Published in final edited form as:

New Genet Soc. 2019 ; 38(2): 165–194. doi:10.1080/14636778.2018.1562327.

Consumer (dis-)interest in Genetic Ancestry Testing: The roles of race, immigration, and ancestral certainty

Adam L. Horowitz,

The Edmond J. Safra Center for Ethics, Tel Aviv University

Aliya Saperstein,

Department of Sociology, Stanford University

Jasmine Little,

Department of Sociology, Stanford University

Martin Maiers,

Bioinformatics Research Department, National Marrow Donor Program

Jill A. Hollenbach

Department of Neurology, University of California, San Francisco

Abstract

Genetic ancestry testing (GAT) is marketed as a way to make up for missing knowledge about one's ancestry. Previous research questions the GAT industry's ability to fulfill this promise in terms of the validity and reliability of test results. We instead explore the demand side of GAT, evaluating who is most and least likely to express interest in GAT. Using data from an original, nationwide survey of over 100,000 American adults, we find that GAT interest is related to both self-identified race and immigrant generation, with Asian Americans and first-generation immigrants expressing the least interest. Our quantitative and qualitative evidence suggests interest is further shaped by a pre-existing sense of ancestral certainty, leading some individuals to decline GAT, even if it were free. How interest and ancestral certainty are patterned has implications for who is included in – and thus for the conclusions that can be drawn from – genetic ancestry databases.

Keywords

Race; Immigration and Assimilation; Consumer Genomics

Introduction

Direct-to-consumer genetic ancestry testing (GAT) is marketed as a way for people to remedy missing knowledge about their ancestry. On television, radio, and the Internet, GAT advertising has become commonplace, offering potential consumers the opportunity to answer the question, “what are you?”¹ by “find[ing] out what your DNA says about you.”²

In the growing literature on GAT, several studies have evaluated the validity of genetic ancestry technologies and the veracity of the industry's marketing. This analysis explores a different side of the equation: if GAT marketing uses uncertainty in order to pique interest, then for whom does that message most resonate?

Our analysis first explores how much a sample of American adults reports knowing about their ancestry, and what characteristics are associated with expressing interest in GAT. To do so, we draw on a survey of over 100,000 U.S. residents who were asked a battery of questions about their race, ancestry, and knowledge of their family background, along with whether or not they would be interested in taking a genetic ancestry test. Leveraging combined quantitative and qualitative data, we identify several characteristics associated with interest in GAT. We find that older and more educated Americans, as well as self-identified white and black Americans, are most likely to be interested in GAT. This supports previous studies' prediction that "genetic genealogy" draws a similar demographic user base to that of traditional genealogy (Greely 2008; Nelson 2008a; Roth and Lyon 2018).

What is obscured by this focus on who uses GAT, however, is who does *not* use GAT. Our results highlight immigrant generation as an important factor that is often implied in previous research but left explicitly unexplored. We find GAT interest is concentrated in the third-or-later generations – especially among self-identified white and black Americans – with first- and second-generation respondents having less interest in GAT. We further find that this pattern is a consequence of the extent to which one feels certain or uncertain about their family origins – what we term "ancestral certainty."

Our results suggest that later-generation residents of long-standing immigrant receiving countries are the most likely consumers of genetic ancestry testing. Their sense of uncertainty about family ancestry is a feeling significantly less often echoed by their first- or second-generation counterparts. More broadly, higher rates of expressed ancestral certainty, combined with higher rates of explicit disinterest in GAT, mean that some populations – especially self-identified Asian Americans, but also first- and second-generation black Americans – are likely to be under-represented in genetic ancestry databases. These patterns of representation have implications for the validity and reliability of the ancestry results, and the generalizability of research drawn from ancestry databases.

Background

The genetic ancestry testing industry has experienced exponential growth since its launch in the early 2000s. Nevertheless, scholarship on the subject has a limited grasp on who is drawn to GAT and why. The literature instead has two major focuses. First, it analyzes the content of GAT marketing. Second, it considers the broader social impacts of the technology. Our study expands empirical knowledge of where GAT interest is concentrated, and illuminates the roles of self-identified race, immigrant generation, and ancestral certainty in driving demand. In doing so, we also advance a theoretical argument about the social processes that produce a market for GAT. We conclude that immigration and assimilation

¹[23andMe.com](#) homepage, January 6, 2018.

²[AncestryDNA.com](#) homepage, January 6, 2018.

patterns influence ancestral certainty – which in turn influences GAT interest – and that migration context is therefore a telling indicator of where a GAT consumer market is likely to thrive.

The market for genetic ancestry testing—The industry for personalized genetic information has rapidly expanded over the past two decades (Borry, Cornel, and Howard 2010; Lee 2013), including the market for genetic ancestry testing. The GAT industry was valued at \$173 million in 2016, up exponentially from its valuing of \$15 million in 2010, and is expected to surpass a \$350 million valuation by 2020 (Keshavan 2016). This growth matches the expansion of a broader industry for direct-to-consumer genetic testing (DTC-GT) services – and still, within the broader industry, GAT is predominant in both supply and demand. According to current estimates, there are 74 companies offering DTC genetic ancestry testing (Phillips 2016) and 30 companies offering DTC genetic services for health (Chow-White et al. 2018). GAT is the second-most common DTC-GT service offered, with the first being parentage confirmation. On the demand side, Roberts et al. (2017) denote that ancestry information is the primary interest of potential DTC consumers', present among 74% of their sample. One company alone, AncestryDNA, reported a revenue of \$683 million in 2015 and, by 2017, had sold its services to over four million consumers (Helft 2017; Williams 2017; Woodall 2016).

In spite of the industry's growth, scholarship provides little information about the characteristics of GAT consumers (Borry and Howard 2008; Wagner et al. 2012). Within scholarship on DTC industries, the composition of the consumer market is considered the least-well understood aspect of direct-to-consumer (DTC) services (Borry and Howard 2008), and, beyond this, GAT is considered to be an "entire sector" that has been "generally omitted" from review (Wagner et al. 2012:586).

There is some evidence suggesting that, overall, DTC-GT consumers tend to be older and wealthier than the general population (Covolo et al. 2015; Goddard et al. 2009; Ostergren et al. 2015), and this is thought to be true of GAT consumers, as well (Greely 2008; Nelson 2008a). Beyond this, a pair of survey analyses provide somewhat inconsistent information about the racial backgrounds of GAT consumers. Hochschild and Sen's (2015) analysis the 2010 General Social Survey – which they assert is the only time a nationally-representative survey has collected data on GAT views – shows that black Americans are the group most likely to be supportive of GAT. Roth and Lyon's (2018) survey of individuals who have previously taken a genetic ancestry test nevertheless indicates that whites may be overrepresented among GAT consumers, with over 70% of their respondents identifying as white, and under 9%, 7%, and 2% identifying as Latino/a, black, or Asian, respectively. While these analyses provide a mixed view of differences in GAT interest by race, it is nevertheless broadly accepted that race plays an important role within the GAT market, particularly in how the industry markets to potential consumers.

GAT and race—GAT companies use genetic sequencing and genotyping to provide a probabilistic assessment of geographic ancestry based on an individual's haplotype (Jobling, Rasteiro, and Wetton 2016). By comparing against a large database of individual DNA samples and the genomic variations and mutations among those, GAT companies promote

their ability to discern the combination of ancestries represented in a consumer's gene pool. Although companies are careful to avoid language of race, how they advertise makes it easy for customers to think the tests are indicating racial identities (e.g., Bolnick et al. 2007; Duster 2011; Fujimura and Rajagopalan 2010; Nordgren and Juengst 2009). In clear reference to American consumers' familiarity with selecting a racial category on standardized forms, for instance, one AncestryDNA commercial features a consumer stating that, after receiving her GAT results, she "look[s] at forms now and wonder[s], 'what do I mark?'"³

Multiple evaluations of GAT advertising strategies further show that the industry gives heavy focus to race in seeking new customers. Companies, for instance, create internet-based advertising targeted towards specific groups (Einsiedel and Geransar 2009; Lee 2013). The industry gives special attention to black consumers, depicting GAT services as the solution to the "brick wall" issue of African genealogy, wherein consumers can locate their African roots with geographic specificity in a way that wouldn't have been previously possible due to information lost in the trans-Atlantic slave trade (Greely 2008; Nelson 2008b).

Popular media similarly promotes this relationship between GAT and race. Hochschild and Sen (2015) show that media accounts focus overwhelmingly and disproportionately on black Americans, with about two-thirds of testimonial accounts featuring black consumers. Roth and Lyon (2018) similarly describe a journalistic tendency to highlight that consumers use GAT to make claims to racial or ethnic groups, such as for the purpose of seeking Native American tribal membership.

The content of advertising and popular coverage is consequential because these media appear to be effective in growing the industry. In Roth and Lyon's (2018) study, over two-thirds of respondents stated that they learned about GAT through company advertisements or popular media, and Chow-White et al. (2018) find that public perceptions (i.e., on social media) of one GAT provider, 23andMe, are overwhelmingly positive. Whether existing interest or sales resulting from targeted advertising is chicken or egg, what is clear is that race is taken as a powerful indicator of potential GAT consumerism.

