Genómica de la individualidad: ¿qué dice mi genoma de mí?

Francesc Calafell
Institut de Biologia Evolutiva (UPF - CSIC)
Departament de Ciències Experimentals i de la Salut
Universitat Pompeu Fabra

francesc.calafell@upf.edu

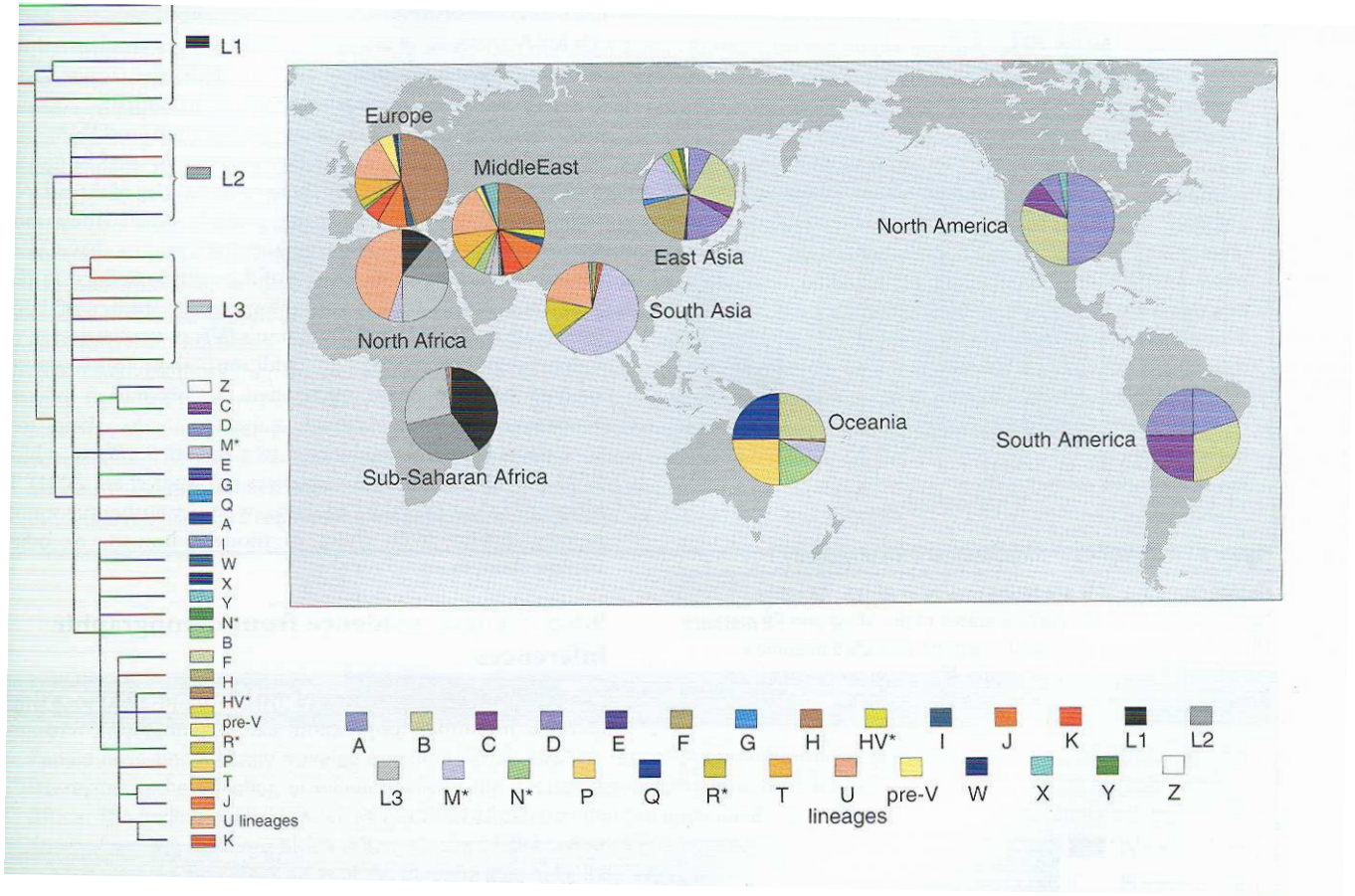
* "Ancestría"

