Genetic variation associated with chronic disease susceptibility in the Portuguese population

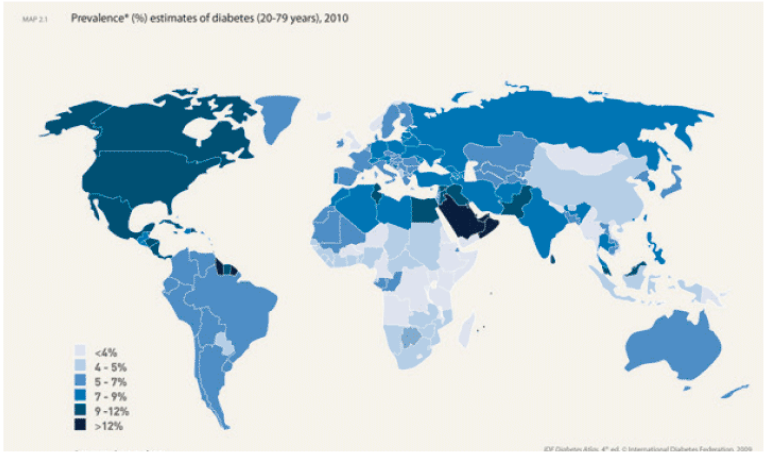
Marta Barreto
Instituto Nacional de Saúde Doutor Ricardo Jorge

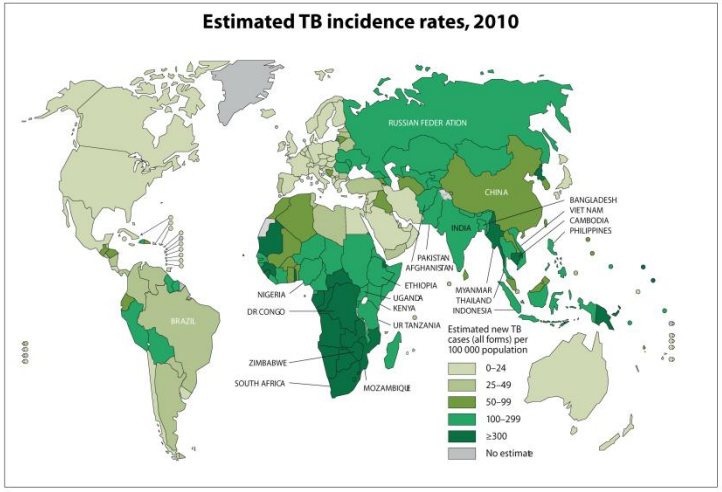
June 27th 2013

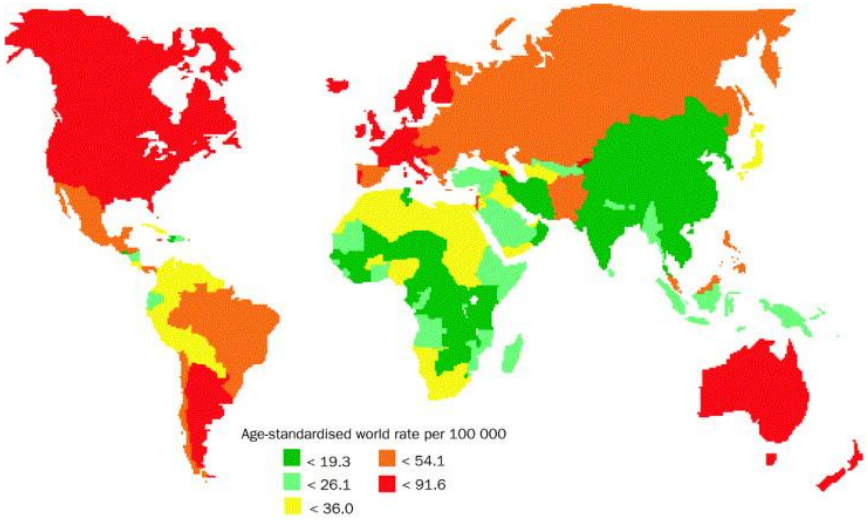
GOVERNO DE PORTUGAL
MINISTÉRIO DA SAÚDE

Instituto_Nacional de Saúde
Doutor Ricardo Jorge

Genetic variation







Objectives

- Identify genetic factors that influence the risk of prevalent chronic diseases in the Portuguese population;
- Characterize the contribution of different genetic factors to chronic disease susceptibility.

Materials and methods

1) Population sample

- 221 participants (95 men and 126 women).
- Pilot study of INSEF – "Inquérito Nacional de Saúde com exame Fisico" - the National Component of the European Health Examination Survey project.
- São Brás de Alportel Health Center
- Random sampling of participants by SNS identification number
- In accordance with EHES procedures to achieve maximum participating rates and quality of the data and samples

Materials and methods

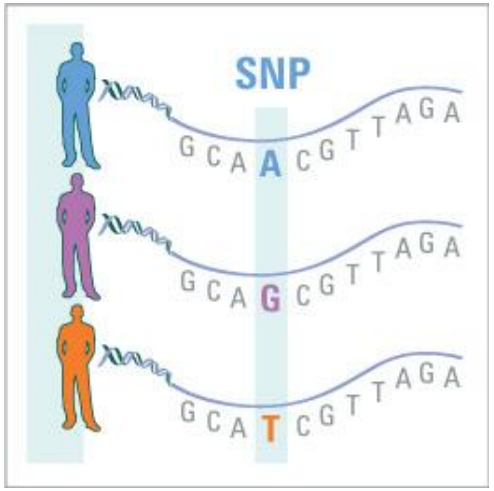
2) Phenotype characterization:

Detailed Questionnaire, (sociodemographics and occupation, medical history and general health, family history of illness focusing on chronic disorders, psychological status, and lifestyle exposures (including smoking, alcohol, physical activity and diet)

Physical exam (weight, height, waist and hip circumpherence, blood pressure)

Blood sample (Glucose, HDL, LDL, Triglycerides, GGT, ALT, AST, Creatinine, C Reactive protein) CBC + Serum, Plasma and DNA for Biobanking

3) Candidate gene analysis



Blood sample → DNA Extraction

↓

SNP analysis

3) Candidate gene analysis

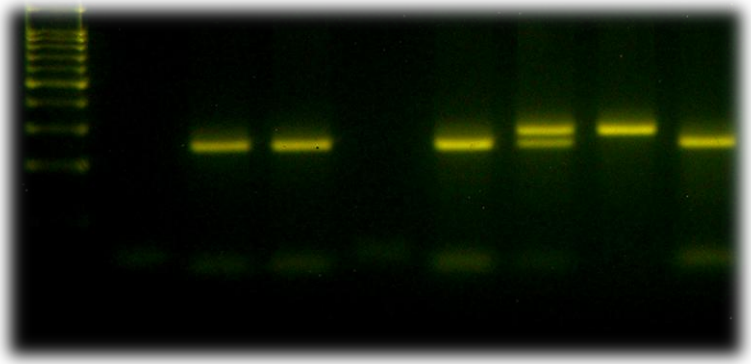
3.1 Candidate Genes SNPs selection (described in literature)

82 diferent Genes associated to:

