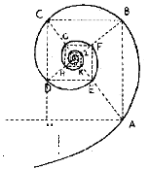
Università degli Studi di Milano

Scuola di Dottorato in Medicina Molecolare



TESI DI DOTTORATO

Ciclo XXVIII

Anno Accademico 2014/2015

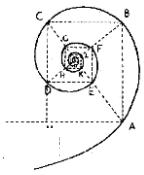
Dottorando: Martina CHITTANI

A GENOME WIDE APPROACH TO IDENTIFY GENETIC VARIANTS ASSOCIATED WITH LEFT VENTRICULAR MASS

Tutore: Dott.ssa Cristina BARLASSINA

Direttore del Dottorato: Ch.mo Prof. Mario Clerici

UNIVERSITÀ DEGLI STUDI DI MILANO



SCUOLA DI DOTTORATO IN MEDICINA MOLECOLARE

CICLO XXVIII

Anno Accademico 2014/2015

TESI DI DOTTORATO DI RICERCA

MED/04

A GENOME WIDE APPROACH TO IDENTIFY GENETIC VARIANTS ASSOCIATED WITH LEFT VENTRICULAR MASS

Dottorando: Martina CHITTANI

Matricola N° R10145

TUTORE: Dott.ssa Cristina BARLASSINA

DIRETTORE DEL DOTTORATO: Ch.mo Prof. Mario CLERICI

SOMMARIO

La massa ventricolare sinistra (MVS) è un importante fenotipo clinico, la cui valutazione è in grado di predire eventi cardiovascolari avversi e morte prematura indipendentemente da sesso, provenienza etnica ed età. Un aumento della MVS si definisce, al di sopra di una determinata soglia, come ipertrofia ventricolare sinistra (IVS), ed è caratterizzata dall'ispessimento del ventricolo sinistro del cuore. In studi community-based, la presenza di IVS e l'aumento di MVS si sono dimostrate in grado di predire lo sviluppo di malattia coronarica, insufficienza cardiaca congestizia, ictus e malattie cardiovascolari. Questo fenotipo, quindi, non viene considerato solo come indicatore della struttura cardiaca, ma anche come fenotipo intermedio per stabilire l'esito clinico di varie malattie cardiovascolari. Diversi studi hanno indicato che la MVS è influenzata da fattori genetici. Allo scopo di identificare i geni che influenzano l'MVS, sono stati effettuati studi di linkage e di associazione genome-wide in diverse popolazioni, ma restano ancora da definire gran parte dell'ereditabilità e l'identità dei pathway e delle varianti funzionali; la promessa di predizione del rischio su base genetica non è quindi ancora realizzabile.

Lo scopo dello studio è stato quello di indagare l'associazione tra varianti genetiche comuni e la massa ventricolare sinistra, mediante un approccio genome-wide in una coorte di 966 soggetti non trattati con ipertensione lieve-moderata. Dall'analisi lineare, sono stati selezionati 85 polimorfismi a singolo nucleotide (SNP), con un p-value suggestivo di associazione ($\leq 10^{-5}$). In particolare, alcuni SNP si trovano in geni con un ruolo nella patogenesi dell'ipertrofia cardiaca già riportato in letteratura, come ROCK1, IGF1, CACNA1D, FGFR1, TRAF5, SOX5, e KSR2. Ciascuno di essi potrebbero giocare un ruolo nel determinare il fenotipo, nonché far parte di pathway direttamente o indirettamente correlati alla fisiopatologia cardiaca.

Per identificare gli alleli di rischio associati ai risultati più interessanti in relazione al fenotipo studiato, è stata effettuata un'analisi caso-controllo dividendo il nostro campione in due sottogruppi in base ai valori di MVS. La maggior parte degli SNP associati alla MVS nella regressione lineare presentano un'associazione significativa, dunque i portatori degli alleli di rischio hanno un odds ratio > 1 di avere una MVS aumentata, vale a dire di essere casi, rispetto ai controlli. Tuttavia, come per la maggior parte dei tratti complessi, gli odds ratio osservati sono modesti, quindi la loro rilevanza dal punto di vista clinico è ridotta. Abbiamo dunque definito uno score di rischio genetico aggregato e ponderato (wGRS), utilizzando l'effetto dell'allele di rischio (beta dell'analisi di regressione lineare) per spiegare la forza dell'associazione genetica di ogni allele. La possibilità di combinare più varianti in uno score di rischio genetico potrebbe essere interessante e aggiungere rilevanza ai risultati ottenuti.

In conclusione, il nostro studio ci ha permesso di individuare dei geni il cui ruolo nella funzione cardiaca e/o ipertrofia cardiaca è stata dimostrata in precedenza da pubblicazioni di diversi autori. Inoltre, abbiamo evidenziato l'utilità di una misura aggregata di rischio di IVS nel discriminare i soggetti ad alto rischio. Tuttavia, i risultati devono essere interpretati nel contesto di alcune limitazioni e potenziali prospettive. Nessuno SNP ha raggiunto il livello di significatività di Bonferroni, probabilmente a causa della dimensione limitata del campione analizzato. Tuttavia,

l'omogeneità fenotipica della coorte analizzata e l'assenza di precedenti trattamenti antipertensivi, sono stati i presupposti dell'analisi per l'identificazione di veri effetti genetici. Una replica in una coorte indipendente è in genere necessaria per confermare ulteriormente i risultati; tuttavia una coorte indipendente con criteri simili non era disponibile per la replica. Inoltre, come spesso accade in studi di questo tipo, gli SNP significativi si trovano in regioni non codificanti, e questo fatto rende difficile spiegare il loro ruolo causale. Queste limitazioni tuttavia non sminuiscono la rilevanza dei geni identificati e confermati da lavori pubblicati in precedenza.

Le prospettive future di questo studio dovrebbero essere la replica dei risultati in coorti indipendenti e la valutazione della capacità predittiva del wGRS di classificare correttamente i veri positivi e i veri negativi in base al loro background genetico.

ABSTRACT

Left ventricular mass (LVM) is an important clinical phenotype, whose assessment can predict adverse cardiovascular events and premature death in all genders, races, and ages. Increase in LVM defines left ventricular hypertrophy (LVH) with the thickening of the left ventricle of the heart. In community-based cohorts, the presence of left ventricular hypertrophy (LVH) and increased LVM predict the development of coronary heart disease, congestive heart failure, stroke, and cardiovascular disease. Thus this trait serves not only as measures of cardiac structure, but also as intermediate phenotype for clinical cardiovascular disease outcome. Several studies have indicated that LVM is influenced by genetic factors. Genome wide linkage and association studies in diverse population have been performed to identify genes influencing LVM, but much of the heritability remains unexplained, the identity of the underlying gene pathways and functional variants remain unknown, and the promise of genetically based risk prediction remains unfulfilled.

The aim of the study was to investigate the association of common genetic variants with left ventricular mass using a genome wide approach in a large cohort of never treated mild-to-moderate essential hypertensive subjects. From the linear analysis, we selected 85 single nucleotide polymorphisms (SNPs), with a suggestive p-value of association with LVM (≤ 10-5). In particular, some SNPs lying in genes previously described as having a role in the pathogenesis of cardiac hypertrophy, such as ROCK1, IGF1, CACNA1D, FGFR1, TRAF5, SOX5, and KSR2. Each of them might play a putative role in determining the LVM phenotype as well as other pathophysiological pathways directly or indirectly linked to cardiac pathophysiology.

To assess the risk alleles associated to the most interesting findings in relation to the phenotype studied, we performed a case-control analysis by dividing our sample in two subsets according to LVM values. Most of the SNPs associated with LVM in linear regression presented a significant association, showing that the carriers of the risk alleles have an odds ratio higher than 1 to have increased LVM, i.e. to be cases respect to controls. Nevertheless as for most of the complex traits, the observed odds ratios are modest (except for those biased by the absence of homozygous risk genotypes), so their relevance for a clinical use is uncertain. Thus, we defined a weighted genetic risk score using the effect size of the risk allele (beta value of the linear regression analysis) to account for the strength of the genetic association with each allele. The possibility to combine more variants in a global genetic risk score could be interesting and could add relevance to the results.

In conclusion, our GWAS allowed us to pinpoint genes whose role in heart function and/or cardiac hypertrophy has been demonstrated in previously publications by different authors. Moreover, we highlighted the usefulness of an aggregate measure of risk of LVH to discriminate high-risk subjects. However, the results must be interpreted within the context of some potential limitations and perspectives. No SNPs reached a Bonferroni's significance level probably due to a limited sample size. However, the phenotypic homogeneity of our cohort and the absence of previous antihypertensive treatment are prerequisites for the identification of true genetic effects. A replication in independent cohorts is needed

to further confirm the findings; however an independent cohort with similar criteria was not available for replication. Moreover, it often happens, as in our study, the significant SNPs map in non-coding regions, making it difficult to explain their causative role. These limitations should not reduce the relevance of the genes identified and confirmed by previously published papers.

Future perspectives of this study should be the replication of the GWAS findings in independent cohorts and the assessment in independent samples of the prediction ability of wGRS to correctly classify true positives and true negatives according to their genetic background.

TABLE OF CONTENTS

LIST OF ACRONYMS AND SYMBOLS

ASE	American Society of Echocardiography
BMI	body mass index
BP	blood pressure
BSA	body surface area
CG	candidate gene
CMR	cardiac magnetic resonance
CVD	cardiovascular disease
DBP	diastolic blood pressure
ECG	echocardiography
EH	essential hypertension
GWA	genome wide association
GWAS	genome wide association study
HR	heart rate
IBD	identity by descent
IVS	ipertrofia del ventricolo sinistro
IVST	interventricular septum thickness
LD	linkage disequilibrium
LV	left ventricle
LVDd	left ventricular diameter in diastole
LVH	left ventricle hypertrophy
LVID	left ventricular internal diameter
LVM	left ventricular mass
$LVMh^{2.7}$	LVM indexed for $height^{2.7}$
LVMI	left ventricular index
MAF	minor allele frequency
MVS	massa del ventricolo sinistro
OR	odds ratio
PC	principal component
PCA	principal component analysis
PWT	posterior wall thickness
QC	quality control
RAS	renin-angiotensin system

SBP	systolic blood pressure
SD	standard deviation
SE	standard error
SNP	single nucleotide polymorphism
sPRA	sitting plasma renin activity
uK24h	urinary potassium 24 hours
uNa24h	urinary sodium 24 hours
wGRS	weighted genetic risk score

LIST OF FIGURES AND TABLES

1 INTRODUCTION

1.1 DEFINING LVM

Left ventricular mass (LVM) is an important clinical phenotype, whose assessment can predict adverse cardiovascular events and premature death in all genders, races, and ages [1]. Increase in LVM defines left ventricular hypertrophy (LVH) with the thickening of the left ventricle of the heart. The diagnosis of LVH has been incorporated in clinical practice as marker of cardiovascular risk [2]. Population-based studies have revealed that increased LVH provide prognostic information in hypertension, and in the general population, as well as in a variety of clinical settings [3]. Furthermore, regression of LVH appears to be a favorable prognostic marker independent of the treatment-induced blood pressure (BP) reduction [3].

In the Sixties, the association of echocardiography LVH with cardiovascular events was first described in clinical and population-based studies. The 20% of patients with LVM exceeding a predefined cut-off had approximately 4-fold higher rate of morbid events (24%), than the patients without LVH (6%). Other studies have subsequently extended these findings by demonstrating that increased LVM strongly predicts cardiac and cerebrovascular morbidity and mortality, independent of traditional risk factors [4]. A report from the Framingham Heart Study showed that increased LVM strongly predicted all-cause of cardiac mortality and coronary heart disease events in adults over 40 years, independently of conventional risk factors. Incidence of stroke and transient ischemic attack was higher in the highest quartile than in the lowest one [5,6]. Similarly, in the Cardiovascular Health Study the incidence of congestive heart failure was higher in the highest quartile than in the lowest one [7]. Another

analysis from the Framingham Heart Study reports a linear relationship between LVM and the rate of future cardiovascular events [8].

Therefore, LVM has been touted not only as measures of cardiac structure, but also as a suitable measure for cardiovascular disease (CVD) risk stratification and a marker for subclinical disease [9].

LVM shows a continuous distribution in the general population, whereas LVH defines the upper limit end of LVM distribution [3]. LVH is a common condition, both in the general population (in the United States occurs in 16% of white and in 33-43% of black people) and in hypertensives (22-60% of US hypertensives) [10]. Although historically considered an adaptive response of the heart to hypertension, it is now recognized that the presence and magnitude of LVH varies substantially among individuals. At equal BP levels, some individuals develop LVH, whereas others do not. Because the major causes of morbidity and mortality among hypertensives are due to the cardiovascular (CV) manifestation of hypertension (i.e. LVH) and not to the level of blood pressure per se, it is important to understand the causes of LVH [10].

1.2 DETERMINANTS OF LEFT VENTRICULAR MASS

In both healthy individuals and patients with disease, LVM is likely to be determined by a combination of genetic factors and adaptive responses to environmental and mechanical factors.

Gender and body size have been identified as predictors of LVM, and are used as covariates to corrected LVM measurement. LVH diagnosis is based on this corrected measure. Many other constitutional factors and exposures may lead to pathophysiological changes in LVM.

1.2.1 Blood pressure and hypertension

Numerous population based studies have shown an association between hypertension and LVM [11,12,13]. The prevalence of LVH varies with severity of hypertension, ranging from < 20% in mild hypertension to almost 100% in severe or complicated hypertension [14]. Even within the normal range, increases in blood pressure are related to an increased LVM, due to both hypertrophic response to increased overload, and to neuro-humoral and genetic factors.

There is some evidence for involvement of the renin-angiotensin system (RAS), with impaired suppression of the RAS or increased sensitivity to angiotensin II appear to act as stimulus for LVH in hypertensive patients. Several studies confirm an association between increasing plasma renin activity levels and increased LVH [14]. Experimental evidences suggest that angiotensin II induces hypertrophy in myocytes and vascular smooth muscle cells and may regulate collagen synthesis. Aldosterone also may stimulate extracellular collagen deposition and myocardial fibrosis [14]. There is much evidence that dietary salt intake plays a role in the development of LVH in hypertensive patients, although the mechanism is not still clear [14].

LVH is an independent risk factor for CVD in patients with hypertension. The underlying mechanisms for this association may include a combination of electrophysiological alterations, anatomical changes, and increased sympathetic RAS activity [14]. LVH is recognized as a hypertension-related target organ damage in several clinical practice guidelines, representing an intermediate unfavorable prognostic marker [2].

1.2.2 Age, gender and ethnicity

Several studies show a significant univariate association between age and LVM [5]. In the Framingham population, echocardiography-detected LVH prevalence ranged from 6% in persons under 30 years of age to 43% in those with ≥ 70 years [14]. LVM progressively increases during aging both in normotensive and hypertensive subjects [2]. However, multivariate analyses suggest that adjusting for lean mass, fat mass and blood pressure removes some or all of the relationship with age, suggesting that age differences in LVM are at least partly due to age differences in lean mass, fat mass and blood pressure [15,16,17]. Nevertheless it appears prudent to adjust for age, gender, and ethnicity in epidemiological investigations related to LVM and LVH [18].

Several studies report strong gender differences in LVM [19,20]. LVH detected by echocardiography was more prevalent in women than in men (2.9% vs. 1.5%), whereas echocardiographically detected LVH was more common in men than in women (17.6% vs. 14.2%) [21]. However, indexing LVM to lean mass or height$^{2.7}$ usually results in similar LVM index values in men and women, suggesting that differences between men and women in lean body mass and body size account for most of the gender gap in LVM [22,23].

Ethnicity also plays a role in the epidemiology of LVH. The prevalence of LVH is higher in African-Americans and Africans, and ancestry has been identified as an independent risk factor for LVH [2,24]. This is probably attributable, in part, to the elevated risk of hypertension in African-Americans and Africans, with one study showing an almost fourfold increase in the incidence of LVH in blacks as compared to whites (odds ratio 3.7, 95% CI 3.2-4.4) [25,26]. Other studies show no race differences in either LVM indexed for body size or in the prevalence of LVH [27,28].

1.2.3 Obesity and diabetes

Obesity has been shown independently associated to LVM [29], particularly in populations with high prevalence of hypertension and other metabolic risk factors [30,31]. Despite this association, the impact of obesity on LVM may be less than expected, thus “uncomplicated obesity” was not a risk factor for LVH indexed by body surface area (BSA) or height$^{2.7}$ [32,33]. Together with obesity and hypertension, diabetes is an important determinant of LVM in most population-based studies [34,35,36,37].

1.2.4 Other risk factors

Environmental exposure, such as alcohol consumption [38], salt intake [39], smoking [11,38], and physical activity in men [40] have been associated to increased LVM. Other factors such as blood lipids, pulmonary function, heart rate, low weight at one year-old, and hematocrit have also been implicated, but with some inconsistency among different studies [7,41,42,43]. Clinical validity and impact of those factors is controversial, but it may be important to consider them as potential confounders in epidemiological studies investigating the role of risk factors in LVH and the role of LVM in disease prediction [18].

1.3 LVM HERITABILITY

Only one-half to two-thirds of the inter-individual variability of LVM can be explained by its clinical and hemodynamic correlates. Several studies have indicated that LVM is influenced by genetic factors [4]. Monozygotic twins have substantially more similar LVM than dizygotic twins [4]. Family-based studies have confirmed that echocardiographic measures of LVM, after adjustments for covariates, are heritable. Heritability of LVM estimated in studies of twins, hypertensive siblings, nuclear, and complex families ranged from 15 to 84% [44].

1.3.1 Previous studies on genetics of LVM

The normal distribution of LVM in populations suggests that this phenotype is a complex trait influenced by multiple genes. Genetic analysis indicates that the segregation of LVM was compatible with an additive polygenic model [45].

Candidate genes studies have shown a potential role of genetic polymorphisms located in angiotensin-converting enzyme (*ACE*) [46,47,48,49], aldosterone synthase (*CYP11B2*) [50], insulin-like growth factor (*IGF1*) [51], neuropeptide Y (*NPY*) [52], guanine nucleotide-binding protein 3 (*GNB3*) [53], endothelial nitric oxide synthase (*eNOS*) [54], peroxisome proliferator-activated receptor-alpha (*PPARA*) [55], and centlein centrosomal protein (*CNTLN*) [56] genes. Genome-wide linkage and association studies have shown an association between LVM and several loci located in different chromosomes. Particularly, in a whole genome linkage study of hypertensive families, three regions (10q23.1, 12q14.1, 17p13.3) were found to approach suggestive evidence of linkage for particular measures of LVH [57]. A genome wide association study (GWAS) on Koreans reported a significative correlation between the skeletal muscle Ca(2+) channel protein *RYR1* gene on chromosome 19 and LVH [58]. The HyperGEN study identified a polymorphism in *KCNB1* gene associated with LVM using a genome wide approach [9]. A large meta-analysis identified loci associated with left ventricular structure on the solute carrier family 35, member F1 (*SLC35F1*) gene, the chromosome 6 open reading frame 203 (*C6orf203*) and the phospholamban (*PLN*) gene [59]. Recently, Barve RA *et al* identified eleven SNPs with a suggestive association with left ventricular mass trait, in a comparative study between M-mode and 2D echocardiography and between raw LVM and BSA-indexed LVM. One SNP lies in *CDH13* gene were confirmed in all the four measures [60]. However, the physiopathological link between genes and

LVM remains unclear and these results explain only a small part of LVM variance. The main limits in association studies of LVM with gene variants in essential hypertension (EH) are a) the limited sample size of the study cohorts, and b) the presence of patients under antihypertensive treatment or after short wash-out periods. As long-term treatment with antihypertensive drugs modifies cardiac mass, large cohorts of never treated essential hypertensives are mandatory to study the association between LVM and gene variants, although we are perfectly aware of the difficulty to recruit large cohorts of such patients [61,62].

1.4 LEFT VENTRICULAR MEASUREMENT

Given the clinical importance of LVM, it is essential to have a reliable method for its estimation. Echocardiography (ECG) has been clinically employed for more than 30 years, becoming one of the most important non-invasive imaging methods in the evaluation of cardiac morphology and dynamics. Despite inherent limitations, conventional echocardiography continues to be the imaging modality of choice for the assessment of LVM in routine clinical practice. Echocardiography has been shown to reliably characterize LVM, and its use has been extensively validated in clinical care and in research [63,2].

A standard echocardiogram is also known as a transthoracic echocardiogram (TTE), or cardiac ultrasound. In this case, the echocardiography probe is placed on the thorax of the subject, and images are taken through the chest wall. This is a non-invasive, highly accurate and quick assessment of the overall health of the heart. An alternative way to perform an echocardiogram is transesophageal echocardiogram (TEE); TEE provides additional and more accurate information than TTE for some patients, for several specific diagnoses and for many catheter-based cardiac interventions [64].

1.4.1 Imaging mode

Motion-mode (M-mode), 2-dimensional echocardiography (2DE) and finally 3-dimensional echocardiography (3DE) are the imaging modalities used to estimate LVM [65]. M-mode ECG was one of the earliest modalities of echocardiography to come into clinical use. It allows better endocardial border definition as it has greater resolution due to higher frame-rate; 2DE imaging, on the other hand, depicts the “real” ventricular shape, but has lower lateral resolution and frame-rate. 3DE is feasible in the clinical setting and provides fast and accurate assessment of LVM, which is superior to conventional echocardiographic methods, especially in distorted hearts [63].

Although 2D or 3D echocardiography can be more accurate, M-mode was the first non-invasive imaging technique developed and remains the recommended method. Most epidemiological reports use M-mode imaging modality; preference for M-mode is based on its technical feasibility and availability at the time when most studies were performed [2].

