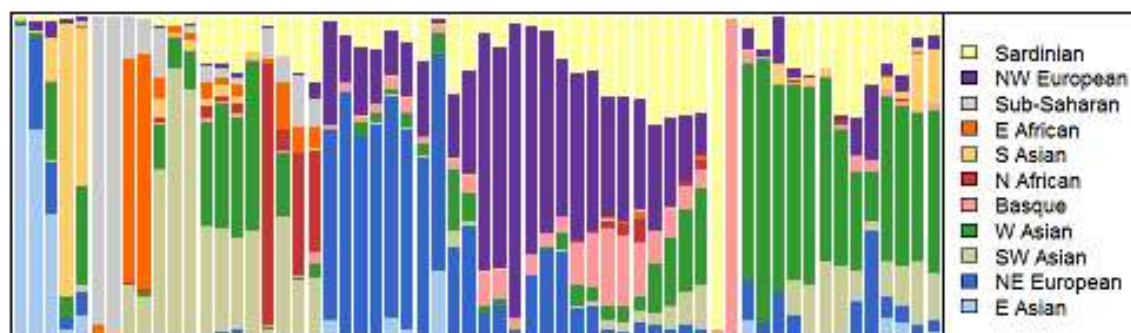
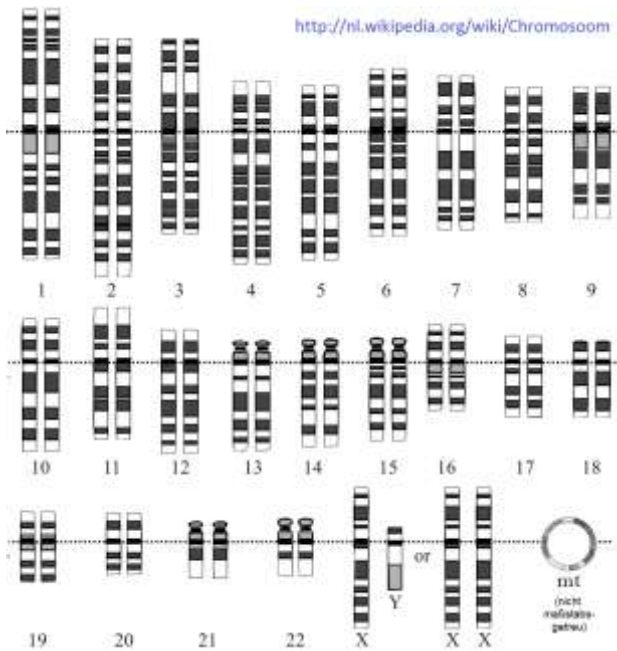


# Ergebnisse der autosomalen Genealogie



Da autosomale DNA und X-DNA mit ihren homologen Chromosomen rekombinieren, somit Gene von den Vorfahren zufällig vererbt werden, sind aufwändige Mutationsanalysen notwendig, um genealogische Rückschlüsse ziehen zu können.



Nach dem erfolgreichen Entziffern der mt-DNA 1987, der daran anschließenden noch informativeren Entschlüsselung des Y-Chromosoms im Jahr 2000 und der Erstellung von genetischen Stammbäumen (Haplogruppen) für beide, wurde es technisch möglich noch größere Datenmengen zu verarbeiten.

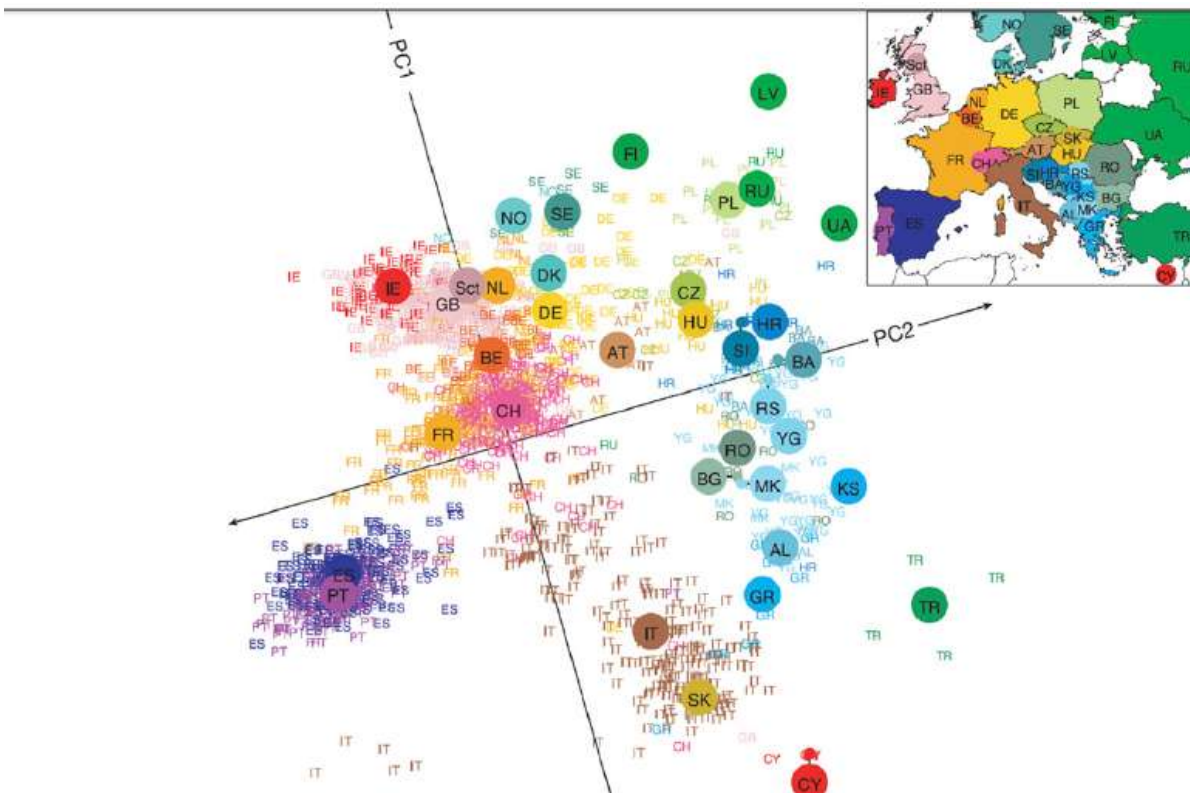
Man begann nun auch auf dem X-Chromosom nach genetischen Mustern Ausschau zu halten. Und da, wie das X-Chromosom die anderen 22 Chromosomen ebenfalls rekombinieren, wurde die Suche auch auf diese erweitert.

Tatsächlich war diese Suche von Erfolg gekrönt, konnte aber auf Grund der Rekombination der Chromosomen nichts Eindeutiges zur genealogischen Abkunft der jeweiligen Träger aussagen.

Hier führte jedoch eine andere Herangehensweise zum Erfolg. Die bisher gefundenen Muster zeigten besondere geografische Auffälligkeiten. In einigen geografischen Regionen kamen bestimmte Muster häufiger vor als in anderen. Es wurden nun die Völker und Gruppen, die in diesen Regionen lebten, gezielt auf diese regionalen Muster hin untersucht.

### Geografische Herkunft bzw. Abstammung

Um die Zusammensetzung der genetischen Abstammung bezogen auf geografische Gruppen (Völker, Clans) analysieren zu können, bedarf es einer ausreichenden Anzahl von Probanden, deren Vorfahren aus einem abgegrenzten bekannten Gebiet stammen. Je nach Analyseform und Wahl der entscheidenden Komponenten, bzw. "Daten-Haufen" (sog. Cluster), sind recht verschiedene Auswertungen und Darstellungen möglich.



Hauptkomponentenanalysen (PCA) werden meist in Diagrammen mit zwei Dimensionen (PC1/PC2) dargestellt. Üblicherweise werden für die vorliegenden Probandendaten die zwei aussagekräftigsten Komponenten der Analyse ausgewählt. Bei multidimensionalen Auswertungen, werden für jeden Probanden die Komponenten angegeben; zumeist in prozentualen Werten. Aus solchen Auswertungen lassen sich auch Landkarten mit den dominierenden Komponenten erstellen.

Auswertungen lassen auch die Erstellung von Stammbäumen zu, welche die Nähe der Verwandtschaft unterschiedlicher Probanden, bzw. Cluster aufzeigen. Besonders bei dieser Darstellungs-Form ist die Wahl der entscheidenden Komponente/n ausschlaggebend und bereits geringe Änderungen an der Auswertungskonstellation, bringen andere Ergebnisse hervor.

[http://en.wikipedia.org/wiki/Genealogical\\_DNA\\_test](http://en.wikipedia.org/wiki/Genealogical_DNA_test)

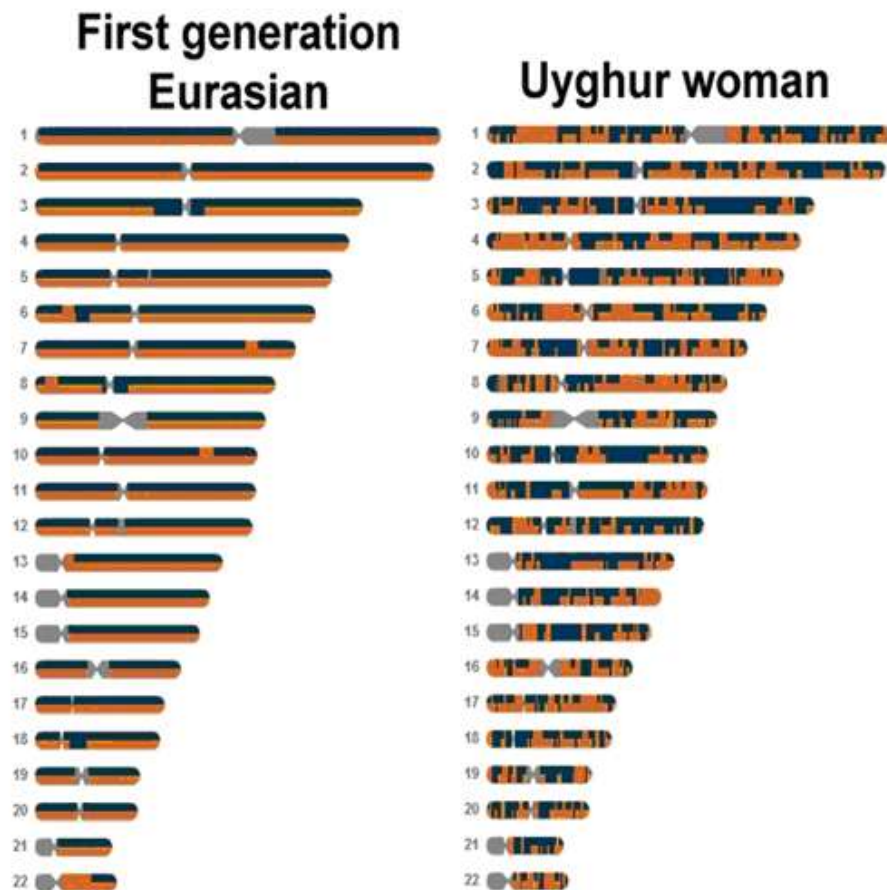
Inzwischen sind autosomale Tests erhältlich, die die miteinander rekombinierenden Chromosomen testen. Diese Tests versuchen den Anteil am gemischt-geografischen Erbe eines Individuums zu messen, indem sie bestimmte Marker zu identifizieren versuchen, welche ancestry informative markers oder AIM genannt werden und mit Bevölkerungen in bestimmten geografischen Gebieten verbunden sind.

