



DNA Tribes® Digest August 29, 2008
Copyright © 2008 DNA Tribes®. All rights reserved.

To request an email subscription to DNA Tribes® Digest, email digest@dnatribes.com with the subject "Subscribe." To request an unsubscribe from DNA Tribes® Digest, email digest@dnatribes.com with the subject "Unsubscribe."

Table of Contents:

Introducing DNA Tribes® Digest	1
Featured Topic: Who are the Americans?	3
Getting the Most from Your Testing	10

Introducing DNA Tribes® Digest

Hello, and welcome to the first issue of DNA Tribes® Digest, a free electronic magazine where we will discuss our work here at DNA Tribes, including regular features with a more in-depth look at topics of interest to our customers. My name is Lucas Martin, and if you have ever contacted us about an order or your results, we might have spoken or corresponded already.

If you have looked through our website at dnatribes.com, you probably know that our work involves not just analysis services for our customers, but also extensive original research to identify the genetic structure of world populations. Our research involves a rigorous mathematical analysis of a uniquely detailed and comprehensive database of human populations around the world. These populations are assembled by researchers according to everyday notions of nationality and ethnicity, which gives us the opportunity to observe how these descriptions of ancestry relate to the actual genetic structure of human populations. In most cases, the genetic structure we see corresponds closely to geography: islands, coastlines, peninsulas, mountain ranges, lakes, and other natural features of the land. This is because human communities form and sustain themselves around natural features: trade, travel, and natural areas of settlement form the basis of human communities, which over time develop localized cultural, linguistic, and genetic characteristics.

For example, one feature of European population structure we identified in our database and embodied in our **DNA Tribes® Europa** analysis (more information is available at <http://dnatribes.com/dnatribes-europa.html>) is that populations located around the Norse Sea and the English Channel share genetic characteristics that distinguish them from other European populations, including Germanic-speaking populations of Central Europe. We labeled this genetic grouping "Norse" to reflect shared historical contact with the famous (or infamous, depending on which historical text is read) Viking societies of the Middle Ages. However, the Norse genetic grouping is likely substantially older than the Viking invasions, which represent

only recent events known to history, against a more ancient background of Northern European settlement and migration dating several thousand years to the end of the recent Ice Age. As illustrated in **Figure 1** below, our analysis indicates this “Norse” region is most closely related to the nearby “Celtic” region that encompasses the British Isles, and these two maritime zones (Norse and Celtic) are together distinguished from all other Northern Europeans, including the Germanic zone of Central Europe (despite shared language characteristics), as illustrated by the neighbor-joining tree diagram below:

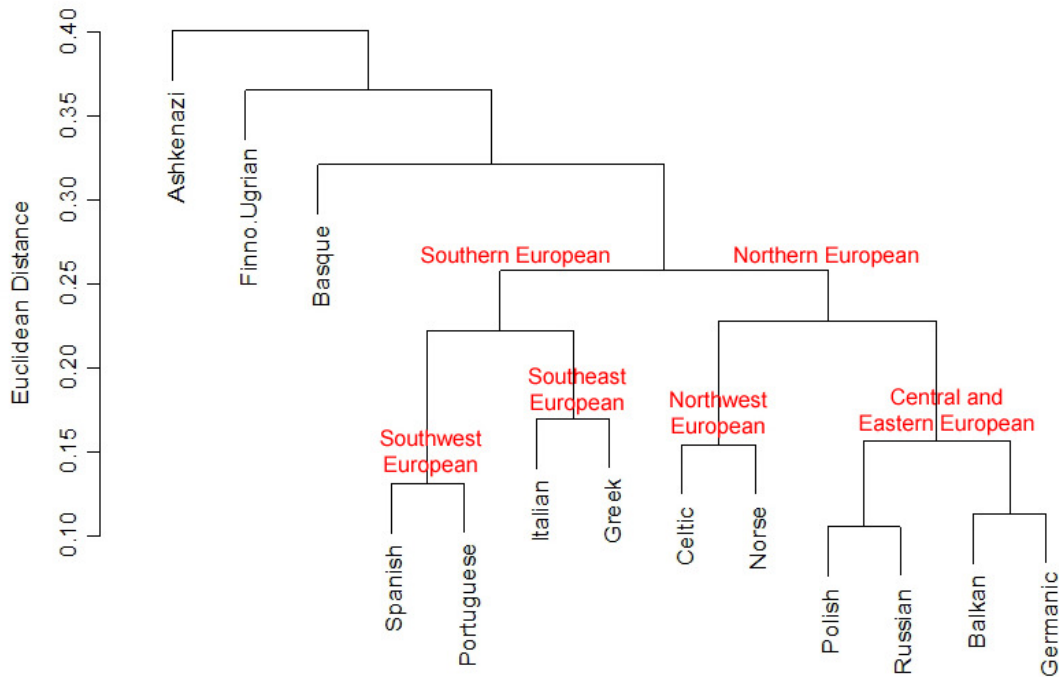


Fig. 1: Neighbor-joining “family tree” diagram of European genetic sub-regions. More closely related sub-regions are joined towards the bottom of the diagram, and more distant relationships are joined towards the top. For more information on this diagram and European genetic structure, see our Global Survey at <http://dnatribes.com/sample-results/dnatribes-global-survey-regional-affinities.pdf>.

Our analysis of Northern European sub-structure was recently confirmed by a large study of European genetic structure using several thousand SNP genetic markers published in the August 7 issue of *Current Biology* (online at <http://www.current-biology.com/> and summarized in a *New York Times* article at <http://www.nytimes.com/2008/08/13/science/13visual.html>). The study reported substantial overlap among the Atlantic populations of Northern Europe, while the two German populations (including one in Kiel at the far north of Germany) were observed to overlap most with nearby Austrian, Czech, and French populations in Central Europe.

Here as elsewhere, geography trumps language and everyday ideas of ethnicity as a predictor of genetic patterns. Why is this? Local genetic patterns develop over long periods of time. Many or most human populations have been settled continuously for several thousands of years (in some cases, tens or hundreds of thousands of years), so the everyday ethnic and national

terms we ordinarily use to describe ancestry just scratch the surface of these cumulative events. Newsworthy events that fill history books, such as invasions, migrations, and imperial expansions can have less impact on genetic patterns than the long term patterns of trade and a culture rooted in natural conditions that determine a population's interactions with its neighbors. For purposes of our analysis, we use familiar terms to label the genetic units identified, while keeping in mind that the actual genetic reality "on the ground" is often older and deeper than any labels we use.

Further, human population structure can change, especially when the underlying geographical relationships shift in importance. This can happen naturally (for instance, when glaciers receded and places like Northern Europe became inhabitable at the end of the last Ice Age), or for historical reasons. Just over 500 years ago, Europeans looking for trade routes stumbled upon North, Central, and South America. This chance encounter between populations previously separated by the Atlantic Ocean introduced catastrophic waves of illness to indigenous communities throughout the Americas, and also brought back American food plants (including the potato and cassava) that fueled population booms in Europe and Africa. These new circumstances facilitated new patterns of human interaction that have re-shaped the genetic structure of world populations living today.

Our first feature Digest article will look at the outcome of these events in two American ethnic groups. Later issues of DNA Tribes® Digest will continue to examine many more topics of interest to our customers based on original DNA Tribes® analysis of world genetic structure. Our goal is to shed new light on human genetic relationships in a way that is interesting and informative, and discuss how this can apply to your personal genetic ancestry research.

Talk to you soon,
Lucas @ DNATribes.com

Featured Topic: Who are the Americans?

At the 2008 Beijing Olympics, people around the world came together in the spirit of athletic competition. Among these Olympians were competitors hailing from a country that came into existence within the last 500 years in the wake of several large migrations: athletes from the United States of America. Americans come from many backgrounds, and two large ethnic groups that have emerged from these recent migrations are the Caucasians and the African-Americans. DNA Tribes® global population database includes several reference samples collected throughout the United States and classified by ethnicity according to everyday usage (for government and medical purposes in most cases). Below, we will examine the genetic relation of these American Caucasian and African-American populations to world genetic structure to measure the outcome of these migration events in communities living today.