1.4.2 Calculating and indexing left ventricular mass

The most common used formulas to estimate LVM are all variations of the same mathematical principle, based on the volume formula; whether using M-mode, 2D, or 3D measurements, LVM estimation by echocardiography is based on subtraction of the left ventricular cavity volume from the volume enclosed by the correspondent epicardium to obtain left ventricle muscle or shell volume. This shell volume is then converted to mass by multiplying the myocardial density (taken to be 1.05 g/ml).

To date, most LVM calculations have been made. The most used formula is that by Deveraux and colleagues, based on American Society of Echocardiography (ASE) criteria of edges definition:

$$LVM\ (ASE) = 0.8 \times \{1.04 \times [(LVID + PWT + IVST)^3 - (LVID)^3]\} + 0.6\ g$$

This formula was validated on the necropsy findings of 52 individuals (r=0.9, p-value < 0.001) [66].

In order to allow comparison of LVM among subjects of different body sizes, different allometric approaches have been suggested to normalize LVM: height, diverse allometric height adjustment, weight, body surface area, body mass index, and free-fat mass. However, the best way for LVM normalization is still controversial and results in different patient classifications. Body surface area was the first anthropometric variable used to index LVM and has shown a stronger statistical correlation than height with LVM, but underestimates the prevalence of obesity-related LVH [67]. Consequently, height has also been used for indexing (either height alone or height raised to an allometric power of 1.7 or 2.7). Indexation of LVM to height raised to an allometric exponent of 2.7 (LVM/height$^{2.7}$) has shown better predictive value for CVD outcomes, better detection of obesity-related LVH, and less variability of LVM among normal individuals [68]. BSA has been widely adopted by the ASE and European Association of Cardiovascular Imaging as the preferred method for indexing LVM [65].

1.4.3 Determining cut-off points

The determination of LVM cut points to define abnormality is a source of controversy and can be driven by different strategies. American Society of Echocardiography recommends reference values for LVM, obtained from an ethnically diverse population of 510 normal-weight, normotensive, and non-diabetic white, African American, and American Indian adults, without recognized cardiovascular disease [65]. Reference upper limits of normal LVM by linear measurements are reported in table 1.

Table 1 Reference limits and partition values for left ventricular mass, determined by linear M-mode imaging.

	Women				Men			
	Reference range	**Mildly abnormal**	**Moderately abnormal**	**Severely abnormal**	**Reference range**	**Mildly abnormal**	**Moderately abnormal**	**Severely abnormal**
LVM (g)	67-162	163-186	187-210	≥ 211	88-224	225-258	259-292	≥ 293
LVM/BSA (g/m^2)	43-95	96-108	109-121	≥ 122	49-115	116-131	132-148	≥ 149
LVM/height (g/m)	41-99	100-115	116-128	≥ 129	52-126	127-144	145-162	≥ 163
$LVM/height^{2.7}$ ($g/m^{2.7}$)	18-44	45-51	52-58	≥ 59	20-48	49-55	56-63	≥ 64

BSA, body surface area; LVM, left ventricular mass; 2D, 2-dimensional (adapted from Ref. 18).

2 AIM OF THE STUDY

The aim of this exploratory study was to investigate the association of common genetic variants with left ventricular mass using a genome wide approach in a large cohort of never treated mild-to-moderate essential hypertensive subjects. After identification of some genetic susceptibility loci, we aimed to create a “weighted genetic risk score” that aggregates the measure of risk of increased LVM.

By targeting intracellular signaling pathways involved in regulation of LVM, it should be possible to define therapeutic strategies for inhibiting hypertrophy of heart cells, and thus to reduce the risk of cardiac morbidity and mortality.

3 MATERIALS AND METHODS

3.1 OVERVIEW

Data presented in the study are part of the data generated within two projects: 1) “HYPERGENES Project” - European Network for Genetic Epidemiological Studies: building a method to dissect complex genetic trait, using essential hypertension as a disease model - which is a collaborative project financed by the European Commission in the 7th Framework Program [69];

2) “InterOmics Project” - Development of an integrated platform for the application of "omic" sciences to biomarker definition and theranostic, predictive and diagnostic profiles (financed by MIUR-CNR http://www.interomics.eu).

3.2 STUDY PARTICIPANTS AND INCLUSION CRITERIA

The analyzed sample consists of 1,029 newly diagnosed mild-to-moderate essential hypertensive patients - i.e. no previous antihypertensive treatment - of Caucasian origin. Patients were recruited using the following inclusion criteria: 1) baseline untreated BP levels in the hypertensive range (systolic BP ≥ 140 mmHg or diastolic BP ≥ 90 mmHg; 2) no previous antihypertensive treatment; 3) absence of comorbidities (i.e. type 1 and type 2 diabetes); 4) glycaemia ≤ 150 mg/dL; 5) serum creatinine < 2 mg/dL.

They were enrolled during several pharmacogenomics studies, performed at the “Hypertension and Related Diseases Centre-AOU” - University of Sassari and at other eleven Clinical Research Centers all over Italy (table A1). The pharmacogenomics studies were developed also in collaboration

with the “Genomic and Bioinformatics Laboratory” of the University of Milan [69,70,71].

Participants who were all untreated underwent a run-in period of eight weeks under standardized dietary conditions to qualify the presence of EH. During this period, BP was measured weekly to meet the inclusion criteria whereas all the other measurements of cardiac, renal and metabolic phenotypes were performed after the run-in period.

The study was approved by the Ethics Committee of the University of Sassari and supported by all the other local Ethics Committees. Written informed consent was obtained for the study and for the DNA analysis, and all clinical investigation was conducted according to the Declaration of Helsinki.

3.3 PHENOTYPES

3.3.1 LVM Quantitative phenotype

Left ventricular dimensions were measured by transthoracic echocardiography according to the guidelines of ASE [18]. The linear internal measurements of the left ventricle (LV) were acquired in the parasternal long-axis view and obtained using M-mode tracing perpendicular to the LV long-axis immediately below the level of the mitral valve leaflet tips. The average of at least four consecutive measurements at the end-diastolic phase (R-wave peak of ECG trace) was considered to determine: interventricular septum thickness (IVST), left ventricular internal diameter (LVID), and posterior wall thickness (PWT). LVM was estimated using the Devereux equation [66], and indexed for height$^{2.7}$ (LVMh$^{2.7}$) [67]. Left ventricular hypertrophy was defined by left ventricular mass index > 48 g/m$^{2.7}$ in men and > 44 g/m$^{2.7}$ in women [18]. TTE measurements were performed by different operators and blindly revised by a second operator:

the presence of different operators in the evaluation of LVM, although blind each other, was considered as a relevant covariate in the association analysis.

3.3.2 Case-control phenotype

We divided the sample according to LVM values: 158 subjects with abnormal LVM (≥ 52 g/m$^{2.7}$ for women and ≥ 56 g/m$^{2.7}$ for men) were identified as cases and 615 subjects with LVM in the normal range (≤ 44 g/m$^{2.7}$ for women and ≤ 48 g/m$^{2.7}$ for men) were identified as controls. We did not consider subjects with intermediate LVM values (between 45 and 51 g/m$^{2.7}$ for women and between 49 and 55 g/m$^{2.7}$ for men). We used as reference limits and partition values of LVM the values indicated by Lang RM and colleagues [18].

3.4 GENETIC CHARACTERIZATION

Genomic DNA was extracted from peripheral blood with standard procedures. Genotyping was performed with the Illumina Human1M-Duo array (Illumina Inc, San Diego, CA, USA) within HYPERGENES project (n=901) and the Illumina HumanOmniExpress array within the InterOmics project (n=139). Illumina Human1M-Duo array captures 1,199,187 genetic loci. In addition to markers necessary for broad genome coverage (of which 672,002 within 10 Kb of RefSeq genes, 21,877 non-synonymous SNPs, and 483 Indel), the chip contains 51,207 markers in sex chromosomes, 138 markers in mitochondrial DNA, 35,969 markers covering copy number variant regions, and 30,908 markers in MHC/ADME regions. Illumina HumanOmniExpress array captures > 713,014 markers. In addition to markers necessary for broad genome coverage (of which 395,094 within 10 Kb of RefSeq genes, 12,286 non-synonymous and 10,854 synonymous), the chip contains 19,485 markers in sex chromosomes and 17,712 markers in HLA/ADME regions. Genotyping was performed at “Genomic and

Bioinformatics Laboratory" of University of Milan. Raw intensity data were analyzed with the Illumina software "GenomeStudio" for genotype calling (http://bioinformatics.illumina.com/informatics/sequencing-microarray-data-analysis/genomestudio.html), using the Illumina reference cluster file.

3.5 DATA QUALITY CONTROLS

After the generation of raw genotyping data, we performed cleaning procedures that are crucial steps to avoid false positive and false negative results. All quality control (QC) steps were performed in accordance with the protocol by Anderson C.A. and colleagues [72], using PLINK software (version 1.07) [73]. SNPs and subjects that passed QCs have been then tested for imputation and further association analyses.

3.5.1 Sex mismatches

The gender information of each individual reported in clinical data records was compared to that estimated using X-chromosome markers. When sex discrepancies could not be resolved by clinicians who conducted phenotyping, the individuals were not considered in the analysis.

3.5.2 Call rate per individual and per SNP

The individual call rate is the proportion of genotypes per subject with non-missing data. Accordingly, for each SNP, the call rate is the percentage of subjects with a non-missing genotype attribution. We filtered-out subjects with call rate < 0.98 and SNPs with call rate < 0.99.

3.5.3 Minor Allele Frequency

Minor allele frequency (MAF) is the lowest allele frequency at a specific locus observed in a particular population. MAF indicates how much frequently the minor allele of a SNP is found in the sample under study. We removed all SNPs with a very low MAF setting a threshold at 1%.

3.5.4 Heterozygosity

We inspected the distribution of mean heterozygosity across all individuals, to identify subjects with excessive or reduced proportion of heterozygote genotypes, which may be indicative of DNA samples contamination or inbreeding respectively. Mean heterozygosity is calculated as (N(NM) − O(Hom))/N(NM), where [N(NM)] denotes the number of non-missing genotypes per individual and [O(Hom)] is the observed number of homozygous genotypes. We excluded all individuals with heterozygosity rate ± 3 standard deviations from the mean.

3.5.5 Homogeneity analysis

In order to identify duplicated or related individuals, we performed the homogeneity analysis of samples using a genome wide identity by descent (IBD) estimation. Two or more alleles are identical by descent if they are identical copies of the same ancestral allele. IBD estimation is based on the average proportion of alleles shared between subjects. We used estimates of pairwise IBD to find pairs of subjects who look too similar to each other, more than we would expect by chance in a random sample.

3.5.6 Population stratification detection

In order to identify subjects with large-scale differences in ancestry, we assessed the population stratification using the principal component analysis (PCA), as implemented in the EIGENSOFT package (version 3.0) [74,75]. For PCA we used SNPs in common between the two genotyping arrays (≈ 400 k). Population stratification is a major confounder in association studies that occurs when allele frequencies differ between subjects of the comparison samples due to ancestry differences, various ethnic backgrounds or even to “hidden” stratification. The presence of substructures in the population can lead to spurious association between a phenotype and unlinked candidate loci, causing either false positive of false

negative results. Based on principal components value for each individual, the software calculates a mean and a standard deviation. We removed genetic outliers, defined as individuals that exceed a default number of standard deviations (6.0) from the whole sample along any of the principal components. The first ten principal components (PCs) were selected to correct for stratification by including them as covariates in the linear regression model.

3.6 IMPUTATION

In order to infer genotypes at not directly typed loci, we performed a two-step imputation method on markers of highest quality that consists of: 1) phasing of the typed genotype using SHAPEIT2 software [76] and 2) imputation of the genotypes on the reference panel, using Minimac [77]. We used the high-density panel of the 1000 Genomes Project as reference (www.1000genomes.org) [78] based on the release of March 2012, which integrates more than 39 million variants.

Through imputation we could increase the overall number of markers available for association testing and we could also combine data from the two arrays used for genotyping. In fact, when a dataset is collected using two or more arrays with different sets of markers, some markers are not assayed across the entire dataset, because they were not present in both arrays. This limits the association analysis at those markers typed with both arrays. Using imputation, we could predict genotypes at loci that are not in common between the two genotyping arrays. Thus imputation increases the sample size at each marker in the total number of individuals genotyped across the entire study.

Imputation accuracy was evaluated using “imputation R-square” (Rsq), a parameter provided by Minimac software, which estimates the squared

correlation between imputed and true genotypes, i.e. the ratio of the variance of imputed and true allele count [79].

3.7 GENOTYPE-PHENOTYPE ASSOCIATION ANALYSES

We performed two types of analyses: 1) a quantitative trait association analysis using the imputed allele dosages, as implemented in Mach2qtl [80]; 2) and a case-control analysis dividing our sample according to LVM values and focusing on SNPs significantly associated to LVM in the linear analysis.

In the quantitative trait analysis, the analyzed phenotype was LVM estimated by Devereux equation, in accordance with the ASE criteria, and indexed for height$^{2.7}$ (LVMh$^{2.7}$), as described in the *paragraph 3.3.1*. To assess the genotype to phenotype association we performed a linear regression on LVMh$^{2.7}$, under an additive model, adjusting for some covariates: sex, systolic and diastolic blood pressure (SBP, DBP), serum creatinine, urinary sodium 24 h (uNa24h), sitting plasma renin activity (sPRA), body mass index (BMI), heart rate (HR), first ten PCs, and TTE operator. The selection of covariates to be included in the model was performed through analysis of variance (ANOVA), as implemented in StataSE. To estimate significance threshold, results was controlled for multiple testing using Bonferroni's adjustment.

For case-control analysis, we divided the sample according to LVM values as described in the *paragraph 3.3.2*. We performed a logistic regression adjusting for the above-mentioned covariates.

3.8 WEIGHTED GENETIC RISK SCORE

We investigated the usefulness of an aggregate measure of risk of LVH based on the selected genetic susceptibility *loci*, weighted with the effect size on the LVM trait. Following the model presented by De Jager PL *et al*

[81], we selected a subgroup of SNPs from the GWA analysis of LVM trait and designed a “weighted genetic risk score” (wGRS). In the wGRS algorithm, we used the beta effects from quantitative GWA results, to ascertain the strength of the genetic association with each allele. The wGRS was calculated by multiplying the number of risk alleles for each SNP by the weight for that SNP and then taking the sum across all SNPs, according to the following formula:

$$wGRS = \sum_{i=1}^{7} (w_i \times X_i)$$

where i is the SNP, w_i is the weight of the SNP i, and X_i is the number of risk alleles (0, 1 or 2). The weight for each SNP is the beta effect for each allele, obtained from the linear regression analysis.

After calculation of wGRS, we tested the distribution of the score in cases and controls separately, using a two-sample t-test. The two groups were defined as in the case-control logistic analysis, described in *paragraph 3.3.2*. All analyses were done using Stata SE (version 11).

Since a continuous score is difficult to interpret on an individual level, we partitioned subjects into different categories of risk. These risk categories were created using the means and standard deviations (SD) of wGRS from the control samples. The seven categories were defined as ± 0.25, ± 0.75 and ± 1.25 SDs from the mean, with the extreme categories being < 1.25 and > 1.25 SDs from the mean. Dividing our score into seven categories provided a robust distribution, allowing us to parse out the highest and lowest risk groups, while assuring that there were statistically sufficient numbers of cases and controls in these extreme categories of interest. We used as reference the “category 4”, which contained the wGRS mean of the control population. The subjects in this category can be considered as

showing the average risk of the assessed population. Within the case-control dataset, we calculated the odds ratio for each category.

3.9 STATISTICAL SOFTWARE

3.9.1 PLINK

PLINK (version 1.07) is a free, open-source, whole genome association analysis toolset, designed to perform a range of basic, large-scale analyses in a computationally efficient manner [73]. PLINK was written in C/C++ language. The focus of PLINK is purely on analysis of genotype/phenotype data. PLINK provides a simple way to handle large GWAS datasets, assesses confounding due to stratification and nonrandom genotyping failure and to produce a range of other summary statistics, performs a variety of standard association tests on very large datasets in populations or families, for disease or quantitative outcomes, allowing for covariates, haplotipic tests, *etc*. PLINK is being developed by Shaun Purcell at the Center for Human Genetic Research, Massachusetts General Hospital, and the Broad Institute of Harvard & MIT, with the support of others (http://pngu.mgh.harvard.edu/purcell/plink).

3.9.2 Eigensoft

The EIGENSOFT package (version 3.0 for Linux platform, Department of Genetics, Harvard Medical School, Boston, USA) uses PCA to correct for population stratification in medical association studies (EIGENSTRAT) [74] and to detect and analyze population structure (SMARTPCA) [75]. It combines functionality from population genetics methods and EIGENSTRAT stratification correction method. The EIGENSTRAT method uses PCA to explicitly model ancestry differences between cases and controls along continuous axes of variation. The method produces several uncorrelated variables from a data matrix containing observation across a

number of potentially correlated variables; in the PCA model of ancestry detection, the observations are the individuals and the potentially correlated variables are the markers. The resulting correction is specific to a candidate marker's variation in frequency across ancestral populations, minimizing spurious associations while maximizing power to detect true associations. The EIGENSOFT package has a built-in plotting script and supports multiple file formats and quantitative phenotypes (http://genepath.med.harvard.edu/~reich/Software.htm).

3.9.3 SHAPEIT2

SHAPEIT2 is a fast and accurate method for estimation of haplotypes (aka phasing) from genotyping or sequencing data [76]. It takes as input a set of genotypes and a genetic map, and produces as output, either a single set of estimated haplotypes, or a haplotype graph that encapsulates the uncertainty about underlying haplotypes (https://mathgen.stats.ox.ac.uk/genetics_software/shapeit/shapeit.html).

3.9.4 Minimac

Minimac (2012-11-16 version) is a low memory, computationally efficient implementation of the MaCH algorithm for genotype imputation [77]. MaCH is a Markov Chain Haplotyping software package for haplotype estimation and genotype imputation (http://www.sph.umich.edu/csg/abecasis/MACH/index.html). Minimac is designed to work on phased genotypes and can handle very large reference panels with hundreds or thousands of haplotypes. It provides imputation results as probabilistic calls and not as discrete genotypes. These probabilistic genotype calls should not be converted into discrete genotypes, as that can result in a substantial loss of information [80]. Imputed allele count for each allele can conveniently be tested for

association with phenotypic traits using appropriate software, such as Mach2qtl (http://genome.sph.umich.edu/wiki/Minimac).

3.9.5 Mach2qtl

Mach2qtl (V1.1.2) software performs quantitative trait association analysis based on allele dosages or genotype probabilities inferred from imputation software [80]. Mach2qtl was developed by Goncalo Abecasis at the University of Michigan (http://www.sph.umich.edu/csg/yli/software.html).

3.9.6 Stata SE

Stata SE (version 11) is a commercial data analysis and statistical software, created in 1985 by “StataCorp” for managing, analyzing, and graphing data. Stata SE is used for analysis of large databases (http://www.stata.com).

3.9.7 R

R (version 2.14.1, copyright 2011, The R Foundation for Statistical Computing) is a free language and environment for statistical computing and graphics. It is a GNU project (http://www.gnu.org) which is similar to the S language and environment, developed at Bell Laboratories (formerly AT&T, now Lucent Technologies) by John Chambers and colleagues. R can be considered as a different implementation of S. R provides a wide variety of statistical (linear and nonlinear modeling, classical statistical tests, time-series analysis, classification, clustering, etc.) and graphical techniques; it is highly extensible and presents some packages implementing statistical methods and algorithms for the analysis of genetic data and for related population genetics studies (http://www.r-project.org).

3.9.8 Locuszoom

LocusZoom is designed to facilitate viewing of local association results together with useful information about a locus, such as the location and orientation of the genes it includes, linkage disequilibrium coefficients and local estimates of recombination rates [82]. LocusZoom provide plot summaries of genome-wide scan interactively. LocusZoom allows for quick visual inspection of the strength of association evidence, the extent of the association signal and linkage disequilibrium (LD), and the position of the associated SNPs relative to genes in the region. LocusZoom plots provide an option to size the data points relative to sample size and can display functional annotation. LocusZoom was written in R using the grid and lattice graphics packages and runs within a Python wrapper (http://csg.sph.umich.edu/locuszoom).

4 RESULTS

The result section is composed by:

1. quality controls;
2. descriptive statistics of the cohort;
3. genotype-phenotype association with LVM quantitative trait;
4. genotype-phenotype case-control analysis;
5. aggregation of findings into a weighted genetic risk score.

4.1 QUALITY CONTROLS

One thousand and twenty-nine subjects underwent quality control (QC) of genetic data.

Six DNA samples were excluded for low call rate (< 98%) and 8 subjects for reduced or increased proportion of heterozygous genotypes. No sex mismatches were identified.

Using genome-wide IBD estimation, we identified and removed from the analysis 2 duplicated and 42 related subjects (34 first-degree and 8 second-degree).

In order to identify individuals with large-scale differences in ancestry, we assessed the population stratification within the data using the principal component analysis. We removed 10 outliers defined as individuals that exceed 6 standard deviations from the whole sample along any of the principal components. Results for the first two PCs are described in figure 1.

Figure 2 shows the subjects' flow from the recruitment to the pre-analysis quality control steps. After quality control the final sample is composed of 966 subjects (633 males and 333 females).

Markers were filtered for call rate (threshold at 0.99) and MAF (threshold at 0.01), leaving 458,953 SNPs, in common between the two genotyping array.

These SNPs were used for imputation. After imputation, markers were filtered for minor allele frequency (threshold at 0.99), and for imputation quality (threshold of Rsq at 0.8), leaving 7,239,388 SNPs for the association analysis.

