CHAPTER

Admixture in North America

Esteban J. Parra*

Abstract

The history of North America has been marked by the encounter of populations from different continents. The discovery of the New World began a period defined by human migrations at a much larger scale than in previous history. This movement of people, voluntary or forced, changed profoundly the human landscape. As populations came into contact, admixture followed in varying degrees depending on the circumstances. Today, many people living in the US, Mexico, Canada and the Caribbean can trace their ancestry to more than one continent. The most important genetic contributions came from the indigenous Native American groups, Western Europeans and West Africans, although there have also been influences from other regions, such as East Asia and South Asia. We can reconstruct and interpret this history of migration and admixture using genetic markers. A very complete perspective can be obtained when analyzing autosomal markers (located on any of the chromosomes other than the sex-determining chromosomes, which are inherited from both parents), maternally transmitted mtDNA markers and Y-chromosome markers, which are transmitted from fathers to sons. In many cases, the maternal and paternal admixture histories are remarkably different, so including mtDNA and Y-chromosome markers can provide a much better picture than that offered by the autosomal markers. Thus, using genetic markers we can reconstruct history at the individual and population level, even in the absence of a historical record. In this chapter, I provide an overview of admixture in North America, with a particular emphasis on the two major admixed groups: African Americans (and African Caribbeans) and Hispanics. I also discuss the implications of the history of admixture for the distribution of the genetic variation involved in drug metabolism and drug response and the potential consequences of population stratification in candidate gene association studies in recently admixed populations.

Brief History of the Main North American Admixed Populations

African Americans and African Caribbeans

The history of African Americans can be traced back to 1619, when the first Africans arrived at the British colonies (Jamestown), although as early as 1526 the presence of African slaves was reported in Spanish expeditions to what would become the United States (South Carolina, Georgia, Florida and New Mexico). Institutional slavery began very soon after. Nevertheless, it was not until the beginning of the 18th century that the arrival of enslaved Africans reached increased rates, in parallel with the demand for workers to cultivate the tobacco, indigo, and rice plantations in the southern colonies. The highest numbers arrived in the decade from 1790-1800 and the early years of the 19th century. In 1808 slave trade became illegal but it continued at a low rate for several more decades. Estimates of the total number of enslaved

*Esteban J. Parra—Department of Anthropology, University of Toronto at Mississauga, 3359 Mississauga Rd. N., Mississauga, Ontario L5L 1C6, Canada. Email: eparra@utm.utoronto.ca

Pharmacogenomics in Admixed Populations, edited by Guilherme Suarez-Kurtz. ©2007 Landes Bioscience.
Africans brought into the United States differ but generally accepted numbers range between 380,000 and 570,000.

Although it is very difficult to precisely determine the ethnic origin of the enslaved Africans, information from shipping lists has provided an approximate picture of their geographic provenance. The slave trade affected a very wide area of western and western-central Africa, mainly the coastline between the present day countries of Senegal in the north and Angola in the south. The most important regions were Senegambia (Gambia and Senegal), Sierra Leone (Guinea and Sierra Leone), Windward Coast (Ivory Coast and Liberia), Gold Coast (Ghana), Bight of Benin (from the Volta river to the Benin river), Bight of Biafra (east of Benin river to Gabon), and Angola (Southwest Africa, including part of Gabon, Congo and Angola). Curtin has offered, based on data on the English trade of the 18th century (the peak of the Atlantic slave trade), estimates of the proportional contribution by areas. His analysis shows that Angola and Bight of Biafra were the regions contributing the highest numbers of slaves imported into the North American mainland (around 25% each). It is important to point out, however, that there were significant differences in ethnic origin depending on the port of entry in the US, and the figures for the Colonies of Virginia and South Carolina differed considerably.

The history of African Americans has been marked not only by their forced migration from Africa, but also by their admixture with the other ethnic groups that they met when they arrived in North America, namely Europeans and Native Americans. However, few historical records address the issue of admixture. Additionally, there have been important factors that, in the time since the abolition of slavery until the present, have configured the present African-American population. Of special interest is the pattern of migrations of African Americans within the US over the past 150 years. In this sense, the redistribution of African Americans in the southern states during the 19th century and the Great Migration from the rural south to the urban areas in the north beginning after World War I are of particular relevance, and have had an enormous impact in defining the present distribution of the African-American population in the US. According to the 2004 American Community Survey, there are approximately 34.8 million people who identify themselves as Black or African American in the US (12.2% of the total population). The majority of the African American population lives in the eastern states of the US and the largest percentages of African Americans with respect to the total population are found in the District of Columbia, Maryland, South Carolina, Georgia, Louisiana and Mississippi (Fig. 1).

Finally, during colonial times the situation in the islands of the Caribbean was quite different from the US mainland. In addition to the Spanish, who concentrated their colonization effort in Cuba, Hispaniola and Puerto Rico, other European powers (mainly England, France, and Holland) established colonies in the Caribbean during the 17th century. The indigenous populations had been decimated by warfare, disease and forced labor, and millions of Africans were forcefully brought to the Caribbean during the slave trade in order to work in the sugarcane plantations, which were the mainstay of the economy of the islands. It has been estimated that the British and French islands received around three million enslaved Africans between 1601 and 1870. In terms of absolute numbers, many more enslaved Africans were brought to the Caribbean than to the US during the slave trade. Additionally, the enslaved Africans in the islands of the Caribbean far outnumbered the white settlers. For example, in 1775 the number of enslaved Africans in Jamaica was approximately 200,000 and the number of white settlers around 12,000. The conditions of slavery were particularly harsh in this region and the death rates were extremely high, surpassing the birth rates. The abolition of the slave trade in 1807 and slavery in 1833 created a shortage of labor in the British West Indies, and tens of thousands of indentured servants (laborers under contract for a specified time in return for payment of travel expenses and maintenance) were brought to the region (mainly to Trinidad, British Guyana and Jamaica). Most of the indentured servants came from South Asia (India), China and the Portuguese islands of Madeira and the Azores, further defining the demographic picture of the islands of the Caribbean.
Admixture in North America

Hispanics

In the US, the terms “Hispanic” or “Latino” are generally employed to identify persons of Latin American origin or descent. Although this definition lumps together people with very different historical, cultural and linguistic backgrounds, this classification is widely used. Even though Central America, the Caribbean, and South America were for centuries under the domination of Spain and Portugal, they have had quite different regional histories, both before and after the Colonial period. Populations from four continents, North and South America, Europe, and Africa have contributed to the formation of contemporary Hispanic populations. My main objective here is to discuss the anthropological background of the three main Hispanic groups currently living in North America: Mexicans (and Mexican Americans), Cubans and Puerto Ricans.

