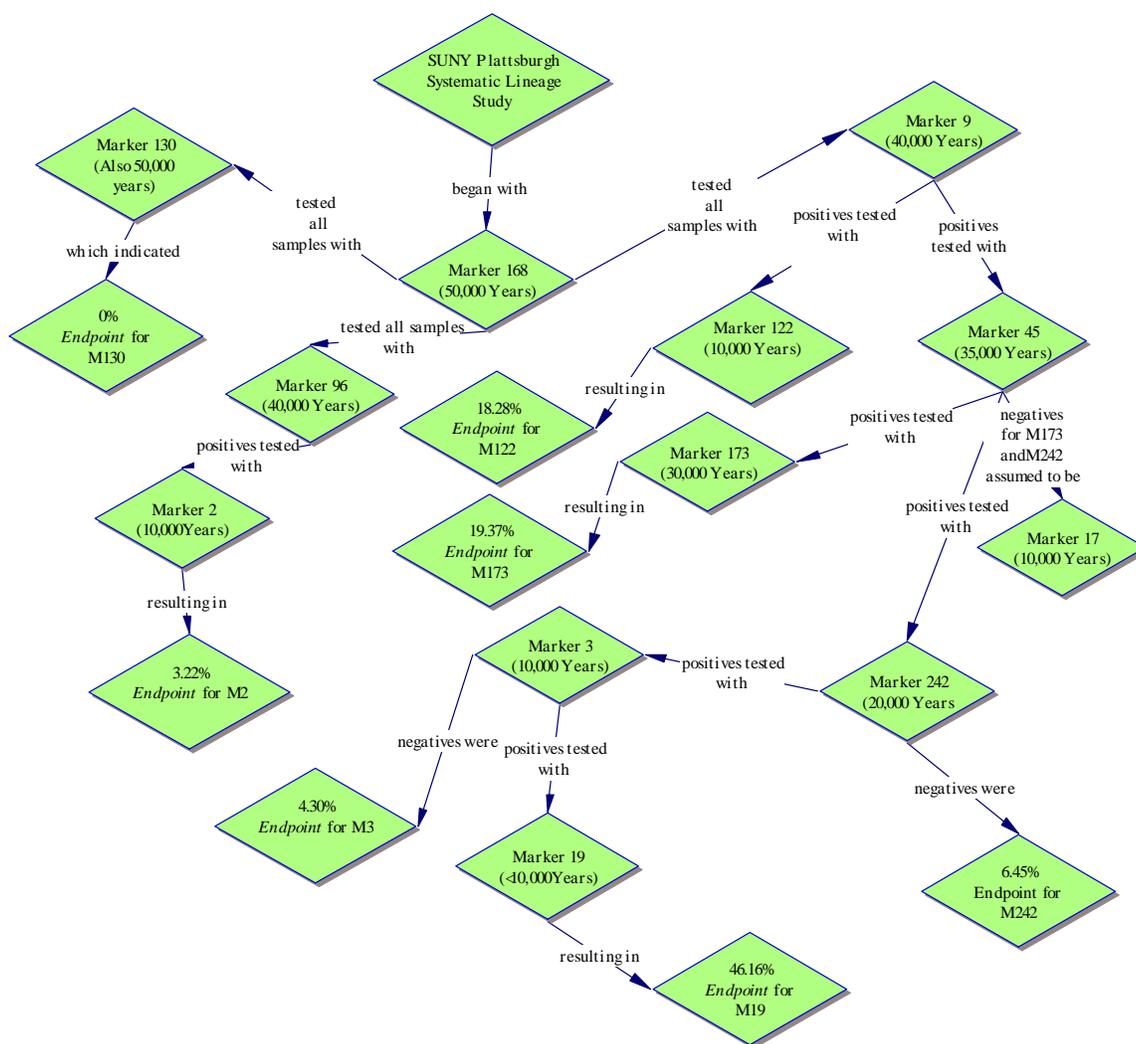


Figure 2. A Concept Map for retracing the movement of genetic lineages for 100+ male students at the State University of New York at Plattsburgh, Plattsburgh, New York. Generated by Sandra Latourelle.

