



LANGELIN - LANguages Genes LINeages
ERC Advanced Grant n. 295733

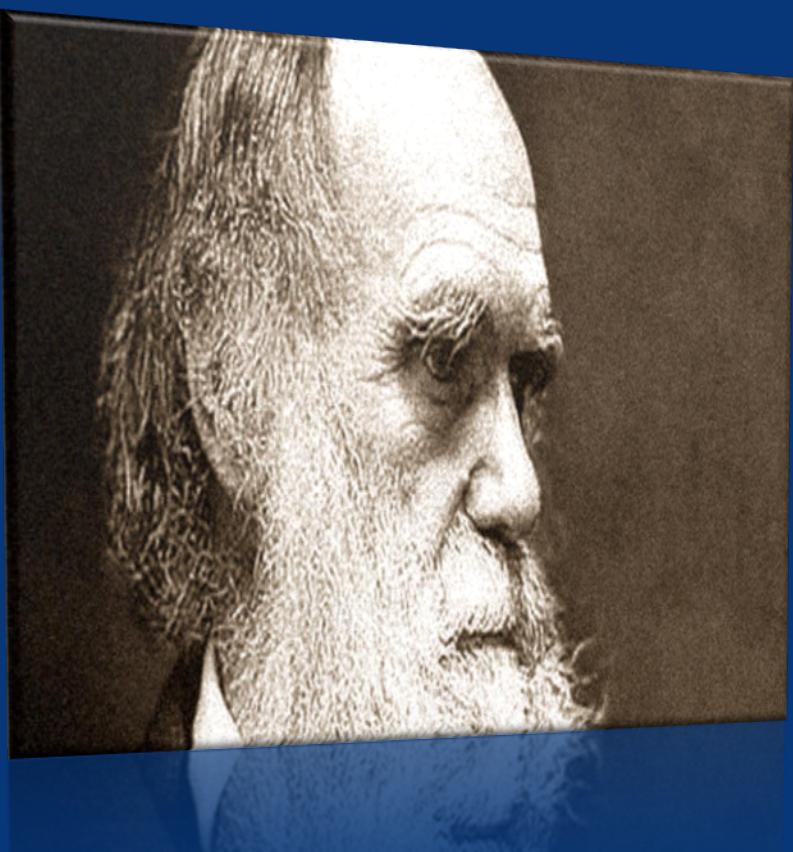


Funded by
the European Union

**Across language families:
Genome mirrors linguistic variation within Europe**

**Giuseppe Longobardi, Cristina Guardiano, Andrea Ceolin,
Silvia Ghirotto, Guido Barbujani**





Darwin's last challenge



*If we possessed a perfect pedigree of mankind, a genealogical arrangement of the races of man would afford **the best classification of the various languages now spoken throughout the world**; and if all extinct languages, and all intermediate and slowly changing dialects, were to be included, such an arrangement would be the only possible one*

Genes mirror geography within Europe

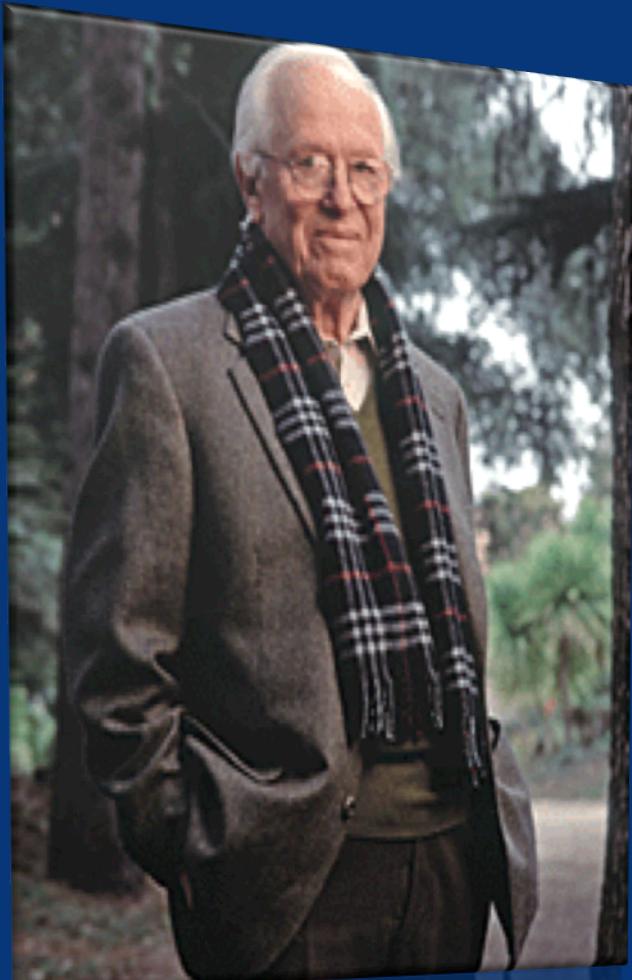
See associated Correspondence: [Detours, Nature 455, 861 \(October 2008\)](#)

John Novembre^{1,2}, Toby Johnson^{4,5,6}, Katarzyna Bryc⁷, Zoltán Kutalik^{4,6}, Adam R. Boyko⁷, Adam Auton⁷, Amit Indap⁷, Karen S. King⁸, Sven Bergmann^{4,6}, Matthew R. Nelson⁸, Matthew Stephens^{2,3} & Carlos D. Bustamante⁷

*...among Europeans,
we find a close correspondence between genetic and
geographic distances ...*

*...a geographical map of Europe
arises naturally as an efficient two-dimensional
summary
of genetic variation in Europeans*

Cavalli Sforza, Menozzi, and Piazza (1994)
The History and Geography of Human Genes



*We believe that the
major breakthrough
in the study of human variation has
been
the introduction of genetic markers,
which are strictly inherited and
basically immune to the problem of
rapid changes induced
by the environment*

Population Genetics as a Model

- **Biology**

Reduction of human ***genetic variation*** to the complex effects of a **finite** and **universal** set of **discrete biological** options
(genetic polymorphisms)

- **Linguistics**

Reduction of the whole space of possible ***grammatical variation*** to the complex effects of a **finite** and **universal** set of **discrete biological** options (parameters, i.e. syntactic polymorphisms)

объектов (параметров, т.е. синтаксических полиморфизмов)

Parametric Comparison Method

Longobardi (2003),
Guardiano and Longobardi
(2005),
Longobardi and Guardiano
(2009),
Longobardi, Guardiano, et
al. (2013)

Parameter values

may appropriately act as *comparanda*
for historical reconstruction

It becomes possible:

- ◆ to precisely calculate the syntactic distance between any two languages
- ◆ to assess the probabilistic value of such distances

IE parameters: Longobardi et al (2013)