In Mexico, intermixture of Spanish men with Indigenous American women began soon after the arrival of Hernán Cortés in 1521 and continued through the three centuries of Spanish domination in “New Spain”, configuring the Mexican population both biologically and culturally. Today, the majority of Mexicans (approximately 60% of the total population of more than 100 million people) are “mestizos”, a term used to describe persons of mixed European/Native American ancestry. It is important to mention that there was also a substantial African presence in Mexico during the Spanish rule. Curtin estimated the number of enslaved Africans brought in Mexico during the entire period of the slave trade to be around 200,000. In fact, in the early colonial period, enslaved Africans outnumbered the Spanish settlers in Mexico. The West African contribution was higher around the Gulf coast (Campeche, Yucatan, Tabasco and Veracruz) and areas of Southwest Mexico (Oaxaca, Guerrero), regions where the largest Afro-Mexican communities in Mexico are located today.

In the Caribbean Colonies (Cuba and Puerto Rico), the Native American population was far smaller than in Mexico, and was decimated by slavery and disease soon after the first contact with the Europeans. Nevertheless, the rate of admixture during the initial phases of the colonization was high enough to result in an appreciable genetic contribution from the Arawaks (Taíno) and Caribs, the original inhabitants of the Hispanic Caribbean. Another distinctive

Figure 1. Percentage of the population in the US who are black or African American. Map showing data by state. Source: US Census Bureau, available at http://www.census.gov/.
feature of this region is a significant African influence, which is also reflected in many aspects of the present societies of countries like Cuba, Puerto Rico, and the Dominican Republic. Enslaved Africans were brought to work in the sugar plantations in large numbers, even outnumbering the population of European origin. Accordingly, the African contribution in contemporary Cubans and Puerto Ricans has been higher than in other Hispanic populations.

Hispanics have become the largest minority in the US. According to the American Community Survey, in 2004 there were around 40.5 million Hispanics in the US (approximately 14% of the total US population). Hispanics constitute a large proportion of the population in the southwest of the US (California, Arizona, New Mexico, and Texas) and also in states such as New York, New Jersey, Illinois, and Florida (Fig. 2). Mexicans are the largest Hispanic group, with more than 25 million people (64% of the total Hispanic population). The Mexican presence in what is now the US has a long history. A substantial portion of southwestern US was Mexican territory, until it was annexed by the US after the US-Mexican war (1846-1848). However, most of the Mexican migration to the US has occurred since the 20th century. By the 1900s, the total Mexican-American population was estimated to be between 380,000 and 560,000. The first major wave of migration to the US occurred after the Mexican revolution, between 1910 and 1930. Mexican immigration, both legal and illegal has continued to increase since then, particularly after the 1950s. In the US, the states with the largest communities of Mexican ancestry are California, Texas, Arizona, and Illinois. The second largest Hispanic group in the US is the Puerto Rican community, with approximately 3.9 million people (around 10% of the total Hispanic population). Migration from Puerto Rico (a self-governing commonwealth in association with the US) to the mainland began in the early 1940s and concentrated mainly in the northeastern states. Currently, the largest Puerto Rican communities in the US are found in New York, Florida, New Jersey, Pennsylvania, and Massachusetts. Finally, there is also a large Cuban community in the US, comprising 1.4 million people (3.6% of the total Hispanic population). There were several waves of migration from Cuba to the US starting in 1959, and the majority of the Cubans have settled in the state of Florida.
Other Admixed Groups in North America

This chapter focuses on African Americans and African Caribbeans and also on the main North American Hispanic groups (Mexican and Mexican Americans, Cubans and Puerto Ricans). These are the admixed groups for which most of the genetic data have been collected. However, it is important to mention that admixture has been widespread in North America and has not been restricted to these groups. In Canada, African Canadians comprise approximately 2% of the population (662,000 people). Many of them are immigrants from the Caribbean, but some trace their roots to Black Loyalists who supported the British during the American Revolution and to fleeing slaves who arrived in Canada through the Underground Railroad. The Métis, one of the three federally recognized Canadian aboriginal groups, is an admixed population resulting from the admixture of indigenous Canadians and Europeans. This admixture can be traced back to the initial colonization of Canada by the French and the British. Currently, there are almost 300,000 Métis living in Canada. Finally, European Americans and Asian Americans also show evidence of genetic contributions from other continental groups, but to a lesser extent than African American or Hispanic populations.