Der Wert und die Verlässlichkeit dieser Tests sind in Frage gestellt worden, bleiben aber weiterhin sehr populär. Anomale Funde resultieren meist aus Datenbasen, die zu klein sind, um Marker mit allen Gebieten, in denen sie in einheimischen Bevölkerungen vorkommen, in Verbindung zu bringen.

Die nicht zufällige Verbindung von Merkmalen (Allelen) auf zwei oder mehr Genorten (Loci), die von einzelnen Vorfahren-chromosomen abstammen, wird in der Populationsgenetik als Verbindungsungleichgewicht (**linkage disequilibrium**) bezeichnet. Dies kann dazu genutzt werden das Alter einer Mischbevölkerung zu bestimmen, die aus zwei oder mehr nicht vermischten Ausgangspopulationen hervor gegangen ist.

Im Beispiel unten ist der Anteil der Ursprungspopulationen bei beiden Probanden 50%. Die Länge der Genabschnitte wurde jedoch im Laufe der Jahrhunderte in der Mischbevölkerung der Uiguren durch Rekombination mit jeder Generation kürzer.

In der Grafik unten sind zwei sogenannte "Ancestry Paintings" der Firma 23andMe zu sehen. Das rechte ist ein Referenzbeispiel, eine Uigurenfrau, das andere ist ein Mann mit einer in der ersten Generation europäisch-asiatisch gemischten Vorfahrenschaft. Der Unterschied im Genom der beiden ist recht offensichtlich.



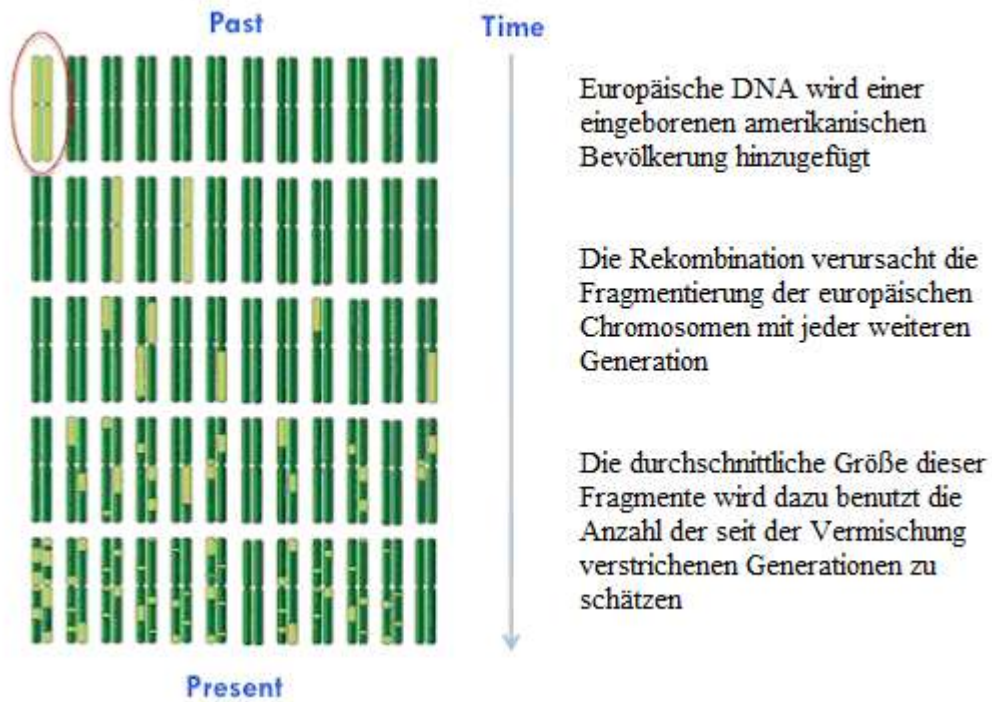
For the record, 23andMe says that the Eurasian man is 50% Asian, 50% European. For the Uyghur woman, 52% European, 48% Asian.

<http://blogs.discovermagazine.com/gvxp/2010/11/eurogens-500k-snp-biogeographicancestry-project/>

Um das Prinzip zu verdeutlichen, ist auf der nächsten Seite das Ergebnis einer Studie zu dieser Problemstellung dargestellt:

Das Ziel der Studie war es, aus der Länge der Chromosomen-Fragmente im Erbgut einer mexikanischen Bevölkerung, die Anzahl der vergangenen Generationen abschätzen zu können, seit dem die europäische DNA Eingang in die ursprüngliche Bevölkerung fand. Die Ergebnisse waren mit der bekannten Geschichte der letzten 500 Jahre in guter Übereinstimmung.

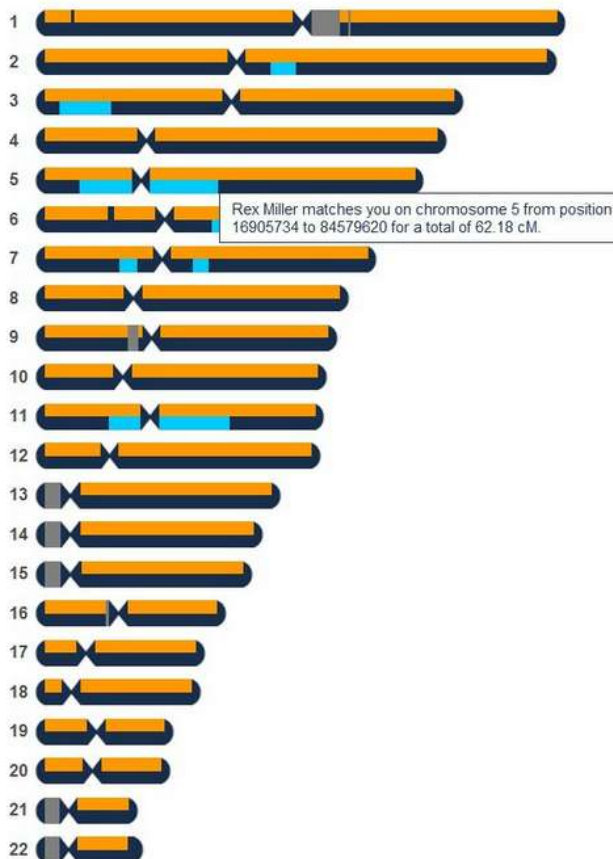
# Estimating Generations Since Admixture



<http://simonsoutherton.blogspot.de/2013/05/could-lamanite-dna-just-disappear.html>

[dna-explained.com/2012/09/03/matches-family-ibd-vs-population-ibs/](http://dna-explained.com/2012/09/03/matches-family-ibd-vs-population-ibs/)

So let's look at an example using the Family Finder Chromosome Browser from [Family Tree DNA](#) which allows you to compare the inherited pieces of DNA of multiple people.



Optional Views: [Download to Excel \(CSV Format\)](#) | [View this data in a table](#)

The graphic above shows the comparison of my mother to me, shown in orange, and then to a Miller cousin, shown in blue. My mother and I share half of all of our DNA, so my orange matches her on every chromosome.

Inzwischen ist es auch Privatpersonen möglich, nahe und entferntere Verwandte mit „Ancestry Painting“ autosomal genetisch zu vergleichen.

Der so genannte Family Finder, eine Programm-Funktion der Firma Family Tree DNA, stellt zu diesem Zweck die jeweils übereinstimmenden Gene, der durch verschiedene Farben gekennzeichneten Personen, für jeden einzelnen Chromosomenort, in einem Genogramm dar.

Das Beispiel links stammt von Roberta Estes, die auf ihrer hervorragenden Webseite [dna-explained.com](#), auf recht unterhaltsame Weise, die neuesten Erkenntnisse der Forschungsgemeinschaft, als auch die Ergebnisse ihrer, teilweise persönlichen, genetischen und genealogischen Familienforschung darstellt.

Die Grafik zeigt den Vergleich der Mutter Robertas in Dunkelblau, mit Roberta in Orange und einem Cousin Robertas, mit dem Namen Rex Miller, in Hellblau.

Da Roberta und ihre Mutter die Hälfte ihrer gesamten DNA gemeinsam, haben zeigen sich auf allen Chromosomen ihre orangefarbenen Übereinstimmungen mit den dunkelblauen der Mutter.

Ihre Mutter und der Miller-Cousin haben den Urgroßvater (John David Miller) gemeinsam, weshalb beide erwarten können jeweils etwa 12,5 % ihrer DNA von diesem geerbt zu haben.

Da aber nicht jedem Nachfahren exakt dieselbe DNA von diesem Vorfahren weitergegeben wurde, kann man damit rechnen, dass statistisch gerechnet, etwa 3,12 % des Erbgutes mit dem des Cousins identisch ist.

[http://en.wikipedia.org/wiki/Race\\_%28human\\_classification%29](http://en.wikipedia.org/wiki/Race_%28human_classification%29)

Eine Kritik an der Vorgehensweise, beim Sammeln der Daten und den Schlüssen aus den Ergebnissen, formulierte Charles Rotimi von [Howard University](#)'s National Human Genome Center 2003 folgendermaßen:

"die Natur der Erscheinung des genetischen Gruppierens (Clustering) von Menschen ist eine Funktion davon, wie die Populationen zusammengetragen werden, wie die Grenzkriterien zwischen Clustern definiert werden und wie hoch die benutzte Auflösung ist."

