



Genetic Mapping of Complex Phenotypes in the Brazilian Population

Alexandre Pereira
alexandre.pereira@incor.usp.br

✓ Brazil's population is very diverse, comprising many races and ethnic groups. In general, Brazilians trace their origins from four sources of migration: Amerindians, Europeans, Africans and Asians.

✓ Historically Brazil's population always experienced large degrees of ethno-religious intermarriage, mutual assimilation of cultures, syncretism and, above all, racial admixture.

Immigration to Brazil, by Ethnic groups, periods from 1500 to 1933

Source: Brazilian Institute for Geography and Statistics (IBGE)

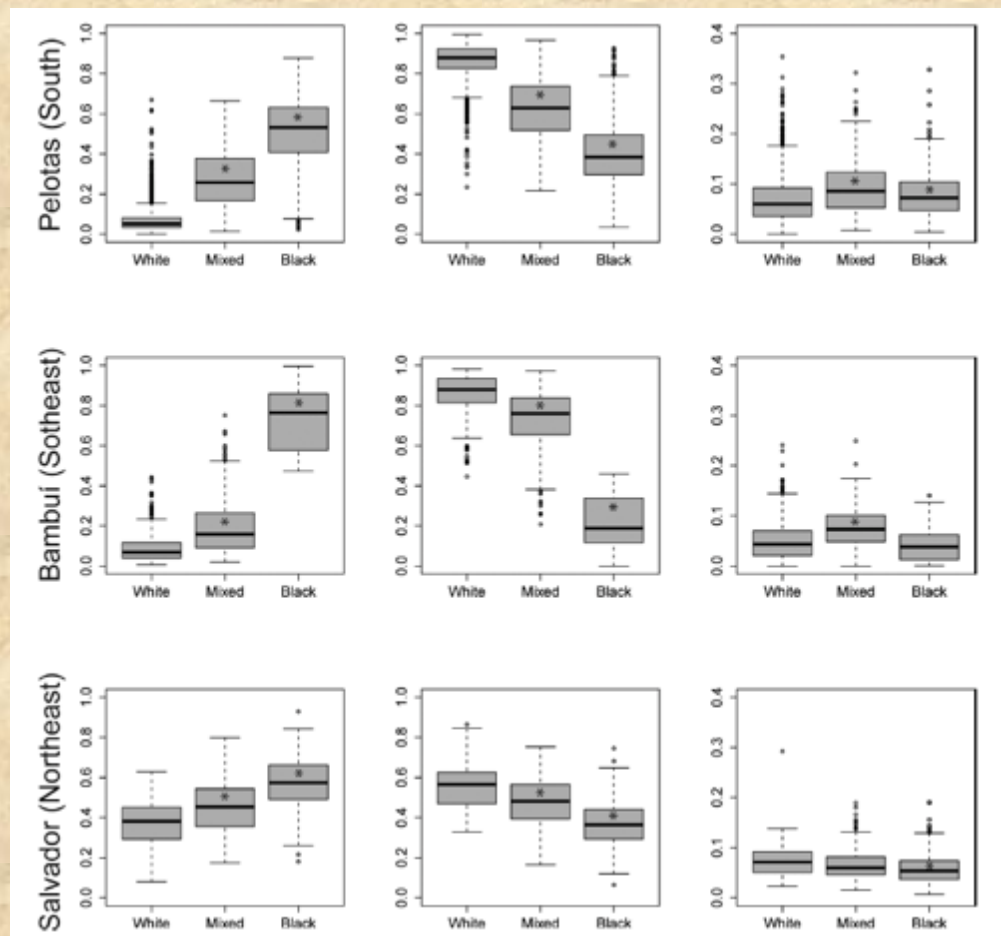
Ethnic group	Period									
	1500– 1700	1701– 1760	1761– 1829	1830– 1855	1856– 1883	1884– 1893	1894– 1903	1904– 1913	1914– 1923	1924– 1933
Africans	510,000	958,000	1,720,000	618,000	—	—	—	—	—	—
Portuguese	100,000	600,000	26,000	16,737	116,000	170,621	155,542	384,672	201,252	233,650
Italians	—	—	—	—	100,000	510,533	537,784	196,521	86,320	70,177
Spaniards	—	—	—	—	—	113,116	102,142	224,672	94,779	52,405
Germans	—	—	5,003	2,008	30,000	22,778	6,698	33,859	29,339	61,723
Japanese	—	—	—	—	—	—	—	11,868	20,398	110,191
Syrians and Lebanese	—	—	—	—	—	96	7,124	45,803	20,400	20,400
Others	—	—	—	—	—	66,524	42,820	109,222	51,493	164,586



IBGE (2010) - cor da pele

- **Brancos – 47,75%**
- **Pardos – 43,05%**
- **Negros – 7,63%**
- **Amarelos – 1,11%**
- **Indígenas – 0,45%**

Genomic ancestry and ethnoracial self-classification based on 5,871 community-dwelling Brazilians (The Epigen Initiative)