Genetic Markers Used to Estimate Admixture

Autosomal Genetic Markers

Obtaining precise estimates of admixture proportions at the individual and global levels requires the use of autosomal genetic markers. The pioneering admixture work of Glass and Li, Workman and Reed employed “classical” genetic markers: blood groups, red cell enzymes and plasma protein polymorphisms. The advent of the PCR technique has made it possible to work directly at the DNA level and currently a wide range of genetic markers are used in admixture studies, including Single Nucleotide Polymorphisms (SNPs), Insertion/Deletion Polymorphisms (Indels), Alu insertions and Short Tandem Repeats (STRs or microsatellites). Irrespective of the type of genetic markers used to estimate admixture, it is advisable to use markers informative for ancestry rather than randomly selected markers because the former maximize the information about admixture while minimizing the genotyping effort. These markers are known as Ancestry Informative Markers (AIMs). Several criteria have been proposed to measure ancestry information content, such as the absolute allele frequency difference between parental populations (delta -δ), the Fisher information content (f) and the Informativeness for assignment (Ia). Although a widely used criterion, delta has some limitations because it doesn’t take into account all available information about allele frequencies and its application to more than two parental populations or multiallelic loci is not straightforward. The f and Ia measures give different absolute values but they rank markers similarly with respect to ancestry information content. It is important to emphasize that precise estimation of individual ancestry proportions requires many more markers than the estimation of the average ancestry proportions in a sample. While two or three dozen AIMs will provide estimates of ancestry proportions in a sample with a reasonable standard error (<5%), would be required to estimate individual admixture proportions with similar precision. In the past, the lack of useful AIMs was an important problem in admixture studies but fortunately this is no longer a limitation and thousands of AIMs have been selected for studying admixture between continental populations. Table 1 shows a list of 100 AIMs that are optimal for studying admixture proportions and dynamics in populations throughout the Americas. The list includes markers showing high frequency differences between Europeans and Native Americans, Europeans and West Africans, and Native Americans and West Africans. A substantial effort is currently underway to develop informative markers to study genetic structure and admixture at the continental level. Due to the more limited genetic differentiation within continental populations, the identification of this type of AIMs is far more challenging than the identification of AIMs to study admixture between the major continental population groups.
Table 1. Panel of Ancestry informative markers to estimate admixture proportions in North American admixed populations

<table>
<thead>
<tr>
<th>rs#</th>
<th>Chr</th>
<th>WAF/ EU</th>
<th>Delta</th>
<th>WAF/ NA</th>
<th>Delta</th>
</tr>
</thead>
<tbody>
<tr>
<td>140864</td>
<td>1</td>
<td>0.117</td>
<td>0.579</td>
<td>0.462</td>
<td>2695</td>
</tr>
<tr>
<td>2225251</td>
<td>1</td>
<td>0.751</td>
<td>0.567</td>
<td>0.184</td>
<td>2695</td>
</tr>
<tr>
<td>725667</td>
<td>1</td>
<td>0.741</td>
<td>0.075</td>
<td>0.666</td>
<td>2695</td>
</tr>
<tr>
<td>963170</td>
<td>1</td>
<td>0.125</td>
<td>0.925</td>
<td>0.800</td>
<td>951308</td>
</tr>
<tr>
<td>2814778</td>
<td>2</td>
<td>0.998</td>
<td>0.000</td>
<td>0.998</td>
<td>1594335</td>
</tr>
<tr>
<td>723822</td>
<td>2</td>
<td>0.032</td>
<td>0.841</td>
<td>0.810</td>
<td>1042602</td>
</tr>
<tr>
<td>6003</td>
<td>3</td>
<td>0.600</td>
<td>0.087</td>
<td>0.687</td>
<td>2366882</td>
</tr>
<tr>
<td>2055160</td>
<td>3</td>
<td>0.434</td>
<td>0.770</td>
<td>0.336</td>
<td>714857</td>
</tr>
<tr>
<td>1506069</td>
<td>3</td>
<td>0.092</td>
<td>0.800</td>
<td>0.708</td>
<td>1487214</td>
</tr>
<tr>
<td>7252</td>
<td>3</td>
<td>0.461</td>
<td>0.209</td>
<td>0.252</td>
<td>720496</td>
</tr>
<tr>
<td>1861498</td>
<td>4</td>
<td>0.808</td>
<td>0.139</td>
<td>0.672</td>
<td>720496</td>
</tr>
<tr>
<td>1063</td>
<td>5</td>
<td>0.032</td>
<td>0.815</td>
<td>0.783</td>
<td>720496</td>
</tr>
<tr>
<td>1526028</td>
<td>5</td>
<td>0.816</td>
<td>0.025</td>
<td>0.793</td>
<td>1042602</td>
</tr>
<tr>
<td>1465648</td>
<td>6</td>
<td>0.821</td>
<td>0.025</td>
<td>0.796</td>
<td>912537</td>
</tr>
<tr>
<td>3172121</td>
<td>6</td>
<td>0.525</td>
<td>0.347</td>
<td>0.765</td>
<td>912537</td>
</tr>
<tr>
<td>938431</td>
<td>6</td>
<td>0.200</td>
<td>0.547</td>
<td>0.300</td>
<td>912537</td>
</tr>
<tr>
<td>1316579</td>
<td>6</td>
<td>0.513</td>
<td>0.212</td>
<td>0.301</td>
<td>912537</td>
</tr>
<tr>
<td>719776</td>
<td>6</td>
<td>0.824</td>
<td>0.025</td>
<td>0.796</td>
<td>912537</td>
</tr>
<tr>
<td>951784</td>
<td>6</td>
<td>0.046</td>
<td>0.763</td>
<td>0.717</td>
<td>912537</td>
</tr>
<tr>
<td>17203</td>
<td>6</td>
<td>0.614</td>
<td>0.020</td>
<td>0.633</td>
<td>912537</td>
</tr>
<tr>
<td>768324</td>
<td>6</td>
<td>0.109</td>
<td>0.753</td>
<td>0.024</td>
<td>912537</td>
</tr>
<tr>
<td>1430909</td>
<td>6</td>
<td>0.065</td>
<td>0.802</td>
<td>0.517</td>
<td>912537</td>
</tr>
<tr>
<td>3287</td>
<td>6</td>
<td>0.571</td>
<td>0.570</td>
<td>0.000</td>
<td>912537</td>
</tr>
<tr>
<td>1350462</td>
<td>6</td>
<td>0.599</td>
<td>0.231</td>
<td>0.368</td>
<td>912537</td>
</tr>
<tr>
<td>1434870</td>
<td>6</td>
<td>0.079</td>
<td>0.846</td>
<td>0.025</td>
<td>912537</td>
</tr>
<tr>
<td>17023</td>
<td>6</td>
<td>0.032</td>
<td>0.841</td>
<td>0.810</td>
<td>912537</td>
</tr>
<tr>
<td>3309</td>
<td>7</td>
<td>0.101</td>
<td>0.306</td>
<td>0.407</td>
<td>912537</td>
</tr>
<tr>
<td>3317</td>
<td>7</td>
<td>0.587</td>
<td>0.699</td>
<td>0.112</td>
<td>912537</td>
</tr>
<tr>
<td>1439148</td>
<td>7</td>
<td>0.868</td>
<td>0.747</td>
<td>0.121</td>
<td>912537</td>
</tr>
<tr>
<td>1461227</td>
<td>7</td>
<td>0.284</td>
<td>0.430</td>
<td>0.714</td>
<td>912537</td>
</tr>
<tr>
<td>3340</td>
<td>7</td>
<td>0.142</td>
<td>0.642</td>
<td>0.499</td>
<td>912537</td>
</tr>
<tr>
<td>2077681</td>
<td>7</td>
<td>0.820</td>
<td>0.928</td>
<td>0.028</td>
<td>912537</td>
</tr>
<tr>
<td>951554</td>
<td>7</td>
<td>0.139</td>
<td>0.511</td>
<td>0.650</td>
<td>912537</td>
</tr>
<tr>
<td>1935946</td>
<td>7</td>
<td>0.620</td>
<td>0.725</td>
<td>0.105</td>
<td>912537</td>
</tr>
<tr>
<td>1881826</td>
<td>7</td>
<td>0.175</td>
<td>0.575</td>
<td>0.750</td>
<td>912537</td>
</tr>
<tr>
<td>2763</td>
<td>7</td>
<td>0.055</td>
<td>0.501</td>
<td>0.446</td>
<td>912537</td>
</tr>
<tr>
<td>2161</td>
<td>8</td>
<td>0.135</td>
<td>0.215</td>
<td>0.350</td>
<td>912537</td>
</tr>
<tr>
<td>2396676</td>
<td>8</td>
<td>0.697</td>
<td>0.672</td>
<td>0.025</td>
<td>912537</td>
</tr>
<tr>
<td>2341823</td>
<td>8</td>
<td>0.796</td>
<td>0.475</td>
<td>0.321</td>
<td>912537</td>
</tr>
<tr>
<td>1320892</td>
<td>8</td>
<td>0.053</td>
<td>0.842</td>
<td>0.789</td>
<td>912537</td>
</tr>
<tr>
<td>285</td>
<td>8</td>
<td>0.478</td>
<td>0.529</td>
<td>0.051</td>
<td>912537</td>
</tr>
<tr>
<td>983271</td>
<td>8</td>
<td>0.588</td>
<td>0.245</td>
<td>0.833</td>
<td>912537</td>
</tr>
<tr>
<td>3176921</td>
<td>8</td>
<td>0.579</td>
<td>0.665</td>
<td>0.085</td>
<td>912537</td>
</tr>
<tr>
<td>1373302</td>
<td>8</td>
<td>0.090</td>
<td>0.722</td>
<td>0.813</td>
<td>912537</td>
</tr>
<tr>
<td>1808089</td>
<td>8</td>
<td>0.186</td>
<td>0.605</td>
<td>0.792</td>
<td>912537</td>
</tr>
<tr>
<td>720966</td>
<td>8</td>
<td>0.842</td>
<td>0.950</td>
<td>0.108</td>
<td>912537</td>
</tr>
<tr>
<td>1987956</td>
<td>8</td>
<td>0.000</td>
<td>0.921</td>
<td>0.921</td>
<td>912537</td>
</tr>
<tr>
<td>1928415</td>
<td>8</td>
<td>0.787</td>
<td>0.912</td>
<td>0.125</td>
<td>912537</td>
</tr>
</tbody>
</table>