	TABLE A	Sic	Cal	It	Sal	Sp	Fr	Ptg	Rm	Grl	BoG	Grk	E	D	Da	Ice	Nor	Big	SC	Slo	Po	Rus	Ir	Wel	Far	Ma	Hi	
1	± gramm. person	FGP	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	x	
2	± gramm. number	+FGP	FGN	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	2	
3	± gramm. gender	+FGN	FGG	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	3	
4	± NP over D	+FGP	NOD	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4		
5	± feature spread to N	+FGN	FSN	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	5		
6	± numb. on N (BNs)	+FSN	FNN	+	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	6		
7	± gramm. partial def	DGP	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	7		
8	± gramm. def	+DGP	DGR	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	0	0	0	0	0	+	+	0	8	
9	± strong person	+FGP, +DGR, -NOD	NSD	+	+	+	+	+	+	+	+	+	+	+	-	-	-	0	0	0	0	0	-	0	0	9		
10	± free null partitive Q	+FNN	DPQ	-	-	-	-	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	x0		
11	± gramm. dist. art.	-FSN or -FNN or +DGR	DDA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0	0	0	0	-	0	0	xi		
12	± def-checking N	+DGR	DCN	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0	0	0	0	-	0	0	x2		
13	± def spread to N	+DCN, -NSD	DSN	0	0	0	0	0	0	0	0	0	0	0	0	-	-	0	0	0	0	0	0	0	0	x3		
14	± def on relatives	+DGR	DOR	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0	0	0	0	-	0	0	x4		
15	± D-controlled infl. on N	+FSN	DIN	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	x5		
16	± plural spread from cardinals	+FSN	CPS	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	x6		
17	± gramm. boundedness	CGB	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	x7		
18	± strong article	+DGR, +FNN, -CGB	CGR	+	+	+	+	+	0	+	+	+	+	+	+	+	+	0	0	0	0	0	-	0	0	x8		
19	± bounded.-checking N	+CGB	CCN	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-	x9		
20	± null-N-licensing art	-FSN or -FNN or -DCN, +NOD or +NSD	DNN	-	-	-	+	-	0	-	-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	
21	± structured APs	AST	AST	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	21	
22	± feature spread to struct. APs	+FSN, +AST	FFS	+	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	+	22	
23	± feature spread to pred. APs	+FGN	FSP	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	+	23	
24	± D-controlled infl. on A	-NSD, +FFS	ADI	0	0	0	0	0	0	0	0	0	0	0	0	+	+	0	0	0	0	0	-	0	0	0	24	
25	± DP over relatives	ADR	ADR	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	25	
26	± relative extrap.	-ADR	AER	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	+	26	
27	± free reduced rel	+AST	ARR	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	+	27	
28	± N-raising with obl. pied-piping	+AST	NPP	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	28		
29	± free Gen	GFR	GFR	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	+	29	
30	± uniform Gen	+GFR	GUN	-	-	-	-	-	-	0	0	0	0	-	-	-	-	0	0	0	0	-	-	-	-	30		
31	± DP over free Gen	+GFR, +ADR	GPR	+	+	+	+	+	+	0	0	0	0	+	+	+	+	0	0	0	0	+	+	+	0	31		
32	± GenO	=+GUN	GFO	-	+	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	32		
33	± Gen-feature spread to N	GFS	GFS	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	33		
34	± D checking poss.	+DGR, +NSD or =+CGR	PDC	-	-	-	-	+	+	?	-	-	-	0	0	0	+	0	-	0	0	0	0	+	+	0	0	34
35	± adjectival poss.	APO	APO	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	35		
36	± post-affix poss.	+DCN	PAP	0	0	0	0	0	0	0	+	0	0	0	-	-	-	+	0	0	0	0	0	0	0	36		
37	± clitic poss.	PCL	PCL	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	37		
38	± N-feat. spr. to pron. poss.	+FFS or +AST, +PAP or +PCL	FHS	0	0	0	0	0	0	0	+	-	-	0	0	0	+	+	0	0	0	0	0	0	0	38		
39	± N-feature spread to free Gen	+FFS, +GFR, -PHS	GSP	-	-	-	-	-	-	0	0	0	0	-	-	-	-	0	0	0	0	-	0	0	+	39		
40	± adjectival Gen	+APO	AGE	-	-	-	-	-	-	0	0	0	0	-	0	0	0	+	+	-	+	0	0	0	0	40		
41	± Poss²-checking N	-GFS	GCN	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	41		
42	± Loc Checking Dem	=+CGR, -FSN or +FNN	TLC	0	0	0	0	0	0	0	0	0	0	0	0	0	+	0	+	+	+	+	+	+	+	42		
43	± Split Locality	=-TLC	TSL	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	43		
44	± D Checking Dem	=-TLC, -FSN or +DGR	TDC	+	+	+	+	+	+	+	+	+	?	-	-	-	-	-	-	-	-	-	-	-	-	0	44	
45	± N over cardinals	NOC	NOC	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	45		
46	± N over ordinals	-NO	NOO	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	46		
47	± N over M ₁ As	-NOO, -NPP	NM ₁	+	+	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	47		
48	± N over M ₂ As	-NM ₁	NM ₂	0	0	+	0	+	+	+	+	0	0	-	-	-	-	-	-	-	-	-	0	0	-	48		
49	± N over As	-NM ₂	NOA	0	0	0	0	0	0	0	0	0	0	-	-	-	-	-	-	-	-	-	-	0	0	-	49	
50	± N over GenO	=-GFO, -NOA or -AST	NGO	0	0	0	0	0	0	0	0	0	0	+	0	+	0	0	+	+	+	0	0	0	0	50		
51	± N over ext. arg.	-NGO or (-GFO, -NOA or -AST)	NOE	0	0	0	0	0	0	0	0	0	0	0	-	0	-	0	-	+	0	0	0	0	+	51		
52	± free MOD	-NOA	AFM	0	0	0	0	0	0	0	0	0	0	0	+	-	-	-	-	-	-	-	-	0	0	-	52	
53	± class MOD	-AFM	ACM	0	0	0	0	0	0	0	0	0	0	0	-	-	-	-	-	-	-	-	-	-	-	53		
54	± def on APs	+DGP, +postnom. APs	DOA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0	0	0	0	-	-	0	0	54	
55	± gramm. AP marker	+postnom. APs	DMO	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0	0	0	0	-	-	+	0	0	55
56	± Cons. Pr. (-NM ₁ , +A-Cpl) or (+NPP or -NM ₂ , +Cpl-A)	ACP	0	0	+	0	+	+	+	+	0	0	-	-	-	-	-	-	-	-	-	-	-	0	0	56		

The structure of linguistic distances

A problem: internal implications

GRAMMAR:

PCM = designed to control for the non-independence of characters.

- Explicit hypotheses about implications among syntactic parameters.
- D_{SYN} : normalized Hamming (or Jaccard) distance = $d/(i+d)$

IE cognates: Bouckaert et al. (2012)

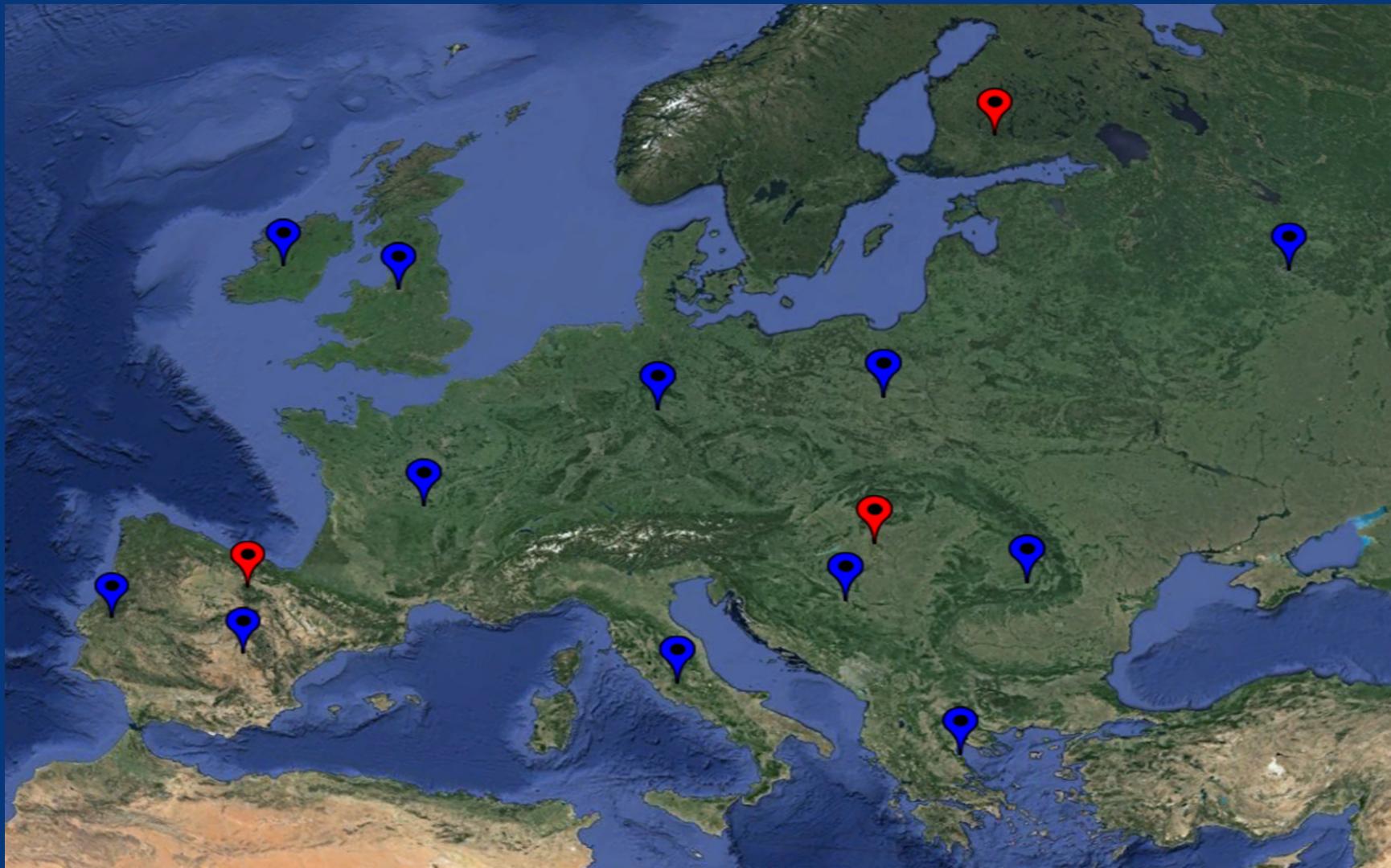
- Expansion/revision of Dyen et al.'s (1992) list of IE cognate words
- Lists actual roots for each meaning (no 1:1 correspondence between meaning and word)
 - Codes for **synonymy**
- Richer device for quantitative experiments on IE lexical diversification