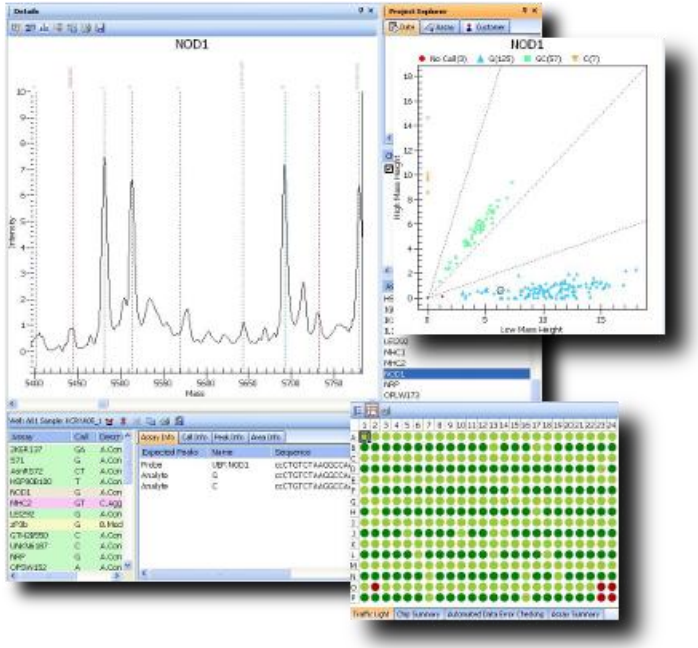
- Cancer
- Drug resistance/metabolism
- Cardiovascular diseases
- Diabetes
- Obesity
- Psychiatric disorders
- Drug addiction

Potential Public Health Impact

3.2 Genotyping:

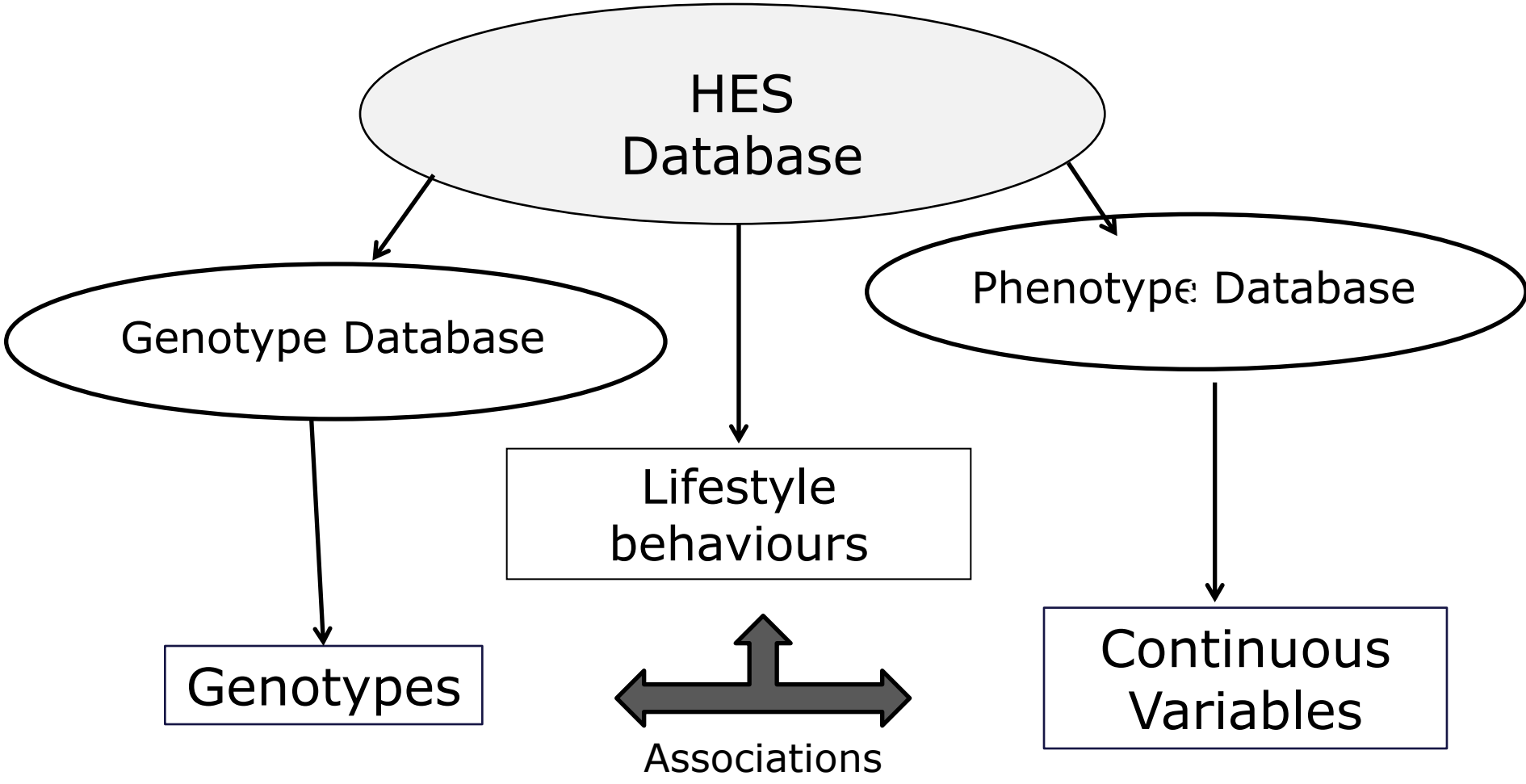


RFLPs



Sequenom-Massarrays

Materials and methods



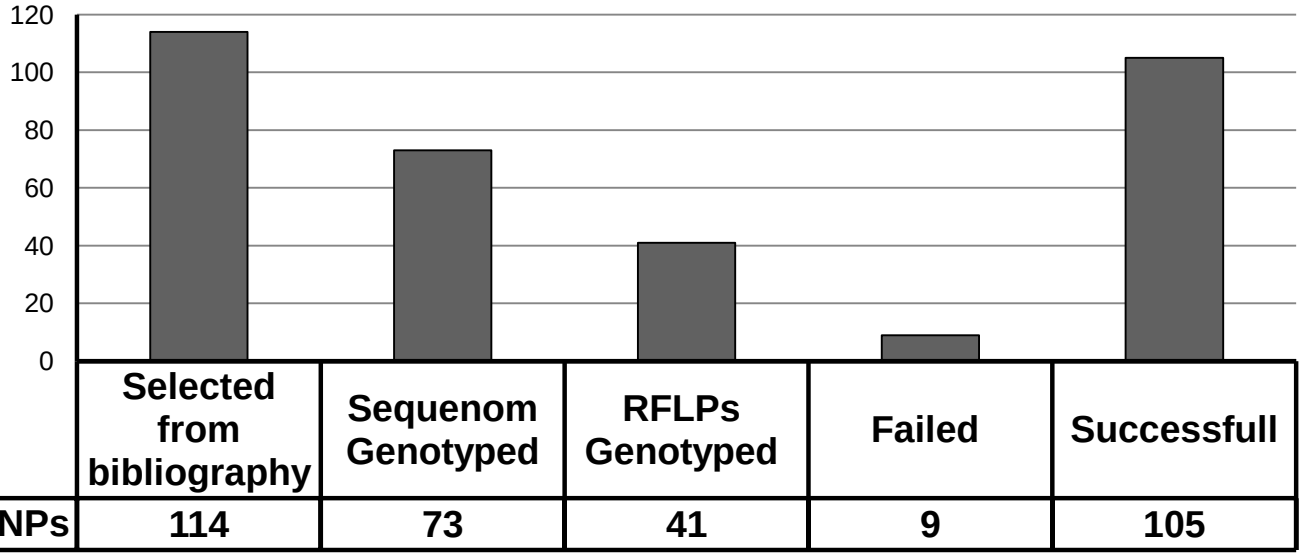
To identify genetic risk factors involved in chronic disease susceptibility, using continuous variables.

Population characterization

Table1. Characteristics of men and women participants (Data are presented as mean±SD for continuous variables and n (%) for proportions).

