

Comparative Analysis of Direct to Consumer Genetic Testing Kits

Kristina Amos, Nate Vance, Sara Dabo*, Suad Ali*, Morgan Hubbard*, & Dr. Kerry Cheesman

Biological Sciences Dept., Capital University, Columbus, OH

Abstract

Direct to consumer (DTC) genetic testing kits are becoming increasingly popular; as a result concerns about what results mean for consumers are showing up in the media. The purpose of this study was to evaluate DTC genetic testing kits to determine inter-test reliability. The companies of interest included Ancestry.com, 23andMe, and MyHeritage. 23 participants, 12 females and 11 males representing a variety of ethnic groups, took the three independent DTC tests for ancestry lineage. Collection procedures specified by each company were used. Results for each participant were compared across companies, with similarities and discrepancies in identical categories being noted. Of the 23 participants in the study, 21 had a significant discrepancy in at least one category, and 8 had large discrepancies in two or more categories. For example, 23andMe reported 40.4% Native American ancestry for one participant while MyHeritage reported 4.2%. These discrepancies point out the unreliability of this consumer science, and can create complications for individuals who assume that the one DTC test they used is an accurate portrayal of their ancestry, and therefore their identity in one or more ethnic groups. Work needs to be done to ensure that consumer DNA reports are indeed accurate reflections of genetic ancestry.

Background

Direct to consumer (DTC) genetic testing kits originated in 2000 by Family Tree DNA, and by the year 2020 these kits had grown into a multi-billion dollar industry with close to 40 individual DTC genetic testing kits available on the market. There are many reasons why an individual might want to take a DTC genetic ancestry test, ranging from curiosity, to a desire to learn about their personal history and culture. Most individuals only purchase one kit and assume that the results obtained are accurate and portray their ancestry appropriately. In some cases individuals use this information to change the perceptions they have of self and family, without knowing the accuracy of the results.

An investigative report performed by The Doctors in 2017, which used three sets of identical triplets to test various ancestry DNA kits, found that two out of the three sets of triplets had at least some results that differed from their siblings even though they share the same genetic makeup¹. This study demonstrates that even with the large advancements made in genetic testing, all results cannot be taken at face value.

The process in which all DTC genetic testing companies compare customer samples are through their own reference databases known as Ancestry Informative Markers (AIMs)². Essentially, the more AIMs that are used for estimates of genome ancestry the more accurate the estimate is presumed to be³.

Jennifer Wagner⁴ described some hypothetical scenarios where major issues may arise from consumers performing the DTC genetic ancestry testing kits. One such scenario explores the cultural implications of an African American man who, growing up, had no knowledge of his paternal lineage. After taking an ancestry testing kit he learns that he has some European ancestry on his paternal side of the family. In this scenario the outcome wasn't too bad, but Wagner notes that given the large numbers of individuals looking at results from their own tests some might become very unhappy about what they see. For instance, what happens when the DNA report of ancestry differs from the oral tradition within a family? Which is more reliable? DNA (even if we don't know how the answers were calculated) or first-hand family knowledge?

The purpose for this study was to compare three of the most popular DTC genetic testing kits, using the same set of volunteers, to determine what similarities and differences exist in the reports produced. Based on the study of triplets published in 2010¹ and a preliminary study reported from this lab⁵, a hypothesis was established that most individuals tested would have significant differences in at least some ethnic backgrounds, and that these differences would be seen regardless of the overall ethnicity of the individuals.



Methods

Data was collected by having 13 females and 9 males take the DTC genetic testing kits from Ancestry DNA, 23andMe, and MyHeritage and report the results back. Subjects followed the instructions that were provided within each kit on how to provide their DNA samples. The subjects then provided the percentage breakdown of their results to the researchers. Data points were chosen by cross referencing each subjects' percentage breakdown from all three companies to identify any possible discrepancies. Discrepancies were based off of the labeling that each DTC genetic testing company used for their regions, if labeling does not match then general geographical location was used to match percentages together. Then once the percentages are lined, any data points that had a difference of 5% or more between the three companies is considered a discrepancy.