LEXICAL DISTANCES

	English	German	
<i>fish</i>	+ (fish)	+ (Fisch)	1 identity
<i>breathe</i>	+ (breathe)	-	1 difference
	-	+ (atmen)	1 difference

- $D_{LEX} : d/(i+d) = 2/3$
- $D_{LEX} : \text{differences} = 0.5 = 1/2$

Europe across language families



Syntactic and Lexical distances

	E	Fr	D	Grk	Ir	It	Po	Ptg	Rm	Rus	SC	Sp
E	0	0.1842	0.0732	0.3	0.1111	0.15	0.2424	0.15	0.2105	0.2424	0.2188	0.175
Fr	0.7193	0	0.1579	0.2895	0.1944	0.0476	0.2333	0.0488	0.1622	0.2667	0.2759	0.0714
D	0.38847	0.73272	0	0.2564	0.1111	0.125	0.1471	0.125	0.25	0.1765	0.1471	0.15
Grk	0.82439	0.82432	0.81124	0	0.2571	0.225	0.1818	0.2564	0.2	0.1818	0.1875	0.25
Ir	0.81074	0.81176	0.81221	0.86456	0	0.2105	0.2143	0.1892	0.2571	0.2143	0.2222	0.1842
It	0.73012	0.23111	0.72889	0.81345	0.80543	0	0.1935	0.0233	0.1026	0.2258	0.2333	0.0682
Po	0.72816	0.77629	0.74944	0.84649	0.83066	0.77489	0	0.1935	0.2333	0.0278	0.0571	0.2258
Ptg	0.76233	0.3431	0.75884	0.83903	0.84615	0.31313	0.78049	0	0.1053	0.2258	0.2333	0.0233
Rm	0.75799	0.44186	0.75476	0.82609	0.84052	0.39468	0.79835	0.44788	0	0.2333	0.2414	0.1026
Rus	0.7201	0.75234	0.74766	0.83562	0.81862	0.75676	0.31519	0.76371	0.77302	0	0.0286	0.2581
SC	0.73298	0.77458	0.76978	0.82028	0.83092	0.76498	0.33488	0.77778	0.78462	0.309	0	0.2667
Sp	0.75	0.30193	0.75589	0.82809	0.82895	0.27122	0.775	0.22505	0.44664	0.76139	0.77728	0

$$r = 0.850$$

$$p = 0.001$$

Genetic Data

ARTICLE

The Population Reference Sample, POPRES: A Resource for Population, Disease, and Pharmacological Genetics Research

Matthew R. Nelson,^{1,*} Katarzyna Bryc,² Karen S. King,¹ Amit Indap,² Adam R. Boyko,² John Novembre,^{3,4} Linda P. Briley,¹ Yuka Maruyama,¹ Dawn M. Waterworth,⁵ Gérard Waeber,⁶ Peter Vollenweider,⁶ Jorge R. Oksenberg,⁷ Stephen L. Hauser,⁷ Heide A. Stirmadel,⁸ Jaspal S. Konner,⁹ John C. Chambers,¹⁰ Brendan Jones,¹ Vincent Mooser,⁵ Carlos D. Bustamante,² Allen D. Roses,¹ Daniel K. Burns,¹ Margaret G. Ehrn,¹ and Eric H. Lai¹

5,886 subjects genotyped at 500,568 loci using the Affymetrix 500K single nucleotide polymorphism (SNP) chip.

Basque

OPEN  ACCESS  Fully available online

PLOS GENETICS

Genomic Ancestry of North Africans Supports Back-to-Africa Migrations

Brenna M. Henn^{1*}, Laura R. Botigué^{2,3}, Simon Gravel¹, Wei Wang³, Abra Brisbin³, Jake K. Byrnes¹, Karima Fadhloui-Zidi⁴, Pierre A. Zalloua⁵, Andres Moreno-Estrada¹, Jaume Bertranpetti², Carlos D. Bustamante^{1,6}, David Comas^{2,7}

¹ Department of Genetics, Stanford University, Stanford, California, United States of America, ² Institute of Evolutionary Biology (CSIC UPF), Universitat Pompeu Fabra, Barcelona, Spain, ³ Department of Biological Statistics and Computational Biology, Cornell University, Ithaca, New York, United States of America, ⁴ Laboratory of Genetics, Immunology, and Human Pathology, University Tunc El Manar, Tunis, Tunisia, ⁵ The Lebanese American University, Beirut, Lebanon