A growing body of work highlights the social consequences of these explicit and implicit links between GAT and racial group membership. Some scholars have raised concern about how contemporary research may be influencing a "molecular reinscription" of race (Duster 2015). These scholars argue that seeing racial classifications as genetically determined reinforces essentialist thinking – notions that individuals' abilities are genetically preordained and encoded along racial lines (Byrd and Hughey 2015; Sarkar 1998). Indeed, experimental evidence shows that being exposed to DTC-GT services heightens the likelihood that American adults will espouse determinist and essentialist attitudes about race, including the belief that racism is genetically justified, in comparison to someone without such exposure (Condit et al. 2004; Hochschild and Sen 2015; Phelan et al. 2014). The potential for renewed perceptions of race as a biological distinction rather than a matter

³AncestryDNA commercial, aired 2016-2018. Accessed through ispot.tv, November 26, 2018: <https://www.ispot.tv/ad/djTa/ancestrydna-testimonial-livie>

of social experience contradicts the understanding of race as socially constructed – a view that has reached near consensus in many disciplines (Alba 1985; Cornell and Hartmann 2007; Feldman 2010; Lewontin 1972; Shim, Rab Alam, and Aouizerat 2018).

Race, immigration, and ancestral certainty—We posit that there is additional nuance to consider in order to reach a full understanding of the processes shaping the GAT industry, particularly with regard to the relationship between race and immigration. Race and immigration function in close relationship within American society, as sociological literature has long established (Gans 1979; Gordon 1964; Waters 1999). We argue that GAT interest is not only a matter of race, per se; instead, immigration and assimilation patterns shape feelings of group membership and ancestral (un-)certainty, which in turn influence GAT interest.

Differences in the timing and composition of immigrant arrivals to the U.S. are key to our argument. The height of European immigration was in the 19th and early-20th centuries (Alba 1990). Over the course of several generations, European-origin immigrant groups “mixed” with people of different immigrant backgrounds and forged new racial and ethnic identities (see, e.g., Hout and Goldstein 1994). Although distinctions between, for instance, “Italian-Americans” and “Irish-Americans” were once robust, many in later generations became “unhyphenated” whites (Lieberson 1985). Each generation of descendants of European immigrants were increasingly likely to intermarry, and specific ancestral attachments became more unknown, more distant, and less salient (Gans 1979; Waters 1990). This knowledge vacuum and its associated sense of uncertainty about one’s precise ancestry is what GAT promises to fill.

By comparison, immigrants in the second half of the 20th century and beginning of the 21st come from a broader swath of countries in Asia, Africa, and Latin America. These immigrants’ relationship to race and ethnicity is markedly different from that of earlier European immigrants, with their process of integration into American life not marked by the same “straight-line” trajectory of assimilation into the native (majority white) population (Portes and Zhou 1993; Lichter, Qian, and Tumin 2015). Consistent immigration flows from Asia and Latin America over the past 50 years also mean that more Asian Americans and Latino/a Americans descend from more recent immigrant generations, compared to their European American and African American counterparts.

However, clearer understandings of family origins may not only stem from greater recency of arrival. For example, same-race, but mixed nativity, marriages have increased among Asian Americans (Qian and Lichter 2007, 2011). This endogamy, and continually replenished waves of immigrants with the same origins, may serve to reinforce non-white status for many (cf. Jiménez 2010). For Asian immigrants and their children, these patterns also serve to maintain group boundaries and cultural knowledge (Qian, Glick, and Batson 2012), which may leave less room for uncertainty about family origins. In contrast, Latino/a immigrants – even those who arrived within the same period and marry endogamously – may experience levels of ancestral uncertainty more similar to that of European Americans because of national narratives in many Latin American countries celebrating their “mixed” populations (e.g., Amado 2012; Smith 1997).

We expect that these variations will produce differences in ancestral certainty that function, in part, in relationship to immigrant generation. Those most generationally-proximate to immigrants and who believe themselves to have the most homogenous ancestries will be most likely to report the greatest ancestral certainty and, in turn, the least interest in GAT. Specifically, we expect that self-identified Asian Americans will be the least interested in GAT, both because of their relative proximity to the immigration experience and stronger endorsement of perceived homogenous ancestry in their origin countries. However, we also expect that ancestral certainty will be a key predictor of GAT interest above and beyond self-identified race or immigration generation.

Materials and methods

Data source

To assess interest/disinterest in genetic ancestry testing, we draw on survey data from nearly 110,000 U.S. adults who were registered as potential volunteer bone marrow donors with the National Marrow Donor Program (NMDP). The survey was conducted as part of a larger study examining concordance between self-reported race and ancestry and genetic measures for the purpose of improving the process of donor-recipient transplant matching (see, e.g., Hollenbach et al. 2015). Participants were recruited through email invitations sent between May and July 2015 to nearly 2 million potential donors, of which approximately 5% completed the survey.⁴

This survey presents a strategic opportunity to assess interest in GAT. The survey collects extensive information about what respondents know about their ancestry, including measuring the degree of respondents' certainty about these details. Respondents reported whether they, their parents, and their grandparents were born in the United States or elsewhere; the countries or parts of the world from whence their ancestors came, along with estimated ancestry proportions from each place reported; specific questions about the ancestral origins of each biological grandparent; how much they know about the maternal and paternal sides of their family; and what actions (if any) they have taken to seek out ancestry information.⁵

The survey further included two simple but direct measures of interest in GAT. First, respondents were asked if they had previously taken a genetic ancestry test. Approximately 5% of respondents said they had. The rest were posed a hypothetical question: "If you were offered a free genetic ancestry test, would you be interested in taking it?" The test was described explicitly as free with the aim of removing financial circumstances as a reason to opt into or out of GAT. Respondents who said they would not be interested in a free genetic ancestry test were then asked to provide the reasons why they were not interested in an open-ended written response.⁶

⁴A low response rate is not uncommon for web-based surveys and does not indicate lower data quality (see Fan and Yan 2010 for a review), but it does necessitate caution in interpreting the generalizability of our results.

⁵The survey design also included several randomized features, including testing four different questionnaire orders and two different email invitations. Controls for these potential design effects, as well as for whether the respondent completed the survey after receiving their initial email invitation or one of two follow-ups, are included in all of our multivariate analyses.

⁶Given that respondents knew they were responding to a survey about genetics and ancestry before they elected to participate, we interpret our results on disinterest in GAT as conservative.

The survey also gathered information on respondent's self-identified race. Respondents were asked to indicate the racial or ethnic group or groups they use to describe themselves, using the categories recommended by the federal government for racial/ethnic data collection (Office of Management and Budget 1997). Compared to the U.S. population as a whole, our sample over-represents people who identify as white alone as well as people who selected multiple racial or ethnic categories (see Table 1). The former reflects the greater proportion of self-identified whites who are registered with the National Bone Marrow Program (Gragert et al. 2014); the latter is likely a result of describing the study as one about race, ancestry, and genetics that aimed to inform how people are matched for transplants.⁷

People over age 65 are not eligible to be bone marrow donors, so the NMDP registry – and thus our sample – only includes people aged 18 to 64. Within that age range, respondents aged 25 to 44 are somewhat over-represented, as might be expected in an online survey (see, e.g., Börkan 2010), as are female respondents and people with advanced degrees (see Table 1). However, the regional distribution of respondents compares favorably to contemporary estimates of the adult population ages 18-65 from the American Community Survey.

Although the demographic makeup of respondents and non-random nature of the sample necessitates caution in generalizing from our results, the data are uniquely suited to explore the connections between GAT interest and self-identified race, nativity, and ancestry. The wide array of information on biogeographic origins, racial identification and ascription, and ancestral knowledge is unavailable in standard nationally representative surveys. The large sample size and overall diversity of the respondents are also strengths. For example, few nationally representative surveys include enough Asian Americans for detailed statistical analysis, including a breakdown by immigrant generation. Finally, drawing on both quantitative and qualitative data provides a multidimensional approach to evaluating GAT interest, allowing us to explain interest both in terms of broad demographic patterns and in the words, thinking, and rationales of potential consumers themselves.

Analytical approach

Our analysis iteratively combines quantitative and qualitative data to explore both the frequency and correlates of ancestral certainty and how these relate to interest (or lack thereof) in genetic ancestry testing. We also estimate a series of logistic regressions predicting: GAT (dis)interest, ancestral certainty, and, finally, GAT disinterest controlling for ancestral certainty.⁸ These multivariate analyses allow us to determine whether ancestral certainty is associated with interest in genetic ancestry testing above and beyond known interest-driving factors such as sex, age, and socioeconomic status. Our analytical sample for these models includes the 91,729 respondents who had no missing data on any of our key characteristics.

⁷Skepticism about race and genetics, especially in a health-related context (e.g., Gamble 1997; Mello and Wolf 2010), may have led some who were invited to take the survey opt out of participation entirely.

⁸We use logistic regression because we coded our outcomes as binary (i.e., expressing disinterest or not, expressing certainty or not), but the results are substantively similar when we estimate linear probability models instead (results available upon request). Thus, our conclusions are not driven by the functional form of our model.

