

RESEARCH ARTICLE

Whole genome structural analysis of Caribbean hair sheep reveals quantitative link to West African ancestry

Gordon L. Spangler¹, Benjamin D. Rosen^{1*}, Moses Babatunde Ilori², Olivier Hanotte², Eui-Soo Kim³, Tad S. Sonstegard³, Joan M. Burke⁴, James L. M. Morgan⁵, David R. Notter⁶, Curtis P. Van Tassell¹

1 Animal Genomics and Improvement Laboratory, Agricultural Research Service, United States Department of Agriculture, Beltsville, Maryland, United States of America, **2** School of Life Sciences, University of Nottingham, Nottingham, United Kingdom, **3** Acceligen of Recombinetics Incorporated, Saint Paul, Minnesota, United States of America, **4** Dale Bumpers Small Farms Research Station, Agricultural Research Service, United States Department of Agriculture, Booneville, Arkansas, United States of America, **5** Katahdin Hair Sheep International, Fayetteville, Arkansas, United States of America, **6** Department of Animal and Poultry Sciences, Virginia Tech, Blacksburg, Virginia, United States of America

* ben.rosen@ars.usda.gov



OPEN ACCESS

Citation: Spangler GL, Rosen BD, Ilori MB, Hanotte O, Kim E-S, Sonstegard TS, et al. (2017) Whole genome structural analysis of Caribbean hair sheep reveals quantitative link to West African ancestry. *PLoS ONE* 12(6): e0179021. <https://doi.org/10.1371/journal.pone.0179021>

Editor: Qin Zhang, China Agricultural University, CHINA

Received: October 24, 2016

Accepted: May 23, 2017

Published: June 29, 2017

Copyright: This is an open access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the [Creative Commons CC0](https://creativecommons.org/licenses/by/4.0/) public domain dedication.

Data Availability Statement: All the data required to replicate this study are publically available. Genotype files unique to this article (West African, St Croix, Bergamesca) can be found on the NAL repository. The dataset (doi: [10.15482/USDA.ADC/1351110](https://doi.org/10.15482/USDA.ADC/1351110)) is now published and can be accessed via <http://dx.doi.org/10.15482/USDA.ADC/1351110>. The remainder of the genotype files used for purpose of comparison are available in the ISGC website (<http://www.sheepmap.org/>). Interested researchers can access the data by

Abstract

Hair sheep of Caribbean origin have become an important part of the U.S. sheep industry. Their lack of wool eliminates a number of health concerns and drastically reduces the cost of production. More importantly, Caribbean hair sheep demonstrate robust production performance even in the presence of drug-resistant gastrointestinal nematodes, a rising concern to the industry. Despite the growing importance of hair sheep in the Americas their genetic origins have remained speculative. Prior to this report no genetic studies were able to identify a unique geographical origin of hair sheep in the New World. Our study clarifies the African and European ancestry of Caribbean hair sheep. Whole-genome structural analysis was conducted on four established breeds of hair sheep from the Caribbean region. Using breeds representing Africa and Europe we establish an objective measure indicating Caribbean hair sheep are derived from Iberian and West African origins. Caribbean hair sheep result from West African introgression into established ecotypes of Iberian descent. Genotypes from 47,750 autosomal single nucleotide polymorphism markers scored in 290 animals were used to characterize the population structure of the St. Croix, Barbados Blackbelly, Morada Nova, and Santa Ines. Principal components, admixture, and phylogenetic analyses results correlate with historical patterns of colonization and trade. These patterns support co-migration of these sheep with humans.

Introduction

The U.S. sheep industry has been in steady decline since the end of WWII. The national flock peaked at about 56 million head in 1942 and dropped to 7 million by 2000[1]. However, religious, ethnic, and organic markets for lamb and sheep products have grown substantially in

agreeing to the Terms of Access (<http://www.sheepmap.org/termsaccess.php>).

Funding: The authors received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

recent years[2]. Many consumers are attracted by the low-input, pasture-based production system of sheep. New producers are leveraging low start-up expenses and modest time commitment to maximize profitability of small production systems.

Hair sheep of Caribbean origin are ideal breeds for organic producers, especially those in sub-tropical regions. Their lack of wool eliminates the cost of shearing and obviates the need for practices such as mulesing or docking to prevent health risks particular to fleeced sheep. Heat tolerance facilitated by the characteristics of a hair coat allows these breeds to thrive in tropical, arid, and semi-arid regions[3]. High temperatures and humidity are environmental conditions ideal for internal parasites such as *Haemonchus contortus*. Caribbean hair sheep demonstrate robust performance even in the presence of drug-resistant gastrointestinal nematodes (GIN)[4,5,6,7]. Taken together, these factors result in reduced input and maintenance costs, decreased labor demands, and an overall greater return on investment than conventional wool breeds.

Despite the growing importance of hair sheep in the Americas their origins remain speculative. Prior to this report no genetic studies were able to identify a unique geographical origin of hair sheep in the New World. European, African, and Caribbean hair sheep share a common mitochondrial lineage[3], but different Y-chromosomal, and autosomal backgrounds[8]. This is logical considering the history of migration in the Old World and the colonization of the New.

During the late colonial period, sheep were transported around the globe as part of the settlement by European powers[9]. They took Merino from Spain and several British breeds to South America, South Africa, Australia and New Zealand[10]. It is often erroneously assumed these sheep mixed with indigenous breeds upon arrival to the Caribbean. However, there are no domestic sheep indigenous to the Americas.

Further confusion surrounds the African component from which these breeds originate. A typical example would be the Dorper, which is a composite breed of Dorset Horn and Blackhead Persian (DorPer)[11]. The Blackhead Persian is a fat-tailed breed and very different from the thin-tailed hair sheep of West Africa. Although classified as a hair breed, the Dorper has been shown to differ in performance from strictly Caribbean hair sheep such as the St. Croix [12], Morada Nova[13], or St. Ines[14, 15]. Another important consideration is the introgression of North African traits into the Iberian breeds developed during the Moorish occupation of Spain[16].

Recently, the greater affordability of high-throughput sequencing and high-density genotyping methodologies has facilitated impressive research into the historic development of domesticated sheep. The first draft of the sheep genome led to the development of a highly informative panel of single nucleotide polymorphisms (SNP)[17]. The International Sheep Genome Consortium (ISGC) used Hapmap data constructed from 50 thousand SNP to identify regions that have undergone positive selection across a global panel of breeds[18].

Researchers have taken the logical next step and applied these tools to elucidate the genetic structure of breeds specific to local geographic, economic, and environmental interests [19,20,21]. Dense genotyping also provides a means of mapping desired traits to specific regions of the genome designated quantitative trait loci (QTL)[22]. This identifies not only phenotypes controlled by single genes, but complex traits produced by the interaction of many genes[1, 4, 23,24].

Careful analysis of genetic structure and population stratification is an important step in understanding the genetic basis of complex traits[25]. The search for QTL depends on the identification of loci displaying outlier behavior and in itself is not proof that selection has taken place unless factors such as human-mediated selection and transportation, climate adaptation, admixture, gene flow, and introgression between breeds are taken into account[26].

With this in mind, our study attempts to clear away the ambiguity of African and European Ancestry in Caribbean hair sheep. In this study we perform whole-genome structural analysis using representative breeds from Africa and Europe to establish a quantitative link between Caribbean hair sheep and their specific origins. This is a preliminary step towards discovery of QTL linked to traits such as heat resistance, forage conversion, and parasite resistance.