	Total	Men	Women	P-value*
Number of participants	206	87 (42,2%)	119 (57,8%)	
Age (years±SD)	56,31 ± 16,37	55,80 ± 16,45	56,67 ± 16,37	0,754
BMI(Kg/m^2)	27,88 ± 4,69	27,44± 4,20	28,20± 5,01	0,336
MetS [1]	95 (46,1%)	40 (46,0%)	55 (46,2%)	0,124
MetS risk factors (mean±SD)				
Waist circumference (cm)	95,50 ± 12,56	97,62 ± 1,72	93,94 ± 12,97	3,8x10^{-2}
DBP (mmHg)	80,67 ± 9,96	80,89 ± 10,02	80,52 ± 9,95	0,793
SBP (mmHg)	131,72 ± 20,02	133,00 ± 16,38	130,79 ± 22,33	0,245
HDL (mg/dL)	53,51 ± 13,33	49,61 ± 12,85	56,35 ± 12,99	2,5x10^{-4}
TG (mg/dL)	107,71 ± 60,29	115,26 ± 74,73	102,19 ± 46,61	0,717
Glucose (mg/dL)	103,29 ± 33,91	109,79 ± 47,10	98,54 ± 1,66	2,6x10^{-4}
MetS related diseases [2]				
Hypertension	54 (26,2%)	20 (23,0%)	34 (28,6%)	0,054
Type 2 Diabetes	15 (7,3%)	8 (9,2%)	7 (5,9%)	0,796
Hypercholesterelomia	26 (12,6%)	7(8,0%)	19 (16,0%)	0,019
TOTAL	95 (46,1%)	35(40,2%)	60 (50,4%)	
Medication				
Hypertension	52 (25,2%)	19 (21,8%)	33 (27,7%)	0,052
Type 2 Diabetes	13 (6,3%)	6 (6,9%)	7 (5,9%)	0,782
Hypercholesterelomia	24 (11,7%)	6 (6,9%)	18 (15,1%)	0,014
TOTAL	89 (43,2%)	31 (35,6%)	58 (48,7%)	
Smoking status				
Current smokers	37 (18,0%)	19 (21,8%)	18 (15,2%)	0,869
Former smokers	42 (20,5%)	32 (36,8%)	10 (8,5%)	0,001
Never smokers	126 (61,5%)	36 (41,4%)	90 (76,3)	1,5x10^{-6}
Regular Physical activity	80 (39,6%)	35 (40,2%)	45 (37,8%)	0,264

Genotype database



ID	SNP1	SNP2	…	SNP105
1	AA	TG	…	CC
2	AT	GG		CA
3	TT	TT	…	AA
…	…	…	…	…
…	…	…	…	…
…	…	…	…	…
208	TT	GG	…	CC

105 SNPs Genotyped SNPs

INSEF sample:208

≈61000 Genotypes

Metabolic Syndrome (MetS)

Alberti *et al*, 2009

1. Waist circumference Men≥94 cm
Women≥80 cm

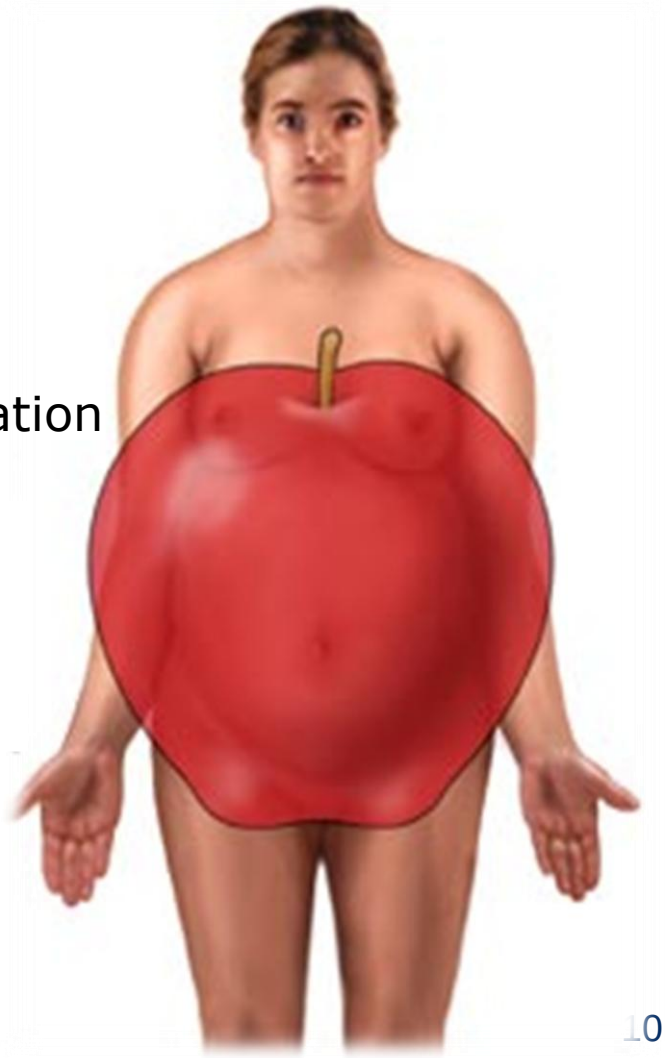
2. Blood pressure ≥130/85mmHg or Medication

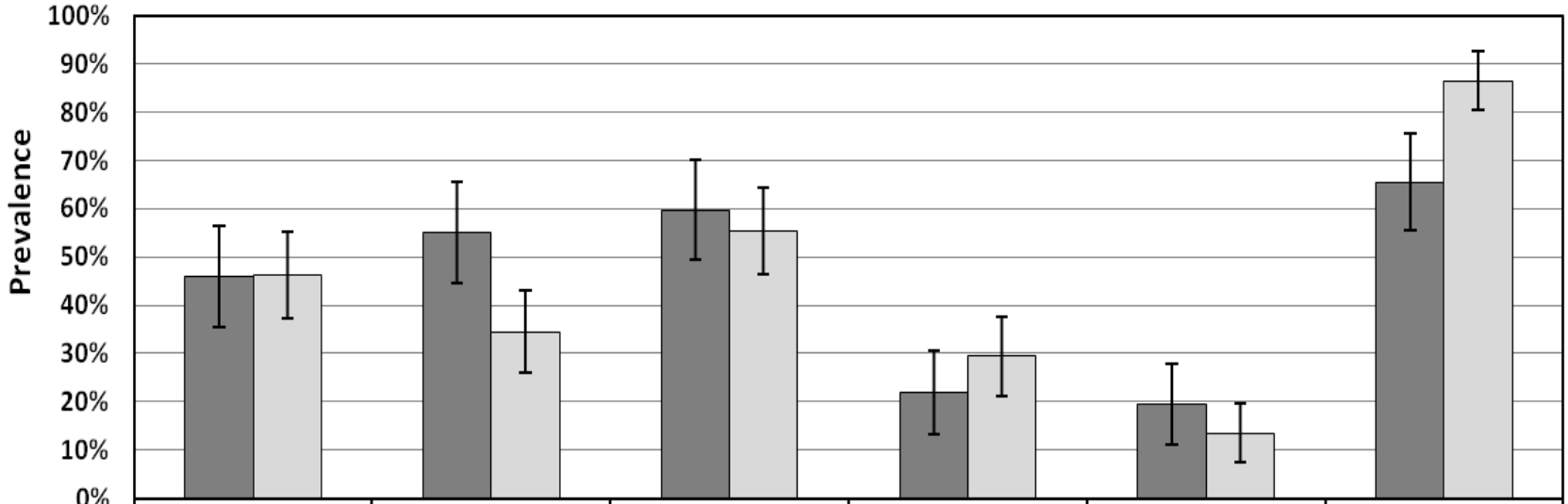
3. TG ≥150mg/dL or Medication

4. Glucose ≥100mg/dL or Medication

5. HDL Men≤40mg/dL or Medication
Women≤50mg/dL

≥3 risk factors: MetS Diagnosis





	MetS	Glucose ≥ 100 mg/dL	BP ≥ 130/85mmHg	HDL ≤ 40 mg/dL in men , 50 mg/dL in women	TG ≥ 150 mg/dL	Waist circumference ≥ 94cm in men and 80cm in women
Men	45,98%	55,17%	59,77%	21,84%	19,54%	65,52%
Women	46,22%	34,45%	55,46%	29,41%	13,45%	86,55%

Figure 1. Prevalence of MetS and its risk factors. Participants medicated for hypertension, hypercholesterelomia and diabetes were also accounted. Error bars represent the 95% confidence intervals. Abbreviations: MetS, metabolic syndrome; DBP, diastolic blood pressure; SBP, systolic blood pressure; HDL, high density lipoprotein cholesterol; TG, triglycerides

Why Metabolic Syndrome (MetS) approach?