Part One: Caucasians in the United States

The first American ethnic group we will examine is Caucasians. The term "Caucasian" comes originally from a broader anthropological term that has been used to describe human populations living in a large zone stretching from Europe to the Near East and India. This inclusive

anthropological use of the term originally referred to the Caucasus, a mountainous area at the geographical intersection of Europe and the Near East. DNA Tribes® analysis of human populations has identified a family of genetic regions that generally fits this original, inclusive definition of “Caucasian” (except that our analysis indicates that all populations of India form their own branch or “family” of Central and South Asian regions that is overall closest to East Asian regions), as illustrated in **Figure 2** below:

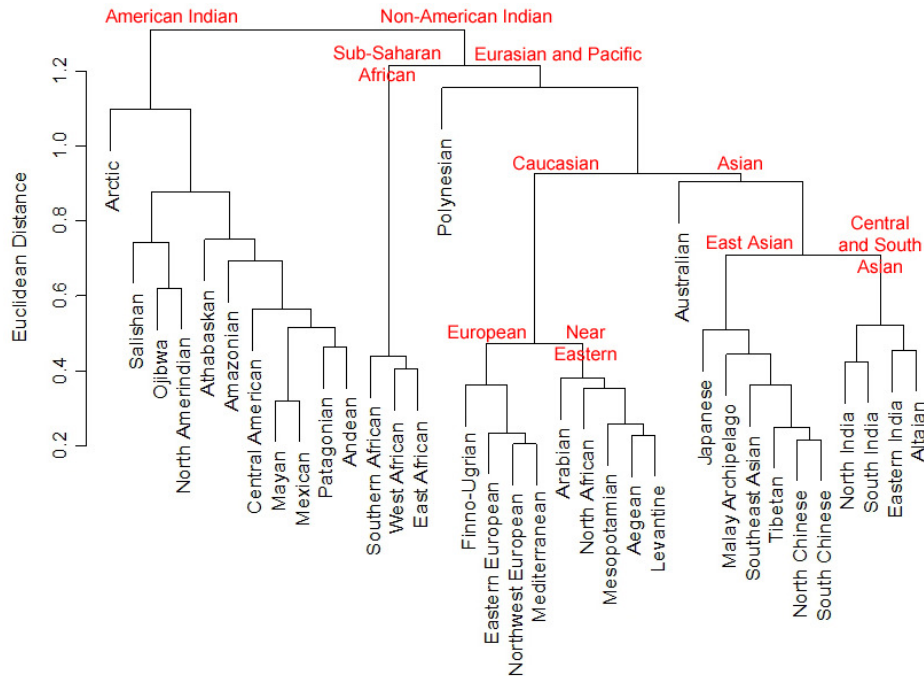


Fig. 2: Neighbor-joining “family tree” diagram of world genetic regions. More closely related regions are joined towards the bottom of the diagram, and more distant relationships are joined towards the top. For more information on this diagram and world genetic structure, see our Global Survey at <http://dnatribes.com/sample-results/dnatribes-global-survey-regional-affinities.pdf>.

In the United States, “Caucasian” generally refers to English-speaking communities first established by settlers from the British Isles during the Colonial era. These founding Colonial British populations have since absorbed other North American settlements (Dutch, German, and French in particular) as well as later waves of European migration (first from northern and western Europe, and later from southern and eastern Europe). For this reason, “Caucasian” in the United States has acquired a more specific ethnic meaning reflecting the particular European origins of Caucasian Americans.