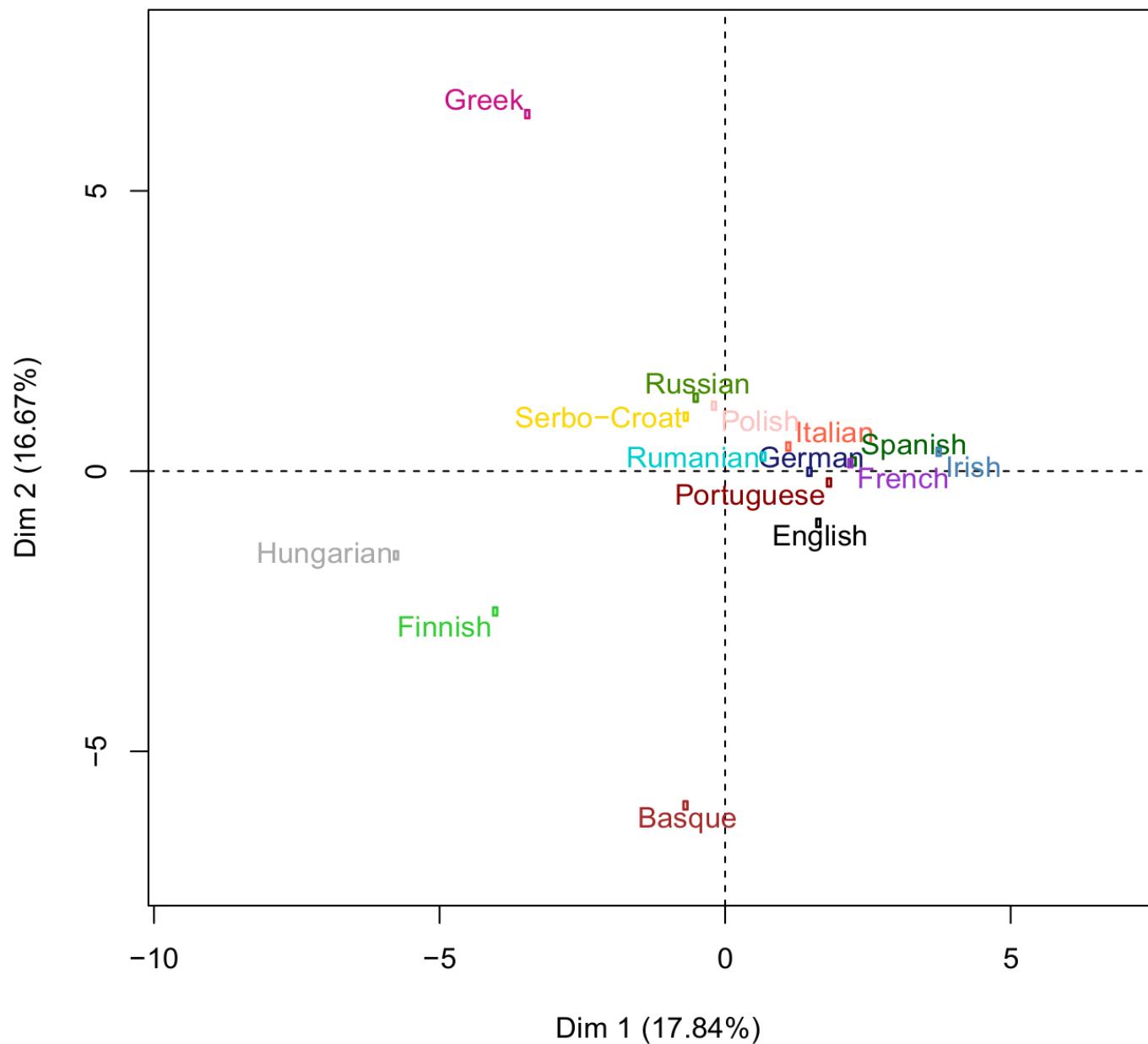


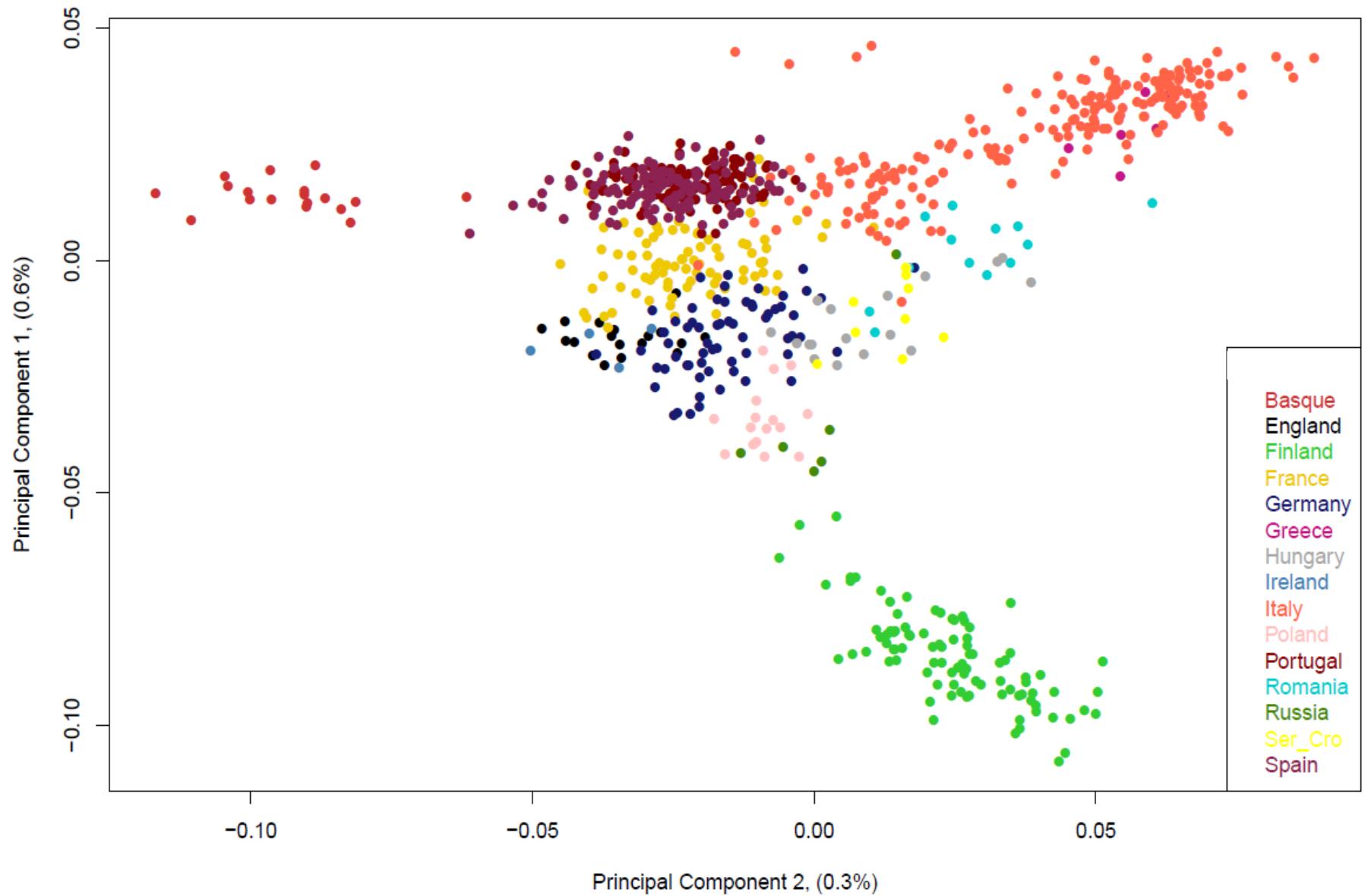
Finnish



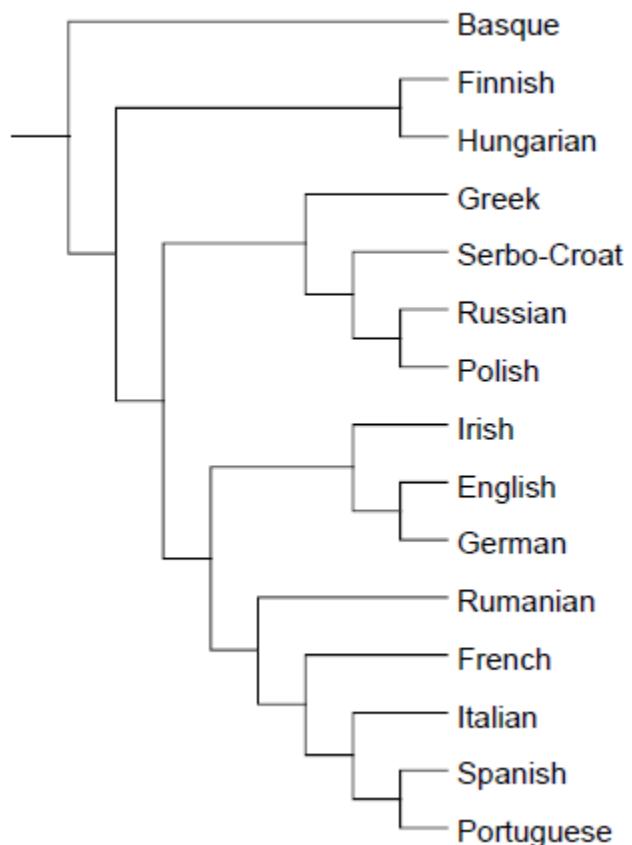
The screenshot shows the homepage of the 1000 Genomes Project. It features a search bar for "Search 1000 Genomes" and a "Start Browsing 1000 Genomes data" section. Below this is a "Browsing results Browser interface" section. The right side of the page contains a "The 1000 Genomes Browser" sidebar with links to "About the 1000 Genomes Project", "Data Sources", "Data Release", "Sample Details", "Sample Browser", and "Help". The footer includes copyright information and links to "1000 Genomes", "Contact Us", and "Privacy Policy".

