



DNA Tribes® Digest September 2, 2014

Copyright © 2014 DNA Tribes®. All rights reserved.

To request an email subscription to DNA Tribes® Digest, email digest@dnatribes.com with the subject heading "Subscribe." To unsubscribe from DNA Tribes® Digest, email digest@dnatribes.com with the subject heading "Unsubscribe." Previous issues of DNA Tribes® Digest are available online at <http://dnatribes.com/library.html>.

Table of Contents:

Introduction.....	1
A Genetic Component Linking Two Worlds: First Farmer (EEF) Ancestry in the Middle East and Europe	2
Background.....	2
Non-Local Components in the Aegean and West Asia (STR)	3
Non-Local Genetic Components in the Aegean and West Asia (SNP)	5
Conclusion	7
<i>DNA Tribes®</i> SNP Personal Genome Analysis: Update for Summer 2014	8

Introduction

Hello, and welcome to the September 2014 issue of DNA Tribes® Digest. This month's article explores the **Early European Farmer (EEF) ancestry that links the Middle East and Europe**, by examining non-local components shared by populations in both Southeastern Europe and West Asia. This might provide clues about the origin of the First Farmers: whether they originated from Europe, the Middle East, or **instead emerged from a more complex Neolithic mixture of European, African, and Asian populations**.

Best regards,
Lucas Martin
DNA Tribes

DNA Tribes® is on Facebook.
Find us at <http://facebook.com/DNATribes>

A Genetic Component Linking Two Worlds: First Farmer (EEF) Ancestry in the Middle East and Europe

Background

Recent landmark studies have unearthed important clues about the peopling of Eurasia, in the form of several ancient genomes spanning the Paleolithic, Mesolithic, and Neolithic periods. These genomes include Paleolithic hunters from Siberia (Mal'ta-1 and Afontova Gora-2); Mesolithic hunter-gatherers from Western Europe (Ajvide, La Brana, Loschbour, Motala, and Stora Förfvar); Neolithic farmers (Stuttgart and Gökhem), and a Chalcolithic Alpine mountaineer ("Ötzi" the Iceman).¹

Genetic analysis has identified all of these ancient genomes as belonging to populations that helped shape the genetic structure of Eurasia and the Americas, in ways that geneticists are just beginning to document (illustrated in **Figure 1**).² **In particular, Early European Farmers (EEF) from Neolithic Southern Germany (Stuttgart) and the Baltic Sea (Gökhem) have been shown to be ancestral to both Middle Eastern and European populations.**³

