POOLED-COHORTS GENOME-WIDE ASSOCIATION STUDY OF METALS: THE METAL-GWAS INITIATIVE

Maria Tellez-Plaza, MD PhD

Senior Scientist,
National Center for Epidemiology,
National Institutes for Health
(Instituto de Salud Carlos III, Madrid, Spain).

Centro Nacional de Epidemiología

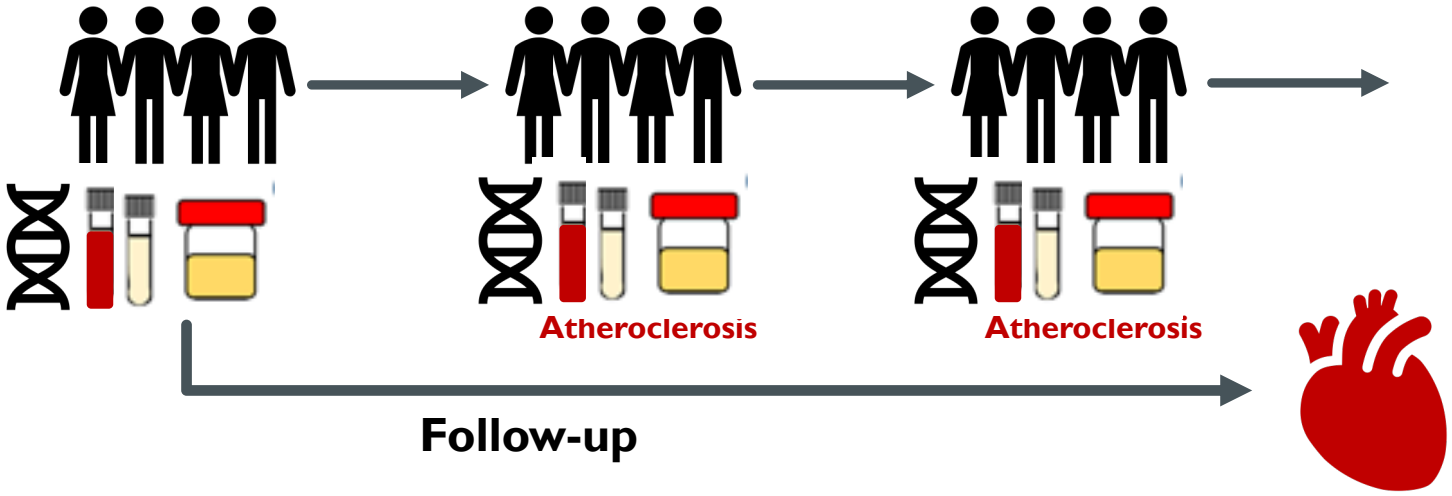
GOBIERNO DE ESPAÑA
MINISTERIO DE CIENCIA E INNOVACIÓN

OUTLINE

- I. A LITTLE BIT OF BACKGROUND
- II. THE METAL GWAS INITIATIVE: RATIONALE AND OBJECTIVES
- III. WHAT WE HAVE SO FAR
- IV. PROSPECT FOR FUTURE COLLABORATIONS IN THE SETTING OF THE METAL GWAS INITIATIVE

I. BACKGROUND: WHY METALS?

LONG-TERM TRAYECTORY ON THE EVALUATION OF METAL-RELATED CVD IN CARDIOVASCULAR COHORTS



Dr. Guallar

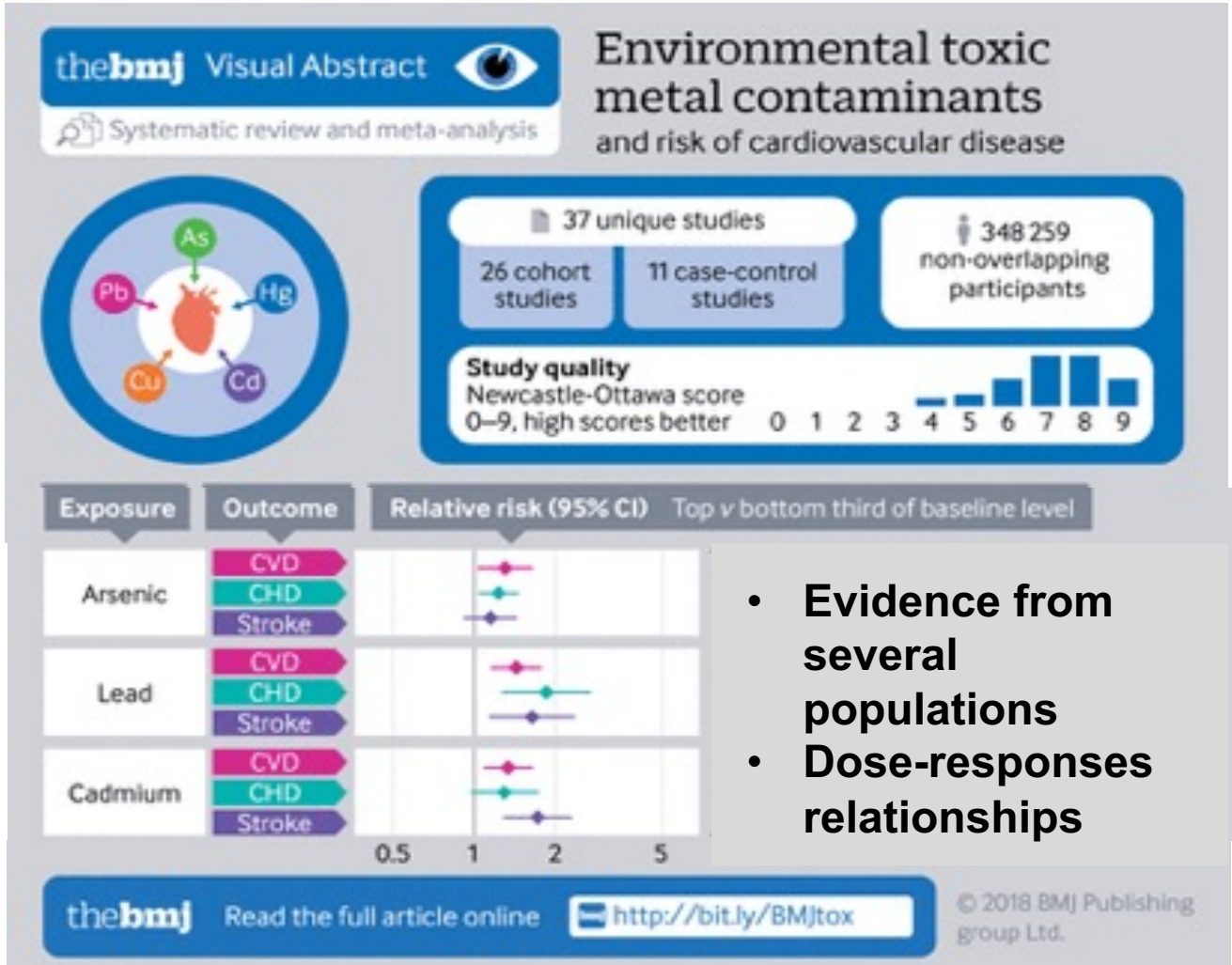
Dr. Navas-Acien

JOHNS HOPKINS UNIVERSITY

nhanes

STRONG HEART STUDY

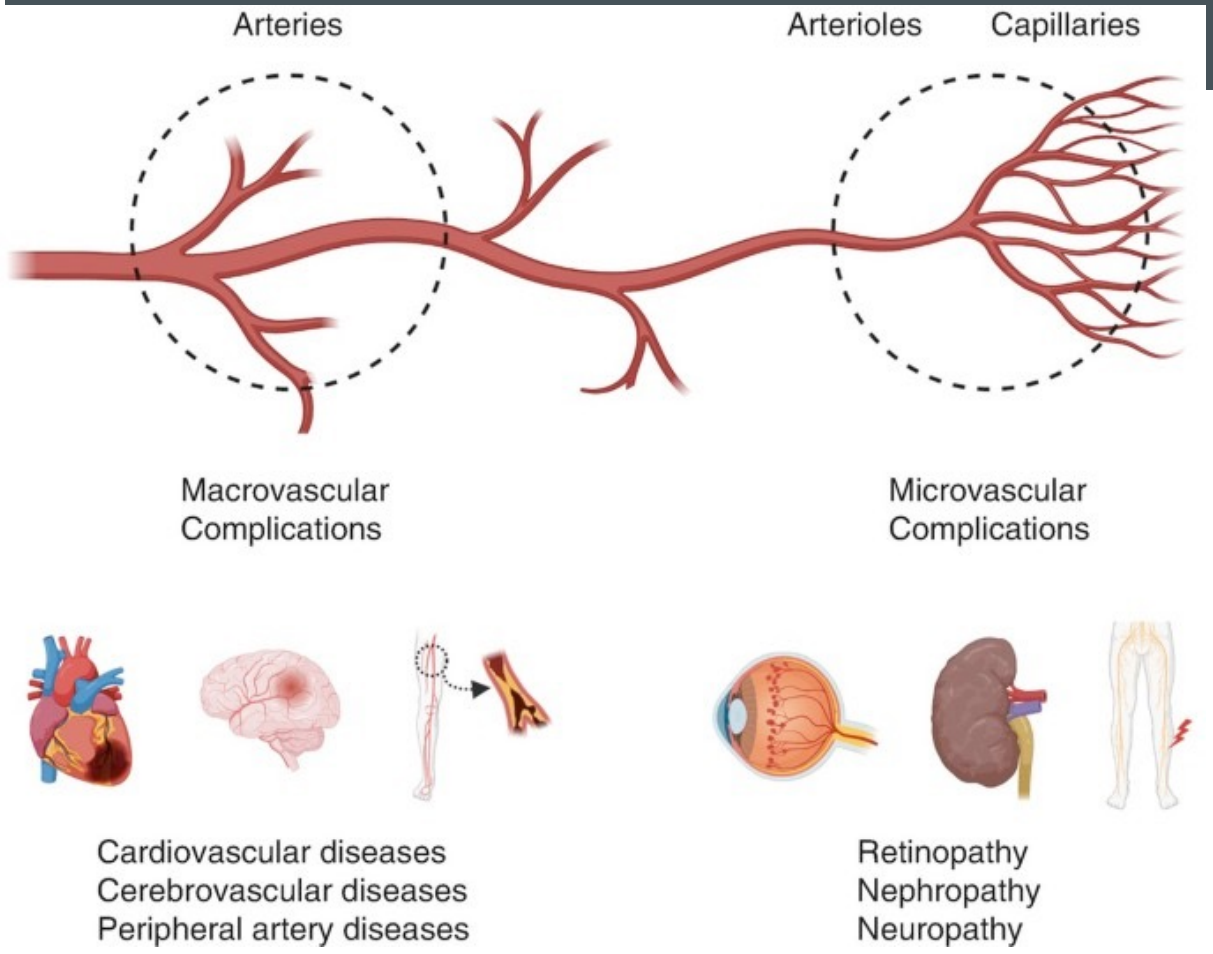
ARE CONTAMINANT METALS CARDIOVASCULAR RISK FACTORS?



- **Evidence from several populations**
- **Dose-responses relationships**

1. Number of high quality studies with low-moderate exposure levels: not sufficient at the time

2. Biological pathways involved: not completely understood



A NEED TO EXPAND THE RANGE OF METAL-RELATED CARDIOVASCULAR ENDPOINTS: FROM EARLY BIOLOGICAL EFFECT MARKERS AND ATHEROSCLEROSIS TO THE DEVELOPMENT OF CLINICAL SYMPTOMS INCLUDING MICROVASCULAR IN ADDITION TO MACROVASCULAR DISEASE

Source: Yang, W. (2021). Stroke Revisited. Springer, Singapore. https://doi.org/10.1007/978-981-16-5123-6_3

Casasnovas *et al. BMC Cardiovascular Disorders* 2012, **12**:45
http://www.biomedcentral.com/1471-2261/12/45

OPEL

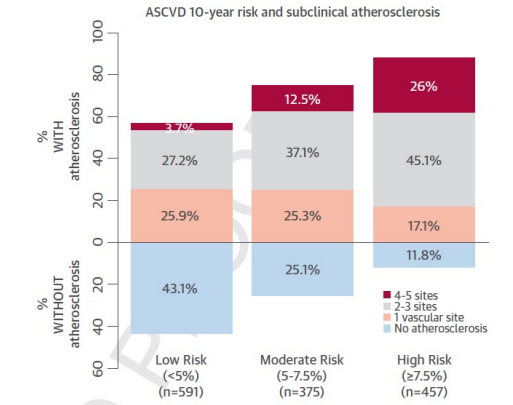
BMC Cardiovascular Disorders

STUDY PROTOCOL Open Access

Aragon workers' health study – design and cohort description

José A Casasnovas[1], Victor Alcaide[2], Fernando Civeira[3], El
Jesús Jiménez Borreguero[4,7], Martin Laclaustra[4], Montserr
Miguel Pocovi[11,12], Ginés Sanz[4] and Valentín Fuster[4,13]

FIGURE 3 Presence of Subclinical Atherosclerosis According to Traditional Risk Equations



Presence and extent of subclinical atherosclerosis according to the 10-year risk for atherosclerotic cardiovascular disease (ASCVD) calculated using the Pooled Cohort Equations, classified as low (<5%), moderate (5% to 7.5%), or high (≥7.5%) risk.