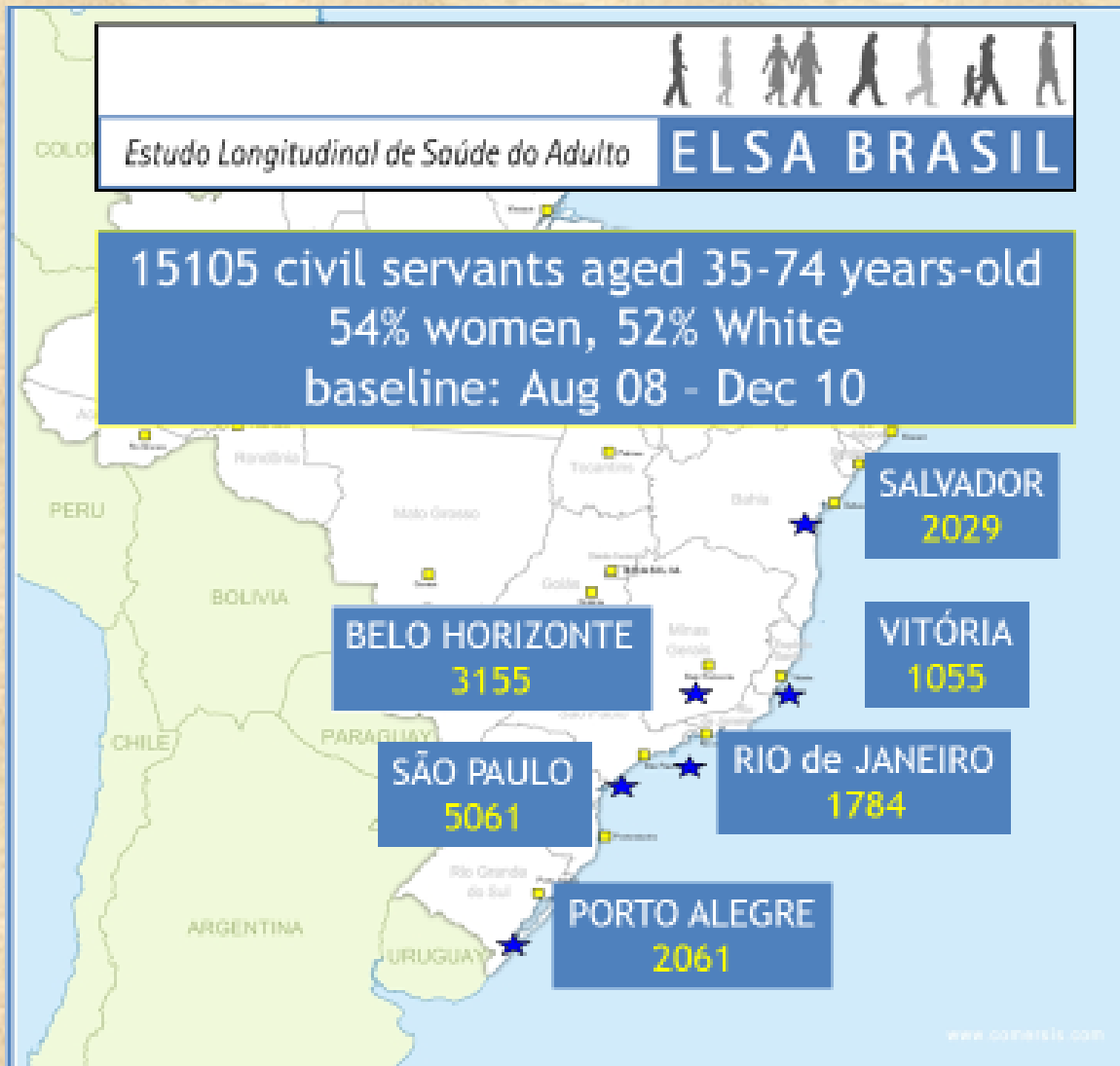




Estudo Longitudinal de Saúde do Adulto

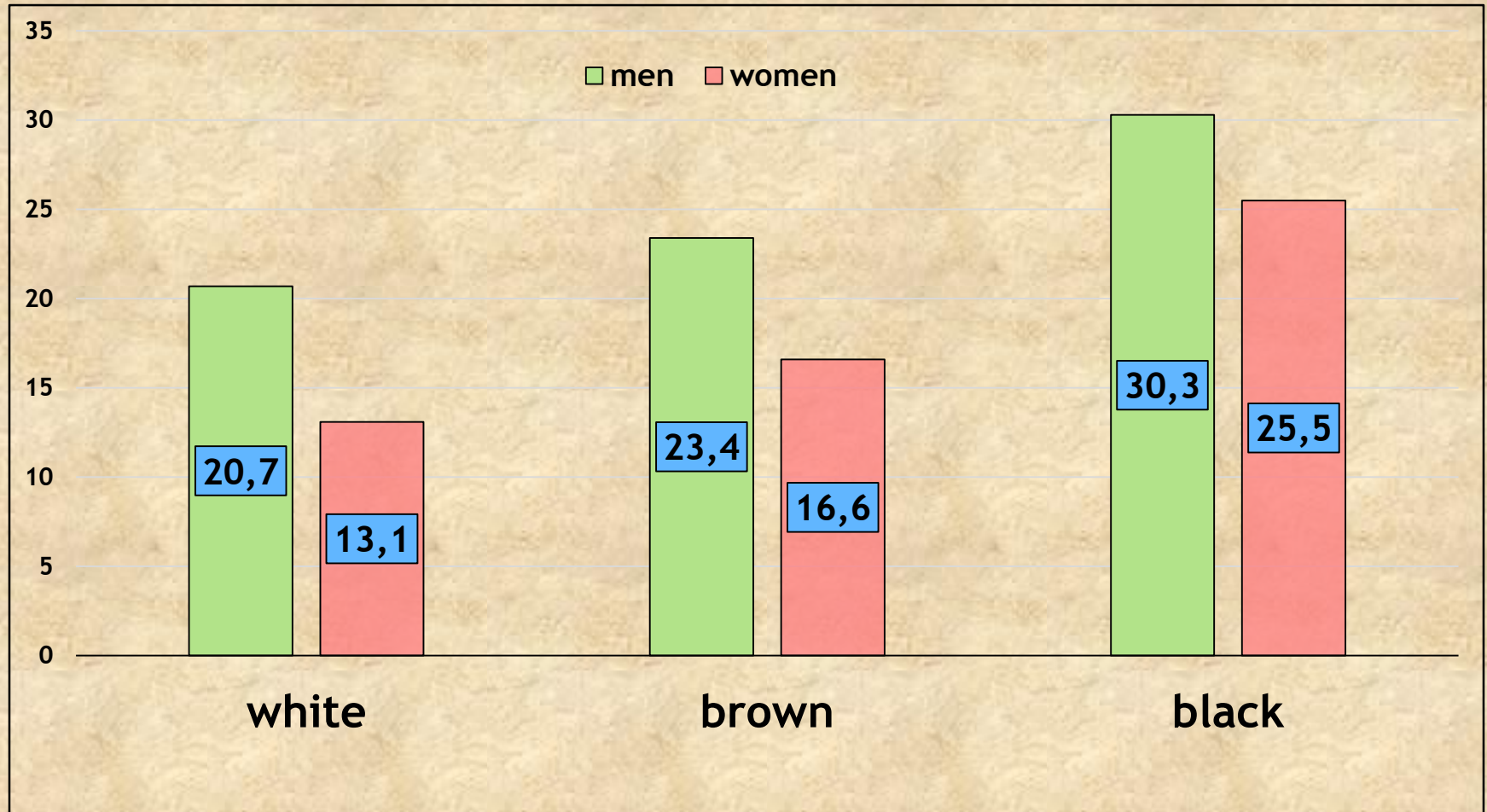
ELSA BRASIL

15105 civil servants aged 35-74 years-old
54% women, 52% White
baseline: Aug 08 - Dec 10

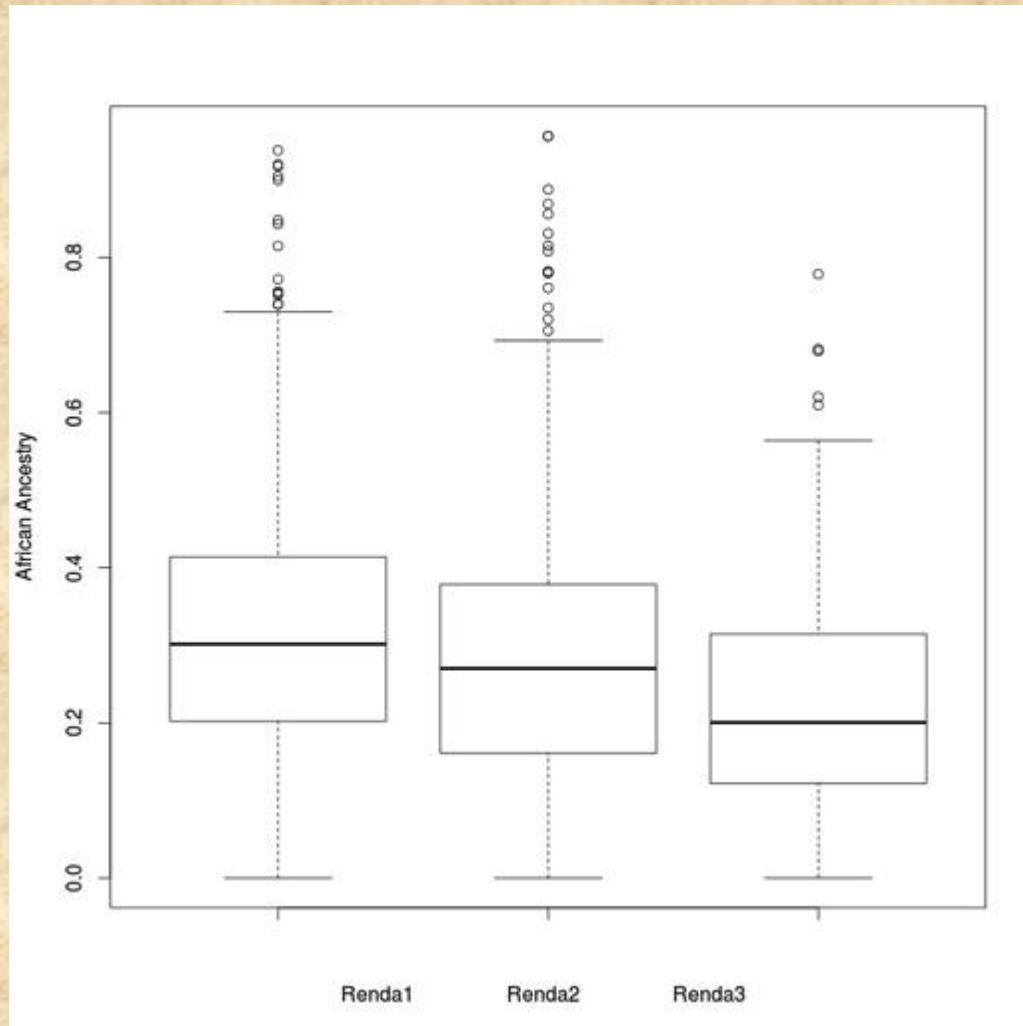




diabetes and race

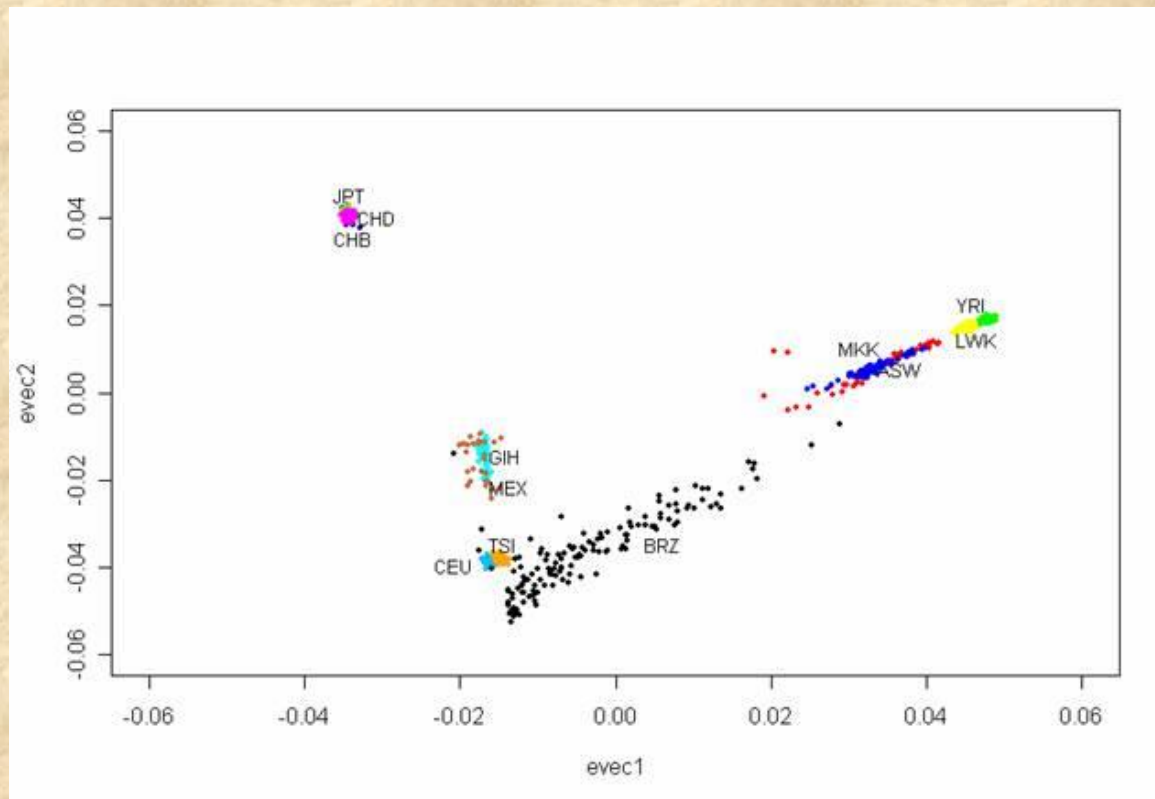


Horizontal Stratification, Income, Race/Color and Genetic Ancestry



ARTICLE

Brazilian urban population genetic structure reveals a high degree of admixture



The Project

EPIGEN-Brasil

Genomic Epidemiology of Complex Diseases in Brazilian Cohorts of Populational Basis