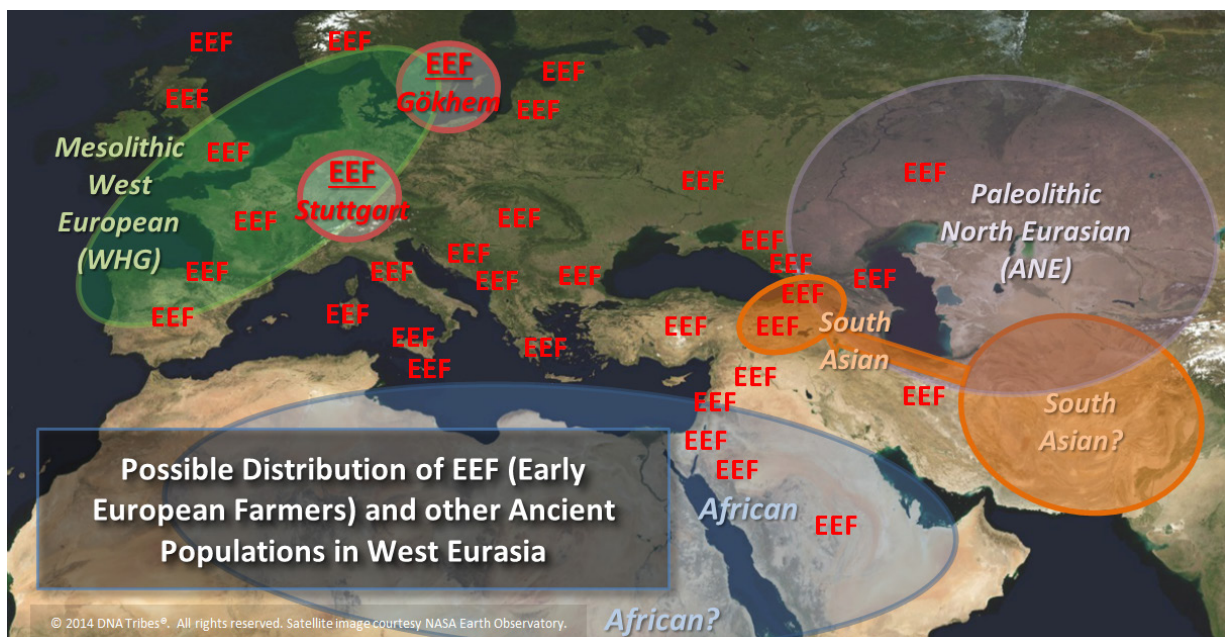


Figure 1: Possible geographic distribution of Early European Farmers (EEF) and other ancient populations in West Eurasia (Europe and the Middle East), based in part on Lazaridis *et al.*⁴

¹ See "Ancient Human Genomes Suggest Three Ancestral Populations for Present-day Europeans" by Lazaridis *et al.* (2013), available at <http://biorxiv.org/content/early/2014/04/05/001552> and "Genomic Diversity and Admixture Differs for Stone-Age Scandinavian Foragers and Farmers" by Skoglund *et al.* at <http://www.sciencemag.org/content/344/6185/747.abstract>.

² For more discussion, see: <http://dnatribes.com/dnatribes-digest-2014-02-01.pdf>; <http://dnatribes.com/dnatribes-digest-2014-03-01.pdf>; and <http://dnatribes.com/dnatribes-digest-2014-04-02.pdf>.

³ See <http://biorxiv.org/content/early/2014/04/05/001552>, Table 1 (p. 16).

⁴ *Ibid.*

Modern populations shown to be descended from **Early European Farmers (EEF)** in Lazaridis *et. al.* (2013) include:

- | | | |
|------------------|--------------|--------------------|
| • Adygei | • Greek | • Moroccan Jewish |
| • Albanian | • Hungarian | • North Ossetian |
| • Armenian | • Iranian | • Palestinian |
| • Balkar | • Jordanian | • Sephardic Jewish |
| • Bedouin | • Kumyk | • Sicilian |
| • Bergamo, Italy | • Lebanese | • Syrian |
| • Bulgarian | • Lezgin | • Turkish |
| • Croatian | • Libyan Jew | • Tuscany, Italy |
| • French | • Maltese | |

This remarkable connection linking Neolithic farmers (EEF) with modern populations living throughout the Caucasus Mountains, the Middle East, Mediterranean, and Northern Europe suggests that EEF populations were part of a larger Eurasian expansion. To fully understand how EEF populations shaped the genetic structure of a large part of West Eurasia will eventually require ancient DNA samples from throughout the Middle East and Caucasus Mountains.

At present, a key question is where these EEF populations originated: whether in **North and Central Europe** (where the only ancient EEF genomes known to date were found), **the Middle East** (where no ancient DNA has been recovered to date), or an **intermediate location** (such as the Mediterranean Sea or Caucasus Mountains).

Past DNA Tribes® Digest articles have identified underlying **Aegean or Greek-like genetic components** throughout Europe and the Middle East, perhaps related to **Ubaid** or **Hassuna-Samara** cultures that appeared around **the East Mediterranean (Levant), in both the Fertile Crescent and Aegean** between 6,700-6,500 BCE.^{5,6} However, the origins of these first “**Old European**” (Balkan Neolithic) and **Ubaidian** (Fertile Crescent proto-Sumerians) city-founders remain obscure and debated by scholars.⁷ In even earlier periods, the Levant had been a contact point between West Asian Hunter-Gatherer populations and Natufian-PPNA cultures during the **PPNB (Pre-Pottery Neolithic B)** period.

Nevertheless, the **geographic location of the Levant as a nexus between Europe, West Asia, and Africa** suggests that ancient populations from one or more of these continents might have been involved in the emergence and expansion of EEF peoples. To explore these underlying connections and search for clues to EEF origins in Europe and/or Middle East, this month’s Digest analysis will use autosomal STR and SNP analysis to examine non-local genetic components in the Aegean and West Asia.

Non-Local Components in the Aegean and West Asia (STR)

Genetic Analysis: Non-local genetic components of genetic regions in the Aegean, Mesopotamian, and Arabian regions were identified using autosomal STR data. Results are listed in **Table 1** and illustrated in **Figure 2**.