JOURNAL OF THE AMERICAN COLLEGE OF CARDIOLOGY VOL. 67, NO. 11, 20
© 2016 BY THE AMERICAN COLLEGE OF CARDIOLOGY FOUNDATION ISSN 0735-1097/$36.
PUBLISHED BY ELSEVIER http://dx.doi.org/10.1016/j.jacc.2015.12.0

ORIGINAL INVESTIGATIONS

Femoral and Carotid Subclinical Atherosclerosis Association With Risk Factors and Coronary Calcium

The AWHS Study

Martín Laclaustra, MD, PhD, MPH,[a,b,c] José A. Casasnovas, MD, PhD,[d] Antonio Fernández-Ortiz, MD, PhD,[a,e]
Valentin Fuster, MD, PhD,[a,f] Monserrat León-Latre, MD,[d] Luis J. Jiménez-Borreguero, MD, PhD,[a,g]

METALS AND ATHEROSCLEROSIS IN WORKERS

ARAGON WORKERS' HEALTH STUDY

Increased atherosclerosis

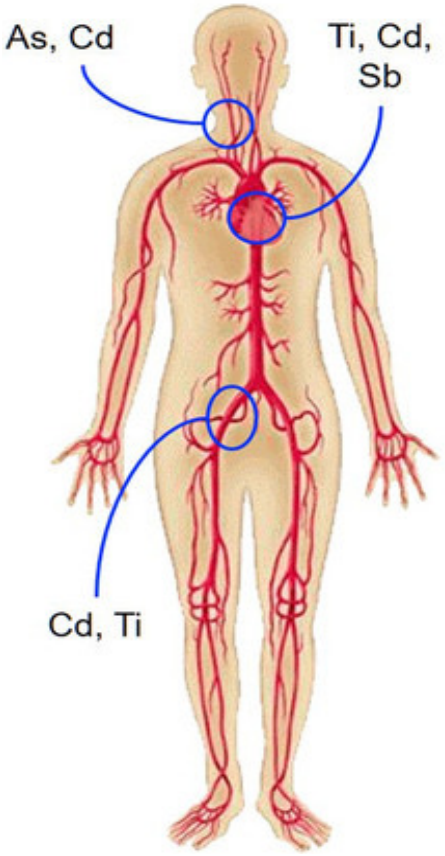


Table 2. OR (95% CI) of Subclinical Atherosclerosis Presence Comparing the 80th to the 20th Percentile of Urine Metal Distribution in Adult Participants From the Aragon Workers Health Study

	At least one territory (cases/noncases=1341/447)*		Carotid (cases/noncases=659/1138)		Femoral (cases/noncases=987/737)		Coronary (cases/noncases=691/1177)	
	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value
As	1.25 (1.03–1.51)	0.02	1.24 (1.05–1.47)	0.01	1.10 (0.92–1.30)	0.30	1.07 (0.91–1.27)	0.40
Cd†	1.67 (1.22–2.29)	0.001	1.38 (1.05–1.81)	0.02	1.72 (1.30–2.28)	<0.001	1.28 (0.98–1.67)	0.07
			1.06 (0.91–1.24)	0.45				
Ti	1.26 (1.04–1.52)	0.02	0.96 (0.81–1.13)	0.63	1.25 (1.05–1.48)	0.01	1.16 (1.00–1.34)	0.05
U	1.00 (0.84–1.19)	1.00	0.92 (0.79–1.07)	0.28				
V	1.06 (0.89–1.25)	0.53	1.00 (0.87–1.17)	0.95	1.08 (0.93–1.26)	0.33	1.04 (0.89–1.20)	0.65
W	0.97 (0.81–1.15)	0.69	1.00 (0.86–1.16)	0.99	1.01 (0.87–1.18)	0.86	1.02 (0.88–1.18)	0.84

Models adjusted for age (years and splines), sex (men and women), body mass index (kg/m^2), education (≤secondary education and >secondary education), smoking status (never, former, and current smoker), estimated glomerular filtration rate (mL/min per 1.73 m^2), diabetes status (no and yes), hypertension status (no and yes), and high cholesterol (no and yes). p80 and p20 cutoffs (μg/g creatinine) were 2.98 and 1.15 for As; 3.66 and 1.03 for Ba; 0.43 and 0.16 for Cd; 2.00 and 0.78 for Cr; 0.11 and 0.03 for Sb; 15.6 and 6.0 for Ti; 0.05 and 0.02 for U; 1.02 and 0.43 for V; and 0.42 and 0.12 for W. As indicates arsenic; Ba, barium; Cd, cadmium; Cr, chromium; IQR, interquartile range; OR, odds ratio; Sb, antimony; Ti, titanium; U, uranium; V, vanadium; and W, tungsten.

*Subclinical atherosclerosis in at least one territory could not be defined in 85 participants because they did not have subclinical atherosclerosis measures in all the 3 vascular territories.

†Nonlinear associations with corresponding metals modeled as restricted quadratic splines.

American Heart Association.

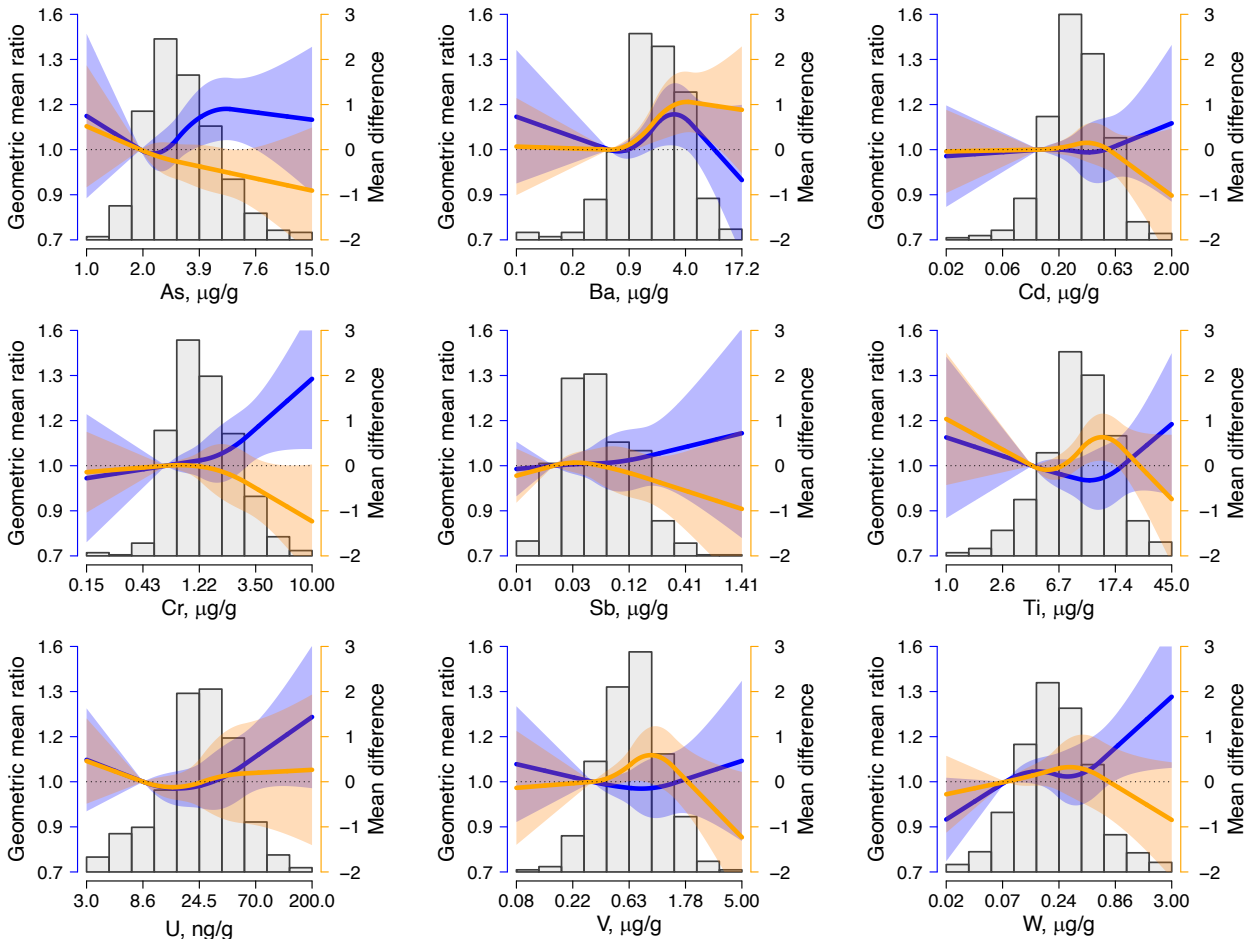
Grau-Perez et al. ATVB 2021.

METALS AND LONGITUDINAL CHANGES IN RENAL DISEASE MARKERS

Albuminuria Estimated Glomerular Filtration Rate

ARAGON WORKERS' HEALTH STUDY

Associations of urine **As, Cd, Cr and, W** with **annual change in albuminuria and estimated glomerular filtration rate** compatible with potential risk factors of renal disease at relatively low exposure levels



Grau-et al. 2022
Under review.

BMJ Open

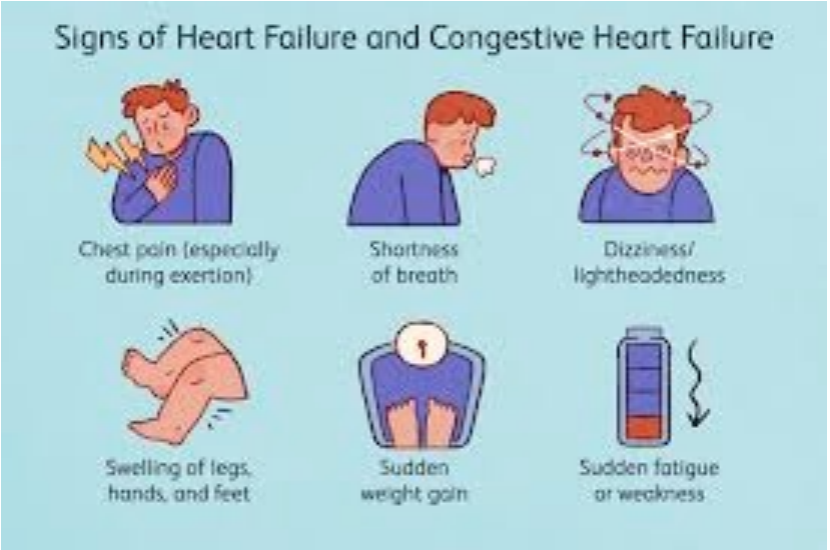
Cohort profile: the Hortega Study for the evaluation of non-traditional risk factors of cardiometabolic and other chronic diseases in a general population from Spain

Maria Tellez-Plaza,[1,2] Laisa Briongos-Figuero,[3] Gernot Pichler,[2] Alejandro Dominguez-Lucas,[2] Fernando Simal-Blanco,[3] Francisco J Mena-Martin,[3] Jesus Bellido-Casado,[3] Delfin Arzua-Mouronte,[3] Felipe Javier Chaves,[4,5] Josep Redon,[2,6,7] Juan Carlos Martin-Escudero[3]

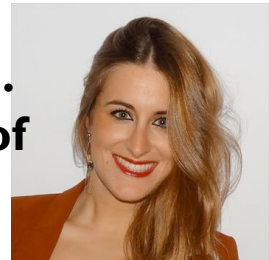
HR for heart failure (95% CI) associated with urine metals in adult participants of the Hortega Study (N=1171)

	Tertile 1, Non-cases/Cases	Tertile 2, Non-cases/Cases	Tertile 3, Non-cases/Cases	p80 vs p20	p-trend
Sb	385/16	374/12	366/18		
	1.00 (Referent)	1.16 (0.50, 2.70)	1.93 (0.92, 4.04)	1.88 (1.10, 3.24)	0.02
Ba	365/11	376/12	384/23		
	1.00 (Referent)	0.93 (0.39, 2.23)	1.26 (0.54, 2.95)	1.27 (0.63, 2.55)	0.50
Cd	382/10	374/16	369/20		
	1.00 (Referent)	2.33 (1.04, 5.23)	3.68 (1.54, 8.75)	3.95 (1.44, 10.88)*	0.008*
Cr	378/18	377/15	370/13		
	1.00 (Referent)	0.82 (0.39, 1.72)	1.34 (0.59, 3.01)	1.43 (0.81, 2.53)	0.21
V	380/19	376/15	369/12		
	1.00 (Referent)	0.80 (0.39, 1.64)	1.29 (0.54, 3.11)	1.39 (0.79, 2.44)	0.25

Models adjusted for sex, education (<high school, >=high school), smoking status (never, former and current smoker), cumulative smoking dose (0, 0-12, >12), urine cotinine (<34, 34–500, and >=500 ng/mL), estimated glomerular filtration rate (mL/minute per 1.73m^2), residence place (urban or rural), HDL cholesterol level (mg/dl), total cholesterol level (mg/dl), dyslipidemia treatment status, hypertension treatment status, diabetes mellitus of type 2 status and systolic pressure (mm Hg). *Non-linear associations with corresponding metals modelled at restricted quadratic splines.