Delta is the absolute frequency difference between the parental populations. WAF: West Africa (Mende from Sierra Leone), EU: Europe (Spanish from Valencia), NA: Native American (Nahua from Mexico).
Admixture in North America

**mtDNA Polymorphisms**

The unique characteristics of mtDNA make this marker a useful tool in anthropological and evolutionary studies. In particular, mtDNA doesn't recombine and shows maternal inheritance, making it possible to trace the pattern of female migrations through time. It is important to mention that a recent study has identified a case of mtDNA paternal transmission in humans. However, to my knowledge, this is the only such case described in humans, and maternal inheritance is supported by thousands of maternal-offspring comparisons and is still regarded as the rule. Another useful feature of mtDNA is its high mutation rate, making this marker ideal to explore recent evolutionary events. Finally, mtDNA has an effective population size of one quarter that of autosomal DNA markers, and it is more susceptible to the effects of genetic drift. As a result, mtDNA polymorphisms show more geographic structure than autosomal markers and this is a useful attribute for admixture studies. Typically, analysis of mtDNA is carried out using PCR-RFLP to characterize polymorphisms defining common Continent-specific mtDNA haplogroups (groups of related sequences defined by shared diagnostic mutations) and/or sequence analysis of the two hypervariable sequences located on the mtDNA control region (HVR1 and HVR2). In terms of admixture studies in the Americas, it is important to mention that macrohaplogroup L (comprising haplogroups L1, L2 and L3) is restricted to Africa, haplogroups H, I, J, K, T, U, V and W to Europe and haplogroups A, B, C and D to Asia and the New World.

**Y-Specific Polymorphisms**

The Y chromosome has been a relative newcomer in evolutionary studies. For decades, efforts to find polymorphic markers in this chromosome were unsuccessful, but the situation has changed dramatically in the past few years. Tens of thousands of SNPs and hundreds of STRs have been identified in the nonpseudoautosomal region of the Y chromosome. Of these markers, at least 200 SNPs and 30 STRs have been reasonably well characterized in human populations. Similarly to the mtDNA, the Y-specific region of the Y-chromosome does not recombine and has an effective population size of one quarter that of the autosomal markers. Due to the strong influence of genetic drift and the effect of patrilocality the Y-specific polymorphisms show a remarkable geographic clustering. The Y-chromosome is inherited from father to son, so studying Y-specific polymorphisms provides a unique perspective on the patterns of male migration, serving as a complement to the information on female migration provided by the mtDNA. However, it is important to realize that due to the pattern of inheritance of the mtDNA and the Y-specific region (uniparental and nonrecombinant) they behave as a single genetic locus and only provide a partial perspective on the history of admixture. The global, multigenerational picture of admixture history can only be obtained by incorporating autosomal markers.