⁵ See <http://dnatribes.com/dnatribes-digest-2013-11-01.pdf>.

⁶ See <http://dnatribes.com/dnatribes-digest-2013-08-01.pdf>.

⁷ See <http://dnatribes.com/dnatribes-digest-2014-03-01.pdf>.

STR Region	Horn of Africa	North African	Basque	Belgic	Balkan	Indus Valley	Other
Aegean	0.6%	11.5%	1.7%	30.8%	47.5%	6.9%	0.9%
Arabian	26.1%	30.3%	0.0%	0.0%	16.0%	26.1%	1.5%
Levantine	6.4%	36.5%	5.6%	0.0%	29.4%	13.7%	8.4%
Mesopotamian	1.7%	31.3%	0.0%	0.0%	32.1%	29.9%	5.0%

Table 1: Non-local genetic components in the Aegean and West Asia based on autosomal STR data.

Discussion: Results in **Table 1** and **Figure 2** express several genetic components in the Aegean and West Asia. **Components found in all studied regions include: North African, Balkan, and Indus Valley percentages.** In contrast, components found in only some studied regions include: Horn of Africa, Basque, and Belgic.

North African percentages are largest in the Levantine region (36.5%) and are smallest in the Aegean region (11.5%). This genetic component, shared by all studied regions, might to some extent reflect North African related contact with the Fertile Crescent, possibly dating to the integration of PPNB populations with (possibly African related) Harifian hunters in the **Circum-Arabian Pastoral Complex**.⁸

Balkan percentages are largest in the Aegean region (47.5%) and smallest in the Arabian region (16.0%). The distribution of these components deep into the Middle East might reflect contacts between early Ubaidians and other Fertile Crescent city founders with Balkan Neolithic populations, **possibly including direct gene flow from Europe** (perhaps following Kurgan related steppe incursions in Europe, displacing some “Old European” populations to the East Mediterranean).⁹ This point is particularly notable, since the known ancient EEF genomes were discovered not in the Middle East, but in Central and Northern Europe.

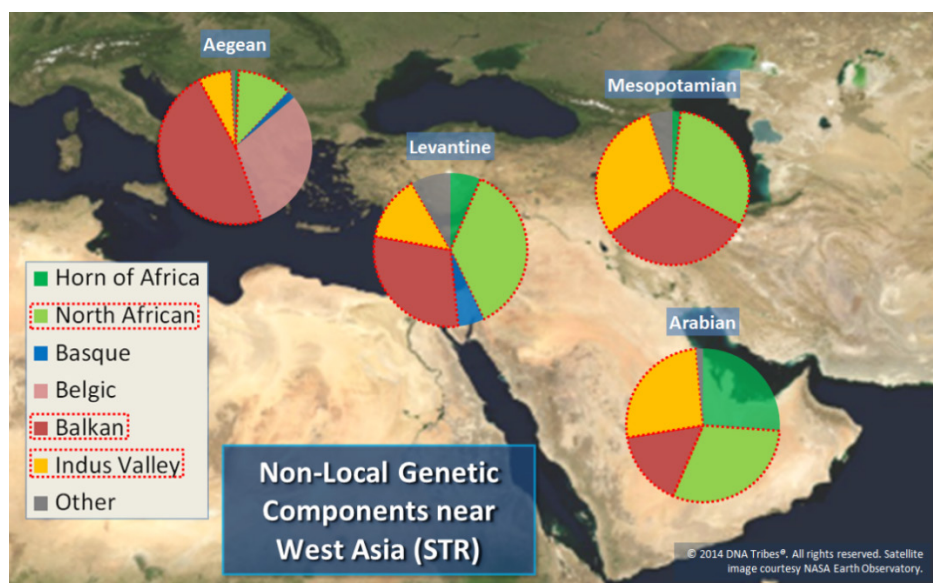


Figure 2: Map of non-local genetic components in the Aegean and West Asia based on autosomal STR data. Components found in all regions are outlined in **red**.

⁸ For more discussion, see <http://dnatribes.com/dnatribes-digest-2014-03-01.pdf>, p. 2.

⁹ See <http://dnatribes.com/dnatribes-digest-2013-11-01.pdf>.