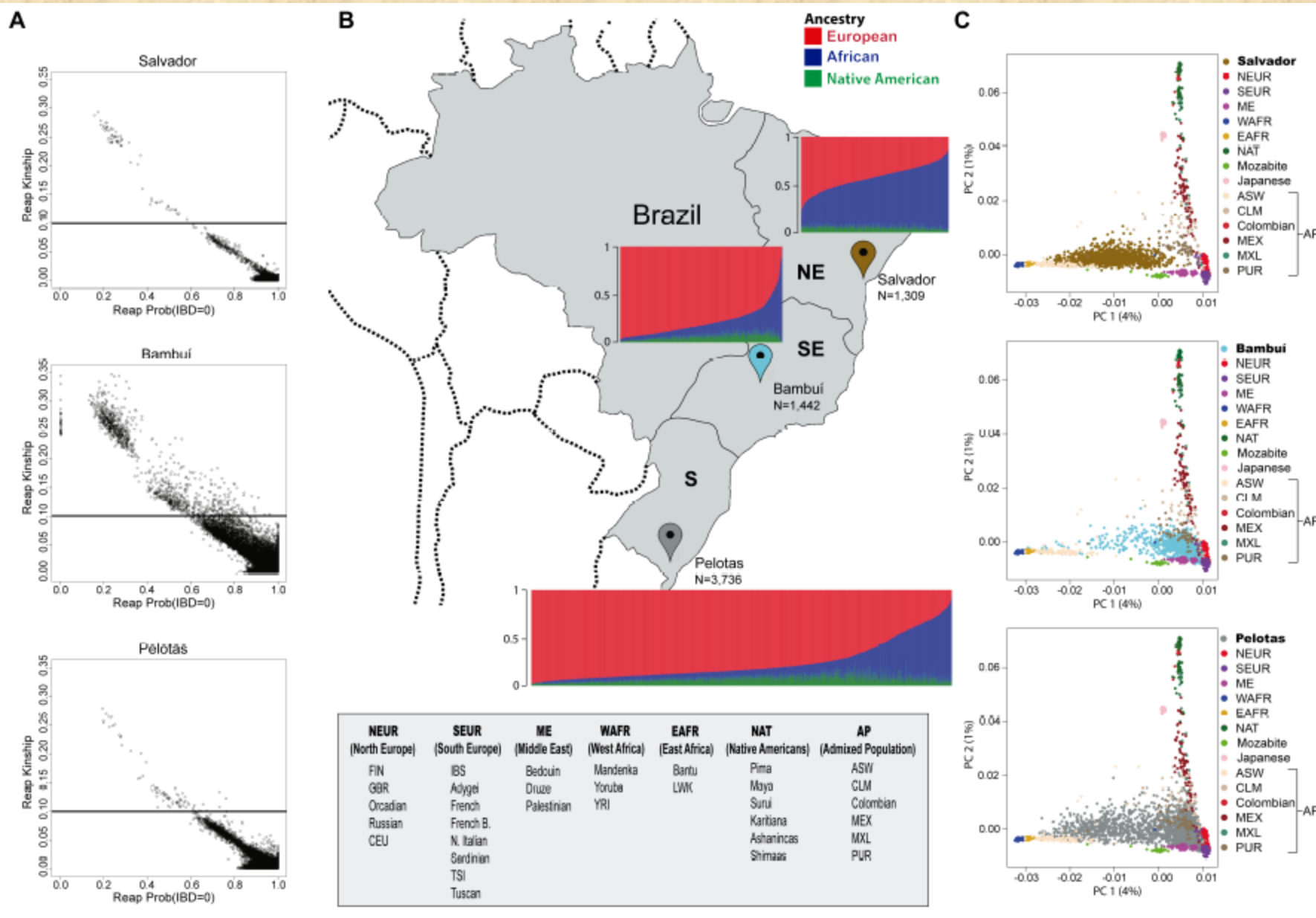


EpiGen-Brasil

ORIGIN AND DYNAMICS OF ADMIXTURE IN BRAZILIANS: A POPULATION BASED FINE-SCALE APPROACH

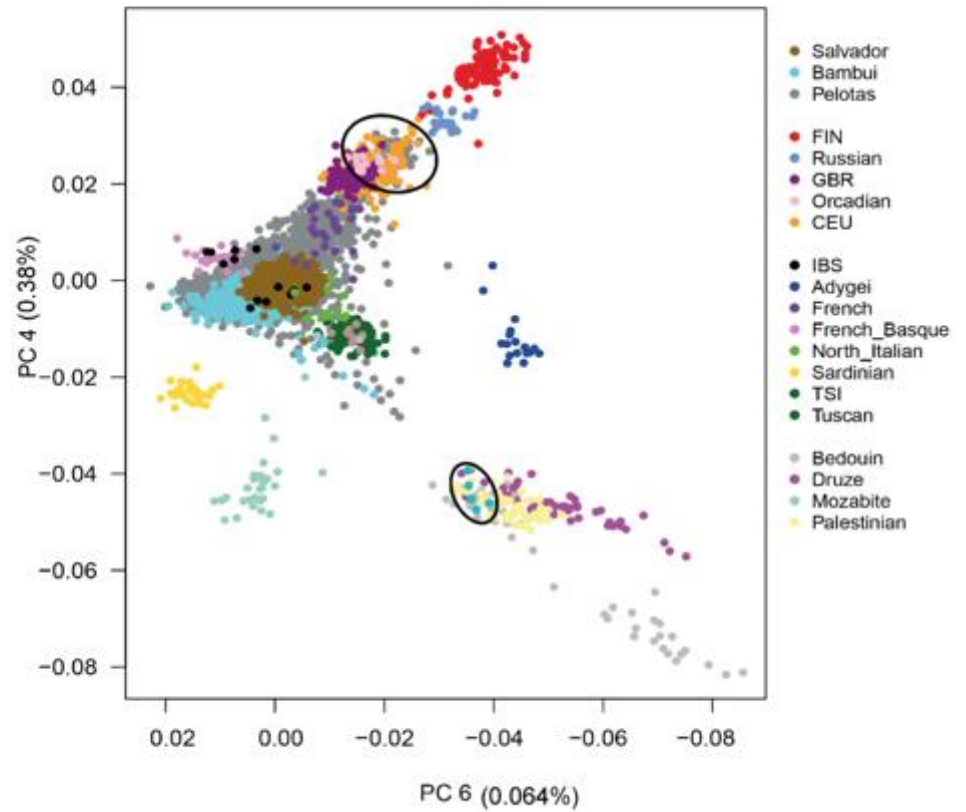
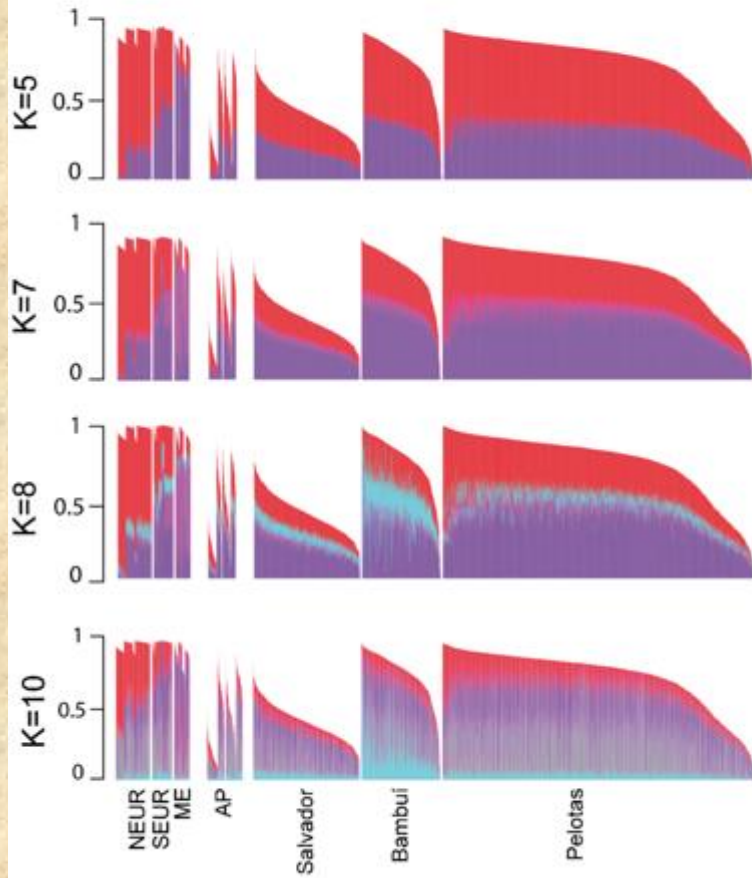