To explore the genetic outcome of these interactions, we examined Caucasian American population samples from several U.S. states, and estimated the genetic contribution (or admixture) from four European regions identified by our analysis (Northwest European, Mediterranean, Finno-Ugrian, and Eastern European) and five non-European zones (Near Eastern, Sub-Saharan African, American Indian, South and Central Asian, and East Asian). (*For more information on these regions and continental zones, see our Global Survey linked above.*)

In each case, these reference populations include at most several hundred individuals classified as “Caucasian” for government or medical purposes, and were referenced “as-is” for this study to reflect everyday ethnic classifications. The results of this study are illustrated in map below:

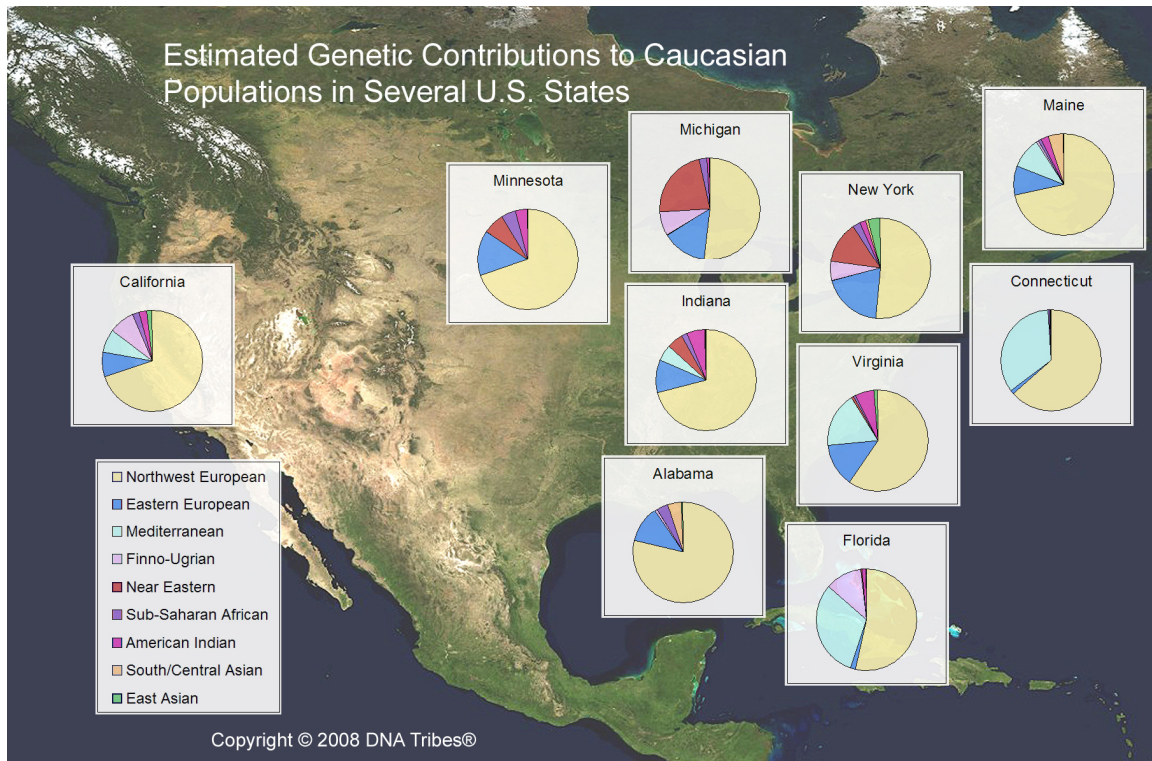


Fig. 3: Map illustrating estimated genetic contributions to Caucasian populations in several U. S. states.