4.2 SAMPLE CHARACTERISTICS

The clinical characteristics of the 966 patients enrolled in the study are reported in table 2. Study participants were essential hypertensive white Caucasians and included 333 women (34.5%). Age averaged 47.8 years (SD ± 9.0); average (± SD) SBP and DBP were 154.8 ± 12.4 mmHg and 100.1 ± 8.0 mmHg, respectively. The mean $LVMh^{2.7}$ observed was 44.5 $g/m^{2.7}$ (SD ± 10.0). $LVMh^{2.7}$ distribution is shown in figure 3 and $LVMh^{2.7}$ distribution according to gender is shown in table 3. To assess statistical difference in mean $LVMh^{2.7}$ between males and females, analysis of variance (ANOVA) was performed, as implemented in StataSE. $LVMh^{2.7}$ is not significantly different between males and females (p-value = 0.3).

For the case-control analysis, we divided the sample according to LVM values. As described in the *Materials and Methods section*, we defined as case a subject with LVM ≥ 52 $g/m^{2.7}$ for women and ≥ 56 $g/m^{2.7}$ for men, and as control a subjects with LVM in normal range (≤ 44 $g/m^{2.7}$ for women and ≤ 48 $g/m^{2.7}$ for men). We excluded from the analysis the subjects with mildly abnormal LVM (LVM 45-51 $g/m^{2.7}$ in women and 49-55 $g/m^{2.7}$ in men). The number of subjects in each group according to their LVM values, and their clinical parameters are shown in table 4. Table 5 shows the characteristics of the two groups (cases and controls). To assess statistical difference in mean $LVMh^{2.7}$ and in the other clinical characteristics between

cases and controls, analysis of variance (ANOVA) was performed, as implemented in StataSE; p-values of comparison are reported in table 5. Case and control subjects were similar for serum creatinine, uNa24h, HR, glycaemia, and uK24h. On the contrary, SBP, DBP, and BMI were significantly higher in cases than in controls (p-value < 0.0001); sPRA was slightly higher in cases than in controls (p-value < 0.05). Age also was slightly different between cases and controls although no evidence exists that a difference in age could affect the evaluation of cardiac mass.

4.3 RESULTS OF QUANTITATIVE TRAIT ASSOCIATION ANALYSIS

A linear regression analysis was performed on $LVMh^{2.7}$, adjusting for the covariates mentioned in *paragraph 3.7*. This analysis had the aim to identify the loci in the genome significantly associated with LVM phenotype.

Although no SNPs achieved genome wide significance for association with LVM (p-values < $5x10^{-8}$), we considered SNPs with p-values ≤ 10^{-5} as suggestive. The choice of this p-value threshold was supported by the q-q plot (Figure 4). The plot displays deviation from the null distribution only in the upper tail, which corresponds to SNPs with the stronger evidence of association. SNPs p-value deviated above the distribution reference line, at a level ≤ 10^{-5}. In order to exclude redundant findings, we filtered out SNPs that were in linkage disequilibrium with each other (r^2 ≥ 0.8); we also excluded SNPs mapping in desert regions. According to the threshold and to these filters, we selected 85 SNPs (table A2, figure 5). Among these, we identified 14 SNPs lying in genes previously associated to LVM (table 6): rs12369523 in Kinase Suppressor of Ras 2 (*KSR2*) gene; rs35996865 in Rho-Associated Coiled-Coil Containing Protein Kinase 1 (*ROCK1*) gene; rs78633628 in WW Domain Containing Oxidoreductase (*WWOX*) gene; rs17068332 and rs76156580 in CUB and Sushi multiple domains protein 1

(*CSMD1*) gene; rs183544012 in Calcium Channel Voltage-Dependent L Type Alpha 1D Subunit (*CACNA1D*) gene; rs10863888 in TNF Receptor-Associated Factor 5 (*TRAF5*) gene; rs6590636 in Contactin 5 (CNTN5) gene; rs79910493 in Insulin-Like Growth Factor (IGF1) gene; rs13023211 in Fidgetin (*FIGN*) gene; rs9284436 in Hyperpolarization Activated Cyclic Nucleotide Gated Potassium Channel 2 (HCN2) gene; rs76472108 in Regulator Of G-Protein Signaling 7 (*RGS7*) gene; rs2288696 in Fibroblast Growth Factor Receptor 1 (*FGFR1*) gene; and rs7137607 in SRY sex determining region Y-box 5 (*SOX5*) gene. Figure 6 shows regional plots for the described genes.

Figure 7 shows LVM average values according to genotypes of the most significant SNPs. The derivative alleles at some of the significant SNPs are very rare in the European ancestry population (1000 Genomes Phase 3 data, http://browser.1000genomes.org/index.html) and in our cohort. They are allele C at rs78633628 (*WWOX* gene), T at rs183544012 (*CACNA1D* gene), T at rs79910493 (*IGF1* gene), C at rs76472108 (*RGS7* gene), and A at rs76156580 (*CSMD1* gene). Due to the low frequency, the homozygous genotypes for the rare alleles were very rare as well, and were not present in our cohort that has a limited size. This justifies the very high beta effect linked to these alleles, as shown in table 6.

The identified polymorphisms map in intronic regions of the mentioned genes, except for rs35996865, that maps 532 bases upstream of the ROCK1 gene. We investigated if the variant maps to the promoter region of ROCK1, through Variant Effect Predictor tool of Ensembl database (http://grch37.ensembl.org/Homo_sapiens/Tools/VEP) [83], TRANSFAC® database [84] and UCSC database (http://genome.ucsc.edu/) [85]. According to all the databases, the variant lies in the promoter region of the gene (figure 8), that spans ≈ 1200 bp upstream the transcription initiation

site. The database analysis suggests several possible cis DNA elements such as AP-1, Sp1, and Oct-1 in the promoter region [86].

4.4 RESULTS OF CASE-CONTROL ASSOCIATION ANALYSIS

The aim of the case-control analysis was to identify the risk alleles associated with increased LVM. We performed a logistic regression analysis on the 14 most interesting results of the quantitative trait analysis (table 6). The logistic regression was performed adjusting for ancestry PCs, sex, SBP, DBP, serum creatinine, uNa24h, sPRA, BMI, HR, and echocardiography operator.

All SNPs associated with LVM in the linear regression showed a significant association with LVM (p-value < Bonferroni's threshold $3.6x10^{-3}$), except for rs79910493 and rs76472108, showing that the carriers of the risk alleles have an odds ratio higher than 1 to have increased LVM, i.e. to be cases respect to controls (table 7).

As for the quantitative analysis some odds ratios are very high, because of the absence of individuals with homozygous genotype of minor alleles. For WWOX variant the OR was not calculable because no cases with homozygous genotype of the C allele exist: cases are all carriers of the homozygous risk allele genotype TT.

4.5 WEIGHTED GENETIC RISK SCORE

We defined a weighted genetic risk score using the effect size of the risk allele (beta value of the linear regression analysis) to account for the strength of the genetic association with each allele. We calculated the wGRS including alleles of seven genes (ROCK1, IGF1, CACNA1D, FGFR1, TRAF5, SOX5, and KSR2) that are chosen for their functional role

in cardiac hypertrophy pathway or LVM determination, supported by data from the literature. The weight of each risk allele was calculated as function of its genotypic effect size (beta) from the linear GWA analysis (table 8).

4.5.1 Distribution of wGRS in cases and controls

The distribution and mean of the wGRS was plotted separately for cases and controls (figure 9) and compared using a two-sample t-test. We defined as case a subject with LVM ≥ 52 g/m$^{2.7}$ for women and ≥ 56 g/m$^{2.7}$ for men, and as control a subjects with LVM in normal range (≤ 44 g/m$^{2.7}$ for women and ≤ 48 g/m$^{2.7}$ for men). Using this model, cases had mean genetic risk score of 0.99±0.26, while controls 0.77±0.24 (two sample t test p-value < 0.00001) (table 9).

4.5.2 Partitioned wGRS

To describe the difference in the distribution of wGRS between cases and controls we partitioned the subjects by defining risk categories; these categories are created using the mean and standard deviation (SD) from the control subjects. The seven categories were defined as ± 0.25, ± 0.75, and ± 1.25 SDs from the mean, with the extreme categories being < 1.25 or > 1.25 SDs from the mean. We used as reference the category 4, which contains the mean of the control population. This category 4 approximated the group of subjects with an average risk of LVM increase. Six subsequent categories (1-3 and 5-7) were defined by the subjects found in increasing intervals of wGRS (table 10, figure 10). The major percentage of control subjects (≈ 19.4%) was in category 3, while the majority of the cases (≈ 35.44%) were located in the risk category 7, showing a higher genetic susceptibility risk for LVH.

Within the case-control dataset, we fitted a single logistic regression analysis, to study the association of wGRS with risk of developing LVH.

Thus we calculated the odds ratio (OR) for each category, taking the category 4 as reference (table 11). We could not consider results of the statistics as reliable on the categories 1, 2, and 3, as their p-value was not significant, whereas subjects in categories 5, 6 and 7 (i.e. those with the highest wGRS) had 2.22, 4.30 and 5.34 times increased odds of developing LVH, respectively, compared with subjects in category 4. The LVM distribution according to risk categories is reported in figure 11 (beta 1.78, p-value 5.03×10^{-27}).

Table 2 Characteristics of the subjects analyzed.

Characteristics	Total sample (n=966)
Sex, M/F	633/333
Age, years	47.8±9.0
$LVMh^{2.7}$, $g/m^{2.7}$	44.5±10.0
SBP, mmHg	154.8±12.4
DBP, mmHg	100.1±8.0
Heart rate, b.p.m.	75.4±10.0
Urine sodium, mEq/24 h	148.0±51.0
Urine potassium, mEq/24 h	56.5±20.3
sPRA, ng/mL/h	1.4±1.1
Glycaemia, mg/dl	89.8±11.0
Serum creatinine, mg/dl	0.9±0.2
BMI, kg/m^2	27.2±3.6

Values are reported as mean±standard deviation. Glossary: $LVMh^{2.7}$, left ventricular mass indexed for $height^{2.7}$; SBP, systolic blood pressure; DBP, diastolic blood pressure; b.p.m, beats per minute; sPRA, sitting plasma renin activity; BMI, body mass index.

Table 3 $LVMh^{2.7}$ distribution according to gender.

Sample	N	mean $LVMh^{2.7}$	SD	min $LVMh^{2.7}$	max $LVMh^{2.7}$
All	966	44.5	10.0	20.0	82.4
Males	633	45.4	10.0	20.5	82.4
Females	333	42.7	9.8	20.0	80.7

$LVMh^{2.7}$ was evaluated as $g/m^{2.7}$.

Table 4 Reference limits of LVM according to gender (from Lang RM *et al* [18]) and characteristics of analyzed subjects for each class.

	Women			
	Reference range	Mildly abnormal	Moderately abnormal	Severely abnormal
LVM, $g/m^{2.7}$	18-44	45-51	52-58	≥ 59
N subjects	198	74	43	18
Age, years	48.3±8.9	47.3±8.8	48.9±7.9	50.4±6.2
LVM, $g/m^{2.7}$	36.2±5.1	47.6±2.3	55.2±2.1	65.0±5.9
SBP, mmHg	154.1±11.6	159.0±15.7	161.6±15.5	159.4±15.1
DBP, mmHg	98.5±8.2	101.7±6.8	100.9±6.1	101.4±6.5
HR, b.p.m.	75.3±9.6	75.3±10.3	76.9±10.2	73.7±9.8
uNa, mEq/24 h	136.6±50.4	142.6±49.3	147.6±36.0	124.2±38.2
uK, mEq/24 h	52.4±15.8	53.1±15.4	54.3±10.1	56.7±18.4
sPRA, ng/mL/h	1.3±1.2	1.9±1.3	1.6±1.0	1.7±0.9
Glycemia, mg/dl	87.8±9.9	90.2±12.4	87.6±10.2	87.8±8.1
Serum crea, mg/dl	0.83±0.15	0.88±0.16	0.82±0.15	0.84±0.24
BMI, kg/m^2	25.4±3.6	28.1±4.7	26.8±3.5	28.6±4.0

	Men			
	Reference range	Mildly abnormal	Moderately abnormal	Severely abnormal
LVM, g/m$^{2.7}$	20-48	49-55	56-63	≥64
N subjects	417	119	67	30
Age, years	47.3±9.3	47.5±9.3	47.9±7.7	52.2±9.2
LVM, g/m$^{2.7}$	39.8±5.8	51.5±2.2	59.3±2.1	69.1±4.9
SBP, mmHg	152.6±10.1	155.0±12.9	156.1±12.8	164.4±16.6
DBP, mmHg	99.0±7.4	102.4±8.9	102.5±8.4	105.6±9.4
HR, b.p.m.	75.7±9.8	76.2±9.2	71.2±10.3	75.9±10.8
uNa, mEq/24 h	150.3±52.7	162.7±55.1	154.5±46.3	146.6±32.0
uK, mEq/24 h	59.6±24.2	57.2±19.7	57.4±17.4	49.4±12.3
sPRA, ng/mL/h	1.3±1.1	1.5±1.3	1.5±1.0	1.7±1.0
Glycemia, mg/dl	90.5±11.3	91.5±11.1	90.7±11.5	87.8±6.4
Serum crea, mg/dl	0.95±0.16	0.97±0.16	0.97±0.14	0.97±0.12
BMI, kg/m^2	27.4±3.0	27.8±3.5	27.9±2.6	30.7±3.6

Values are reported as mean±standard deviation. Glossary: LVMh$^{2.7}$, left ventricular mass indexed for height$^{2.7}$; SBP, systolic blood pressure; DBP, diastolic blood pressure; HR, heart rate; b.p.m beats per minute; uNa, urine sodium; uK, urine potassium; sPRA, sitting plasma renin activity; crea, creatinine; BMI, body mass index.

Table 5 Characteristics of the subjects analyzed in cases and controls.

Characteristic	Cases (n=158)	Controls (n=615)	p-value
Sex, M/F	97/61	417/198	NS
Age, years	49.3 ± 8.0	47.6 ± 9.1	3.53E-02
$LVMh^{2.7}$, $g/m^{2.7}$	60.7 ± 6.0	38.6 ± 5.8	< 0.0001
SBP, mmHg	159.5 ± 14.8	153.1 ± 10.6	< 0.0001
DBP, mmHg	102.5 ± 8.0	99.0 ± 8.0	< 0.0001
Heart rate, b.p.m.	73.9 ± 10.5	75.6 ± 9.7	NS
Urine sodium, mEq/24 h	147.7 ± 41.0	145.9 ± 52.3	NS
Urine potassium, mEq/24 h	54.9 ± 15.1	57.2 ± 22.1	NS
sPRA, ng/mL/h	1.6 ± 1.0	1.3 ± 1.1	1.92E-03
Glycaemia, mg/dl	89.0 ± 10.0	89.6 ± 11.0	NS
Serum creatinine, mg/dl	0.9 ± 0.2	0.9 ± 0.2	NS
BMI, kg/m^2	28.2 ± 3.5	26.7 ± 3.4	< 0.0001

Values are reported as mean±standard deviation. P-values of comparison among groups are reported. Glossary: $LVMh^{2.7}$, left ventricular mass indexed for $height^{2.7}$; SBP, systolic blood pressure; DBP, diastolic blood pressure; b.p.m., beats per minute; sPRA, sitting plasma renin activity; BMI, body mass index; NS, non significative.

Table 6 Functional SNPs associated to left ventricular mass ($LVMh^{2.7}$) in linear regression.

SNP	Chr	Position (bp)	Gene	Location	Risk/other Alleles	Allele freq	Beta	SE	p-value
rs12369523	12	118309008	KSR2	intron	A/G	0.57	2.25	0.43	1.37E-07
rs35996865	18	18692344	ROCK1	upstream	G/T	0.3	2.33	0.46	3.17E-07
rs78633628	16	78971380	WWOX	intron	T/C	0.99	10.95	2.46	8.66E-06
rs17068332	8	3833181	CSMD1	intron	C/T	0.88	2.63	0.63	3.05E-05
rs183544012	3	53747902	CACNA1D	intron	T/C	0.02	6.46	1.57	3.93E-05
rs10863888	1	211502769	TRAF5	intron	G/A	0.59	1.71	0.42	4.23E-05
rs6590636	11	100047729	CNTN5	intron	A/C	0.44	1.71	0.42	4.30E-05
rs79910493	12	102843754	IGF1	intron	T/C	0.02	5.94	1.46	4.60E-05
rs13023211	2	164504320	FIGN	intron	A/G	0.86	2.55	0.63	4.85E-05
rs9284436	19	607108	HCN2	intron	C/T	0.43	1.75	0.44	5.89E-05
rs76472108	1	241304791	RGS7	intron	C/G	0.02	6.07	1.54	7.71E-05
rs76156580	8	4474130	CSMD1	intron	A/C	0.01	8.53	2.16	7.98E-05
rs2288696	8	38286225	FGFR1	intron	G/A	0.81	2.06	0.53	8.79E-05
rs7137607	12	23778584	SOX5	intron	A/C	0.55	1.61	0.41	9.63E-05

$LVMh^{2.7}$ association was evaluated using a linear regression analysis under an additive model, adjusted for ancestry PCs, sex, SBP, DBP, serum creatinine, uNa24h, sPRA, BMI, HR, and ECG operator. Allele frequency and Beta effect are referred to risk allele. To retrieve information about SNPs and their genomic context (the nearest gene) we used the hg19 assembly (National Center for Biotechnology Information 37). Glossary: SNP, Chr, chromosome; bp, base pair; and SE, standard error.

Table 7 Functional SNPs associated to left ventricular mass ($LVMh^{2.7}$) in case-control analysis.

SNP	Chr	Position (bp)	Gene	Location	Risk/other Alleles	Allele freq	OR	p-value	95% CI
rs12369523	12	118309008	KSR2	intron	A/G	0.57	1.7	3.62E-04	1.27-2.27
rs35996865	18	18692344	ROCK1	upstream	G/T	0.3	1.7	5.02E-04	1.26-2.30
rs78633628	16	78971380	WWOX	intron	T/C	0.99	/*	/	/
rs17068332	8	3833181	CSMD1	intron	C/T	0.88	2.46	2.62E-04	1.52-3.99
rs183544012	3	53747902	CACNA1D	intron	T/C	0.02	3.92	9.92E-04	1.74-8.83
rs10863888	1	211502769	TRAF5	intron	G/A	0.59	1.99	8.34E-06	1.47-2.69
rs6590636	11	100047729	CNTN5	intron	A/C	0.44	1.81	3.84E-05	1.37-2.41
rs79910493	12	102843754	IGF1	intron	T/C	0.02	2.77	1.10E-02	1.27-6.06
rs13023211	2	164504320	FIGN	intron	A/G	0.86	1.86	7.00E-03	1.18-2.94
rs9284436	19	607108	HCN2	intron	C/T	0.43	1.62	1.00E-03	1.22-2.17
rs76472108	1	241304791	RGS7	intron	C/G	0.02	3.72	1.10E-02	1.68-8.24
rs76156580	8	4474130	CSMD1	intron	A/C	0.01	4.62	5.00E-03	1.58-13.54
rs2288696	8	38286225	FGFR1	intron	G/A	0.81	1.73	5.00E-03	1.18-2.54
rs7137607	12	23778584	SOX5	intron	A/C	0.55	1.69	3.23E-04	1.27-2.26

$LVMh^{2.7}$ association was evaluated using a logistic regression analysis under an additive model, adjusted for ancestry PCs, sex, SBP, DBP, serum creatinine, uNa24h, sPRA, BMI, HR, and echocardiography operator. Allele frequency and Beta effect are referred to risk allele. Glossary: SNP, single nucleotide polymorphism; Chr, chromosome; OR odds ratio; and 95% CI, 95% confidence intervals. *not calculable (cases with homozygous genotype of the C allele are missing)

Table 8 SNPs that compose the weighed genetic risk score and weights assigned to each marker.

Chr	SNP	Risk Allele	Gene	Beta effect (weight)
12	rs12369523	A	KSR2	2.25
18	rs35996865	G	ROCK1	2.33
3	rs183544012	T	CACNA1D	6.46
1	rs10863888	G	TRAF5	1.71
12	rs79910493	T	IGF1	5.94
8	rs2288696	G	FGFR1	2.06
12	rs7137607	A	SOX5	1.61

Glossary: Chr, chromosome; and SNP, single nucleotide polymorphism.

Table 9 Left ventricular mass mean values and weighted genetic risk score distribution in cases and controls.

	Mean LVM ($g/m^{2.7}$)	SD	number	wGRS	SD	95% CI	
Controls	38.63	5.83	615	0.77	0.24	0.75	0.79
Cases	60.71	6.01	158	0.99	0.26	0.95	1.03
Total	43.15	10.66	773	0.82	0.26	0.80	0.84

Table 10 Distribution of subjects in the risk categories defined by the weighed genetic risk score.