**Admixture in African American and African Caribbean Populations**

The availability of AIMs has made it possible to gain new insights on admixture in African American and African Caribbean groups. The history of these populations is complex for a number of reasons: the enslaved Africans came from various regions within the African continent and there were differences in ethnic origins depending on the North American port of entry. Additionally, there have been regional differences in the history of migration and admixture within the US and the Caribbean. The research carried out during the last decade reports substantial geographic variation in admixture levels, but several trends can be identified. The European contribution tends to be lower in African Caribbean communities than in US African Americans. There have been several admixture studies in the Caribbean, and the estimates of the European contribution are typically below 15%. The lowest contribution has been described in Tobago, with an average European ancestry of 4.6%. Our estimate of European ancestry in Jamaica was slightly higher (6.8%), while the European contribution to African Caribbeans living in Britain was 10.2%. The highest European contributions described in this region correspond to Trinidad, with an average value around 16%.
Pharmacogenomics in Admixed Populations

In the continental US, the European contribution to African American communities also shows a substantial variation, but in general is higher than in the Caribbean. Figure 3 depicts a map showing the percentage of European ancestry observed in several African American samples throughout the US. These estimates were obtained using especially selected AIMs and are quite precise. The pioneering admixture studies that took place in the 1950s and 1960s seemed to indicate that northern populations have a higher level of European ancestry than southern populations. However, the recent data summarized in Figure 3 show that the situation is much more complex than previously thought. Although the European admixture proportions in some southeastern African American communities are relatively low (e.g., South Carolina and Georgia), there is substantial variation in admixture levels within northern and southern populations. Additionally, the west coast communities tend to show the highest admixture estimates, although the number of samples analyzed in this area is still very limited and additional studies will be required to confirm this point.

Among the African American groups, the Gullah-speaking Sea Islanders from South Carolina show the lowest European contribution (3.5%). This is in agreement with historical, cultural and anthropological data indicating that the Gullah have been relatively isolated throughout history and have retained numerous African characteristics in language, social organization, religion, magic, art, folklore and music. The proportion of European ancestry is also relatively low in the “Low Country” around Charleston, SC (ranging from 9.9% and 14%) and in Georgia (10%). The European contribution tends to be higher in other regions of the US, ranging from 11.0% to 22.5% in other southern states, from 12.8% to 20.2% in the northeast and from 20% to 35% in the west coast. It is clear that the differences in admixture proportions observed between African American groups cannot simply be explained in terms of geography (e.g., differences between northern and southern states). For example, the admixture proportions in New Orleans (22.5%) are higher than in most African American communities in the northeast. This could be due to the unique history of Louisiana. This area was under French rule for a substantial period, until it became part of the Spanish territory in 1763.
and finally part of the US in 1803. Both the geographic origin of the enslaved Africans imported to Louisiana and their status during the French domination have been distinct from what occurred in the British Colonies, and there have been historical accounts of substantial intermarriage in the New Orleans area. Finally, it is also important to consider the recent demographic history of the US African American populations. In the period following World War I there were significant changes in the distribution of African Americans in the US. In the largest internal migration in the history of North America, southern African Americans, constituting the immense majority of the total African-American population (around 90%), left the rural south in search of new opportunities in the urban areas of the north. It is known that big cities like Chicago, Detroit, New York, Philadelphia, Pittsburgh and Baltimore, experienced a very significant increase in the number of African-American residents, both in absolute and in relative terms. These internal migrations have probably had a major effect in defining the current distribution of admixture proportions depicted in Figure 3.

In addition to the data on autosomal markers, further insights on the nature and dynamics of admixture can be obtained using maternally and paternally transmitted markers (mtDNA and the nonpseudoautosomal region of the Y chromosome, respectively). As described in the previous section, the effect of genetic drift is much more pronounced in mtDNA and the nonpseudoautosomal region of the Y chromosome than in the autosomes. Consequently, in human populations the mtDNA and the Y-specific region show continent-specific polymorphisms that can provide useful insights on the admixture process. In a study published in 1998, we estimated male and female European contributions to eight African American groups throughout the US and one African Caribbean sample (Jamaica). The results of this analysis are shown in Table 2. The first column of the table depicts the estimate of maternal European ancestry based on the L macrohaplogroup and the H haplogroup. The second column depicts the estimate based on the YAP insertion. There is strong evidence of a sex-biased European contribution. In every population analyzed there is evidence of a higher European male contribution, as shown by the Y/mtDNA ratios. Therefore, even if marriages between African-American males and European-American females are presently more common than marriages between African-American females and European-American males,

<table>
<thead>
<tr>
<th>Sample</th>
<th>mtDNA (%)</th>
<th>Y Chromosome (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maywood, Ill.</td>
<td>8.31</td>
<td>24.32</td>
</tr>
<tr>
<td>Detroit</td>
<td>0.00</td>
<td>30.33</td>
</tr>
<tr>
<td>New York</td>
<td>9.11</td>
<td>18.58</td>
</tr>
<tr>
<td>Philadelphia 1</td>
<td>11.02</td>
<td>22.94</td>
</tr>
<tr>
<td>Philadelphia 2</td>
<td>2.84</td>
<td>23.55</td>
</tr>
<tr>
<td>Pittsburgh</td>
<td>9.90</td>
<td>23.87</td>
</tr>
<tr>
<td>Baltimore</td>
<td>14.94</td>
<td>22.79</td>
</tr>
<tr>
<td>Charleston, SC</td>
<td>6.46</td>
<td>NA</td>
</tr>
<tr>
<td>New Orleans</td>
<td>7.04</td>
<td>46.88</td>
</tr>
<tr>
<td>Houston</td>
<td>6.80</td>
<td>8.55</td>
</tr>
<tr>
<td>Jamaica</td>
<td>12.93</td>
<td>17.89</td>
</tr>
</tbody>
</table>

Source: Parra et al.11
it seems clear that during a substantial part of African-American history European males have made a more significant genetic contribution to the African-American gene pool than European females. This is in accordance with the historical data regarding the period of slavery in the U.S.42