<http://epigen.grude.ufmg.br/>

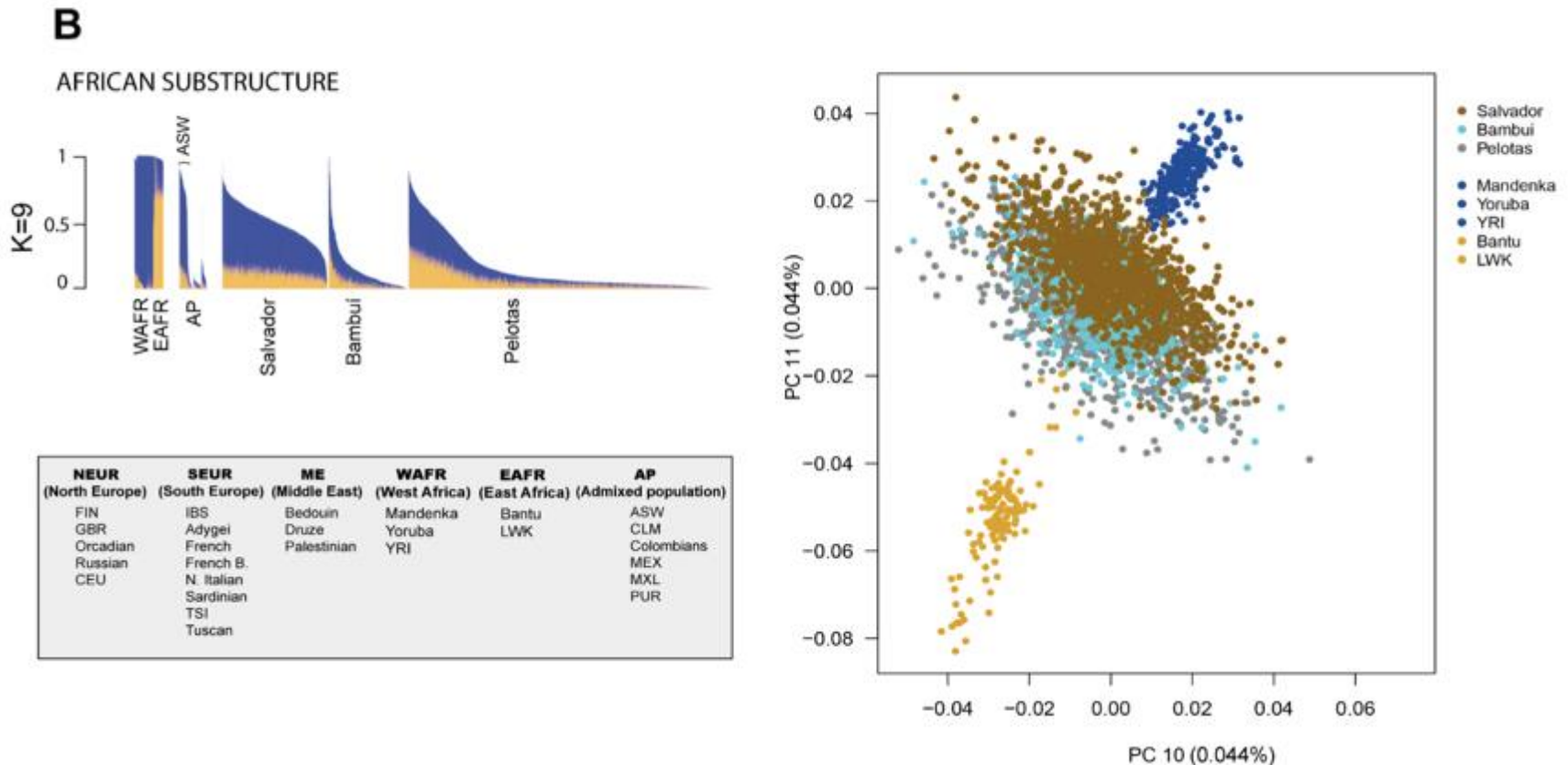


A

EUROPEAN SUBSTRUCTURE



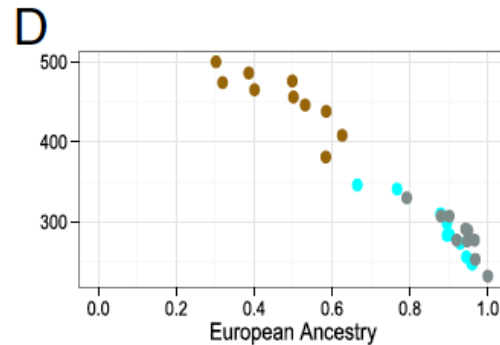
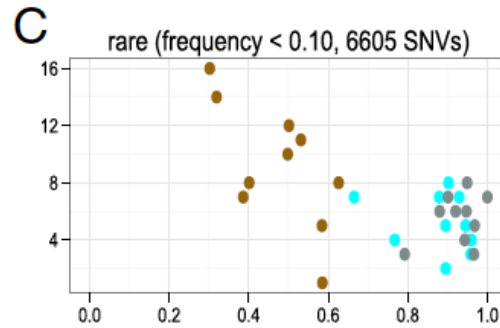
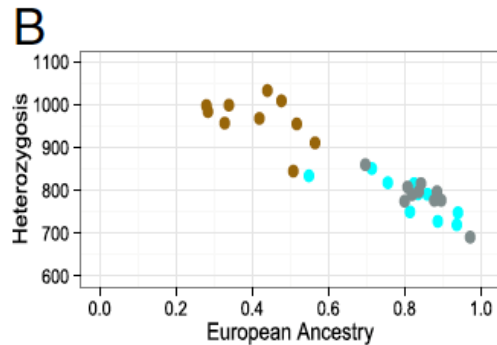
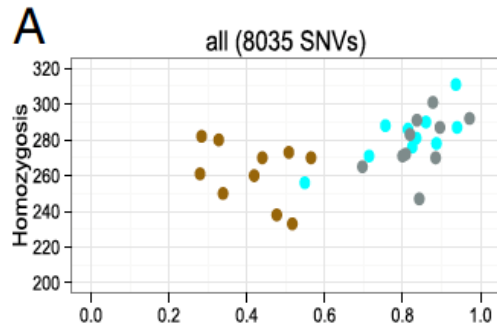
Slave trade also from Mozambique (Eastern Africa)



Western/non-Bantu associated cluster



Eastern/Bantu associated cluster



Disease / Trait	N.SNPs	SNP list (38)	Fst
Cognitive performance	3	rs2807580 rs2229741 rs4751674	0.0941 0.0707 0.0703
Crohn's disease	3	rs7702331 rs7517847 rs6556412	0.0750 0.0603 0.0599
Inflammatory bowel disease	3	rs477515 rs2382817 rs7517847	0.1261 0.0683 0.0603
Multiple sclerosis	2	rs12466022 rs533259	0.0688 0.0688
Obesity related	2	rs7964120 rs7784447	0.1322 0.0957
Emphysema-related traits	1	rs641525	0.1469
Epstein-Barr virus immune response	1	rs477515	0.1261
Liver enzyme levels	1	rs4547811	0.1108
<u>Schizophrenia</u>	1	rs1635	0.0862
Myopia (pathological)	1	rs4142248	0.0825
Alzheimer's disease	1	rs610932	0.0822
F-cell distribution	1	rs7565301	0.0738
Amyotrophic lateral sclerosis	1	rs2819332	0.0726
Menopause	1	rs11889862	0.0725
Eosinophil counts	1	rs4143832	0.0719
Obsessive-compulsive disorder	1	rs9652236	0.0717
HIV related	1	rs1020064	0.0716
Sphingolipid levels	1	rs1000778	0.0689
IgE levels in asthmatics	1	rs10404342	0.0673
Economic and political preferences	1	rs210648	0.0667
Bladder cancer	1	rs2294008	0.0660
Duodenal ulcer	1	rs2294008	0.0660
Nasopharyngeal carcinoma	1	rs6774494	0.0660
Non-alcoholic fatty liver disease histology	1	rs887304	0.0658
Prostate cancer	1	rs4242382	0.0652
Resp.to irinotecan/platinum-based chemo. lung cancer	1	rs344924	0.0647
Sudden cardiac arrest	1	rs5762311	0.0637
Type 1 diabetes	1	rs1004446	0.0634
Bipolar disorder	1	rs7250872	0.0626
Response to gemcitabine in pancreatic cancer	1	rs1901440	0.0625
Mean platelet volume	1	rs12526480	0.0625
Pancreatic cancer	1	rs10088262	0.0620
Breast size	1	rs7104745	0.0612

* Bold indicates unique entries and underline indicate co-occurrence in OMIM disease results.



Corações de Baependi





- The Baependi Heart Study is a family-based cohort;