Measuring interest in GAT—Our measure of interest in genetic ancestry testing distinguishes three categories of respondents: those who said they have already taken a GAT, those who expressed interest in taking a GAT if it were free, and those said they were not interested in GAT. Our descriptive analyses explore variation across these three groups, while our multivariate models alternately compare people who have already taken a GAT to those who have not, or compare people who said they are not interested in GAT to those who are. Approximately 2% of the sample said they were not interested in taking GAT, even if it were free. Of those respondents, 77% (1,448 of 1,875) went on to provide one or more reasons for their disinterest.⁹

Throughout, we supplement our quantitative analysis using these written responses, highlighting patterns and themes in the explanations respondents provided for their disinterest. We analyzed the data provided in these open-ended responses following the iterative coding model of Lofland and Lofland (1995). The process began by identifying themes to capture the full spectrum of responses; we then returned to each response to code it according to one or more themes. To facilitate analyzing thematic predominance, we assigned each response a primary code based on the most prominent rationale for declining provided (with at least two coders agreeing on prominence in a small number of ambiguous cases). Quoted examples for each primary code are provided in Figure A.

Measuring immigrant generation—We coded immigrant generation based on the respondents' self-reported nativity information for themselves, their parents, and their grandparents. We define “first-generation” as someone who was not born in the U.S., whose parents and grandparents were also not born in the U.S. We define “second-generation” as someone who was born in the U.S. who has at least one foreign-born parent and at least one foreign-born grandparent. “Third-generation” was a U.S.-born respondent with U.S.-born parents who has at least one foreign-born grandparent, and “fourth-generation and up” refers to respondents who listed everyone in their immediate family as U.S.-born. Respondents who had missing data on one or more of the associated nativity questions, or gave responses that were inconsistent with this scheme (e.g., because they reported having foreign-born parents but only native-born grandparents), were dropped from the analysis.¹⁰

Measuring ancestral certainty—Our measures of ancestral certainty draw on closed-ended survey responses to questions about the respondents' ancestry and that of their immediate biological relatives. We created a set of four indicators to capture different aspects of ancestral certainty or knowledge of family history, partly inspired by our reading of the qualitative data.

Two of these measures focus on respondents' confidence in their ancestral knowledge. Respondents were asked to report their own ancestry, before they provided responses for select biological relatives, using a series of check-all-that-apply geographical origin

⁹Using the written responses, we recoded 50 respondents because they had already taken a genetic ancestry test or stated that they inadvertently selected the wrong answer to the GAT interest question.

¹⁰Respondents with missing data on immigrant generation (N=9,184) are more likely to be interested in GAT, all other measured characteristics being equal. Supplemental analyses that retain these cases, and include an indicator for their uncertain immigration generation in our multivariate models, provide substantively similar results to those present here (available upon request).

categories such as “Western Europe,” “Sub-Saharan Africa,” or “East Asia.”¹¹ Respondents could also select “Unknown.”¹² Our first measure of ancestral certainty indicates the respondents who did not select “Unknown” among their responses. The second is a direct measure of ancestral certainty based on a pair of questions about how much respondents know about their family history on their biological mother’s and biological father’s side. We combine responses to determine which respondents say they know “a lot” about both sides of their family (as opposed to knowing “a little” or “nothing at all”).

The other two measures of ancestral certainty are based on the information respondents provided about their biological grandparents. Our third measure indicates whether respondents were able to provide complete information on their grandparents’ ancestries, with respondents who provided one or more ancestries for all four of their biological grandparents considered to have more knowledge than others. Finally, we constructed a measure that reflects both whether or not all four grandparents were assigned a single origin category and whether those singular responses all matched; that is, whether respondents gave the exact same response for each grandparent. Although this may be more accurately considered a measure of ancestry homogeneity, we include it here because the qualitative responses suggested that having (or believing one had) singular ancestral origins was linked to a sense of ancestral certainty and disinterest in GAT.¹³

Of course, taking GAT is just one of several methods to which people can turn to learn about their ancestry – people interested in genealogy also take part in activities such as looking through family documents, using ancestry websites, and searching or sending away for material from library archives. The survey also asked whether respondents had taken these or any other actions to seek out ancestry information. Responses to this question allow us to account for respondents’ prior knowledge-seeking behavior, as we highlight the relationship between ancestral certainty and GAT interest.

Results

We begin by presenting the frequency of expressed interest in GAT, compared across various demographic characteristics. We then introduce our measures of ancestral certainty, and explore both their relationship to GAT (dis)interest and their demographic patterning in our sample, before weighing the relative importance of ancestral certainty alongside other respondent characteristics using multivariate models. Taken together, our analysis paints a nuanced picture of the predictors of GAT interest: our hypothesized mechanisms of immigrant generation and ancestral certainty exhibit significant associations with GAT in addition to, not instead of, other factors identified in previous research.

¹¹Each answer option included examples to both define the regions and help with borderline cases. For example, the Southern Europe examples were “(Italy, Spain, Turkey, etc.)” while the Middle East examples were “(Iran, Lebanon, Saudi Arabia, etc.)”

¹²A pilot study indicated this was an important option to offer, both for people who were adopted and had little or no knowledge of their biological relatives, as well as people of African descent whose specific family origins were obscured by the slave trade.

¹³We considered two additional measures of uncertainty that we do not present here because of their substantive similarity to the four measures we do discuss (results available upon request). These include a measure for reporting no unknown nativity and the proportion of unknown ancestry respondents reported for themselves.

Who is interested in GAT, and who is not?

We find that who has already taken GAT, who reports interest in GAT, and who reports disinterest in GAT, is strongly patterned by a host of demographic characteristics, including race, immigrant generation, age, and education. For example, Table 1 shows that, in our sample, self-identified Asian respondents and first-generation immigrants report no interest in GAT with the highest frequency. Meanwhile, self-described white and black respondents, along with those in third and higher immigrant generations not only report disinterest less frequently, but also have most frequently already taken a genetic ancestry test. The oldest respondents and those with advanced degrees report already taking GAT at high rates, confirming the presumptions of previous literature (Greely 2008; Nelson 2008a; Roth and Lyon 2018). Interestingly, though, older and more educated respondents more frequently express disinterest in GAT than those in other educational and age groups, as well.

The multivariate analysis in Table 2 confirms many of the bivariate patterns of GAT interest and disinterest seen in Table 1. This is particularly true in the case of self-identified Asian respondents: even when comparing people of similar sex, age, education, immigrant generation, and region of residence, Asian respondents remain significantly less likely to have taken a genetic ancestry test, and significantly more likely to state they are not interested in doing so. Otherwise similar respondents who identify as black alone, white alone, or as multiple minorities (i.e., “Two races, not including white”) are all equally likely to have taken a GAT or be interested in GAT if it were free. Multiracial respondents in our sample who also identify as white, or who chose three or more racial/ethnic categories to describe themselves, are significantly more likely than monoracial white respondents to have already taken GAT and significantly less likely to say they have no interest, while self-identified Latino/a respondents are significantly less likely both to have taken GAT and to express disinterest. These results support the idea that self-identified race is a key predictor of interest in genetic ancestry testing.

However, we also see patterns in Table 2 that were less obvious from the bivariate descriptive statistics. For example, echoing the relationships described for self-identified Latino/a respondents, female respondents and people living in the South are significantly less likely to have already taken GAT but also less likely to express disinterest (compared otherwise similar male respondents and people living in the Northeast, respectively). On the other hand, when we control for other factors that predict GAT interest, first- and second-generation respondents are significantly more likely to express disinterest in GAT than their counterparts who are third-generation or above, while first-, second-, and third-generation respondents are less likely than those in the fourth and later generations to have already taken a GAT.¹⁴ These results indicate that baseline patterns work in combination with (and may be moderated by) immigrant generation in shaping GAT interest.

¹⁴The estimate for second-generation respondents is similar in size and magnitude to the other two but is not statistically significant at conventional levels.

What are the sources of GAT disinterest?

We continue our analysis by exploring why certain respondents are more or less likely to be interested in GAT. To do so, we first draw inspiration from the written responses provided by those who said they would decline a free GAT, before turning back to our quantitative data.

Our coding of the qualitative responses revealed a variety of rationales for declining the hypothetical offer of a free genetic ancestry test. Respondents provided such explanations as having privacy concerns, too little time, insufficient information about what GAT would involve, or a general lack of interest. A relatively brief expression of disinterest was the most common response, while privacy issues and skepticism about test accuracy were among the least common responses. (See Figure A, ordered by response frequency.) Particularly prominent, expressed in approximately a quarter of the responses, was respondents' estimation that GAT simply would not reveal any new information. Respondents wrote that they declined the GAT offer, for instance, "because I know my ancestry definitively," because "I know my ancestry to a very high degree of precision," or because "I know the details of my ancestors with 100% confidence." What these respondents are expressing, in our parlance, is ancestral certainty.