Domingo-Relloso et al. International Journal of Epidemiology, 2019



METALS AND CVD: POTENTIAL MECHANISMS

- Increase oxidative stress and inflammation
- Affect endothelial function (interference with calcium signaling pathways)
- Endocrine disruption
- Epigenetics…

Metals and oxidative stress biomarkers

		GSSG/GSH		MDA		8-Oxo-dG	
		GMR (95%CI)	P-trend	GMR (95%CI)	P-trend	GMR (95%CI)	P-trend
Essential	Co	0.98 (0.89, 1.08)	0.70	0.99 (0.94, 1.03)	0.543	1.03 (0.98, 1.09)	0.195
	Cu	1.06 (0.95, 1.18)	0.30	1.04 (0.98, 1.09)	0.188	1.04 (0.99, 1.10)	0.142
	Mo	1.14 (1.03, 1.27)	0.01	1.01 (0.96, 1.07)	0.611	1.04 (0.94, 1.15)	0.201
	Zn	0.99 (0.88, 1.11)	0.85	1.07 (1.01, 1.14)	0.019	1.07 (1.01, 1.13)	0.02
Non-essential	**Sb**	0.99 (0.90, 1.09)	0.84	1.00 (0.95, 1.05)	0.883	1.05 (0.96, 1.15)*	0.001*
	Ba	1.17 (1.05, 1.31)	0.006	1.02 (0.96, 1.08)	0.561	1.00 (0.95, 1.06)	0.962
	Cd	1.07 (0.97, 1.19)	0.19	1.12 (1.02, 1.23)	0.02	1.09 (0.99, 1.20)*	0.08
	Cr	1.23 (1.04, 1.46)*	0.002*	0.98 (0.92, 1.03)	0.413	1.04 (0.99, 1.10)	0.131
	V	1.18 (1.00, 1.40)*	<0.001*	0.97 (0.92, 1.03)	0.288	1.03 (0.98, 1.09)	0.256

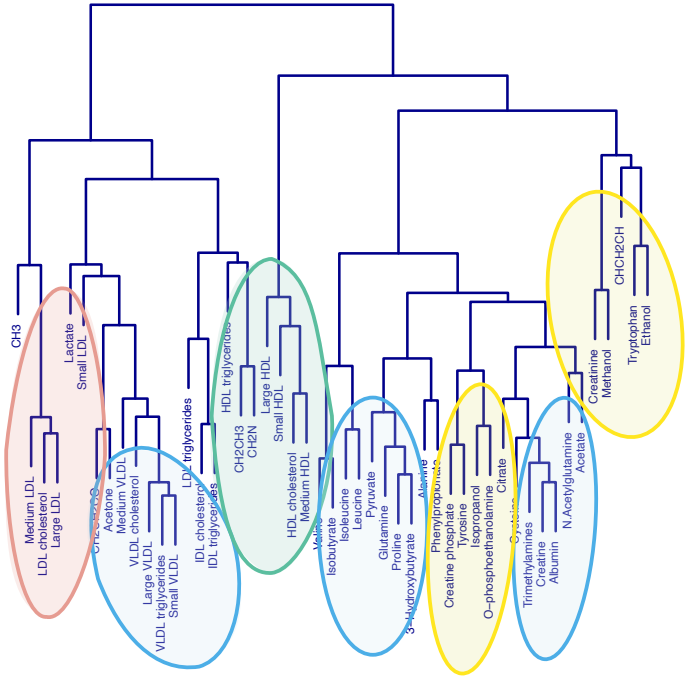
A. Domingo

M. Grau

Galvez-Fernandez, Environment International 2019.

Gene environment interaction of redox-related metals and metabolomic compounds

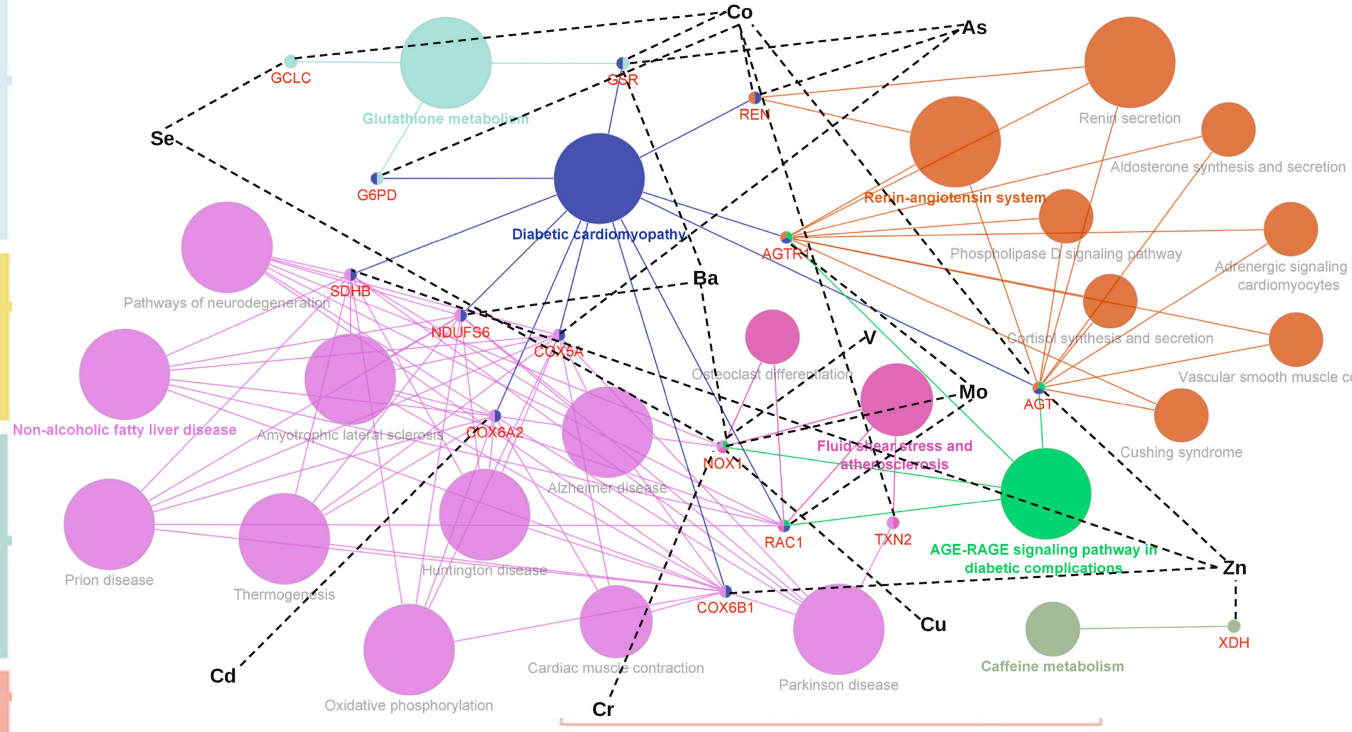
Dendrogram of metabolite cluster



Metabolic patterns (mPC1, mPC2, mPC3 and mPC4)

- **Cu and Sb positively associated with mPC1** (↑ non-essential amino acids, branched-chain amino acids, bacterial co-metabolism and ↓ VLDL subclasses and fatty acids)
- **Se, Zn and Cd inversely associated with mPC2** (↑ aromatic amino acids and bacterial co-metabolism)
- **Se and Zn positively associated with mPC3** (↑ LDL subclasses)
- **Co inversely associated with mPC3** (↑ LDL subclasses)
- **Zn inversely associated with mPC4** (↑ HDL subclasses)
- **Sb associated with mPC4** (↑ HDL subclasses)

Carriers of variants in redox genes with a role in diverse human disease pathways show differential metal-related metabolic changes

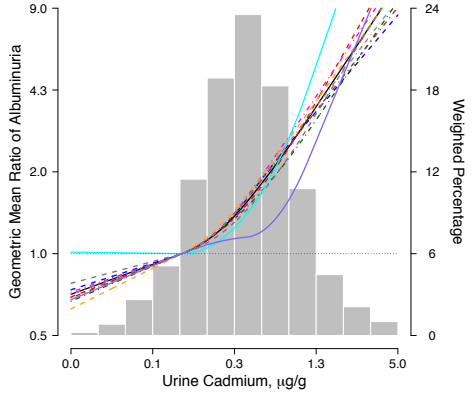


- Co and Zn showed the highest number of statistical interactions
- The most SNP-metal interacting genes were *NOX1*, *GSR*, *AGT*, *GCLC* and *REN*

Galvez-Fernandez, Redox Biology 2021.

Gene-environment interactions: Cadmium and Albuminuria

Objective: To test the hypothesis that carriers of specific genotypes are at increased cadmium potential effects. **Gene-by-environment interactions can point to key pathways and provide biological insight**



GMR of albuminuria levels comparing 80th to 20th percentile of cadmium distribution, by genotypes of SNPs with significant interactions at the Bonferroni level

SLC30A4	RAC1	N	GMR (95% CI)	P - int
rs3087816	rs4720672			<0.001
T/T + T/C (ref)	T/T (ref)	892	1.82 (1.65, 2.01)	
T/T + T/C (ref)	T/C + C/C	344	3.02 (1.85, 4.94)	
C/C	T/T (ref)	34	2.43 (2.07, 2.85)	
C/C	T/C + C/C	14	19.1 (8.02, 45.49)	

Source: Grau-Perez et al. Environment International 2017

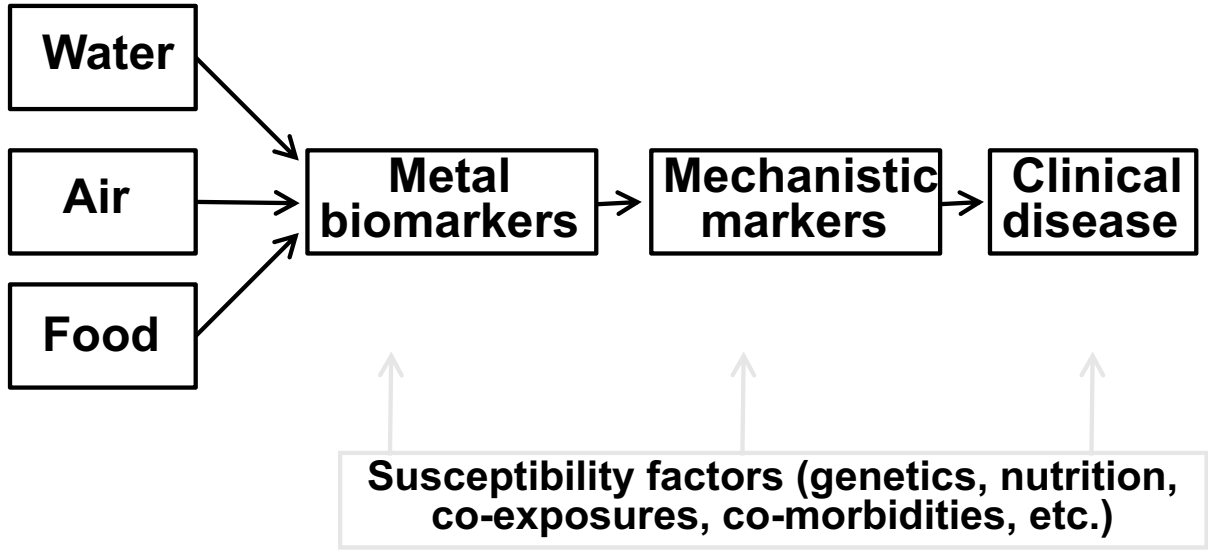
SLC30A4: Endosomal zinc transporter. No known role in albuminuria.

RAC1: Rho-family small GTP-ase, with a role in the maintenance of podocytes and proximal tubules integrity. Involved in ROS overproduction in endothelium and endothelial dysfunction

II. THE METAL GWAS INITIATIVE

MAIN OBJETIVE

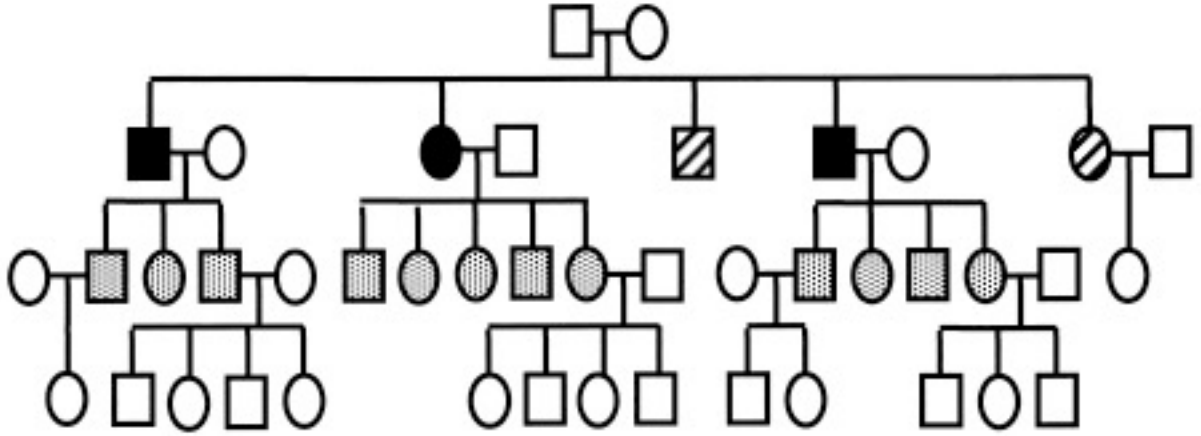
To use the largest possible set of high-quality cohort studies with data on metal and metalloid biomarkers to evaluate how levels of metal biomarkers, which are considered markers of internal dose (depending on uptake, metabolism and clearance), are influenced by genetic variation.



EARLIER GENETIC STUDIES OF METALS IN THE STRONG HEART FAMILY STUDY (SHFS)

- Goal: to map and identify genes that contribute to cardiovascular risk in American Indians

- Prototype family included in the SFHS



- SHFS participants had ~400 genome-wide genetic markers (microsatellites) genotyped

North et al. Am J Epidemiol 2003;157:303–314

Heritability

$$y_i = X_i\beta + \varepsilon_i \; ; \; \varepsilon_i \sim N(0, \Omega)$$

$$\Omega = 2\Phi\sigma_g^2 + I\sigma_e^2$$

$$h^2 = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_e^2}$$

r	relationship
0.5 (½)	parent-offspring
0.25 (¼)	grandparent-grandchild
0.125 (⅛)	great grandparent-great grandchild
1	identical twins
0.5 (½)	full siblings
0.25 (¼)	half siblings
0.125 (⅛)	first cousins

*Adjusted for age, age^2, age x sex, age^2 x sex, urine creatinine (log g/L), smoking status (never, former, current) and pack years categories (0, 0-10, 10-20, >20). Residual kurtosis was 0.79.