Study of cardiovascular diseases and their genetic associations

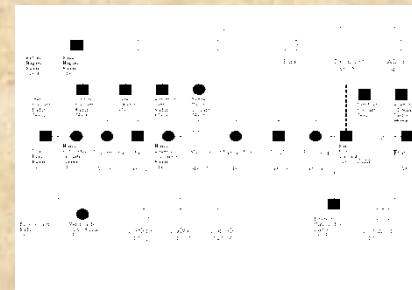
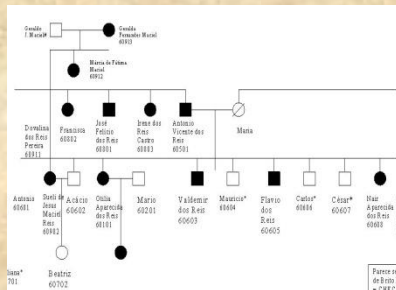
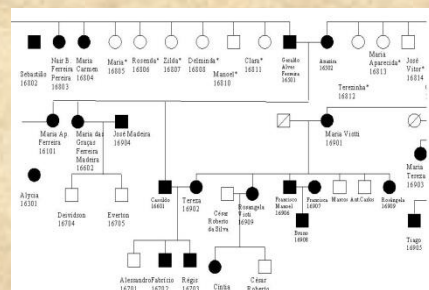
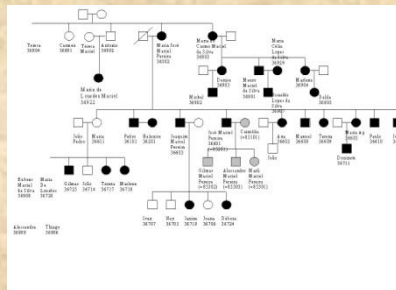
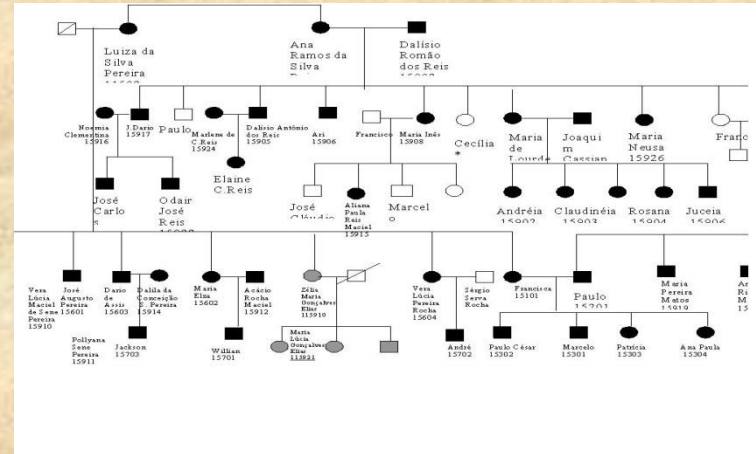
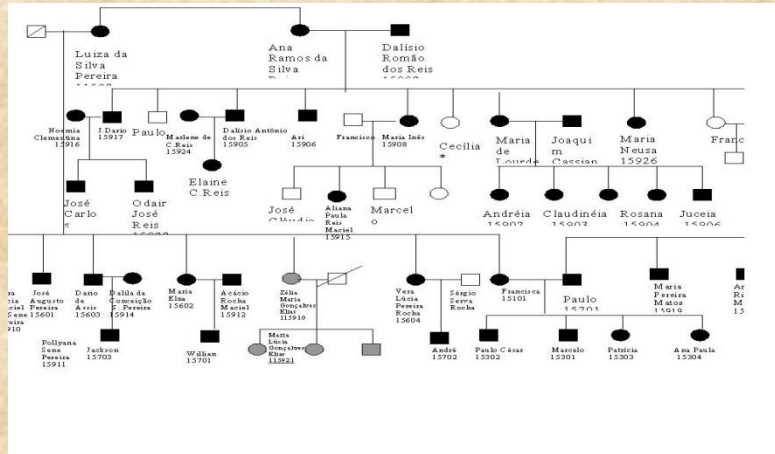
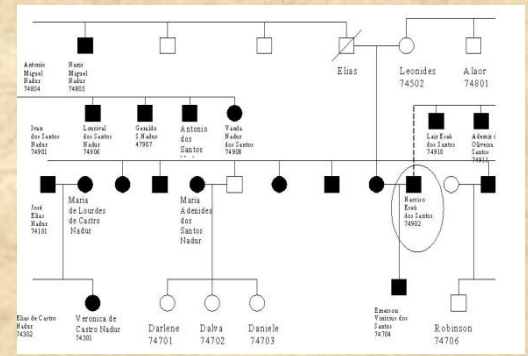
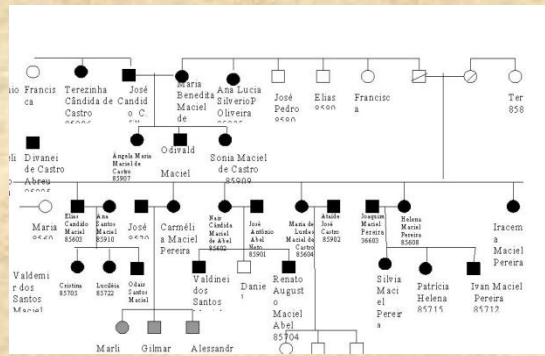
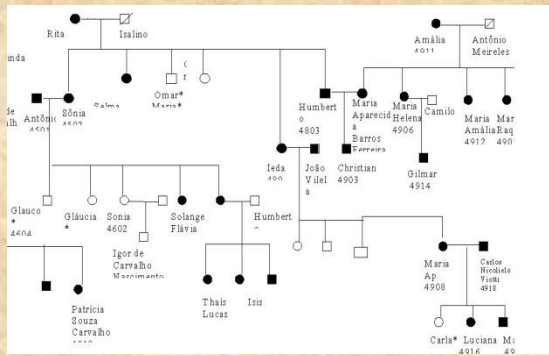
- The project began in December 2005;
- Follow-up every 5 years.



Baependi, Minas Gerais State, BR.



Baependi



Baependi - MG



✓ **Area:** 752 Km²

✓ **Population:** 18.273 inhabitants

}	68,4% Urban Area
	31,6% Rural Area



RELEASE 1

(December 2005)

1,695 individuals



81 families

RELEASE 2

(December 2010)

2,279 individuals



125 families

Blood collection

- Lipid profile;
- Creatinine;
- Uric acid;
- Glucose;
- HbA1c hemoglobin;
- Urea.

Interview

- Questionnaire;
 - Family relationships ;
 - Medical history;
 - Environmental risk factors

Physical exam

- Weight;
- Height;
- Blood Pressure, and others.

Several cardiovascular phenotypes

Genetic Mapping

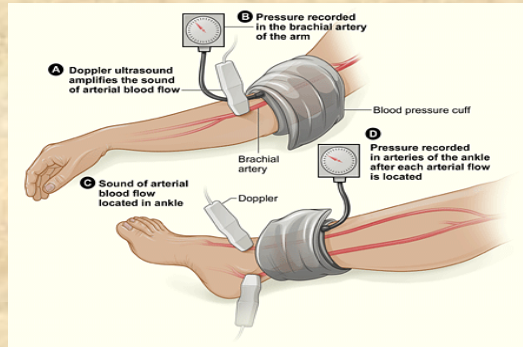
- Affymetrix platform (approach 1,000.000 SNPs)

Physiological Phenotypes

Heart Rate Variability



Ankle Brachial Index



24hr Blood pressure monitoring



Pulse wave velocity



Carotid US



Sleep study

What is the Genetic Contribution for CV Phenotypes?

Table 2: Heritability (h^2) estimates, and standard error, for each quantitative trait for cardiovascular risk factors (in percentage), unadjusted (Model 1), adjusted for age, sex, age², and age × sex interaction (Model 2), and adjusted for age, sex, age², and age × sex interaction, and current medication (Model 3).

	Model 1*	Model 2*	Model 3*
Waist circumference (cm)	26.1(0.042)	40.1(0.049)	-
Fasting glucose (mg/dL) [?]	32.8(0.046)	34.5(0.047)	32.3(0.047)
Triglycerides (mg/dL)	25.7(0.044)	28.8(0.048)	28.4(0.048)
HDL-cholesterol (mg/dL)	31.2(0.045)	32.0(0.046)	32.3(0.046)
Systolic blood pressure (mmHg) [?]	15.0(0.036)	25.9(0.042)	20.6(0.042)
Diastolic blood pressure (mmHg) [?]	16.4(0.037)	26.2(0.044)	21.0(0.044)
Total Cholesterol (mg/dL) [?]	28.6(0.047)	29.2(0.049)	30.4(0.049)
LDL-cholesterol (mg/dL) [?]	26.3(0.048)	26.3(0.049)	27.2(0.050)
BMI (kg/m ²)	39.1(0.045)	51.0(0.048)	-