Methods

Sample collection

Djallonke sheep are the dominant breed of West and Southwest Africa and not estimated to be endangered. Blood samples from Nigerian sheep were collected commensally during routine veterinary treatments, and all owners agreed to the analysis; no further specific permissions were required from the University of Nottingham Animal Care and Use Committee (IACUC) for this study. Ten Djallonké blood samples were taken via jugular venipuncture into Vacutainer tubes (Becton Dickinson, Franklin Lakes, NJ) containing EDTA from a privately owned flock in southwestern Nigeria with the owner's permission. Genomic DNA was isolated from blood samples using standard procedures and stored at -20°C . Animals were genotyped using the OvineSNP50 BeadChip [Illumina, Inc., San Diego, CA]. Genotyping was performed by GeneSeek, Inc. (Geneseek/Neogen, Lincoln, NE). Raw data were analyzed using the GenomeStudio V2011.1 and Genotyping Module v.1.9.4. All call rates were greater than 99%. The genotype file was then exported to Golden Helix SNP & Variation Suite v8.x (Golden Helix, Inc., Bozeman, MT, <http://www.goldenhelix.com>) for further filtering and analysis.

Illumina 50K ovine genotypes of St. Croix sheep were obtained from the National Animal Germplasm Program, National Center for Genetic Resources Preservation, ARS, USDA, Ft. Collins, CO. Sample collection and DNA purification were as described by Blackburn et al [27].

Breed selection

Using the method of Price et al [28], we first assessed the ordination of the Djallonké genome in relation to the structural variation of a global panel of 74 breeds provided by ISGC [18] (<http://www.sheephapmap.org>). This included three Caribbean hair sheep breeds: Barbados Blackbelly, Morada Nova, and Santa Inés and four wool breeds from the America's: Gulf Coast Native, Brazilian Creole, Navajo Churra, and St. Elizabeth. To this we added ten St Croix genotypes obtained from the National Animal Germplasm Program (http://nrrc.ars.usda.gov/A-GRIN/main_webpage).

We then conducted a model-based Bayesian analysis [29], in conjunction with a cursory examination using SNPhylo [30] and Treemix [31] on the complete set of 76 breeds of sheep to identify only those that may possibly have contributed to the pedigree of Caribbean hair sheep. We normalized this set to 10 individuals each to coincide with the number of Djallonké and St. Croix sheep to which our collection was limited. We finally selected only those breeds that could reasonably have contributed to the development of hair sheep through colonization and subsequent migration throughout the Caribbean (see Table 1).

In addition to breeds representing the respective colonial influences in the region, we included several breeds as models for drift and admixture (see Table 2). The Navajo Churro is a direct descendant of the Spanish Churra brought to the Americas during the period of colonization. The Australian Merino and Australian Suffolk correspond to their British counterparts arriving in Australia sometime during the 19th century [32].

The Australian Suffolk is also used to demonstrate admixture. In the U.S. the polled Dorset was the result of a mutation, but in Australia it was developed from introgression of the poll

Table 1. List of breeds and summary statistics.

Breed	Location	P _n	H _o	H _e	f	A _r	*HS _{WAD}
African Group							
Namaqua Afrikaaner	South Africa	0.688	0.288	0.254	0.269	1.688	14.5
Ronderib Afrikaaner	South Africa	0.811	0.322	0.296	0.187	1.880	14.0
Djallonké	Nigeria	0.769	0.286	0.277	0.277	1.769	-
African Dorper	South Africa	0.884	0.340	0.333	0.137	1.884	11.7
Egyptian Barki	Egypt	0.948	0.369	0.371	0.066	1.948	13.2
Red Maasai	Kenya	0.884	0.326	0.325	0.173	1.884	15.6
Ethiopian Menz	Ethiopia	0.871	0.326	0.326	0.176	1.871	15.2
Iberian Group							
Australian Merino	Australia	0.952	0.368	0.376	0.067	1.952	11.1
Bergamasca	Italy	0.940	0.364	0.369	0.078	1.939	10.2
Castellana	Spain	0.951	0.381	0.377	0.082	1.951	10.8
Ojalada	Spain	0.953	0.378	0.377	0.045	1.953	11.2
Rasa Aragonesa	Spain	0.965	0.385	0.386	0.022	1.965	11.0
Spanish Churra	Spain	0.938	0.364	0.366	0.076	1.938	11.2
Spanish Merino	Spain	0.951	0.376	0.374	0.046	1.951	10.7
British Group							
Dorset Horn	UK	0.811	0.325	0.286	0.182	1.811	7.4
Dorset Poll	Australia	0.893	0.347	0.342	0.118	1.893	7.9
Soay	UK	0.745	0.269	0.268	0.317	1.745	10.0
Australian Suffolk	Australia	0.941	0.376	0.371	0.049	1.941	8.7
Irish Suffolk	UK	0.864	0.315	0.328	0.199	1.864	9.4
Wiltshire Horn	UK	0.752	0.272	0.263	0.309	1.752	8.0
New World							
Brazilian Creole	Brazil	0.946	0.341	0.371	0.136	1.946	10.8
Navajo Churro	SW USA	0.935	0.372	0.365	0.260	1.935	10.3
Gulf Coast Native	Alabama	0.951	0.373	0.373	0.055	1.951	10.7
St Elizabeth	Jamaica	0.946	0.374	0.372	0.050	1.946	13.0
Caribbean Hair							
Barbados Blackbelly	Barbados	0.880	0.331	0.332	0.159	1.880	23.8
Morada Nova	Brazil	0.860	0.311	0.319	0.208	1.860	21.1
Santa Inés	Brazil	0.919	0.352	0.353	0.107	1.919	23.8
St Croix	Virgin Islands	0.832	0.285	0.307	0.274	1.831	16.4
Asia							
Norduz	SW Asia	0.891	0.356	0.339	0.102	1.891	12.0

Intrapopulation summary statistics for 38, 337 SNP loci from 29 breeds used in this study: coefficient of inbreeding (f), fraction of polymorphic loci (P_n), observed heterozygosity (H_o), expected heterozygosity (H_e), allelic richness (A_r), and mean IBD shared between Djallonké and sample breed in Mb.

<https://doi.org/10.1371/journal.pone.0179021.t001>

gene into the Dorset Horn from two “Down” breeds; the Morrisdale and Ryeland[33]. We use the Australian Suffolk to represent those breeds in this study. The Dorper is a composite breed developed in South Africa by crossing the Blackhead Persian from Somalia with the Dorset Horn. It is used here to demonstrate a known East African/European cross, and assist in distinguishing breeds of West African/European ancestry. We use the Ethiopian Menz and Red Maasai to represent the Blackhead Persian, which was unavailable at the time of this study. The St. Inés is a production hair sheep in Brazil developed by crossing the Morada Nova and

Table 2. Control models used in this study.

Drift		
Spanish Churra	→	Navajo Churro
Spanish Merino	→	Australian Merino
Irish Suffolk	→	Australian Suffolk
Admixture		
Dorset Horn + *Corriedale/Ryland	=	Australian Dorset Poll
Dorset Horn + †Blackhead Persian	=	Dorper
Morada Nova+Bergamesca	=	Santa Inés

*The Corriedale/Ryland is represented by the Australian Suffolk.

†Blackhead Persian is represented by the Red Maasai/Ethiopian Menz.

<https://doi.org/10.1371/journal.pone.0179021.t002>

Bergamesca[34]. The Bergamesca is a relative of the Merino imported to Brazil from Italy in the 1930s[35].

Finally, the Norduz is used in this analysis to root the phylogeny of the African breeds to the assumed Mesopotamian center of sheep domestication[36]. The Egyptian Barki is used to connect Mediterranean influence with North Africa and Iberia.

Quality measures, summary statistics, and structural analysis

After merging all data sets we were left with 38,337 autosomal SNP from 290 individuals representing 29 breeds. Relevant diversity statistics of the coefficient of inbreeding (f) and percentage polymorphic loci (P_n), Heterozygosity Expected (H_e) and Observed (H_o), and Allelic Richness (A_r) were calculated using the PLINK version 1.9 (Purcell & Chang <https://www.cog-genomics.org/plink2>)[37].

Haploype sharing (HS_{WAD}) was calculated using the mean size of the consensus haplotypes estimated between all pairs of individuals and between Djallonké and other breeds[38]. The haplotype phase was inferred using Beagle[38] and the summary statistics of the haplotype sharing was computed using Perl.