Another relevant issue to African American history is the extent of the Amerindian contribution. There have been accounts of substantial contact among Native Americans and people of African descent in specific periods of US history, especially in regions like the Mississippi delta and Florida.46 Some early anthropological reports emphasized the high proportion of African-American college students claiming some Amerindian ancestry.47,48 The same can be said of some African Caribbean populations. In fact, the Garifuna, also known as Black Caribs, living in Guatemala, Honduras and Belize trace their ancestry to the indigenous Caribbean populations (Caribs and Arawaks) and West Africans.49 Several studies have evaluated the extent of the Native American contribution to African Americans and African Caribbeans, using a panel of AIMs that incorporate markers showing large frequency differences between West African and Native American populations and European and Native American groups. These studies indicate that the Native American contribution has been relatively low, with values generally lower than 5%. Using a panel of 38 AIMs, we reported that the average Native American ancestry in African Americans living in Washington D.C. was 2.7% and the Native American contribution to British African Caribbeans was 1.9%.12 A recent study in three African American populations (Winston-Salem, Sacramento and Pittsburgh)41 also reported low Native American contributions (3.9%, 4.9% and 0.9%, respectively). The Native American ancestry in African Caribbean samples from Trinidad and Tobago was estimated as 10% and 1.4% respectively.35,36 Research using the maternally inherited mtDNA shows general agreement with the autosomal data. We tested 10 African-American samples and a sample from Jamaica for the presence of the common Amerindian-specific mtDNA haplogroups (A, B, C and D) and detected just four individuals with an Amerindian haplogroup among more than 1,000 individuals.11 Therefore, the available data seem to indicate that the Native American contribution to African Americans and to African Caribbeans has been quite limited. However, it is important to emphasize that the information remains very incomplete, and given the diverse history of these populations, it is quite possible that the Native American ancestry will be higher in some groups. For example, the Native American contribution seems to be relatively high in the Black Caribs, a finding that is consistent with the unique history of this Caribbean population. A study in a population of Black Caribs of Guatemala using classic genetic markers (erythrocyte antigens, serum proteins and enzymes) reported 75.2% West African, 22.4% Native American and 2.4% European contributions.50 Additionally, two mtDNA studies were carried out in Black Carib communities from Belize and Honduras. In Belize, the majority of the mtDNA corresponded to African lineages, and only one out of 28 individuals showed the 9-bp deletion characteristic of Native American populations. In Honduras,51 the maternal Native American contribution was estimated to be around 16%.

Admixture in Hispanic Populations

As mentioned in the introduction, the term “Hispanic” or “Latino” is used to refer to individuals with very diverse cultural and historical backgrounds. One of the characteristics that most of the “Hispanic” populations share is a history of recent admixture, which began soon after the arrival of Columbus to America in 1492. However, this history of admixture has been quite heterogeneous. In this section, I summarize information available for the three major North American Hispanic groups: Mexicans (and Mexican Americans), Puerto Ricans and Cubans.

The majority of the contemporary Mexican population consists of “mestizos” or admixed individuals. According to the 2000 Mexican census, 60% of the population are mestizos, 30% are Native Americans and 9% are people of European ancestry.52 The available genetic
Admixture in North America
data indicate that there is a wide dispersion of admixture proportions in mestizo populations throughout Mexico, with the Native American contribution ranging from 28% to 76%, the European contribution from 16% to 71%, and the West African contribution from 1% to 40% (reviewed in Bonilla et al ref. 53). There are several reasons for the wide range of ancestry estimates reported in the literature: the type and number of genetic markers used in the different studies, differences in the regional histories of the Mexican states, and differences in the characteristics of the samples (e.g., socioeconomic status). Some of the early admixture studies were based on blood groups, serum or red cell enzyme polymorphisms, which are not nearly as informative for inferring admixture proportions as the AIMs used in the most recent research. It is also important to emphasize regional differences in population history within Mexico. For example, the high West African contributions that have been reported in the states located on the east coast of Mexico (e.g., Campeche, Yucatan, Tabasco and Veracruz), where West African admixture proportions range between 20% and 40%53 are consistent with historical reports indicating a substantial West African presence around the Gulf coast and areas of southwest Mexico (Oaxaca, Guerrero), regions where the largest Afro-Mexican communities in Mexico are located today.7 Additionally, there is evidence that socioeconomic status is strongly related to individual admixture proportions in Mexico. In a recent study in Mexico City, we observed a positive association of education and European ancestry. In a logistic model with education as a dependent variable, the odds ratio for higher educational status associated with an increase from 0 to 1 in European admixture proportions was 9.4 (95% credible interval: 3.8-22.6). These data indicate that, in agreement with previously published studies, in Mexico not everyone has the same access to education.54 Therefore, we would expect estimates of mean admixture proportions to vary between studies that have sampled different socioeconomic groups.

There have also been several admixture studies in Mexican-American populations. The results are consistent with our earlier discussion focused on Mexican populations. Native American ancestry ranges from 18% to 52%, West African ancestry is typically lower than 10%, and the balance of the gene pool is of European origin.23,55-61 There is also evidence of population stratification in Mexican Americans. For example, Relethford et al59 demonstrated the effect of social class subdivision on admixture levels in Mexican Americans living in San Antonio, Texas. The wealthier socioeconomic group, residing in the transitional neighborhoods, exhibited the highest European ancestry (82%), while European ancestry levels for individuals from the low-income barrios were significantly lower (54%).