The number of pair relationships in the overall sample is distributed as follows: 66 parent-offspring; 847 siblings; 206 avuncular; 57 half siblings; 3 double 1st cousins; 3 grand avuncular; 26 half avuncular; 134 1st cousins; 6 half 1st cousins & half 2nd cousins; 8 1st cousins and 1 rem; 9 half 1st cousins; and 12 half 1st cousins and 1 rem and half 2nd cousins and 1 rem. Mean f is 2.158

Heritability of Urine Cadmium in the Strong Heart Study

	Heredabilidad, % (SE)	P-valor	% Varianza explicada por covariables
	Total (N=1936)		
Modelo 1	**0.30 (0.05)**	<0.001	23.8%
Modelo 2	**0.28 (0.05)**	<0.001	27.8%
	Arizona (N=200)		
Modelo 1	0.60 (0.20)	0.001	31.7%
Modelo 2	0.60 (0.20)	0.001	31.9%
	Oklahoma (N=822)		
Modelo 1	0.29 (0.07)	<0.001	22.2%
Modelo 2	0.25 (0.07)	<0.001	29.6%
	N. y S. Dakota (N=914)		
Modelo 1	0.29 (0.07)	<0.001	24.3%
Modelo 2	0.27 (0.07)	<0.001	25.2%

The 30 % of unexplained variability in urine cadmium levels could be attributed to additive genetic effects (p-value <0.001)*

Modelo 1 ajustado por edad, sexo, región, tabaquismo y componentes principales de estructura poblacional.
Modelo 2 ajustado además por zinc en orina.

Abreviaturas: SE: error estándar.

22/11/2022

QUANTITATIVE TRAIT LOCUS (QTL) LINKAGE SCAN: PRELIMINARY ANALYSIS (N = 341)

$$y_i = X_i\beta + \varepsilon_i \; ; \; \varepsilon_i \sim N(0, \Omega)$$

$$\Omega = \boxed{\Pi\sigma_q^2} + 2\Phi\sigma_g^2 + I\sigma_e^2$$

$$h_q^2 = \frac{\sigma_q^2}{\sigma_q^2+\sigma_g^2+\sigma_e^2}$$

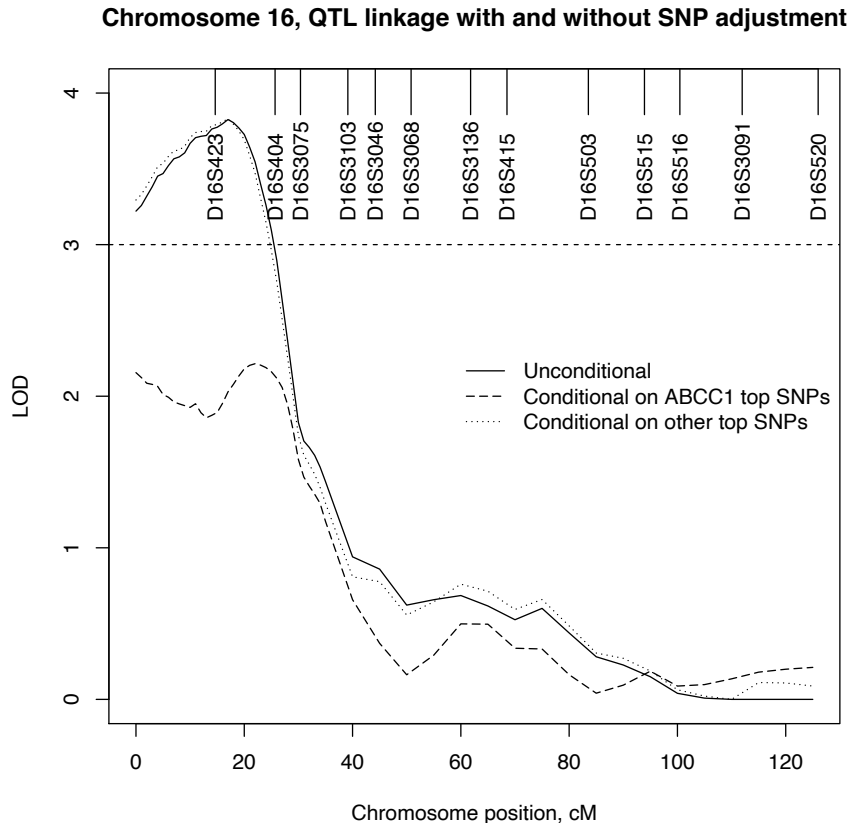
At each marker location (~400 markers):

LOD score = log10 likelihood ratio of the model with estimated σ_q versus the model with σ_q constrained to be 0

The genetic variant rs215106 in ABCC1 can explain the QTL signal

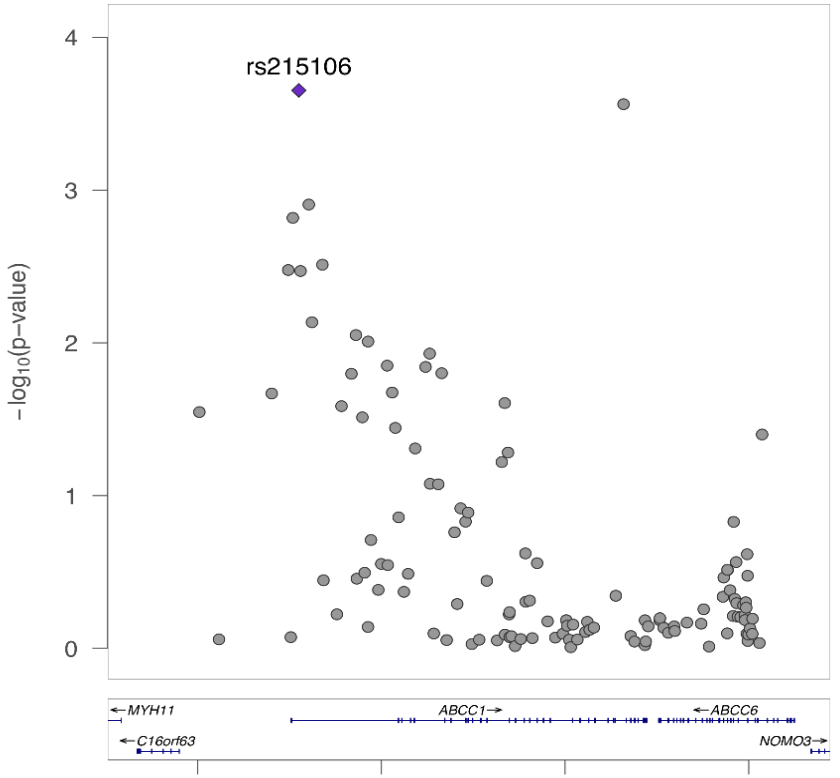
b. QTL linkage scan + conditional models

Cromosoma 16 (LOD score = 3.83)



c. Regional association plot

Gen *ABCC1* en cromosoma 16. Top SNP rs215106



Dotted lines represent LOD scores for loci with causal genes influencing urine cadmium concentrations in residuals extracted from models adjusted for age, age^2, interaction age x sex, interaction age^2 x sex, urine creatinine (log g/L), smoking status (never, former, current) and pack-years categories (0, 0-10, 10-20, >20) and body mass index. Continues lines represent LOD scores for loci with causal genes influencing urine cadmium concentrations, in residuals extracted from models further adjustment for study region.

Heritability and Preliminary Genome-Wide Linkage Analysis of Arsenic Metabolites in Urine

Maria Tellez-Plaza,[1,2,3,4] Matthew O. Gribble,[1,2] V. Saroja Voruganti,[5] Kevin A. Francesconi,[6] Walter Goessler,[6] Jason G. Umans,[7,8] Ellen K. Silbergeld,[2] Eliseo Guallar,[1,3,9,10] Nora Franceschini,[11] Kari E. North,[11] Wen H. Kao,[1,9] Jean W. MacCluer,[5] Shelley A. Cole,[5] and Ana Navas-Acien[1,2,9,10]

[1]Department of Epidemiology, and [2]Department of Environmental Health Sciences, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA; [3]Area of Epidemiology and Population Genetics, National Center for Cardiovascular Research (CNIC), Madrid, Spain; [4]Fundacion de Investigacion del Hospital Clinico de Valencia-INCLIVA, Valencia, Spain; [5]Department of Genetics, Texas Biomedical Research Institute, San Antonio, Texas, USA; [6]Institute of Chemistry–Analytical Chemistry, Karl-Franzens University, Graz, Austria; [7]MedStar Health Research Institute, Hyattsville, Maryland, USA; [8]Georgetown–Howard Universities Center for Clinical and Translational Science, Washington, DC, USA; [9]Welch Center for Prevention, Epidemiology and Clinical Research, and [10]Department of Medicine, Johns Hopkins Medical Institutions, Baltimore, Maryland, USA; [11]Department of Epidemiology, School of Public Health, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

Background: Arsenic (III) methyltransferase *(AS3MT)* has been related to urine arsenic metabolites in association studies. Other genes might also play roles in arsenic metabolism and excretion.

Objective: We evaluated genetic determinants of urine arsenic metabolites in American Indian adults from the Strong Heart Study (SHS).

Methods: We evaluated heritability of urine arsenic metabolites [percent inorganic arsenic (%iAs), percent monomethylarsonate (%MMA), and percent dimethylarsinate (%DMA)] in 2,907 SHS participants with urine arsenic measurements and at least one relative within the cohort. We conducted a preliminary linkage analysis in a subset of 487 participants with available genotypes on approximately 400 short tandem repeat markers using a general pedigree variance component approach for localizing quantitative trait loci (QTL).

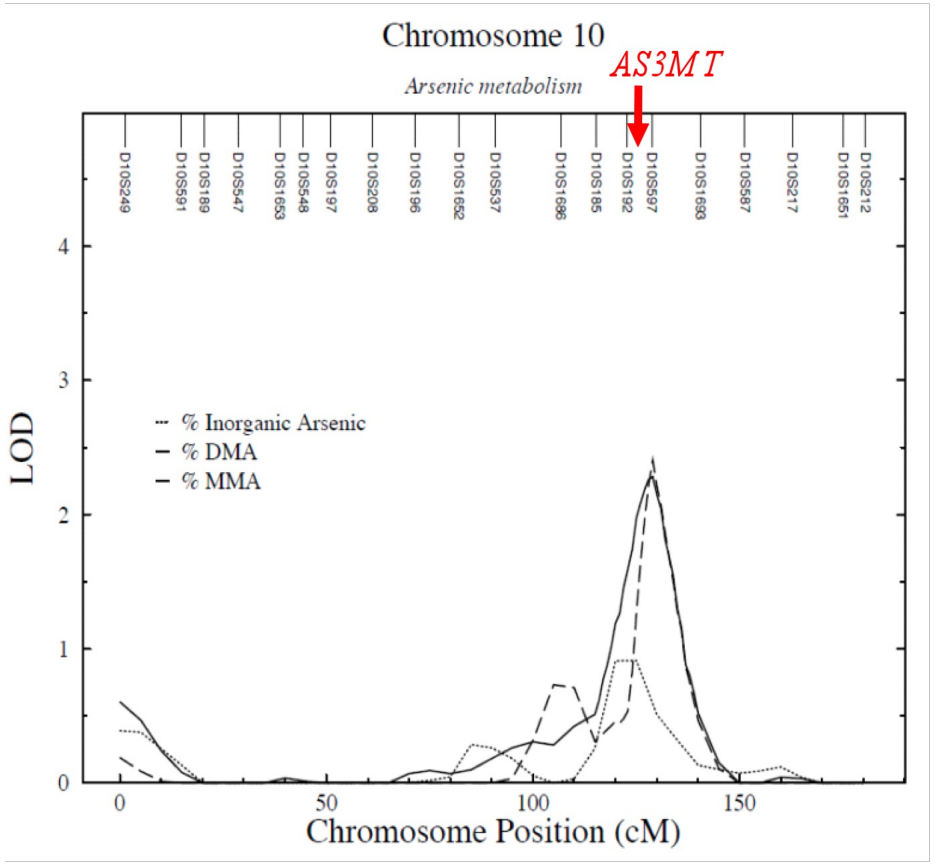
Results: The medians (interquartile ranges) for %iAs, %MMA, and %DMA were 7.7% (5.4–10.7%), 13.6% (10.5–17.1%), and 78.4% (72.5–83.1%), respectively. The estimated heritability was 53% for %iAs, 50% for %MMA, and 59% for %DMA. After adjustment for sex, age, smoking, body mass index, alcohol consumption, region, and total urine arsenic concentrations, LOD [logarithm (to the base of 10) of the odds] scores indicated suggestive evidence for genetic linkage with QTLs influencing urine arsenic metabolites on chromosomes 5 (LOD = 2.03 for %iAs), 9 (LOD = 2.05 for %iAs and 2.10 for %MMA), and 11 (LOD = 1.94 for %iAs). A peak for %DMA on chromosome 10 within 2 Mb of *AS3MT* had an LOD of 1.80.