	N.W. Eur.	E. Eur.	Medit.	Finn.	N. East.	S-S. Afr.	Amerind.	S/C. Asian	E. Asian
Alabama	78.7%	11.6%	0.0%	0.8%	0.0%	3.9%	0.0%	4.5%	0.5%
California	69.8%	7.9%	7.5%	8.6%	0.0%	2.0%	2.4%	0.0%	1.8%
Connecticut	63.4%	1.2%	34.5%	0.0%	0.0%	0.8%	0.2%	0.0%	0.0%
Florida	53.3%	1.8%	31.2%	11.9%	0.4%	0.0%	1.3%	0.0%	0.0%
Indiana	70.9%	10.6%	5.3%	0.0%	5.2%	1.9%	5.9%	0.3%	0.0%
Maine	71.5%	9.3%	10.0%	0.8%	0.0%	1.0%	2.5%	5.0%	0.0%
Michigan	51.7%	14.4%	0.0%	7.7%	22.8%	2.6%	0.8%	0.0%	0.0%
Minnesota	69.3%	15.2%	0.0%	0.0%	6.8%	4.7%	3.9%	0.0%	0.0%
New York	51.3%	19.6%	0.0%	6.0%	13.9%	2.6%	1.9%	0.8%	3.9%
Virginia	59.5%	13.7%	17.9%	0.0%	0.7%	0.8%	6.2%	0.0%	1.1%
Average	63.9%	10.5%	10.6%	3.6%	5.0%	2.0%	2.5%	1.1%	0.7%

Table 1: Table listing estimated genetic contributions to Caucasian populations in several U. S. states.

Upon examination of these results, several points emerge:

- In all U.S. state population samples, the largest genetic contribution identified was Northwest European, ranging from a minimum of 51.3% in Michigan to a maximum of 78.7% in Alabama, with an average of 63.9%. This reflects the influx of peoples from the British Isles, Scandinavia, and adjacent parts of Northern Europe (in particular, the Netherlands, Northern France, and to some extent Germany).
- Substantial contributions of 34.5% and 31.2% from the Mediterranean genetic region (which includes Spain, Portugal, and Northern Italy) were identified for Connecticut and Florida populations. This appears to reflect the large Italian-American community in Connecticut and Hispanic-American community in Florida, which includes a substantial number of Cuban-Americans of predominantly Spanish origins.
- A sizeable Mediterranean contribution of 17.9% was identified in Virginia, where a larger than expected American Indian contribution of 6.2% was also observed. This could reflect local variation in ethnic classification, in this case classifying people of Latin American origins as “Caucasian” by government agencies when compiling Virginia data.
- Substantial Eastern European contributions were identified for most states, with an average of 10.5%. This reflects not only immigrants from Slavic-speaking countries, but also to some extent Eastern European genetic patterns brought from Germany, which is genetically intermediate between Northwestern and Eastern Europe.
- Significant Near Eastern contributions of 22.8% and 13.9% were identified for Michigan and New York, respectively. This could reflect the large Arab-American community in Michigan and to some extent the Southern Italian-American community (affiliated with the Aegean genetic region) in New York State.
- In most state populations, smaller but substantial Sub-Saharan African (of an average of 2.0%) and American Indian contributions (of an average of 2.5%) were identified.
- Little East Asian genetic contribution was identified for most states, with the primary exceptions of New York (3.9%) and California (1.8%).

In summary, results indicate a predominantly Northwest European genetic contribution to Caucasian populations in all U.S. states, which is consistent with historical migrations largely from the British Isles, Scandinavia, and neighboring parts of Europe, with substantial contributions from other populations as well.

Part Two: African-Americans in the United States

The second United States ethnic group we will examine is African-Americans. African-American communities primarily trace their descent to peoples brought to the United States from West Africa by the Atlantic slave trade, who later came in contact with both European migrants and indigenous populations. African-Americans share a predominantly West African heritage with the larger African Diaspora also prominent in the Caribbean and Latin America, which we will explore further in the future.

To assess the genetic outcome of these interactions, we examined African-American population samples from several U.S. states, and estimated the genetic contribution from three African regions (West African, Southern African, and East African) and five non-African zones (European, American Indian, South and Central Asian, and East Asian). (*For more information on these regions and continental zones, see our Global Survey at <http://dnatribes.com/sample-results/dnatribes-global-survey-regional-affinities.pdf>.)* Additionally, we examined the Lumbee, a community in North Carolina named for the Lumbee River and historically classified as “free people of color” but now considered a tri-racial community.