* Fenotipo

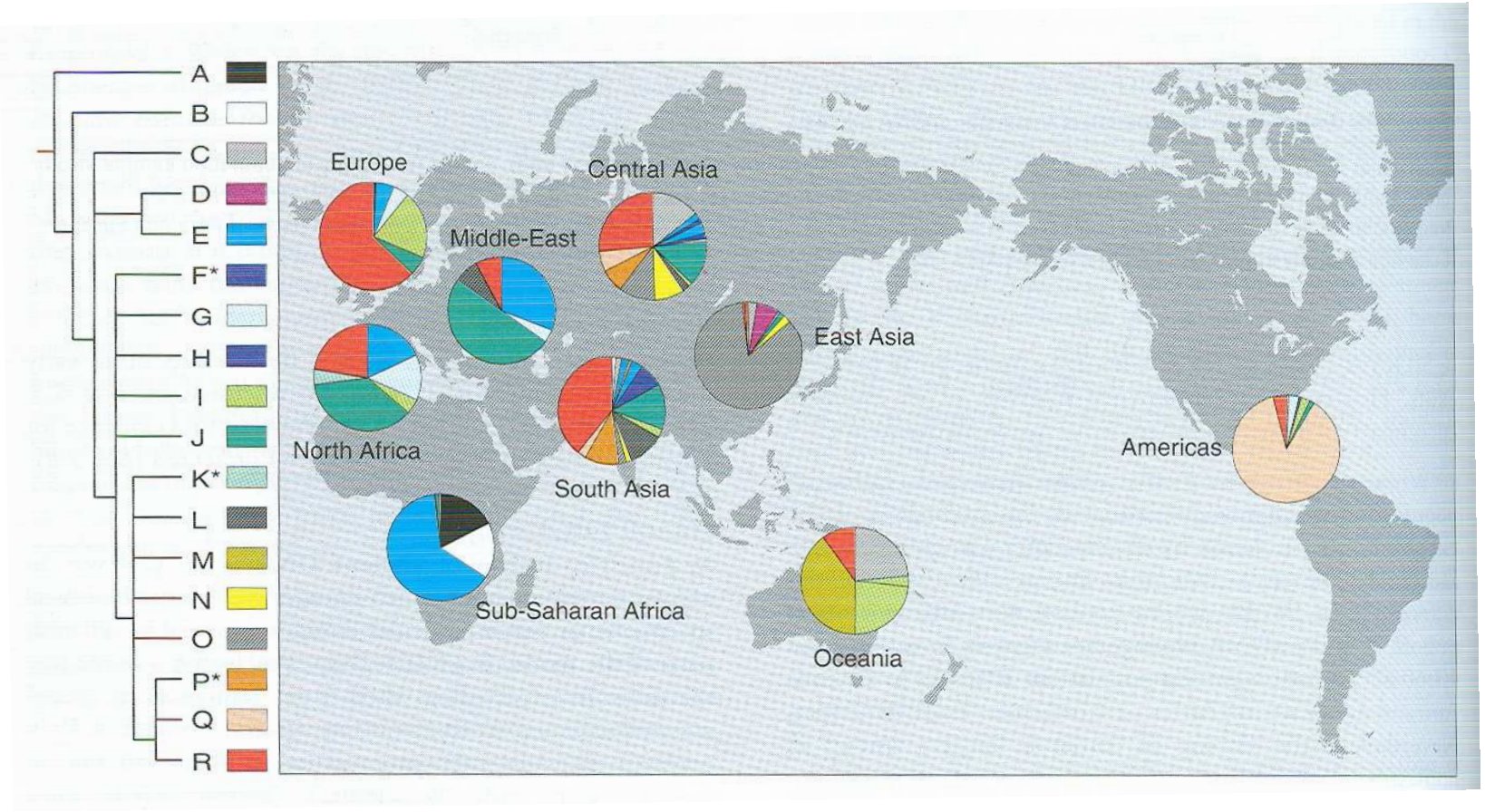
* Riesgos para enfermedades

Orígenes geográficos (*ancestría*) de mi genoma

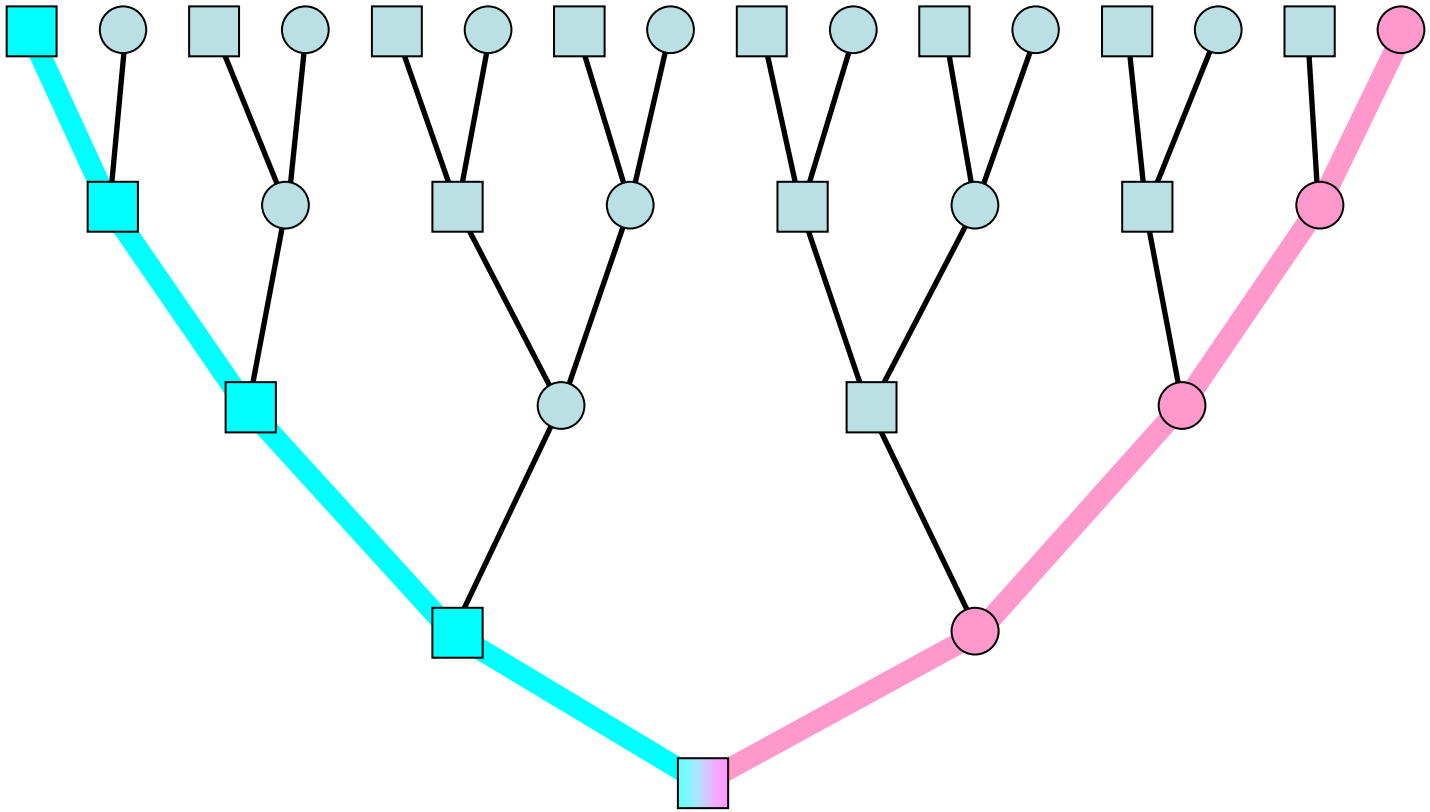
El ADN mitocondrial presenta una estructura geográfica muy marcada...



... y el cromosoma Y también



...pero estos marcadores representan sólo una pequeña parte de nuestro genoma, y de nuestros antepasados



Diversidad genética en Brasil

Diversidad genética en Brasil

...pero un 28% de los brasileños blancos llevan secuencias de mtDNA de origen africano!

Pero ahora podemos acceder al genoma entero (o casi)

Infinium HD BeadChips	Samples per BeadChip	Markers per Sample
HumanOmni1-Quad	4	> 1 million*
Human1M-Duo	2	> 1 million
HumanOmniExpress	12	> 700,000
Human660W-Quad	4	> 658,000
HumanCytoSNP-12	12	~ 300,000
Semi-Custom Human1M-Duo+, and HumanHap550-Quad+	2 / 4	standard content and up to 60,800 customize SNPs per sample

*Completely redesigned array with aggressively selected SNPs and probes from all three HapMap phases, the 1,000 Genome: Project, and published studies.

Genome-Wide Human SNP Array 6.0

Printer-Friendly Email

Overview | Literature | How to Order | Tools & Data

Pure Power and Performance

The new Affymetrix Genome-Wide Human SNP Array 6.0 features 1.8 million genetic markers, including more than 906,600 single nucleotide polymorphisms (SNPs) and more than 946,000 probes for the detection of copy number variation. The SNP Array 6.0 is the only platform with analysis tools to truly bridge copy number and association, including a new, high-resolution reference map and a copy number polymorphism (CNP) calling algorithm developed by the Broad Institute. The SNP Array 6.0 demonstrates industry-leading performance and represents more genetic variation on a single array than any other product, providing maximum panel power and the highest physical coverage of the genome.