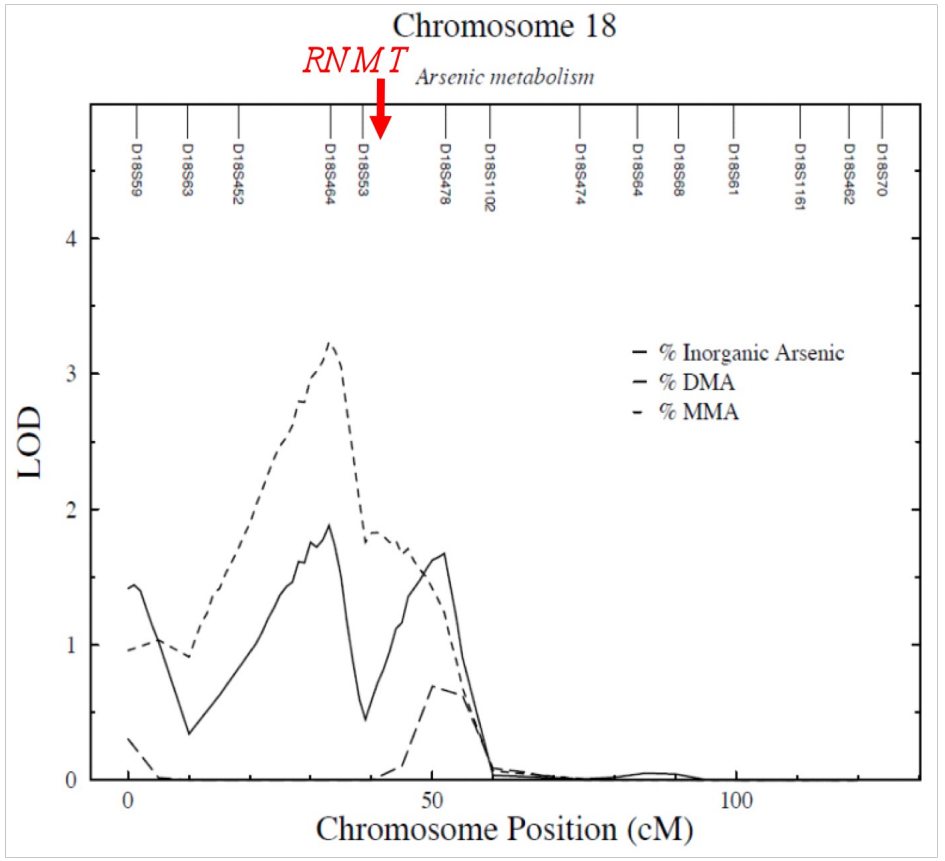
Conclusions: This population-based family study in American Indian communities supports a genetic contribution to variation in the distribution of arsenic metabolites in urine and, potentially, the involvement of genes other than *AS3MT*.

Key words: American Indians, arsenic metabolism, arsenic species, determinants, heritability, linkage scan, Strong Heart Study. *Environ Health Perspect* 121:345–351 (2013). http://dx.doi.org/10.1289/ehp.1205305 [Online 15 January 2013]

human tissues, although the genes involved remain largely unknown. Genome-wide genetic approaches may contribute to the discovery of genes related to variation in urine arsenic metabolites. Moreover, while arsenic metabolism shows evidence for familial aggregation (Chung et al. 2002), the heritability of urine arsenic methylation patterns has not been evaluated.

The Strong Heart Study (SHS) is a population-based prospective cohort study funded by the National Heart, Lung, and Blood Institute (NHLBI) to evaluate cardiovascular disease and its risk factors, including genetic and environmental determinants, in 13 U.S. American Indian communities from Arizona, Oklahoma, and North and South Dakota (Lee et al. 1990). Some of these communities are known to be exposed to arsenic in drinking water (Navas-Acien et al. 2009). In this study, we first evaluated the heritability of urine arsenic methylation patterns in SHS participants who had at least one relative within the cohort. In a subset of the popula-





Red arrows indicate approximate locations for *AS3MT* and *RNMT* based on genes positions relative to the microsatellite markers.

RATIONALE FOR THE METAL GWAS INITIATIVE

- Genome-wide association studies (GWAS) have evolved into a powerful tool for investigating the genetic architecture of human disease, especially for common complex disease
- Previous studies which have implemented this methodology to study the association of metal exposures to single-nucleotide polymorphisms (SNPs) have been modest in sample size
- Biological mechanisms underlying metal toxicokinetic and gene-environment effects of metal exposures are not yet well understood

SPECIFIC AIMS:

- **Specific Aim 1:** investigate the associations between genome-wide SNPs and metal biomarkers using several methods:

  - a) **Candidate-gene approach** by extracting SNPs annotated to metal and metalloid transporters, and
  - b) **Genome-wide exploratory approach**

  By using **traditional** single-SNP association and **recently developed** methods to simultaneously model highly dimensional SNPs matrices to identify most relevant SNPs in a way less affected by the multiple comparisons

SPECIFIC AIMS:

- **Specific Aim 2:** Investigate *In silico* potential **biological implications** of the genetic variants associated with metal biomarker levels

By an **integrative bioinformatic analysis** to characterize interconnected physio-pathological pathways from well-established bioinformatics datasets including IntAct, KEGG, Wikipaths and others.

SPECIFIC AIMS:

- **Specific Aim 3:** Create a repository of estimated coefficients and standard errors from the findings of the metal-GWAS Initiative for future Mendelian Randomization studies of metals and health outcomes.

III. WHAT WE HAVE DONE SO FAR

PROJECT 1: MENDELIAN RANDOMIZATION STUDY OF **SELENIUM AND DIABETES-RELATED ENDPOINTS**

Funding for:

- Selenoprotein speciation (University of Huelva)
- **Genotyping** (Illumina GSA microarrays) **in AWHS and Hortega Studies** (National Genotyping Core).
- **1 Predoctoral Student** (Zulema Rodriguez-Hernandez with a background in **Bioinformatics**)

Goal: To examine the causal role of exposure to Selenium (Se) and Se metabolism on diabetes and related conditions (i.e. insulin resistance, kidney disease), and to explore the associated metabolic mechanisms, with a clear focus on informing precision medicine.

Dr. Pastor-Barriuso

Dr. Tellez-Plaza

Se

SelenOmics

Spanish Agency for Research
PID2019-108973RB-C21

Centro Nacional de Epidemiología

PROJECT 11: THE CAUSAL ROLE OF **METAL EXPOSURES** ON SUBCLINICAL **ATHEROSCLEROSIS PROGRESSION**

Approved dbGAP accession (#26157): data from **11 cardiovascular cohorts** funded by the US NIH for the **SNP-Atherosclerosis stage** of mendelian randomization analysis

Funding for:

- Measuring **metals in AWHS from an additional visit** (University of Huelva)
- Subcontracting processing raw **subclinical atherosclerosis measures from repeated study visits**, including advanced imaging
- **1 Research Assistant** (Biostatistics/Epidemiology background to be hired starting in January 2023)

Goal: To examine the causal role of exposure to low doses of metals and metal trajectories on atherosclerosis changes and exploring associated metabolic mechanisms

Dr. Tellez-Plaza

MetalOmic Heart

Spanish Funds for Health Studies, PI22/00029

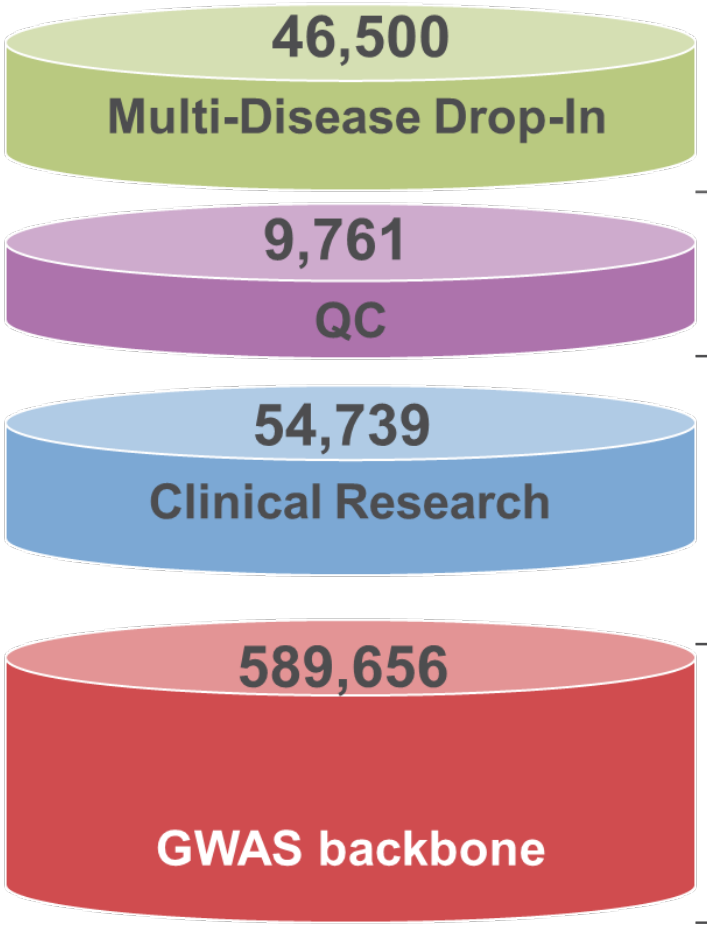
Centro Nacional de Epidemiología

Genotyping of the AWHS and Hortega Studies using the Illumina Global Screening Array (GSA) at National Genotyping Core (CNIO)

ARAGON WORKERS' HEALTH STUDY

GSA Manifest variant count

Predictive, Clinical Research, and QC



- User defined custom content OR
- Pre-designed multi-disease drop in

- Sample tracking and stratification

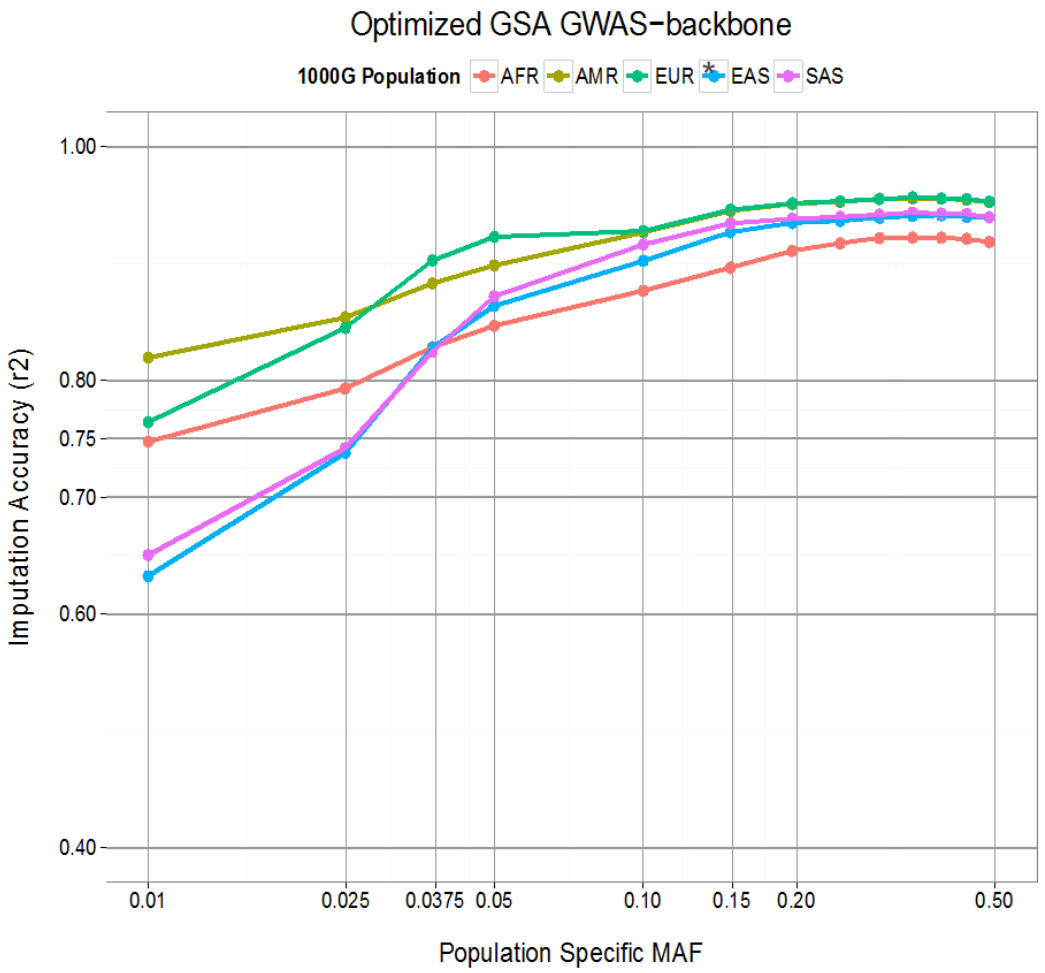
- Up-to-date known clinical associations
- Pharmacogenomics
- Well-curated exome content
- NHGRI-GWAS and HLA content

- Cross-Population and population-specific
- Enriched for low-frequency variants (1-5%)
- High imputation accuracy for ALL populations