- Simple clinical tool for predicting diabetes and CVD and the conceptual basis for understanding at least part of the pathopgysiological link between metabolic risk, future diabetes and CVD;
- Provides a framework for reserach exploring a possible unifying pathophysiological basis for the observed cluster of risk factors;
- It can guide relative risk prediction and clinical management decisions;
- It provides na easily comprehensive public health message and reminds health professionals of the need to assess related risk factors when one risk factor is detected.

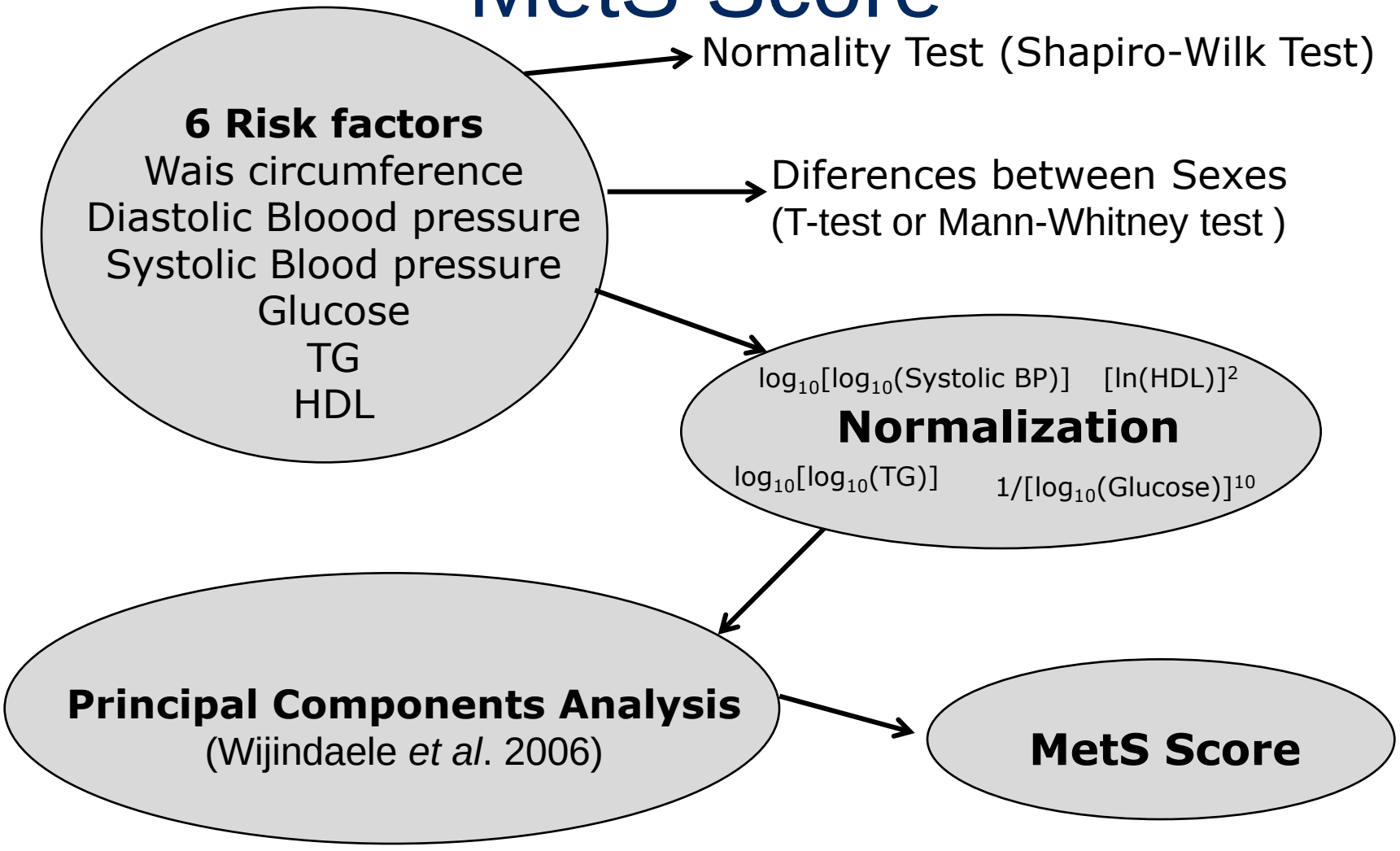
Dichotomized MetS definition vs Continuous MetS Score

Dichotomized definition of MetS enabling a yes or no diagnosis remains useful to clinical practice

BUT

in genetic epidemiological approaches , it reduces the statistical power of the MetS association studies and a continuous MetS score will be a more appropriate alternative

MetS Score



Higher MetS score
=
Less favorable MetS profile

Continuous MetS score calculation by PCA analysis

Men

- PC1 and PC2 explain 35,9% and 27,4% of MetS score variance

- measured correlations

PC1 [PC2]:
Waist circumference 0,650[0,255]
Systolic blood pressure 0,826[0,057]
Diastolic blood pressure 0,771[0,320]
Glucose -0,598[0,147]
HDL 0,079[-0,885]
Triglycerides 0,305[0,818].

Women

- PC1 and PC2 explain 36,7% and 25,1% of MetS score variance

- measured correlations

PC1[PC2]:
waist circumference 0,491[0,381]
Systolic blood pressure 0,891[0,019]
diastolic blood pressure 0,812[0,018]
glucose -0,661 [-0,251]
HDL 0,047 [-0,838]
triglycerides 0,266 [0,770]

The result MS score was 0,00±1,41 in both genders

5. Results

MetS score increases progressively with increasing numbers of risk factors (ANOVA test p<0,001).