The high price-performance value of the SNP Array 6.0 enables researchers to design association studies with larger sample sizes in the initial scan and replication phases, thereby significantly increasing the overall genetic power of their studies.

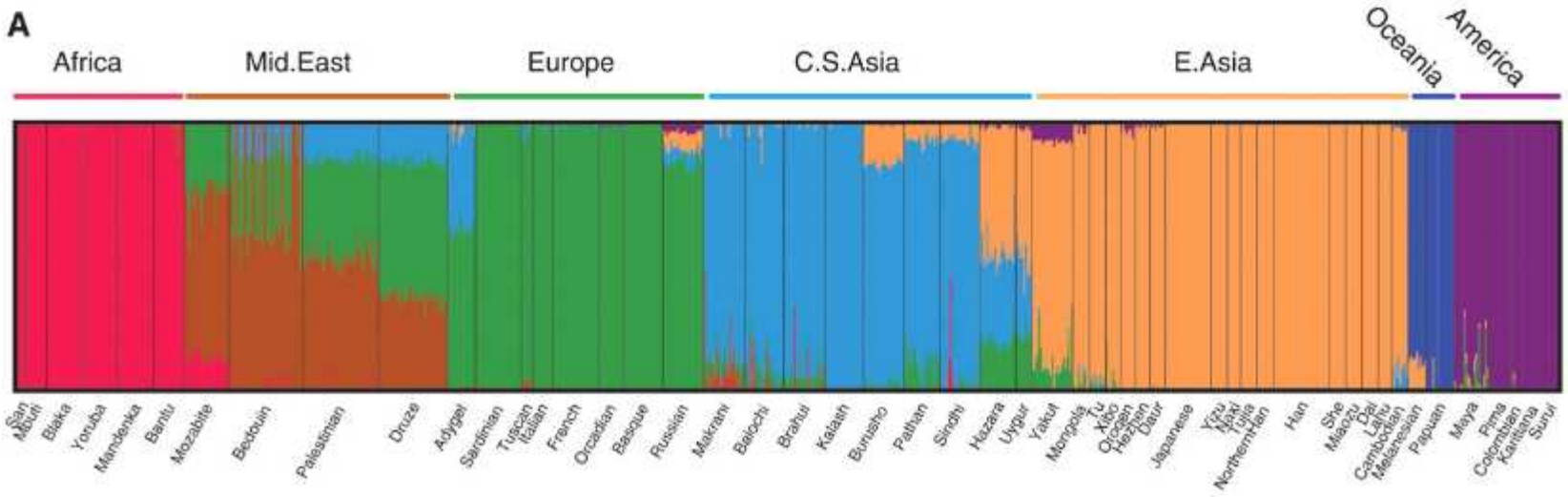
New! Affymetrix Cytogenetics Research Solution

Affymetrix Copy Number Solution on SNP 6.0

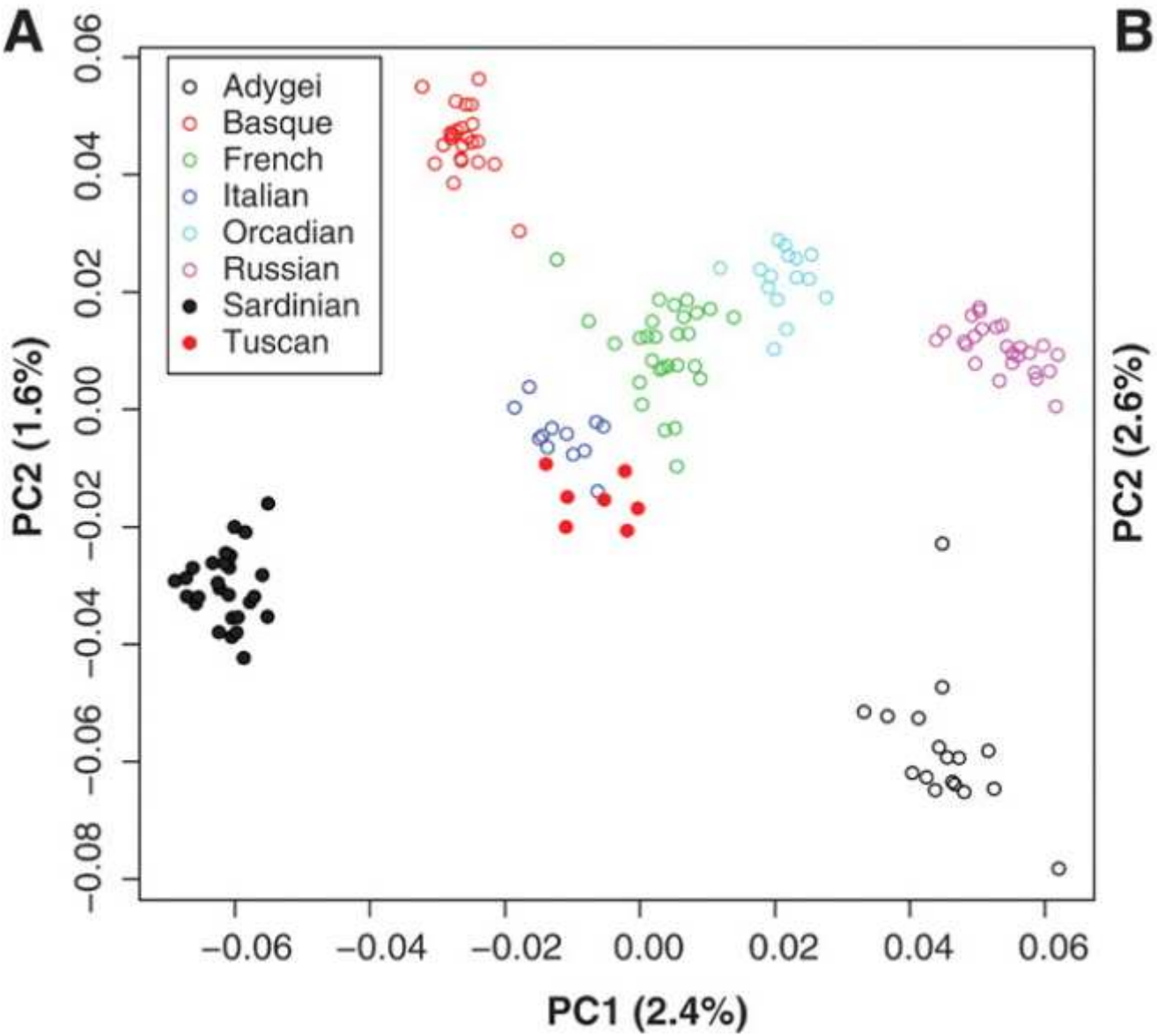
Copy Number and LOH analysis tools in Genotyping Console