These responses led us to focus our quantitative analyses on measures of certainty/uncertainty. Table 3 shows descriptively how several metrics of ancestral certainty are related to GAT (dis)interest. Among those who have not already done GAT, there are notable differences between those who express interest in GAT and those who do not. Those who are not interested in GAT more frequently said they know "a lot" about both sides of their family and reported having no unknown ancestry (i.e., they know their full ancestry), versus respondents who were interested in a free test. Moreover, compared to respondents interested in GAT, disinterested respondents more frequently reported ancestries for all four of their biological grandparents and reported that all four of their grandparents share the same single geographic origin.

The story is somewhat different for those who have done GAT in the past. Those who have already taken GAT report having homogenous ancestries less frequently than those who expressed disinterest, but report having no unknown ancestry at a similar rate, and even more frequently report all four grandparents ancestries and knowing "a lot" about both sides of their family. On the one hand, these patterns suggest that GAT likely increases both an individual's ancestral knowledge and their sense of certainty. Yet, some of these same elements can be impetus for seeing GAT as unnecessary.

We compared among respondents who declined GAT and highlight key distinctions between people who said they already know their ancestry and all other decline reasons. T-tests demonstrate that respondents who said they declined because they already know their ancestry, compared to those who provided any other reason for declining, are significantly more likely to report no unknown ancestry, say they know "a lot" about both sides of their family, have specific knowledge of all four grandparents' origins, and report that all four grandparents' origins are homogenous.

To be sure, the full group who declined GAT – no matter what written explanation they provided – was more likely to express ancestral certainty than those who reported interest in GAT. This suggests that ancestral certainty might have an influence on interest even if not stated directly. Combined, these results provide initial, descriptive evidence that ancestral certainty is related to GAT interest.

Who expresses ancestral certainty?

Our data further illuminate how individual demographic characteristics correlate with ancestral certainty. In their open-ended decline responses, those who asserted that they already know their ancestry frequently mentioned specific regions or countries of origin, and nearly half of those responses referenced an Asian country or region. For example, respondents said they were not interested in GAT because “I already know I’m full Japanese” or “I’m 100% Chinese.” The qualitative responses also included mentions of immigrant generation. One person said, for example, “My parents were the first in my ancestry to leave India, so I know that all my ancestors were Indian.” Further, some responses allude to the effect of homogeneity on interest. As one respondent wrote, “My entire family including myself is from China. It is extremely unlikely that I would have any other ethnicities mixed in.”

Following this lead from the qualitative responses, we return to our quantitative data to assess the connections between race, immigrant generation, and ancestral certainty. Self-identified Asian respondents were overrepresented in the decline group (5.1%) compared to the full sample (3.5%) and they were more likely to say they already know their ancestry compared to others who declined the test. Half of the survey’s Asian respondents are first-generation immigrants and 58% of Asian Americans in the decline group were foreign-born. First-generation Asian respondents also had more knowledge of their grandparents, with 80% reporting all four grandparents’ ancestries and 73% reporting grandparents with a single matching ancestry. These rates are not only higher than for people who self-identify with other races, but also higher even than the average rates among all foreign-born (i.e., first-generation) respondents.

We also use logistic regression to examine whether self-reported race and nativity are predictors of reporting ancestral certainty, net of other measured demographic characteristics. We find that respondents who self-identify as Asian, black, or Latino/a are all significantly more likely to express ancestral certainty than self-identified Whites, regardless of the certainty measure we consider. (See Table 4). The magnitude is particularly notable for self-reported Asian respondents, with odds of expressing ancestral certainty ranging from 1.2 to 4.6 times greater than otherwise-similar self-identified white respondents.

As hypothesized, we see significant variation in ancestral certainty by nativity. Compared to those in the fourth-and-higher immigrant generations, respondents in the first, second, and third immigrant generations are significantly more likely to express ancestral certainty on each certainty measure. Here, again, one group stands out, with those in the first generation having greater odds of expressing ancestral certainty than all of their native-born counterparts.

Does ancestral certainty drive GAT interest?

Thus far, we have shown qualitative and descriptive quantitative evidence suggesting that ancestral certainty is a meaningful factor shaping whether someone is likely to be interested or disinterested in GAT. Our preliminary multivariate analyses also suggest that self-identified race and nativity are associated with both ancestral certainty and GAT interest. These findings raise the question of whether all of the factors operate independently of one another, or whether – for example – ancestral certainty helps to explain the observed relationships between race, nativity, and GAT interest.

To speak to the relative importance of these factors, we estimate a final series of logistic regressions predicting disinterest in GAT that build to a full model, controlling for ancestral certainty, previous ancestral knowledge seeking, and our various measured demographic characteristics. We begin by examining our measures of ancestral certainty separately (Models 1-3) and then collectively (Model 4). Our final model, introduces controls for the types of non-genetic genealogical research respondents reported (Model 5). We do this to ensure that our predictions about GAT disinterest are not better explained by a general disinterest in knowing one's family ancestry.¹⁵

Overall, the models in this series demonstrate that ancestral certainty continues to be positively associated with GAT disinterest, net of respondents' demographic characteristics. (See Table 5.) We find that three of the four certainty measures (not having unknown ancestry, knowing all four grandparents' ancestry, and grandparent ancestry homogeneity) have statistically significant associations with disinterest in GAT, above and beyond self-identified race and immigrant generation. These patterns hold regardless of whether the certainty measures are included separately or simultaneously. Once we are also comparing between people who have done equal amounts of previous genealogical research (see Model 5), we find that reporting ancestries for all for grandparents, in and of itself, is not significantly associated with GAT disinterest; however, we also find that otherwise similar respondents who say they already know "a lot" about their family ancestry are significantly more likely to express disinterest in GAT.

Notably, in these models, some categories of both self-identified race and immigrant generation also retain significant associations with GAT disinterest. First- and second-generation immigrant respondents are significantly more likely to decline a free GAT even after controlling for reported ancestral certainty (see Models 1-4). Similarly, self-identified Asian respondents continue to be more likely to decline GAT, while Latino/a respondents and some multiracial respondents remain significantly more likely to express interest than self-identified white respondents, all else being equal. However, in the full model, when we introduce controls for previous ancestral knowledge seeking, the association between immigrant generation and GAT disinterest – though still positive for first- and second-generation respondents (i.e., with odds ratios greater than 1) – is no longer statistically significant at conventional levels ($p=.13$ and $p=.17$, respectively).

¹⁵Results are substantively similar if we exclude respondents who have already taken genetic ancestry tests from these models (results available upon request).

To interpret these findings it is helpful to return to the descriptive patterns related to ancestral knowledge-seeking shown in Table 3. Respondents who had already taken GAT reported the most knowledge-seeking activities, by far – suggesting that GAT is just one among many tools used by Americans interested in genealogy (Nelson 2008a; Nelson 2008b). Respondents who expressed no interest in GAT reported having engaged in the fewest knowledge-seeking activities on average. Yet, those disinterested in GAT also reported more ancestral certainty despite less knowledge seeking, indicating that their ancestral knowledge is not derived from greater genealogical research efforts.

Further, when we examined who reported any knowledge seeking compared to none, by a combination of self-identified race and immigrant generation, we found that knowledge seeking is not only negatively and monotonically associated with immigrant generation (first-generation does the least, followed by the second, and so on), but this pattern also holds regardless of self-identified race (results available upon request). This helps to explain why, when we control for knowledge seeking in the multivariate analysis (Table 5, Model 5), the relationship between immigrant generation and GAT disinterest is significantly attenuated. Among people with similar levels of previous knowledge seeking, ancestral certainty better predicts interest in GAT than immigrant generation. Put another way, greater expressed disinterest in GAT among first- and second-generation immigrants likely stems from a sense of ancestral certainty attained without engaging directly in genealogical research.

The relationship between ancestral certainty, ancestral knowledge seeking, and self-identified race is somewhat different than for immigrant generation. For instance, even in our final model, self-identified Latino/a respondents continue to be significantly less likely to decline (i.e., are more interested in) GAT than white respondents, and self-identified Asian respondents are still significantly more likely to decline (i.e., are less interested in) GAT. Before controlling for ancestral knowledge seeking, these estimates have nearly identical patterns in sign and statistical significance regardless of whether the ancestral certainty measures are included separately or simultaneously. When also controlling for knowledge seeking, the estimates are attenuated somewhat (e.g., Asian respondents who have done similar amounts of knowledge seeking as their non-Asian peers decline GAT somewhat less frequently), but they remain in the same direction and of generally similar magnitude. Thus, although ancestral certainty is significantly patterned by self-reported race (as shown in Table 4), neither their relationship nor other measured factors can fully account for either of their associations with GAT interest. Our multivariate models suggest that ancestral certainty and racial self-identification predict GAT interest largely – but not entirely – independently of one another.

Discussion

Our analysis illuminates the relationships between interest in genetic ancestry testing and various sociodemographic characteristics. The results support previous research suggesting that GAT consumerism draws heavily from the upper ends of age and education groups, and that there is some relationship between GAT interest and self-identified race (Greely 2008; Nelson 2008a; Roth and Lyon 2018). Yet, our findings also point toward a more nuanced

understanding of GAT interest. While interest is associated with traditional demographic factors, and works along with, and potentially through, differences in immigrant generation, ancestral certainty also plays an important role in shaping who participates in the GAT market and who does not.