Individuals factor map (PCA)

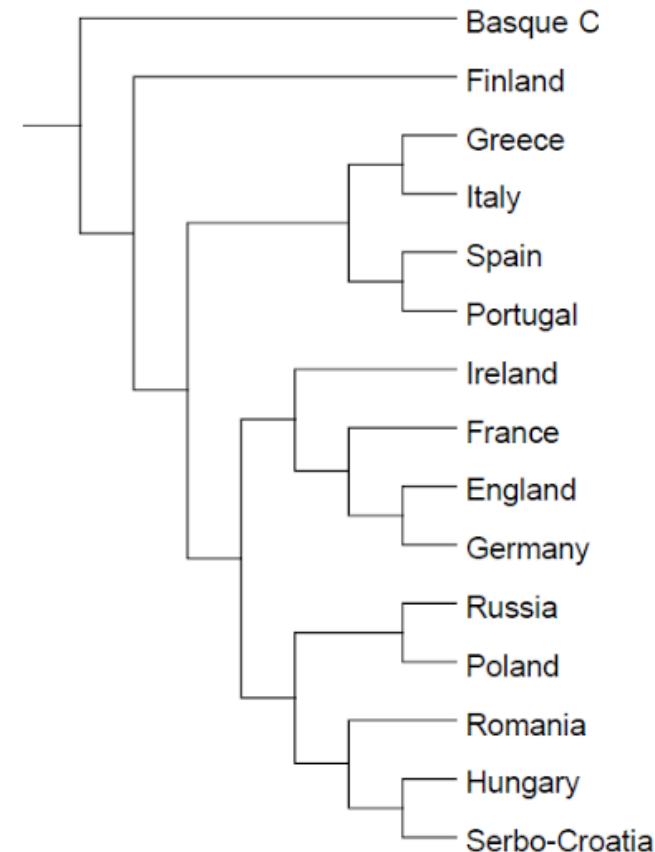




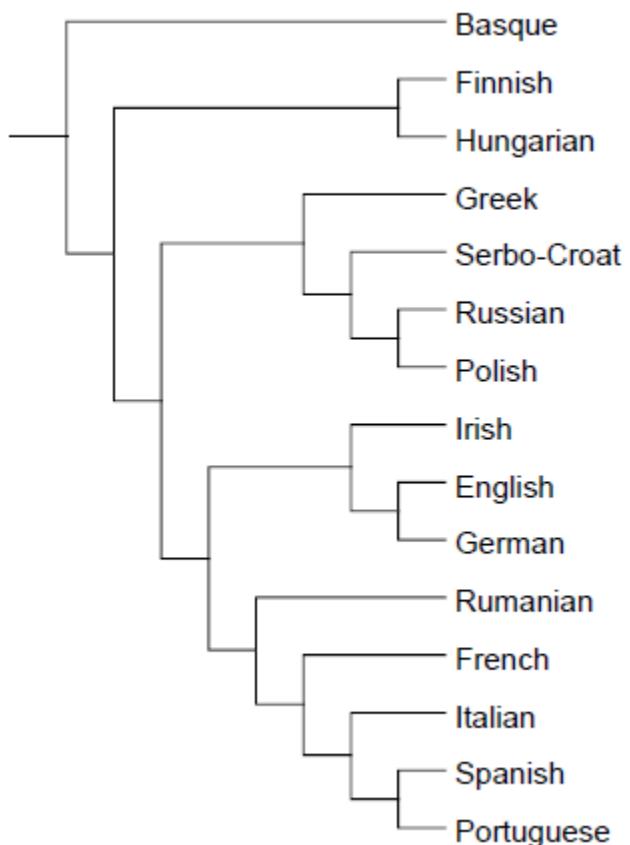
Syntactic distances



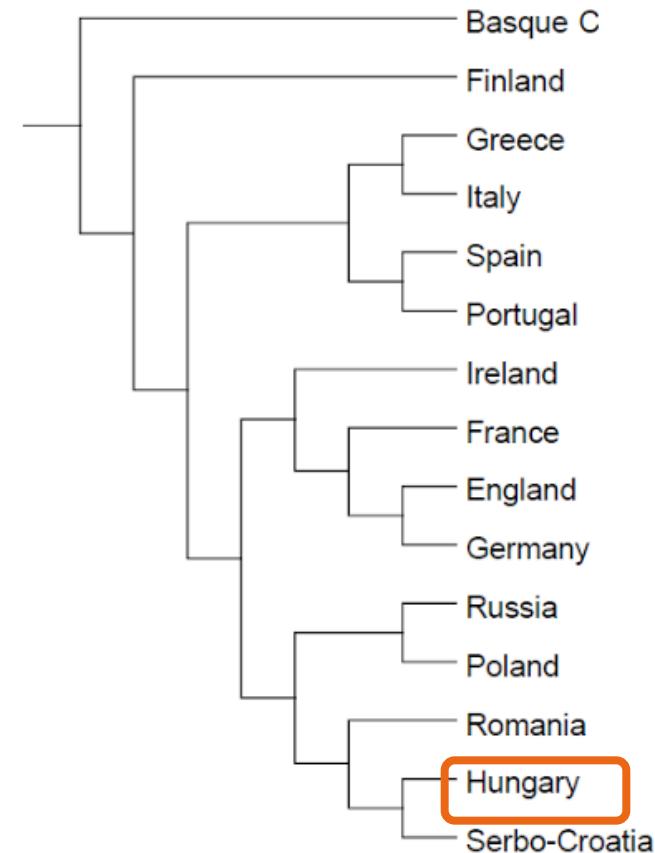
Genetic distances

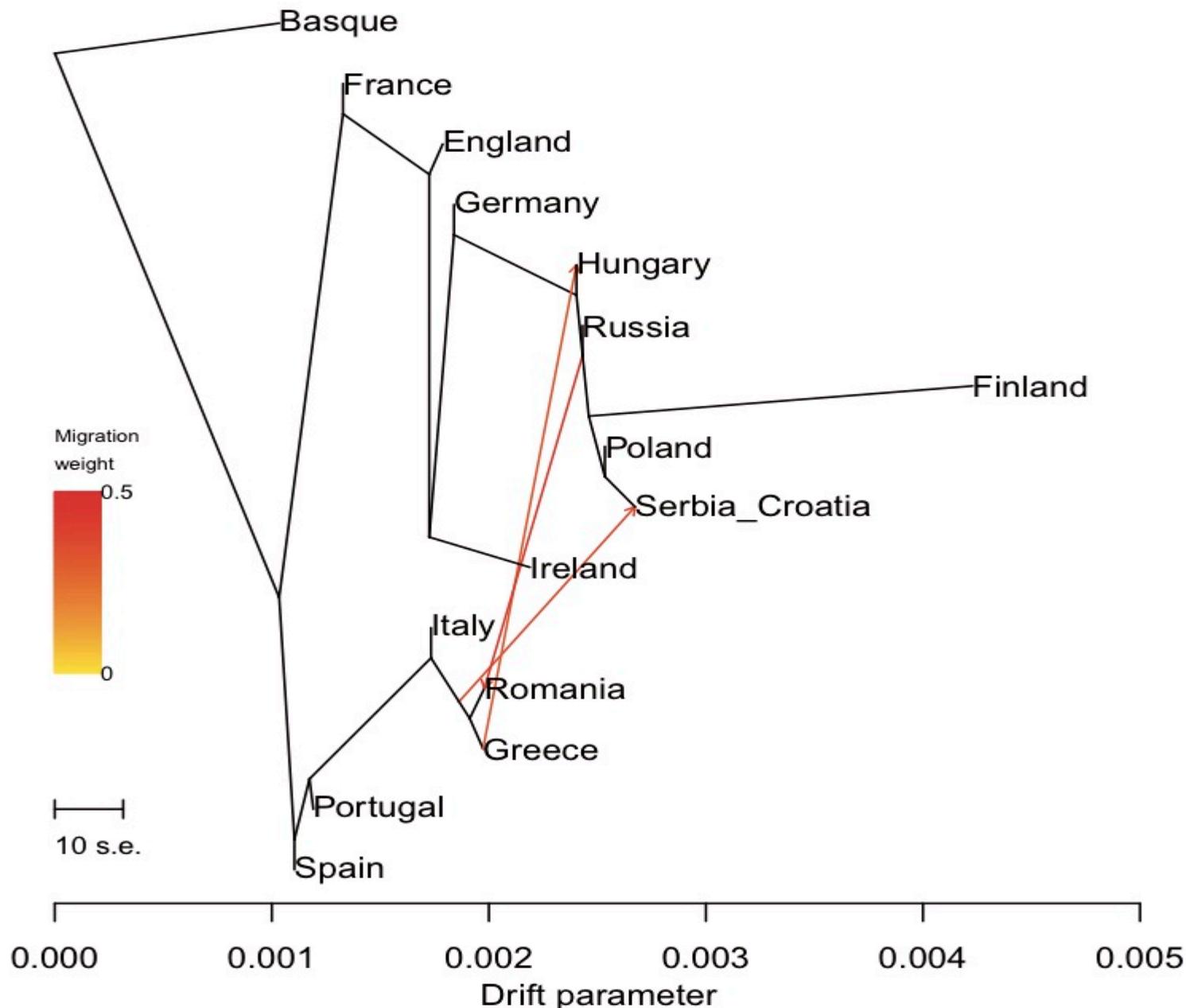