As for Caucasian population samples discussed above, we have used these population samples of individuals classified as “African-American” for ordinary government and medical purposes “as-is,” to explore the relationship of everyday ethnic classifications to human genetic structure. The results of this study are illustrated in **Figure 4** below:

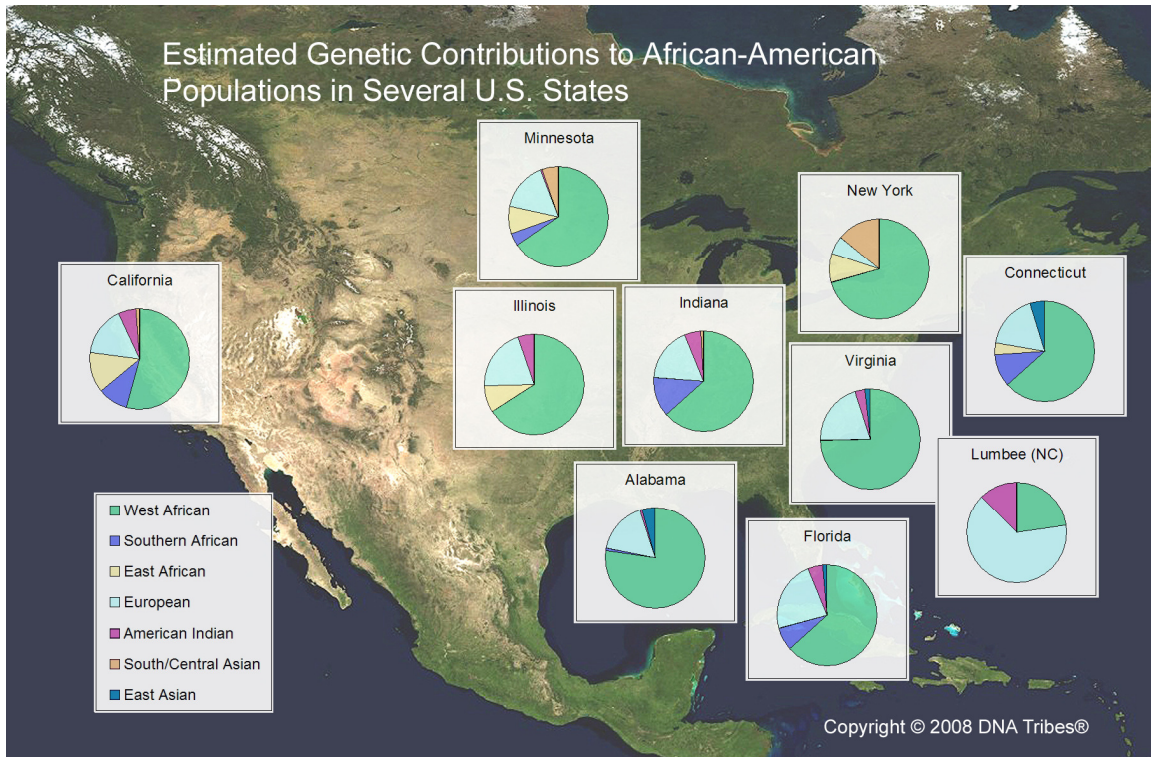


Fig. 4: Map illustrating estimated genetic contributions to African-American populations in several U. S. states.

	W. Afr.	S. Afr.	E. Afr.	Euro.	Amerind.	S./C. Asian	E. Asian
Alabama	77.2%	0.8%	0.0%	17.1%	0.7%	0.0%	4.1%
California	54.1%	9.8%	13.1%	16.1%	5.8%	1.0%	0.0%
Connecticut	63.3%	10.7%	3.7%	17.5%	0.0%	0.0%	4.8%
Florida	63.3%	7.6%	0.0%	23.0%	4.7%	0.0%	1.4%
Illinois	65.9%	0.0%	8.4%	20.5%	5.1%	0.0%	0.0%
Indiana	63.4%	12.7%	0.0%	17.9%	5.3%	0.8%	0.0%
Lumbee (NC)	22.8%	0.0%	0.0%	64.6%	12.6%	0.0%	0.0%
Minnesota	65.4%	3.9%	9.1%	15.7%	0.8%	5.2%	0.0%
New York	70.6%	0.0%	9.5%	5.9%	0.0%	14.0%	0.0%
Virginia	74.6%	0.0%	0.0%	20.4%	3.2%	0.0%	1.7%
Average*	66.4%	5.1%	4.9%	17.1%	2.9%	2.3%	1.3%

Table 2: Table listing estimated genetic contributions to African-American populations in several U. S. states. *Average does not include Lumbee, who are discussed separately.