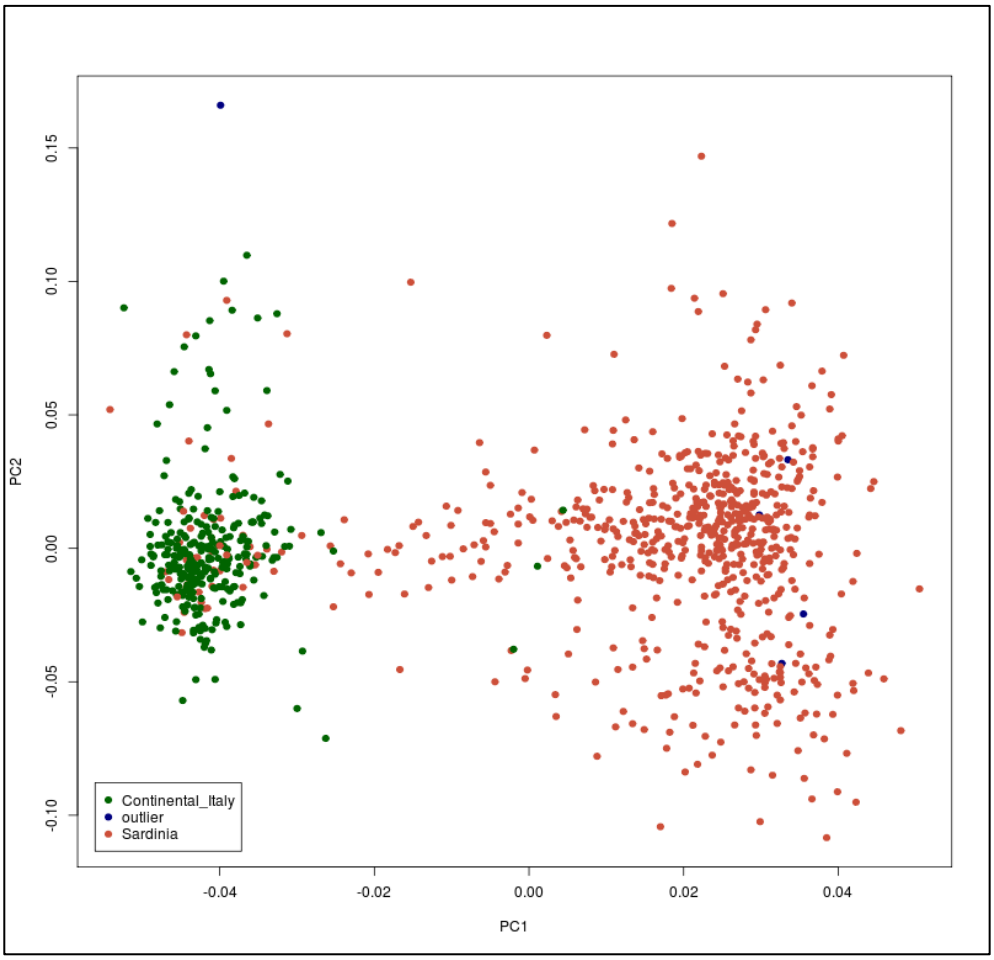
Categories	Formula	Score	N (%) controls	N (%) cases
1	mean -1.25 SD	≤ 0.470	61 (9.92)	2 (1.27)
2	mean -0.75 SD	0.470-0.591	101 (16.42)	6 (3.80)
3	mean -0.25 SD	0.59-0.713	122 (19.84)	20 (12.66)
4 (Ref)	*mean*	*0.714-0.834*	*116 (18.86)*	*16 (10.13)*
5	mean +0.25 SD	0.83-0.955	85 (13.82)	26 (16.46)
6	mean +0.75 SD	0.956-1.077	54 (8.78)	32 (20.25)
7	mean +1.25 SD	> 1.077	76 (12.36)	56 (35.44)
TOT			615 (79.56)	158 (20.44)

In the table, for each category, is reported the number and percentage of controls and cases. The frequency is calculated according to the total number of individuals for each status.

Table 11 Weighed genetic risk score scores and odds ratios of left ventricular mass in each risk category.

wGRS category	Controls	Cases	OR	p-value
1	61	2	0.24	0.06
2	101	6	0.43	0.09
3	122	20	1.19	0.63
4	*116*	*16*	*1 (reference)*	-
5	85	26	2.22	0.022
6	54	32	4.30	< 0.0001
7	76	56	5.34	< 0.0001

Figure 1 Principal component plot.



Samples are marked according to ancestry cluster based on genotyping data: continental Italy samples in green, Sardinian samples in red and outliers in blue; PC principal component.

Figure 2 Subjects' flow from recruitment to pre-analysis quality control steps.

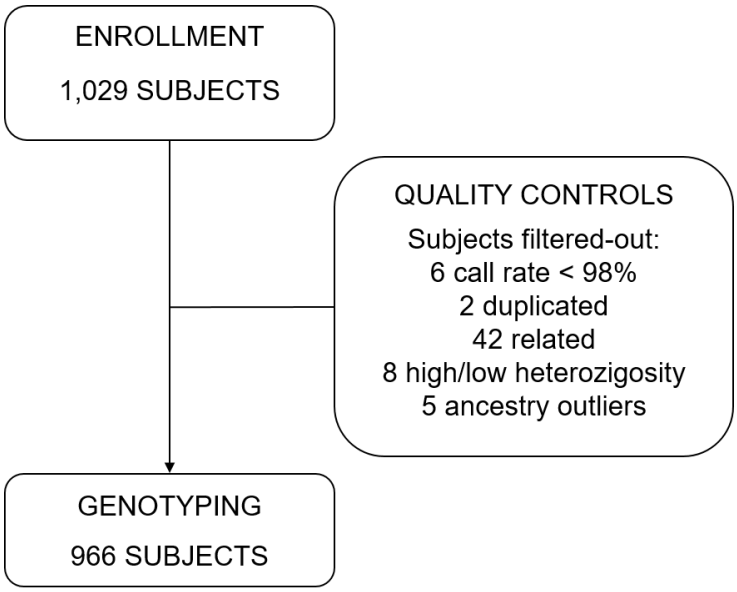


Figure 3 Histogram of left ventricular mass ($LVMh^{2.7}$) distribution in the analyzed sample (normal curve superimposed); $LVMh^{2.7}$ was expressed in $g/h^{2.7}$.

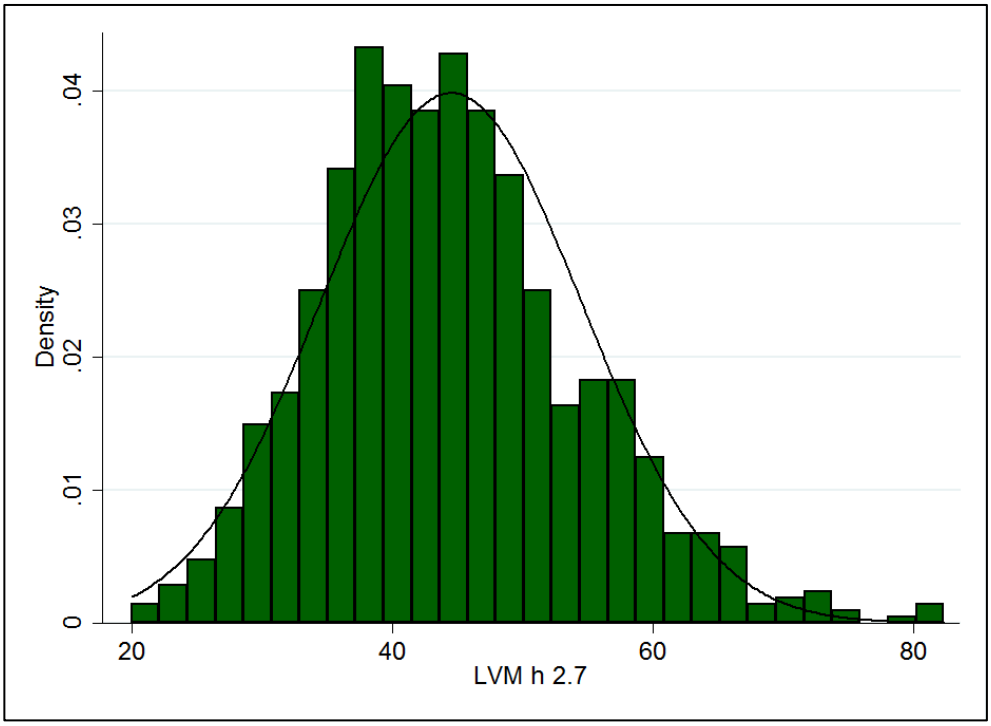
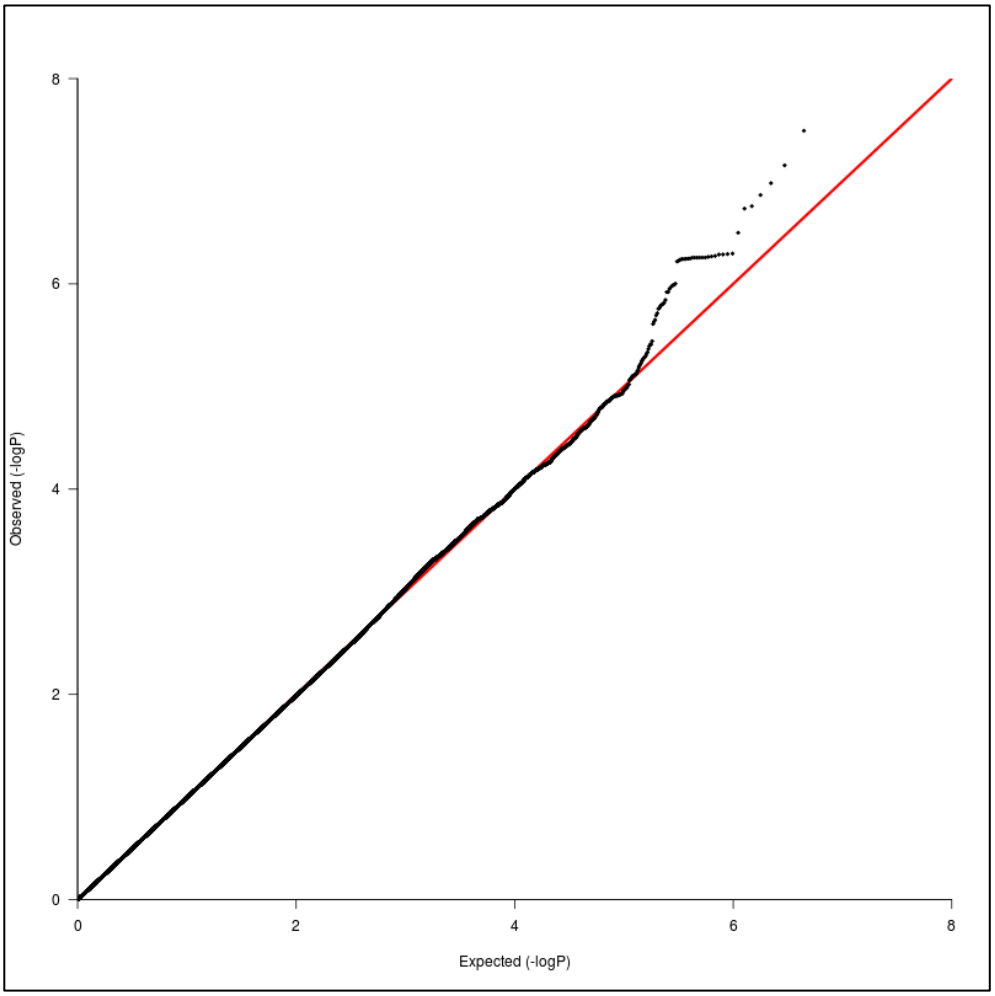
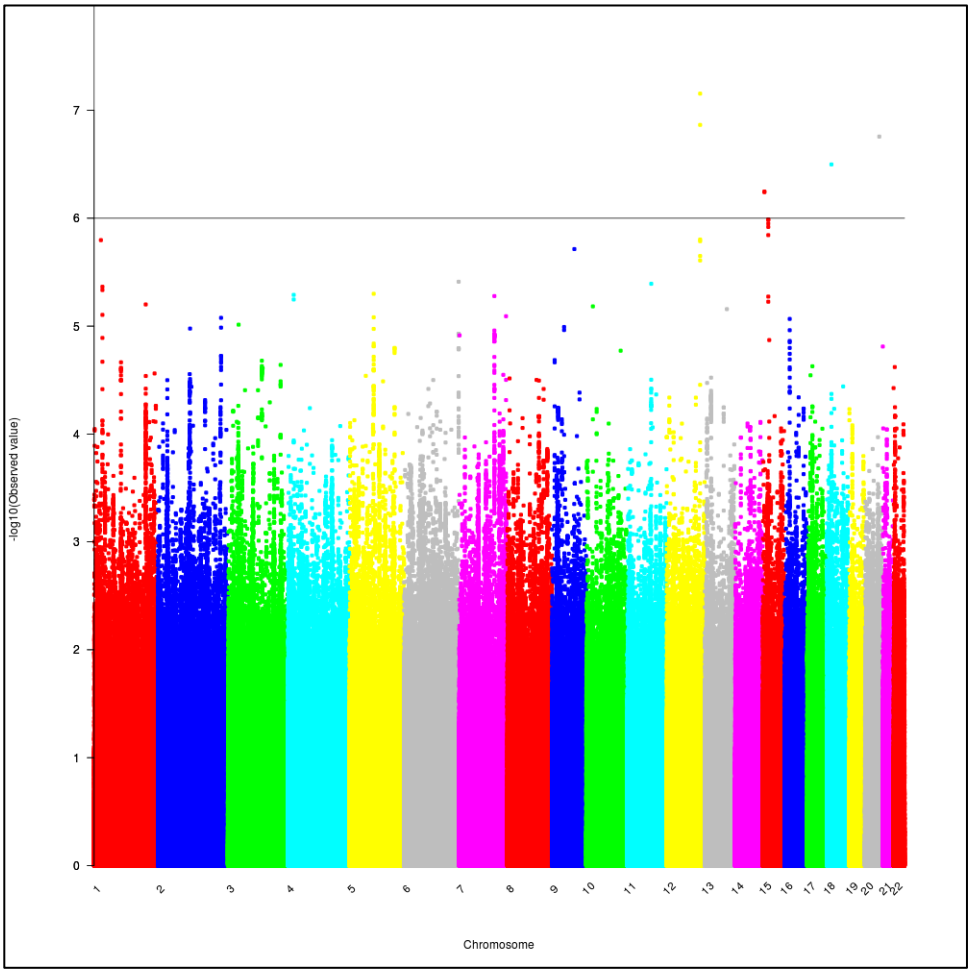


Figure 4 Quantile-quantile plot of single nucleotide polymorphism p-values from genome wide association analysis of left ventricular mass ($LVMh^{2.7}$).



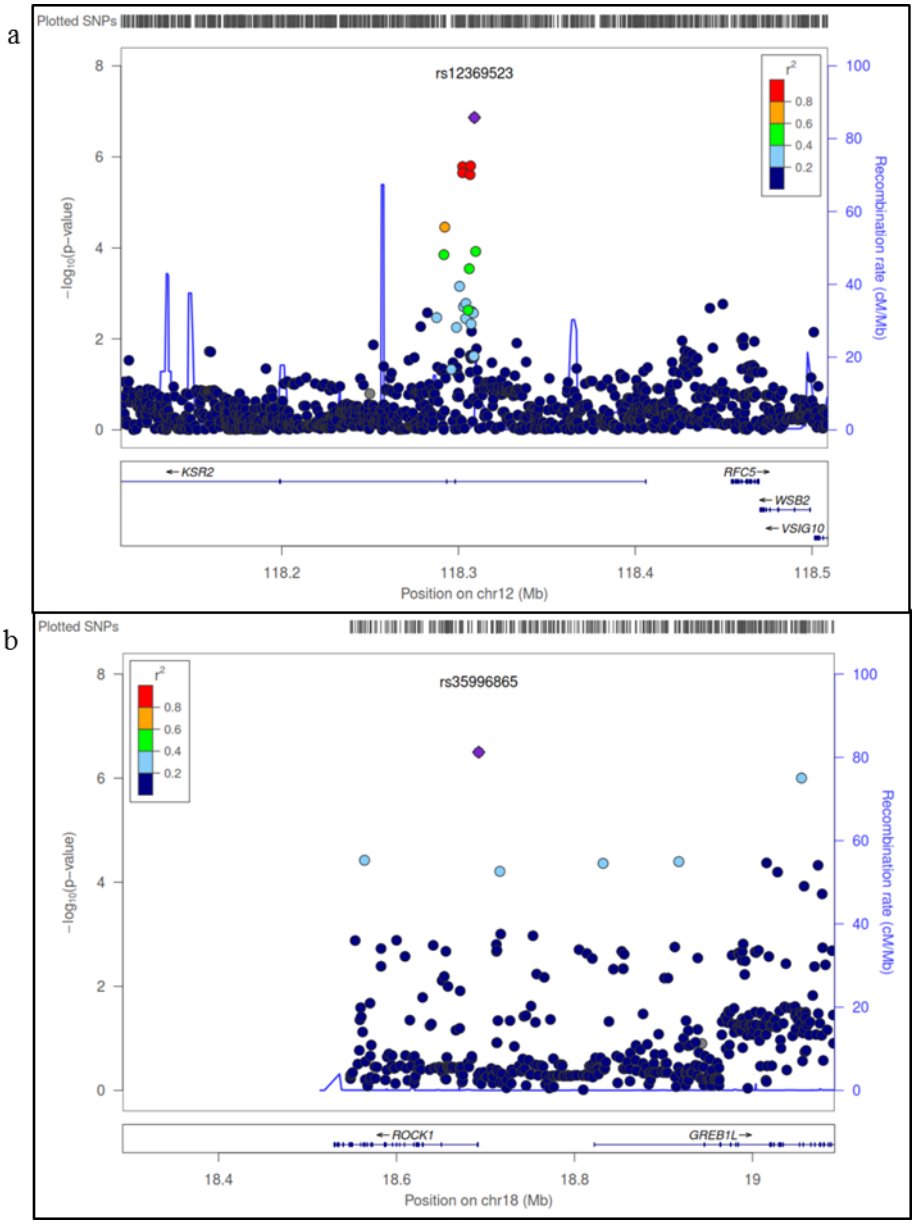
The red line indicates the middle of the first and third quartile of the expected distribution of test statistics; the dashed line marks the 95% confidence interval of the expected distribution of the test statistics.

Figure 5 Manhattan plot of single SNP linear regression test statistics for $LVMh^{2.7}$.

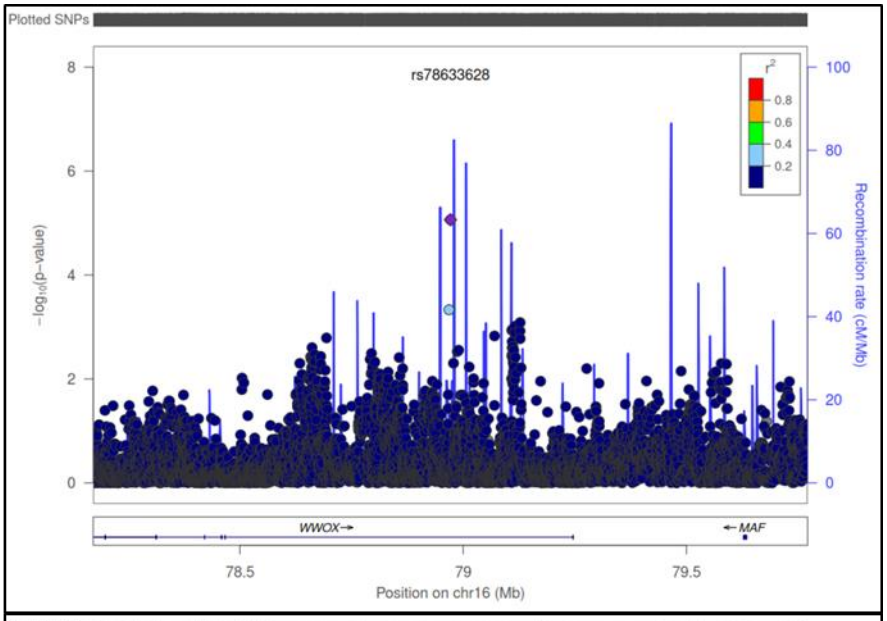


Regression analysis was adjusted for gender, systolic and diastolic blood pressure, serum creatinine, urinary sodium 24 h, sitting plasma renin activity, body mass index, heart rate, first ten principal components, and echocardiography operator. Markers were filtered for imputation quality (Rsq, threshold 0.8) and minor allele frequency (threshold 0.01). Results are reported as -log10 (p-value) by genomic position chromosomal location. Values for each chromosome are shown in different colors.

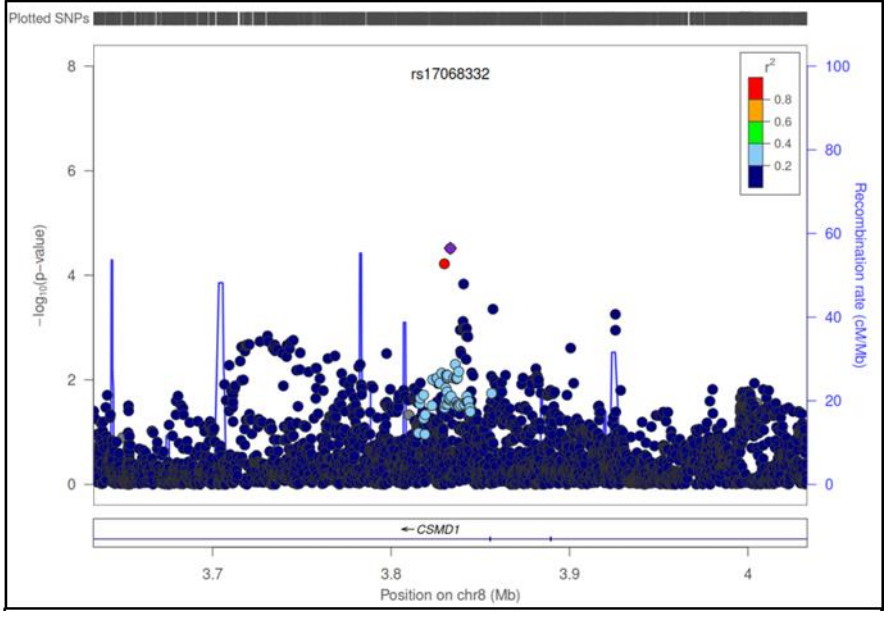
Figure 6 Local Manhattan plot for *KSR2* (a), *ROCK1* (b), *WWOX* (c), *CSMD1* (d), *CACNA1D* (e), *TRAF5* (f), *CNTN5* (g), *IGF1* (h), *FIGN* (i), *HCN2* (j), *RGS7* (k), *FGFR1* (l), and *SOX5* (m) genes.