Automation Solution

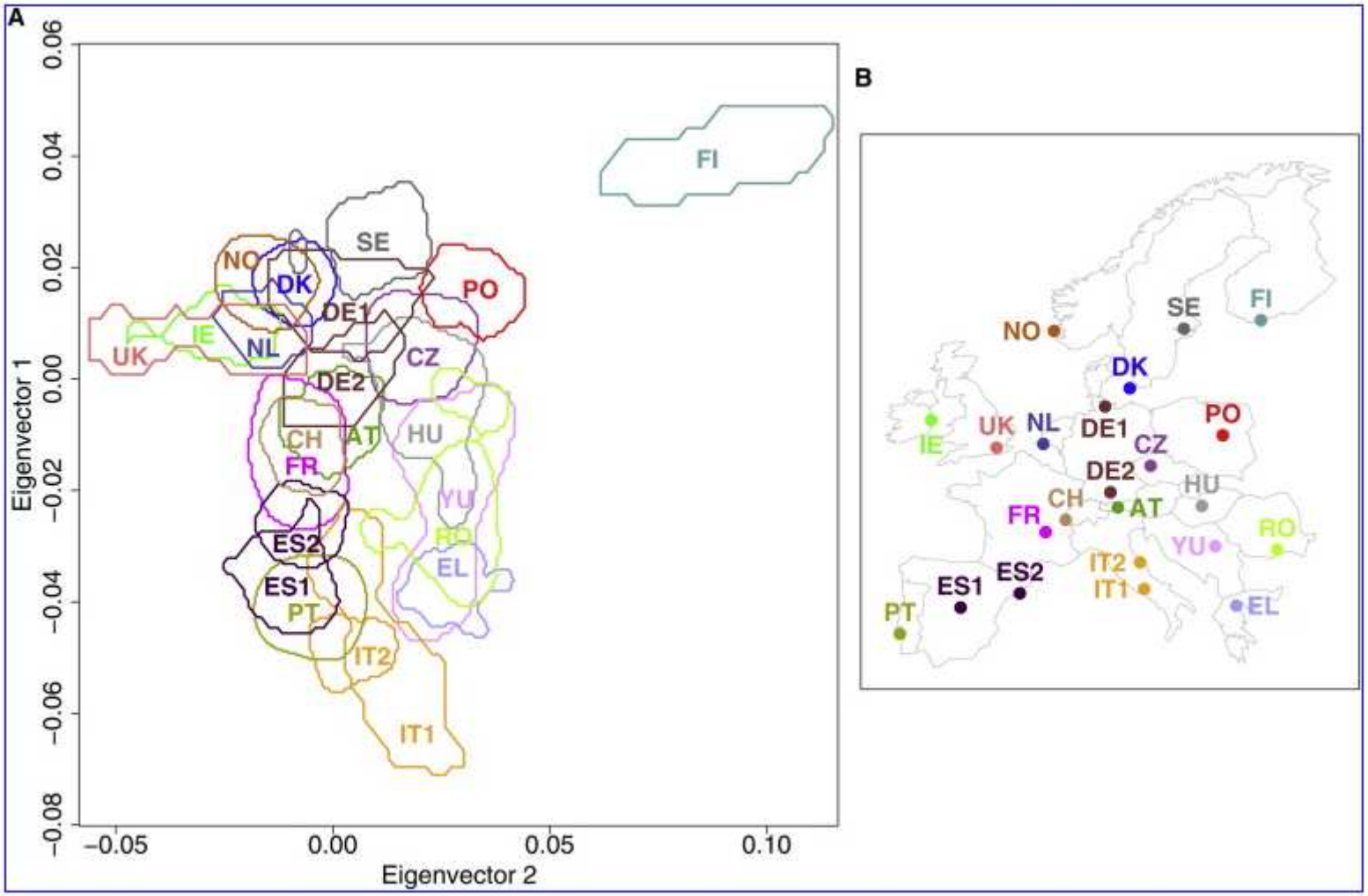
...lo que ha llevado el análisis al nivel individual



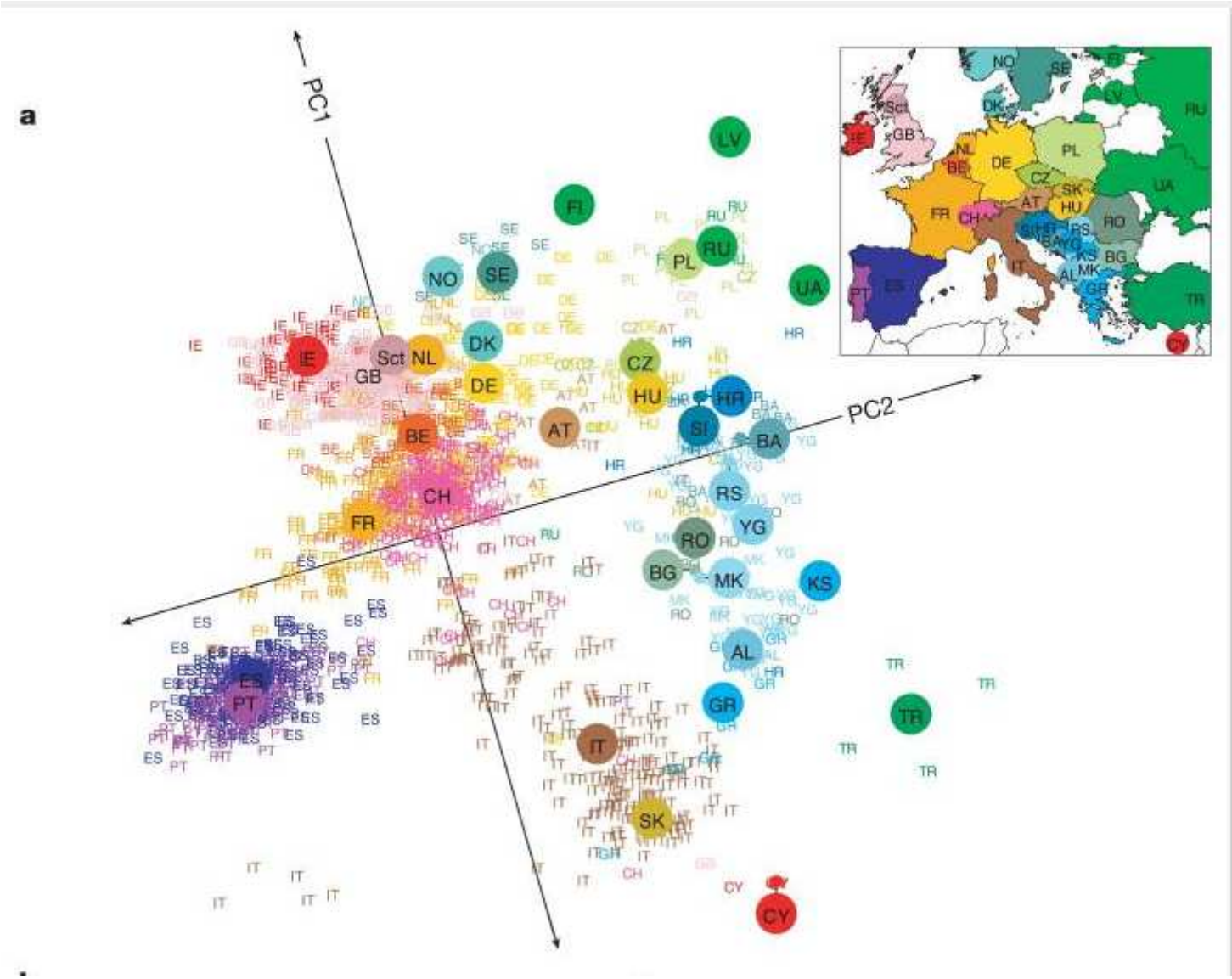
Li et al. Science. 2008 Feb 22;319(5866):1100-4.