Removal of SNP in high linkage disequilibrium (LD) was used to counter the effect of ascertainment bias and generate meaningful comparisons among our breed populations[39]. We applied a stringent threshold of 0.05 for initial PCA analysis resulting in 5,790 sites. Generation of Eigen values from allele frequencies was performed by Golden Helix software using the Additive Model (dd)→(Dd)→(DD). A second analysis to remove outliers and increase resolution was performed by recomputing principal components 5 times removing outliers more than 3 standard deviations from 5 components. Admixture, phylogeny analysis, and phylogeny checks using f -statistics were again performed on the final dataset as described above.

Results

Summary statistics

All statistics for the ISGC breeds were in accordance with those reported. P_n for Djallonké was 0.74 (see Table 1). The f -statistic for Djallonké (0.24) and St. Croix (0.29) were slightly higher than other breeds. This may be due to the single flock sampling of individuals used in this study. Caribbean hair sheep show the highest degree of haplotype sharing with the Djallonké (HS_{WAD}) of any other group studied including the wool breeds from the same geographical region ($p < 0.03$) (see New World group Table 1).

Principal components analysis

We first placed the West African Djallonké, St. Croix, Navajo Churro, and Bergamesca into the complete set of genotypes provided by ISGC [18]. The Bergamesca and Navajo Churro clustered indistinguishably within the group designated “Iberian/Italian/Merino” by ISGC. The St. Croix clustered with “Brazilian/Caribbean” group, and Djallonké with the “African” group. Reducing the number of individuals in each breed to 10 concentrated the clusters and did not significantly change the spatial relationship of the groups. However, this was not significant enough to resolve the relationship of the breeds of interest.

We therefore reduced the number of breeds to those which could have reasonably contributed to Caribbean ancestry and reran the analysis (Fig 1A). The cline demonstrated by the African breeds in PC01 and PC02 concurs with known migration of fat-tailed, course-wooled sheep into the continent from the origin of domestication [40], and the spatial orientation of these breeds roughly matches the geographic location of the sampling locations [41]. The Djallonké skews slightly higher in PC02 aligning it in a position of admixture with the Caribbean breeds and the Iberian. The position of the Dorper directly between the Menz/Maasai and Dorset Horn suggests this is a valid assumption. Further reinforcement of this assumption can be seen in the gradient demonstrated among Dorset Horn, Australian Poll Dorset, and Australian Suffolk. A shift between the Irish and Australian Suffolk is most likely due to admixture rather than drift, as the Spanish and Australian Merino are indistinguishable within the Iberian group. This is also the case with the Spanish Churra and Navajo Churro.

To further increase resolution we reran the analysis this time eliminating outliers (Fig 1B and 1C). This eliminated the Namaqua Afrikaaner, Wiltshire, and Dorset Horn from the data set. PC01 versus PC02 clustered all the European breeds except for the Irish Suffolk within the Iberian group. The New World fleeced breeds also clustered within this group. PC01 distinctly separated all of the remaining East African breeds from the West African placing it squarely within the quadrant containing only the Caribbean hair breeds. The Dorper which is a composite of approximately 70% Blackhead Somali (Maasai/Menz) clustered with the East African breeds.

PC02 versus PC03 (Fig 1C) maintained the separation of East and West Africa but PC03 separated the European group into British and Iberian breeds. The Morada Nova and Santa Inés clustered on the Spanish side, while the St. Croix and Barbados Blackbelly clustered to the British side. The African Dorper, Gulf Coast Native, and St. Elizabeth were also clustered to the British side while the Navajo-Churro and Brazilian Creole remained clearly within the Iberian group.

Admixture

The genetic structure of Caribbean hair sheep breeds was investigated by model-based Bayesian analysis using the software package fastStructure [29]. Results from $K = 3$ to 14 are shown in Fig 2. LD restrictions were relaxed to threshold level of 0.8 leaving 17,138 SNP. We tested theoretical population values from 2 to 24. After which a maximum likelihood value of optimal differentiation was calculated as $K = 14$. The pattern of differentiation generally follows that of Kijas et al [18] for breeds obtained from ISGC. However, the addition of the Djallonké and St. Croix clearly indicates a significant West African component in the Caribbean breeds.

$K = 3$ to 6 shows differentiation of the ancient breeds Dorset, Afrikaaner, Wiltshire, and Soay. However, we see a marked separation of the African groups with the addition of the Djallonké to this panel beginning at $K = 6$. Namaqua, Ronderib, Maasai, Menz, and Dorper segregate to the East African side, and all the Caribbean hair breeds cluster with the West African Djallonké. The fleece breed from St. Elizabeth (Jamaica) also shows some admixture with the