Da jedes dieser Kriterien die Resultate beeinflussen kann, zog er den Schluss, dass die Probanden sehr vorsichtig sein sollten, eine Verbindung zwischen genetischen Linien oder Gruppen und ihrer persönlichen Vorstellung von Identität herzustellen.

Andererseits argumentierte Rosenberg (2005), wenn genügend genetische Marker und Personen analysiert würden, wären die gefundenen Cluster beständig. Wie viele genetische Marker ein kommerzielles Unternehmen nutzt variiert dabei stark, obwohl neue Technologien es stetig erlauben, noch größere Daten-Mengen zu analysieren. Schließlich gründen Menschen üblicherweise ihre Identität mehr auf Familie und persönliche Beziehungen, als auf Daten.

<https://genographic.nationalgeographic.com>

Das Prinzip der Fragmentanalyse und der prozentualen Zuordnung der Fragmentmuster zu regionalen Populationen, die als Modellpopulationen bzw. Ursprungsregionen postuliert werden, wurde von Genographic Project als Muster und Grundlage ihrer autosomalen Forschung ausgearbeitet.

Dazu wurden neun Ursprungsregionen beispielhaft ausgewählt, deren Fragmentmuster als typisch für diese Populationen gelten. Je nachdem, wie viele Fragment-Anteile einer untersuchten Person zu einer bestimmten Ursprungsregion gehören, wird der jeweilige Anteil als Prozentzahl der Region im Erbgut der Person ausgedrückt. Zuerst folgt nun ein Überblick über die Ursprungsregionen, danach werden heutige Bevölkerungen in ihrer genetischen Zusammensetzung grafisch dargestellt.

## Who Am I: Regions Overview

Indem von Genographic Project eine Analysemethode benutzt wurde, die auf Ursprungsbevölkerungen der Vorfahren schließen lässt, wurden neun Ursprungs-Regionen von Genographic Project postuliert, die in unterschiedlicher Kombination erlauben, jedermanns Genom darzustellen. Jeder Mensch trägt eine Mischung dieser Komponenten zwischen 0% und 100% in sich, aufgrund der Interaktionen dieser Weltregionen während der letzten 50 000 Jahre. Die eingeborenen Indianer Amerikas besitzen ihre eigene Gruppe, wie auch die Ozeanier und die KhoiSan Südafrikas. An diesen Orten ist es sogar möglich 100% der jeweiligen Region zu finden. Die meisten anderen Menschen in der Welt besitzen eine Mischung aus diesen neun Komponenten, die alte und neue Wanderungsmuster widerspiegelt. Diese Gruppen sind:

**Mediterranean:** Diese Komponente wird mit der größten Häufigkeit in Südeuropa und der Levante gefunden - Sardinien, Italien, Libanon, Griechenland, Ägypten und Tunesien sind sog. Referenzpopulationen. Obwohl nicht auf diese Gruppen begrenzt, wird sie weniger häufig im Rest Europas, dem Mittleren Osten, in Zentral- und Südasien gefunden. Die Komponente ist wahrscheinlich ein Signal der jungsteinzeitlichen Expansion aus dem Mittleren Osten nach Europa, die vor etwa 8 000 Jahren im westlichen Teil des Fruchtbaren Halbmondes begann.

**Southwest Asian:** Diese Komponente wird mit der größten Häufigkeit in Indien und benachbarten Populationen, einschließlich Iran und Tadschikistan als Referenzpopulationen im Genographic-Datensatz, gefunden. Mit geringerer Häufigkeit findet man sie auch in Europa und Nordafrika. Wie auch die Mediterrane Komponente, geschah ihre Ausbreitung während der jungsteinzeitlichen Expansion, vielleicht aus dem östlichen Teil des Fruchtbaren Halbmondes. Auch Personen mit stark europäischem Einfluss zeigen Spuren dieser Komponente, da alle Europäer sich über Jahrtausende mit Leuten aus Südwestasien (Schwarzmeerregion) vermischt haben.

**Northern European:** Diese Komponente wird mit der größten Häufigkeit in nordeuropäischen Bevölkerungen gefunden – Großbritannien, Dänemark, Finnland, Rußland und Deutschland sind Referenzpopulationen. Mit geringerer Häufigkeit findet man sie auch im Rest Europas. Diese Komponente ist höchst wahrscheinlich ein Signal der Jäger und Sammler, der frühesten Einwohner Europas, die als letzte vor ca. 8 000 Jahren während der Jungsteinzeit, den Übergang zu Ackerbau und Viehzucht vollzogen.

Weitere Ursprungsregionen sind:

**Subsaharan African, Northeast Asian, Southern African, Native American, Oceanian and Southeast Asian.**

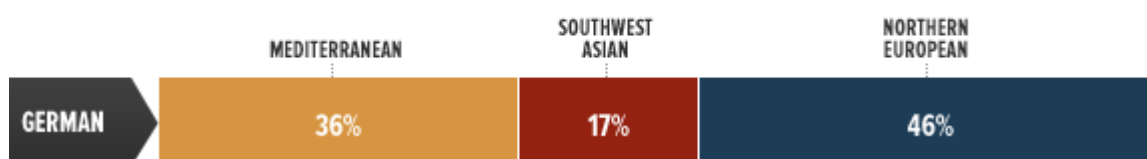
Auf der nächsten Seite sind einige Beispiele heutiger Bevölkerungen, aufgeschlüsselt nach den Ursprungsregionen in der für Genographic typischen Darstellungsweise aufgelistet. Dazu wurden die autosomalen Muster, der untersuchten Personen einer Bevölkerung aufsummiert und als prozentualer Anteil des Musters im Genom der Population ausgewiesen:

## BRITISH (UNITED KINGDOM)



This reference population is based on samples collected from populations in the United Kingdom. The dominant 50% northern European component likely reflects the earliest settlers in Europe, hunter-gatherers who arrived there more than 35,000 years ago. The 33% Mediterranean and 17% Southwest Asian percentages arrived later, with the spread of agriculture from the Fertile Crescent in the Middle East, over the past 10,000 years. As these early farmers moved into Europe, they spread their genetic patterns as well. Today, northern European populations retain their links to both the earliest Europeans and these later migrants from the Middle East.

## GERMAN



This reference population is based on samples collected from people native to Germany. The dominant 46% Northern European component likely reflects the earliest settlers in Europe, hunter-gatherers who arrived there more than 35,000 years ago. The 36% Mediterranean and 17% Southwest Asian percentages probably arrived later, with the spread of agriculture from the Fertile Crescent in the Middle East over the past 10,000 years. As these early farmers moved into Europe, they spread their genetic patterns as well. Today, northern and central European populations retain links to both the earliest Europeans and these later migrants from the Middle East.

## EGYPTIAN



This reference population is based on samples collected from native Egyptians. As ancient populations migrated from Africa, they passed first through southwest Asia. The 65% Mediterranean and 18% Southwest Asian components in Egypt are representative of that ancient migratory route, as well as later migrations from the Fertile Crescent in the Middle East with the spread of agriculture over the past 10,000 years, and migrations in the 7th century with the spread of Islam from the Arabian peninsula. The 14% sub-Saharan African indicates intermixing with African populations to the south.

Sie und weitere heutige Bevölkerungen können auf der Internetseite <https://genographic.nationalgeographic.com>, des Genographic Project nachgeschlagen werden. Ebenso sind dort weitere interessante Forschungsergebnisse in englischen Artikeln dargestellt.

Grundsätzliche Informationen über das Projekt sind auf [http://de.wikipedia.org/wiki/The\\_Genographic\\_Project](http://de.wikipedia.org/wiki/The_Genographic_Project) zu finden.

## Autosomal Admixture – Aufschlüsselung von Bevölkerungen nach Ursprungsregionen

Den Hintergrund der nun folgenden Darstellungen möchte ich anhand eines Beispiels erläutern:

Will man eine Gruppe von Menschen nach ihren Merkmalen ordnen, so bestimmt man wohl zuerst die augenscheinlichsten Unterschiede. Sagen wir, dies wäre die Haarfarbe und wir lassen nur zwei Gruppierungen oder Kategorien (K) zu: Braun und Blond.

Ein Computerprogramm wird nun alle Menschen in diese Kategorien einteilen, egal wie die tatsächliche Haarfarbe auch aussehen mag. Schwarz und kräftiges Rot wird bei Braun registriert, Rotblond, Grau oder Weiß bei Blond. Es ist eine sehr grobe Einteilung der zugrunde liegenden tatsächlichen Farben. Erhöht man die Kategorien, so erhöht man die Genauigkeit der Ergebnisse.

Mit fünf Kategorien: Schwarz, Braun, Rot, Blond und Weiß erhält man wohl die beste Information. Aber natürlich existieren auch Schattierungen und Übergänge. Also erhöht man die Kategorien immer weiter, bis der Computer keine sinnvollen Ergebnisse mehr hervor bringt.

Auf der Seite von Genographic Project wurde dieses Prinzip, statt mit Haarfarben, mit neun geografischen Ursprungsregionen und ihren vom Computerprogramm „erkannten“ spezifischen genetischen Mustern angewandt, um die untersuchten Bevölkerungen miteinander zu vergleichen. Das Programm, das zur Untersuchung autosomaler Gruppierungen meistens verwendet wird, heißt ADMIXTURE.

Dieses Programm arbeitet mit den autosomalen Daten (Zehntausenden SNP's pro Person) von einigen wenigen Personen bis zu mehreren Tausend, teilt sie in K Kategorien, die vor der eigentlichen Analyse gewählt werden und gibt sie als prozentuale Anteile je Individuum oder Population zurück.

Weiter unten werden die Ergebnisse für eurasische Bevölkerungen dargestellt und die Schlussfolgerungen daraus erläutert.

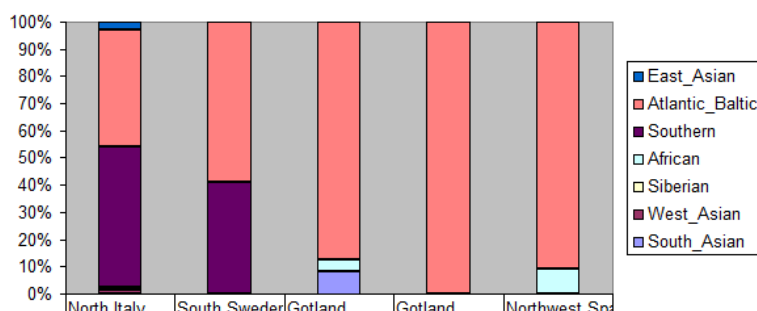
Die genauere Erklärung ist auf <http://dienekes.blogspot.com.es/2012/07/bronze-age-indo-european-invasion-of.html> im Kommentar von Kostas zu finden:

Die K's repräsentieren einige tausend Datenpunkte in einem sehr hochdimensionalen Vektorraum. Manche Punkt-Gruppen sind einander ähnlicher als andere, weshalb man Gruppierungen oder „Cluster“ bilden kann. Innerhalb dieser Cluster kann man wieder Unter-Cluster finden etc.