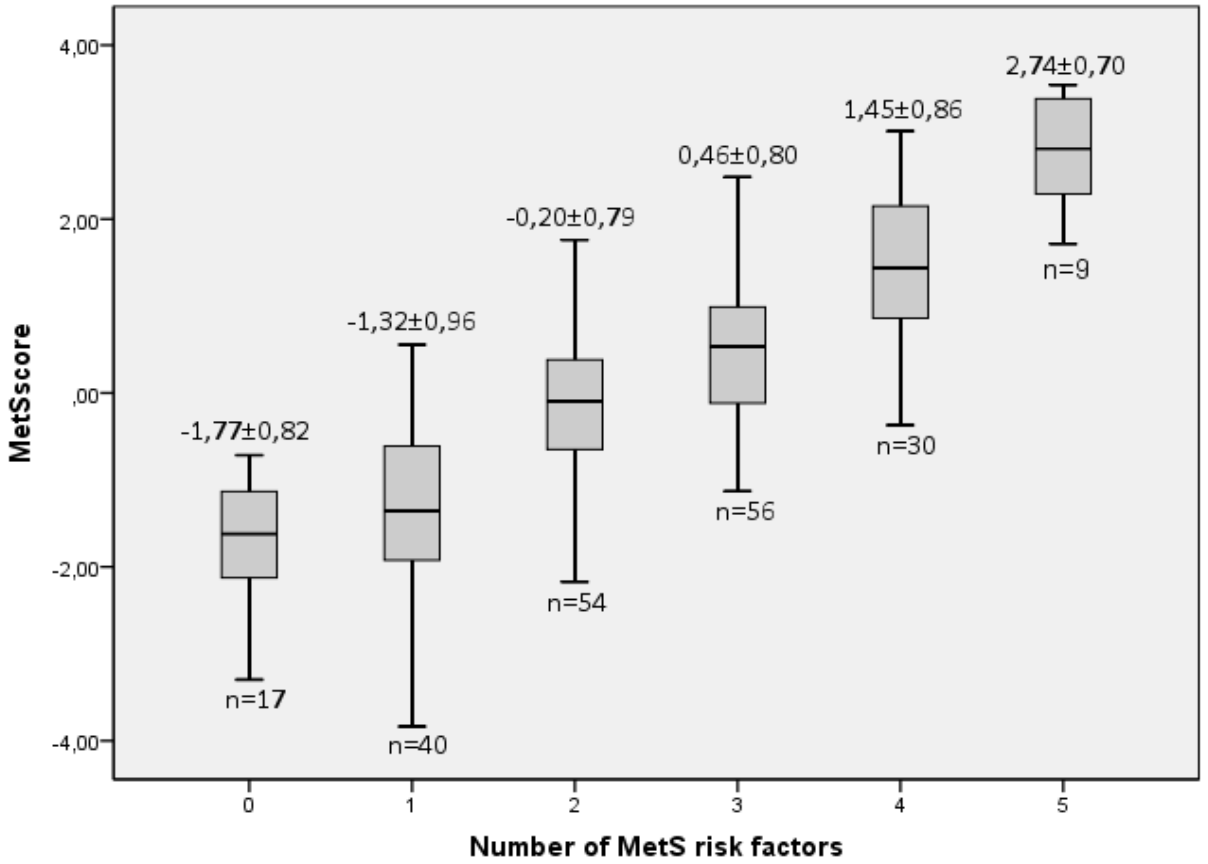
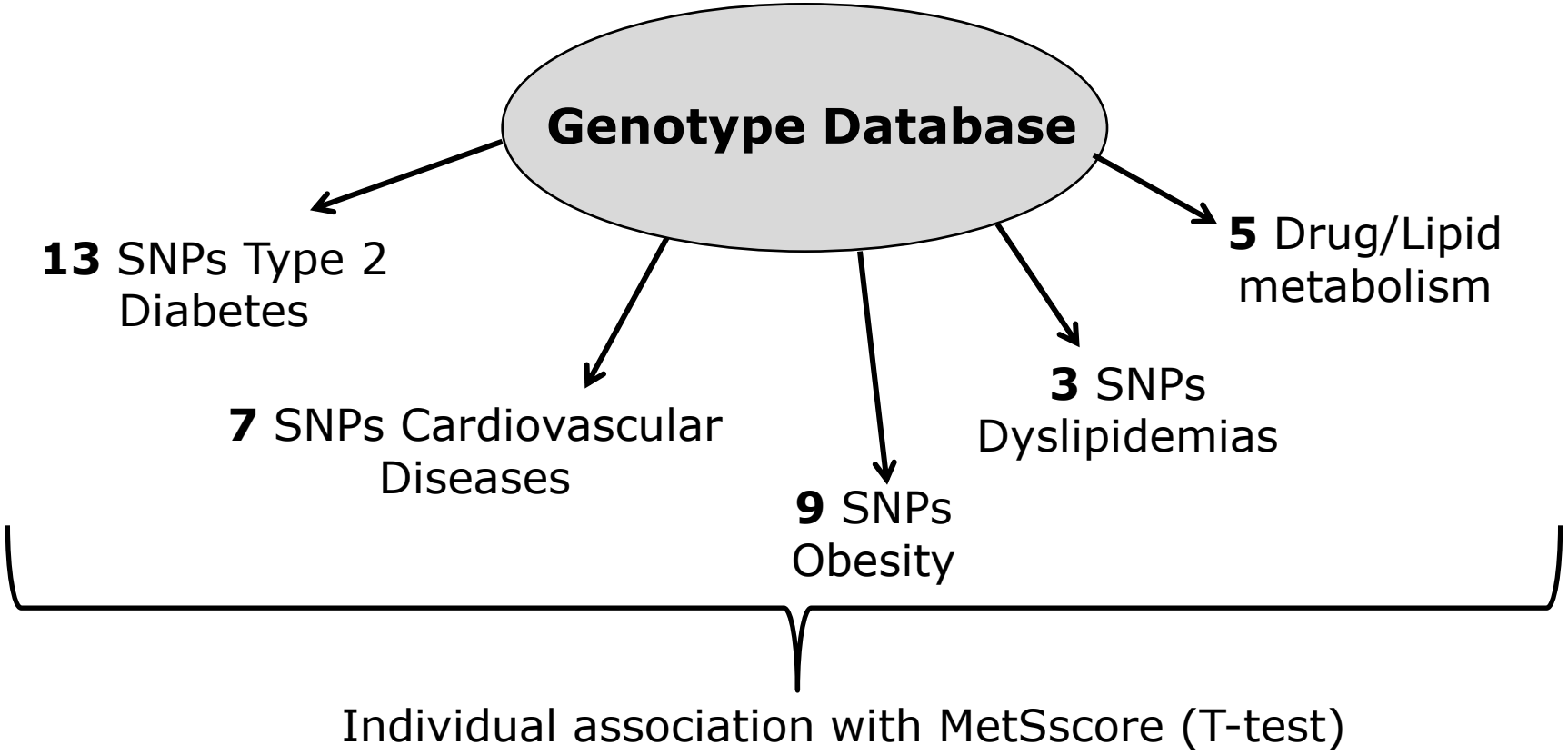


Figure 2- Variation of the MSscore descriptive statistics according to the number of risk factors. The 5 risk factors considered are those presented in the consensus MetS definition

Candidate gene analysis



Multiple testing correction (Bonferroni Test)

5. Results

Gene	NCBI ID	Alteration	Related traits	European MAF[1]	Obtained MAF
CDKAL1	rs7754840	C→G	Type 2 Diabetes	0,336	0,286
CDKN2A/B	rs10811661	C→T		0,199	0,201
HHEX	rs1111875	A→G		0,416	0,371
IGF2BP2	rs4402960	G→T		0,280	0,272
IL6	rs1800795	C→G		0,465	0,337
KCNJ11	rs5219	C→T		-	0,333
KCNQ1	rs2237892	C→T		0,075	0,051
MTNR1B	rs10830963	C→G		0,300	0,223
PPARG	rs1801282	C→G		0,076	0,093
SLC30A8	rs13266634	C→T		0,239	0,286
TCF7L2	rs7903146	C→T		0,279	0,302
ADCY5	rs11708067	A→G		0,226	0,199
KCNQ1	rs231362	C→T		0,482	0,234
ACE	rs4646994	Ins/Del	Cardiovascular diseases	-	0,420
NOS1AP	rs12143842	C→T		0,188	0,265
ADRB1	rs1801252	A→G		-	0,108
ADRB2	rs1042714	C→G		0,467	0,407
ADRB2	rs1042713	A→G		0,358	0,362
NOS3	rs1799983	G→T		0,342	0,417
NOS3	rs2070744	C→T		-	0,451

Table 2-List of SNPs selected in the present study. (Abreviations: MAF,Minor allele frequency

5. Results

Gene	NCBI ID	Alteration	Related traits	European MAF[1]	Obtained MAF
GNPDA2	rs10938397	A→G	Obesity	0,446	0,481
MTCH2	rs10838738	A→G		0,363	0,282
NPC1	rs1805081	A→G		0,467	0,288
PTER	rs10508503	C→T		0,092	0,075
SH2B1	rs7498665	A→G		0,382	0,303
FTO	rs9939609	A→T		0,449	0,361
ADRB3	rs4994	C→T		0,088	0,090
GABRA2	rs279871	A→G		-	0,434
TMEM18	rs6548238	C→T		0,150	0,127
APOE	rs7412	C→T	Dyslipidemia	-	0,027
LDLR	rs2228671	C→T		0,106	0,124
NPY	rs16147	A→G		0,491	0,450
CYP2C8	rs10509681	C→T	Drug/Lipid metabolism	0,137	0,129
CYP2C9	rs1799853	C→T		0,104	0,138
CYP2D6	rs16947	A→G		-	0,393
CYP2C19	rs4244285	G→A		0,155	0,129
TPMT	rs1142345	A→G		0,027	0,032

Table 2-List of SNPs selected in the present study (continuation).