As seen in Table 3 and Figure 2, the results indicated all markers tested were present with the exception of M130. When looking at endpoint genetic markers, the largest number for all samples studied was found when using M19 (Table 3, Figures 2 & 4). 29.02% of all samples tested positive for one of the 10,000-year markers: M17, M3, M2, and M122 (Figures 2, 4, 5). The largest percentage of samples, 46.16%, tested positive for M19 (Table 3, Figures 2 & 4), a marker for less than 10,000 years. It should be noted that all tests were conducted using DNA from female donors as controls.

Marker	Years ago	% in samples
M130	50,000 years	0%
M17	10,000 years	3.22%
M173	30,000 years	19.37%
M242	20,000 years	6.45%
M3	10,000 years	4.30%
M2	10,000 years	3.22%
M122	10,000 years	18.28%
M19	<10,000 years	46.16%

Table 3. Endpoint percentage results from the Genetic Markers used.

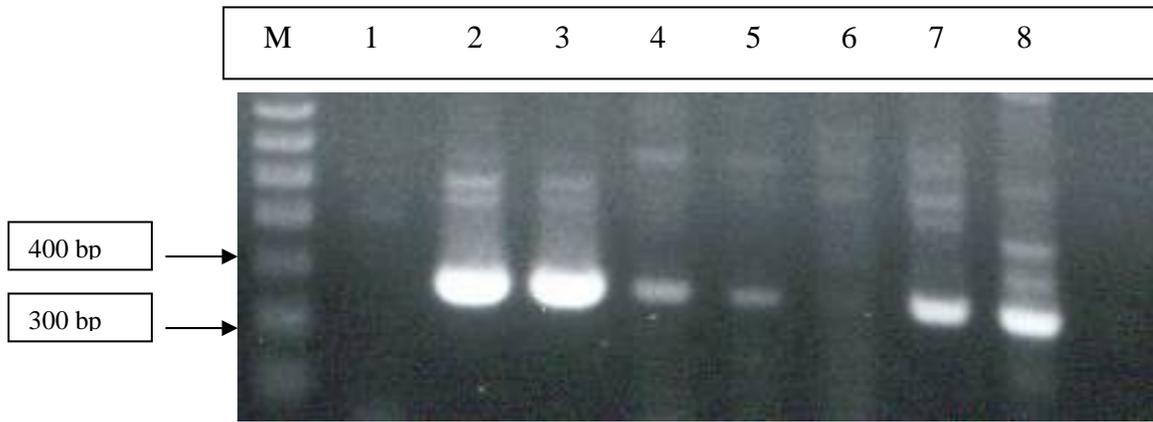


Figure 3. Results from 8 samples amplified with the M45 primers (35,000 years). The expected polymerase chain reaction product is 352 base pairs. Lane M is the 100 base pair (bp) standard. The samples in lanes 1 and 6 did not test positive for the M45 genetic marker. The sample in lane 1 was our negative control (female DNA).

Discussion

What our research students found were glimpses into the past revealed by the mirror like palindromes, markers of mystery — The DNA from our male donors had preserved their paternal legacy permitting inference of human evolution and demographic history [Underhill et al, 2000]. Our four-month study during the course actually allowed us to do a ‘long range paternity test’. It was surprising for the students at first to see that 46.16% of their samples tested positive for the M19 marker (Figure 4). This was a marker that would have made its way through Asia, across through Alaska and then into the Americas. The students were expecting a higher percentage of their results to make their migration westward through Europe. This is the beauty of science; the data is the data no matter what the expectations. Reflecting back the students realized that a large number of their sample pool were both minority and international students.

If this research project were to be expanded and not just one part of a bioinformatics course, additional Y chromosome genetic markers would be added. This would hopefully address the 19.37% of our samples that ended their migration at 30,000 years (M173). Presently, there are 100+ genetic markers for the Y chromosome that could be used. We just used the ones that we hoped would give us a basic picture of our 100 students migration.

Assessment

As mentioned this project was only part of the requirements for a junior/senior level bioinformatics course. For this part the students were assessed on their laboratory notebooks, review and presentation of the literature, laboratory techniques and poster presentation of their results at an on-campus research symposium.

Concluding Remarks

Young men willing to donate some cheek cells for a scientific endeavor provided our original samples. We have no way of knowing whom they are and if they ever heard about the French philosopher, Simone Weil. In his words, “To be rooted is perhaps the most important and least recognized need of the human soul [Smolenyak, 2004].” Perhaps some of these young men will become involved in what Megan Smolenyak refers to in her book, *Trace Your Roots with DNA*, genealogy- a marriage of genealogy and genetics [Smolenyak 2004].