Indus Valley components are largest in the Mesopotamian (29.9%) region and smallest in the Aegean (6.9%) region. This shared component suggests that early **Central Asians or South Asians might have interacted with EEF farmers**, perhaps near the Caucasus Mountains and Iranian Plateau. In particular, this might reflect West Asian interactions involved in **Secondary Uruk Expansions** and the emergence of the **Indus Valley Civilization** during the second millennium BCE.¹⁰

Because these components are found in all studied Aegean and West Asian regions, they might to some degree reflect traces of EEF populations ancestral to both Europeans and West Asians.

Non-Local Genetic Components in the Aegean and West Asia (SNP)

Genetic Analysis: Non-local genetic components of East Mediterranean and West Asian clusters identified by DNA Tribes® SNP Personal Genome Analysis¹¹ were identified using autosomal SNP data. Results are listed in **Table 2** and illustrated in **Figure 3**.

SNP Cluster	Horn of Africa	Berber-North African	Basque	Spanish-Portuguese	NW Europe	Central European	Balto-Slavic	Tatar-Chuvash	Central Asian	Burusho-Punjab-Sindhi	Kalash-Balochi
Aegean-Balkan-Italian	0.0	11.6	8.5	21.3	22.1	18.4	8.8	0.0	3.9	0.0	5.4
Ashkenazi Jewish	0.0	21.0	7.3	21.5	15.4	12.7	4.8	0.0	6.6	0.0	10.8
Azeri-Kurdish-Persian-Turkey	0.0	15.4	3.7	17.8	6.1	9.7	0.0	0.0	14.6	11.7	21.0
Bedouin-Gulf Arabs	9.1	40.2	5.5	17.4	1.5	5.2	0.0	0.0	4.7	0.3	16.0
Daghestan-North Caucasus	0.0	2.2	0.4	10.2	17.2	14.1	7.7	1.2	15.2	14.1	17.7
Lebanese-Cyprus	2.9	25.2	5.7	21.1	6.4	9.7	0.0	0.0	9.2	3.6	16.1
Mideast Jewish	0.0	20.3	5.0	18.9	4.2	8.9	0.0	0.0	13.0	8.7	21.0
Pontic-West Caucasus	0.0	8.0	2.3	15.7	10.8	13.2	3.1	0.0	18.9	11.2	16.7
Red Sea Yemen-Egypt	18.0	31.7	4.2	16.3	1.0	6.3	0.0	0.0	7.1	3.3	12.1
Sephardic-Sicilian	0.0	24.2	7.1	23.9	13.8	12.9	0.0	0.0	6.4	0.5	11.2

Table 2: Non-local genetic components (%) in the Aegean and West Asia based on autosomal SNP data.

Discussion: Results in **Table 2** and **Figure 3** express several genetic components in the Aegean and West Asia. **Components found in all studied regions include: Berber-North African, Spanish-Portuguese, Central European, Central Asian, and Kalash-Balochi percentages.** In contrast, components found in only some studied regions include: Horn of Africa, Basque, NW Europe, Balto-Slavic, Tatar-Chuvash, and Burusho-Punjabi-Sindhi.

As with STR components (discussed in the previous section of this article), **Berber-North African** percentages are expressed for all studied SNP clusters. Berber-North African percentages are

¹⁰ For more detailed discussion, see <http://dnatribes.com/dnatribes-digest-2013-08-01.pdf> and <http://dnatribes.com/dnatribes-digest-2013-10-01.pdf>.

¹¹ For information about DNA Tribes® SNP Personal Genome Analysis, see <http://dnatribes.com/snp.html>.

largest in Bedouin-Gulf Arabs, which (as discussed in the previous section) might reflect contacts between nomadic North African and Arabian cultures dating in part to the emergence of the **Circum-Arabian Pastoral Complex**. Further, these North African genetic components might more fundamentally reflect early contacts between **Natufians** and **PPNA-B** populations, possibly involved in First Farmer cultures that spread throughout Europe and the Middle East during the Neolithic period.

Spanish-Portuguese percentages are largest in the Sephardic-Sicilian cluster (23.9%), but are also found throughout West Asia, reaching a still substantial local minimum in the Daghestan-North Caucasus (10.2%). The widespread distribution of this cluster suggests widespread Mediterranean related contacts with the Middle East (**possibly originating with EEF populations and preserved in the Iberian Peninsula, but not necessarily originating in the West Mediterranean**).

Central European percentages in all studied populations provide additional evidence for widespread involvement between Europe and the Middle East, possibly dating to EEF contacts involving the **Balkan-Aegean Neolithic and related Ubaid cultures** of the Fertile Crescent discussed in the Background section of this article.