The information from the uniparentally transmitted markers (mtDNA and Y-specific markers) indicates that the admixture process has been sex-biased in Mexico, with the Native-American contribution coming mainly from the females and the European contribution from the males. In our recent study in Mexico City, we analyzed 69 autosomal AIMs and several mtDNA and Y-specific polymorphisms. The average autosomal Native American contribution was 65%. In contrast, the Native American genetic contribution for the maternally inherited mtDNA was estimated as approximately 90%, and the paternal Native American genetic contribution around 40%. The European-specific markers showed the reverse picture, with a European maternal contribution of around 7% and a paternal contribution of 60% (the average autosomal European contribution was 30%). This sex-biased contribution has already been described in many other admixture studies throughout the Americas62-68 and is consistent with historical reports indicating that during colonial times Spanish men embarking on the conquest of America commonly practiced unions with Native American women.69

The other major Hispanic groups, Puerto Ricans and Cubans, have not been studied as extensively as Mexican Americans. Hanis et al57 were the first authors to report admixture estimates for these two groups. Their estimates of the European, Native American and West African genetic contributions in a sample of Puerto Ricans were 45%, 18% and 37%, respectively. For the Cuban sample, the estimates were 62%, 18% and 20%, respectively. Recent studies in Puerto Ricans show general agreement with Hanis’ estimates. Using a
Pharmacogenomics in Admixed Populations

A panel of 35 AIMs, an analysis of a sample of Puerto Rican women living in New York reported ancestry proportions of 53% European, 18% Native American and 29% West African. Choudhry et al. recently studied a group of Puerto Rican asthmatics and controls with 44 AIMs and observed European contributions of 65.5% (cases) and 59.7% (controls), Native American contributions of 18.3% (cases) and 19.1% (controls) and West African contributions of 16.2% (cases) and 21.3% (controls). Therefore, the available data indicate that the proportion of West African ancestry in Hispanic Caribbean populations is, on average, higher than in Mexican American groups. In contrast, Mexican Americans typically have higher Native American proportions than Puerto Ricans and Cubans. This is not surprising given the historical evidence described above, although it is important to remember that within Mexico there are substantial regional differences in admixture. The known disparities in the history of admixture between Mexico and the Caribbean, combined with the differential geographic distribution of the major Hispanic groups in the US, explain the east/west differences in admixture proportions observed in the Hispanic population as a whole. As Bertoni et al. reported in a recent systematic study of admixture in Hispanics, in the west of the US, where Hispanics are predominantly of Mexican origin, the Native American contribution tends to be higher than in the east, where the majority of Hispanics are of Puerto Rican or Cuban origin. Conversely, the West African contribution tends to be higher in the east of the US.

Finally, I discuss the issue of sex-biased gene flow in the Caribbean Spanish colonies. In this regard, there have been several mtDNA studies in Puerto Rico that can shed some light on this topic. Bonilla et al. directly compared autosomal and mtDNA ancestral contributions to a sample of Puerto Rican women from New York (ref. 72 and unpublished data). The ancestry proportions estimated using the autosomal markers were 53.3% European, 29.1% West African and 17.6% Native American. The ancestry proportions estimated using the mtDNA haplogroups were approximately 9.4% European, 26.4% West African and 64.2% Native American. Similarly to what has been reported in many other Latin American populations, there is strong evidence of sex-biased gene flow. The majority of the mtDNA contribution is of Native American origin, while the autosomes show a major European component. It is important to mention that the estimates of maternal contributions obtained by Bonilla et al. show remarkable agreement with a systematic mtDNA analysis recently carried out by Martinez-Cruzado et al. in Puerto Rico. Their estimates of maternal contributions were 61.3% Native American, 27.2% sub-Saharan African and 11.5% West Eurasian. Therefore, although historical reports indicate that the Taino disappeared from Puerto Rico late in the 16th century, the majority of present-day Puerto Ricans have Native American ancestry in their maternal lineages. The information compiled in the last decade using autosomal and uniparental genetic markers clearly indicates that the current Mexican and Puerto Rican populations have been defined by the admixture process that took place between European males and Native American females. In Mexico, approximately 90% of the maternal lineages are of Native American ancestry, implying that there has been very little European female contribution throughout colonial and post-colonial history. In Puerto Rico the situation is slightly different. The initial admixture process that took place in a relatively short period (mainly the 16th century) between the Spanish males and the Taino women and the subsequent growth of the admixed population explain the large Native American maternal contribution observed today (~60%), but there has been a higher West African and European female gene flow than in Mexico. Unfortunately, the information for Cuba is very limited. An mtDNA study carried out by Torroni et al. in the Cuban province of Pinar del Rio suggested a European maternal contribution of 50%, a West African contribution of 46% and a Native American contribution of 4%. These values are very different to what has been described in Puerto Rico. However, further studies will be required to confirm this point.
Implication of Admixture for Pharmacogenomics

Admixture and Distribution of Genetic Variation Involved in Drug Metabolism and Drug Response

Anthropological and genetic evidence indicate that the human species has a recent origin, which can be traced back to Africa around 200,000 years ago.\textsuperscript{74} Given this recent origin it is not surprising that the genetic differences between continental groups are, in general, small. Numerous studies indicate that in humans the average autosomal Fst value (the percentage of the genetic variance explained by differences between continental populations) is typically between 0.10-0.20.\textsuperscript{75} However, Fst values show a wide dispersion around the mean (Fig. 4). For most loci, the differences between geographic regions are small, but there is a subset of loci for which the diversity between regions is high. This subset may include loci that are related to drug metabolism and drug response. In 2001, Wilson et al\textsuperscript{76} showed that four out of six drug-metabolizing enzymes had significant frequency differences between three ethnically labeled groups. Similarly, Tate and Goldstein\textsuperscript{77} reported that out of 42 genetic variants that have been significantly associated with drug response in at least two studies, more than two-thirds show significant differences in frequency across populations and nearly one-third have allele frequency differences higher than 0.20. Obviously, these findings have important implications in pharmacogenomics.\textsuperscript{78} In this sense, when there are large differences in the frequency of genetic variants involved in the biological response to drug treatment between continental populations, this information can be taken into account during the therapeutic decision-making process. With respect to the North American admixed groups, the history of admixture and the relative contributions of the parental populations will determine the distribution of the relevant variants. It is important to emphasize that, although knowledge about the population distribution of susceptibility variants is helpful, what really is critical is the individual genotype.

Figure 4. Histogram and cumulative distribution of Fst values (Weir and Cockerham’s unbiased estimator) for 8,525 autosomal SNPs. Source: Shriver et al.\textsuperscript{25}
Pharmacogenomics in Admixed Populations

Due to the recent separation of human populations there will always be, to some extent, overlap in genotype frequencies between populations. The ultimate goal will be to have individualized treatment according to the genetic makeup of each person, but this goal is still beyond reach, so at this point increasing the information available at the population level is extremely important.