Li et al. Science. 2008 Feb 22;319(5866):1100-4.



Lao et al Curr Biol. 2008 Aug 26;18(16):1241-8



Novembre et al. Nature. 2008 Nov 6;456(7218):98-101

Fenotipos dentro de la variación normal:

- Estatura
- Color de piel, ojos, cabello
- Pelo liso / rizado
- Sabor amargo PTC
- Reacción a la ingesta de alcohol
- Persistencia de la lactasa
- Tipo de cerumen
- Síndrome *ACHOO*
- Anosmia del espárrago

.
.
.

Aplicación clara en genética forense: ¿Podríamos reconstruir un retrato robot del sospechoso a partir de una muestra biológica dejada en el escenario del crimen?

Estatura: aunque se han enconrado algunos QTLs, la mejor manera de predecir la estatura de un individuo sigue siendo la media de la de sus padres.

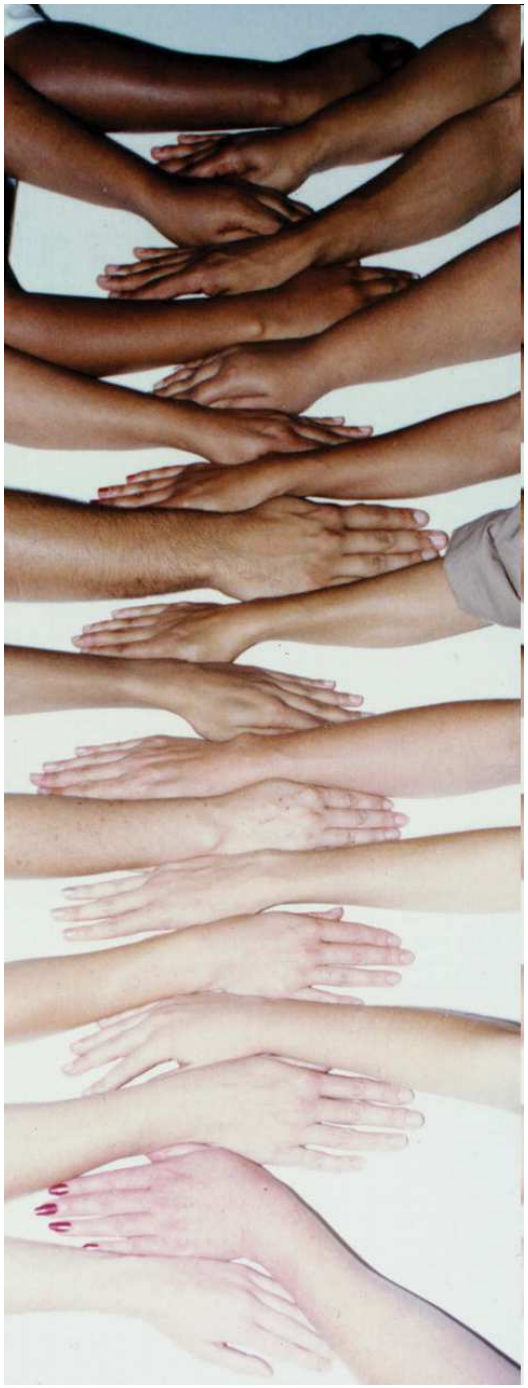
Table 1 SNPs showing strongest correlation with height in meta-analysis