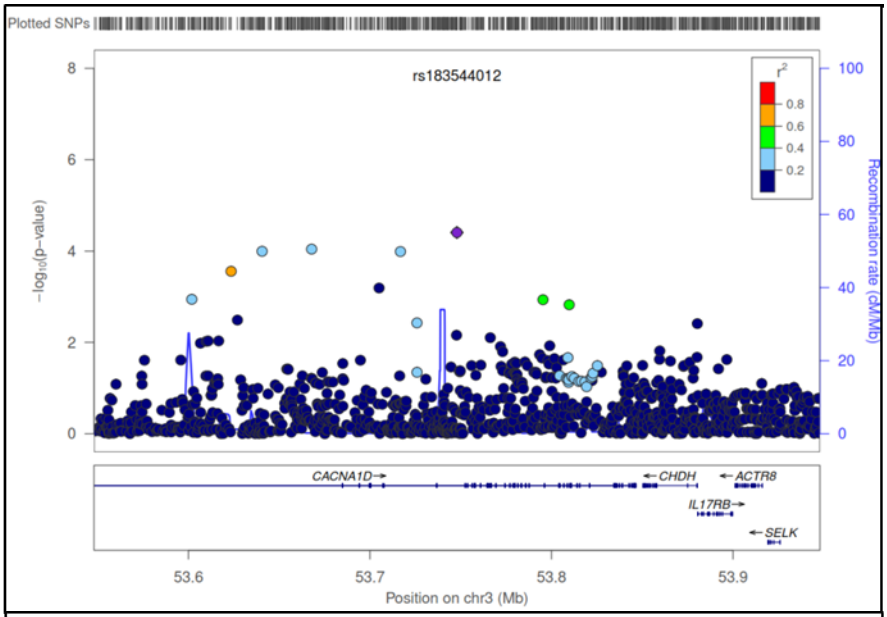
c


Plotted SNPs
rs78633628
r²
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
WWOX→
←MAF
78.5
79
79.5
Position on chr16 (Mb)

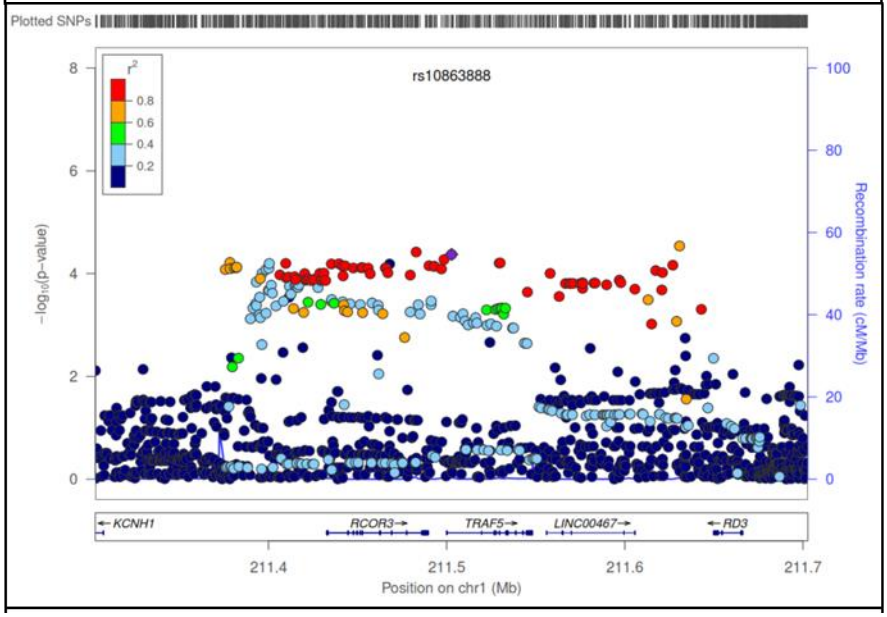
d


Plotted SNPs
rs17068332
r²
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
←CSMD1
3.7
3.8
3.9
4
Position on chr8 (Mb)

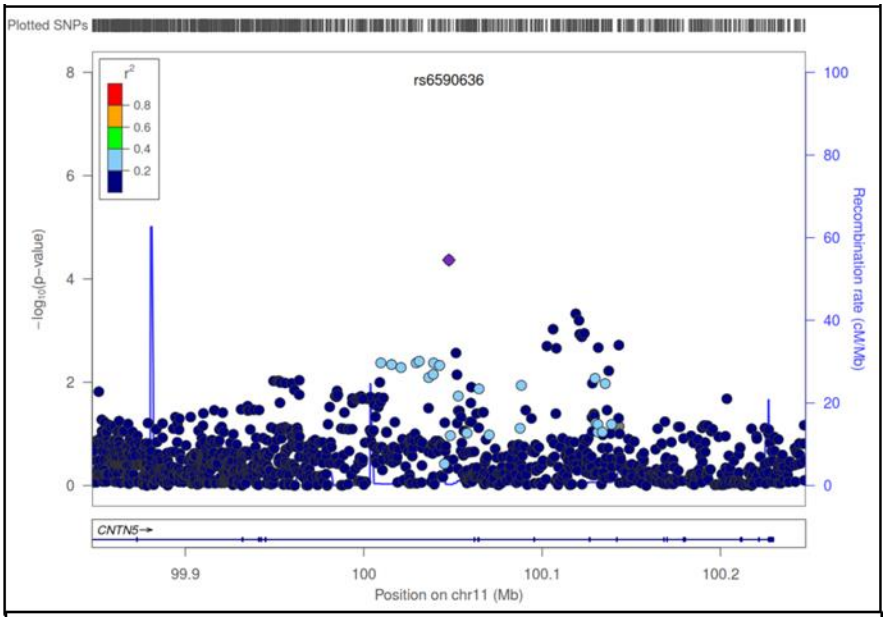
e


Plotted SNPs
rs183544012
r²
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
CACNA1D
CHDH
ACTR8
IL17RB
SELK
53.6
53.7
53.8
53.9
Position on chr3 (Mb)

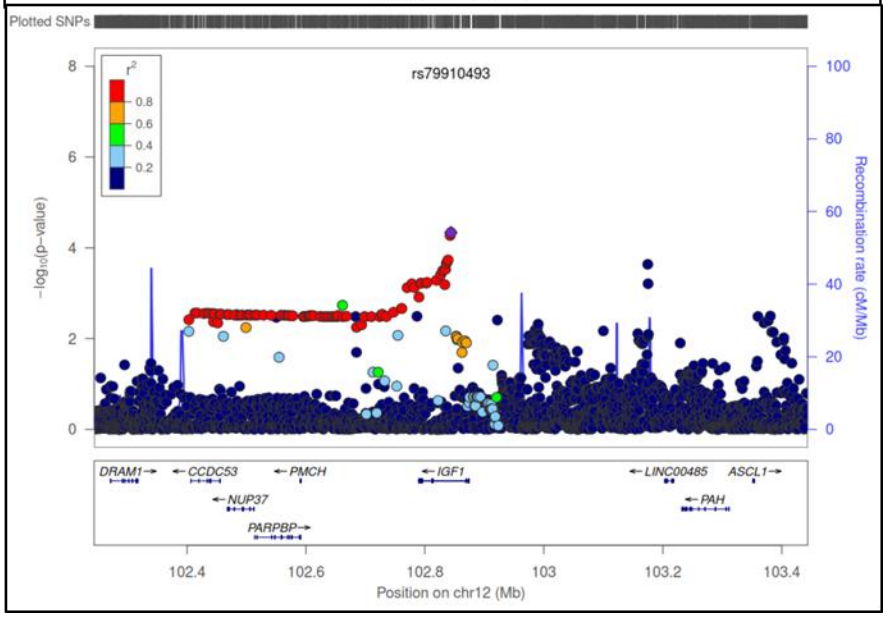
f


Plotted SNPs
rs10863888
r²
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
KCNH1
RCOR3
TRAF5
LINC00467
RD3
211.4
211.5
211.6
211.7
Position on chr1 (Mb)

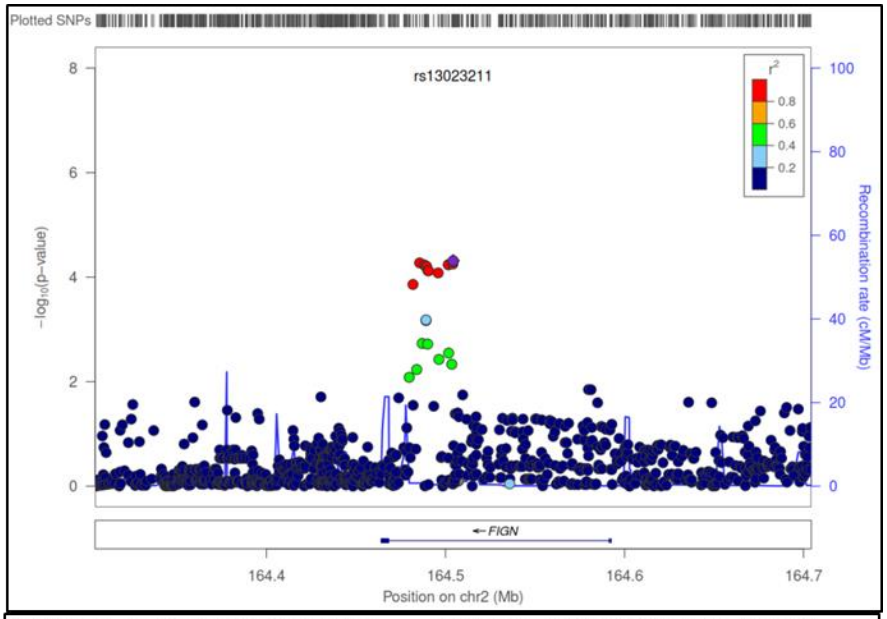
g


Plotted SNPs
rs6590636
r²
0.8
0.6
0.4
0.2
−log10(p-value)
Recombination rate (cM/Mb)
CNTN5
99.9
100
100.1
100.2
Position on chr11 (Mb)

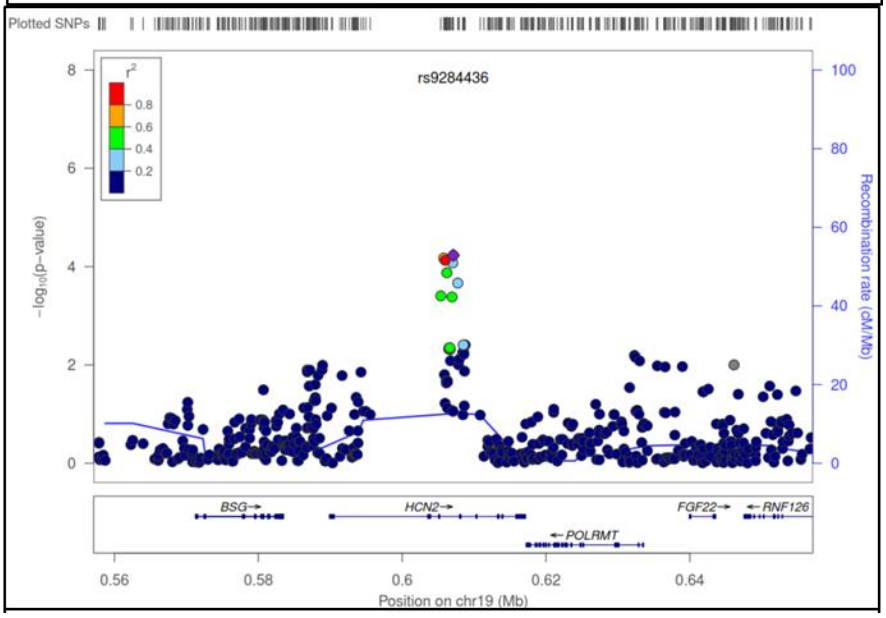
h


Plotted SNPs
rs79910493
r²
0.8
0.6
0.4
0.2
−log10(p-value)
Recombination rate (cM/Mb)
DRAM1
CCDC53
PMCH
IGF1
LINC00485
ASCL1
NUP37
PAH
PARPBP
102.4
102.6
102.8
103
103.2
103.4
Position on chr12 (Mb)

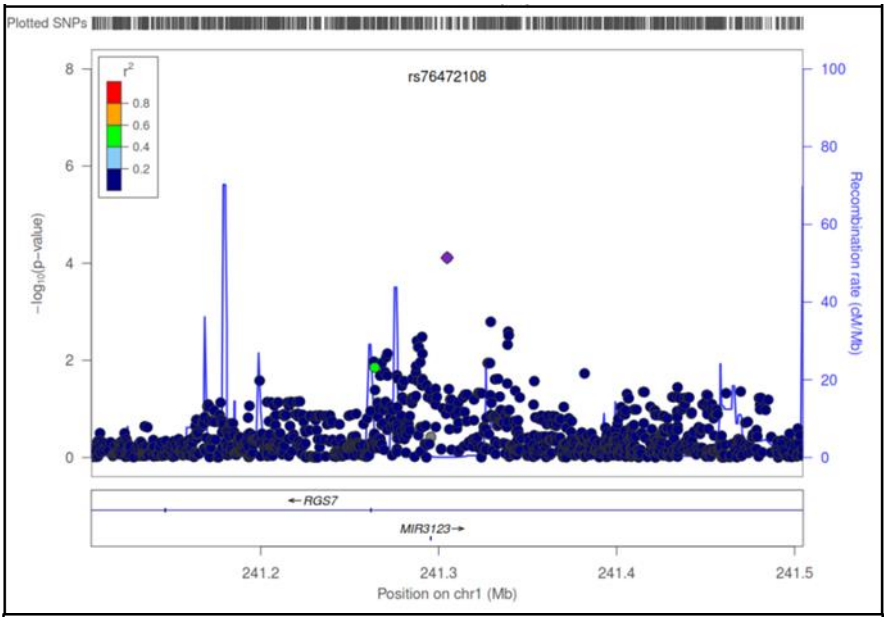
i


Plotted SNPs
rs13023211
r²
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
FIGN
164.4
164.5
164.6
164.7
Position on chr2 (Mb)

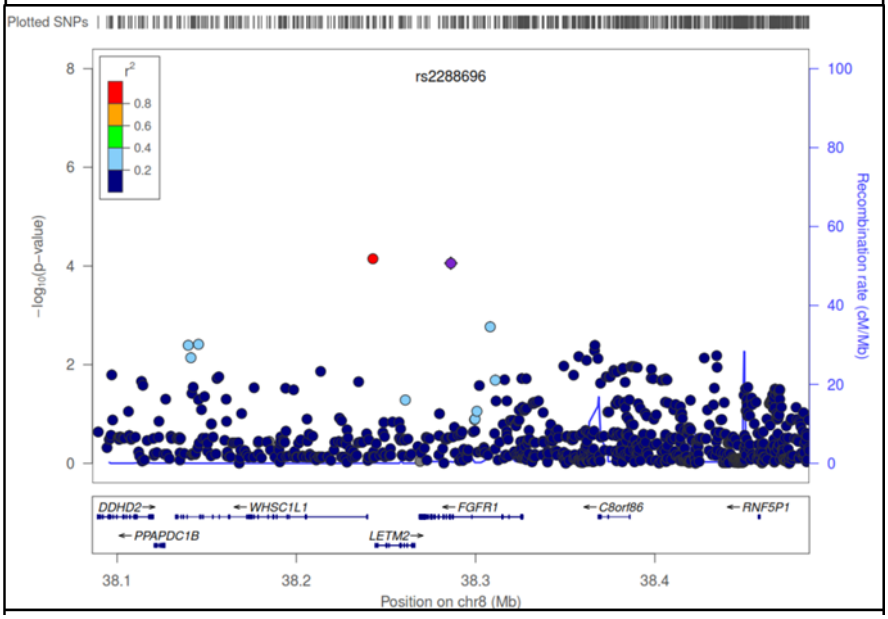
j


Plotted SNPs
rs9284436
r²
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
BSG
HCN2
FGF22
RNF126
POLRMT
0.56
0.58
0.6
0.62
0.64
Position on chr19 (Mb)

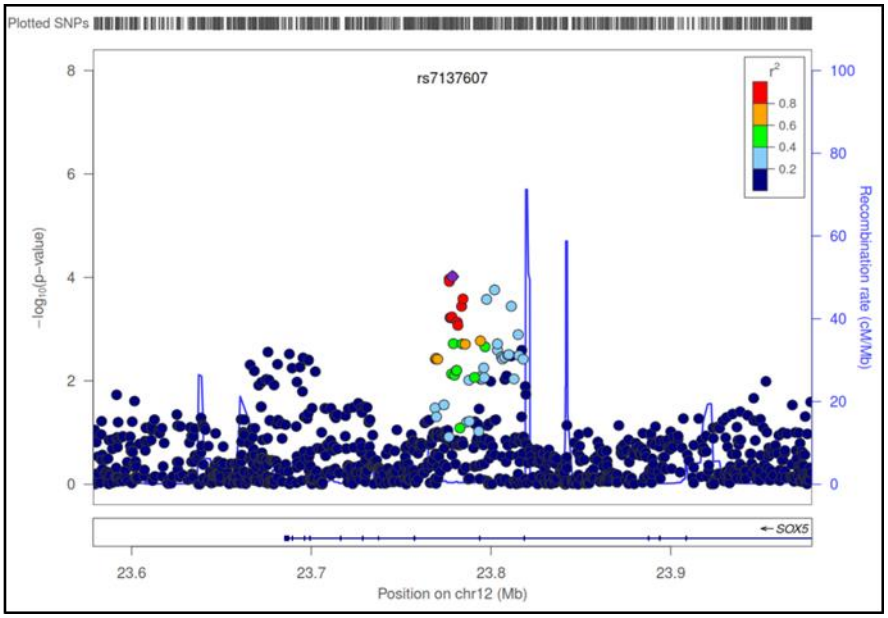
k


Plotted SNPs
rs76472108
r²
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
RGS7
MIR3123
241.2
241.3
241.4
241.5
Position on chr1 (Mb)

l

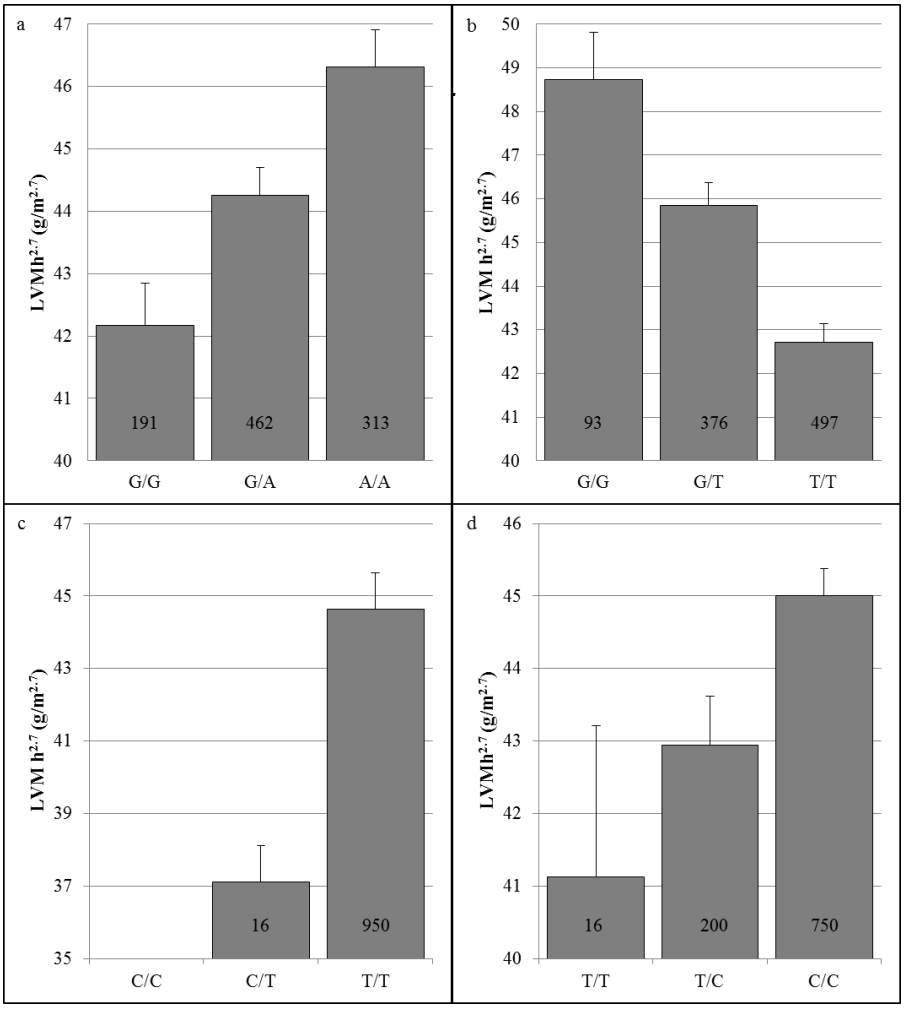

Plotted SNPs
rs2288696
r²
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
DDHD2
WHSC1L1
FGFR1
C8orf86
RNF5P1
PPAPDC1B
LETM2
38.1
38.2
38.3
38.4
Position on chr8 (Mb)