Manchmal ist es nützlich Daten mit drei Clustern zu analysieren, während man bei anderen Gelegenheiten 10 oder 20 K für die Gesamtzahl der Cluster benötigt. Wenn Dienekes K7 erwähnt, splittet er seine Gen-Datenbasis in 7 Cluster auf und untersucht dann, welche Teile des Genoms von Ötzi besser zu jedem Cluster passen. Wie man an den Ergebnissen sehen kann, passen 51,9% von Ötzis Genom „besser“ zu dem Cluster der Menschen aus Südeuropa, während 43,1% besser zu dem Atlantic\_Baltic genannten Cluster der Jäger und Sammler passen.

Man kann es sich so vorstellen: die Bezeichnungen East\_Asian, Atlantic\_Baltic, etc. sind analog zu der Basis, mit der ein Vektorfeld erzeugt wird. Das eigentliche Problem besteht jedoch darin, dass in der Genetik ein Computerprogramm mit Namen ADMIXTURE verwendet wird, dessen Source-Code nicht frei zugänglich ist. Die genaue Prozedur ist also unbekannt und könnte genauso gut durch Magie funktionieren.

**Ancient European DNA (K7b)**

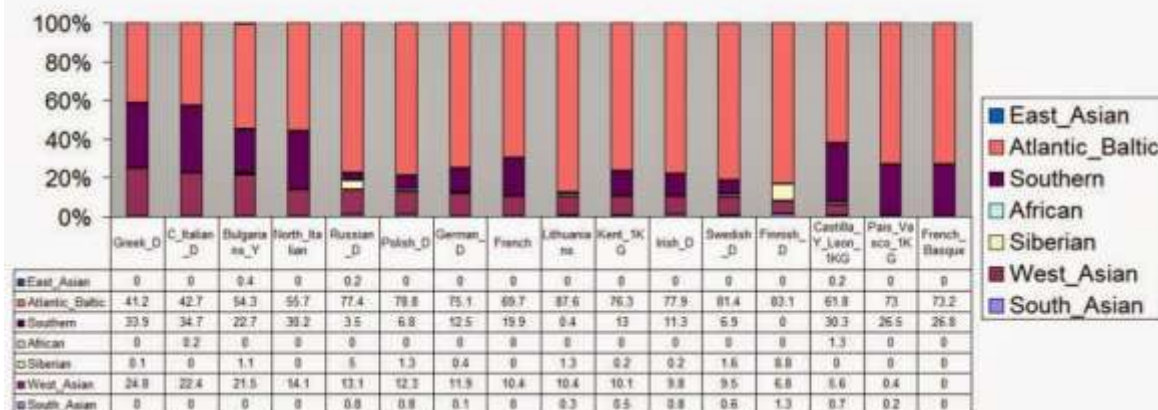


East_Asian	2.8				
Atlantic_Baltic	43.1	59.1	87.6	100	90.7
Southern	51.9	40.9			
African	0.8		4.3		9.3
Siberian					
West_Asian	1.4				
South_Asian			8.1		

Die Ancestral-Component-Analyse kommt der linearen Zusammensetzung einer Matrix genetischer Daten am nächsten (jede mit Zehntausenden SNP's), mit einer bestimmten Anzahl Individuen (bis zu einigen Tausend), aufgeteilt in K Eigenvektoren (wobei der Wert von K vor der Analyse gewählt wird und zwischen zwei und bis zu einem Dutzend oder mehr Eigenvektoren liegen kann). Diese werden in unterschiedlichen Anteilen vermischt, um eine Matrix für jedes Individuum darzustellen.

Die Eigenvektoren repräsentieren bestimmte historische oder prähistorische Wellen massiver Bevölkerungswanderungen, die nach Vermischung einen stabilen genetischen Bevölkerungsmix ergeben, der sich lokal stabilisiert hat. Die größten Kontroversen entzündeten sich an den Details darüber, wie viel Vermischung mit welchen Wellen zu welcher Zeit geschah.

**K7b (Modern Populations)**

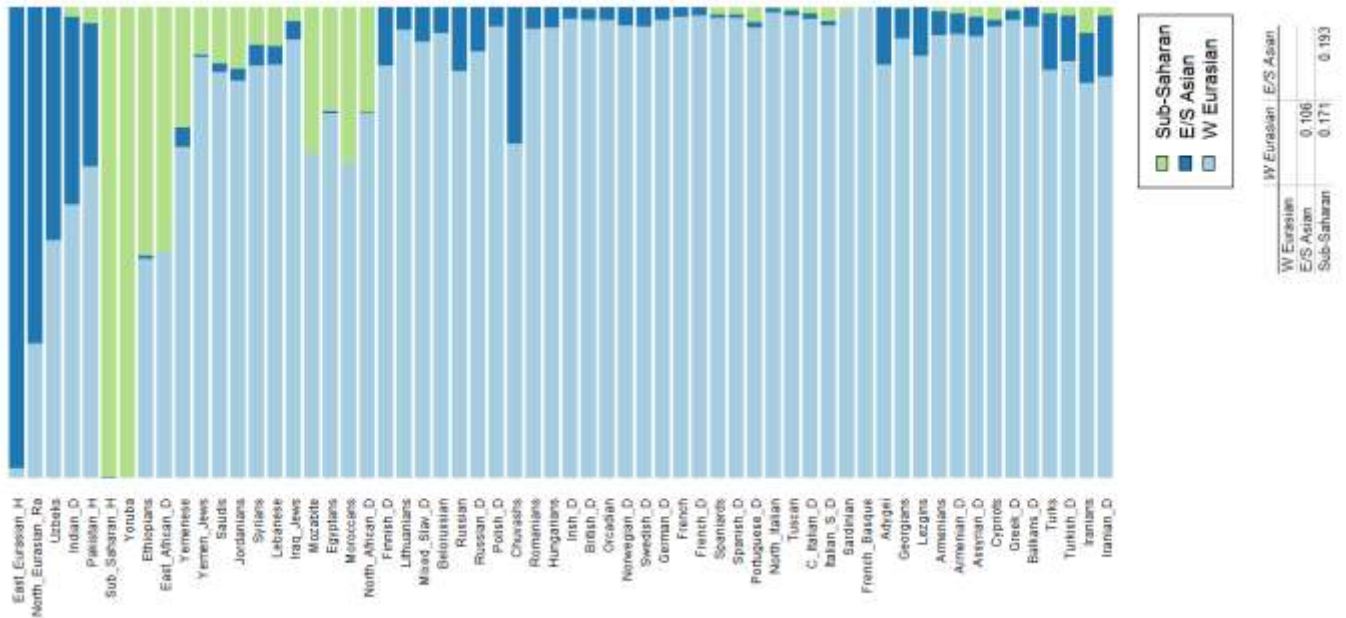


**Genetische Struktur West-Eurasischer Bevölkerungen**

Dienekes Pontikos hat auf seinem Blog die Ergebnisse der ADMIXTURE-Programmdurchläufe dargestellt.

Die technischen Einzelheiten sind durch obigen [Link](#) zu erreichen und nachzulesen. Um die Lesbarkeit des Textes zu erhöhen wurden sie hier weggelassen, da sie zum Verständnis auch nicht unbedingt notwendig sind. Die vollständigen Ergebnisse und Quell-Bevölkerungen können dort heruntergeladen werden. Die unten stehenden Texte wurden aus dem Englischen möglichst wortgetreu übersetzt.

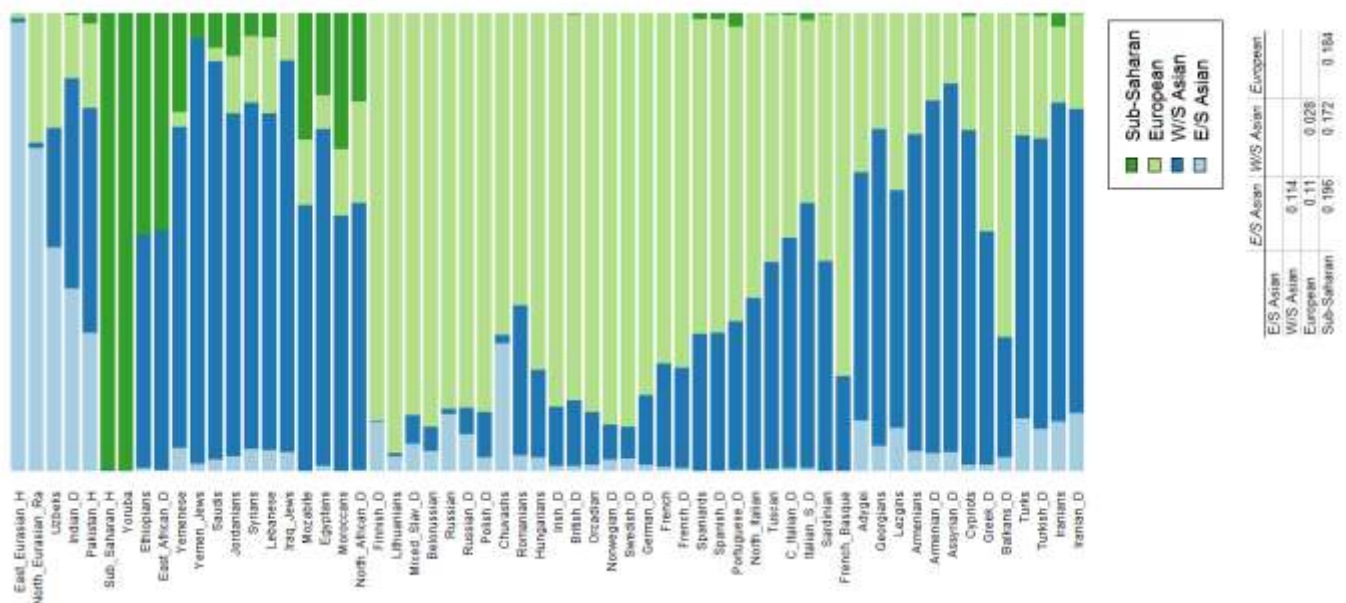
**Die Resultate:**



Bei **K=3** beobachten wir das Erscheinen von West-Eurasischen, Sub-Saharischen, and Ost/Süd-Asiatischen Komponenten in den untersuchten Bevölkerungen.