* p-value < 0.0001

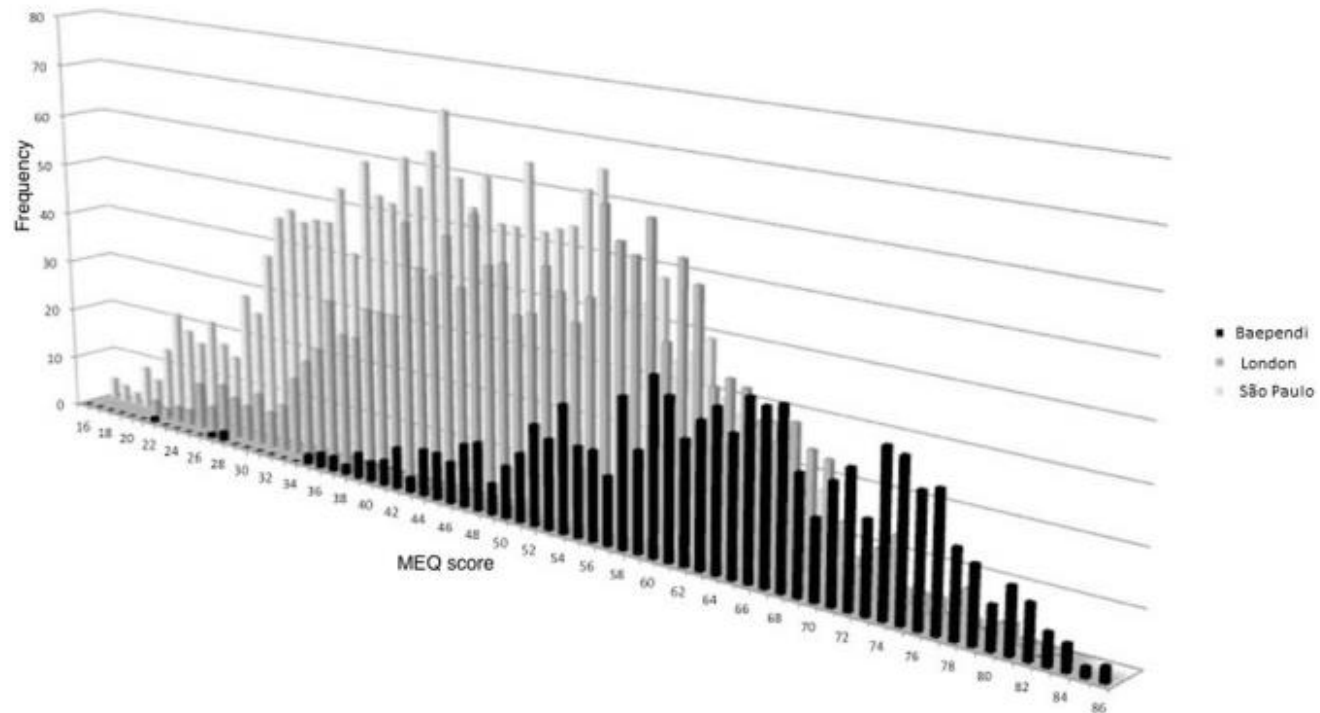
? using tdist procedure in SOLAR

Heritability

Característica	h^2 (s.e.)	Household (s.e.)	Covariáveis	Var (cov)
Weekly PA	0.40 (0.12)	-	-	-
Sedentarism	0.11 (0.06)	0.08 (0.06) <i>n.s.</i>	idade, idade x sexo, idade ² , idade ² x sexo	0.022
Smoking	0.32 (0.14)	-	idade	0.013
Heart Rate	0.22 (0.07)	0.05 (0.06) <i>n.s.</i>	idade, sexo	0.035
Sokolow_ly	0.25 (0.05)	0.06 (0.04)	idade, sexo, idade x sexo, idade ² x sexo	0.078
Cornell	0.11 (0.07)	0.03 (0.07) <i>n.s.</i>	idade, sexo, idade x sexo	0.088
Duration_qrs	0.15 (0.04)	0.07 (0.04) <i>n.s.</i>	idade, sexo, idade ²	0.043
Interval_qt	0.22 (0.05)	-	idade ²	0.004
Interval_qtc	0.11 (0.06)	-	idade, sexo	0.092
Interval_pr	0.37 (0.05)	-	idade, sexo, idade x sexo	0.077
Pr_segment duration	0.34 (0.05)	-	idade, idade x sexo, idade ²	-
Menarche	0.38 (0.07)	-	-	-

Distribution and heritability of diurnal preference (chronotype) in a rural Brazilian family-based cohort, the Baependi study

Malcolm von Schantz^{1,2}, Tamara P. Taporoski^{2,3}, Andréa R. V. R. Horimoto², Nubia E. Duarte², Homero Vallada³, José E. Krieger², Mario Pedrazzoli⁴, André B. Negrão^{2,3} & Alexandre C. Pereira²



821k UKBiobank Markers

388k Markers Removed

- 376.7 Genome-Wide Coverage Markers for Caucasian Europeans (~280K rare EUR markers)
- 11.3 k Neanderthal

335K Markers added

Genome-Wide Coverage Markers to target MAF > 5% for YRI

433k UKBiobank Markers Added

59K Markers Added
exome content polymorphic in YRI and MXL

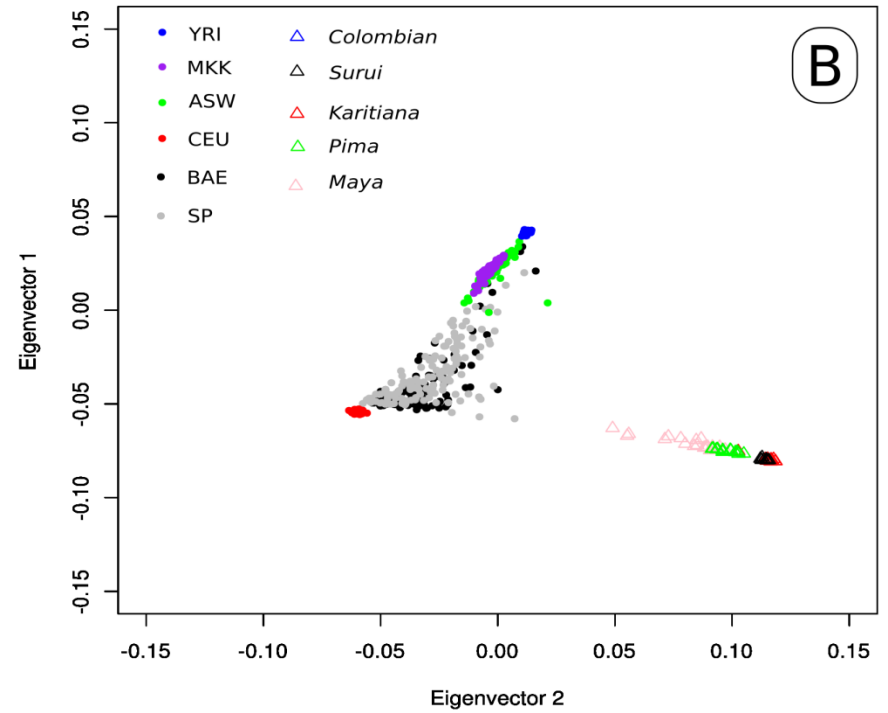
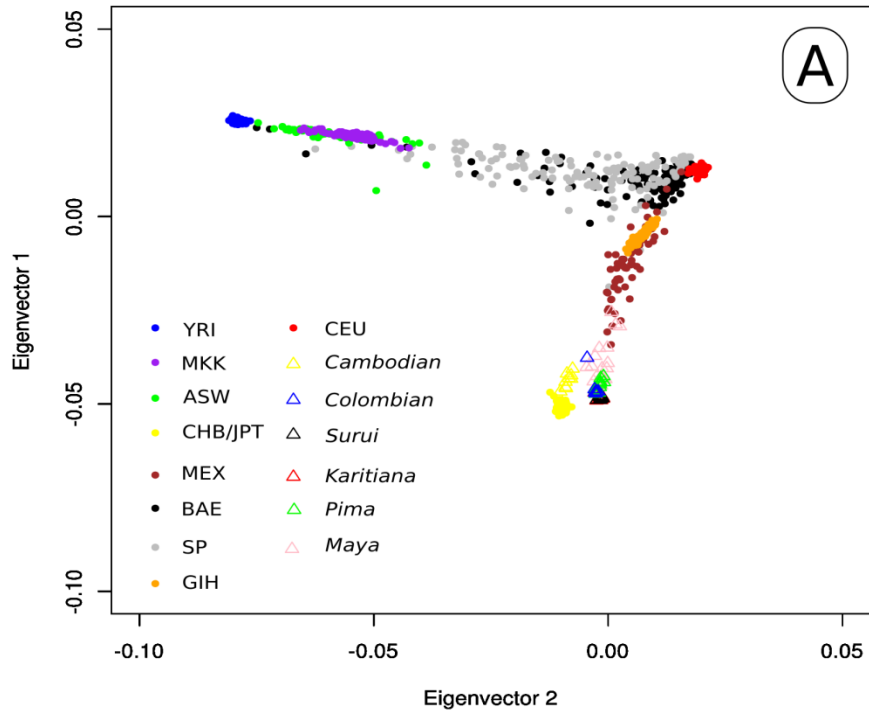
827K Markers
Axiom_InCor_BB array



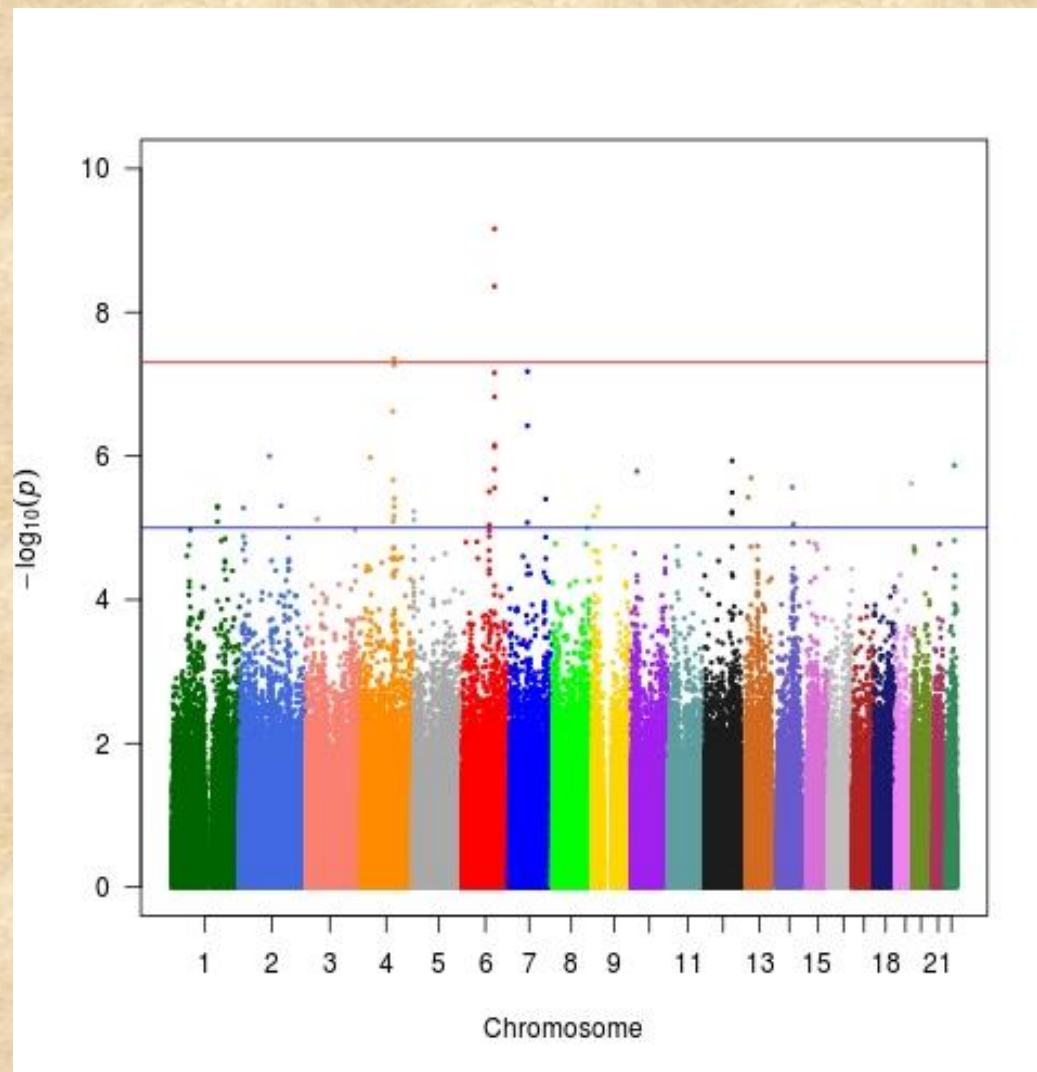
Category	Number of markers
Markers of Specific Interest	
Alzheimer's Disease	803
ApoE	1,147
Autoimmune/Inflammatory	258
Blood Phenotypes	2,545
Cancer common variants	343
Cardiometabolic	377
eQTL	17,115
Fingerprint	262
HLA	7,348
KIR	1,546
Lung function phenotypes	8,645
Common mitochondrial DNA variants	180
Neurological disease	19,791
NHGRI GWAS catalog	8,136
Pharmacogenetics/ADME	2,037
Tags for Neanderthal ancestry	11,507
Y chromosome markers	807
Rare variants in cancer predisposition genes	6,543
Rare variants in cardiac disease predisposition genes	1,710
Rare, possibly disease causing, mutations	13,729
CNV regions for developmental delay, neuropsychiatric disorders and lung function	2,369
Rare coding variants	
Protein truncating variants	30,581
Other rare coding variants	80,581
Genome-wide coverage	
Genome-wide coverage for common variants	348,569
Genome-wide coverage for low frequency variants	280,838
Total number of markers	820,967