Acknowledgements

We thank all of the DNA donors for making this project possible. This project was supported by a generous grant from Delta Kappa Gamma Society International, a professional honorary society of women educators

RETRACING THE MOVEMENT OF GENETIC LINEAGES AROUND THE WORLD

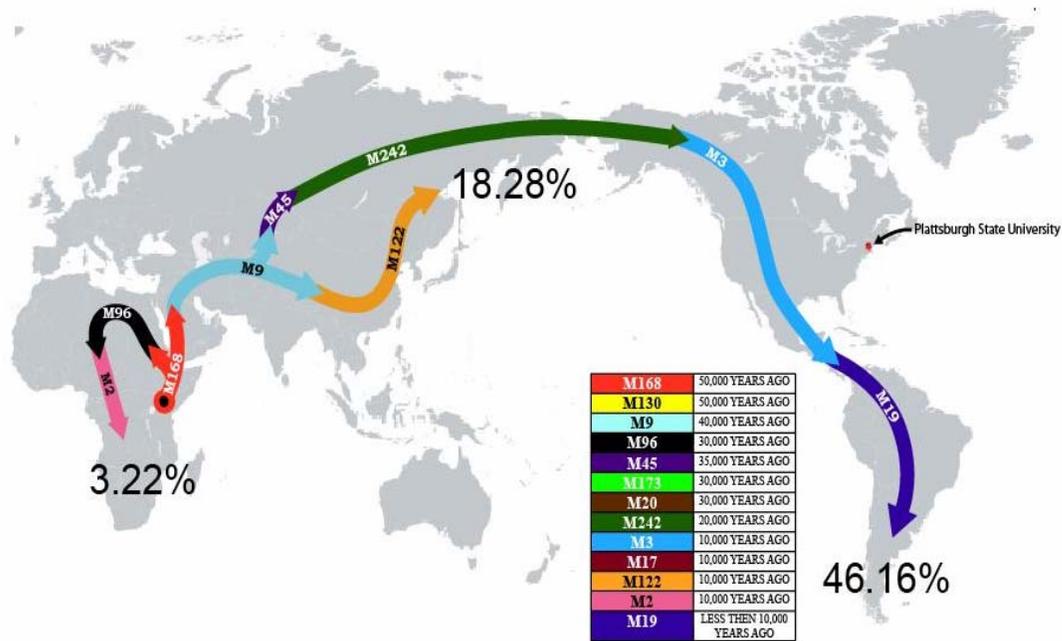


Figure 4. Results for three of the proposed migration routes from our collected samples. The percentages represent the percentage of samples that reached an endpoint for that particular migration marker.

RETRACING THE MOVEMENT OF GENETIC LINEAGES AROUND THE WORLD

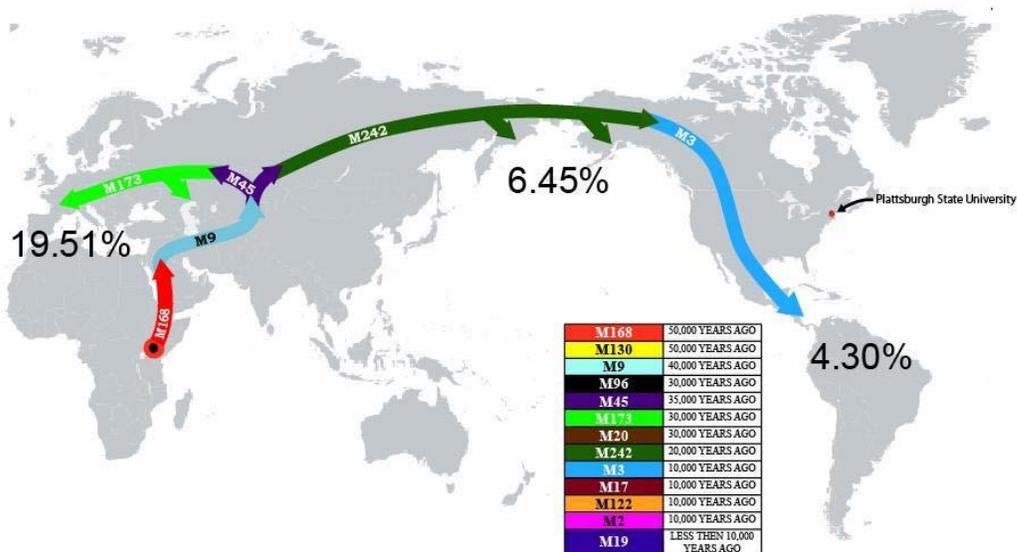


Figure 5. Results for three of the proposed migration routes from our collected samples. The percentages represent the percentage of samples that reached an endpoint for that particular migration marker.