Syntactic distances



Genetic distances

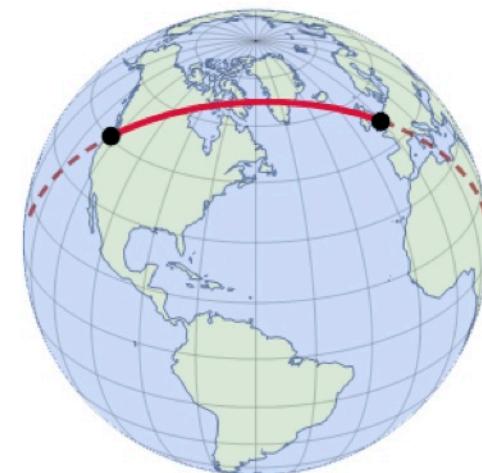


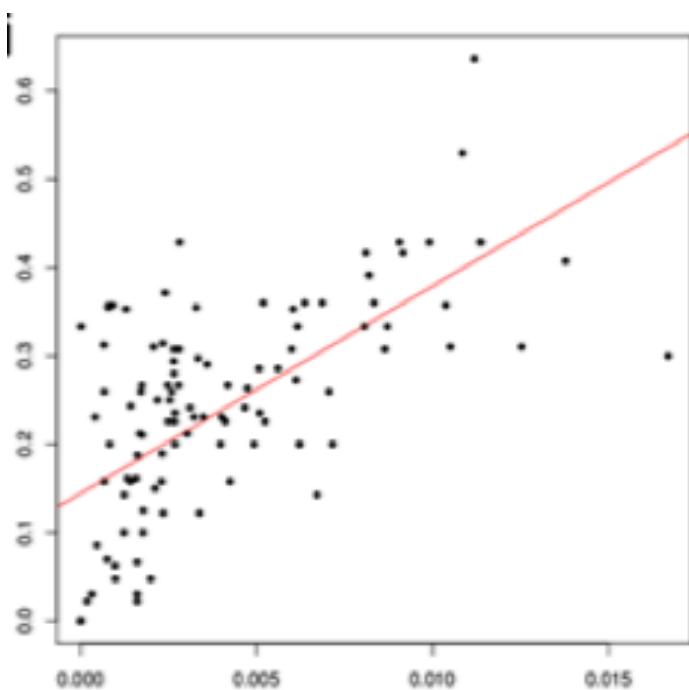
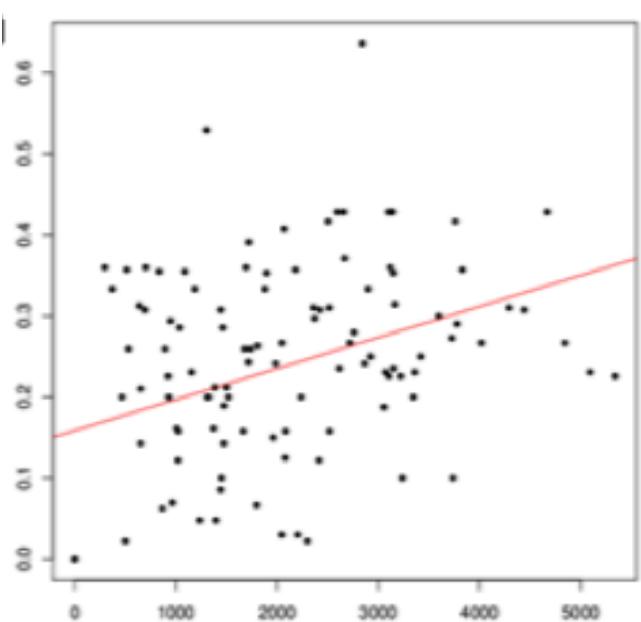
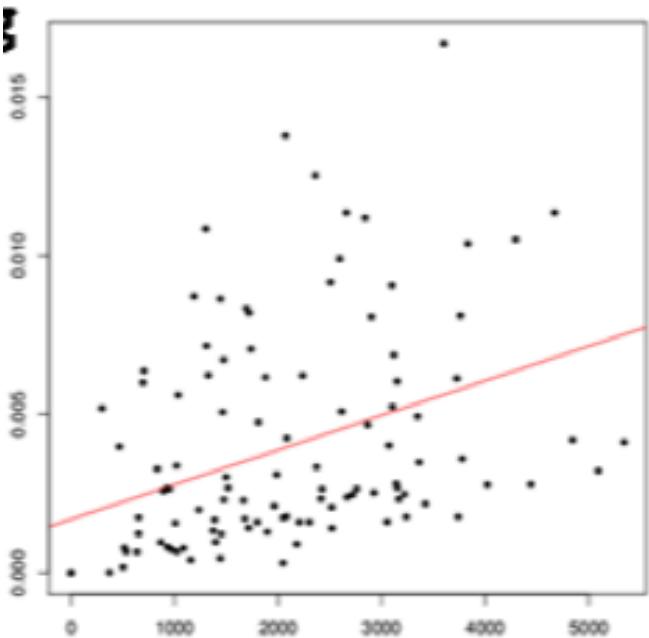