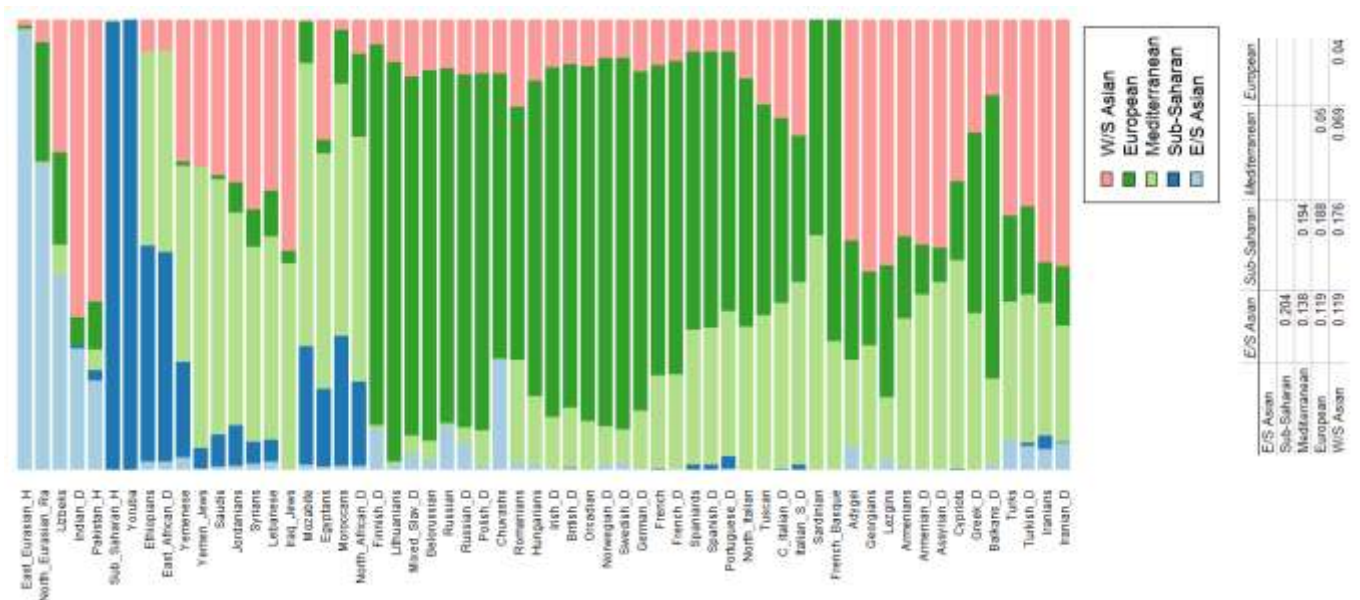
Der Einfluss der subsaharischen Komponente kommt am deutlichsten in Nordafrika und dem Nahen Osten zum Ausdruck, im Besonderen unter Arabern. Der Einfluss der Ost/Süd-asiatischen Komponente erscheint in Westasien (Kaukasus/Anatolien) und in Nordosteuropa besonders unter Finnisch- und Türkisch-sprachigen Bevölkerungen.

Interessant erscheint auch, dass 39,8% der indischen Gruppe zur Ost/Süd-Asien-Komponente gehören. D.h. dass 1/3 der indischen Vorfahrenschaft stärker mit Ostasiatischen als Westeurasischen Bevölkerungen verwandt ist.



Bei **K=4**, erscheint die erste Teilung innerhalb der Kaukasischen Gruppe: Eine Komponente in Europa und eine deren Zentrum in West-/Süd-Asien liegt. Viele Bevölkerungen besitzen diese beiden Komponenten in unterschiedlichen Anteilen.

Die europäische Komponente schrumpft bis zur Unkenntlichkeit in Arabern, wie den Saudis und den Jemeniten. Die West-/Südasiatische Komponente verschwindet in Nordosteuropäern, wie Finnen, Litauern, Nordrussen und Tschuwaschen.



Bei  $K=5$  erscheint eine neue mediterrane Komponente. Dies zeigt sich besonders stark in den Bevölkerungen, die im Norden, im Süden und im Osten des Mittelmeeres an dieses grenzen.

Bemerkenswert ist die Abwesenheit dieser Komponente in Indien und Nordosteuropa.

In Nordosteuropa gibt es die mediterrane Komponente praktisch überhaupt nicht, während die west-/südasiatische Komponente, nun von ihren mediterranen  $K=4$  Komponenten befreit, auf der Bildfläche erscheint.

Umgekehrt schwindet im Westlichen Mittelmeer, unter Basken Sardiniern, Marokkanern und Mozabiten, die west-/südasiatische Komponente bis zur Nichtexistenz.

An dieser Stelle werden die weiteren Programmdurchläufe, aus Platzgründen, abgebrochen. Die vollständige Analyse kann unter dem Link <http://dienekes.blogspot.com/2011/04/genetic-structure-of-west- Eurasians.html> eingesehen werden.

<http://www.rtf1.de/news.php?id=7196>

Der Molekularbiologe und Anthropologe Dr. Wolfgang Haak aus Künzelsau schrieb im Jahr 2006 an der Johannes Gutenberg-Universität Mainz seine Dissertation über „Die Populationsgenetik der ersten Bauern Mitteleuropas“.

Heute arbeitet er an der Universität Adelaide in Australien als Senior Research Associate mit einem Team aus deutschen, australischen und amerikanischen Forschern an den Rätseln der europäischen Frühgeschichte.

Eine bahnbrechende Arbeit auf dem Gebiet der autosomalen Genetik, die Herkunft der europäischen Sprachen betreffend, wurde im März 2015 im naturwissenschaftlichen Magazin [Nature](http://www.nature.com) veröffentlicht.

## Massive migration from the steppe was a source for Indo-European languages in Europe

Fast drei Milliarden Menschen sprechen heute eine der 445 Sprachen, die der indoeuropäischen Sprachfamilie zugerechnet werden. Dazu gehören auch Deutsch, Englisch, Französisch, Italienisch, Griechisch, Iranisch und Russisch. Die Ähnlichkeit dieser in Eurasien weitverbreiteten Sprachen wurde bereits vor 200 Jahren erkannt. Ihre Herkunft und Ausbreitung liegt jedoch noch weitestgehend im Dunkeln. In einer nun im Wissenschaftsmagazin Nature erschienenen Studie, an der auch die Universität Tübingen beteiligt war, fand ein internationales Forschungsteam jetzt Hinweise: Demnach gab es vor 4.500 Jahren eine massive Wanderungsbewegung von Menschen aus den eurasischen Steppengebieten nach Europa. Das könnte einen entscheidenden Einfluss auf die Entwicklung der indoeuropäischen Sprachen gehabt haben.

Das Team unter Leitung der Harvard Medical School in Boston, USA, und des Australian Centre for Ancient DNA der Universität Adelaide fand Hinweise auf eine massive Wanderungsbewegung von Menschen aus den eurasischen Steppengebieten vor circa 4.500 Jahren, die Einfluss auf die Verbreitung einiger indoeuropäischer Sprachgruppen gehabt haben muss. Die Proben für die Studie wurden von einem internationalen Team unter maßgeblicher Beteiligung der Universitäten Mainz, Basel und Tübingen, dem Landesamt für Vorgeschichte in Halle sowie dem neuen Max-Planck-Institut für Menschheitsgeschichte in Jena zusammengestellt.

Besonders viele Proben stammen aus Sachsen-Anhalt, wo beim Bau von ICE-Trassen und Bundesstraßen wertvolle Neufunde gemacht wurden, deren genetische Analyse auch von der Deutschen Forschungsgemeinschaft (DFG) finanziell unterstützt wurde.

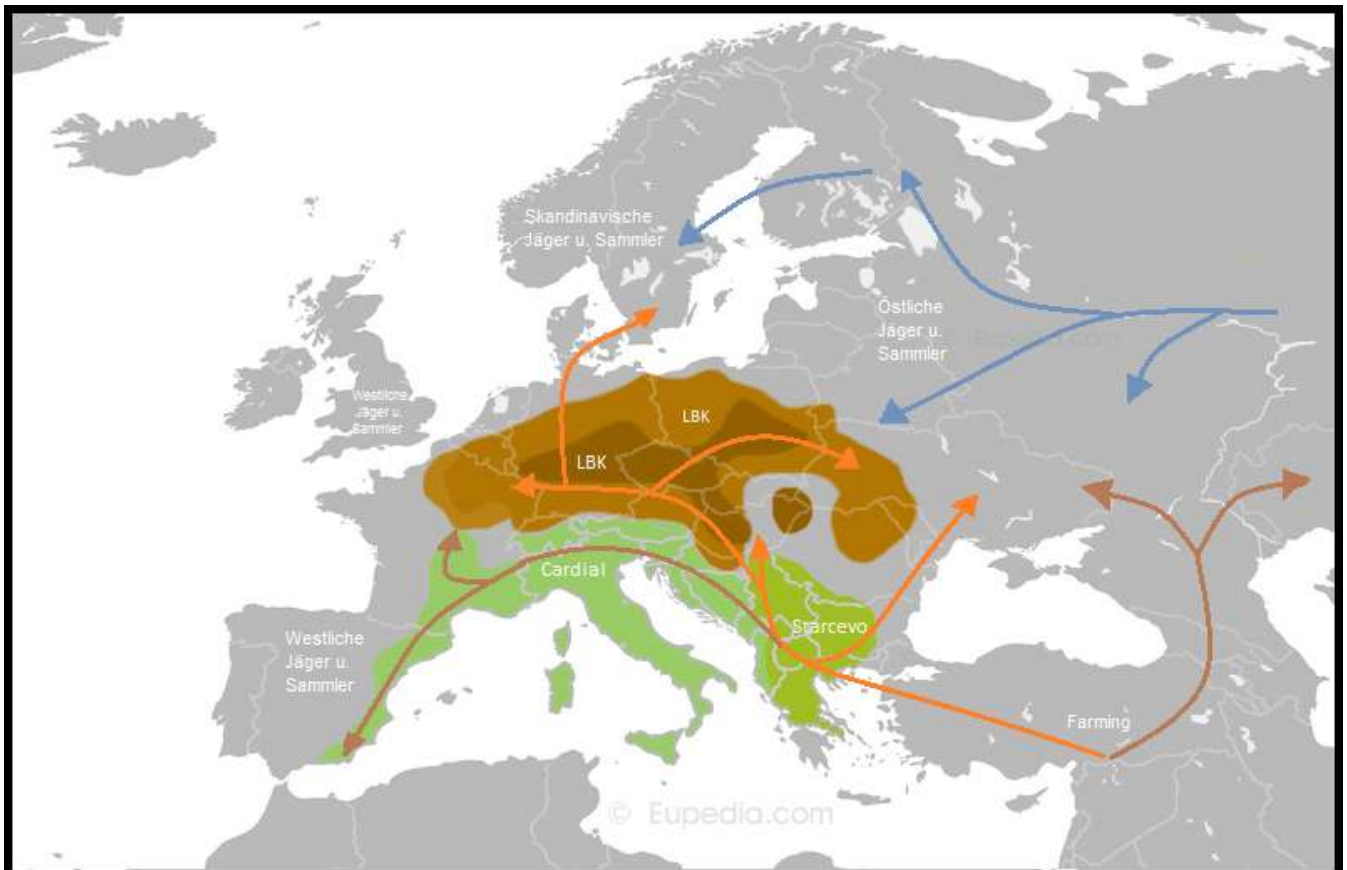
## Gemeinsame Herkunft der frühen Bauern