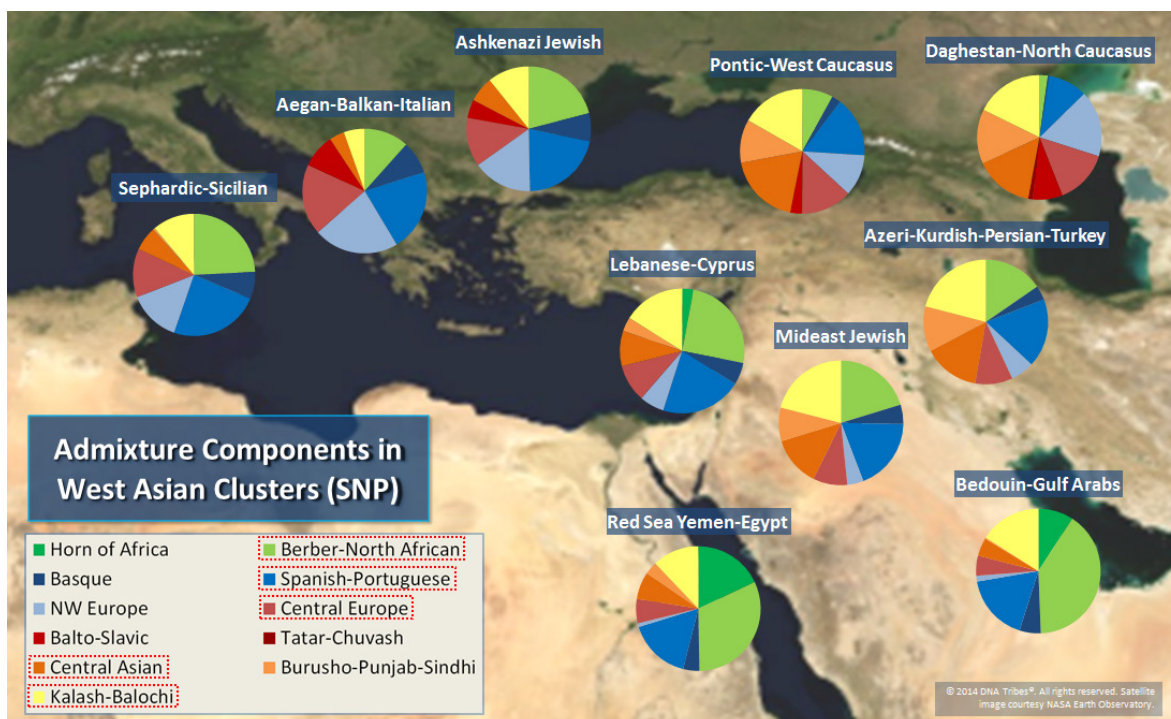


Figure 3: Map of non-local genetic components in the Aegean and West Asia based on autosomal SNP data. Components found in all regions are outlined in red.

Central Asian components (most representative of present day Turkmen and Tajik populations) are surprisingly also expressed in all studied regions. These are highest in the Pontic-West Caucasus (18.9%), Daghestan-North Caucasus (15.2%), and Azeri-Kurdish-Persian-Turkey (14.6%) clusters. This might reflect late Uruk expansions near Central Asia, possibly related to the formation and dispersion of the **Bactro-Margiana Archaeological Complex (BMAC)**.¹²

¹² For more discussion, see <http://dnatribes.com/dnatribes-digest-2012-05-01.pdf>.

Finally, **Kalash-Balochi** percentages are largest in Mideast Jewish and Azeri-Kurdish-Persian-Turkey (both 21.0%) clusters. This suggests a possible connection with the **Indo-Iranian languages** and the BMAC or Indus Valley Civilization.¹³ However, more data will be needed to explore this in more detail.

Conclusion

Both autosomal STR and SNP results express North African, European, and Central-South Asian related genetic components found throughout studied Aegean and West Asian populations, when local components are excluded. Some or all of these shared components might in part reflect genetic traces of the **EEF ancestry found in these modern populations**.

¹³ <http://dnatribes.com/dnatribes-digest-2012-09-01.pdf>, <http://dnatribes.com/dnatribes-digest-2013-10-01.pdf>.