References

- Cavalli-Sforza, LL. Genes, people, and languages. *Pro. Natl. Acad. Sci.* 1997, 94:7719-7724.
- Jobling, MA, and Tyler-Smith, C. The Human Y Chromosome: An Evolutionary Marker comes of age. *Nature reviews* 2003, August, 4:598-612
- Kunzig, R. History of men. *Discover.* 2004 Dec; 25 (12):32-39
- Lewis, R. Y envy. *The Scientist* (2003). 17 (15): 64.
- Seielstad MT, Hebert JM, Lin AA, Underhill PA, Ibrahim M, Vollrath D, Cavalli-Sforza LL. Construction of human Y-chromosomal haplotypes using a new polymorphic A to G transition. *Hum Mol Genet.* 1994 Dec;3(12):2159-61
- Seielstad, M, Yuldasheva, N., Singh, N., Underhill, P., Oefner, P., Shen, P., Well, R.S., A novel Y-chromosome variant puts an upper limit on the timing of first entry into the Americas. *Am J Hum Genet* 73:700-705 (2003).
- Shnayerson, M. The Map of Us All. *National Geographic Adventure.* Pages 80-83; 89. August 2005.
- Smolenyak, MS, Turner A. (2004). Trace Your Roots with DNA. United States of America: Holtzbrinck Publishers.
- Underhill PA, Jin L, Zemans R, Oefner PJ, Cavalli-Sforza LL. A pre-Columbian Y chromosome-specific transition and its implications for human evolutionary history. *Proc Natl Acad Sci U S A.* 1996 Jan 9;93(1):196-200.
- Underhill PA, Jin L, Lin AA, Mehdi SQ, Jenkins T, Vollrath D, Davis RW, Cavalli-Sforza LL, Oefner PJ. Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. *Genome Res.* 1997 Oct;7 (10):996-1005.
- Underhill, PA, Shen P, Lin A, Jin L, Passarino G, Yang W, Kauffman E, Bonne-Tamir B, Bertranpetit J, Francalacci P, Ibrahim M, Jenkins T, Kidd, J, Mehdi Q, Seielstad, M, Wells, S, Piazza A, Davis R, Feldman M, Cavalli-Sforza, LL, Oefner P. Y chromosome sequence variation and the history of human populations. *Nature Genetics.* 2000 Nov; 26: 358-361
- Underhill A, Passarino G, Lin AA, Shen P, Mirazon Lahr M, Foley RA, Oefner PJ, Cavalli-Sforza LL. The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. *Ann Hum Genet.* 2001 Jan;65(Pt 1):43-62
- Wells RS, Yuldasheva N, Ruzibakiev R, Underhill PA, Evseeva I, Blue-Smith J, Jin L, Su B, Pitchappan R, Shanmugalakshmi S, Balakrishnan K, Read M, Pearson NM, Zerjal T, Webster MT, Zholoshvili I, Jamarjashvili E, Gambarov S, Nikbin B, Dostiev A, Aknazarov O, Zalloua P, Tsoy I, Kitaev M, Mirrakhimov M, Chariev A, Bodmer WF The Eurasian heartland: a continental perspective on Y-chromosome diversity. *Proc Natl Acad Sci U S A.* 2001 Aug 28; 98 (18):10244-9.
- Yuehai Ke, Bing Su, Xiufeng Song, Daru Lu, Lifeng Chen, Hongyu Li, Chunjian Qi, Sangkot Marzuki, Ranjan Deka, Peter Underhill, Chunjie Xiao, Mark Shriver, Jeff Leil, Douglas Wallace, R Spencer Wells, Mark Seielstad, Peter Oefner, Dingliang Zhu, Jianzhong Jin, Wei Huang, Ranajit Chakraborty, Zhu Chen, and Li Jin. African Origin of Modern Humans in East Asia: A Tale of 12,000 Y Chromosomes. *Science* May 11 2001: 1151-1153