Several observations can be made based on these results:

- The predominant genetic contribution identified for African-American populations in all states was West African (excluding Lumbee), with an average of 66.4%. This is consistent with predominantly West African origins described by historical sources.
- Substantial European genetic contributions were also observed for all African-American population samples, with an average of 17.1%.
- Some American Indian genetic contributions were observed for some African-American population samples, ranging from a low of 0% in New York to a high of 5.8% in California (excluding Lumbee).
- Substantial contributions from Southern Africa and East Africa were identified for some state population samples. These observations indicate the possibility that a substantial proportion of people with origins outside of coastal West Africa entered the United States via the Trans-Atlantic Slave Trade is not excluded. In particular, these observations could reflect genetic structure within Africa that is not yet fully represented in our database: in particular, in the geographically interior zone that is now the Democratic Republic of Congo, which historians identify as a substantial source of people for the Atlantic slave trade. We will continue to examine this possibility in future issues of DNA Tribes® Digest, including a broader comparison to African Diaspora communities in other parts of the Americas.
- A sizeable Central and South Asian contribution of 14.0% was identified for the New York State reference sample and, to a lesser extent, for the Minnesota reference sample. This could reflect classification practices when these data were compiled, or else could reflect local features of the African-American community in New York State related to specific origins in the African Diaspora. This second possibility will be explored in the

future, when we examine genetic patterns in communities of the African Diaspora (including Afro-Caribbean Diaspora) further.

- Along with African-American communities, the Lumbee of North Carolina were also included in this analysis. The Lumbee people have been identified as “free people of color” in the past, but are presently described as a tri-racial isolate community with European, American Indian, and African origins. Our analysis of a Lumbee population sample indicated substantial genetic contributions from Europeans (64.6%), West Africans (22.8%), and American Indians (12.6%), consistent with tri-racial community origins.

In summary, results for African-American population samples from several U.S. states are consistent with predominantly West African genetic origins, along with substantial European and American Indian contributions. Smaller contributions from other sources, including other African genetic regions, are also not excluded by these results.

How Population History Relates to Individual Ancestry

Analysis of entire populations provides information about the overall composition of an ethnic group. Populations who have recently migrated from a homeland elsewhere typically retain genetic characteristics of their place of origin. In the above examples, U.S. Caucasian and African-American populations retain genetic patterns from Northwest Europe and West Africa. In the United States, each of these diaspora ethnic groups has come in contact with other migrant or indigenous populations, in a process sometimes described as admixture.

In the case of U.S. Caucasians populations founded initially by Colonial British communities, many of the other ethnic groups encountered in North America have been from nearby Northwest Europe, continuing the process of interactions that have taken place in and near the British Isles in the past. New contacts less frequent in the past have also taken place in North America: for instance, between West Africans and Europeans, contributing to the development of new local genetic characteristics in African-American communities in the United States.

Individuals whose family origins are in these United States populations can inherit ancestry characteristic of the predominant founder populations (Northwest Europeans or West Africans) and their genetic relatives (Caucasians for Northwest Europeans; East and Southern Africa for African-Americans). In some cases, genetic material characteristic of other populations met in North America (such as Europeans for African-Americans or American Indians for both Caucasians and African-Americans) can also be inherited.

Conclusion

The genetic contributions identified in this study of Caucasian and African-American ethnic populations from several U.S. states are consistent with the histories of these ethnic groups. More specifically, U.S. Caucasian populations are predominantly Northwest European, and U.S. African-American populations are predominantly West African. In future issues, we will continue to explore topics relevant to both indigenous populations and diaspora communities



around the world. The latter will include the global European Diasporas and African Diasporas, which share important characteristics with their more local components in the United States.