Insgesamt wurden die Gene von mehr als 90 Individuen sequenziert, die zwischen etwa 3.000 und 8.000 Jahren vor heute in Europa lebten. Bei der Analyse der Datensätze kristallisierten sich zwei wesentliche Bevölkerungsbrüche heraus:

Der erste Umbruch geht auf die Ausbreitung der frühen Bauern über ganz Europa zurück. Sie zogen vor mehr als 9.000 Jahren aus dem Nahen Osten gen Westen und wurden bereits vor rund 7.500 Jahren in Mittel- und Westeuropa sesshaft.

Hier lebten seit der letzten Eiszeit Wildbeutergruppen über ganz Europa verteilt. Die zweite Bevölkerungsgruppe der Bauern unterscheidet sich genetisch deutlich von den damals in Europa lebenden Jägern und Sammlern.

Die genetischen Daten der Jäger und Sammler formen einen osteuropäischen Cluster (EHG) am einen Ende und einen westeuropäischen Cluster (WHG) am anderen Ende eines Jäger und Sammler-Spektrums, während die schwedischen Jäger und Sammler genetisch, wie geografisch dazwischen liegen. Die Gene der Ackerbauern und Viehzüchter hingegen sind selbst über große Distanzen erstaunlich homogen.



Archäologisch wurden bisher zwei unterschiedliche Wanderungsrouten der Bauern beschrieben, welche im Wesentlichen auf Unterschieden im materiellen Fundgut wie zum Beispiel Cardial-Keramiken aus dem Mittelmeerraum und dem mittel- und nordeuropäischen Inland (Linearbandkeramik) beruhen.

„Die genetischen Daten bestätigen dies jedoch nicht“, äußert der Erstautor der Studie, Dr. Wolfgang Haak von der Universität Adelaide. „Die frühen Ackerbauern aus Spanien, Deutschland und Ungarn sind genetisch nahezu identisch, was auf einen gemeinsamen Ursprung im Nahen Osten schließen lässt.“

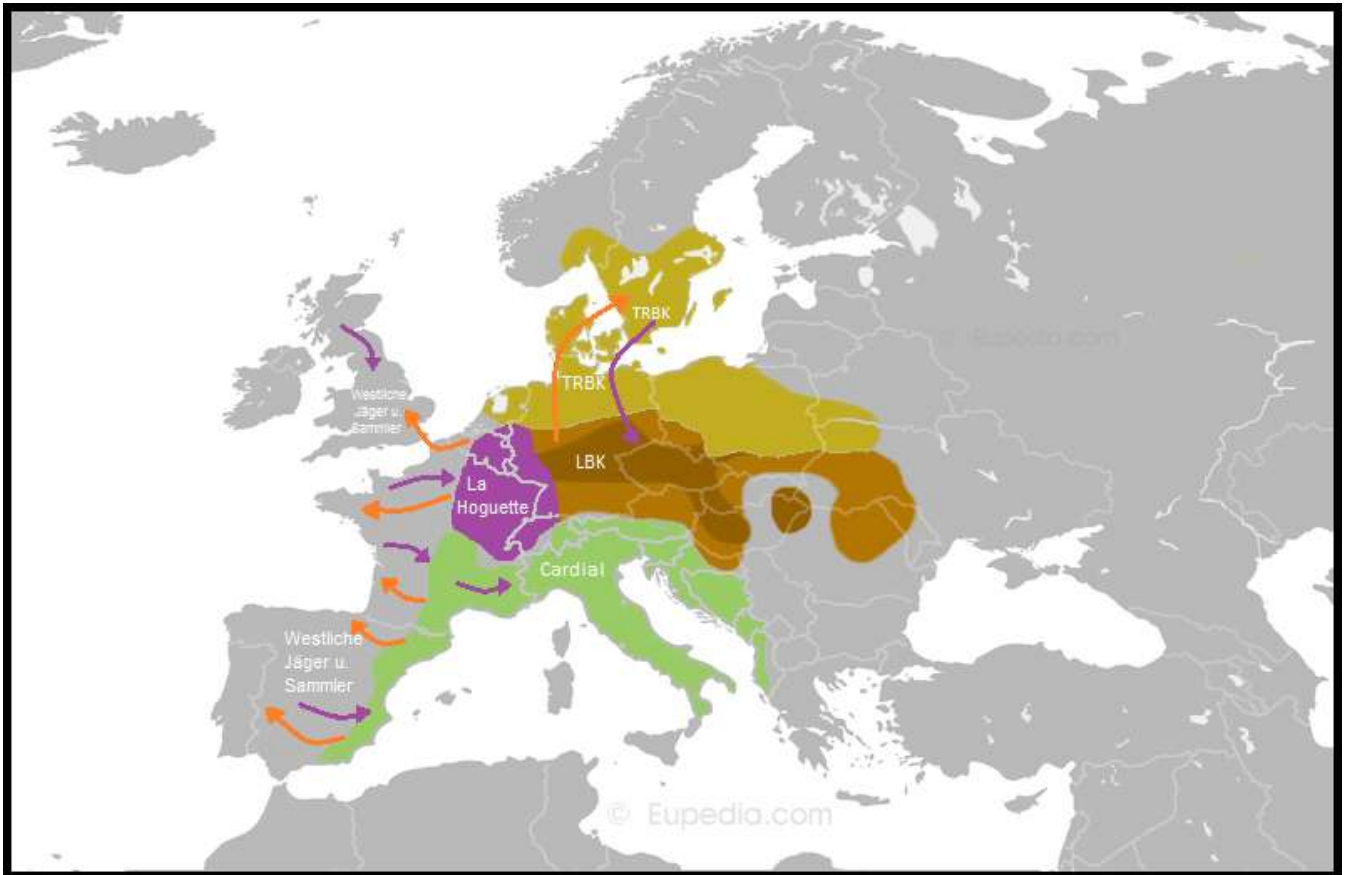
Etwa 2000 Jahre lang lebten die Jäger und Sammler und die Ackerbauern nebeneinander, ohne die Lebensweise der anderen zu übernehmen. Die Funde zeigen, dass der Anteil der Ackerbauern an der Bevölkerung in dieser Zeit stetig stieg, bis sie mit ca. 3/4 die Mehrheit stellten.

Die Jäger und Sammler sind jedoch nicht komplett verschwunden. „Um 6.000 bis 5.000 Jahren vor heute sehen wir einen Wiederanstieg des Jäger-Sammler-Anteils im Genom“, sagt der Co-Erstautor Dr. Iosif Lazaridis von der Harvard Medical School. „Das bedeutet, dass Jäger-Sammler-Gesellschaften noch lange nach Ankunft der Bauern bestanden haben müssen.“

Der Autor Professor Kurt Alt von der Universität Basel und der Privatuniversität Krems fügt hinzu: „Es zeigt aber auch, dass die Jäger und Sammler nach und nach in bäuerliche Gemeinschaften integriert wurden.“

## Die Wanderungsbewegung aus dem Osten

In ihren früheren Arbeiten hatten einige derselben Autoren auf die Zusammensetzung der Gene aller Europäer aus hauptsächlich drei Linien von Ahnen hingewiesen: ein bisschen Jäger und Sammler, ein bisschen frühe Bauern, und eine dritte Komponente, die Ähnlichkeit zu Sibirern und gar den ersten Indianern Amerikas zeigte. Genomdaten von Bauern hatten gezeigt, dass der dritte Anteil zu dieser ganz frühen Zeit noch nicht in Europa vorhanden war und daher erst später hinzugekommen sein konnte – wann und wie, das war jedoch unklar.



Im Mittel-Neolithikum zwischen 4.000 v.C. und 3.000 v.C. zeigen die genetischen Daten der West- und Mitteleuropäer eine leichte Zunahme von WHG-Genen. Aus Berechnungen ergibt sich ein erhöhter Anteil der WHG-Gene zwischen 18% bis 34%. Dies zeigt, dass die ursprüngliche Bevölkerung nun die Lebensweise der eingewanderten Bauern übernommen hatte und nur die Gene noch ihre frühere Herkunft verrieten.

Im Gegensatz dazu erfuhren die Steppe-Hirten der Yamnaya-Kultur im gleichen Zeitraum eine Abnahme an EHG-Genen und einen Zuzug von Genen aus dem Nahen Osten bzw. dem Kaukasus. Allerdings war die Struktur der neuen Gene anders als die der ersten Bauern in West- und Mitteleuropa. Sie hatten eine größere Ähnlichkeit mit den Genen der heutigen Armenier im Gebiet des Kaukasus. Diese Steppe-Hirten waren auch Träger des dritten Anteils, dessen Ursprung wohl in Nordeurasien lag.