Making sense of race, immigration, and GAT interest

At first glance, it is tempting to think that ancestral certainty (and, correspondingly, GAT interest) is neatly organized by race and nativity. Asian Americans in our sample report high ancestral certainty, in comparison to other populations, and a high degree of ancestral homogeneity. Third-or-later-generation Americans, meanwhile, report significantly less homogeneity and significantly more overall uncertainty about their family ancestry – a feeling that is less often echoed by their first- or second-generation counterparts. To be sure, self-identified white and black respondents are over-represented among third or later generation Americans, while the vast majority of self-identified Asian respondents are either first- or second-generation immigrants. But those compositional differences do not account for our results.

Instead, a more complex pattern is evident in Table 6, which presents observed reporting of ancestral certainty and GAT disinterest by both race and immigrant generation. In general, when comparing across rows, both ancestral certainty and GAT disinterest tend to decrease with each generation removed from the immigrant experience. This pattern is clearest for self-reported white and black respondents, although starting levels vary somewhat by racial self-identification. However, rates of certainty and GAT disinterest are strikingly similar among self-reported Asian respondents in every immigrant generation. Latino/a respondents also report both less ancestral certainty and more GAT interest than would be expected by immigrant generation alone.

These departures from the main patterns shown in our multivariate models can only be understood by considering both the general trends in immigration to the U.S. and the perspectives on race and ancestry that have the most currency in the countries from whence contemporary immigrants come. Thus, although we find that ancestral certainty tends to be increased by proximity to the immigrant experience, this is likely only the case when immigrants bring with them to the U.S. a sense that their ancestry is homogenous. As one respondent wrote in explaining their disinterest in GAT, “Chinese people have been considered the same group of ethnicity for thousands of years. I doubt we’ll get any results back other than ‘you are Chinese and so are your ancestors.’” For some contemporary Asian immigrants, this sense of homogeneity may be further strengthened through endogamous coupling facilitated by continually replenished waves of immigrants with the same geographic origins (Qian and Lichter 2007, 2011). Yet, immigrants from many Latin American countries arrive with the opposite expectation of their ancestry: a history of colonial and immigration-produced mixing in Central and South America inspired the widely celebrated – and critiqued – concept of *mestizaje*, which maintains that mixed ancestry rather than homogeneity is a source of national strength (e.g., Amado 2012; Smith 1997). This valorization came at the expense of maintaining ties to indigenous and African ancestry and traditions, leaving people with only a generic sense of being “mixed” and thus

a sense of uncertainty about the precise nature of their family origins. A similar process of broken ancestral ties likely also explains the popularity of GAT among third-and-later generation “white ethnics,” as well.

Thus, we argue, broader social and historical factors shape the uneven alignment of GAT interest by self-identified race and immigrant generation. Rather than race or nativity alone being the most consequential predictor of GAT interest, as we show in our analysis, one’s sense of ancestral certainty – no matter its source – must also be taken into account.

GAT as both product and producer

We have pointed to the context of immigration and assimilation as shaping ancestral uncertainty that, in turn, defines GAT interest. Simultaneously, it may be this context that situates the United States as a ripe environment for GAT consumerism. If the degree of ancestral “mixing” in a population correlates with the presence of uncertainty, then the extent to which American immigration has been shaped by a broad variety of sending countries over several hundred years, combined with assimilation patterns that increase intermarriage, would indicate a high presence of uncertainty within the US populace. This high presence of uncertainty, in turn, would suggest that the United States is an economic market with a large consumer base for GAT.¹⁶

This also means that GAT could have greater influence on conceptions of race and ancestry in the United States. Although we have spoken primarily of how GAT interest is patterned by self-identified race, several studies have shown that racial self-identification may also be an outcome of taking GAT (Golbeck and Roth 2012; Hirschman and Panther-Yates 2008). Thus, race can function recursively in the context of GAT. At another level, the racial-identity-related information that some consumers believe they receive is partly determined by the race and ancestry of those who previously chose to take the test. As Jobling and colleagues (2016) describe, one strategy by which companies infer ancestry proportions is by drawing comparisons among the genetic profiles of those in their databases. Delineating ancestry according to variation in the existing consumer base means that the composition of that consumer base affects where and how ancestry assignments are made, and the resulting proportions reported to consumers. If a population is over- or under-represented in a given database, as our results indicate, then that has implications for the results characterizing genetic ancestry that everyone receives.

Ancestral certainty should be viewed with a similar eye. As with race, there is a recursive relationship between GAT and ancestral certainty. The claim in GAT marketing that tests can alleviate ancestral uncertainty is not off the mark: as we have shown, uncertainty is indeed a consequential driver of interest, and those who have already taken a GAT report less ancestral uncertainty, on average, than those who expressed interest but have not taken a GAT. Yet, one’s sense of ancestral certainty is also subjective. Respondents’ reports of all four grandparents having the same ancestry, for instance, are not necessarily evidence that

¹⁶To that end, we expect other countries that already have multiple generations of immigration – like the United Kingdom, Canada, and Australia – to similarly produce GAT markets. There is indication that there are growing GAT markets in the aforementioned countries (Harris, Wyatt, and Kelly 2013), though a direct comparison based on the size of existing markets is mediated by differences in DTC regulation by country (Borrey and Howard 2008).

they do have identical and homogenous ancestry – only that the respondents believe that to be the case.

A respondent's perception of certainty may or may not be accurate, but it nevertheless funnels certain people into or out of GAT. This is consequential because genetic databases – including those compiled by GAT companies – are used for more than selling information about ancestry. Databases collected from consumers interested in ancestry information have been used, as Tutton (2007) describes, for “improving scientific understanding of the etiology of common diseases and developing the prescription of targeted pharmaceuticals, to securing valuable economic resources for countries seeking to compete in a global economy” (173). Genetic databases are routinely-used sources of information for academic genome-wide association studies, as well. A focus on uncertainty produces a form of self-sorting among GAT consumers that influences the composition of genetic databases – if those who (believe themselves to) have greater ancestral certainty are more likely to opt out of GAT, then GAT databases will underrepresent not only individuals with more perceived ancestral certainty but perhaps broader populations in which perceived certainty is more widespread. Our analysis suggests this is likely to be the case for both self-identified Asian Americans, and first- and second-generation immigrants from beyond the Americas. The patterns in GAT consumerism revealed in our sample – related to self-identified race, nativity, and ancestral certainty, but also age, sex, region, and education – give reason to proceed with caution when those GAT databases are used for other purposes. Although we cannot pinpoint the magnitude of these skews in actual GAT databases, the patterns we have identified highlight additional reasons for caution when attempting to make generalizable claims from ancestry databases.

In effect, the factors that produce interest in GAT are consequential because they are inextricably intertwined what GAT sends back into society as a product. People with particular backgrounds may be more or less interested in GAT, and GAT also shapes how people consider and report their race and ancestry. Likewise, ancestral certainty may push one to becoming a consumer of GAT, while the fact of being a consumer influences ancestral certainty. Yet, what occurs on the back-end, as ancestry databases are increasingly used to generate new knowledge, may be more far-reaching. Characteristics that shape GAT interest skew who puts their money into GAT, but the information drawn from GAT databases has consequences for everyone.

References

- Alba Richard D., ed. 1985 *Ethnicity and Race in the U.S.A.: Toward the Twenty-First Century*. London: Routledge & Kegan Paul.
- Alba Richard D. 1990 *Ethnic Identity: The Transformation of White America*. New Haven, CT: Yale University Press.
- Amado María L. 2012 “The ‘New Mestiza,’ the Old Mestizos: Contrasting Discourses on Mestizaje.” *Sociological Inquiry* 82(3):446–59.
- Bolnick Deborah A. et al. 2007 “The Science and Business of Genetic Ancestry Testing.” *Science* 318(5849):399–400. [PubMed: 17947567]
- Börkan Bengü. 2010 “The Mode Effect in Mixed-Mode Surveys: Mail and Web Surveys.” *Social Science Computer Review* 28(3):371–80.