			Effect						Effect		Effect		
SNP / Allele	Freq.	Region	ICE	DUT	USe	DGI	USa	Scan *P*	DAN	DAN *P*	Comb.	Comb. *P*	Neighboring genes
rs11205277 G	0.44	1q12	4.8	5.5	10.6	–	2.2	7.3×10^{-9}	5.4	0.0052	5.1 (0.8)	1.4×10^{-10}	Histone class 2A, *MTMR11, SV2A, SF3B4*
rs678962 G	0.22	1q24	4.4	11.3	8.0	6.5	0.8	3.2×10^{-8}	–	–	5.4 (1.0)	3.2×10^{-8}	*DNM3*
rs2274432 T	0.37	1q25	5.3	1.2	9.8	–	11.4	9.7×10^{-8}	4.5	0.027	5.3 (0.9)	7.8×10^{-9}	*C1orf19,GLT25D2*
rs3791679 T	0.81	2p16	5.9	4.5	6.6	2.9	18.8	1.1×10^{-7}	9.1	3.0×10^{-5}	5.8 (0.9)	5.9×10^{-11}	*EFEMP1, PNPT1*
rs6763931 A	0.45	3q23	7.4	9.2	6.5	6.4	3.7	2.6×10^{-18}	13.5	9.9×10^{-13}	7.4 (0.7)	1.4×10^{-27}	*ZBTB38*
rs6842303 T	0.26	4p15	3.8	5.0	15.7	–	14.8	5.7×10^{-6}	9.2	2.0×10^{-5}	4.8 (0.8)	4.2×10^{-9}	*LCORL, NCAPG*
rs6830062 T	0.89	4p15	4.8	11.9	21.6	3.6	8.8	7.6×10^{-8}	10.5	0.00018	6.3 (1.0)	1.3×10^{-10}	*LCORL, NCAPG*
rs1812175 C	0.86	4q31	7.9	8.3	7.8	11.0	10.4	2.8×10^{-10}	6.9	0.010	8.3 (1.2)	9.7×10^{-12}	*HHIP*
rs12198986 A	0.50	6p24	7.9	2.7	2.2	–	1.0	6.4×10^{-11}	3.3	0.080	6.8 (1.0)	2.4×10^{-11}	*BMP6*
rs10946808 A	0.70	6p22 B	5.6	6.1	7.7	–	–0.2	1.1×10^{-7}	7.4	0.00085	5.6 (0.9)	5.8×10^{-10}	Histone class 1, Butyrophilin genes
rs2844479 T	0.67	6p21 A	7.0	6.7	4.0	0.7	8.8	8.4×10^{-9}	–	–	6.3 (1.1)	8.4×10^{-9}	HLA class III
rs3130050 G	0.19	6p21 A	8.5	0.3	0.0	–	3.6	5.1×10^{-7}	6.3	0.021	7.4 (1.3)	3.2×10^{-8}	HLA class III
rs185819 T	0.52	6p21 B	5.2	5.2	5.9	–	3.6	1.9×10^{-6}	6.4	0.0048	5.2 (0.9)	3.2×10^{-8}	HLA class III
rs1776897 C	0.07	6p21 D	9.8	2.3	7.3	–	8.4	1.1×10^{-7}	6.9	0.043	8.8 (1.6)	1.4×10^{-8}	*HMGA1, LBH*
rs4713858 G	0.86	6p21 E	6.4	11.2	1.1	8.6	2.5	4.9×10^{-8}	3.9	0.22	6.8 (1.2)	3.5×10^{-8}	*ANKS1, TCP11, ZNF76, DEF6, SCUBE3*
rs3748069 A	0.74	6q24	5.7	9.7	11.4	8.0	3.2	4.9×10^{-11}	8.3	0.00015	6.5 (0.9)	4.5×10^{-14}	*GPR126*
rs798544 G	0.72	7p22	5.0	9.8	13.1	4.8	10.5	4.5×10^{-12}	7.7	0.00029	5.9 (0.8)	6.5×10^{-15}	*GNA12*
rs2282978 C	0.29	7q21	5.5	10.4	–0.5	7.5	1.4	9.8×10^{-9}	–	–	5.8 (1.0)	9.8×10^{-9}	*CDK6, PEX1, GATAD1, ERVWE1*
rs11765954 C	0.24	7q21	6.0	8.3	–0.2	8.9	–0.1	1.2×10^{-7}	2.9	0.17	6.1 (1.1)	6.9×10^{-8}	*CDK6, PEX1, GATAD1, ERVWE1*
rs10958476 C	0.23	8q12	5.5	3.0	9.9	–	4.6	1.4×10^{-6}	5.7	0.015	5.4 (1.0)	6.6×10^{-8}	*PLAG1, MOS, CHCHD7, RDHE2, RPS20, LYN, TGS1, PENK*
rs7846385 C	0.27	8q21	5.1	9.9	2.6	2.7	–4.5	5.3×10^{-7}	4.6	0.030	5.0 (0.9)	4.7×10^{-8}	*PXMP3, ZFHX4*
rs4743034 A	0.23	9q31	4.9	5.2	10.9	7.9	0.0	5.2×10^{-7}	5.8	0.012	5.3 (0.9)	2.1×10^{-8}	*ZNF462*
rs8756 C	0.52	12q14	6.0	5.0	8.8	10.7	9.4	2.2×10^{-13}	7.2	0.00016	6.6 (0.8)	1.8×10^{-16}	*HMGA2*
rs7153027 A	0.52	14q32	4.6	10.1	4.3	6.9	18.1	1.1×10^{-10}	–	–	5.7 (0.9)	1.1×10^{-10}	*TRIP11, FBLN5, ATXN3, CPSF2*
rs4533267 A	0.28	15q26	6.2	7.1	–0.4	2.6	1.5	3.8×10^{-7}	4.8	0.030	5.6 (1.0)	3.3×10^{-8}	*ADAMTS17*
rs3760318 C	0.63	17q11	5.9	6.8	4.1	8.8	–0.8	1.8×10^{-9}	–	–	6.0 (1.0)	1.8×10^{-9}	*CRLF3, ATAD5, CENTA2, RNF135*
rs4794665 A	0.48	17q23 A	3.4	4.3	3.8	3.3	7.8	9.2×10^{-6}	6.2	0.0016	3.6 (0.7)	9.9×10^{-8}	*NOG, DGKE, TRIM25, COIL, RISK*
rs757608 T	0.35	17q23 B	4.7	3.4	10.8	3.4	–7.4	3.0×10^{-7}	3.7	0.073	4.4 (0.8)	6.3×10^{-8}	*BCAS3, NACA2, TBX2, TBX4*
rs4800148 A	0.79	18q11	5.2	10.5	16.3	–	8.7	9.4×10^{-8}	5.9	0.012	6.4 (1.1)	3.7×10^{-9}	*CABLES1, RBBP8*, C18orf45
rs967417 C	0.53	20p12	4.3	4.3	5.5	4.6	1.3	4.6×10^{-6}	7.1	0.00025	4.3 (0.8)	1.5×10^{-8}	*BMP2*

© 2008 Nature Publishing Group http://www.nature.com/naturegenetics

"The estimated effects per allele of these variants ranged between 0.3 and 0.6 cm and, taken together, they explain around 3.7% of the population variation in height. "