„Es war ein echtes Aha-Erlebnis, als wir uns die ersten Daten ansahen“, berichtet Lazaridis. „Der dritte Anteil war in jedem Individuum zu finden, das jünger als 4.500 Jahre war, jedoch in keiner der älteren Proben aus Mitteleuropa.“ Haak geht sogar noch weiter: „Das Signal ist so stark, dass man fast von einer genetischen Datierung sprechen könnte, basierend auf dem Vorkommen von ein, zwei oder allen drei Komponenten.“

Tatsächlich fanden sich unter den Proben auch einige Ausreißer, die bisher archäologisch als älter datiert waren, allerdings die dritte Komponente aufwiesen. Zur Klärung wurden Datierungen mit der Radiokarbonmethode in Auftrag gegeben.

Danach musste das Alter dieser Individuen deutlich nach unten korrigiert werden. „Sie wurden in der Tat jünger als 4.500 Jahre vor heute datiert“, ergänzt Professor Harald Meller, der Direktor des Landesmuseums für Vorgeschichte in Halle/Saale.

In Deutschland seien es die Menschen der schnurkeramischen Kultur (Corded Ware) aus dem dritten Jahrtausend v. Chr., am Übergang zwischen Jungsteinzeit und Bronzezeit, bei welchen erstmals die dritte Komponente auftaucht und damit einen zweiten Bevölkerungsumbruch markiert, sagt Haak, „und zwar mit lautem ‚Hallo!‘“.

„Basierend auf einem direkten Vergleich mit den Individuen der Yamnaya-Kultur, das sind Viehhirten aus den eurasischen Steppengebieten, konnten wir den genetischen Anteil der Menschen aus der Steppe auf beträchtliche 75 Prozent des Genoms der Schnurkeramiker aus Sachsen-Anhalt errechnen,“ sagt Lazaridis, und fügt hinzu, dass sich die Gene der Schnurkeramiker und der Yamnaya-Population trotz der geografischen Distanz von 2.600 Kilometern erstaunlich ähnlich sind.

## Indoeuropäisch aus der Steppe?

Bei einem genetischen Einschlag dieser Größenordnung drängt sich die Frage auf, ob diese Expansion auch einen Einfluss auf die Verbreitung von Sprachen hatte. „Die Ergebnisse legen nahe, dass die Schnurkeramiker nicht nur genetisch eng mit den Hirten aus der Steppe verwandt waren, sondern möglicherweise auch eine ähnliche Sprache hatten“, sagt Lazaridis. „Da sämtliche Mittel- und Nordeuropäer heutzutage einen hohen genetischen Anteil der damaligen Steppenbewohner in sich tragen und zudem eine indoeuropäische Sprache sprechen, lässt sich zumindest ein deutlicher Beitrag der Steppe nicht ausschließen“, bemerkt Haak.



Das deckt sich mit der Ansicht von Linguisten, die argumentieren, dass die Sprachentwicklung schneller voranschreitet als die der Gene und für die eine hypothetische Verbreitung der indoeuropäischen Sprache mit den ersten Bauern daher einige tausend Jahre zu früh angesetzt wäre. Reich fügt hinzu: „Unsere Ergebnisse stellen die Theorie der Sprachverbreitung im Zusammenhang mit der Einwanderung der ersten Bauern in Frage. Mit der spätneolithischen Auswanderung von Menschen aus der Steppe, die einen beachtlichen Einfluss auf andere Bevölkerungsgruppen hatte, erscheint eine spätere, das heißt mehr rezente Verbreitung des Indoeuropäischen plausibel.“

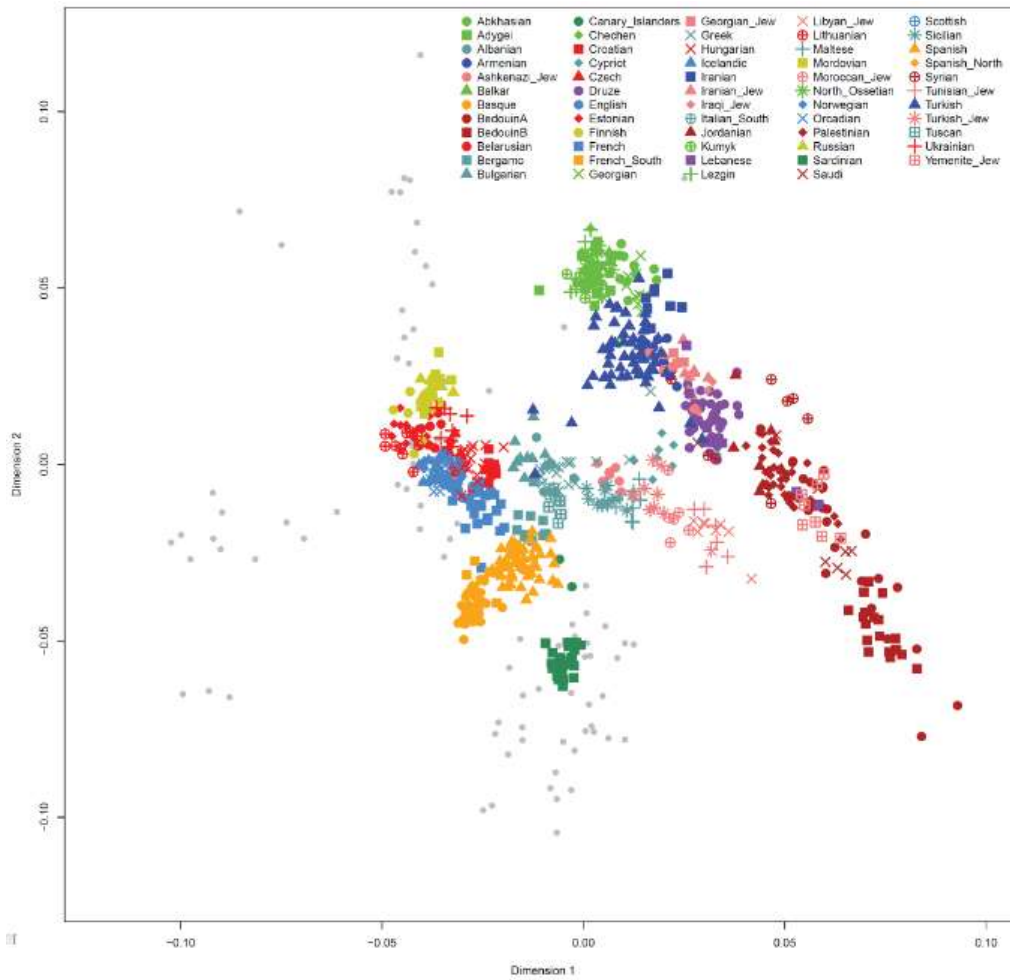
Die fachübergreifende Interpretation von archäologischen, linguistischen und genetischen Daten ist jedoch kontrovers. „Das ist ein heikles Thema und muss mit Bedacht angegangen werden“, warnt Mitautor Johannes Krause, Direktor des neuen Max-Planck-Instituts für Menschheitsgeschichte in Jena, der zuvor an der Universität Tübingen geforscht hat. „Allerdings haben wir mit alter DNA endlich die zeitliche und genetische Auflösung, die uns hier weiterbringen kann.“ Im Oktober sei in Jena ein Workshop geplant, in dem sich Experten aus allen drei Fachrichtungen diesen Fragen widmen wollen.

Eine nach wie vor ungelöste Frage ist die nach dem Ursprung der indoeuropäischen Sprachfamilie. Reich, Haak und ihre Kollegen sind trotz der monumentalen Aufgabe optimistisch, dass man sich der Lösung annähern wird. Haak sagt: „Die Hauptaufgabe besteht nun darin, nach und nach die Fundlücken in unserer genetischen Kartierung zu stopfen. Wir wollen verstehen, wie ähnlich sich Bevölkerungsgruppen aus Europa, Anatolien, dem Kaukasus, Iran und Indien vor 3.000 bis 6.000 Jahren waren, um so dem potenziellen Ursprung der indoeuropäischen Sprachen näher zu kommen.“

### Originalpublikation:

Wolfgang Haak, Iosif Lazaridis, Nick Patterson, Nadin Rohland, Swapan Mallick, Bastien Llamas, Guido Brandt, Susanne Nordenfelt, Eadaoin Harney, Kristin Stewardson, Qiaomei Fu, Alissa Mittnik, Eszter Bánffy, Christos Economou, Michael Francken, Susanne Friederich, Rafael Garrido Pena, Fredrik Hallgren, Valery Khartanovich, Aleksandr Khokhlov, Michael Kunst, Pavel Kuznetsov, Harald Meller, Oleg Mochalov, Vayacheslav Moiseyev, Nicole Nicklisch, Sandra L. Pichler, Roberto Risch, Manuel A. Rojo Guerra, Christina Roth, Anna Szécsényi-Nagy, Joachim Wahl, Matthias Meyer, Johannes Krause, Dorcas Brown, David Anthony, Alan Cooper, Kurt Werner Alt and David Reich: **Massive migration from the steppe is a source for Indo-European languages in Europe**. Nature, Advance Online Publication, 2. März 2015

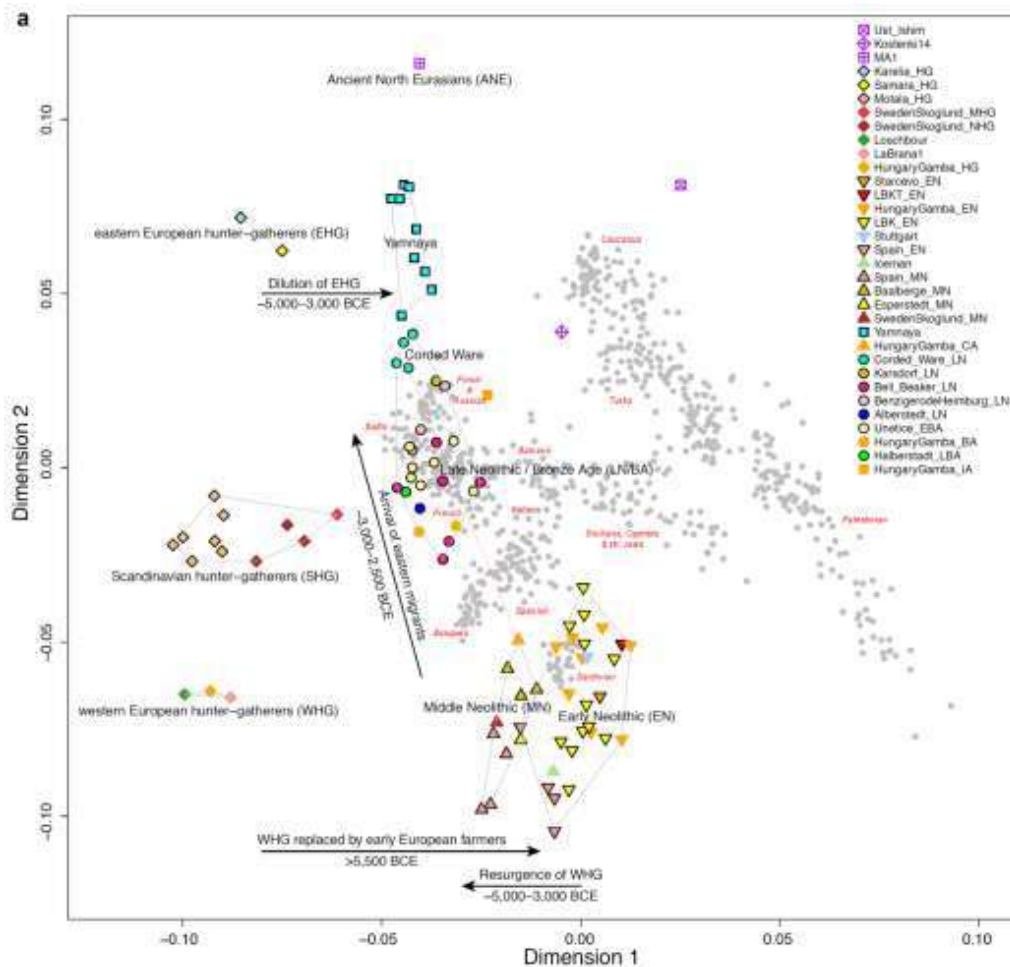
Die folgenden PCA's sollen die Verhältnisse der modernen europäischen Bevölkerungen zu ihren Ursprüngen darstellen.