5. Results

Table 3. Polymorphism significantly associated with MetS score. (MetS score are presented as mean±SD).

Gene	SNP ID	Genotype	n	MetSscore	P-value*	Corrected P-value*
CYP2C19	rs4244285	GG	156	0,19±1,37	0,00044	**0,016**
		AA+GA	50	-0,6±1,36		
GABRA2	rs279871	AA	63	0,37±1,35	0,013	0,487
		GG+GA	143	-0,16±1,41		
NPY	rs16147	AA	58	0,38±1,63	0,017	0,612
		GG+GA	148	-0,15±1,29		
TPMT	rs1142345	AA	192	-0,07±1,38	0,0098	0,360
		GA	13	0,97±1,63		

*P-value were obtanied by T-test and Corrected P-value were obtained by Bonferroni Correction.

- **Multiple linear regression Models**

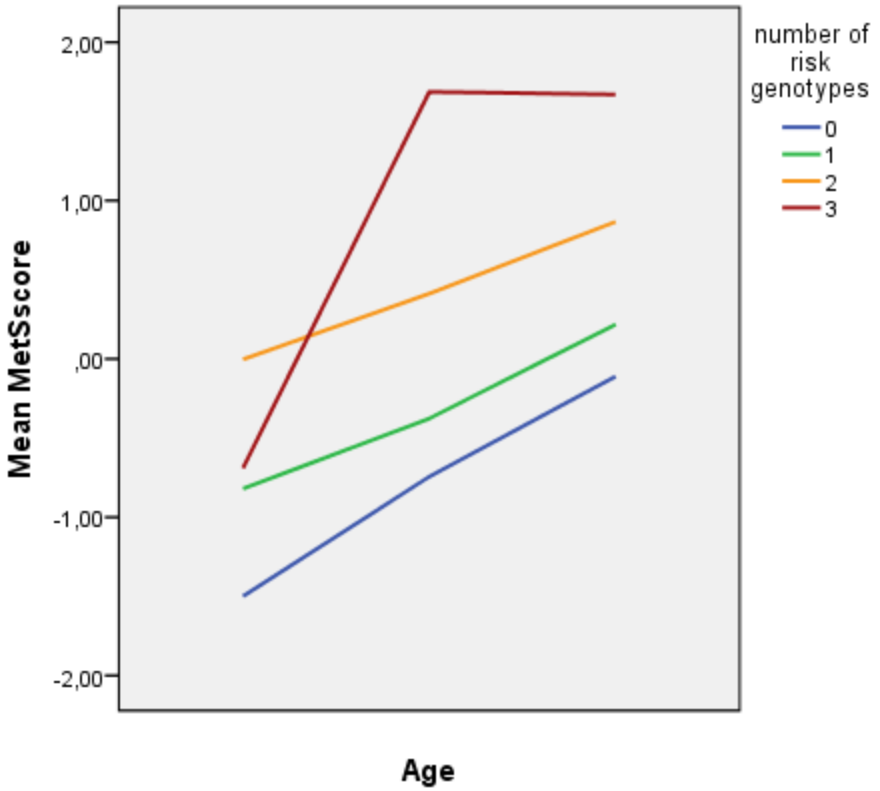
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,355	,126	,122	1,32334
2	,418	,175	,167	1,28904
3	,455	,207	,195	1,26702
4	,472	,223	,208	1,25706
5	,500	,250	,231	1,23807

1 Age
2 Age+CYP2C19
3 Age+CYP2C19+GABRA2
4 Age+CYP2C19+GABRA2+NPY
5 Age+CYP2C19+GABRA2+NPY+TPMT

ANOVA test $p<0,001$ for all the models

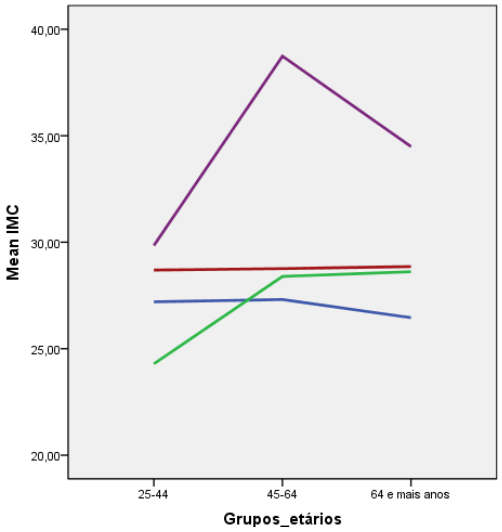
- Age is as an important risk factor.
- Age + 4 genetic variants explain 23% of the MetS score variation.

Additive genetic effects

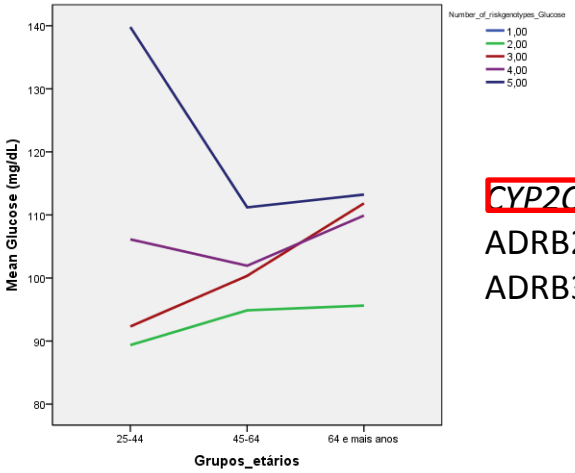


No association was found with environmental factors – lack of statistical power

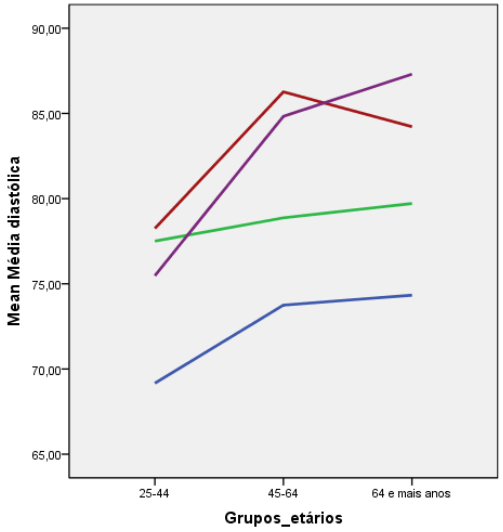
Underlying phenotypes



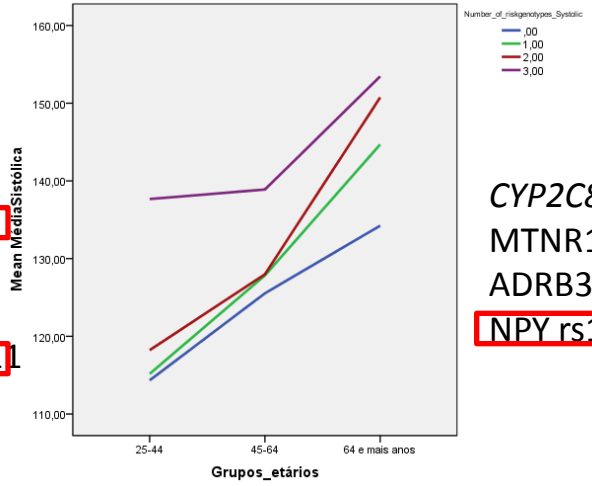
GABRA2 rs279871 *P*=0.026
NPC1 rs1805081 *P*=0.039
NPY rs16147 *P*=0.024