Gudbjartsson et al. (2008) Nat Genet 40:609-615



Pigmentación: piel

Genes y SNPs de pigmentación en humanos

Table 1. Human pigmentation gene polymorphisms and population allele frequencies

Mouse coat colour/KO	Human	Protein	Null	Variation	SNP	Amino acid	Population[a] SNP frequencies (allele 1/allele 2)[b] European	Chinese	Japanese	African
Albino	*TYR*	Tyrosinase 529 aa	OCA1	Skin colour	rs1042602 C>A	Ser192Tyr	**0.583/0.417**	1.0/0	1.0/0	1.0/0
				Skin colour	rs1800422 G>A	Arg402Gln	**0.604/0.396**[c]	1.0/0[c]	–	0.935/0.065[c]
					rs1126809 G>A	Arg402Gln	**0.783/0.217**	1.0/0	1.0/0	1.0/0
Brown	*TYRP1*	TYRP1 537 aa	OCA3	Eye colour, skin colour	rs1408799 C>T	–	**0.30/0.70**	0.989/0.011	0.978/0.022	0.775/0.225
					rs2733832 C>T		**0.367/0.633**	0.989/0.011	0.977/0.023	0.933/0.067
Pink-eyed dilution	*OCA2*	P-protein 838 aa	OCA2	Eye colour, skin colour	rs1800401 C>T	Arg305Trp	0.935/0.065[d]	–	–	0.979/0.021[d]
					rs1800407 G>A	Arg419Gln	**0.933/0.067**	1.0/0	1.0/0	1.0/0
					rs1800414 A>G	His615Arg	1.0/0	**0.367/0.633**	**0.477/0.523**	1.0/0
				Brown/blue eye colour	rs12913832 T>C	–	**0.208/0.792**	1.0/0	1.0/0	1.0/0
Underwhite	*SLC45A2*	MATP 530 aa	OCA4	Skin colour	rs26722 G>A	Glu272Lys	1.0/0	**0.611/0.389**	**0.591/0.409**	0.95/0.05
					rs16891982 G>C	Leu374Phe	**0.017/0.983**	0.989/0.011	1.0/0	1.0/0
Ocular albinism	*SLC24A5*	NCKX5 500 aa	–	Skin colour	rs1426654 G>A	Ala111Thr	**0/1.0**	0.989/0.011	0.989/0.011	0.975/0.025
Extension/recessive yellow	*MC1R*	MC1R 317 aa	Red hair	r	rs1805005 G>T	Val60Leu	**0.122**[e]			
				R	rs1805006 C>A	Asp84Glu	**0.012**[e]			
				r	rs2228479 G>A	Val92Met	**0.097**[e]			
				R	rs11547464 G>A	Arg142His	**0.004**[e]			
				R	rs1805007 C>T	Arg151Cys	**0.11**[e]			
				R	rs1110400 T>C	Ile155Thr	**0.009**[e]			
				R	rs1805008 C>T	Arg160Trp	**0.07**[e]			
				R	rs885479 G>A	Arg163Gln	**0.047**[e]			
				R	rs1805009 G>C	Asp294His	**0.027**[e]			
Agouti	*ASIP*	ASIP 132 aa	–	Eye colour	rs4911442 A>G		**0.93/0.07**	1.0/0	1.0/0	1.0/0
					rs6058017 (8818 A>G)		**0.758/0.242**[d]			0.396/0.604[d]
					rs1015362 A>G		0.233/0.767	0.122/0.878	0.261/0.739	**0.833/0.167**
					rs4911414 G>T		0.725/0.275	0.878/0.122	0.739/0.261	0.867/0.133
Steel	*KITLG*	KITLG 273 aa	–	Skin colour	rs642742 A>G		0.136/0.864	0.267/0.733	0.114/0.886	**0.922/0.0778**
				Hair colour	rs12821256 T>C		**0.858/0.142**	1.0/0	1.0/0	1.0/0
	IRF4		–	Hair colour	rs12203592 C>T		**0.167/0.833**	1.0/0	1.0/0	1.0/0
				Hair colour	rs1540771 G>A		**0.424/0.576**	0.756/0.244	0.644/0.356	0.942/0.058
	SLC24A4	SLC24A4 605 aa	–	Hair colour, eye colour	rs12896399 G>T		**0.4/0.6**	0.789/0.211	**0.42/0.58**	0.992/0.008
	TPCN2	TPC2 752 aa	–	Hair colour	rs35264875 A>T	Met484Leu	**0.825/0.175**	1.0/0	1.0/0	1.0/0
				Hair colour	rs3829241 G>A	Gly734Gln	**0.55/0.45**	0.822/0.178	0.727/0.273	0.975/0.025

[a]Hapmap population descriptors: CEU: CEPH (Utah residents with ancestry from northern and western Europe); JPT: Japanese in Tokyo, Japan; CHB: Han Chinese in Beijing, China; YRI: Yoruba in Ibadan, Nigeria; http://www.hapmap.org/.
[b]Population allele frequencies that are most divergent are indicated in bold type.
[c]Population data recorded in Perlegen database http://genome.perlegen.com/.
[d]Population data recorded in SNP500Cancer panel NCI.
[e]Allele frequency from (56).

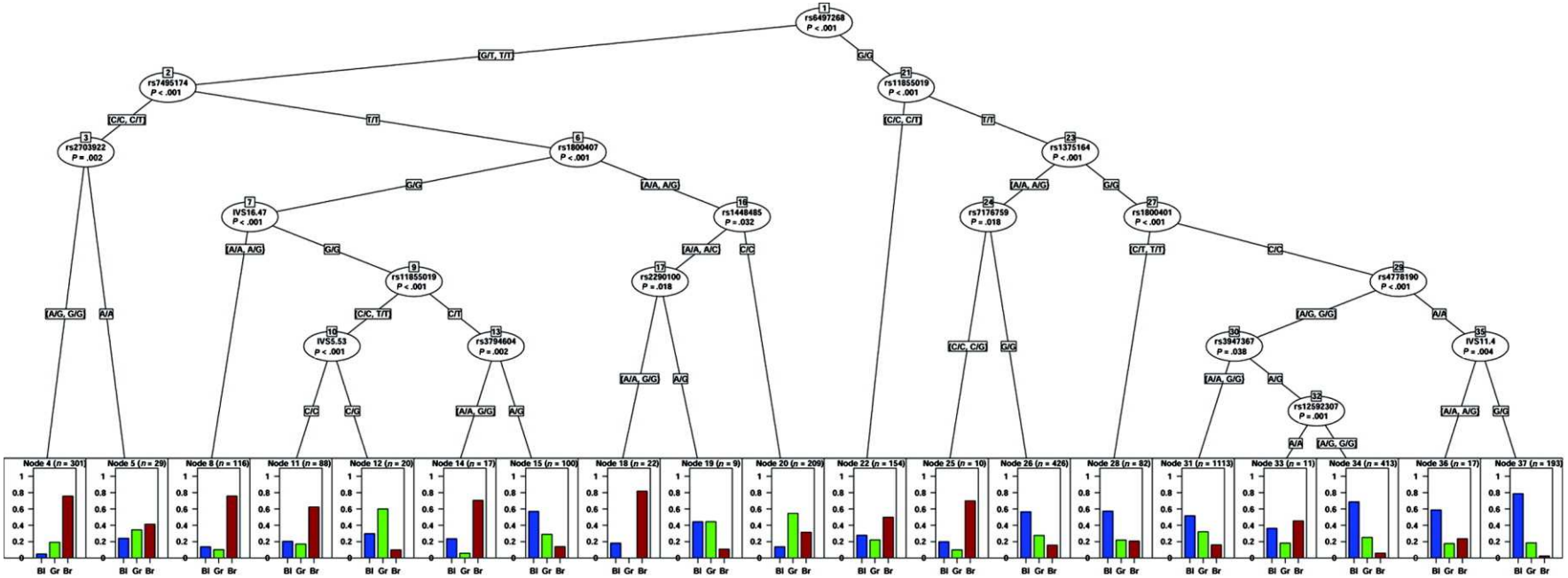
Sturm et al. Human Molecular Genetics, 2009, Vol. 18, Review Issue 1 R9–R17

Color del cabello: sólo se conocen con certeza las bases genéticas del pelo rojo

MC1R↓ ⇒ feomelanina (no eumelanina)
pelirrojos o rubios rojizos
piel clara y pecosa
riesgo cáncer de piel↑
sensibilidad al dolor↓

Color de ojos:

básicamente determinado, aunque de manera muy compleja, por la variación en el gen OCA1.