- Borry Pascal, Cornel Martina C., and Howard Heidi C.. 2010 "Where Are You Going, Where Have You Been: A Recent History of the Direct-to-Consumer Genetic Testing Market." *Journal of Community Genetics* 1(3):101–6. [PubMed: 21475669]
- Borry Pascal and Howard Heidi. 2008 "DTC Genetic Services: A Look across the Pond." *The American Journal of Bioethics* 8(6):14–16.
- Byrd W. Carson and Hughey Matthew W.. 2015 "Biological Determinism and Racial Essentialism: The Ideological Double Helix of Racial Inequality." *The ANNALS of the American Academy of Political and Social Science* 661(1):8–22.
- Chow-White Peter et al. 2018 "'Warren Buffet Is My Cousin': Shaping Public Understanding of Big Data Biotechnology, Direct-to-Consumer Genomics, and 23andMe on Twitter." *Information, Communication & Society* 21(3):448–64.
- Condit Celeste M., Parrott Roxanne L., Harris Tina M., Lynch John, and Dubriwny Tasha. 2004 "The Role of 'Genetics' in Popular Understandings of Race in the United States." *Public Understanding of Science* 13(3):249–72. [PubMed: 15372732]
- Cornell Stephen E. and Hartmann Douglas. 2007 *Ethnicity and Race: Making Identities in a Changing World*. 2nd ed Thousand Oaks, CA: Pine Forge Press.
- Covolo Loredana, Rubinelli Sara, Ceretti Elisabetta, and Gelatti Umberto. 2015 "Internet-Based Direct-to-Consumer Genetic Testing: A Systematic Review." *Journal of Medical Internet Research* 17(12):e279. [PubMed: 26677835]
- Duster Troy. 2011 "Ancestry Testing and DNA: Uses, Limits, and Caveat Emptor" Pp. 99–115 in *Race and the Genetic Revolution: Science, Myth, and Culture*, edited by Krinsky S and Sloan K. New York: Columbia University Press.
- Duster Troy. 2015 "A Post-Genomic Surprise: The Molecular Reinscription of Race in Science, Law and Medicine." *The British Journal of Sociology* 66(1):1–27. [PubMed: 25789799]
- Einsiedel Edna and Geransar Rose. 2009 "Framing Genetic Risk: Trust and Credibility Markers in Online Direct-to-Consumer Advertising for Genetic Testing." *New Genetics and Society* 28(4): 339–62.
- Fan Weimiao and Yan Zheng. 2010 "Factors Affecting Response Rates of the Web Survey: A Systematic Review." *Computers in Human Behavior* 26(2):132–39.
- Fujimura Joan H. and Rajagopalan Ramya. 2010 "Different Differences: The Use of 'Genetic Ancestry' versus Race in Biomedical Human Genetic Research." *Social Studies of Science* 41(1): 5–30.
- Gamble Vanessa Northington. 1997 "Under the Shadow of Tuskegee: African Americans and Health Care." *American Journal of Public Health* 87(11):1773–78. [PubMed: 9366634]
- Gans Herbert J. 1979 "Symbolic Ethnicity: The Future of Ethnic Groups and Cultures in America." *Ethnic and Racial Studies* 2(1):1–20.
- Goddard Katrina A. B. et al. 2009 "Public Awareness and Use of Direct-to-Consumer Genetic Tests: Results From 3 State Population-Based Surveys, 2006." *American Journal of Public Health* 99(3): 442–45. [PubMed: 19106425]
- Golbeck Natasha and Roth Wendy D.. 2012 "Aboriginal Claims: DNA Ancestry Testing and Changing Concepts of Indigeneity" Pp. 415–32 in *Biomapping Indigenous Peoples: Towards and Understanding of the Issues*, edited by Berthier-Foglar S, Collingwood-Whittick S, and Tolazzi S. Amsterdam: Rodopi.
- Gordon Milton M. 1964 *Assimilation in American Life: The Role of Race, Religion, and National Origins*. New York: Oxford University Press.
- Gragert Loren et al. 2014 "HLA Match Likelihoods for Hematopoietic Stem-Cell Grafts in the U.S. Registry." *New England Journal of Medicine* 371(4):339–48. [PubMed: 25054717]
- Greely Henry T. 2008 "Genetic Genealogy: Genetics Meets the Marketplace" Pp. 215–34 in *Revisiting Race in a Genomic Age*, edited by Koenig BA, Lee SS-J, and Richardson SS. New Brunswick, NJ: Rutgers University Press.
- Harris Anna, Wyatt Sally, and Kelly Susan E.. 2013 "The Gift of Spit (and the Obligation to Return It): How Consumers of Online Genetic Testing Services Participate in Research." *Information, Communication & Society* 16(2):236–57.

- Helft Michael. 2017 “[Ancestry.com](https://www.ancestry.com) DNA Database Tops 3M, Sales Rise To \$850M Ahead Of Likely 2017 IPO.” *Forbes.com*, 1 10.
- Hirschman Elizabeth C. and Panther-Yates Donald. 2008 “Peering Inward for Ethnic Identity: Consumer Interpretation of DNA Test Results.” *Identity* 8(1):47–66.
- Hochschild Jennifer L. and Sen Maya. 2015 “To Test or Not? Singular or Multiple Heritage? Genomic Ancestry Testing and Americans’ Racial Identity.” *Du Bois Review* 12(2):321–47.
- Hochschild Jennifer and Sen Maya. 2015 “Genetic Determinism, Technology Optimism, and Race: Views of the American Public.” *The ANNALS of the American Academy of Political and Social Science* 661(1):160–80.
- Hollenbach Jill A. et al. 2015 “Race, Ethnicity and Ancestry in Unrelated Transplant Matching for the National Marrow Donor Program: A Comparison of Multiple Forms of Self-Identification with Genetics.” *PLoS ONE* 10(8).
- Hout Michael and Goldstein Joshua R.. 1994 “How 4.5 Million Irish Immigrants Became 40 Million Irish Americans: Demographic and Subjective Aspects of the Ethnic Composition of White Americans.” *American Sociological Review* 59:64.
- Jiménez Tomás R. 2010 *Replenished Ethnicity: Mexican Americans, Immigration, and Identity*. Berkeley, CA: University of California Press.
- Jobling Mark A., Rasteiro Rita, and Wetton Jon H.. 2016 “In the Blood: The Myth and Reality of Genetic Markers of Identity.” *Ethnic and Racial Studies* 39(2):142–61.
- Keshavan Meghana. 2016 “20 Key Players in the Direct-to-Consumer Lab Testing Market.” *MedCityNews*, 1 20.
- Lee Sandra Soo-Jin. 2013 “American DNA.” *Current Anthropology* 54(S7):S77–86.
- Lewontin Richard. 1972 “The Apportionment of Human Diversity.” *Evolutionary Biology* 6:381–98.
- Lichter Daniel T., Qian Zhenchao, and Tumin Dmitry. 2015 “Whom Do Immigrants Marry? Emerging Patterns of Intermarriage and Integration in the United States.” *The ANNALS of the American Academy of Political and Social Science* 662(1):57–78.
- Liebersohn Stanley. 1985 “Unhyphenated Whites in the United States.” *Ethnic and Racial Studies* 8(1): 159–80.
- Lofland John and Lofland Lyn H.. 1995 *Analyzing Social Settings: A Guide to Qualitative Observation and Analysis*. 3rd ed Belmont, CA: Wadsworth.
- Mello Michelle M. and Wolf Leslie E.. 2010 “The Havasupai Indian Tribe Case — Lessons for Research Involving Stored Biologic Samples.” *New England Journal of Medicine* 363(3):204–7. [PubMed: 20538622]
- Nelson Alondra. 2008a “Bio Science: Genetic Genealogy Testing and the Pursuit of African Ancestry.” *Social Studies of Science* 38(5):759–83. [PubMed: 19227820]
- Nelson Alondra. 2008b “The Factness of Diaspora: The Social Sources of Genetic Genealogy” Pp. 253–68 in *Revisiting Race in a Genomic Age*, edited by Koenig BA, Lee SS-J, and Richardson SS. New Brunswick, NJ: Rutgers University Press.
- Nordgren A and Juengst ET. 2009 “Can Genomics Tell Me Who I Am? Essentialistic Rhetoric in Direct-to-Consumer DNA Testing.” *NEW GENETICS AND SOCIETY* 28(2):157–72.
- Office of Management and Budget. 1997 “Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity.” *Federal Register* 62(210):58782–90.
- Ostergren Jenny E. et al. 2015 “How Well Do Customers of Direct-to-Consumer Personal Genomic Testing Services Comprehend Genetic Test Results? Findings from the Impact of Personal Genomics Study.” *Public Health Genomics* 18(4):216–24. [PubMed: 26087778]
- Perez Anthony Daniel and Hirschman Charles. 2009 “The Changing Racial and Ethnic Composition of the US Population: Emerging American Identities.” *Population and Development Review* 35(1):1–51. [PubMed: 20539823]
- Phelan Jo C., Link Bruce G., Zelner Sarah, and Yang Lawrence H.. 2014 “Direct-to-Consumer Racial Admixture Tests and Beliefs About Essential Racial Differences.” *Social Psychology Quarterly* 77(3):296–318. [PubMed: 25870464]