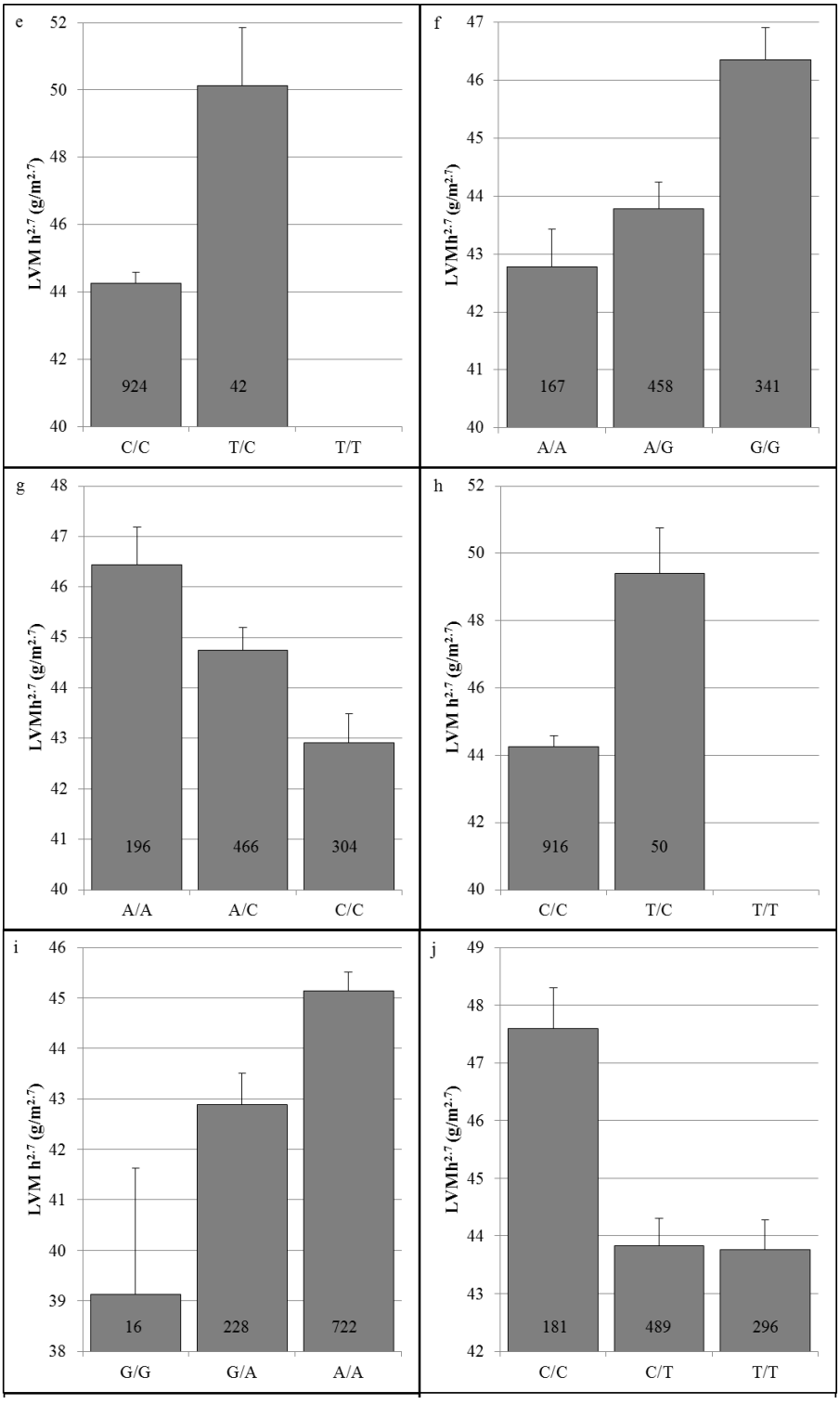
m

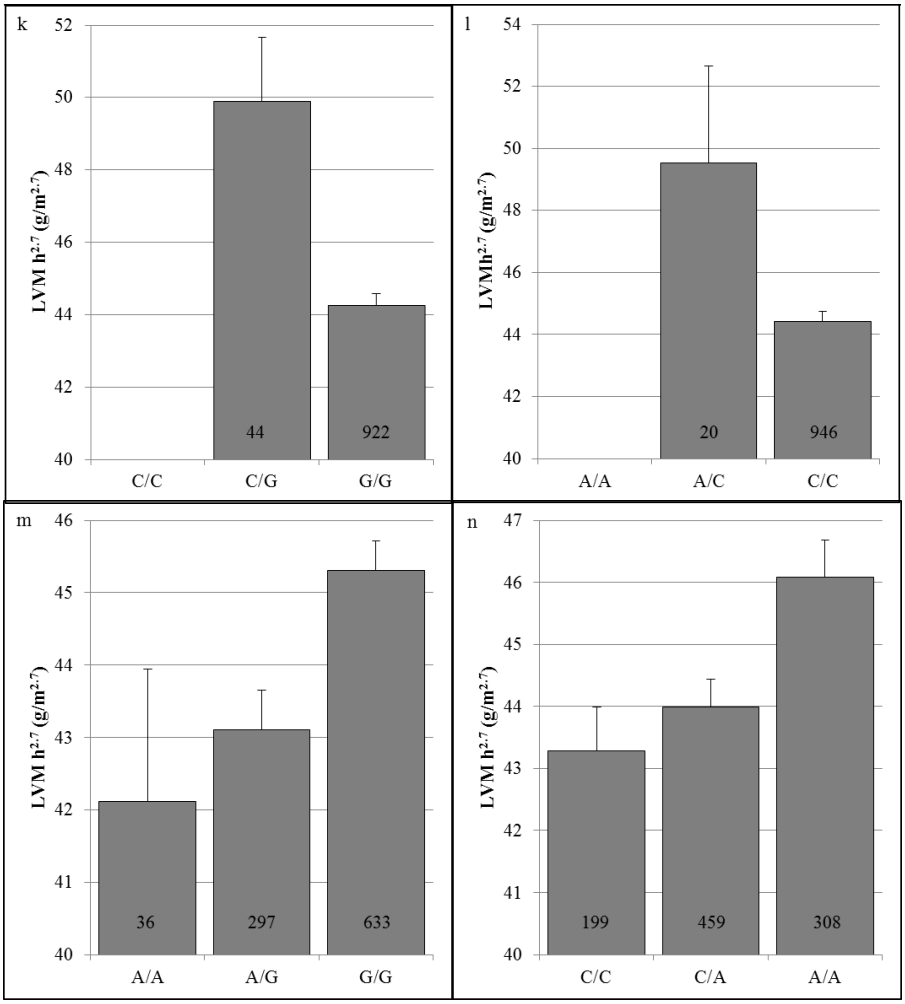


Each circle represents a single nucleotide polymorphism; y-axis is the -log10 LVM association p-value, and x-axis represents the physical position on the chromosome (build 37, hg 19). The circles are filled with colors according to the linkage disequilibrium (LD, r2) between the given SNPs and the lead SNP (purple square).

Figure 7 Plot of LVM average by *KSR2* rs12369523 (a), *ROCK1* rs35996865 (b), *WWOX* rs78633628 (c), *CSMD1* rs17068332 (d), *CACNA1D* rs183544012 (e), *TRAF5* rs10863888 (f), *CNTN5* rs6590636 (g), *IGF1* rs79910493 (h), *FIGN* rs13023211 (i), *HCN2* rs9284436 (j), *RGS7* rs76472108 (k), *CSMD1* rs76156580 (l), *FGFR1* rs2288696 (m), and *SOX5* rs7137607 (n) genotypes of participants.




e
LVM h2.7 (g/m2.7)
52
50
48
46
44
42
40
924
42
C/C
T/C
T/T
f
LVMh2.7 (g/m2.7)
47
46
45
44
43
42
41
40
167
458
341
A/A
A/G
G/G
g
LVMh2.7 (g/m2.7)
48
47
46
45
44
43
42
41
40
196
466
304
A/A
A/C
C/C
h
LVM h2.7 (g/m2.7)
52
50
48
46
44
42
40
916
50
C/C
T/C
T/T
i
LVM h2.7 (g/m2.7)
46
45
44
43
42
41
40
39
38
16
228
722
G/G
G/A
A/A
j
LVMh2.7 (g/m2.7)
49
48
47
46
45
44
43
42
181
489
296
C/C
C/T
T/T



In each box the number of subjects per genotype is indicated. P-values and Beta effects for each variant are reported in table 6.

Figure 8 Genomic region of *ROCK1* gene on chromosome 18 (a) and region in detail around rs35996865 variant (b) (http://browser.1000genomes.org/).

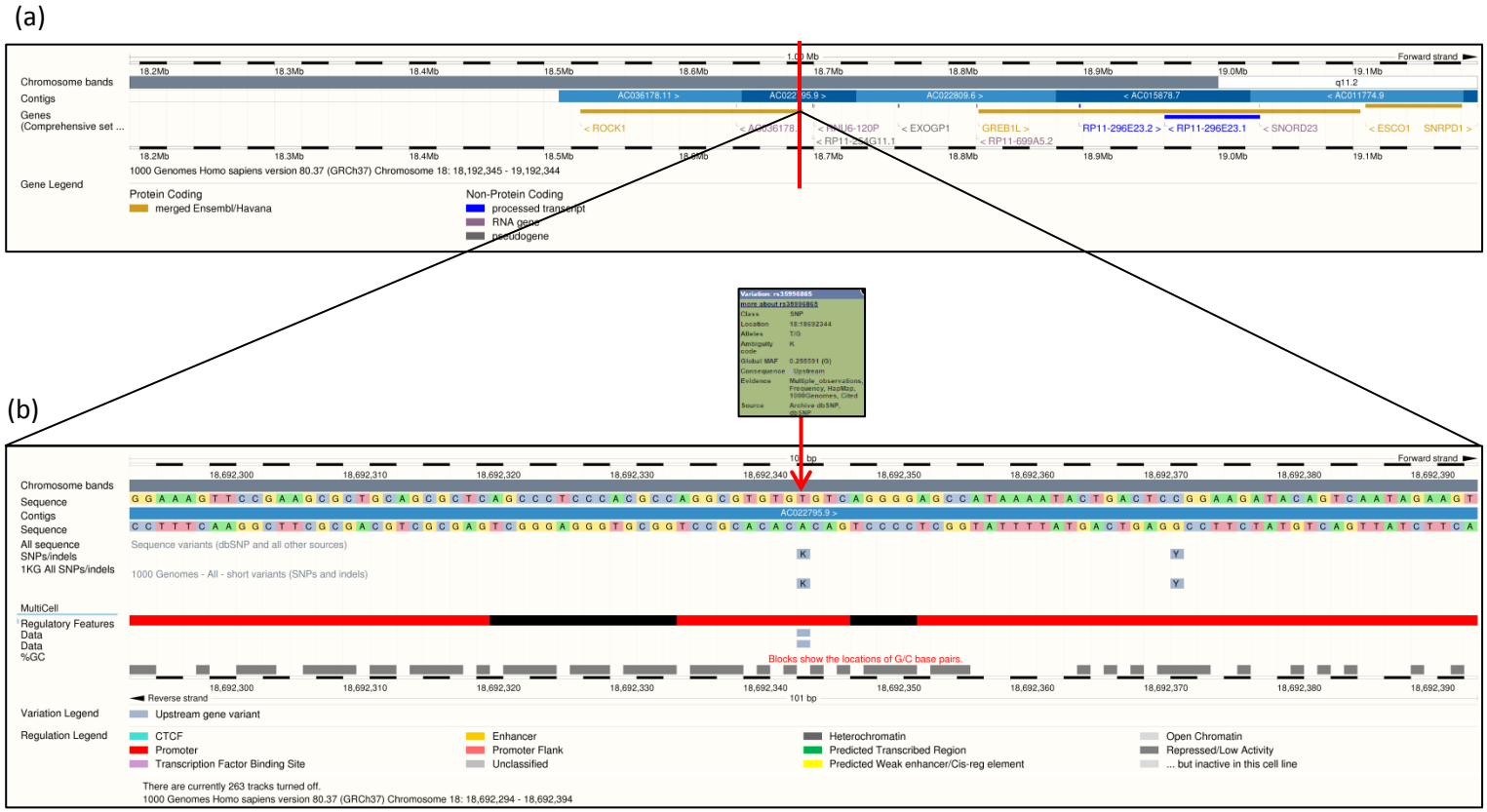


Figure 9 Box plot of wGRS average value and distribution in control and in case groups.

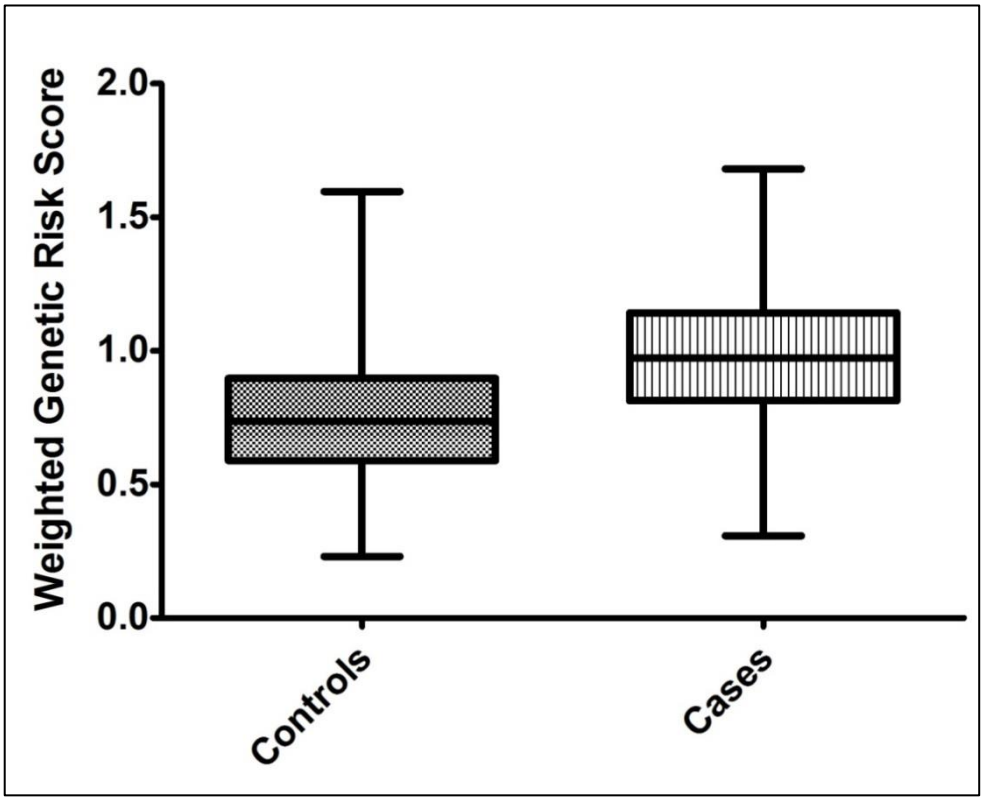
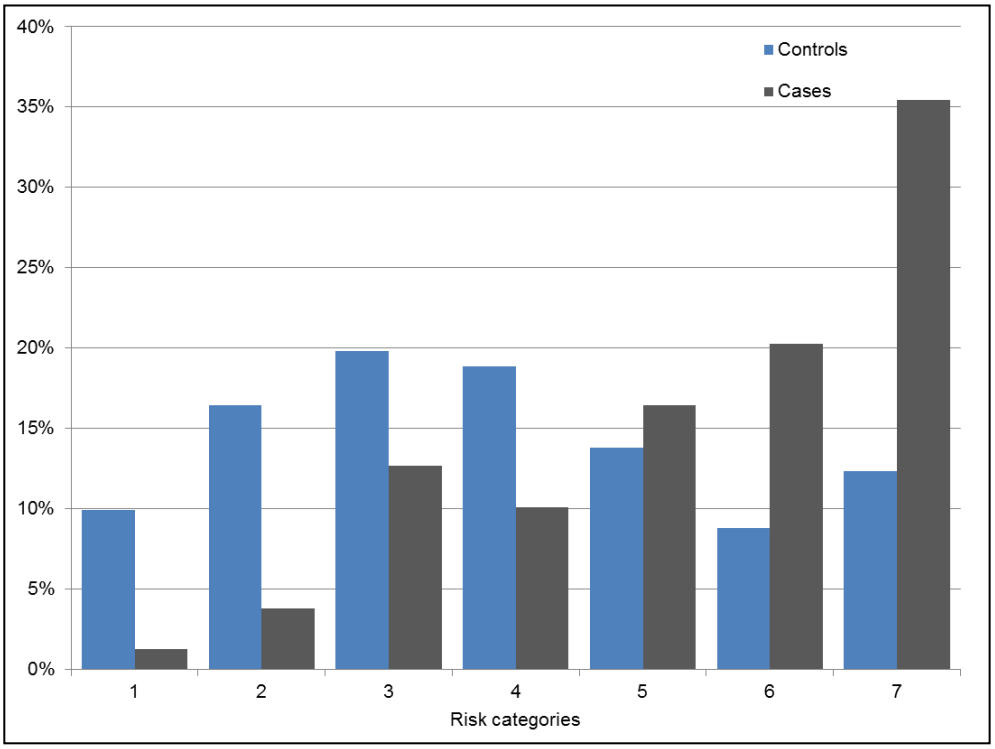
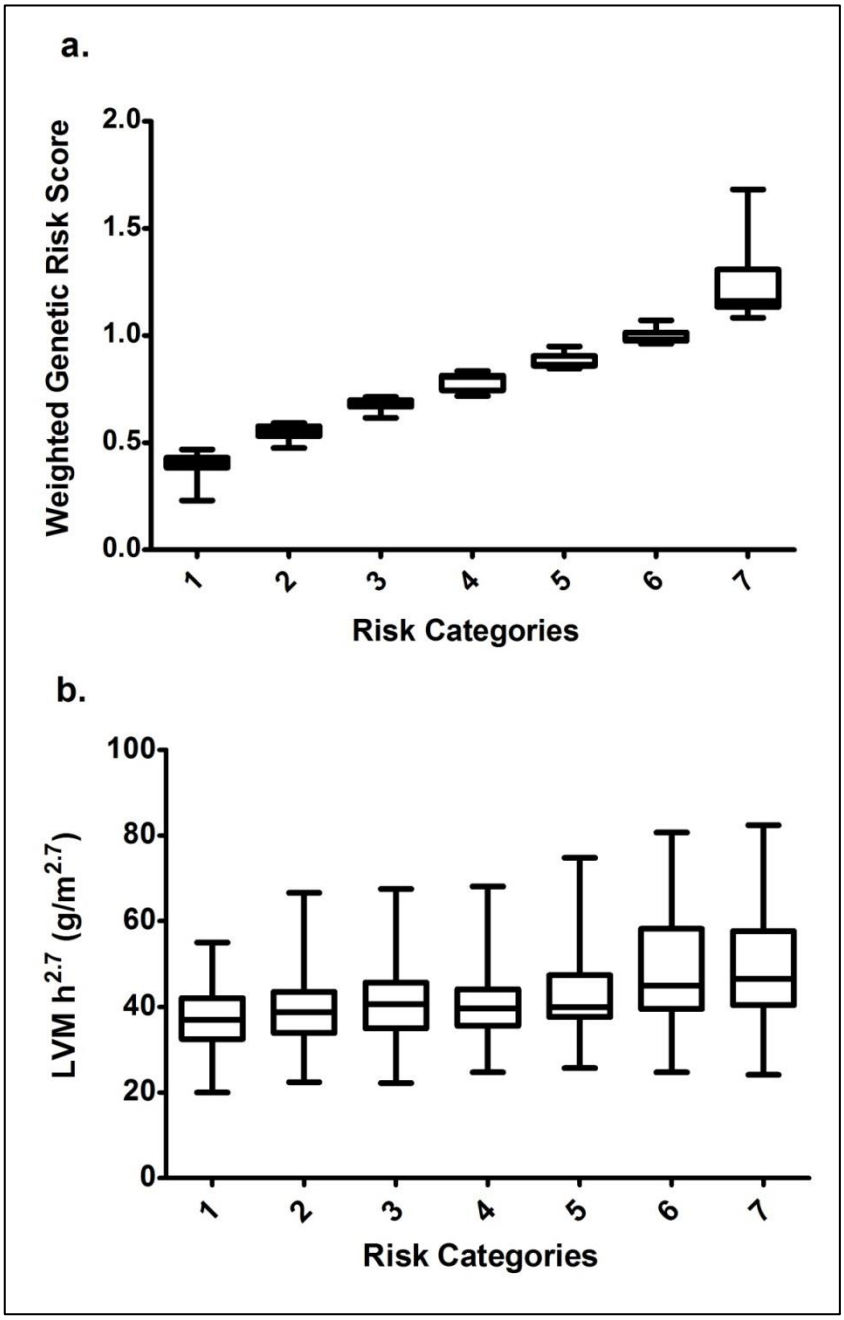


Figure 10 Distribution of percentage of controls (blue) and cases (grey) subjects in each risk category defined by wGRS.



The frequency is calculated according to the total number of individuals for each status.

Figure 11 Box plot of wGRS (a) and LVM (b) distribution for each risk categories.



$LVMh^{2.7}$ was expressed in $g/h^{2.7}$.

5 DISCUSSION

We performed a genome wide association analysis in an Italian cohort of never treated mild-to-moderate essential hypertensive subjects to search for genetic variants predictive of LVM trait. The absence of previous antihypertensive treatment, i.e. absence of unpredictable interferences by antihypertensive drugs on LVM, is the pivotal characteristic of our cohort to warrant a “clean reference phenotype” thus minimizing the “background noise” that often represents a bias in genotype-phenotype association studies. The antihypertensive treatments commonly used to control blood pressure affect cardiac mass, thus the reliability of the findings in association studies on cardiac mass in EH must be based on never treated EHs [87,88,89].