FINAL Product : 700, 656

HOSPITAL UNIVERSITARIO RÍO HORTEGA

Dr. Gonzalez-Neira

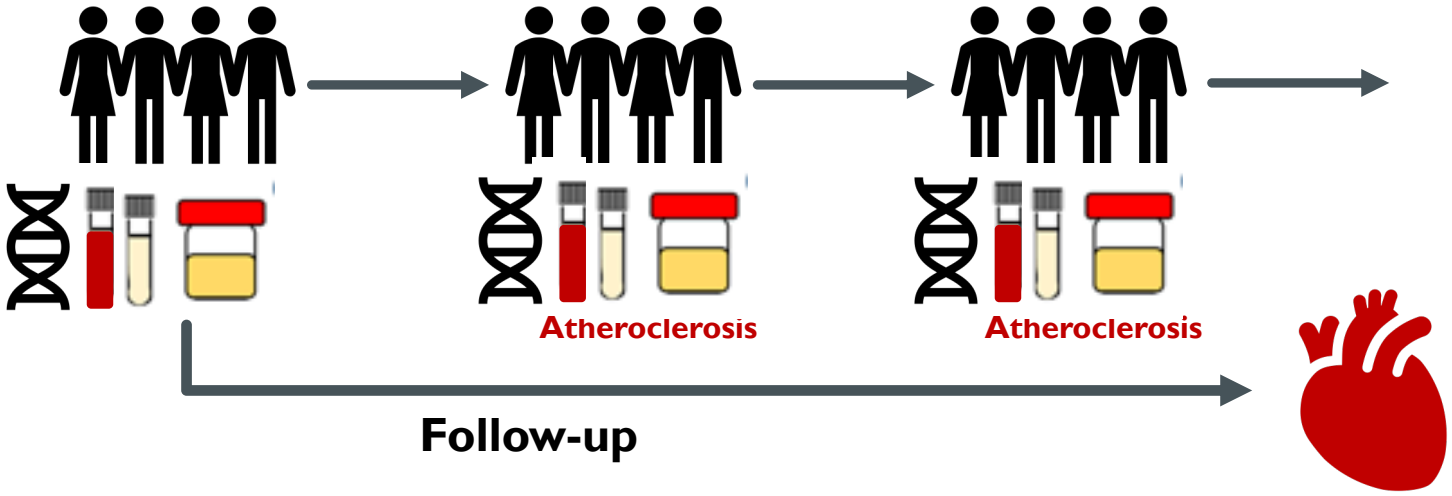


Mean Imputation Accuracy

>1 % Minor allele frequency

Population	GSA
AFR	0.95
AMR	0.97
EUR	0.97
EAS	0.96
SAS	0.96

COHORTS WITH AVAILABLE (OR ONGOING) METAL BIOMARKERS DETERMINATION AND SNPS (N~6000)



AWHS ARAGON WORKERS' HEALTH STUDY

ortega

Gac Sanit. 2015;29(4):308–315

Special article

Population-based multicase-control study in common tumors in Spain (MCC-Spain): rationale and study design

Gemma Castaño-Vinyals[a,b,c,q,*,u], Nuria Aragonés[d,q,r,u], Beatriz Pérez-Gómez[d,q,r], Vicente Martín[e,q], Javier Llorca[f,q], Victor Moreno[g,h,q], Jone M. Altzibar[i,q], Eva Ardanaz[j,q], Sílvia de Sanjosé[g,q], José Juan Jiménez-Moleón[k,q], Adonina Tardón[l,q], Juan Alguacil[m,q], Rosana Peiró[n,q], Rafael Marcos-Gragera[o,q], Carmen Navarro[p,q,s], Marina Pollán[d,q,r,u], Manolis Kogevinas[a,b,c,q,t,u], MCC-Spain Study Group

ICO Institut Català d'Oncologia

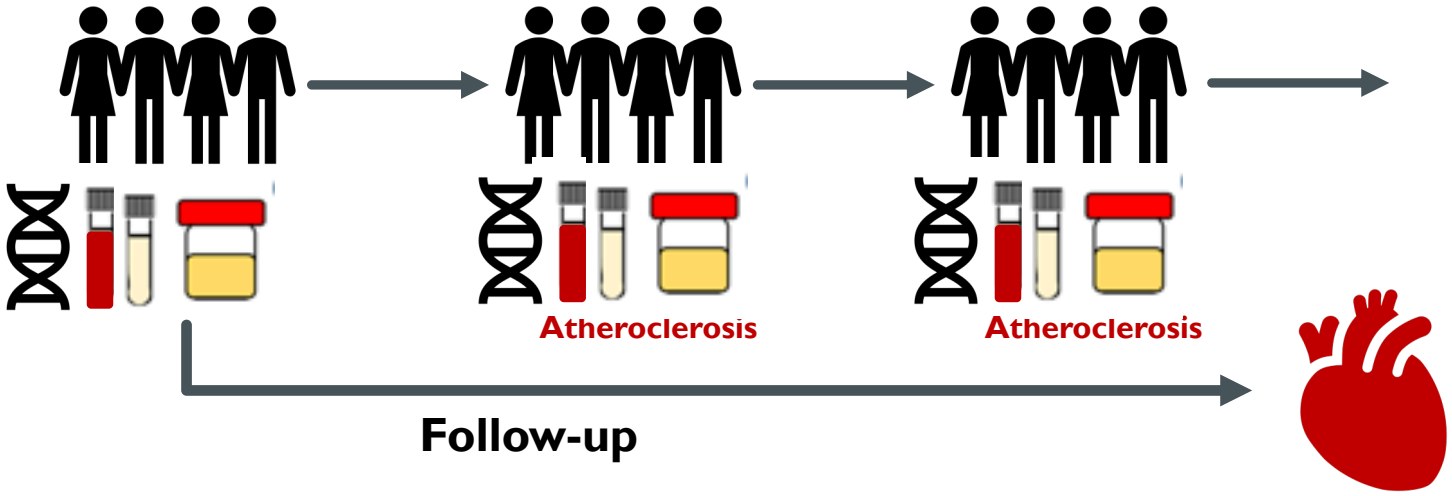
MINISTERIO DE CIENCIA E INNOVACIÓN

Instituto de Salud Carlos III

MULTI-CASO CONTROL - SPAIN

COHORTS WITH AVAILABLE (OR ONGOING) METAL BIOMARKERS DETERMINATION AND SNPS (N~10000)

Centro Nacional de Epidemiología



MULTI-CASO CONTROL - SPAIN

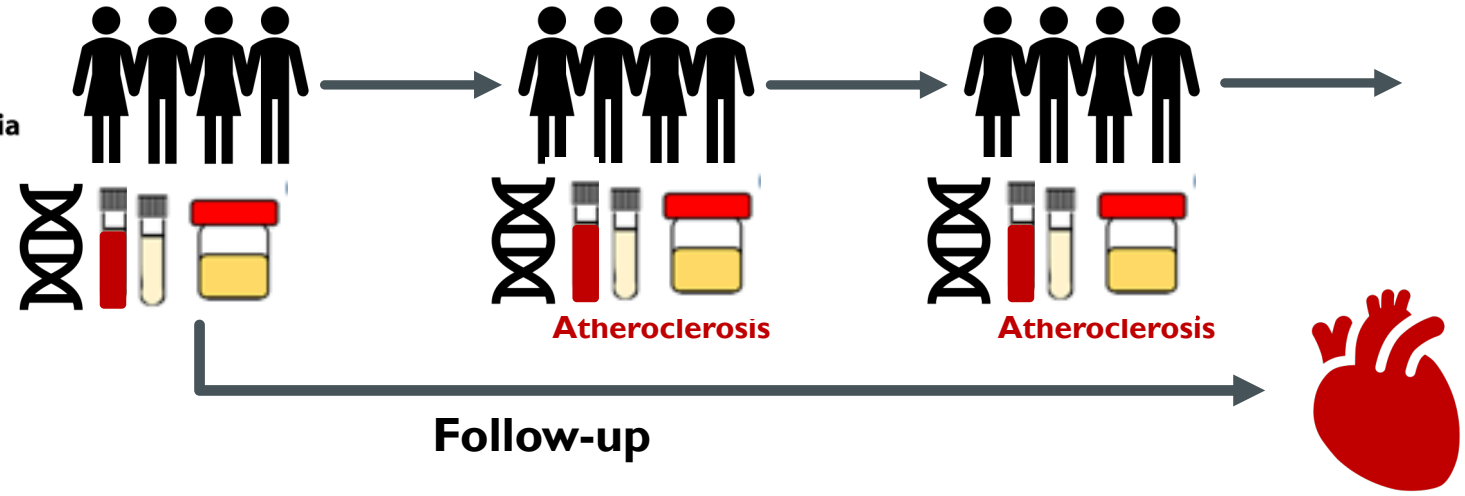
ARAGON WORKERS' HEALTH STUDY

ortega

COHORTS WITH AVAILABLE (OR ONGOING) METAL BIOMARKERS DETERMINATION AND SNPS (N~18000)

Centro Nacional de Epidemiología

ICO
Institut Català d'Oncologia



COLUMBIA UNIVERSITY

Dr. Navas-Acien

MULTI-CASO CONTROL - SPAIN

AWHS ARAGON WORKERS' HEALTH STUDY

ortega

STRONG HEART STUDY

MESA

RESOURCES IN PLACE AT THE INSTITUTO DE SALUD CARLOS III

HIGH PERFORMANCE COMPUTING SERVER- XTUTATIS

Número de nodos	34
Cores/CPUs totales	768
Memoria total	12,6 Tb
GPGPU	4 x NVIDIA Tesla P 100
Conexión almacenamiento	Ethernet de 10 Gbps
Almacenamiento principal	100 Tb
Almacenamiento interno nodo	800 Gb
Sistema operativo	Distribución Linux CentOs-8

Centro Nacional de Epidemiología

¡GRACIAS A BRUNO Y OMAR DEL HPC en el ISCIII !

PIPELINE

Se SelenOmics

MetalOmic Heart

Datos Genéticos crudos

PLINK

QC

Non valid SNPs excluded:
- Missings SNPs > 5%
- SNPs ∉ chr 1:22, X
- SNPs ≠ A, T, C o G alelos
- SNPs multiple mapping

Non valid samples excluded:
- Heterozigosidad (mean + 4sd)
- Missings simples > 10%
- Sex discordance
- Duplicates
- Cryptic relations

Transformar a versión 38

perl

Preparar ficheros para imputación

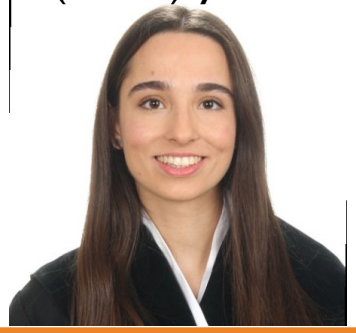
vcftools

bcftools

Imputación

TOPMED

- Zulema Rodriguez-Hernandez, PhD student in Biotecnology and strong backgroun in epidemiology and biostats
- Inestimable ayuda: Pablo Fernandez (ISCIII) y Ana Villanueva (ICO)



Centro Nacional de Epidemiología

+ 1 more person with a Biostats/Epi/Bioinfo profile to be hired in the coming months

Bibliographic search for candidate genes

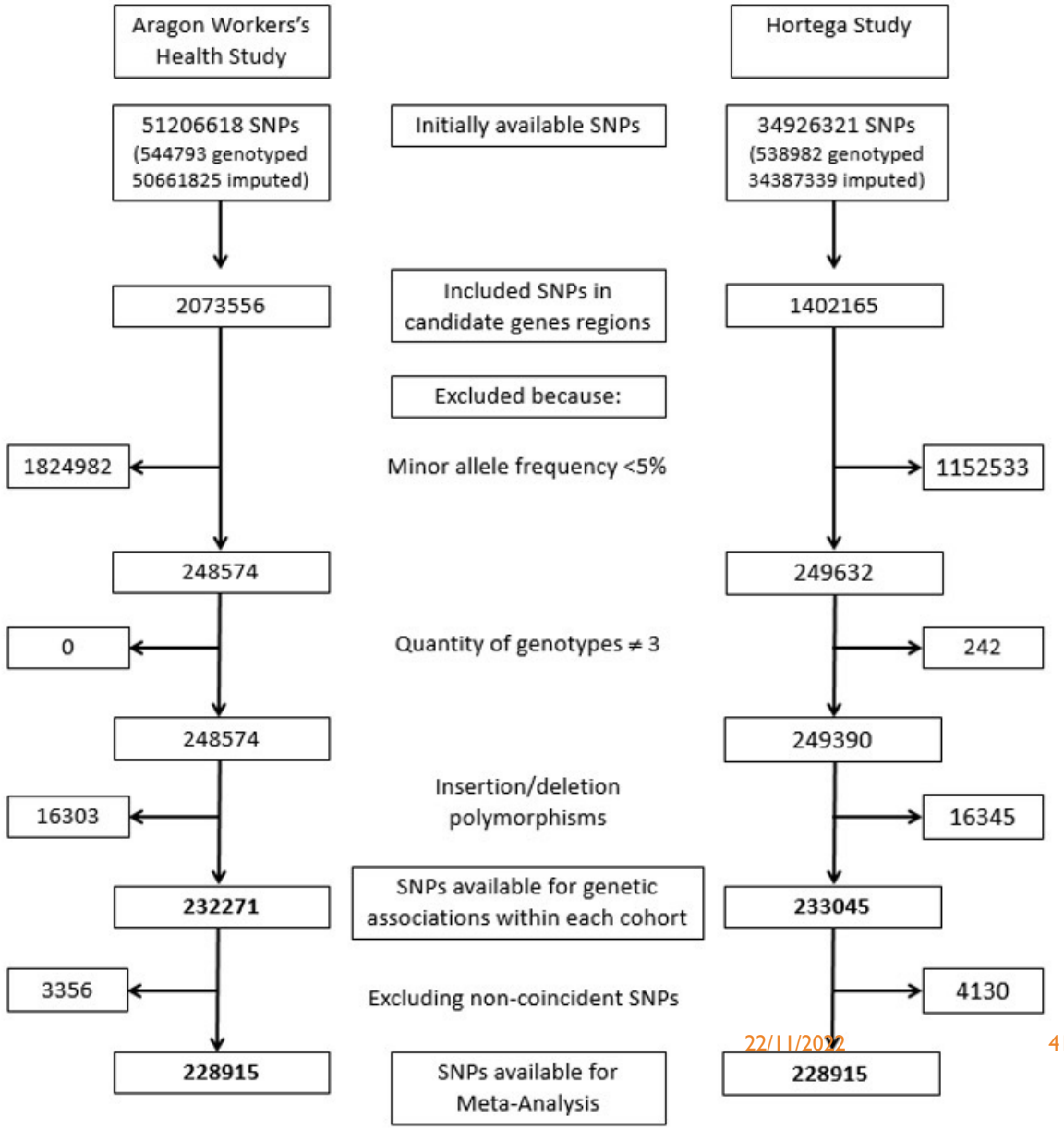
"Candidate-gene approach" + meta-analysis