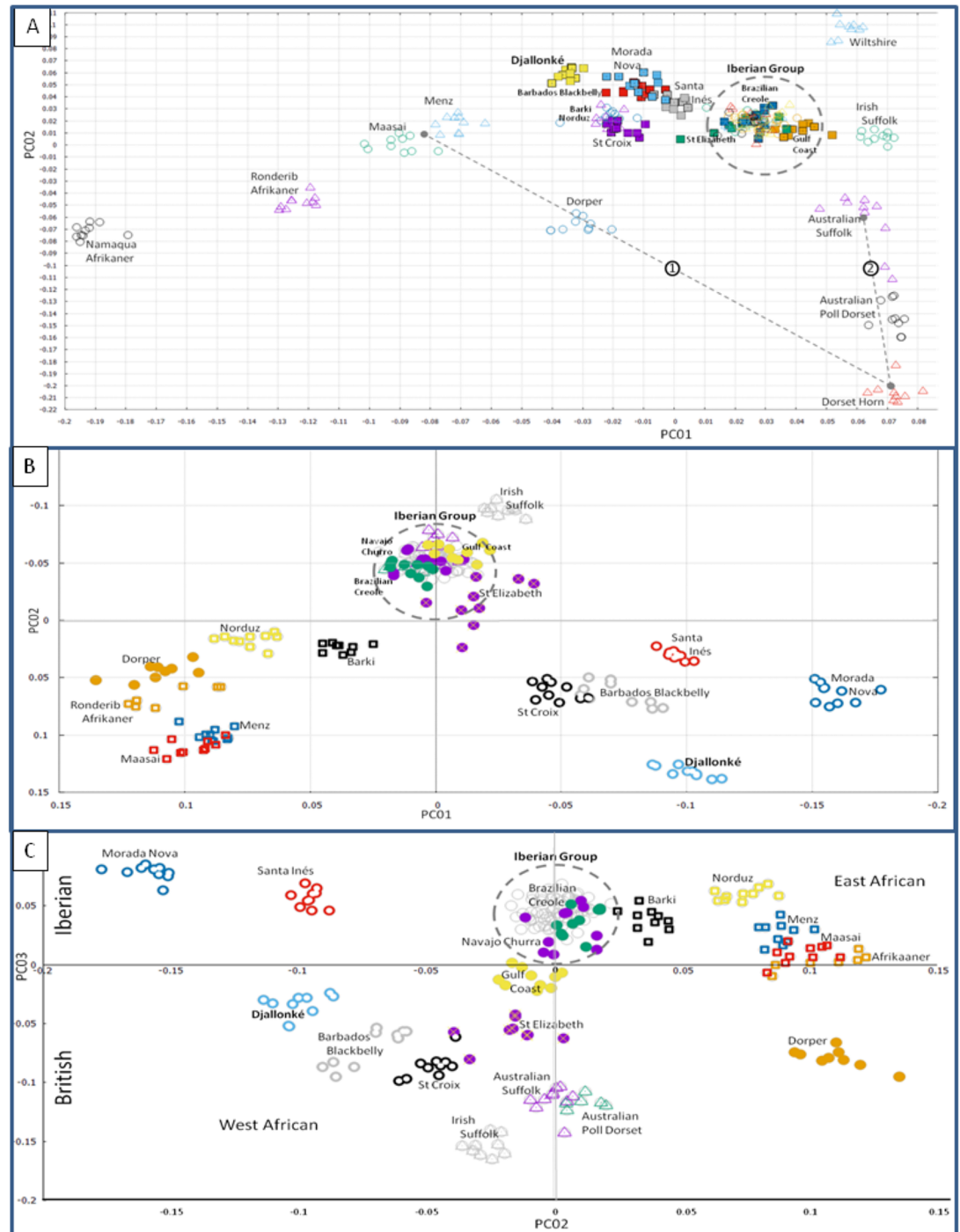


Fig 1. Principal components analysis. **A.** Multi-Dimensional Scaling (MDS) plot of the pair-wise Identity by State distances among the 27 breeds used in this study. The Iberian Group is indicated by a dashed circle and contains the New World Breeds: Navajo Churro, Brazilian Creole, Gulf Coast Native, and St Elizabeth. The Spanish and Australian Merino are indistinguishable in this group, as well as the Spanish Churra, Ojalada, Rosa Aragonesa, Castellana, and the Italian Bergamesca. Dashed lines indicate breeds of known admixture: (1)- The Dorper is a composite breed of the Dorset Horn and Blackhead Somali (Menz/Maasai used as substitutes). (2)- The Australian Poll Dorset was developed by inserting the poll gene into the Dorset Horn by crossing with Corriedale and Ryeland, a breed similar to the Australian Suffolk. **B.** MDS plot of Components 01 vs 02 after outlier removal. Component 1 clearly separates the West African Djallonké and the four Caribbean Hair sheep from the East African group: Menz, Maasai, and Afrikaaner. PC02 separates the Hair sheep from the European breeds leaving the Caribbean Hair sheep in a quadrant alone with the

West African Djallonké. The Dorper is approximately 70% East African admixture and clusters with this group. All of the European breeds cluster in or near the Iberian Group along with the New World wool breeds. **C.** *MDS plot of Components 02 vs 03 after outlier removal.* As in the above plot PC02 divides East and West African breeds. PC03 further separates the British from the Iberian.

<https://doi.org/10.1371/journal.pone.0179021.g001>

West African samples. The American wool breeds Navajo Churro, Brazilian Creole, and Gulf Coast Native remain with the Iberian group.

At K = 10 the genetic components of the Dorper agree with the documented make-up of this composite breed of approximately 70% East African and 30% Dorset Horn. At K = 12 we see indications of the mixtures used between the Australian Suffolk and Dorset Horn to create the Australian Poll Dorset. This mixture of shared genetic components demonstrates the contribution of the Australian Suffolk relative to the Irish Suffolk rather than geographic separation of the two. For this reason we also included the Spanish and Australian Merino and the Spanish and Navajo Churro, which show no cross mixture of components.

At K = 12 we see a divergence of the Caribbean Hair breeds St. Croix and Barbados Blackbelly from the Morada Nova and Santa Inés. At K = 14 we see the full elucidation of the Santa Inés, which was developed from the Morada Nova using Bergamesca and Blackhead Somali [34].

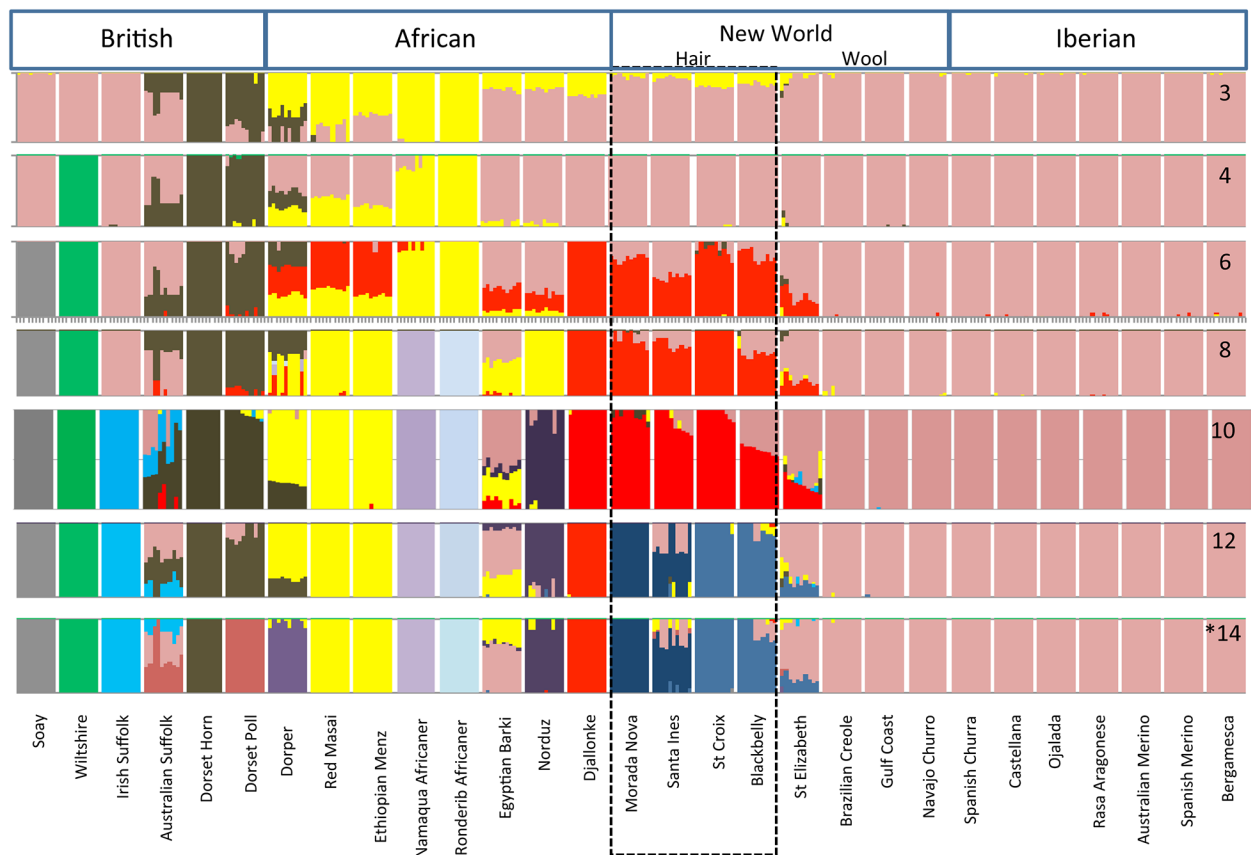


Fig 2. Admixture analysis of 29 breeds used in this study. Breeds are separated into blocks. Lines within each block represent individuals within the breed. Graphs for K = 3–14 show the admixture ratio of each theoretical population [K] within the individuals in a block. Groups are indicated at the top. Caribbean hair sheep are inscribed by dashed lines. Norduz is included with the African group for simplicity.

<https://doi.org/10.1371/journal.pone.0179021.g002>

Treemix

TreeMix uses allele frequencies to infer relationships among populations through the construction of a graph. The graph allows modeling of population splits as in traditional bifurcating tree methods but has the added benefit of indicating gene flow among groups (Fig 3)[31]. We used *f*-statistics to validate migration edges in the tree (Supplemental Table 1). The *f*₃ and *f*₄ statistics test for the validity of connections in three or four population trees respectively. The *f*₃ statistic showed significance for admixture in the Australian Suffolk, however all migrations identified showed significance with *f*₄ statistics.

We rooted our phylogeny on Norduz, which is the breed sampled closest to the center of domestication geographically. The nearest neighbor to the Norduz is the Barki of North Africa reflecting the accepted migration pattern of fat-tailed sheep into the continent[42]. After this the tree bifurcates into African and Iberian breeds.

The African branch splits East and West. The Ronderib and Namaqua Afrikaner are a more ancient branch, and Namaqua demonstrates considerable drift. The Dorper is in a clade with Menz and Maasai and receives a highly weighted migration event from the Dorset family. This correlates to the known composition of Blackhead Somali and Dorset to develop this breed [11]. The West African Djallonké shares a branch with the Barbados Blackbelly and St. Croix proving its significant influence on these hair sheep breeds and a strongly weighted migration event into the Morada Nova/Santa Inés branch.