DNA Tribes® SNP Personal Genome Analysis: Update for Summer 2014

We are pleased to announce a new update for DNA Tribes® SNP analysis. This includes:

- New Populations
- Updated Admixture Analysis
- Updated MDS Analysis
- Enhanced Population Admixture (*Best Fits for Your Genome Composition*)
- New Population Admixture (*Alternative Iterative Fits for Your Genome Composition*)
- Enhanced World Region Grid (*Best Fits for Your Genome Composition*)
- Enhanced Population Similarity (*Best Fits for Your Genome Composition*)
- Updated Admixture Tables

New Caucasus Mountains Populations:

- Azerbaijan
- Imer Georgia
- Kabardin
- Laz Georgia
- Tabasaran

New Diasporic Populations:

- Jewish Algeria
- Jewish France
- Jewish Italy
- Jewish Kurdistan
- Jewish Libya
- Jewish Syria
- Jewish Tunisia

New European Populations:

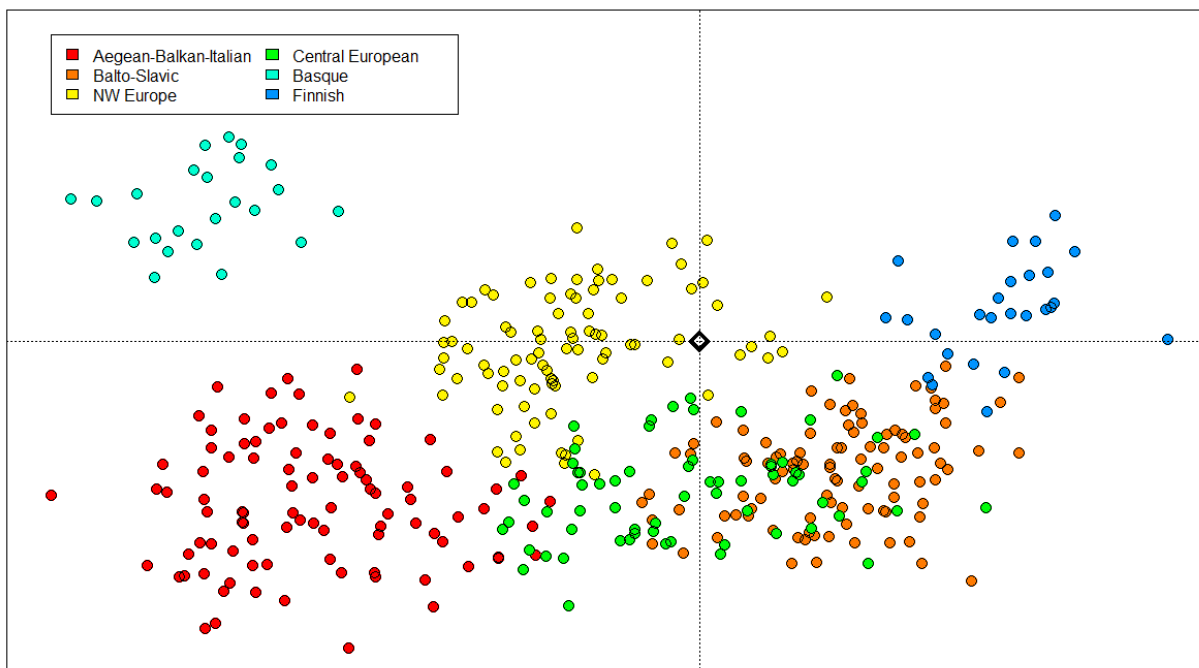
- Abruzzo, Italy
- Agrigento, Sicily
- Central Greece
- Central Sicily
- Croatia
- Kursk, Russia
- Moldavia
- Oryol, Russia
- Smolensk, Russia
- Sweden (*new samples*)
- Tatar Russian Federation
- Thessaly, Greece
- Voronezh, Russia
- Western Sicily

Updated Admixture Analysis based on DNA Tribes® proprietary analysis of world population structure, based on **five continental cores** and **44 detailed regional clusters**:

Continental Core	Percentage
Basque	75.3%
Dravidian South India	17.5%
Mesoamerican	6.6%
Aka-Mbuti-Hadza	0.6%
South China	0.0%

Regional Cluster	Percentage
NW Europe	44.2%
Balto-Slavic	29.8%
Central European	10.4%
Finnish	7.3%
Basque	5.7%
Aegean-Balkan-Italian	2.8%