1541 candidate genes

- 1469 transportes
- 64 metallothioneins
- 8 hand search

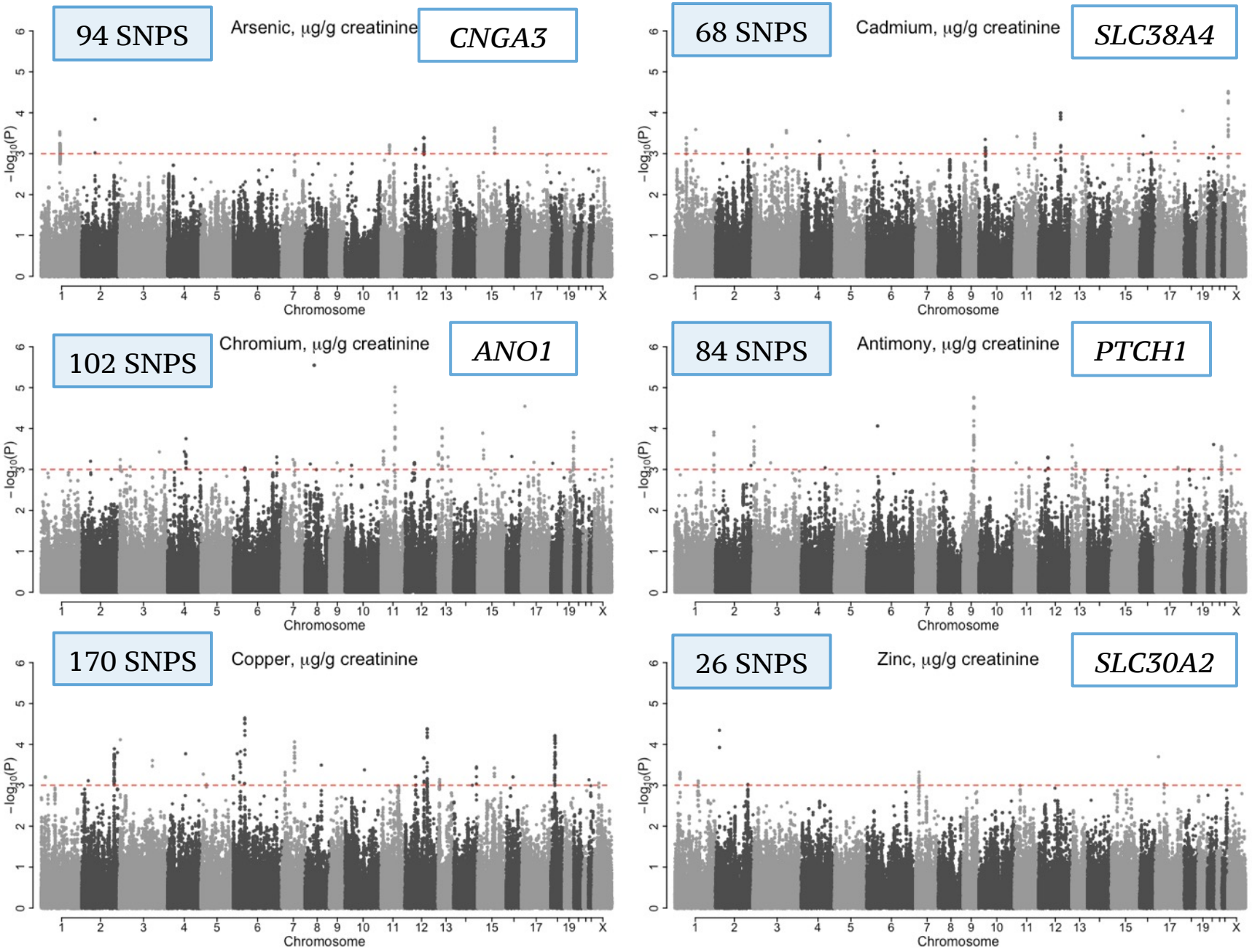
Meta-analysis

- 1867 AWHS participants
- 1401 Hortega Study participants



Association of candidate genetic variants and metals

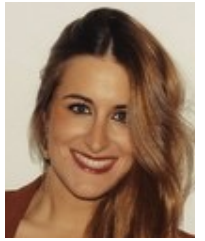
Statistically suggestive SNPs at a p-value <0.001



VARIABLE SELECTION IN THE SETTING OF OMICS DATA: EXTENDED "SIS" PACKAGE

Centro Nacional de Epidemiología

COLUMBIA UNIVERSITY

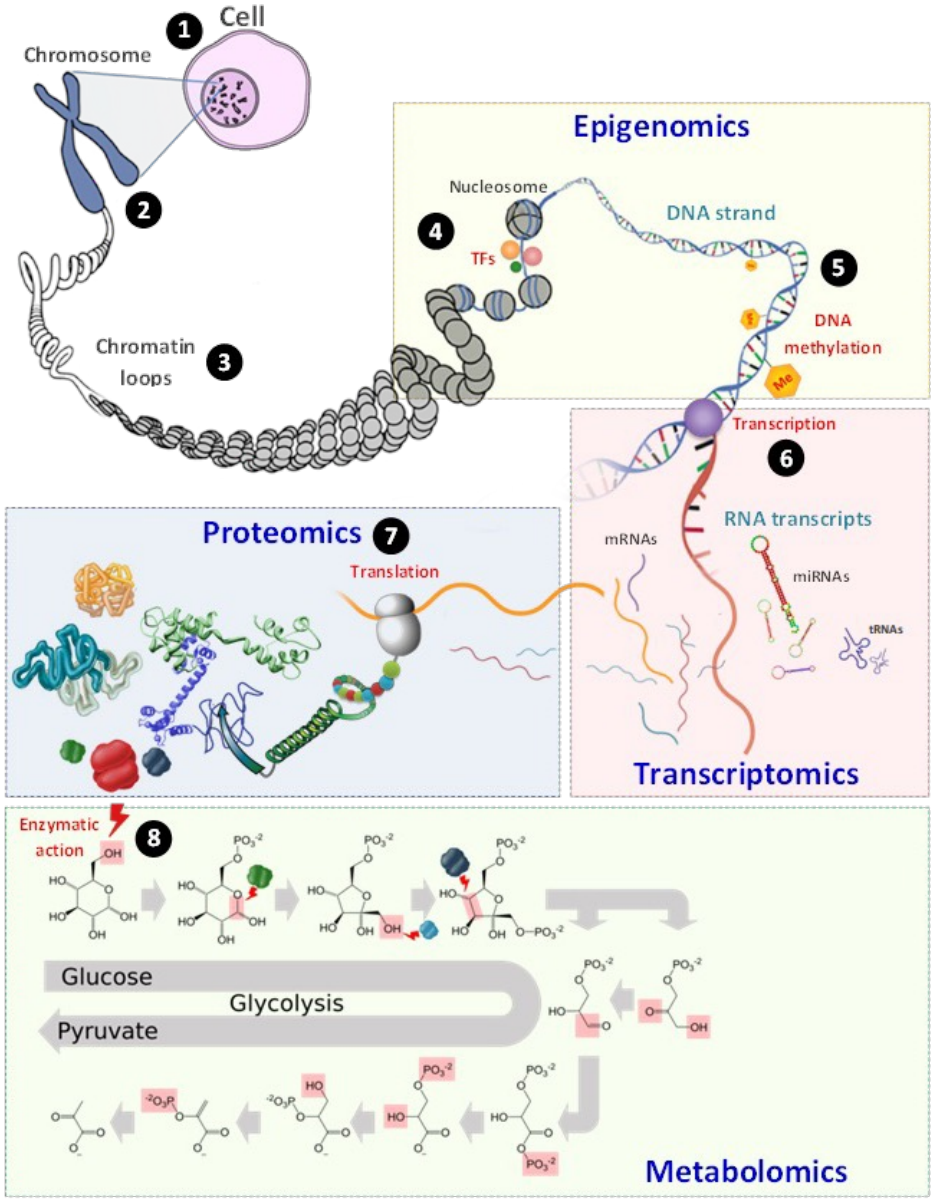

Arce Domingo


Dr. Yang Feng

- **Extended version of the R package *SIS* (publicly available) :**
  - **Iterative Sure Screening (ISIS)**: seleccionar las variables relevantes de un grupo de variables correlacionadas
  - **Adaptive Elastic-Net (AENET).** estimación de parámetros menos sesgada, que puede favorecer el descubrimiento biológico, así como una mejor precision predictiva

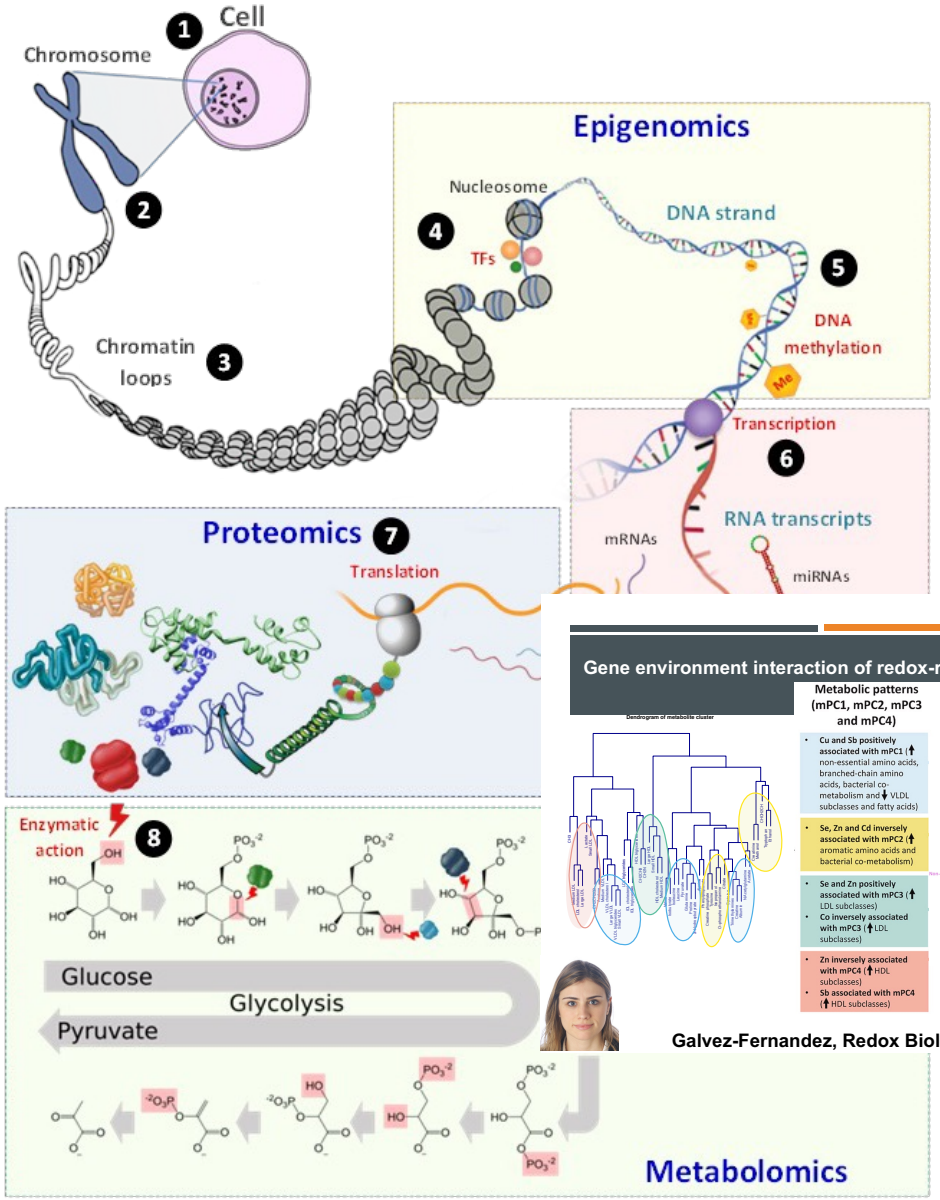
COMPARISON OF METHODS FOR VARIABLE SELECTION HIGH DIMENSIONAL SETTINGS

	Aenet	LASSO	Elastic-net (α=0.05)	MSAenet	SCAD	MCP	Regresión linear
N variables selected	222	221	222	113	93	42	-
Medidas de rendimiento							
Error c. medio entrenamiento	22.89	14.50	15.73	24.25	14.71	20.18	15.97
Error c. medio validación	**31.78**	**38.29**	**35.38**	**42.0**	**41.65**	**33.77**	**35.84**
Time de computación (días)	9.9	9.8	8.7	7.3	1.7	0.9	Immediate

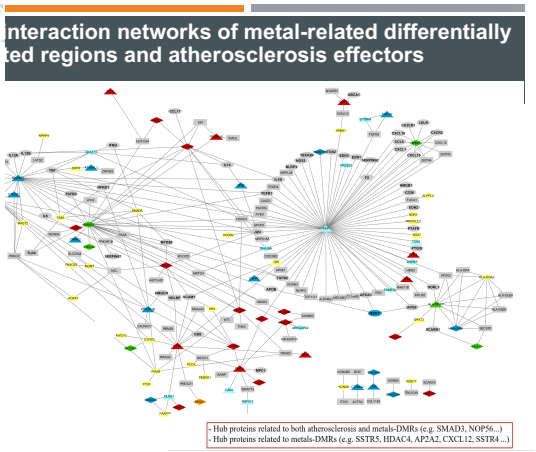


INTEGRATIVE BIOINFORMATIC ANALYSIS OF RESULTS WITH AVAILABLE BIOLOGICAL EVIDENCE FROM PUBLIC DATABASES

INTEGRATIVE BIOINFORMATIC ANALYSIS OF RESULTS WITH AVAILABLE BIOLOGICAL EVIDENCE FROM PUBLIC DATABASES



Galvez-Fernandez, Redox Biology 2021.