Castellan leads to the Rasa Aragonesa, Ojalada, and the Spanish Churra. The Spanish Merino and Australian Merino share the same branch, which supports the idea that the PCA and fastStructure analyses are biased relative to Suffolk. The Navajo Churro, Brazilian Creole, and Gulf Coast Native are in the Iberian branch and no indication of African gene flow into these breeds is seen. Following in this clade are the Bergamesca, Morada Nova, and Santa Inés, which agrees with Paiva et al[34].

The British group ties into the Iberian group through St. Elizabeth and the Gulf Coast Native. Sheep of the British Isles developed through a northern migration from the center of domestication and therefore display less admixture with those of Africa or Iberia[36]. Jamaica (St. Elizabeth) and the Gulf Coast were colonized by the British and this would logically be where the two influences meet.

Discussion

We have established the first quantifiable link between Caribbean hair sheep and their West African ancestors. Previous investigation using mitochondrial and microsatellite markers were unable to specify a single source for the origin of hair sheep in the region[43,44]. Most agree the major influence on Caribbean hair sheep is Iberian, but fail to account for introgression of West African traits[45, 46].

We have shown which breeds share African genetic components and to what degree. Furthermore, we have eliminated some breeds previously assumed to contain African admixture. Our whole genome structural analysis shows that Caribbean hair sheep are descendants of Spanish breeds, most likely the Churra, and West African breeds. The New World wool breeds are of Spanish descent without West African introgression.

Hair sheep are important for the economic viability they bring to the sheep industry, and especially for their remarkable resilience in the presence of multi-drug resistant parasites such as *H. contortus*. Many in-depth studies searching for genomic loci responsible for these traits have produced inconclusive results. We chose instead to clarify the ancestry of these breeds using structural genomic analysis taking into account the dynamic historical colonization of the region.

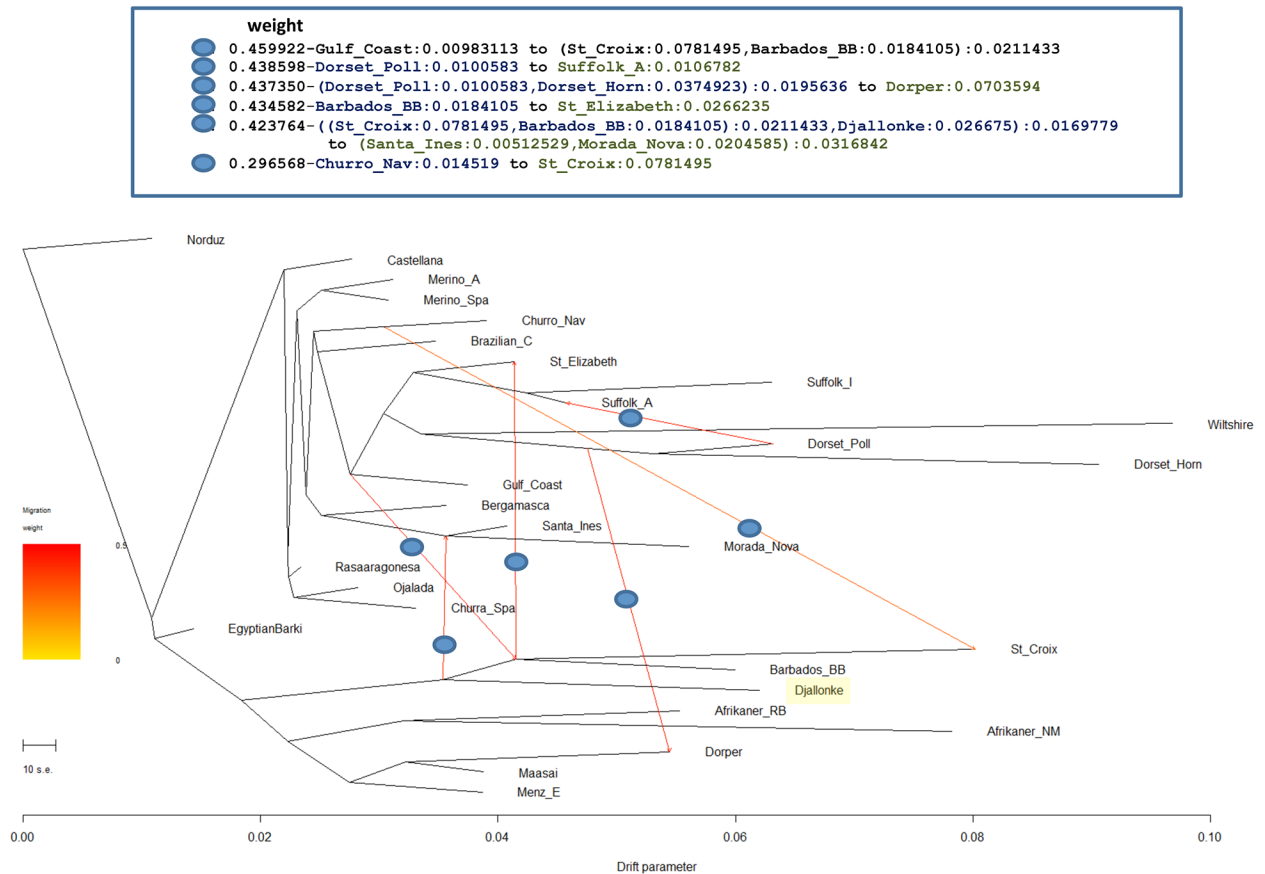


Fig 3. Tree topology with six migration events inferred by Treemix. Drift parameter is shown at the bottom and weights of inferred gene flow shown at the top in modified Newick format.

<https://doi.org/10.1371/journal.pone.0179021.g003>

Spanish Churra arrived in the Caribbean at the beginning of the 16th century with the second voyage of Columbus [47] (Fig 4). It is a hardy, coursewooled, thin-tailed breed accustomed to long cold winters and hot dry summers. It was developed in Spain along with its close cousin the Merino during the Arab Agricultural revolution from the 8th to the 13th centuries [16,48]. However, due to the extreme value of its wool at the time, Merinos were banned from exportation during the early stages of colonization.

The Churra were spread ubiquitously throughout the Caribbean, co-migrating with explorers and settlers (Fig 4, Orange arrows). Livestock, including sheep, were often left to fend for themselves as a food source for subsequent expeditions. Colonization towards mainland America followed shortly after, spreading north and south from the Yucatan peninsula. Northward expansion into present day southwestern U.S. was taken up by the indigenous peoples and resulted in the Navajo Churro. Northeastern migration around the Gulf of Mexico met with the southwestern expansion of the British colonies. Southward migration from Spanish Mexico proceeded down the Pacific coast of South America with a retrograde movement into southern Brazil.

There is virtually no difference in our structural analysis between the Navajo Churro and its accepted ancestor the Spanish Churra. Its principal components oriented indistinguishably from the other breeds in the Iberian group under all conditions. It also clustered identically with the Spanish Churra in admixture analyses. Nearest neighbor trees placed it in the same