CYP2C19 rs4244285 *P*=0.014
ADRB2 rs1042713 *P*=0.027
ADRB3 rs4994 *P*=0.040



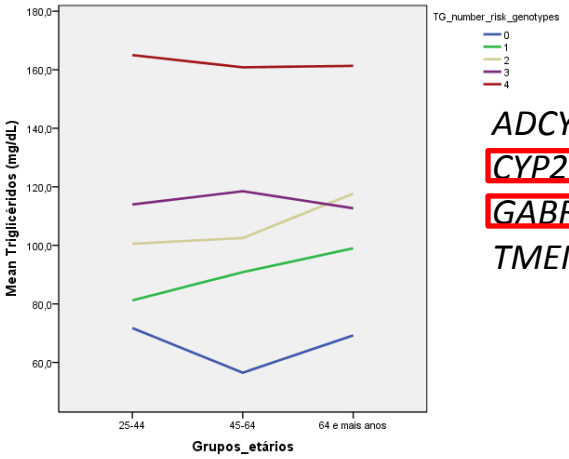
GABRA2 rs279871 *P*=0.014
ADRB2 rs1042713 *P*=0.027
ADRB3 rs4994 *P*=0.040
CYP2C19 rs4244285 *P*=0.011



CYP2C8 rs10509681 *P*=0.017
MTNR1B rs10830963 *P*=0.048
ADRB3 rs4994 *P*=0.040
NPY rs16147 *P*=0.005

Multiple regression models only explain 5-10% of the phenotype variance

Additive genetic effects

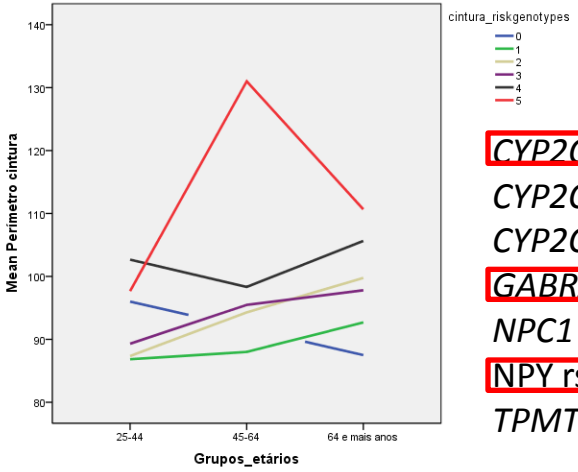


ADCY rs11708067 *P*=0.0056

CYP2C19 rs4244285 *P*=0.0046

GABRA rs279871 *P*=0.0455

*TMEM*18 rs6548238 *P*=0.0078



CYP2C19 rs4244285 *P*=0.0280

CYP2C8 rs10509681 *P*=0.0400

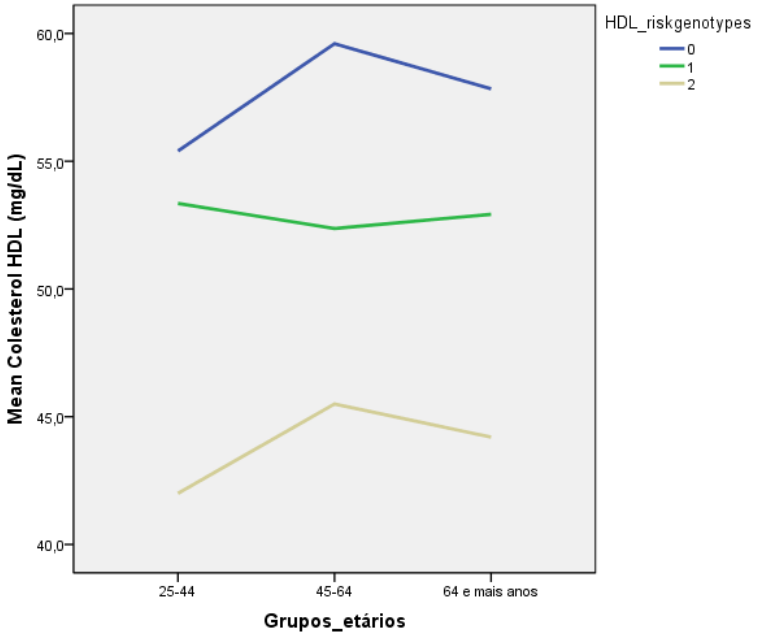
CYP2C9 rs1799853 *P*=0.0130

GABRA rs279871 *P*=0.0200

NPC1 rs1805081 *P*=0.0240

NPY rs16147 *P*=0.0350

TPMT *P*=0.0240



CYP2C19 rs4244285 *P*=0.0366

TPMT *P*=0.0202

Multiple regression models only explain 5-10% of the phenotype variance

Conclusions

- The quatitative MetS score has more power to detect association than the tradicional MetS dichotomous definition;
- We have found a significant association between genetic variants in the *CYP2C19, GABRA2, NPY* and *TPMT* genes and the MetS quantitave score;
- Age + 4 genetic variants explain 23% of the MetS score variation;
- These genes are possibly involved in a pathophysiological mechanism responsible for the clustering of metabolic risk factors;
- No association is found between the phenotype using the traditional MetS definition and the analysed genetic variants;
- No association was found with environmental factors, likely due to lack of statistical power.

Genetic susceptibility to *Influenza* infection

- Infectious disease mortality risk has a heritable component. Children of parents who died of an infectious disease are 6x more likely to die from an infectious cause compared with the general population;
- An investigation of the influenza death records over the past 100 years in the population of Utah provided evidence for an increased risk in close and distant related relatives
- In some recent familial clusters of H5N1 infection, fatal cases curiously clustered among relatives .
- More recently, a pilot study of host genetic variants associated with influenza-related deaths among children and young adults has revealed that individuals who died of influenza had low producing Mannose-binding lectin 2 (MBL2) genotypes conferring increased risk for Methicillin-resistant Staphylococcus aureus (MRSA) co-infection .

Olsen et al., 2005
Ferdinands et al., 2011
Horby et al., 2010

Objectives

- To identify and characterize host and virus genetic factors that influence susceptibility, severity and outcome of 2009 pandemic influenza A (H1N1) and to identify host-virus additive and non-additive interactions that would lead to increased susceptibility, severity or outcome of this infection.