Updated Multi-Dimensional Scaling (MDS) to visualize your genotype's position among continent cores and genetic clusters:

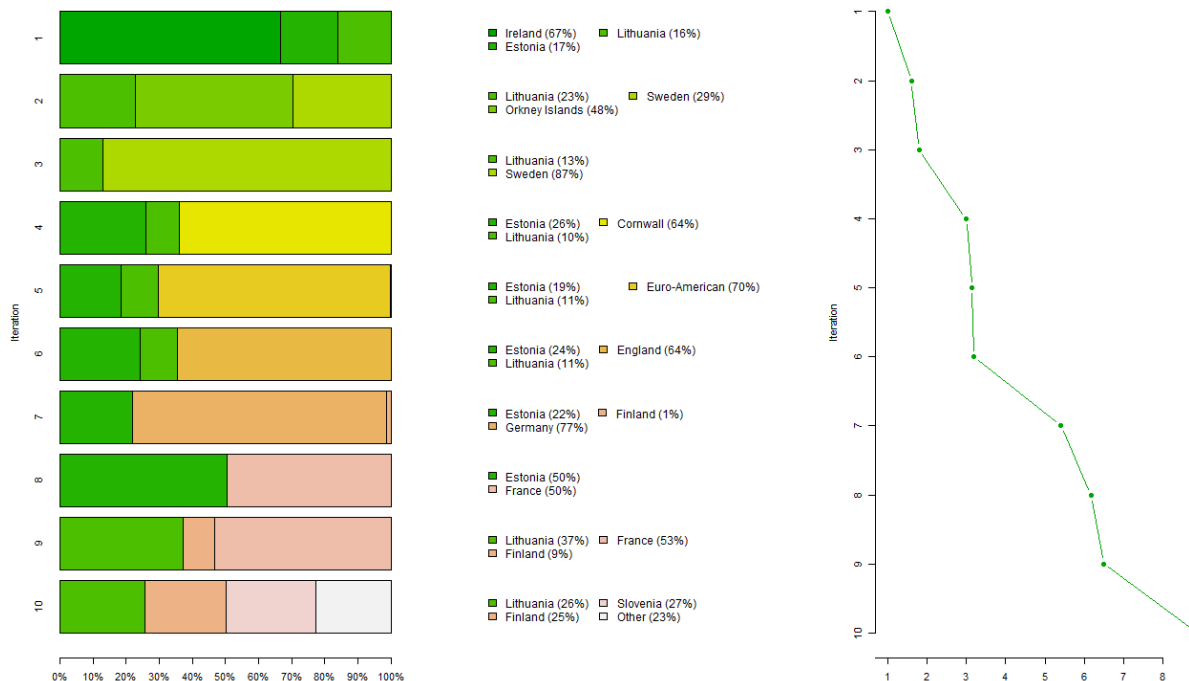


Enhanced Population Admixture (Best Fit for Your Genome Composition) based on the combination of populations that best fits your genome's ancestral composition:

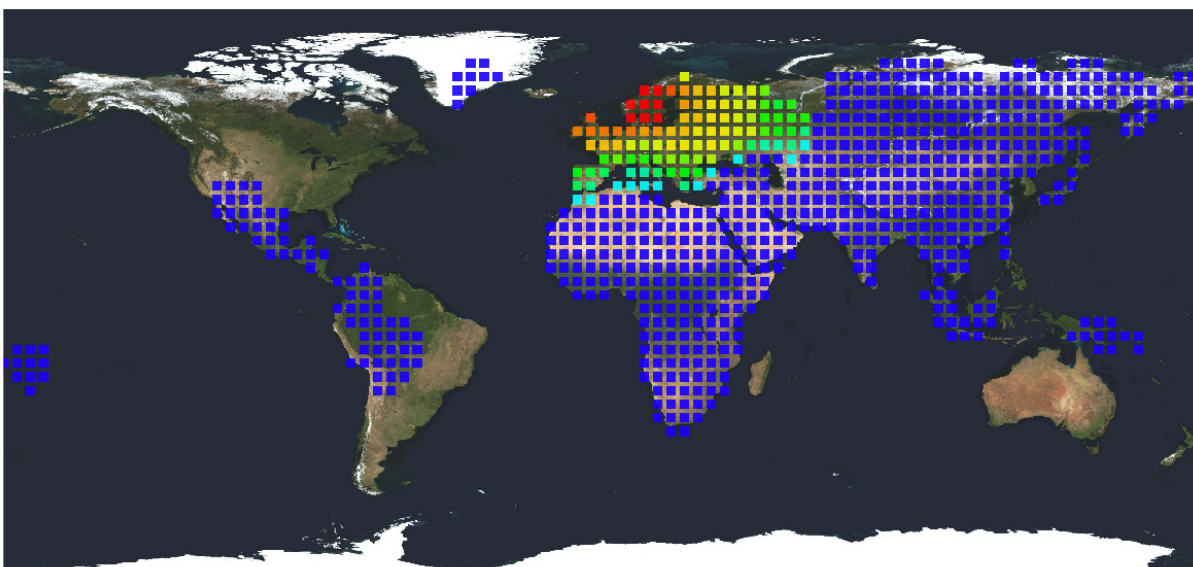
Population (Native Only)	Percentage
Ireland	66.5%
Estonia	17.5%
Lithuania	16.0%