Getting the Most from Your Testing

Once your core DNA Tribes® 15-Marker or 21-Marker testing is complete, we offer several options to keep your report current and to customize your analysis to deliver the information you want. (Prices are listed as of August 26, 2008 and are subject to change.)

Updating Your Analysis:

We incorporate new data in our algorithms on a periodic basis. This includes not only new reference data (our uniquely detailed and comprehensive database now includes 842 reference populations in all inhabited continents around the world), but also refinements to our algorithms and enhanced world region definitions.

New data also allow us to refine our genetic world region definitions. Based on a rigorous mathematical analysis of hundreds of ethnic and national reference samples, world regions express the actual genetic structure we have identified in world populations. These regional definitions both enhance your own analysis, and provide the necessary picture of world genetic structure to put your own results in a global context.

A map illustrating the populations and genetic regions currently identified in our analyses can be viewed at:

<http://dnatribes.com/populations.html>

Our current analyses now include a frequency-based match analysis as well as U.S. patent-pending DNA Tribes® TribeScore system that measures how well your own DNA fits in a population compared to members of that ethnic group or region. This unique scoring system provides the information you need to see how well your genetic fit is around the globe, providing detailed information about your genetic relationships to hundreds of populations.

Once lab testing is complete, your analysis can be updated at any times of your choice for \$24.99 through our secure online checkout at: <http://dnatribes.com/order.html>

Customizing Your Analysis:

DNA Tribes® offers several \$24.99 Add-On reports to customize your analysis for the information you want:



Extended Match Results: This is a comprehensive 22-page listing of your DNA match scores for all 842 reference populations in our database. We recommend this Add-On for customers who enjoy combing through lots of information.

Native American Panel: This two-page report lists your DNA match scores for all Native American tribal reference populations in our database. This can provide a closer look to complement Parts B and D of core results for customers who have substantial Native American origins.

African Panel: This two-page report lists your DNA match scores for all Sub-Saharan African reference populations in our database. This can provide a closer look to complement Parts B and D of core results for customers who have substantial African origins.

Central Asian Panel: This two-page report lists your DNA match scores for all Central Asian reference populations in our database, including Turkic populations as well as European Roma (Gypsy) populations. This can provide a closer look to complement Parts B and D of core results for customers who have substantial European, Near Eastern, South Asian, or East Asian origins.

Once lab testing is complete, Add-Ons can be selected at any time of your choice through our secure online checkout at: <http://dnatribes.com/order.html>

DNA Tribes® Europa: A Detailed Comparison to European Sub-Regions:

DNA Tribes® Europa provides the most detailed and complete analysis of European autosomal genetic structure available. DNA Tribes® Europa provides your DNA match scores for 13 genetic sub-regions of Europe, which is substantially more robust than the individual population matches in Parts B – C of core results and more detailed than the European world regions referenced in Part D of core results.

More information about DNA Tribes® Europa is available at:
<http://dnatribes.com/dnatribes-europa.html>

Expanding Your Analysis with a 21 Marker Upgrade:

For customers who have completed 13-marker or 15-marker testing with DNA Tribes®, we now offer a 21 Marker Upgrade. This includes lab testing of additional STR marker systems, which provides a closer comparison of your own DNA to world populations for greater accuracy and power of exclusion. The incorporation of additional marker systems can confirm or clarify your initial 13 or 15-marker results, and includes an update to all Add-Ons previously ordered for your kit.



21 Marker Upgrades are available for \$119.99 through our secure online checkout at:
<http://dnatribes.com/order.html>

Researching Your Results:

Each person's DNA Tribes® results are one of a kind and express their own unique collection of genetic material inherited from both paternal and maternal ancestors. However, human genetic relationships involve a complex hierarchy of relationships, from individual to family to ethnic group to genetic region to global population structure. Detailed information to allow you to interpret your own results in a broader context is available in our Global Survey at:

<http://dnatribes.com/sample-results/dnatribes-global-survey-regional-affinities.pdf>