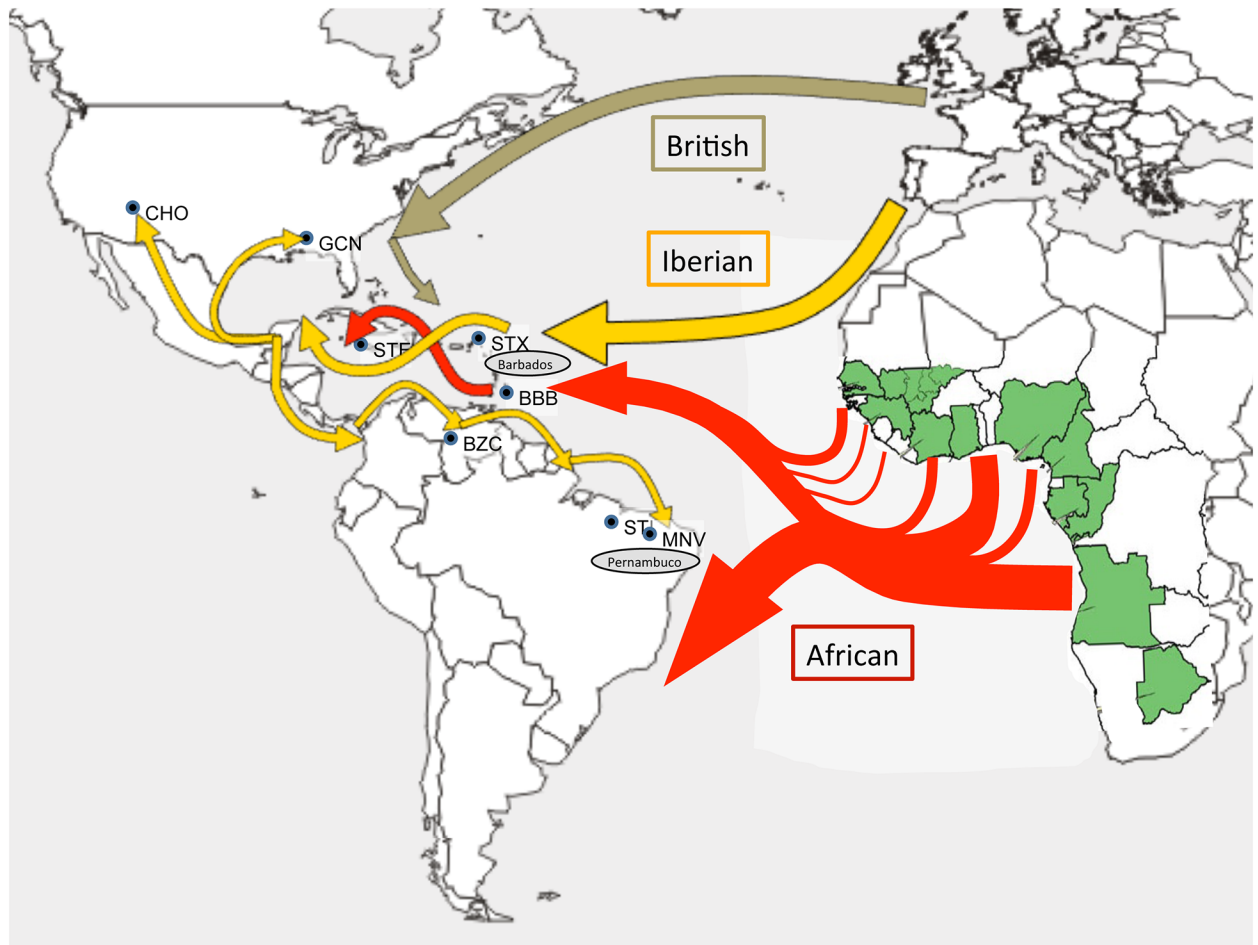


Fig 4. Historical and geographical corroboration of the results of whole genome structural analysis. Map indicates migration routes of British (olive) and Iberian (Orange) colonization and locations of New World and hair sheep breeds used in this study; Navajo Churro (CHO), Gulf Coast native (GCN), St. Elizabeth (STE), St. Croix (STX), Barbados Blackbelly (BBB), Brazilian Creole (BZC), Morada Nova (MNV), Santa Ines (STI). Red arrows indicate routes and relative magnitude of the Trans-Atlantic Slave Trade taken from Galloway. Gray ovals (Barbados and Pernambuco) indicate regions of heavy sugar production and centers of slave importation. Green highlighted African nations indicate range and distribution of the West African Djallonké sheep (FAO).

<https://doi.org/10.1371/journal.pone.0179021.g004>

clade with the Iberian sheep showing only slight drift. This is remarkable considering a difference of 500 years between the two breeds.

Our results agree with the accepted belief that the Brazilian Creole is a product of the retrograde movement along the Pacific and Northeastern coasts of South America[49,50]. The timing and pattern of its migration never brought it into contact with the introgression of West African sheep into the region. The Gulf Coast Native reflects a similar pattern; populating the North American coast of the Gulf of Mexico it would have met with the additional influence of the British. Our PCA shows it leaning in the direction of the British breeds. Although our scaled down Bayesian analysis indicated no British admixture (Fig 2), more comprehensive analyses do show an approximate 10% British component[18]. British breeds evolved and migrated along different lines than Mediterranean sheep[36,51,52]. Our nearest-neighbor tree shows the British branch joining at the junction of the Gulf Coast Native and St. Elizabeth indicating these breeds as the closest link between the British and Iberian groups in this study.

Our genomic analysis clearly indicates a major influence of West African lineage on Caribbean hair sheep. This is strongly supported by the pattern of colonization, economic structure, and trade routes of the period. The introduction of West African sheep into Caribbean populations probably began approximately 100 years after the arrival of the Iberian sheep as the search for silver and gold gave way to sugar production. The rise of the sugar industry created a desperate demand for labor. Historical as well as genomic data shows that this demand was filled by the enslavement of peoples from the West Coast of Africa [53]. The amount of sugar produced in a region roughly correlates to the number of slaves imported to that region (Fig 4 red arrows adapted from [54]). The nature of the Triangle Trade and the reliance on the trade winds prescribed the region of Central West Africa known historically as The Slave Coast. It roughly stretches from the modern day countries of Senegal to Angola. This is the exclusive range of the West African Dwarf or Djallonké breeds of sheep (Fig 4 green areas. Map adapted from [55]).

Main centers of importation in the Caribbean were the Northeast coast of Brazil and the Eastern Islands of the Caribbean, especially those ports connected to centers of sugar production such as Parambuco and Bahiain Northeastern Brazil, and the islands of Barbados and Jamaica in the British controlled Caribbean (Fig 4 gray ovals) [54]. It is at these junctions where the influx of West African sheep carried coincidentally along with the slave trade met and mated with the well established descendants of the Iberian Churra. The Morada Nova, Santa Inés, St. Croix, and Barbados Blackbelly are breeds named after the locations of their development. These correspond exactly to regions of highest slave importation and by correlation, West African sheep.

All of our Principal Component analyses indicate a link to the West African Djallonké. Outlier removal places the West African squarely in the quadrant exclusively occupied by the Caribbean hair sheep. Admixture analysis from $K = 6$ to 10 shows each to be a clear product of Iberian and West African ancestry. Although, these breeds differentiate above $K = 10$ it has been shown previously that theoretical population values of 10 or less are best indicators of ancestry in this type of analysis [25]. Treemix places the St. Croix and Barbados Blackbelly in the same clade as the Djallonké. Although, Morada Nova, and Santa Inés are in line with the Spanish Churra they show a strong indication of gene flow from the West African clade.

Caribbean hair sheep are a result of the introgression of West African traits into the ancestors of the Spanish Churra deposited by Iberian colonists from the time of Columbus. Further differentiation among the Hair sheep breeds may be a result of the influence of British breeds on the Northern Caribbean populations, the degree of husbandry, or simple climate-mediated adaptation. Caribbean hair sheep with West African heritage are ecotypes that combine the benefits of hair coat and disease resistance with economic viability. These locally adapted breeds represent reservoirs of adaptive fitness traits that can contribute to the future of sustainable sheep farming in the face of economic change, global warming, and the increasing pressure of drug-resistant parasites.

Supporting information

S1 Table. Table of f3 and f4 tests for validation of Treemix gene migrations.
(XLSX)

Acknowledgments

The authors would like to thank the International Sheep Genomic Consortium for providing the genotype data to make this study possible. We would also like to thank Harvey Blackburn

of the Center for Agricultural Resources Research / Plant and Animal Genetic Resources Preservation for providing the St. Croix genotypes, Andrea Talenti, Paolo Ajmone Marsan, and Paola Crepaldi of University of Milan, Dip. Medicina Veterinaria for the Bergamesca genotypes, and Magda Benevides for processing the Djallonké samples.

Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the US Department of Agriculture. The USDA is an equal opportunity provider and employer.

Author Contributions

Conceptualization: Gordon L. Spangler.

Data curation: Gordon L. Spangler, Benjamin D. Rosen.