Population (Native and Global)	Percentage
Ireland	66.5%
Estonia	17.5%
Lithuania	16.0%

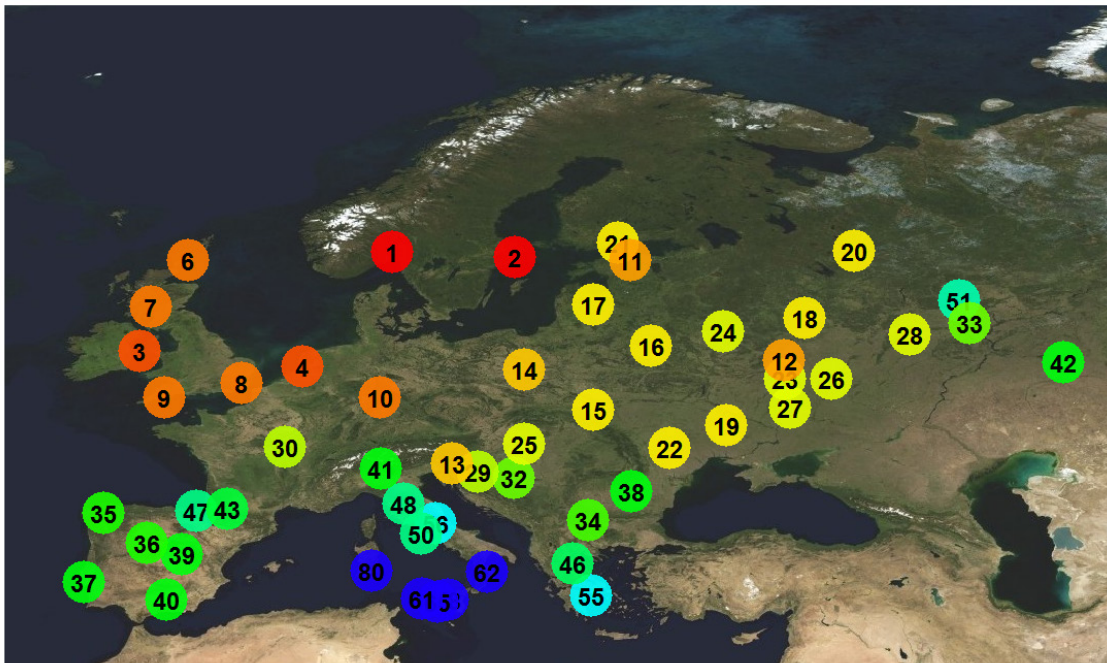
Enhanced Population Admixture (Alternative Iterative Fits) illustrating several alternative combinations of populations that fit your genome's ancestral composition:



Enhanced World Grid Genome Fit Map showing the geographical distribution of your combination of 44 ancestral components on a shaded world map:



Enhanced Population Genome Fit Rankings individually listing the countries and ethnic groups in our SNP database where your combination of 44 ancestral components fits best:



Updated World Admixture Tables (including average **five continent core** and **44 region cluster** percentages for samples populations in our SNP database) are available at <http://dnatribes.com/dnatribes-snp-admixture-2014-06-03.pdf>.

Update Your Personal SNP Analysis New SNP analysis orders (**Sale Price \$39.99** with submitted grandparent information form) and SNP updates (**\$19.99**) be ordered at <http://dnatribes.com/snp.html>.