Left ventricular hypertrophy is considered as a powerful, independent risk factor for cardiovascular disease. Therefore, LVM can be considered not only as measures of cardiac structure, but also can offer prognostic information for assessing CVD risk [1]. Some evidences have demonstrated that regression of LVH is associated with a favorable prognosis [12,90]. Given the clinical importance of this trait, it can be useful to target intracellular signaling pathways involved in the regulation of LVM. In our study, overall we identified 85 SNPs associated to LVM with a suggestive p-value < 10^{-5}. Some of the genes in which our best SNPs map, *ROCK1*, *IGF1*, *CACNA1D*, *FGFR1*, *TRAF5*, *SOX5*, and *KSR2* were previously described as having a role in the pathogenesis of cardiac hypertrophy. Each of them might play a putative role in determining the LVM phenotype as well as other pathophysiological pathways directly or indirectly linked to cardiac pathophysiology.

Rho-Associated Coiled-Coil Containing Protein Kinase 1 (ROCK1) is a downstream mediator of the small GTP-binding protein RhoA. The

RhoA/Rho-kinase pathway is now widely known to play important roles in many cellular functions, including smooth muscle contraction, motility, proliferation, and apoptosis, and its excessive activity induces oxidative stress and promotes the development of cardiovascular disease [91]. A beneficial effect of long-term inhibition of Rho-kinase has been demonstrated in animal models for the treatment of various cardiovascular diseases [92]. In the vascular smooth muscle cells, the activation of RhoA/Rho-kinase pathway modules the expression of hypertrophy genes (*PAI-1*, *MCP-1*, etc.) [91] (figure 12). In vivo animal studies, using *ROCK* inhibitors, Y27632 and fasudil, suggested a role for ROCK in mediating cardiac hypertrophy and remodeling [93]. Shi J. and colleagues demonstrated that *ROCK1* deletion prevented or attenuated a variety of pathological characteristic of Gαq mice (cardiac hypertrophy animal model), such as induction of hypertrophic markers [94].

Insulin-Like Growth Factor 1 (IGF1) is a neurohumoral factor, member of a family of proteins involved in mediating growth and development. In cardiomyocytes, IGF1 activates multiple downstream signaling pathways for controlling cell death, metabolism, autophagy, differentiation, transcription, and protein synthesis [95]. These pathways involve the extracellular signal-regulated kinase (ERK) cascade, associated with the pro-hypertrophic and pro-survival actions [95], and the phosphatidylinositol-3 kinase (PI3K)/Akt pathway, associated with physiological cardiac hypertrophy [96, 97]. Moreover, (PI3K)/Akt signaling axis is critical for transducing physiological and adaptive hypertrophy, but also the overstimulation of the pathway can result in pathological hypertrophy [98]. In addition, the two pathways activated by IGF1 in cardiomyocytes (ERK and Akt pathway) crosstalk with each other leading to hypertrophy [98].

Calcium Channel Voltage-Dependent L-Type Alpha 1D Subunit (*CACNA1D*) gene encodes a member of the alpha-1 subunit family of

voltage-dependent calcium channels, also known as dihydropyridine-sensitive receptor (DHPR). These calcium channels mediate the entry of calcium ions into excitable cells and are involved in a variety of calcium-dependent processes, including cardiac and vascular smooth muscle contraction, hormone and neurotransmitter release, gene expression, cell motility, cell division, and cell death [99]. CACNA1D has a role in hypertrophic cardiomyopathy, in the pacemaker activity of the hearth and in atrial fibrillation [100,101].

The pathways activated by IGF1 lead also to Ca2+ influx and, in cardiomyocytes, IGF1 and CACNA1D protein are in the same cascade of Ca2+ signaling, leading to transcriptional upregulation and cardiac hypertrophy [95,98]. IGF1 and CACNA1D are in the same pathway of hypertrophic cardiomyopathy (HCM), a primary myocardial disorder with an autosomal dominant pattern of inheritance, characterized by hypertrophy of left ventricle. IGF-1, with other molecules (such as such as ACE1, angiotensin II, TGF-β, TNF-α, IL-6, and endothelin) increases the entrance of calcium into the cells, through L-type Ca2+ voltage-gated channels, as CACNA1D, and the activation of transcriptional pathways leads to the diverse histological and structural phenotypes of HCM including cardiac hypertrophy, interstitial fibrosis, and myocyte disarray [100] (Figure 13).

Variants at *IGF1* and *CACNA1D* genes have been also identified as hypertension susceptibility variants in two genome-wide association studies in Chinese cohorts [99,102].

Fibroblast Growth Factor Receptor 1 (*FGFR1*) gene encodes a receptor of fibroblast growth factors (FGFs). FGFs are secreted proteins with diverse functions mainly in development and metabolism. Some FGFs, such as FGF2, FGF16, FGF21, and FGF23 are secreted from the heart, are referred to as cardiomyokines and have an important role in heart function. In particular, FGF2 promotes cardiac hypertrophy and remodeling by

activating MAPK signaling; FGF16 and FGF21 may prevent cardiac hypertrophy and remodeling through the same signaling pathway; FGF23 promotes cardiac hypertrophy and remodeling through calcineurin/NFAT pathway. The biological effects of FGF2, FGF16, FGF21, and probably FGFR23 in the myocardium are mediated by the high-affinity tyrosine kinase receptor FGFR1, the major FGF receptor in the heart. These findings support the pathophysiological roles of FGFs and their receptors in the heart [103].

TNF Receptor-Associated Factor 5 (*TRAF5*) gene encodes for a signal transducer, member of the tumor necrosis factor (TNF) receptor family. In vivo experiments in TRAF5-deficient mice showed that TRAF5 deficiency aggravated cardiac hypertrophy, fibrosis and inflammation, and markedly promotes the activation of MEK-ERK1/2 pathway. Therefore, TRAF5 is an intrinsic suppressor of cardiac hypertrophy through the negative regulation of the MEK-ERK1/2 pathway. This signaling pathway plays a role in the progression of cardiac hypertrophy through the phosphorylation of intracellular targets, including transcription factors, under stress stimuli [104].

Kinase Suppressor of Ras 2 (KSR2) functions as an essential scaffolding protein to coordinate the activation of MEK-ERK cascade in response to calcium signals [105]. Evidence from in vitro studies, revealed that KSR2 is a relevant effector of Ca^{2+}-signaling and depletion of KSR2 significantly impaired ERK activation [105]. As previously mentioned, several studies provide strong evidence for an important role of the MEK-ERK cascade in the heart: a protective anti-apoptotic function as well as a hypertrophic function [106].

Three of the identified genes, *TRAF5*, *IGF1*, and *KSR2*, activate MEK-ERK cascade, well known as involved in the pathophysiological process of cardiac hypertrophy [106].

SRY sex determining region Y-box 5 (*SOX5*) gene encodes a transcription factor. Its expression is modulated by nitric oxide (NO) and guanylyl cyclase after stress in endothelial cells [107]. Endothelial NO levels are significantly correlated with LVM [108]. In addition, SOX5 plays a pivotal role in the expression of the muscle L-type Ca^{2+} channel, as CACNA1D [109]. Variation in the expression of these channels is associated with cardiac hypertrophy [110].

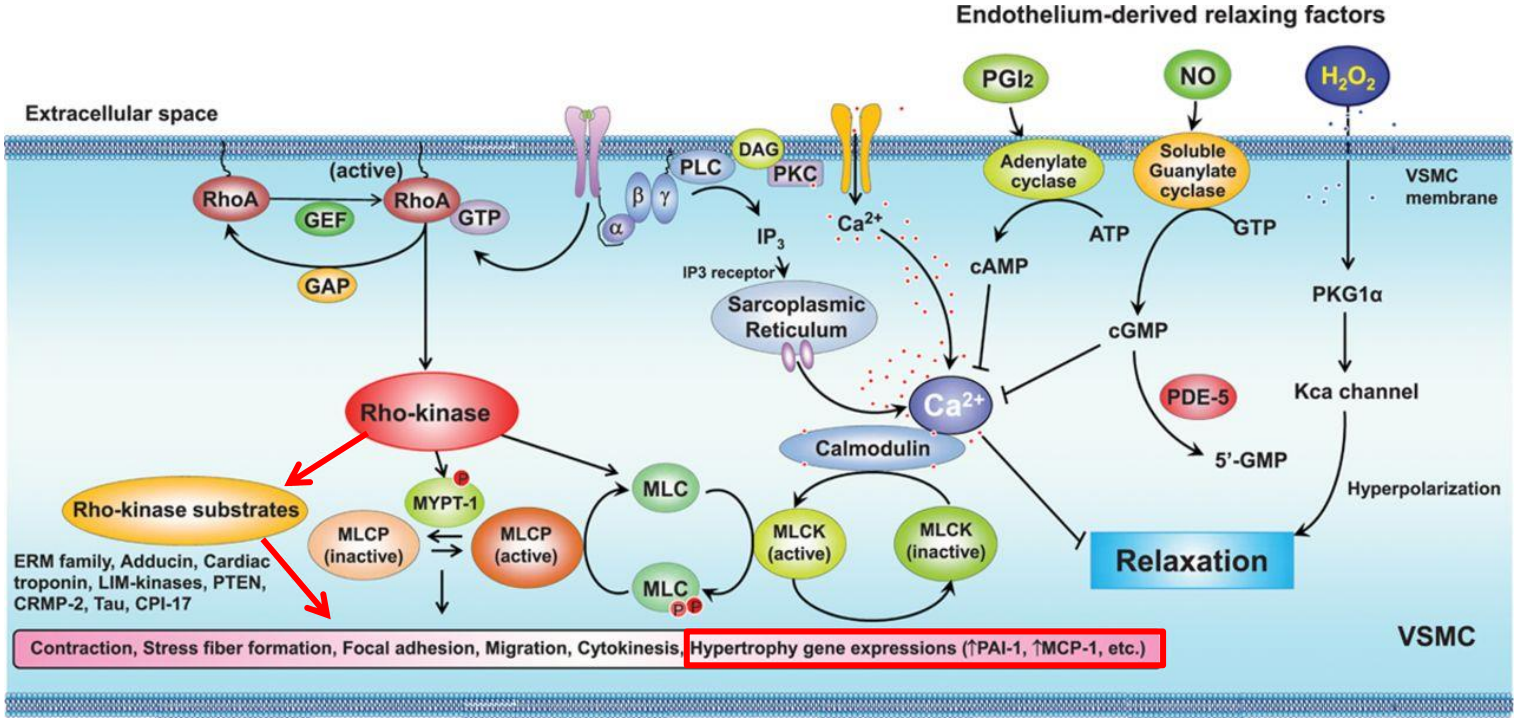
We also identified other SNPs significantly associated to LVM, in genes with a less clear or direct involvement in heart function: *WWOX*, *RGS7*, *CSMD1*, *CNTN5*, *FIGN*, and *HCN2*. Some of these genes have been identified as hypertension susceptibility genes in Han Chinese (*WWOX* [111]) and in two case-control studies in Korean cohorts (*CSMD1* [112]). *WWOX* has been also associated to LV wall thickness [9]; *RGS7* is associated with cardiac disease, including hypertrophy [113], as *HCN2* [114]; *CNTN5* has been associated with atrial fibrillation and heart failure [115]; *FIGN* is involved in blood pressure regulation [116] and with pulse pressure and mean arterial pressure [117]. Table A3 reports more details about their functional role in experimental setting.

The polymorphisms that are found significantly associated to LVM as quantitative trait were also tested for their association to LVM with a case-control approach. The odds ratios obtained were statistically significant (except for two variants). Nevertheless as for most of the complex traits, the observed odds ratios are modest (except for those biased by the absence of homozygous risk genotypes), so their relevance for a clinical use is uncertain. Thus, the possibility to combine more variants in a global genetic risk score could be interesting and could add relevance to the results. Further assessment of this score is therefore warranted, and implementation with newly discovered LVM related loci, and with non-

genetic susceptibility factors, such as environmental risk factors, could improve the predictive ability of the algorithm.

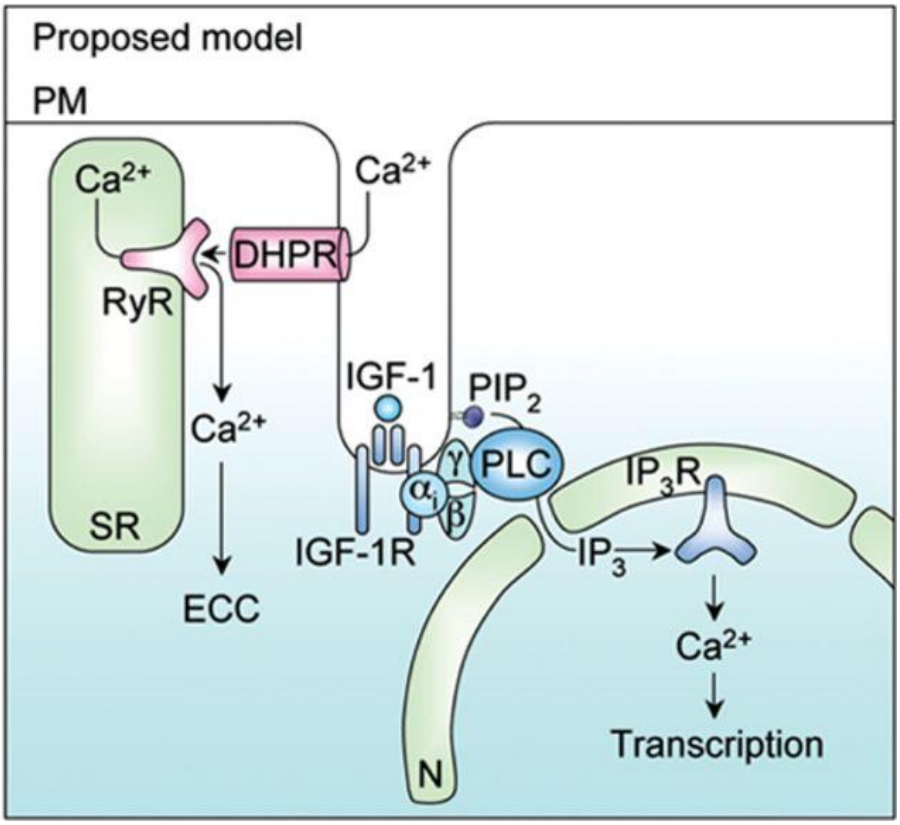
Genome-wide linkage and association studies have been performed by different groups to identify genetic loci associated to LVM. In a whole genome linkage study on hypertensive families, three regions (10q23.1, 12q14.1, 17p13.3) were found to approach suggestive evidence of linkage for LVH [57]. A GWAS on Koreans reported a significative correlation between the skeletal muscle Ca^{2+} channel protein RYR1 on chromosome 19 and ECG-LVH [58]. The HyperGEN study identified a polymorphism in *KCNB1* gene associated with LVM using a GWA approach [9]. A large meta-analysis identified loci associated with left ventricular structure on the solute carrier family 35, member F1 (*SLC35F1*) gene, chromosome 6 open reading frame 203 (*C6orf203*), and phospholamban (*PLN*) gene [59]. Recently, Barve RA *et al* identified eleven SNPs with a suggestive association with left ventricular mass trait, in a comparative study between M-mode and 2D echocardiography and between raw LVM and body surface area-indexed LVM. One SNP lies in *CDH13* gene was confirmed in all the four measurements [60]. None of the genes recognized by these studies achieves a p-value < 10^{-5} in our study. Candidate genes studies have been widely used before the advent of GWAS to explore the genetic basis of LVM. The genes identified with this approach were: angiotensin-converting enzyme (*ACE*) [46,47,48,49], aldosterone synthase (*CYP11B2*) [50], insulin-like growth factor (*IGF1*) [51], neuropeptide Y (*NPY*) [52], guanine nucleotide-binding protein 3 (*GNB3*) [53], endothelial nitric oxide synthase (*eNOS*) [54], peroxisome proliferator-activated receptor-alpha (*PPARA*) [55], and centlein centrosomal protein (*CNTLN*) [56] genes. Neither these genes achieves a significant p-value in our study, except *IGF1*.

Figure 12 Rho/Rho-kinase signaling in endothelial cells and vascular smooth muscle cells (VSMC) interaction. In red is highlighted the pathway that enhances expression of hypertrophic genes.



Adapted from Ref. 91.

Figure 13 Nuclear calcium signaling in cardiomyocytes.



DHPR is a calcium channel (CACNA1D is a member of this family of calcium channels) that mediates the entry of calcium ions into cardiomyocytes. IGF1R activation by IGF1 binding leads to nuclear Ca^{2+} signals and to expression of genes associated to development of cardiomyocyte hypertrophy. Abbreviations: RyR, ryanodine receptor; ECC, excitation–contraction coupling; PLC, phospholipase C; DHPR, dihydropyridine receptor. Adapted from Ref. 99 and 108.

6 CONCLUSION

Our GWAS allowed us to pinpoint genes whose role in heart function and/or cardiac hypertrophy has been demonstrated in previously publications by different authors. Moreover, we highlighted the usefulness of an aggregate measure of risk of LVH to discriminate high risk subjects. However, the present results must be interpreted within the context of the following potential limitations and perspectives.

We did not reach a Bonferroni's significance level probably due to a limited sample size. However, the phenotypic homogeneity of our cohort and the absence of previous antihypertensive treatment are prerequisites for the identification of true genetic effects.

GWAS are exploratory analyses and their replication represents the gold standard for assessing whether the findings are true-positive. In the present study, an independent cohort of hypertensive patients with LVM phenotype and enrolled with similar criteria was not available for replication.

GWA studies may establish significant genomic regions, though the real cause-effect relationship remains difficult to clarify. Moreover it often happens, as in our study, that significant SNPs map in non-coding regions, making it difficult to explain their causative role.

These limitations should not reduce the relevance of the genes identified and confirmed by previously published papers.

Future perspectives of this study should be:

1. Replication of the GWAS findings in independent cohorts to further confirm the genomic regions as true positives and to identify the putative genetic variant with respect to LVM.

2. Assessment in independent samples of the prediction ability (sensibility and specificity) of wGRS to correctly classify true positives (i.e. subjects with high risk of developing LVH) and true negatives (i.e. subjects with low risk) according to their genetic background.

REFERENCES

1 Kinno M, Waller AH, Gardin JM, “Approaches to echocardiography assessment of left ventricular mass: what does echocardiography add?” 2016 (http://www.acc.org/).

2 Foppa M, Duncan BB, Rohde LE. Echocardiography-based left ventricular mass estimation. How should we define hypertrophy? Cardiovasc Ultrasound. 2005; 3:17.

3 Schillaci G, Verdecchia P, Porcellati C, Cuccurullo O, Cosco C, Perticone F. Continuous relation between left ventricular mass and cardiovascular risk in essential hypertension. Hypertension. 2000; 35:580-6.

4 Bella JN, Göring HH. Genetic epidemiology of left ventricular mass. Am J Cardiovasc Dis. 2012; 2(4): 267-78.

5 Levy D, Garrison RJ, Savage DD, Kannel WB, Castelli WP. Prognostic implications of echocardiographically determined left ventricular mass in the Framingham Heart Study. N Engl J Med. 1990; 322:1561-6.

6 Bikkina M, Levy D, Evans JC, Larson MG, Benjamin EJ, Wolf PA et al. left ventricular mass and risk of stroke in an elderly cohort. JAMA. 1994; 272:33-6.

7 Gardin JM, McClelland R, Kitzman D, Lima JAC, Bommer W, Klopfenstein S et al. M-mode echocardiography predictors of six- to seven-year incidence of coronary heart disease, stroke, congestive heart failure and mortality in an elderly cohort (The Cardiovascular Health Study). Am J Cardiol 2001; 87:1051-7.

8 Vasan RS, Larson MG, Levy D, Evans JC, Benjiamin E. Distribution and categorization of echocardiography measurements in relation to reference

limits: the Framingham Heart Study: formulation of a height- and sex-specific classification and its validation. Circulation. 1997; 96: 1863-73.

9 Arnett DK, Li N, Tang W, Rao DC, Devereux RB, Claas SA, Kraemer R, Broeckel U. Genome-wide association study identifies single-nucleotide polymorphism in KCNB1 associated with left ventricular mass in humans: the HyperGEN Study. BMC Med Genet. 2009; 10:43.

10 Arnett DK, de las Fuentes L, Broeckel U. Genes for left ventricular hypertrophy. Curr Hypertens Rep. 2004; 6:36-41.

11 Gardin JM, Arnold A, Gottdiener JS, Wong ND, Fried LP, Klopfenstein HS, et al. Left ventricular mass in the elderly. The Cardiovascular Health Study. Hypertension. 1997; 29:1095-103.

12 Kannel WB. Left ventricular hypertrophy as a risk factor: the Framingham experience. J Hypertens Suppl. 1991; 9:S3-8; discussion S8-9.

13 Daniels SR, Kimball TR, Morrison JA, Khoury P, Witt S, Meyer RA. Effect of lean body mass, fat mass, blood pressure, and sexual maturation on left ventricular mass in children and adolescents. Statistical, biological, and clinical significance. Circulation. 1995; 92:3249-54.

14 Ruilope LM, Schmieder RE. Left ventricular hypertrophy and clinical outcomes in hypertensive patients. Am J Hypertens. 2008; 21:500-8.

15 de Simone G, Devereux RB, Daniels SR, Meyer RA. Gender differences in left ventricular growth. Hypertension. 1995; 26:979-83.