mean.r2			
maf.range	pop	UKbiobank	Axiom_InCor_B B
[0.05,0.5]	YRI	0.812	0.901
[0.05,0.5]	CEU	0.925	0.92
[0.05,0.5]	MXL	0.897	0.894
mean.r2			
maf.range		UKbiobank	Axiom_InCor_BB
[0.01,0.05)	YRI	0.741	0.73
[0.01,0.05)	CEU	0.875	0.684
[0.01,0.05)	MXL	0.844	0.697
frac.r2.ge.0.8			
maf.range		UKbiobank	Axiom_InCor_BB
[0.05,0.5]	YRI	0.639	0.869
[0.05,0.5]	CEU	0.903	0.89
[0.05,0.5]	MXL	0.855	0.847
frac.r2.ge.0.8			
maf.range		UKbiobank	Axiom_InCor_BB
[0.01,0.05)	YRI	0.541	0.538
[0.01,0.05)	CEU	0.821	0.521
[0.01,0.05)	MXL	0.768	0.539

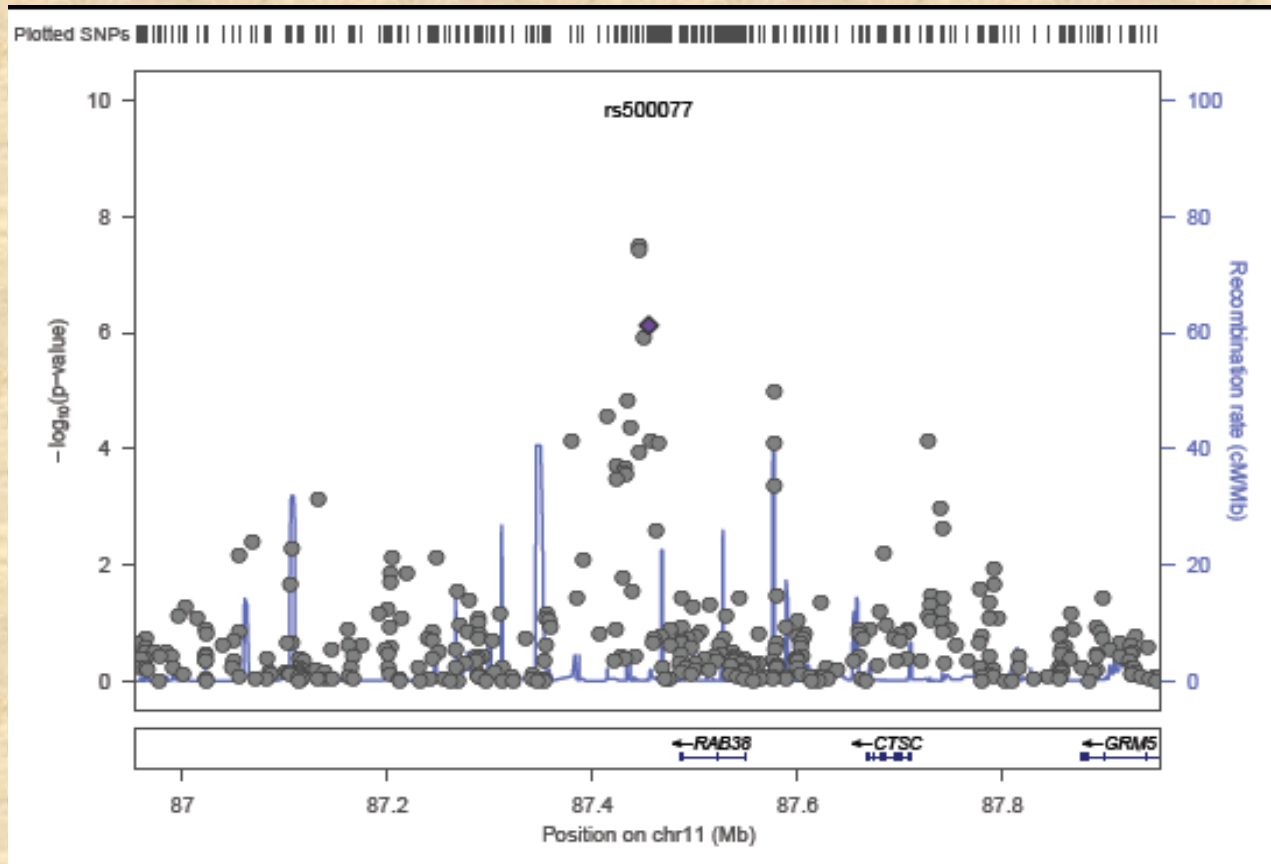
Population Structure