Population Stratification in Admixed Populations: Imlications for Candidate Gene Association Studies

Candidate gene association studies have been widely used to understand the genetics of drug response. Although genome-wide association studies are becoming increasingly feasible in terms of cost and genome coverage, candidate gene studies will remain useful in the near future. However, the presence of population stratification can cause spurious associations between a phenotype and unlinked candidate loci in population-based association analysis. This problem is particularly important in recently admixed populations because individual admixture proportions often vary significantly within the population. This variation in admixture proportions indicates a departure from random mating expectations and an increased potential for false positive results in association studies. Figure 5 shows a triangular plot depicting the variation in individual admixture proportion in three North American admixed samples: Mexicans, Puerto Ricans and African Americans from Washington DC. Note the substantial dispersion of individual admixture estimates around the average estimate for each sample. Using pigmentation as a model phenotype, we have shown that this “admixture stratification” can have a profound effect on population-based association studies. We measured quantitatively skin pigmentation and estimated admixture proportions using 33 AIMs in two admixed samples (African Americans and African Caribbeans). Approximately one-half of the AIMs were significantly associated with skin pigmentation in these samples, even though most of the AIMs are located in genomic regions with no pigmentation candidate genes. In these samples, many markers gave a significant result not because they have a functional effect on pigmentation, but because they are informative for ancestry and pigmentation and ancestry show a strong correlation due to admixture stratification. After adjusting for individual admixture proportions, most of the significant effects disappeared and only two AIMs located within two pigmentation candidate genes (TYR and OCA2) remained significant. In admixed populations, admixture stratification is also reflected in a larger than expected proportion of significant associations between unlinked AIMs. We have detected the presence of admixture stratification in all the African American and Hispanic admixed population samples that we have analyzed but there is substantial variation in the degree of stratification.

In admixed populations, two factors can be responsible for the observed admixture stratification: continuous gene flow and assortative mating. Continuous gene flow refers to an admixture model in which there has been an ongoing contribution from one or more parental populations to the admixed population over a period of time extending into the recent past. Obviously, the process of admixture in North America has been taking place for several centuries, and it is still actively configuring our present societies, so this undoubtedly explains some of the stratification present in admixed populations. In this sense, recent gene flow is of particular significance because recombination does not have enough time to break the associations between unlinked markers. Assortative mating can be defined as nonrandom mating according to phenotypic characteristics. If there is assortative mating based on any factor (e.g., socioeconomic status, education, skin pigmentation) that is correlated with ancestry, any population structure originally present in the admixed population will be maintained through the generations. There is evidence indicating that this may explain some of the stratification observed in Hispanic populations, in particular populations of Mexican ancestry. As mentioned above, Relethford described a strong association of social status with ancestry in Mexican Americans living in San Antonio, TX, and we have recently observed a strong association of education status with ancestry in a large sample from Mexico City.
Because in most societies mating is assortative with respect to socioeconomic status, if socioeconomic status is correlated with ancestry, admixture stratification will be present in the population. The implications of these associations are profound. For example, if there are differences in socioeconomic status between a group of cases and controls, there will also be differences in ancestry proportions between both groups, and any

Figure 5. Triangular representation of individual ancestry estimates. A. The figure shows how to interpret triangular plots depicting the relative Native American (N-AM), European (EUR) and West African (W-AF) genetic contributions in any individual. Shown is a hypothetical example of an individual with 50% N-AM contribution, 35% W-AF contribution and 15% EUR contribution. B. Triangular plot showing individual ancestry proportions in African Americans from Washington DC (blue circles), Puerto Ricans (yellow squares) and Mexicans (red triangles). Average ancestry proportions in the samples are indicated with blue, yellow and red arrows, respectively.
marker showing large differences in frequency between the parental populations will be associated with the disease, even if it is not functionally relevant. In order to avoid these spurious associations, it is critical to take measures to control for potential confounding due to admixture stratification. In admixed populations, a good strategy to avoid confounding is to genotype a large panel of AIMs to estimate individual admixture proportions and include individual ancestry as a variable in the analysis.\textsuperscript{86,87} The important issue of population stratification in admixed populations is the topic of another chapter in this book (Genomic control in admixed populations, by Eduardo Tarazona-Santos).

**Conclusion**

In the previous sections, I have provided a comprehensive review of admixture in the main North American admixed groups. The use of autosomal markers, in particular the powerful AIMs, in combination with mtDNA and Y-chromosome polymorphisms provides a complete picture of the history and dynamics of the admixture process. Although there are still many gaps in our knowledge, and some admixed groups and geographical areas have been poorly covered in the genetic studies, one of the clear messages from this research is that there is substantial heterogeneity in admixture proportions and population stratification within and between the main admixed groups (e.g., African Americans or Mexicans). This heterogeneity has the following implications in the field of pharmacogenomics: 1/ the frequency of some alleles that may be important in drug metabolism or drug response may show variation within and between admixed groups and 2/ in candidate gene association studies there will be differences in the extent of confounding due to population stratification. For these reasons, in pharmacogenetic studies in admixed populations, it is advisable to characterize the admixed sample with a panel of AIMs in order to estimate admixture proportions and to detect and correct for admixture stratification.

**Acknowledgements**

I wish to express my gratitude to Mark Shriver for the population data on AIMs to study admixture in North America, to Rick Kittles for the map of European admixture in African American communities, to Carolina Bonilla for the unpublished mtDNA data from Puerto Rico and to Laura Simmonds for research assistance. Esteban J. Parra is supported by the Natural Sciences and Engineering Research Council of Canada (NSERC), the Canadian Institutes of Health Research (CIHR) and the Banting and Best Diabetes Centre of the University of Toronto.

**References**

http://presidencia.gob.mx/mexico/.
77. Tate SK, Goldstein DB. Will tomorrow’s medicines work for everyone? Nat Genet 2004; 36:S34-S42.