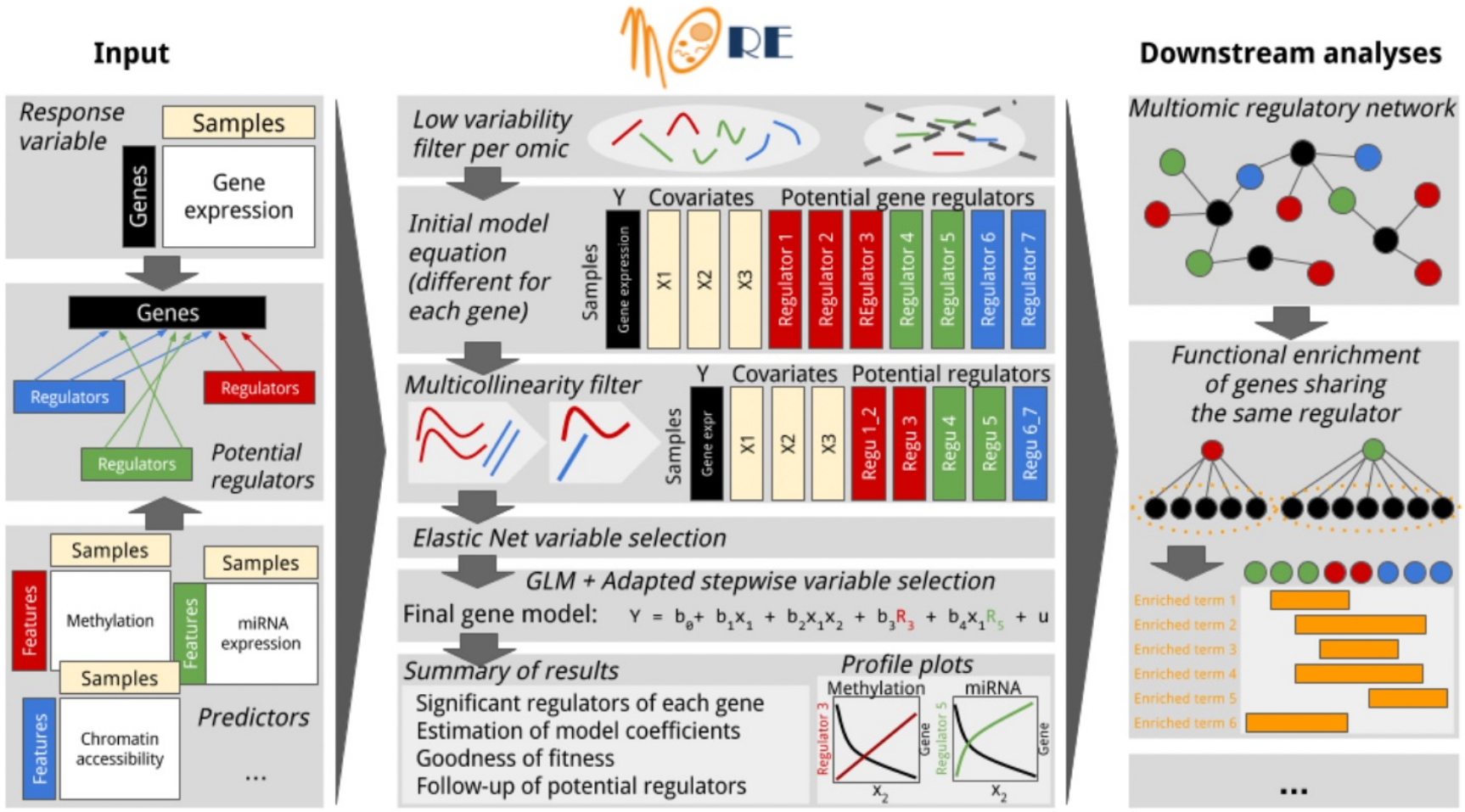


Source: Riffo-Campos et al. Philosophical Transactions B. 2017

https://github.com/ConesaLab/MORE

Genomics of Gene Expression Lab

Centro Nacional de Epidemiología



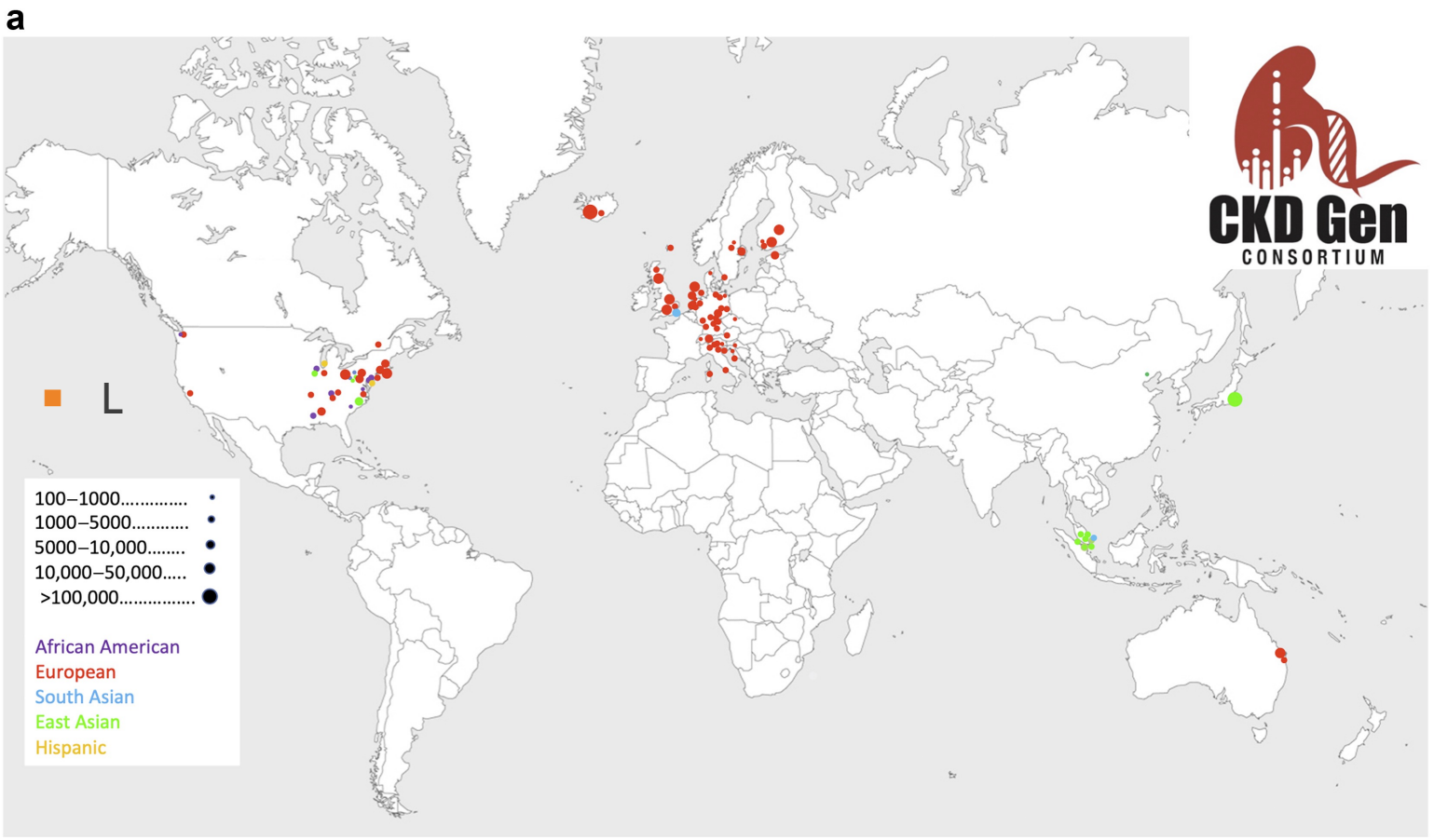
UNIVERSITAT POLITÈCNICA DE VALÈNCIA

"more" package extension to fully accommodate GWAS data

Sonia Tarazona Campos

IV. PROSPECT FOR FUTURE COLLABORATIONS IN THE SETTING OF THE METAL GWAS INITIATIVE

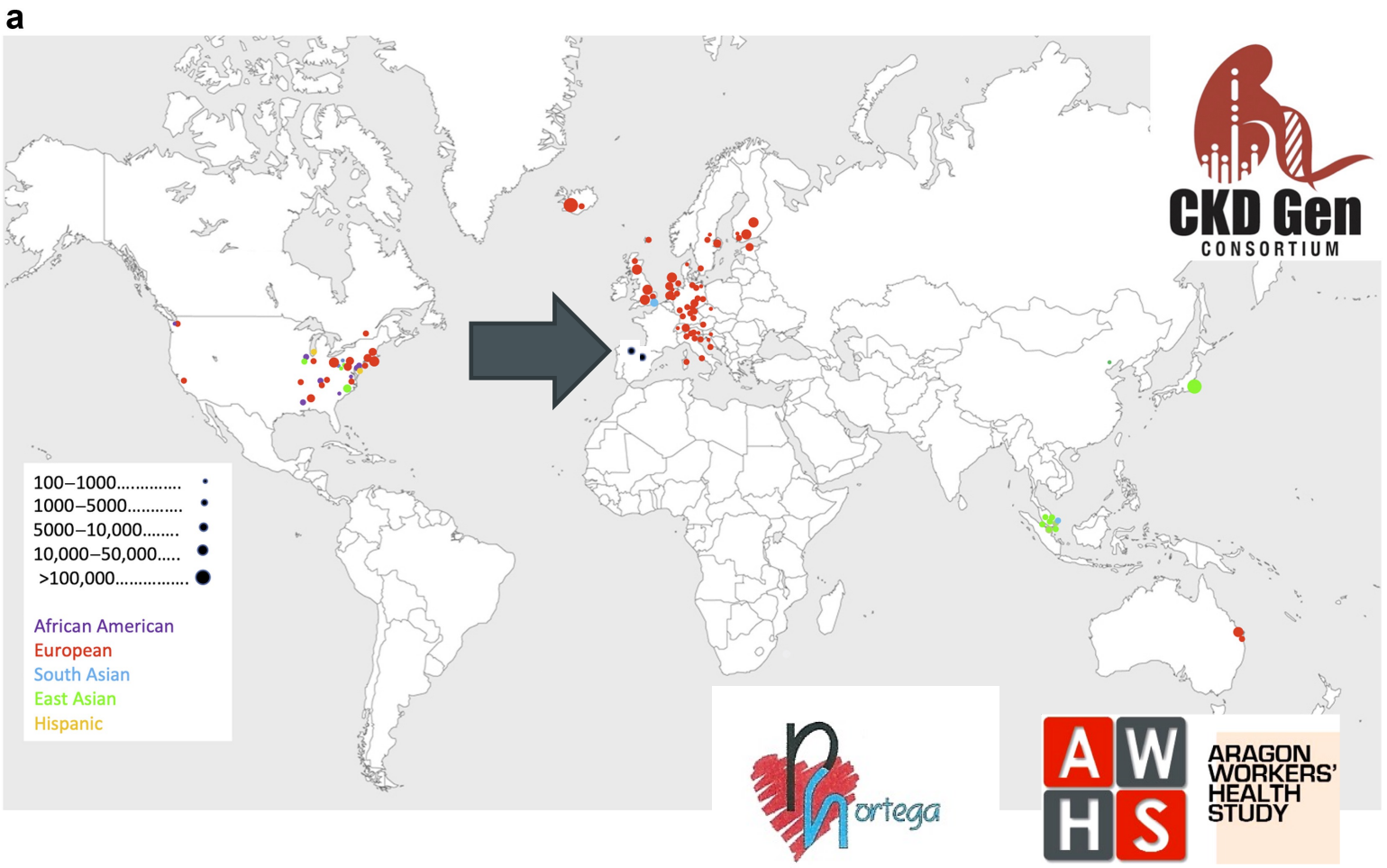
THE EXAMPLE OF THE CKDGEN CONSORTIUM: THE HORTEGA AND AWHS STUDIES CONTRIBUTING COHORTS FOR ROUND 5



CKD Gen CONSORTIUM

Leaded by Drs. Anna Kottgen and Christian Pattaro

THE EXAMPLE OF THE CKDGEN CONSORTIUM: THE HORTEGA AND AWHS STUDIES CONTRIBUTING COHORTS FOR ROUND 5



Leaded by Drs. Anna Kottgen and Christian Pattaro

THE EXAMPLE OF THE CKDGEN CONSORTIUM: THE HORTEGA AND AWHS STUDIES CONTRIBUTING COHORTS FOR ROUND 5

The analysis is a 3-step process:

- first, we are asked to generate the phenotypes in an specific format and to send the **resulting diagnostic files** (not raw data);
- after they confirm that the phenotypes are OK, we are asked to run the GWAS and upload the **results** (not raw data).
- The CKDGen analysists meta-analyse the results from all the contributing cohorts

Software for download that covers phenotype preparation and generation, genotype preparation, preparation of submission files to run GWAS, formatting scripts, as well as log collection and diagnostic plots, is provided.

We do not have to re-invent the wheel in the Metal-GWAS Iniative !!

Z. Rodriguez-Hernandez

ortega

AWHS ARAGON WORKERS' HEALTH STUDY

CONCLUDING REMARKS

The metal GWAS Initiative can enable:

1. The robust evaluation of genetic determinants of metal biomarkers

2. Point to interesting biological pathways for potential metal-related health effects, through integrative post-GWAS analysis

3. Create the necessary structure to enable the participating cohorts to entertain large collaborative efforts to assess, in the most possible robust and reproducible way, potential metal-related health effects and associated gene-environment interactions.

THANK YOU!