Formal analysis: Gordon L. Spangler, Benjamin D. Rosen, Eui-Soo Kim.

Funding acquisition: Tad S. Sonstegard, Curtis P. Van Tassell.

Investigation: Gordon L. Spangler, Benjamin D. Rosen, Moses Babatunde Ilori, Olivier Hanotte, Joan M. Burke, James L. M. Morgan, David R. Notter.

Methodology: Gordon L. Spangler, Benjamin D. Rosen, Tad S. Sonstegard, Curtis P. Van Tassell.

Project administration: Curtis P. Van Tassell.

Resources: Moses Babatunde Ilori, Olivier Hanotte, Curtis P. Van Tassell.

Supervision: Gordon L. Spangler, Curtis P. Van Tassell.

Validation: Gordon L. Spangler, Benjamin D. Rosen, Curtis P. Van Tassell.

Visualization: Gordon L. Spangler, Benjamin D. Rosen.

Writing – original draft: Gordon L. Spangler, Benjamin D. Rosen.

Writing – review & editing: Gordon L. Spangler, Benjamin D. Rosen, Curtis P. Van Tassell.

References

1. Benavides MV, Sonstegard TS, Van Tassell C. Genomic Regions Associated with Sheep Resistance to Gastrointestinal Nematodes. *Trends Parasitol.* 2016 Jun; 32[6]:470–80. <https://doi.org/10.1016/j.pt.2016.03.007> PMID: 27183838
2. Jones K. Marketing Could Boost the U.S. Sheep Industry [Internet]. [cited 2016 Sep 20]. <http://www.ers.usda.gov/amber-waves/2004-february/marketing-could-boost-the-us-sheep-industry.aspx#.V-FJITWUKYw>
3. Álvarez I, Capote J, Traoré A, Fonseca N, Pérez K, Cuervo M, et al. Mitochondrial analysis sheds light on the origin of hair sheep. *Anim Genet.* 2013 Jun 1; 44[3]:344–7. <https://doi.org/10.1111/j.1365-2052.2012.02398.x> PMID: 23020288
4. Salle G, Jacquiet P, Gruner L, Cortet J, Sauve C, Prevot F, et al. A genome scan for QTL affecting resistance to *Haemonchus contortus* in sheep. *J Anim Sci.* 2012 Dec 1; 90[13]:4690–705. <https://doi.org/10.2527/jas.2012-5121> PMID: 22767094
5. Amarante AFT, Bricarello PA, Rocha RA, Gennari SM. Resistance of Santa Ines, Suffolk and Ile de France sheep to naturally acquired gastrointestinal nematode infections. *Vet Parasitol.* 2004 Feb 26; 120[1–2]:91–106. <https://doi.org/10.1016/j.vetpar.2003.12.004> PMID: 15019147
6. Aumont G, Gruner L, Hostache G. Comparison of the resistance to sympatric and allopatric isolates of *Haemonchus contortus* of Black Belly sheep in Guadeloupe [FWI] and of INRA 401 sheep in France. *Vet Parasitol.* 2003 Oct 8; 116[2]:139–50. PMID: 14519318
7. Notter DR, Andrew SA, Zajac AM. Responses of hair and wool sheep to a single fixed dose of infective larvae of *Haemonchus contortus*. *Small Rumin Res.* 2003 Mar; 47[3]:221–5.

8. Muigai AWT, Hirbo J, Sharkey S, Rege JEO, Blackburn H, Hanotte O. Genetic Diversity and relationships of hair sheep breeds of the Americas: First Results.
9. Morris ST. Economics of sheep production. *Small Rumin Res.* 2009 Oct; 86[1–3]:59–62.
10. Zygyiannis D. Sheep production in the world and in Greece. *Small Rumin Res.* 2006 Mar; 62[1–2]:143–7.
11. Milne C. The history of the Dorper sheep. *Small Rumin Res.* 2000 May; 36[2]:99–102. PMID: [10760446](#)
12. Burke JM, Miller JE. Relative resistance of Dorper crossbred ewes to gastrointestinal nematode infection compared with St. Croix and Katahdin ewes in the southeastern United States. *Vet Parasitol.* 2002 Nov 11; 109[3–4]:265–75. PMID: [12423938](#)
13. Malhado CHM, Carneiro PLS, Affonso PRAM, Souza AAO, Sarmento JLR. Growth curves in Dorper sheep crossed with the local Brazilian breeds, Morada Nova, Rabo Largo, and Santa Inês. *Small Rumin Res.* 2009 Jun; 84[1–3]:16–21.
14. McManus C, Paiva SR, Araújo de RO. Genetics and breeding of sheep in Brazil. *Rev Bras Zootec.* 2010 Jul; 39:236–46.
15. Paiva SR, Mariante A da S, Blackburn HD. Combining US and Brazilian Microsatellite Data for a Meta-Analysis of Sheep [*Ovis aries*] Breed Diversity: Facilitating the FAO Global Plan of Action for Conserving Animal Genetic Resources. *J Hered.* 2011 Nov 1; 102[6]:697–704. <https://doi.org/10.1093/jhered/esr101> PMID: [22013019](#)
16. Pereira F, Davis SJM, Pereira L, McEvoy B, Bradley DG, Amorim A. Genetic Signatures of a Mediterranean Influence in Iberian Peninsula Sheep Husbandry. *Mol Biol Evol.* 2006 Jul 1; 23[7]:1420–6. <https://doi.org/10.1093/molbev/msl007> PMID: [16672283](#)
17. Jiang Y, Xie M, Chen W, Talbot R, Maddox JF, Faraut T, et al. The sheep genome illuminates biology of the rumen and lipid metabolism. *Science.* 2014 Jun 6; 344[6188]:1168–73. <https://doi.org/10.1126/science.1252806> PMID: [24904168](#)
18. Kijas JW, Lenstra JA, Hayes B, Boitard S, Neto LRP, Cristobal MS, et al. Genome-Wide Analysis of the World's Sheep Breeds Reveals High Levels of Historic Mixture and Strong Recent Selection. *PLOS Biol.* 2012 Feb 7; 10[2]:e1001258. <https://doi.org/10.1371/journal.pbio.1001258> PMID: [22346734](#)
19. Lv F-H, Agha S, Kantanen J, Colli L, Stucki S, Kijas JW, et al. Adaptations to Climate-Mediated Selective Pressures in Sheep. *Mol Biol Evol.* 2014 Sep 23;msu264.
20. Yang J, Li W-R, Lv F-H, He S-G, Tian S-L, Peng W-F, et al. Whole-Genome Sequencing of Native Sheep Provides Insights into Rapid Adaptations to Extreme Environments. *Mol Biol Evol.* 2016 Oct 1; 33[10]:2576–92. <https://doi.org/10.1093/molbev/msw129> PMID: [27401233](#)
21. Ciani E, Crepaldi P, Nicoloso L, Lasagna E, Sarti FM, Moiola B, et al. Genome-wide analysis of Italian sheep diversity reveals a strong geographic pattern and cryptic relationships between breeds. *Anim Genet.* 2014 Apr 1; 45[2]:256–66. <https://doi.org/10.1111/age.12106> PMID: [24303943](#)
22. Legarra A, Croiseau P, Sanchez MP, Teyssède S, Sallé G, Allais S, et al. A comparison of methods for whole-genome QTL mapping using dense markers in four livestock species. *Genet Sel Evol.* 2015; 47:6. <https://doi.org/10.1186/s12711-015-0087-7> PMID: [25885597](#)
23. Matika O, Riggio V, Anselme-Moizan M, Law AS, Pong-Wong R, Archibald AL, et al. Genome-wide association reveals QTL for growth, bone and in vivo carcass traits as assessed by computed tomography in Scottish Blackface lambs. *Genet Sel Evol.* 2016; 48:11. <https://doi.org/10.1186/s12711-016-0191-3> PMID: [26856324](#)
24. Kijas JW, Hadfield T, Naval Sanchez M, Cockett N. Genome-wide association reveals the locus responsible for four-horned ruminant. *Anim Genet.* 2016 Apr 1; 47[2]:258–62. <https://doi.org/10.1111/age.12409> PMID: [26767438](#)
25. Kijas JW, Townley D, Dalrymple BP, Heaton MP, Maddox JF, McGrath A, et al. A Genome Wide Survey of SNP Variation Reveals the Genetic Structure of Sheep Breeds: e4668. *PLoS One* [Internet]. 2009 Mar [cited 2016 Jun 7]; 4[3]. Available from: <http://search.proquest.com/naturalsciencollection/docview/1291063031/abstract/8C3821F09A944B4APQ/1>
26. Fariello M-I, Servin B, Tosser-Klopp G, Rupp R, Moreno C, Consortium ISG, et al. Selection Signatures in Worldwide Sheep Populations. *PLOS ONE.* 2014 Aug 15; 9[8]:e103813. <https://doi.org/10.1371/journal.pone.0103813> PMID: [25126940](#)
27. Blackburn HD. Genetic structure and diversity among sheep breeds in the United States: Identification of the major gene pools—sheep microsats jas.pdf [Internet]. [cited 2017 Feb 21]. <https://www.ars.usda.gov/ARSUserFiles/30123005/animal/sheep%20microsats%20jas.pdf>
28. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet.* 2006 Aug; 38[8]:904–9. <https://doi.org/10.1038/ng1847> PMID: [16862161](#)