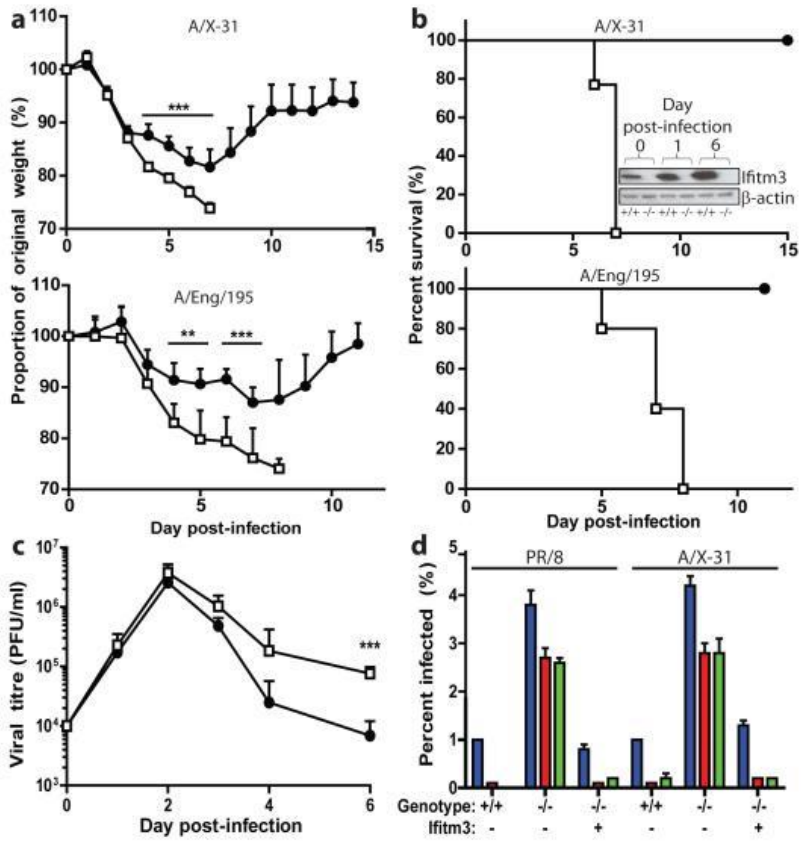
Materials and methods

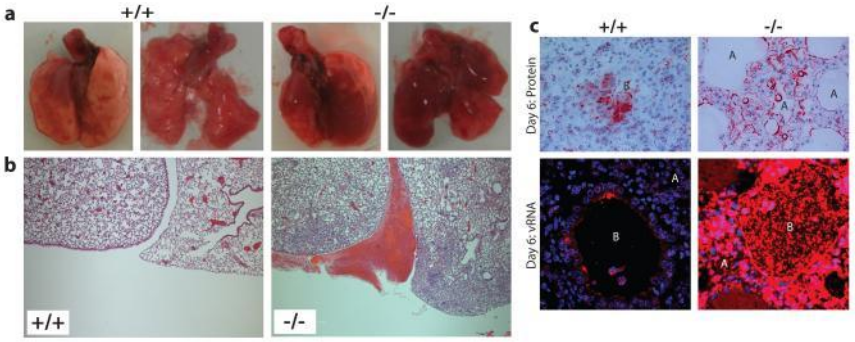
1. Perform a case-control genetic association study using the nasal swab samples that have been collected and sent to the INSARJ for diagnostic purposes during the A (H1N1) pandemic on the context of the National Influenza Surveillance Program (NISP) – targeted sequencing of candidate genes;
2. Dissect viral genetic diversity by sequencing genomic segments of 150 virus present in randomly selected samples in each of the previously established groups of influenza cases (mild and severe);
3. Analyze how host and viral genetic variation interact to influence disease susceptibility and/or severity.

Study design
2009 Pandemics

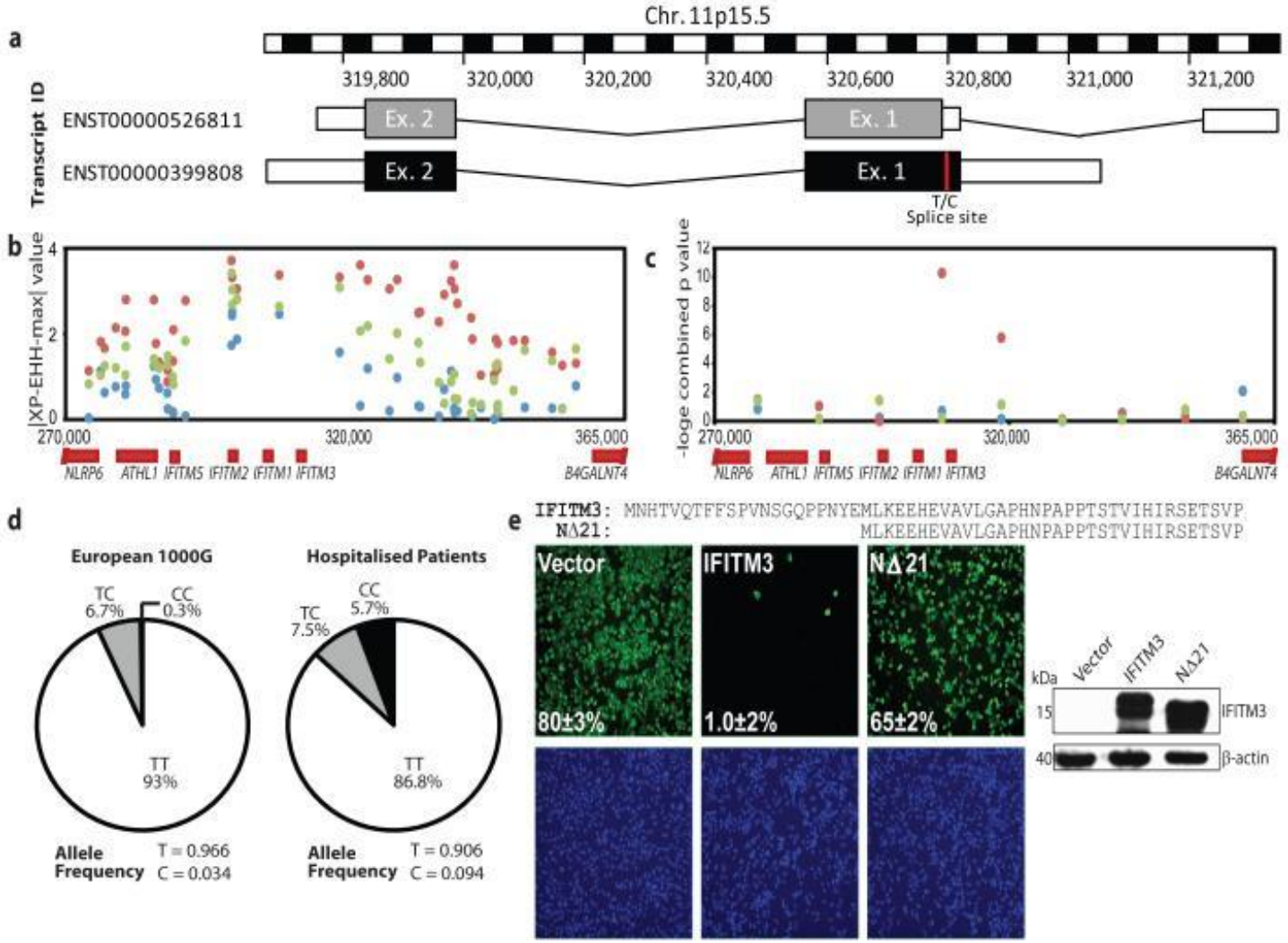
	Severe cases (Hospitalized)	Mild cases (Non hospitalized)
ILI H1N1 influenza virus positive	96 (56M+40F)	212(115M+97F)
ILI - H1N1 influenza virus negative	198 (110M+88F)	403 (217M+186F)

IFITM3 and influenza infection





Everitt et al., 2012



Everitt et al., 2012

Acknowledgements

INSA

- Vânia Gaio
- Vânia Francisco
- Ana Paula Gil
- Carlos Dias
- Mafalda Bourbon
- Astrid Vicente
- Baltazar Nunes
- Pedro Pechirra
- Raquel Guiomar

ARS Algarve

- Álvaro Beleza
- Francisco Mendonça
- Filomena Horta Correia
- Aida Fernandes

BioFIG
Center for Biodiversity, Functional and Integrative Genomics

FCT
Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA EDUCAÇÃO E CIÊNCIA

PTDC/SAU-ESA/101743/2008