Results

Percentage composition results for each subject was the data that was collected. For the purpose of this study the data that was most closely analyzed was the major discrepancy within each participants set of results between the three DTC genetic ancestry companies represented in Table 1. Figure 1 is an example of one of the participants geographical breakdown of their DNA ancestry results from one of the companies along with their percentage breakdown.

Table 1- Major findings in DNA Ancestry tests among participants

Participant	Sex	Self-Reported Ethnicity	23andMe	Ancestry.com DNA	MyHeritage
1	F	African American		13.0% Nigerian	46.6% Nigerian
2	M	African American		9.0% Nigerian	46.5% Nigerian
3	F	Hispanic	76.5% Iberian	42.0% Iberian	59.6% Iberian
4	M	Hispanic	40.4% Native American		4.2% Native American
5	F	Asian	7.8% Japanese	100% East Asia	49.4% Japanese
6	M	Asian		100% East Asia	?
7	F	Hispanic/Caucasian	63.4% European	11.0% European	
8	F	African American/Caucasian	37.0% West African		22.5% West African
9	M	African American/Caucasian	?	22.0% Irish/Scottish/Wales	0.9% Irish/Scottish/Wales
10	F	Caucasian	93.7% North and West Europe	25.0% West European	88.5% North and Western European
11	F	Caucasian	68.7% British and Irish	69% British, 14.0% Irish/Scottish/Wales	
12	M	Caucasian		2.0% Scandinavian	23.6% Scandinavian
13	F	Caucasian	35.6% British and Irish	51% England Wales, and Northwestern Europe	23.3% English
14	F	West African/ Guinean	47.4% Coastal West African	92% Mali	56.9% West African
15	F	East African/ Somali	99.5% Somali	38% Africa South Central Hunter Gatherers	91.4% Somali
16	M	West African/ Senegalese	88.1% Guinean	50% Mali	71.4% West African
17	M	West African Guinean	68% Guinean	79% Mali	54.6% West African
18	F	East African/ Tanzanian & Kenyan	85.9% East African	53% Eastern Africa	70% Kenyan
19	M	Central African/ Cameroon	37.3% Congolese	100% Cameroon	69.1% Nigerian
20	F	African American	67.8% West African	13% Nigerian	46.6% Nigerian
21	M	African American	81.8% West African	9% Nigerian	46.5% Nigerian
22	F	African American	62% Sub-Saharan Africa	30% Cameroon, Congo and Southern Bantu	45.8% West African

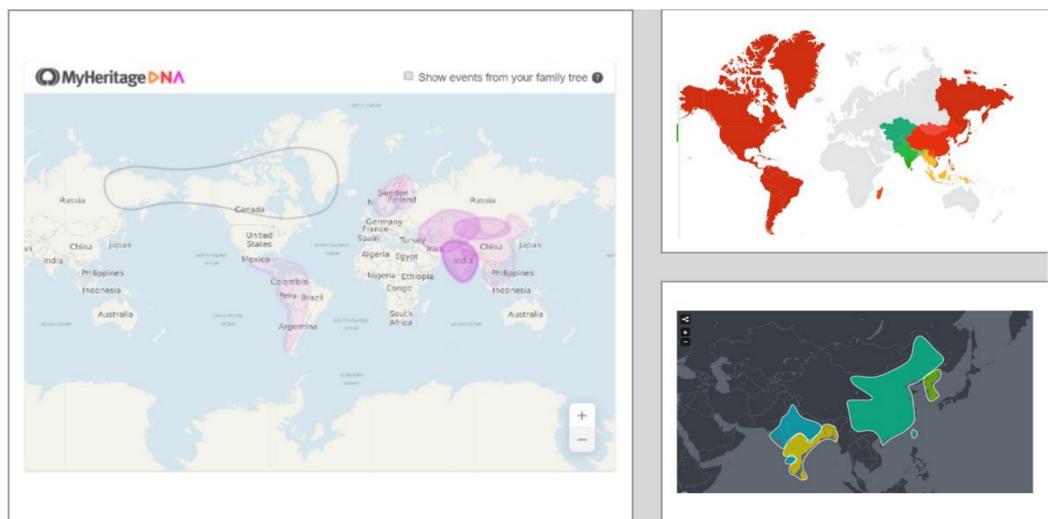


Figure 1- Example of Participant 6's geographic genetic ancestry between all three DTC genetic testing companies. MyHeritage (left), 23andMe (top right), and Ancestry.com (bottom right)