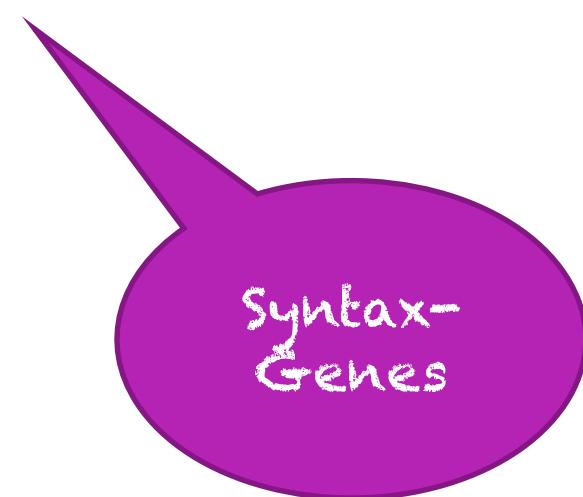
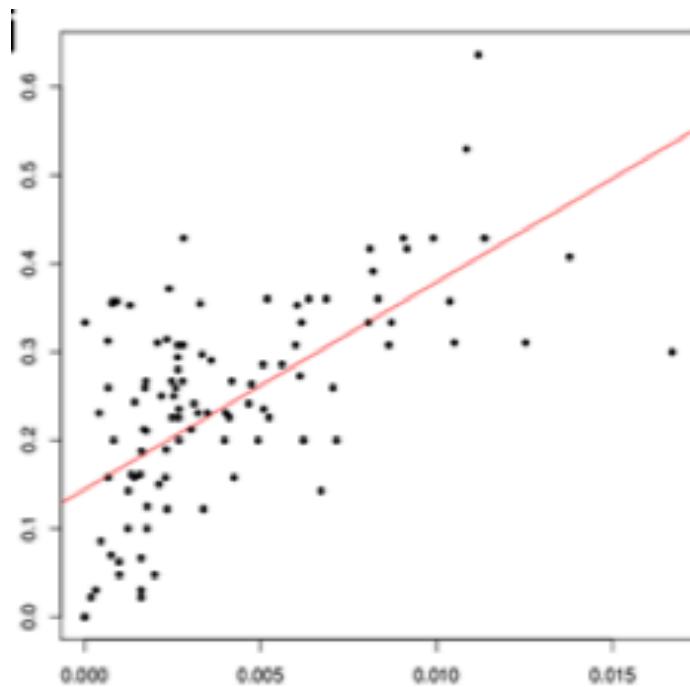
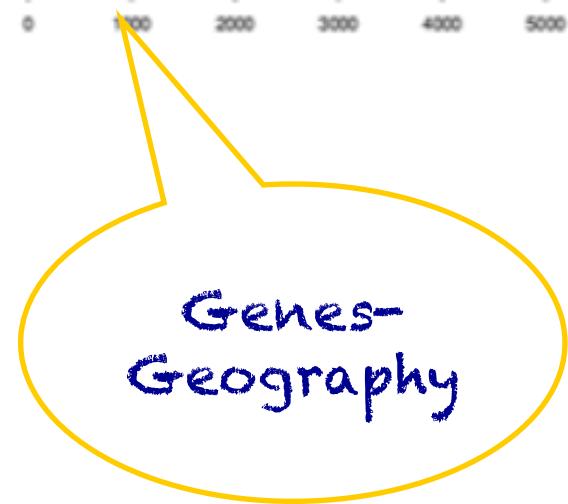
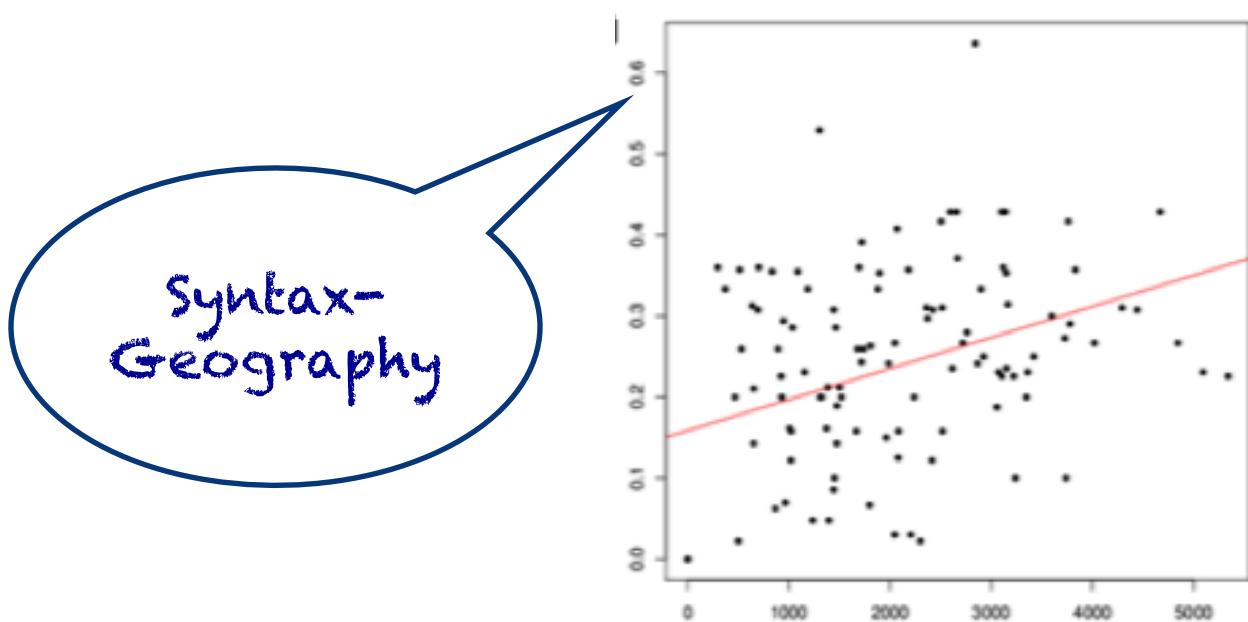
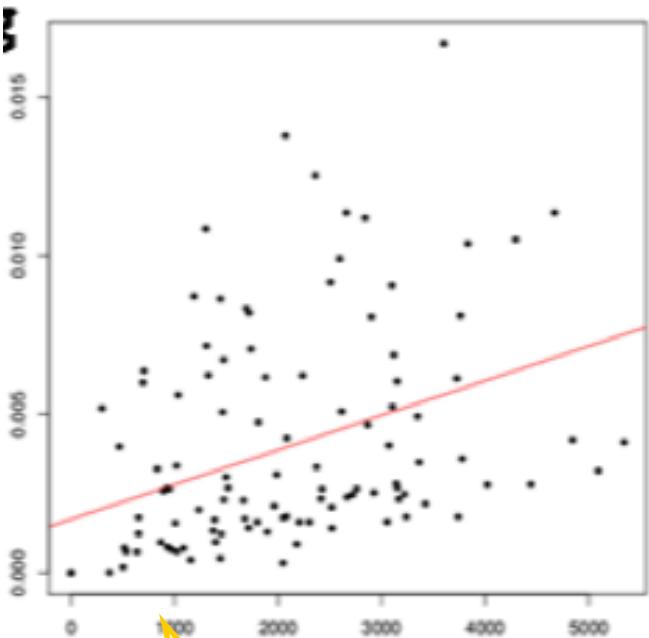
Geographic distances

Great Circle Distances (the shortest distance between two points on the surface of a sphere)

	Basque	England	Finland	France	Germany	Greece	Hungary	Ireland	Italy	Poland	Portugal	Romania	Russia	Ser_Cro	Spain
Basque	0														
England	1187.32	0													
Finland	3598.33	3116.53	0												
France	703.27	930.48	2900.12	0											
Germany	1720.55	1441.89	1879.06	1021.06	0										
Greece	2842.08	3148.58	2070.09	2370.7	1808.53	0									
Hungary	2504.1	2515.3	1472.55	1894.75	1086.01	833.19	0								
Ireland	1301.76	652.09	3760.08	1386.19	2081.45	3726.67	3140.04	0							
Italy	1692.39	2083.82	2359.24	1234.44	1019.94	1155.1	948.34	2614.46	0						
Poland	2659.05	2423.91	1035.1	1986.03	1007.66	1309.18	514.5	3072.68	1325.44	0					
Portugal	693.88	1665.57	4292.19	1395.39	2414.33	3420.15	3164.26	1498.94	2299.75	3346.22	0				
Romania	3099.24	3151.21	1442.74	2518.15	1715.71	652.92	637.63	3777.63	1450.7	892.46	3738.87	0			
Russia	4669.32	4440.39	1463.4	4020.28	3053.41	2236.19	2181.1	5092.48	3103.42	2046.74	5340.66	1678.13	0		
Ser_Cro	2594.25	2759.21	1739.13	2048.29	1372.57	466.32	370.15	3361.71	924.19	868.04	3220.66	531.06	2204.72	0	
Spain	298.17	1473.97	3830.7	965.2	1962.6	2924.66	2666.35	1517.52	1798.83	2864.87	501.42	3237.63	4845.04	2719.27	0