29. Raj A, Stephens M, Pritchard JK. fastSTRUCTURE: Variational Inference of Population Structure in Large SNP Datasets. *Genetics*. 2014 Jan 1;genetics.114.164350.
30. Lee T-H, Guo H, Wang X, Kim C, Paterson AH. SNPhylo: a pipeline to construct a phylogenetic tree from huge SNP data. *BMC Genomics* [Internet]. 2014 [cited 2016 Sep 15]; 15. Available from: <http://search.proquest.com/naturalscicollection/docview/1505023551/abstract/BC71B23800A94FCCPQ/1>
31. Pickrell JK, Pritchard JK. Inference of Population Splits and Mixtures from Genome-Wide Allele Frequency Data. *PLOS Genet*. 2012 Nov 15; 8[11]:e1002967. <https://doi.org/10.1371/journal.pgen.1002967> PMID: 23166502
32. Henzell T. Australian Agriculture: Its History and Challenges. Csiro Publishing; 2007. 321 p.
33. New Zealand Sheepbreeders' Association—Poll Dorset & Dorset Horn [Internet]. [cited 2016 Sep 30]. <http://www.nzsheep.co.nz/index.php?page=poll-dorset>
34. Paiva SR, Silvério VC, Egito AA, McManus C, Faria de DA, Mariante A da S, et al. Genetic variability of the Brazilian hair sheep breeds. *Pesqui Agropecuária Bras*. 2005 Sep; 40[9]:887–93.
35. Ciani E, Lasagna E, D'Andrea M, Alloggio I, Marroni F, Ceccobelli S, et al. Merino and Merino-derived sheep breeds: a genome-wide intercontinental study. *Genet Sel Evol*. 2015; 47:64. <https://doi.org/10.1186/s12711-015-0139-z> PMID: 26272467
36. Chessa B, Pereira F, Arnaud F, Amorim A, Goyache F, Mainland I, et al. Revealing the History of Sheep Domestication Using Retrovirus Integrations. *Science*. 2009 Apr 24; 324[5926]:532–6. <https://doi.org/10.1126/science.1170587> PMID: 19390051
37. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*. 2015 Dec 1; 4[1]:1–16.
38. Browning BL, Browning SR. A Fast, Powerful Method for Detecting Identity by Descent. *Am J Hum Genet*. 2011 Feb 11; 88[2]:173–82. <https://doi.org/10.1016/j.ajhg.2011.01.010> PMID: 21310274
39. Herráez DL, Bauchet M, Tang K, Theunert C, Pugach I, Li J, et al. Genetic Variation and Recent Positive Selection in Worldwide Human Populations: Evidence from Nearly 1 Million SNPs: e7888. *PLoS One* [Internet]. 2009 Nov [cited 2016 Sep 7]; 4[11]. Available from: <http://search.proquest.com/naturalscicollection/docview/1292305704/abstract/508D04B763924412PQ/1>
40. Gifford-Gonzalez D, Hanotte O. Domesticating Animals in Africa: Implications of Genetic and Archaeological Findings. *J World Prehistory*. 2011 Feb 6; 24[1]:1–23.
41. Laloë D, Moazami-Goudarzi K, Lenstra JA, Marsan PA, Azor P, Baumung R, et al. Spatial Trends of Genetic Variation of Domestic Ruminants in Europe. *Diversity*. 2010 Jun 17; 2[6]:932–45.
42. Muigai AWT, Hanotte O. The Origin of African Sheep: Archaeological and Genetic Perspectives. *Afr Archaeol Rev*. 2013 Mar 6; 30[1]:39–50.
43. Álvarez I, Capote J, Traoré A, Fonseca N, Pérez K, Cuervo M, et al. Mitochondrial analysis sheds light on the origin of hair sheep. *Anim Genet*. 2013 Jun 1; 44[3]:344–7. <https://doi.org/10.1111/j.1365-2052.2012.02398.x> PMID: 23020288
44. Álvarez I, Capote J, Traoré A, Fonseca N, Pérez K, Cuervo M, et al. Genetic relationships of the Cuban hair sheep inferred from microsatellite polymorphism. *Small Rumin Res*. 2012 May; 104[1–3]:89–93.
45. Muigai AWT, Hirbo J, Sharkey S, Rege JEO, Blackburn H, Hanotte O. Genetic Diversity and relationships of hair sheep breeds of the Americas: First Results.
46. Delgado JV, Perezgrovas R, Camacho ME, Fresno M, Barba C. The Wool-Less Canary Sheep and their relationship with the present breeds in America. *Anim Genet Resour Génétiques Anim Génétiques Anim*. 2000 Apr; 28:27–34.
47. Major RH. Select Letters of Christopher Columbus, with Other Original Documents, Relating to His Four Voyages to the New World. Hakluyt Society; 1870. 428 p.
48. Watson AM. The Arab Agricultural Revolution and Its Diffusion, 700–1100. *J Econ Hist*. 1974 Mar; 34 [01]:8–35.
49. McManus C, Paiva SR, Araújo de RO. Genetics and breeding of sheep in Brazil. *Rev Bras Zootec*. 2010 Jul; 39:236–46.
50. Mariante A da S, Albuquerque M do SM, do Egito AA, McManus C. Advances in the Brazilian animal genetic resources conservation programme. *Anim Genet Resour Resour Génétiques Anim Recur Génétiques Anim*. 1999 Apr; 25:107–121.
51. Lv F-H, Peng W-F, Yang J, Zhao Y-X, Li W-R, Liu M-J, et al. Mitogenomic Meta-Analysis Identifies Two Phases of Migration in the History of Eastern Eurasian Sheep. *Mol Biol Evol*. 2015 Oct 1; 32[10]:2515–33. <https://doi.org/10.1093/molbev/msv139> PMID: 26085518
52. Tapio M, Marzanov N, Ozerov M, Činkulov M, Gonzarenko G, Kiselyova T, et al. Sheep Mitochondrial DNA Variation in European, Caucasian, and Central Asian Areas. *Mol Biol Evol*. 2006 Sep 1; 23 [9]:1776–83. <https://doi.org/10.1093/molbev/msl043> PMID: 16782761

53. Bryc. Genome-wide patterns of population structure and admixture in West Africans and African Americans [Internet]. [cited 2016 Jun 7]. <http://www.pnas.org/content/107/2/786.full?sid=5fe0c125-049d-446b-8371-4522b23f25a2>
54. Galloway JH. The Sugar Cane Industry: An Historical Geography from Its Origins to 1914. Cambridge University Press; 2005. 288 p.
55. Djallonke [Internet]. [cited 2016 Sep 20]. http://agtr.ilri.cgiar.org/index.php?option=com_content&view=article&id=134&Itemid=148