Discussion

DTC genetic ancestry testing kits have only been available to the public for about twenty years. During this time each company has gathered their own reference samples and produced their own computer algorithms for determining ethnic and/or geographic ancestry. Various assumptions have been made by scientists and computer programmers; since the reference DNA comes from living people and not ancient ancestors, these assumptions are very important and are derived from the available pool of evidence, which changes over time.

DTC genetic ancestry kits are not inexpensive (ranging from \$49. to over \$500.). Although no data can be found about consumer purchases, it is assumed that the vast majority of consumers purchase (and therefore rely on the reporting from) a single test. If all of the tests were equivalent that would not be a problem, but if there are differences in assumptions made within the algorithms used by each company then that can become an issue. In the current study it was found that discrepancies occurred between reports received by 21 or 23 participants, thus making this a very real concern. Even if one looks at all three reports together it is often hard to get a true picture of ancestry.

Because these tests are advertised as scientifically-based and accurate, it is safe to assume that most consumers will accept whatever the report shows even if it does not match what has been handed down through oral tradition within the family. This is evidenced by Ancestry.com's own television commercial, in which a gentleman who thought he was Scottish (from family tradition) found out he was German (from the test results) and "changed his kilt for a lederhosen". In other cases the results might cause confusion or even in-fighting within a family. In other words, the way in which individuals or even families see themselves ethnically may indeed depend on which particular ancestry DNA test is used.

Another implication of these genetic tests and their reported results may have to do with legal status. For example, the 23andMe report for Participant 4 shows the individual's Native American ancestry to be 40.4%, whereas MyHeritage reported Native American ancestry at 4.2%. If this individual were to use the former statistic they would be able to apply for tribal membership and apply for multiple, if not all, federal aid programs for schooling or health programs for Native Americans. However, if this individual used only the MyHeritage report (4.2% Native American), he would not qualify for any of these benefits.

As individuals fill out Census Bureau forms, college applications, and many other federal or state forms, deciding which ethnicity box to check may depend not only on family heritage but, increasingly, on reports received from various DTC genetic testing companies. Since these reports can lead individuals to new and sometimes very different conclusions about themselves and their families, it is important that consumers know and understand the limitations of these tests and the still growing science of genetic ancestry testing. It is also important that these companies share data so that the best, scientifically-accurate, information is available to all.

Acknowledgements

Funding for this research was provided by the Robert M. Geist Endowed Chair in the Biological Sciences at Capital University. Additional support came from the Biological and Environmental Sciences Department. The authors would also like to extend thanks to the faculty advisor for providing guidance and support throughout the project.

References

1. Triplets Put Ancestry Kits to the Test, The Doctors, (March 2017). Retrieved from <https://www.thedoctorstv.com/articles/3863-the-triplets-put-ancestry-kits-to-the-test>.
2. Krinsky, S., Johnston, D.C. (2017). Ancestry DNA Testing and Privacy: A Consumer Guide. Council for Responsible Genetics, (March 2017):1-43.
3. Pardo-Seco, J., Martinon-Torres, F., Salsa, A. (2014). Evaluating the Accuracy of AIM Panels at Quantifying Genome Ancestry. BMC Genomics 15:543-554.
4. Wagner, J.K. (2010). Interpreting the Implications of DNA Ancestry Tests. Perspectives in Biology and Medicine 53:231-248
5. Dabo, S., Ali, S., Cheesman, K. (2019). Analysis of DNA ancestry results for individuals of African descent. Ohio Academy of Science Annual Mtg, (Poster Presentation)