- Phillips Andelka M. 2016 “Only a Click Away - DTC Genetics for Ancestry, Health, Love...and More: A View of the Business and Regulatory Landscape.” *Applied & Translational Genomics* 8:16–22. [PubMed: 27047755]
- Portes Alejandro and Zhou Min. 1993 “The New Second Generation: Segmented Assimilation and Its Variants.” *The Annals of the American Academy of Political and Social Science* 530(1):74–96.
- Qian Zhenchao, Glick Jennifer E., and Batson Christie D.. 2012 “Crossing Boundaries: Nativity, Ethnicity, and Mate Selection.” *Demography* 49(2):651–75. [PubMed: 22350840]
- Qian Zhenchao and Lichter Daniel T.. 2007 “Social Boundaries and Marital Assimilation: Interpreting Trends in Racial and Ethnic Intermarriage.” *American Sociological Review* 72(1):68–94.
- Qian Zhenchao and Lichter Daniel T.. 2011 “Changing Patterns of Interracial Marriage in a Multiracial Society.” *Journal of Marriage and Family* 73(5):1065–84.
- Roberts J. Scott et al. 2017 “Direct-to-Consumer Genetic Testing: User Motivations, Decision Making, and Perceived Utility of Results.” *Public Health Genomics* 20(1):36–45. [PubMed: 28068660]
- Roth Wendy D. and Lyon Katherine A.. 2018 “Genetic Ancestry Tests and Race: Who Takes Them, Why, and How Do They Affect Racial Identities?” Pp. 133–69 in *Reconsidering Race: Cross-Disciplinary and Interdisciplinary Approaches*, edited by Suzuki K and von Vocano D. New York: Oxford University Press.
- Sarkar Sahotra. 1998 *Genetics and Reductionism*. Cambridge: Cambridge University Press.
- Shim Janet K., Rab Alam Sonia, and Aouizerat Bradley E.. 2018 “Knowing Something versus Feeling Different: The Effects and Non-Effects of Genetic Ancestry on Racial Identity.” *New Genetics and Society*. Advance online release.
- Smith Carol A. 1997 “The Symbolics of Blood: Mestizaje in the Americas.” *Identities* 3(4):495–521.
- Tutton Richard. 2007 “Constructing Participation in Genetic Databases.” *Science, Technology, & Human Values* 32(2):172–95.
- Wagner Jennifer K., Cooper Jill D., Sterling Rene, and Royal Charmaine D.. 2012 “Tilting at Windmills No Longer: A Data-Driven Discussion of DTC DNA Ancestry Tests.” *Genetics in Medicine* 14(6):586–93. [PubMed: 22382803]
- Waters Mary C. 1990 *Ethnic Options: Choosing Identities in America*. Berkeley: University of California Press.
- Waters Mary C. 2014 “Ethnic Identities in the Future: The Possible Effects of Mass Immigration and Genetic Testing.” *Ethnic and Racial Studies* 37(5):766–69.
- Williams Mark. 2017 “The Lucrative Rise of DNA Testing: ‘We Created the Market for What We Do.’” *The Guardian*, 5 25.
- Woodall Angela. 2016 “AncestryDNA Sees Huge Growth in Demand for Genetic Testing.” *MedCityNews*, 8 9.

<i>Categorical code (Reason for declining GAT)</i>	<i>Sample quote</i>
General disinterest	"I'm just not very interested in it."
Already knowing ancestry	"My family and I are all from Korea, which is comprised of only Koreans traditionally, so I know I'm 100% Korean."
Wanting more information	"I would need to do more research to determine what was involved. I'm not familiar with these tests."
Lack of relevance	"It doesn't matter to me what my past is. How I live my life today is what is important."
Time commitment issue	"Too busy."
Uncertainty about test value	"I don't know really understand the true point of the test. Once I find out where I come from, then what?"
Concern about data privacy or misuse	"I am not interested in having any genetic testing done as I believe in the not too distant future, this could impact job prospects, health insurance, etc."
Skepticism of test accuracy	"I don't think there is sufficient evidence in genome research for genetic ancestry testing to be accurate and applicable to current medical knowledge."
Would consider only if useful to others	"If it would help someone who needed a transplant I might be interested in helping if it wasn't too much of a pain in the butt or too invasive, but for my own sake, I don't really care."
Apprehension about discovering diseases	"I've lived with and taken care of people in my family who have health issues. I'm not sure I would want to know what lies ahead."

Figure A.
Qualitative response sample quotes

Table 1.

Interest in Genetic Ancestry Testing by demographic characteristics

	<i>Interest in Genetic Ancestry Test (%)</i>			<i>Sample composition</i>	
	Not interested	Interested in free test	Already taken	N	% of total sample
<i>Immigrant Generation</i>					
First	3.1	92.5	4.4	6,127	6.7
Second	2.0	93.3	4.8	10,448	11.4
Third	1.9	93.0	5.2	16,063	17.5
Fourth and up	1.6	93.2	5.2	59,091	64.4
<i>Self-identified race</i>					
Asian only	5.1	91.6	3.3	3,204	3.5
Black only	1.3	93.1	5.6	2,564	2.8
Latino/a only	1.0	95.9	3.1	4,422	4.8
White only	1.8	93.1	5.1	69,862	76.2
American Indian only	1.7	93.0	5.4	243	0.3
Pacific Islander only	1.0	93.9	5.1	98	0.1
Other only	1.8	91.7	6.6	1,016	1.1
Two races incld. White	1.0	93.2	5.8	7,910	8.6
Two races not incld. White	1.9	92.6	5.6	1,077	1.2
Three or more	0.6	88.5	10.9	1,333	1.5
<i>Sex</i>					
Female	1.7	93.4	5.0	72,732	79.3
Male	2.3	92.0	5.7	18,997	20.7
<i>Age</i>					
18 to 24	1.5	95.0	3.5	11,978	13.1
25 to 34	1.3	93.9	4.8	31,343	34.2
35 to 44	1.7	92.9	5.4	24,742	27.0
45 to 54	2.6	91.6	5.8	16,129	17.6
55 to 64	2.9	90.4	6.7	7,537	8.2
<i>Education</i>					
Did not finish HS	1.0	95.2	3.9	310	0.3
High School Degree	1.6	95.0	3.4	15,695	17.1
Associate's Degree	1.2	94.8	4.0	13,732	15.0
Bachelor's Degree	1.6	93.1	5.2	34,491	37.6
Grad/Prof. Degree	2.4	91.1	6.5	27,501	30.0
<i>Region</i>					
Midwest	2.2	93.7	4.2	21,879	23.9
Northeast	2.4	92.3	5.4	15,815	17.2
South	1.5	93.7	4.9	31,392	34.2
West	1.5	92.3	6.2	22,643	24.7
<i>Total</i>	1.8	93.1	5.1		

	<i>Interest in Genetic Ancestry Test (%)</i>			<i>Sample composition</i>	
	Not interested	Interested in free test	Already taken	N	% of total sample
N	1,643	85,397	4,689	91,729	100.0

Table 2.

Baseline logistic regressions predicting GAT interest

	(1)	(2)
	Already taken	Not interested
<u>Immigrant gen. (ref. Fourth+)</u>		
First generation	0.81 ** (0.06)	1.53 *** (0.16)
Second generation	0.90 (0.05)	1.32 ** (0.12)
Third generation	0.89 ** (0.04)	1.01 (0.07)
<u>Self-id race (ref. White only)</u>		
Asian only	0.62 *** (0.07)	2.30 *** (0.26)
Black only	1.16 (0.10)	0.74 (0.13)
Latino/a only	0.72 *** (0.07)	0.55 *** (0.09)
American Indian only	1.06 (0.30)	1.00 (0.51)
Pacific Islander only	1.06 (0.49)	0.56 (0.56)
Other only	1.28 (0.17)	0.79 (0.19)
Three or more races	2.57 *** (0.24)	0.39 ** (0.14)
Two races, including White	1.25 *** (0.07)	0.61 *** (0.07)
Two races, not incld. White	1.22 (0.17)	1.04 (0.24)
Female	0.88 *** (0.03)	0.79 *** (0.04)
Age 18 to 24	0.53 *** (0.04)	0.59 *** (0.06)
25 to 34	0.64 *** (0.04)	0.43 *** (0.04)
34 to 44	0.76 *** (0.04)	0.57 *** (0.05)
45 to 54	0.87 * (0.05)	0.88 (0.08)
No HS Degree	0.55 * (0.16)	0.45 (0.26)
HS only	0.51 *** (0.03)	0.70 *** (0.05)
Associate's Degree	0.59 *** (0.03)	0.56 *** (0.05)
Bachelor's Degree	0.82 *** (0.03)	0.71 *** (0.04)
Midwest	0.77 *** (0.04)	1.00 (0.07)
South	0.91 * (0.04)	0.67 *** (0.05)
West	1.21 *** (0.06)	0.65 *** (0.05)
Constant	0.11 *** (0.01)	0.05 *** (0.01)

Note: N= 91,729. Coefficients presented as odds ratios. Standard errors in parentheses. The reference category for sex is male, age is 55 to 64, education is a graduate degree, and region is living in the Northeast. Controls for survey design effects not shown.

p<0.001,

**
p<0.01,

*
p<0.05

Table 3.

Cross tabulations of GAT Interest with Measures of Ancestral Certainty

	<i>Interest in Genetic Ancestry Test</i>			<i>Reasons for disinterest</i>	
	Not interested	Interested in free test	Already taken	“Already know ancestry”	All other reasons
% Reporting no unknown ancestry (self)	91%	85%	91%	98%	89% ***
% Knows “a lot” about both sides of family	20%	18%	33%	33%	17% ***
% Reporting 4 grandparents ancestries	53%	47%	60%	74%	50% ***
% Grandparents ancestries homogenous	31%	20%	22%	50%	27% ***
% Reporting no knowledge seeking	13%	6%	1%	8%	13% **
Knowledge seeking behaviors	1.4	1.8	3.8	1.5	1.3 ***
Avg. count (0-7)					

Notes:

p<0.001,**
p<0.01,*
p<0.05

All t-tests two-tailed. In analytical sample (N=91,729), “Already know” N=342; All other reasons N=977.