In dieser Grafik sind die heutigen Populationen des Nahen Ostens und Europas in einer PCA (Principal Component Analyse) dargestellt.

Die sich schräg rechts befindende Häufung repräsentiert dabei die Bevölkerungen des Nahen Ostens.

Die Ansammlung links repräsentiert Europa; die Punkte dazwischen zeigen die Bevölkerungen, die als Brückenpopulationen fungieren.



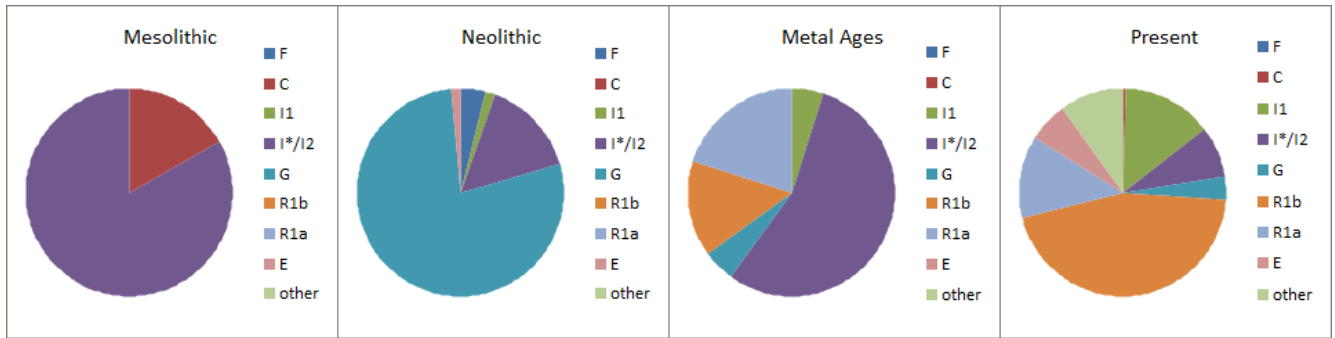
Über die Grafik, nun in grau dargestellt, wurde das, durch genomweite Sequenzierung gewonnene Datenmaterial von 91 Europäern gelegt.

Diese Europäer lebten von vor 8 000 bis vor ca. 3 000 Jahren, vom Mesolithikum bis zur Eisenzeit.

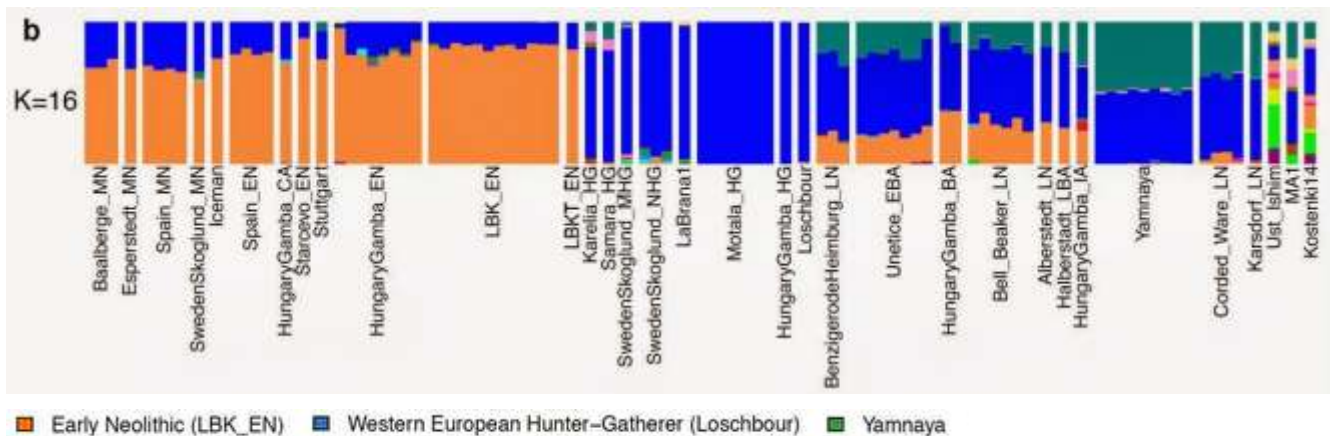
Hinzu kamen noch drei weitere Proben aus der Zeit vor 43 000 bis 22 000 Jahren, also aus dem Paläolithikum.

Aus den Daten konnte auf die Abfolge und Herkunft der, in Europa heimischen und zugewanderten Anteile, für die heute in Europa existierenden Genome, geschlossen werden.

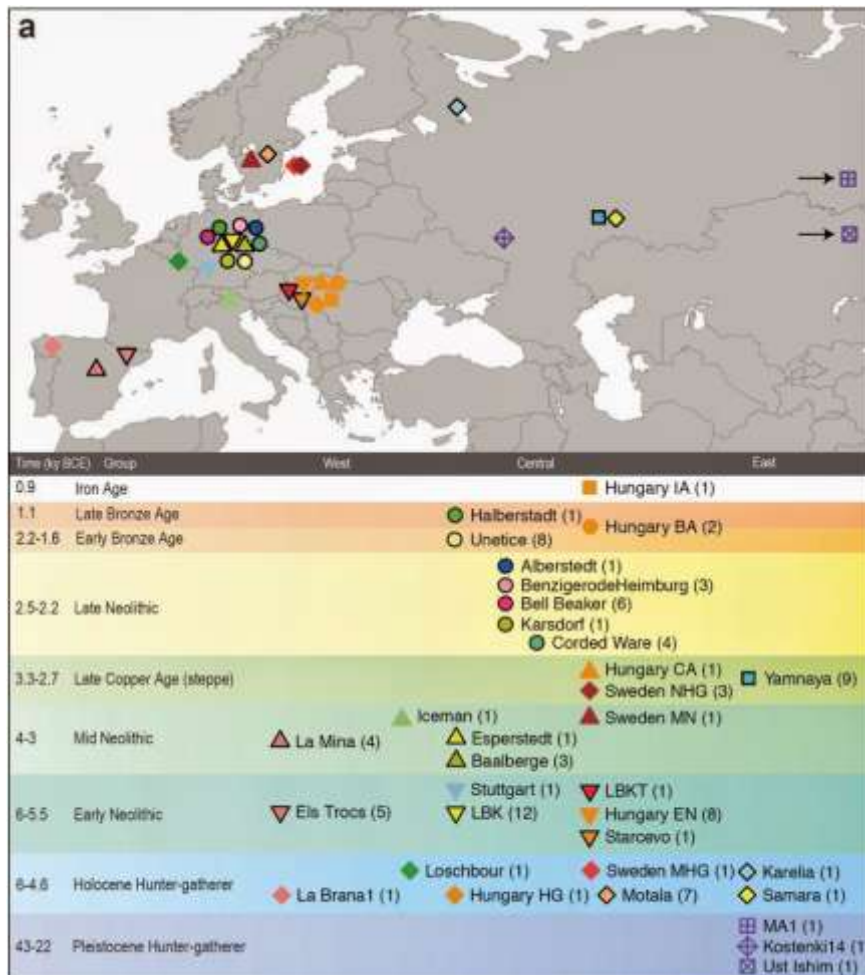
Hier eine Darstellung der paternalen Haplogruppen Europas in den verschiedenen Zeitaltern. Man erkennt deutlich, dass die Bevölkerungen nicht kontinuierlich aus einander hervorgehen.



Zum Vergleich eine Darstellung der bei Ausgrabungen gewonnenen autosomalen Daten, in der Form eines ADMIXTURE-Diagramms mit K=16. Die Daten zeigen bis zu drei verschiedene Bevölkerungseinflüsse für jede europäische Population. Nur die drei paläolithischen Proben aus Sibirien (ganz rechts) zeigen starke zusätzliche Einflüsse.



Die Karte unten zeigt die Ausgrabungsstellen und darunter die zeitliche Einordnung der jeweiligen Skelettfunde.



Als Fazit für die autosomale Genealogie möchte ich bemerken, dass sie durchaus hervorragende Ergebnisse bei den frühen Bevölkerungsbewegungen in Europa und der Aufklärung der Frühgeschichte zeigt.

Eine persönliche Anwendung, wie sie von manchen Vorreitern zur detaillierten Suche nach Vorfahren und deren Genom genutzt wird, scheint mir jedoch übertrieben und viel zu aufwändig. Die Ergebnisse sind zwar teilweise beeindruckend, es hängt jedoch vom Einzelfall und der Menge getesteter Personen ab, ob der Erfolg die Hoffnungen rechtfertigt.

Die Zukunft der autosomalen Genealogie scheint mir im wissenschaftlichen Bereich die besten Erfolge zu zeitigen und ich bin gespannt, auf die noch folgenden Erkenntnisse zur Vergangenheit der im Dunkeln liegenden Geschichte Europas.