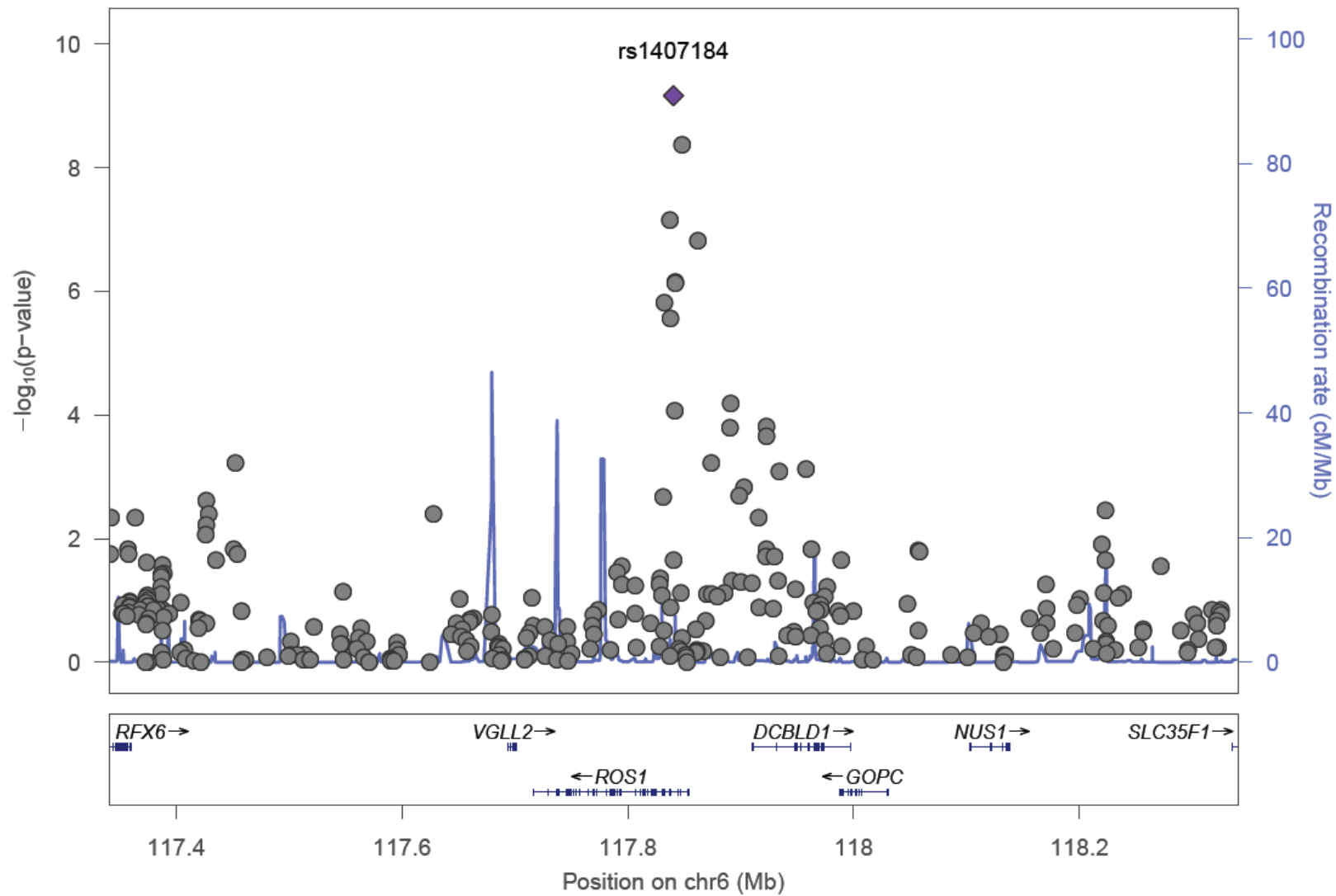
Baependi GWAS - Height



SNP do cromossomo 11

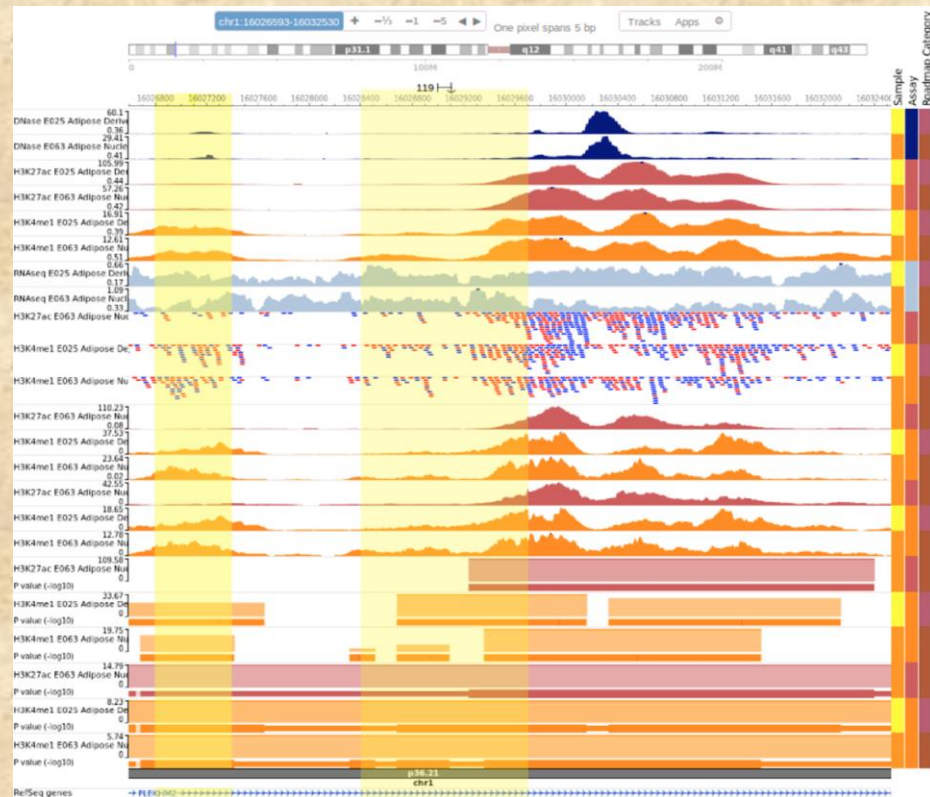
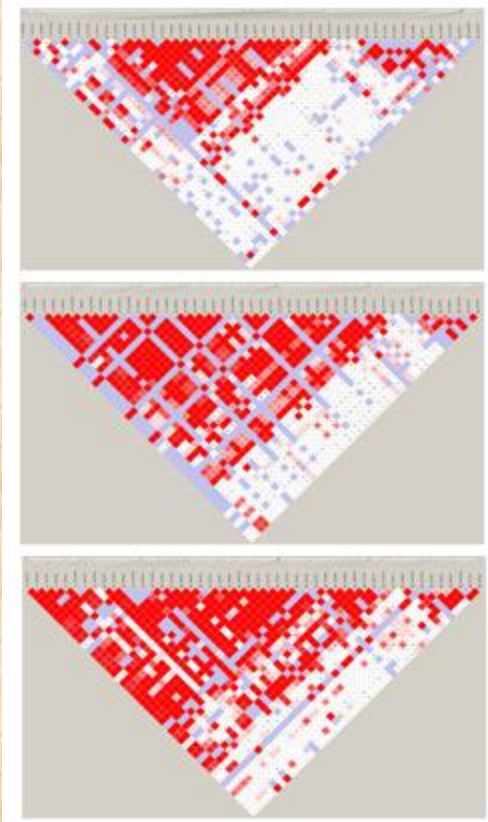


Plotted SNPs

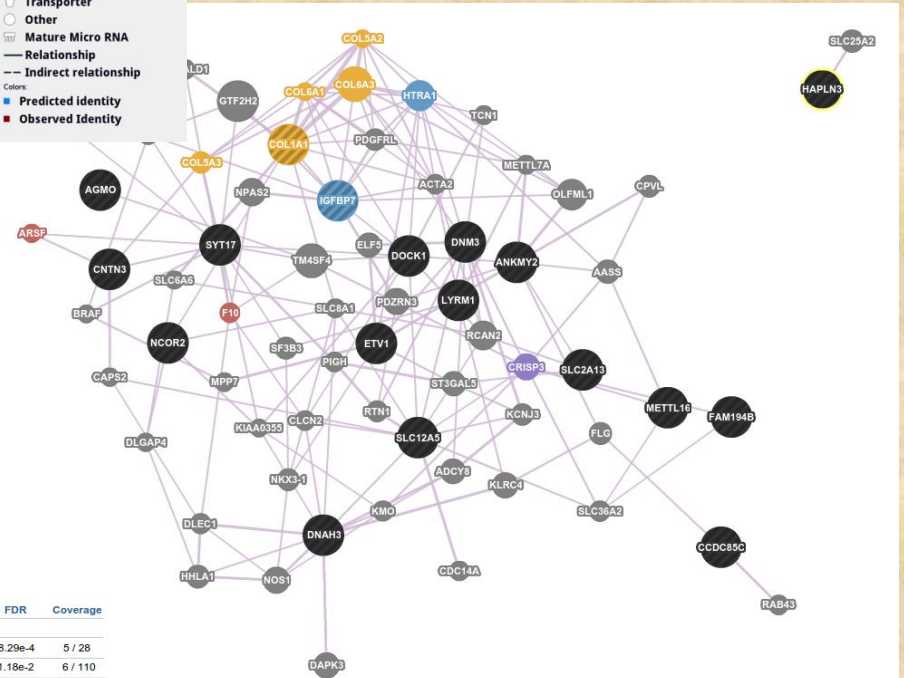
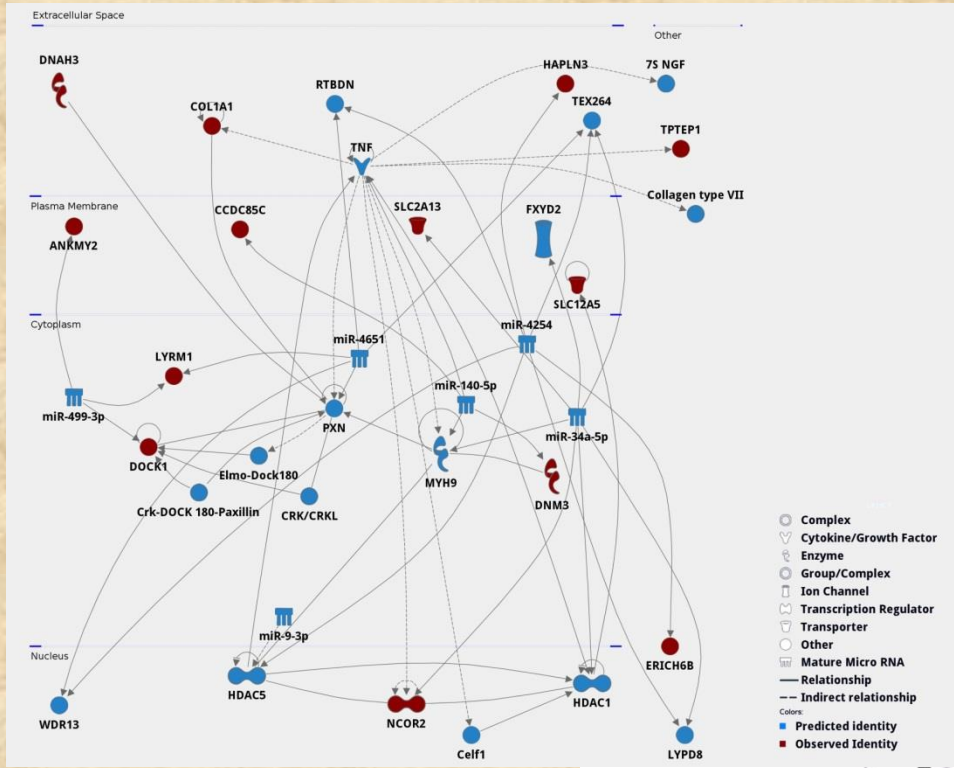


Downstream Functional Characterization of Associated SNPs:

rs548838



Systems Biology



On going actions

- NMR Brain
- 3D Face reconstruction
- Skin colour (colormeter)
- Actigraphy (ActTrust)
- 24 hour Holter studies



Acknowledgements:





Laboratório de
Genética e
Cardiologia
Molecular

Alexandre Pereira: e-mail: alexandre.pereira@incor.usp.br