Table 4.

Logistic regressions predicting ancestral certainty by demographic characteristics

	(1)	(2)	(3)	(4)
	No unknown ancestry	Knows “a lot” of family history	Report 4 grandparents	4 GP ancestries homogenous
<u>Immigrant gen. (ref. Fourth+)</u>				
First generation	3.28 *** (0.20)	1.69 *** (0.06)	2.24 *** (0.07)	3.46 *** (0.12)
Second generation	2.23 *** (0.09)	1.23 *** (0.04)	1.78 *** (0.05)	2.14 *** (0.07)
Third generation	1.75 *** (0.05)	1.01 (0.02)	1.17 *** (0.02)	0.78 *** (0.02)
<u>Self-id race (ref. White only)</u>				
Asian only	3.89 *** (0.54)	1.24 *** (0.06)	2.31 *** (0.11)	4.56 *** (0.21)
Black only	0.29 *** (0.01)	0.46 *** (0.03)	0.51 *** (0.02)	0.57 *** (0.03)
Latino/a only	0.81 *** (0.04)	0.59 *** (0.03)	0.68 *** (0.02)	0.75 *** (0.03)
American Indian only	1.07 (0.20)	1.45 * (0.22)	1.16 (0.15)	0.83 (0.15)
Pacific Islander only	2.46 (1.15)	1.30 (0.31)	1.76 ** (0.38)	0.96 (0.23)
Other only	0.91 (0.10)	1.52 *** (0.11)	1.50 *** (0.10)	1.28 *** (0.09)
Three or more races	0.49 *** (0.03)	0.95 (0.07)	1.01 (0.06)	0.04 *** (0.01)
Two races, including White	0.69 *** (0.02)	0.86 *** (0.03)	1.02 (0.03)	0.15 *** (0.01)
Two races, not incl. White	0.56 *** (0.05)	0.73 *** (0.06)	0.86 * (0.06)	0.38 *** (0.04)
Female	0.78 *** (0.02)	1.00 (0.02)	0.83 *** (0.01)	0.72 *** (0.01)
Age 18 to 24	0.44 *** (0.02)	0.55 *** (0.02)	0.49 *** (0.02)	0.31 *** (0.01)
25 to 34	0.52 *** (0.02)	0.60 *** (0.02)	0.51 *** (0.01)	0.35 *** (0.01)
34 to 44	0.64 *** (0.03)	0.76 *** (0.02)	0.64 *** (0.02)	0.47 *** (0.01)
45 to 54	0.80 *** (0.04)	0.92 * (0.03)	0.84 *** (0.02)	0.71 *** (0.02)
No HS Degree	0.27 *** (0.04)	0.37 *** (0.07)	0.35 *** (0.04)	0.53 *** (0.09)
HS only	0.50 *** (0.01)	0.49 *** (0.01)	0.48 *** (0.01)	0.57 *** (0.02)
Associate’s Degree	0.64 *** (0.02)	0.61 *** (0.02)	0.56 *** (0.01)	0.61 *** (0.02)
Bachelor’s Degree	0.87 *** (0.02)	0.80 *** (0.01)	0.81 *** (0.02)	0.82 *** (0.02)
Midwest	0.83 *** (0.03)	0.92 ** (0.02)	0.76 *** (0.02)	0.83 *** (0.02)
South	0.52 *** (0.02)	0.79 *** (0.02)	0.61 *** (0.01)	0.73 *** (0.02)
West	0.75 *** (0.03)	0.90 *** (0.02)	0.74 *** (0.02)	0.71 *** (0.02)
Constant	23.79 *** (1.46)	0.42 *** (0.02)	2.70 *** (0.10)	1.06 (0.05)

Note: N= 91,729. Coefficients presented as odds ratios. Standard errors in parentheses. The reference category for sex is male, age is 55 to 64, education is a graduate degree, and region is living in the Northeast. Controls for survey design effects not shown.

p<0.001,

**
p<0.01,

*
p<0.05

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 5.

Predicting GAT disinterest by ancestral certainty and ancestral knowledge

	(1) + Unknown ancestry	(2) + Knows “alot”	(3) + GP ancestry reports	(4) All certainty measures	(5) All certainty + knowledge
<u>Ancestral certainty measures</u>					
No unknown ancestry (self)	1.23 * (0.11)			1.26 ** (0.11)	1.35 *** (0.12)
Knows “a lot” about family		0.92 (0.06)		0.91 (0.06)	1.17 * (0.08)
Reports 4 grandparents			0.82 ** (0.05)	0.81 ** (0.05)	0.95 (0.06)
4 GP ancestries homogenous			1.38 *** (0.10)	1.38 *** (0.10)	1.31 *** (0.10)
<u>Immigrant gen. (ref. Fourth+)</u>					
First generation	1.50 *** (0.15)	1.54 *** (0.16)	1.48 *** (0.15)	1.47 *** (0.15)	1.17 (0.12)
Second generation	1.29 ** (0.12)	1.32 ** (0.12)	1.30 ** (0.12)	1.28 ** (0.12)	1.14 (0.10)
Third generation	1.00 (0.07)	1.01 (0.07)	1.03 (0.07)	1.02 (0.07)	1.00 (0.07)
<u>Self-id race (ref. White only)</u>					
Asian only	2.28 *** (0.25)	2.30 *** (0.26)	2.13 *** (0.24)	2.13 *** (0.24)	1.63 *** (0.18)
Black only	0.77 (0.14)	0.74 (0.13)	0.74 (0.13)	0.76 (0.14)	0.71 (0.13)
Latino/a only	0.55 *** (0.09)	0.55 *** (0.09)	0.55 *** (0.09)	0.55 *** (0.09)	0.48 *** (0.08)
American Indian only	1.00 (0.50)	1.01 (0.51)	1.02 (0.52)	1.02 (0.52)	0.93 (0.47)
Pacific Islander only	0.55 (0.56)	0.56 (0.56)	0.57 (0.58)	0.57 (0.57)	0.54 (0.54)
Other only	0.79 (0.19)	0.79 (0.19)	0.79 (0.19)	0.80 (0.19)	0.72 (0.18)
Three or more races	0.39 ** (0.14)	0.39 ** (0.14)	0.41 * (0.15)	0.42 * (0.15)	0.48 * (0.17)
Two races, including White	0.62 *** (0.07)	0.61 *** (0.07)	0.65 *** (0.08)	0.66 *** (0.08)	0.68 ** (0.08)
Two races, not incld. White	1.05 (0.24)	1.03 (0.24)	1.07 (0.25)	1.08 (0.25)	1.03 (0.24)
<u>Knowledge seeking (ref. none)</u>					
Asked relatives					0.63 *** (0.05)
Consulted family documents					0.71 *** (0.04)
Consulted website					0.35 *** (0.03)
Sent away for official records					0.84 (0.11)
Went to a library or archive					0.85 (0.11)
Other					0.56 *** (0.09)
Constant	0.04 *** (0.01)	0.05 *** (0.01)	0.05 *** (0.01)	0.04 *** (0.01)	0.08 *** (0.01)

Note: N=91,729. Coefficients presented as odds ratios. Standard errors in parentheses. Demographic and survey design controls not shown.

p<0.001**
p<0.01*
p<0.05

Table 6.

GAT interest and ancestral certainty by self-identified race and immigrant generation

	Not interested in GAT				No unknown ancestry				Reports 4 grandparents			
	1 st Gen.	2 nd Gen.	3 rd Gen.	4 ^{th+} Gen.	1 st Gen.	2 nd Gen.	3 rd Gen.	4 ^{th+} Gen.	1 st Gen.	2 nd Gen.	3 rd Gen.	4 ^{th+} Gen.
Asian only	6%	4%	6%	6%	98%	98%	100%	99%	80%	76%	82%	84%
Black only	4%	2%	0%	1%	92%	84%	70%	58%	57%	48%	30%	27%
Latino/a only	1%	1%	1%	1%	89%	88%	87%	81%	44%	43%	40%	43%
White only	3%	2%	2%	2%	97%	92%	92%	84%	74%	61%	52%	44%
American Indian only	-	0%	0%	2%	-	91%	80%	84%	-	82%	38%	46%
Native Hawaiian only	0%	3%	0%	0%	94%	94%	93%	95%	59%	72%	57%	51%
Other only	2%	2%	2%	2%	97%	94%	90%	79%	77%	66%	57%	57%
Two races, including White	2%	2%	1%	1%	90%	88%	85%	78%	61%	55%	49%	42%
Two races, not including White	4%	1%	2%	1%	90%	85%	85%	71%	61%	53%	46%	35%
Three or more races	0%	1%	1%	1%	81%	80%	81%	71%	57%	53%	47%	40%
Total	3%	2%	2%	2%	95%	91%	90%	83%	67%	57%	51%	43%

Note: Cell sizes less than 10 not shown.