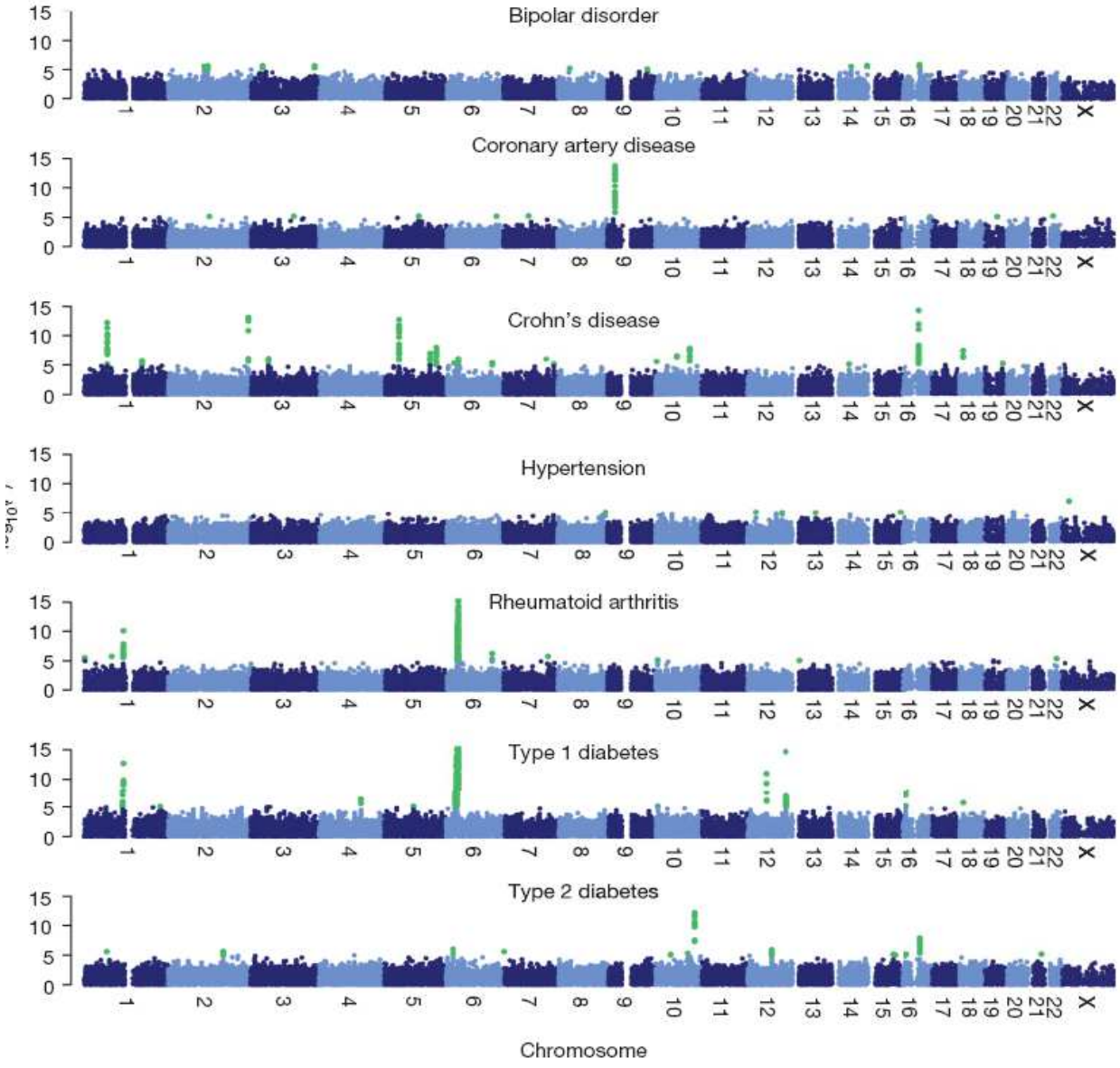
Duffy et al., *Am. J. Hum. Genet.*, 80:241-252, 2007

Enfermedades:

- mendelianas: raras, pero altamente predictibles. Ligamiento

- complejas: muy comunes, pero poco predictibles. Asociación

GWAS con 3000 casos y 3000 controles



Los resultados de los GWASs explican sólo una parte relativamente pequeña de la heredabilidad de las enfermedades complejas

Table 1 | Estimates of heritability and number of loci for several complex traits

Disease	Number of loci	Proportion of heritability explained
Age-related macular degeneration[72]	5	50%
Crohn's disease[21]	32	20%
Systemic lupus erythematosus[73]	6	15%
Type 2 diabetes[74]	18	6%
HDL cholesterol[75]	7	5.2%
Height[15]	40	5%
Early onset myocardial infarction[76]	9	2.8%
Fasting glucose[77]	4	1.5%

* Residual is after adjustment for age, gender, diabetes.

Manolio et al. Nature. 2009 Oct 8;461(7265):747-53

Genómica individualizada y empresa: el modelo "directo al consumidor"

Al menos existen 32 empresas que ofrecen estimación de ancestría con mtDNA y/o cromosoma Y y/o algunos marcadores autosómicos

Genetic Genealogy Testing Firms

African Ancestry
AncestryByDNA
Argus Biosciences
Cambridge DNA
Chromosomal Laboratories, Inc.
The DNA Ancestry Project
DNA Bioscience
DNA Canada
DNA Consultants
DNA Diagnostics Center
DNA Fingerprint
DNA Heritage
DNA Roots
DNA Tribes
DNA Worldwide
Earth Ancestry
EthnoAncestry
Family Genetics
FamilyTreeDNA
Gen by Gen
Genelex
GeneTree
The Genetic Testing Laboratories, Inc.
The Genographic Project
GeoGene
My Genetic Heritage
Oxford Ancestors
Relative Genetics
roots for real
SMGF
Trace Genetics

Y tres empresas ofrecen ancestría y salud con análisis de >1.000.000 SNPs

23andMe deCODEme

Navigenics®