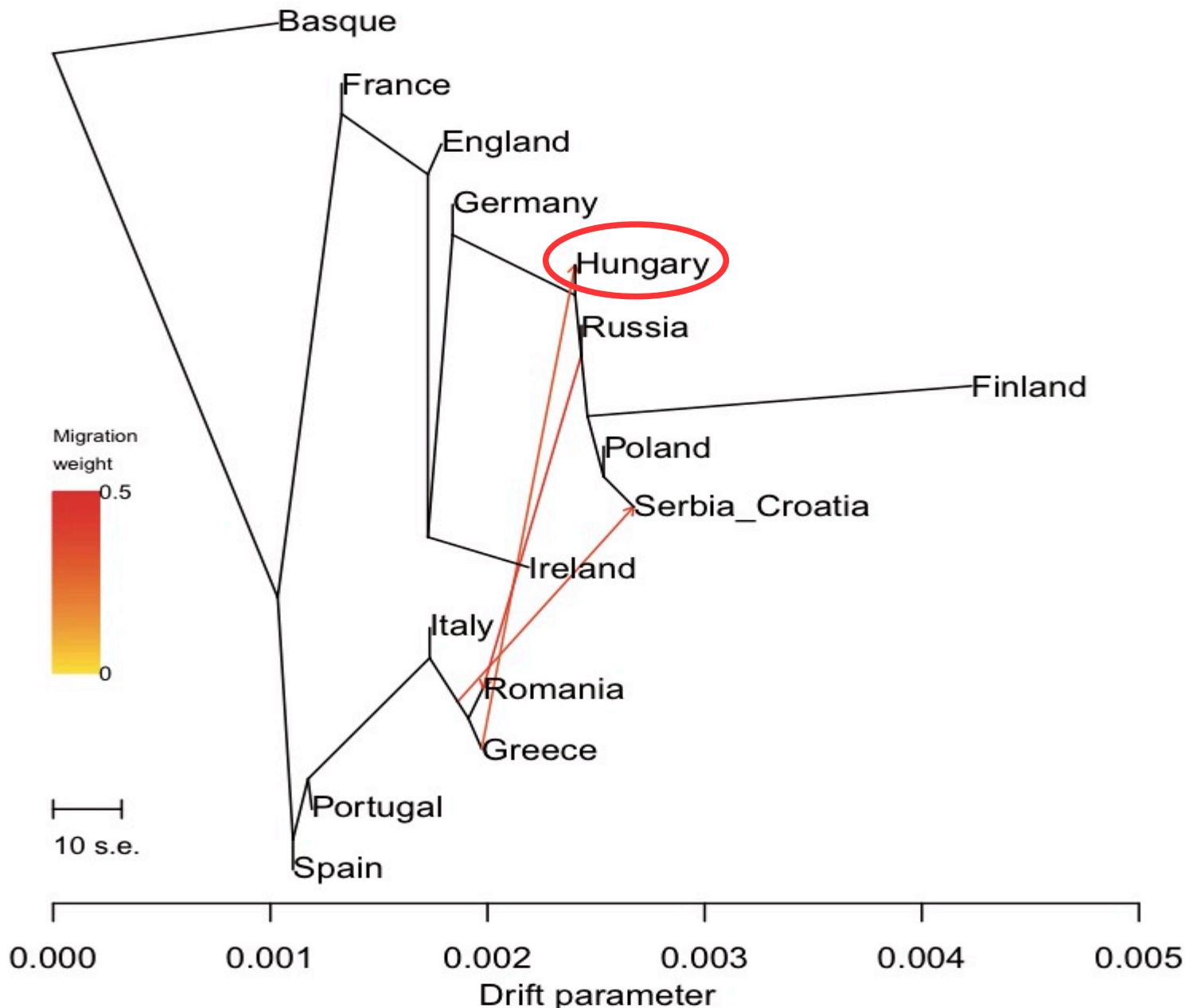
Syntax-
Geography

Genes-
Geography

Syntax-
Genes

Distance matrices	<i>r</i>	<i>P</i>
$d_{\text{GEN}} d_{\text{GEO}}$ Genetic - Geographic	0.299	0.030
$d_{\text{SYN}} d_{\text{LEX}}$ Syntactic - Lexical	0.850	0.001
$d_{\text{SYN}} d_{\text{GEO}}$ Syntactic - Geographic	0.240	0.039
$d_{\text{LEX}} d_{\text{GEO}}$ Lexical - Geographic	0.084	0.264
$d_{\text{SYN}} d_{\text{GEN}}$ Syntactic - Genetic	0.599	0.001
$d_{\text{LEX}} d_{\text{GEN}}$ Lexical - Genetic	0.537	0.001
$d_{\text{GEN}} d_{\text{GEO}} (d_{\text{SYN}})$ Genetic - Geographic (Syntax held constant)	0.200	0.114
$d_{\text{GEN}} d_{\text{GEO}} (d_{\text{LEX}})$ Genetic - Geographic (Lexicon held constant)	0.302	0.035
$d_{\text{SYN}} d_{\text{GEO}} (d_{\text{GEN}})$ Syntactic - Geographic (Genetics held constant)	0.079	0.264
$d_{\text{LEX}} d_{\text{GEO}} (d_{\text{GEN}})$ Lexical - Geographic (Genetics held constant)	-0.095	0.736
$d_{\text{SYN}} d_{\text{GEN}} (d_{\text{GEO}})$ Syntactic - Genetic (Geography held constant)	0.570	0.002
$d_{\text{LEX}} d_{\text{GEN}} (d_{\text{GEO}})$ Lexical - Genetic (Geography held constant)	0.538	0.001

Distance matrices	<i>r</i>	<i>P</i>
$d_{\text{GEN}} d_{\text{GEO}}$ Genetic - Geographic	0.299	0.030
$d_{\text{SYN}} d_{\text{LEX}}$ Syntactic - Lexical	0.850	0.001
$d_{\text{SYN}} d_{\text{GEO}}$ Syntactic - Geographic	0.240	0.039
$d_{\text{LEX}} d_{\text{GEO}}$ Lexical - Geographic	0.084	0.264
$d_{\text{SYN}} d_{\text{GEN}}$ Syntactic - Genetic	0.599	0.001
$d_{\text{LEX}} d_{\text{GEN}}$ Lexical - Genetic	0.537	0.001
$d_{\text{GEN}} d_{\text{GEO}} (d_{\text{SYN}})$ Genetic - Geographic (Syntax held constant)	0.200	0.114
$d_{\text{GEN}} d_{\text{GEO}} (d_{\text{LEX}})$ Genetic - Geographic (Lexicon held constant)	0.302	0.035
$d_{\text{SYN}} d_{\text{GEO}} (d_{\text{GEN}})$ Syntactic - Geographic (Genetics held constant)	0.079	0.264
$d_{\text{LEX}} d_{\text{GEO}} (d_{\text{GEN}})$ Lexical - Geographic (Genetics held constant)	-0.095	0.736
$d_{\text{SYN}} d_{\text{GEN}} (d_{\text{GEO}})$ Syntactic - Genetic (Geography held constant)	0.570	0.002
$d_{\text{LEX}} d_{\text{GEN}} (d_{\text{GEO}})$ Lexical - Genetic (Geography held constant)	0.538	0.001



Europe's diversity without Hungary

G. Tömöry, B. Csányi, E. Bogácsi-Szabó, T. Kalmár, A. Czibula, A. Csosz, K. Priskin, B. Mende, P. Langó, C. S. Downes, and I. Raskó. (2007) **Comparison of maternal lineage and biogeographic analyses of ancient and modern Hungarian populations.** *American Journal of Physical Anthropology* 134:354-368

Ancient DNA in Hungary shows a **predominance of European mitochondrial haplotypes** in burials attributed to the **lower classes**, and a **high incidence of Asian haplotypes in high-status individuals** of the tenth century, which points to the **Asian immigrants** as representing a **social elite**, rather than the bulk of the population

Distance matrices	<i>r</i>	<i>P</i>
$d_{GEN} d_{GEO}$ Genetic - Geographic	0.275	0.048
$d_{SYN} d_{LEX}$ Syntactic - Lexical	0.850	0.001
$d_{SYN} d_{GEO}$ Syntactic - Geographic	0.291	0.026
$d_{LEX} d_{GEO}$ Lexical - Geographic	0.152	0.144
$d_{SYN} d_{GEN}$ Syntactic - Genetic	0.740	0.001
$d_{LEX} d_{GEN}$ Lexical - Genetic	0.687	0.001
$d_{GEN} d_{GEO} (d_{SYN})$ Genetic - Geographic (Syntax held constant)	0.093	0.254
$d_{GEN} d_{GEO} (d_{LEX})$ Genetic - Geographic (Lexicon held constant)	0.238	0.083
$d_{SYN} d_{GEO} (d_{GEN})$ Syntactic - Geographic (Genetics held constant)	0.135	0.178
$d_{LEX} d_{GEO} (d_{GEN})$ Lexical - Geographic (Genetics held constant)	-0.053	0.615
$d_{SYN} d_{GEN} (d_{GEO})$ Syntactic - Genetic (Geography held constant)	0.717	0.001
$d_{LEX} d_{GEN} (d_{GEO})$ Lexical - Genetic (Geography held constant)	0.679	0.001

CONCLUSIONS

- Syntactic distances proved to correlate well with Lexical distances. This means that when evidence for the lexicon is unavailable (e.g. cross-family comparison) we can safely rely on syntactic information.
- Once precise measures are employed for measuring syntactic, genetic and geographic variation, the claim that geography is the best predictor for genetic variation in Europe reveals to be wrong.



THANKS!



Selected References:

- Barbujani, G. & Sokal, R.R. Zones of sharp genetic change in Europe are also linguistic boundaries. *Proc. Natl. Acad. Sci. USA* 87, 1816-1819 (1990).
- Bouckaert, R. et al. Mapping the origins and expansion of the Indo-European language family. *Science* 337, 957-960 (2012).
- Cavalli-Sforza, L. L., Menozzi, P., Piazza, A. *The History and Geography of Human Genes*. Princeton University Press, Princeton (1994).
- Chomsky, N. *Lectures on Government and Binding*. Foris, Dodrecht (1981).
- Darwin, C. *On the Origin of Species*. John Murray, London (1859).
- Longobardi, G., Guardiano, C. Evidence for syntax as a signal of historical relatedness. *Lingua* 119, 1679-1706 (2009).
- Longobardi, G., Guardiano, C., Silvestri, G., Boattini, A., Ceolin, A. Toward a syntactic phylogeny of modern Indo-European languages. *J. Histor. Linguistics* 3, 122-152 (2013).
- Novembre, J. et al. Genes mirror geography within Europe. *Nature* 456, 98-101 (2008).
- Tömöry, G., et al. Comparison of maternal lineage and biogeographic analyses of ancient and modern Hungarian populations." *Am. J. Phys. Anthropol.* 134, 354-368 (2007).