16 Ganau A, Devereux RB, Roman MJ, de Simone G, Pickering TG, Saba PS, et al. Patterns of left ventricular hypertrophy and geometric remodeling in essential hypertension. J Am Coll Cardiol. 1992; 19:1550-8.

17 Hammond IW, Devereux RB, Alderman MH, Lutas EM, Spitzer MC, Crowley JS, Laragh JH. The prevalence and correlates of echocardiographic left ventricular hypertrophy among employed patients with uncomplicated hypertension. J Am Coll Cardiol. 1986; 7:639-50.

18 Lang RM, Bierig M, Devereux RB, Flachskampf FA, Foster E, Pellikka PA, Recommendations for chamber quantification: a report from the American Society of Echocardiography's Guidelines and Standards Committee and the Chamber Quantification Writing Group, developed in conjunction with the European Association of Echocardiography, a branch of the European Society of Cardiology. J Am Soc Echocardiogr. 2005; 18:1440-63.

19 Kaess BM, Gona P, Larson MG, Cheng S, Aragam J, Kenchaiah S, et al. Secular trends in echocardiographic left ventricular mass in the community: the Framingham Heart Study. Heart. 2013; 99:1693-8.

20 Yeon SB, Salton CJ, Gona P, Chuang ML, Blease SJ, Han Y, et al. Impact of age, sex, and indexation method on MR left ventricular reference values in the Framingham Heart Study offspring cohort. J Magn Reson Imaging. 2015; 41:1038-45.

21 Levy D, Lahib SB, Anderson KM, Chrinstiansen JC, Kannel WB, Castelli WP. Determinants of sensitivity and specificity of electrocardiographic criteria for left ventricular hypertrophy. Circulation. 1990; 81:815-820.

22 de Simone G, Devereux RB, Daniels SR, Meyer RA. Gender differences in left ventricular growth. Hypertension. 1995; 26:979-83.

23 de Simone G, Devereux RB, Roman MJ, Ganau A, Chien S, Alderman MH, et al. Gender differences in left ventricular anatomy, blood viscosity and volume regulatory hormones in normal adults. Am J Cardiol. 1991; 68:1704-8.

24 de Simone G. Left ventricular hypertrophy in blacks and whites: different genes or different exposure? Hypertension. 2005; 46:23-4.

25 Drazner MH, Dries DL, Peshock RM, Cooper RS, Klassen C, Kazi F, Willett D, Victor RG. Left ventricular hypertrophy is more prevalent in blacks than whites in the general population: the Dallas Heart Study. Hypertension. 2005; 46:124-9.

26 Arnett DK, Strogatz DS, Ephross SA, Hames CG, Tyroler HA. Greater incidence of electrocardiographic left ventricular hypertrophy in black men than in white men in Evans County, Georgia. Ethn Dis. 1992; 2:10-7.

27 Materson BJ. Lessons on the interaction of race and antihypertensive drugs from the VA cooperative study group on antihypertensive agents. Am J Hypertens. 1995; 8:91s-93s.

28 Savage DD, Henry WL, Mitchell JR, Taylor AA, Gardin JM, Drayer JI, Laragh JH. Echocardiographic comparison of black and white hypertensive subjects. J Natl Med Assoc. 1979; 71:709-12.

29 Kuch B, Hense HW, Gneiting B, Doring A, Muscholl M, Brockel U, Schunkert H. Body composition and prevalence of left ventricular hypertrophy. Circulation. 2000; 102:405-10.

30 Kuperstein R, Hanly P, Niroumand M, Sasson Z: The importance of age and obesity on the relation between diabetes and left ventricular mass. J Am Coll Cardiol. 2001; 37:1957-62.

31 De Simone G, Palmieri V, Bella JN, Celentano A, Hong Y, Oberman A, et al. Association of left ventricular hypertrophy with metabolic risk factors: the HyperGEN study. J Hypertens. 2002; 20:323-31.

32 Gardin JM, Siscovick D, Anton-Culver H, Lynch JC, Smith VE, Klopfenstein HS, et al. Sex, age, and disease affect echocardiographic left

ventricular mass and systolic function in the free-living elderly. The Cardiovascular Health Study. Circulation. 1995; 91:1739-48.

33 Iacobellis G, Ribaudo MC, Leto G, Zappaterreno A, Vecci E, Di Mario U, Leonetti F. Influence of excess fat on cardiac morphology and function: study in uncomplicated obesity. Obes Res. 2002; 10:767-73.

34 Chobanian AV, Bakris GL, Black HR, Cushman WC, Green LA, Izzo JL, et al.: The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: the JNC 7 report. JAMA. 2003; 289:2560-2572.

35 Devereux RB, Roman MJ, Paranicas M, O'Grady MJ, Lee ET, Welty TK, et al. Impact of diabetes on cardiac structure and function: the strong heart study. Circulation. 2000; 101:2271-6.

36 Lee M, Gardin JM, Lynch JC, Smith V-E, Tracy RP, Savage PJ, et al. Diabetes mellitus and echocardiographic left ventricular function in free-living elderly men and women: The Cardiovascular Health Study. Am Heart J. 1997; 133:36-43.

37 Feener EP, King GL, Vascular dysfunction in diabetes mellitus. Lancet. 1997; 350(Suppl 1):SI9-13.

38 Manolio TA, Levy D, Relation of alcohol intake to left ventricular mass: The Framingham Study. J Am Coll Cardiol. 1991; 17:717-21.

39 Schmieder RE, Messerli FH, Garavaglia GE, Nunez BD: Dietary salt intake. A determinant of cardiac involvement in essential hypertension. Circulation. 1988; 78:951-6.

40 Savage DD, Levy D, Dannenberg AL, Garrison RJ, Castelli WP: Association of echocardiographic left ventricular mass with body size, blood

pressure and physical activity (the Framingham Study). Am J Cardiol. 1990; 65:371-6.

41 Gardin JM, Wagenknecht LE, Anton-Culver H, Flack J, Gidding S, Kurosaki T, Wong ND, Manolio TA: Relationship of cardiovascular risk factors to echocardiographic left ventricular mass in healthy young black and white adult men and women. The CARDIA Study. Circulation. 1995; 92:380-7.

42 Jullien V, Gosse P, Ansoborlo P, Lemetayer P, Clementy J: Relationship between left ventricular mass and serum cholesterol level in the untreated hypertensive. J Hypertens. 1998; 16:1043-7.

43 Bella JN, Devereux RB, Roman MJ, O'Grady MJ, Welty TK, Lee ET, Fabsitz RR, Howard BV: Relations of left ventricular mass to fat-free and adipose body mass: the strong heart study. The Strong Heart Study Investigators. Circulation. 1998; 98:2538-44.

44 Jin Y, Kuznetsova T, Bochud M, Richart T, Thijs L, Cusi D, Fagard R, Staessen JA. Heritability of left ventricular structure and function in Caucasian families. Eur J Echocardiogr. 2011; 12:326-32.

45 Garner C, Lecomte E, Visvikis S, Abergel E, Lathrop M, Soubrier F. Genetic and environmental influences on left ventricular mass. A family study. Hypertension. 2000; 36:740-6.

46 Schunkert H, Hense HW, Holmer SR, Stender M, Perz S, Keil U, Lorell BH, Riegger GA. Association between a deletion polymorphism of the angiotensin-converting-enzyme gene and left ventricular hypertrophy. N Engl J Med. 1994; 330:1634-8.

47 Celentano A, Mancini FP, Crivaro M, Palmieri V, Ferrara LA, De Stefano V, et al. Cardiovascular risk factors, angiotensin-converting enzyme gene

I/D polymorphism, and left ventricular mass in systemic hypertension. Am J Cardiol. 1999; 83:1196-200.

48 Doolan G, Nguyen L, Chung J, Ingles J, Semsarian C. Progression of left ventricular hypertrophy and the angiotensin-converting enzyme gene polymorphism in hypertrophic cardiomyopathy. Int J Cardiol. 2004; 96:157-63.

49 Kuznetsova T, Staessen JA, Thijs L, Kunath C, Olszanecka A, Ryabikov A, et al. Left Ventricular mass in relation to genetic variation in angiotensin II receptors, renin system genes, and sodium excretion. Circulation. 2004; 110:2644-50.

50 Kupari M, Hautanen A, Lankinen L, Koskinen P, Virolainen J, Nikkila H, et al. Associations between human aldosterone synthase (CYP11B2) gene polymorphisms and left ventricular size, mass, and function. Circulation. 1998; 97:569-75.

51 Nagy Z, Busjahn A, Bähring S, Faulhaber HD, Gohlke HR, Knoblauch H, et al. Quantitative trait loci for blood pressure exist near the IGF-1, the Liddle syndrome, the angiotensin II-receptor gene and the renin loci in man. J Am Soc Nephrol. 1999; 10:1709-16.

52 Kuch-Wocial A, Slubowska K, Kostrubiec M, Pasierski T, Januszewicz W, et al. Plasma neuropeptide Y immunoreactivity influences left ventricular mass in pheochromocytoma. Clin Chim Acta. 2004; 345:43-7.

53 Semplicini A, Siffert W, Sartori M, Monari A, Naber C, Frigo G, et al. G protein beta3 subunit gene 825T allele is associated with increased left ventricular mass in young subjects with mild hypertension. Am J Hypertens. 2001; 14:1191-5.

54 Xin Y, Song X, Xue H, Liu Z, Wang X, Wang H, et al. A common variant of the eNOS gene (E298D) is an independent risk factor for left ventricular

hypertrophy in human essential hypertension. Clin Sci (Lond). 2009; 117:67-73.

55 Flavell DM, Jamshidi Y, Hawe E, Pineda Torra I, Taskinen MR, Frick MH, et al. Peroxisome proliferator-activated receptor alpha gene variants influence progression of coronary atherosclerosis and risk of coronary artery disease. Circulation. 2002; 105:1440-5.

56 Menni C, Boffi L, Cesana F, Viviani Anselmi C, Nava S, Bertola F, et al. Variant on chromosome 9p is associated with left ventricular mass: results from two cohorts of essential hypertensive individuals. J Hypertens. 2012; 30:2144-50.

57 Mayosi BM, Avery PJ, Farrall M, Keavney B, Watkins H. Genome-wide linkage analysis of electrocardiographic and echocardiographic left ventricular hypertrophy in families with hypertension. Eur Heart J. 2008; 29:525-30.

58 Hong KW, Shin DJ, Lee SH, Son NH, Go MJ, Lim JE, et al. Common variants in RYR1 are associated with left ventricular hypertrophy assessed by electrocardiogram. Eur Heart J. 2012; 33:1250-6.

59 Vasan RS, Glazer NL, Felix JF, Lieb W, Wild PS, Felix SB, et al. Genetic variants associated with cardiac structure and function: a meta-analysis and replication of genome-wide association data. JAMA. 2009; 302:168-78.

60 Barve RA, Gu CC, Yang W, Chu J, Dávila-Román VG, de las Fuentes L. Genetic association of left ventricular mass assessed by M-mode and two-dimensional echocardiography. J Hypertens. 2016; 34:88-96.

61 Schmieder RE, Martus P, Klingbeil A. Reversal of left ventricular hypertrophy in essential hypertension. A meta-analysis of randomized double-blind studies. JAMA 1996; 275:1507-1513.

62 Liebson PR, Grandits GA, Dianzumba S, Prineas RJ, Grimm RH Jr, Neaton JD, et al. Comparison of five antihypertensive monotherapies and placebo for change in left ventricular mass in patients receiving nutritional-hygienic therapy in the Treatment of Mild Hypertension Study (TOMHS). Circulation 1995; 91:698-706.

63 Celebi AS, Yalcin H, Yalcin F. Current cardiac imaging techniques for detection of left ventricular mass. Cardiovasc Ultrasound. 2010; 8:19.

64 Douglas PS, Garcia MJ, Haines DE, Lai WW, Manning WJ, Patel AR, et al. ACCF/ASE/AHA/ASNC/HFSA/HRS/SCAI/SCCM/SCCT/SCMR 2011 Appropriate Use Criteria for Echocardiography. A Report of the American College of Cardiology Foundation Appropriate Use Criteria Task Force, American Society of Echocardiography, American Heart Association, American Society of Nuclear Cardiology, Heart Failure Society of America, Heart Rhythm Society, Society for Cardiovascular Angiography and Interventions, Society of Critical Care Medicine, Society of Cardiovascular Computed Tomography, and Society for Cardiovascular Magnetic Resonance Endorsed by the American College of Chest Physicians. J Am Coll Cardiol. 2011; 57:1126-66.

65 Lang RM, Badano LP, Mor-Avi V, Afilalo J, Armstrong A, Ernande L, et al. Recommendations for cardiac chamber quantification by echocardiography in adults: an update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. J Am Soc Echocardiogr. 2015; 28:1-39.

66 Devereux RB, Alonso DR, Lutas EM, Gottlieb GJ, Campo E, Sachs I, et al. Echocardiographic assessment of left ventricular hypertrophy: comparison to necropsy findings. Am J Cardiol. 1986; 57:450-8.

67 De Simone G, Daniels SR, Devereux RB, Meyer RA, Roman MJ, de Divitiis O, et al. Left ventricular mass and body size in normotensive children and adults: assessment of allometric relations and impact of overweight. J Am Coll Cardiol. 1992; 20:1251-60.

68 Chirinos JA, Segers P, De Buyzere ML, Kronmal RA, Raja MW, De Bacquer D, et al. Left ventricular mass: allometric scaling, normative values, effect of obesity, and prognostic performance. Hypertension. 2010; 56:91-8.

69 Salvi E, Kutalik Z, Glorioso N, Benaglio P, Frau F, Kuznetsova T, et al. Genomewide association study using a high-density single nucleotide polymorphism array and case-control design identifies a novel essential hypertension susceptibility locus in the promoter region of endothelial NO synthase. Hypertension. 2012; 59:248-55.

70 Frau F, Zaninello R, Salvi E, Ortu MF, Braga D, Velayutham D, et al. Genome-wide association study identifies CAMKID variants involved in blood pressure response to losartan: the SOPHIA study. Pharmacogenomics. 2014; 15:1643-52.

71 Chittani M, Zaninello R, Lanzani C, Frau F, Ortu MF, Salvi E, et al. TET2 and CSMD1 genes affect SBP response to hydrochlorothiazide in never-treated essential hypertensives. J Hypertens. 2015; 33:1301-9.

72 Anderson CA, Pettersson FH, Clarke GM, Cardon LR, Morris AP, Zondervan KT. Data quality control in genetic case-control association studies. Nat Protoc. 2010; 5:1564-73.

73 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81:559-75.

74 Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet. 2006; 38:904-9.

75 Patterson N, Price AL, Reich D. Population structure and eigenanalysis. PLoS Genet. 2006; 2:e190.

76 Delaneau O, Marchini J, 1000 Genomes Project Consortium, 1000 Genomes Project Consortium. Integrating sequence and array data to create an improved 1000 Genomes Project haplotype reference panel. Nat Commun. 2014; 5:3934.

77 Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. Nat Genet. 2012; 44:955-9.

78 1000 Genomes Project Consortium, Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, et al. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012; 491:56-65.

79 Porcu E, Sanna S, Fuchsberger C, Fritsche LG. Genotype imputation in genome-wide association studies. Curr Protoc Hum Genet. 2013; Chapter 1:Unit 1.25.

80 Li Y, Willer C, Sanna S, Abecasis G. Genotype imputation. Annu Rev Genomics Hum Genet. 2009; 10:387-406.

81 De Jager PL, Chibnik LB, Cui J, Reischl J, Lehr S, Simon KC, et al. Integration of genetic risk factors into a clinical algorithm for multiple sclerosis susceptibility: a weighted genetic risk score. Lancet Neurol. 2009; 8:1111-9.

82 Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, et al. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics. 2010; 26:2336-7.

83 McLaren W, Pritchard B, Rios D, Chen Y, Flicek P, Cunningham F. Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. Bioinformatics. 2010; 26:2069-70.

84 Matys V, Kel-Margoulis OV, Fricke E, Liebich I, Land S, Barre-Dirrie A, et al. TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. Nuc Ac Res. 2006; 34:D108-110.

85 Karolchik D, Hinrichs AS, Furey TS, Roskin KM, Sugnet CW, Haussler D, et al. The UCSC Table Browser data retrieval tool. Nucleic Acids Res. 2004; 32:D493-6.

86 Shimokawa H, Takeshita A. Rho-kinase is an important therapeutic target in cardiovascular medicine. Arterioscler Thromb Vasc Biol. 2005; 25:1767-75.

87 Gosse P, Dallocchio M. Left ventricular hypertrophy: epidemiological prognosis and associated critical factors. Eur Heart J. 1993; Suppl D:16-21.

88 Devereux RB, Koren MJ, de Simone G, Okin PM, Kligfield P. Methods for detection of left ventricular hypertrophy: application to hypertensive heart disease. Eur Heart J. 1993; Suppl D:8-15.

89 Fagard RH, Celis H, Thijs L, Wouters S. Regression of left ventricular mass by antihypertensive treatment: a meta-analysis of randomized comparative studies. Hypertension. 2009; 54:1084-91.

90 Gardin JM, Lauer MS. Left ventricular hypertrophy: the next treatable, silent killer? JAMA. 2004; 292:2396-8.

91 Shimokawa H, Sunamura S, Satoh K. RhoA/Rho-Kinase in the Cardiovascular System. Circ Res. 2016; 118:352-66.

92 Dong M, Yan BP, Liao JK, Lam YY, Yip GW, Yu CM. Rho-kinase inhibition: a novel therapeutic target for the treatment of cardiovascular diseases. Drug Discov Today. 2010; 15:622-9.

93 Shi J, Zhang YW, Summers LJ, Dorn GW 2nd, Wei L. Disruption of ROCK1 gene attenuates cardiac dilation and improves contractile function in pathological cardiac hypertrophy. J Mol Cell Cardiol. 2008; 44:551-60.

94 Shi J, Zhang YW, Yang Y, Zhang L, Wei L. ROCK1 plays an essential role in the transition from cardiac hypertrophy to failure in mice. J Mol Cell Cardiol. 2010; 49:819-28.

95 Troncoso R, Ibarra C, Vicencio JM, Jaimovich E, Lavandero S. New insights into IGF-1 signaling in the heart. Trends Endocrinol Metab. 2014; 25:128-37.

96 McMullen JR, Izumo S. Role of the insulin-like growth factor 1 (IGF1)/phosphoinositide-3-kinase (PI3K) pathway mediating physiological cardiac hypertrophy. Novartis Found Symp. 2006; 274:90-111.

97 Dorn GW 2nd. The fuzzy logic of physiological cardiac hypertrophy. Hypertension. 2007; 49:962-70.

98 Ibarra C, Vicencio JM, Estrada M, Lin Y, Rocco P, Rebellato P, et al. Local control of nuclear calcium signaling in cardiac myocytes by perinuclear microdomains of sarcolemmal insulin-like growth factor 1 receptors. Circ Res. 2013; 112:236-45.

99 Lu X, Wang L, Lin X, Huang J, Charles Gu C, He M, et al. Genome-wide association study in Chinese identifies novel loci for blood pressure and hypertension. Hum Mol Genet. 2015; 24:865-74.

100 Cambronero F, Marín F, Roldán V, Hernández-Romero D, Valdés M, Lip GY. Biomarkers of pathophysiology in hypertrophic cardiomyopathy: implications for clinical management and prognosis. Eur Heart J. 2009; 30:139-51.

101 Catterall WA. Voltage-gated calcium channels. Cold Spring Harb Perspect Biol. 2011; 3:a003947.

102 Yang HC, Liang YJ, Chen JW, Chiang KM, Chung CM, Ho HY, et al. Identification of IGF1, SLC4A4, WWOX, and SFMBT1 as hypertension susceptibility genes in Han Chinese with a genome-wide gene-based association study. PLoS One. 2012; 7:e32907.

103 Itoh N, Ohta H. Pathophysiological roles of FGF signaling in the heart. Front Physiol. 2013; 4:247.

104 Bian Z, Dai J, Hiroyasu N, Guan H, Yuan Y, Gan L, et al. Disruption of tumor necrosis factor receptor associated factor 5 exacerbates pressure overload cardiac hypertrophy and fibrosis. J Cell Biochem. 2014; 115:349-58.

105 Dougherty MK, Ritt DA, Zhou M, Specht SI, Monson DM, Veenstra TD, et al. KSR2 is a calcineurin substrate that promotes ERK cascade activation in response to calcium signals. Mol Cell. 2009; 34:652-62.

106 Lorenz K, Schmitt JP, Vidal M, Lohse MJ. Cardiac hypertrophy: targeting Raf/MEK/ERK1/2-signaling. Int J Biochem Cell Biol. 2009; 41:2351-5

107 Della-Morte D, Beecham A, Rundek T, Wang L, McClendon MS, Slifer S, et al. A follow-up study for left ventricular mass on chromosome 12p11 identifies potential candidate genes. BMC Med Genet. 2011; 12:100.

108 Hua L, Li C, Xia D, Qu P, Li Z, Zhang W, et al. Relationship between hypertensive left ventricular hypertrophy and levels of endothelin and nitric oxide. Hypertens Res. 2000; 23:377-80.

109 Zheng Z, Wang ZM, Delbono O. Charge movement and transcription regulation of L-type calcium channel alpha(1S) in skeletal muscle cells. J Physiol. 2002; 540:397-409.

110 Benitah JP, Alvarez JL, Gómez AM. L-type Ca(2+) current in ventricular cardiomyocytes. J Mol Cell Cardiol. 2010; 48:26-36.

111 Yang HC, Liang YJ, Chen JW, Chiang KM, Chung CM, Ho HY, et al. Identification of IGF1, SLC4A4, WWOX, and SFMBT1 as hypertension susceptibility genes in Han Chinese with a genome-wide gene-based association study. PLoS One. 2012; 7:e32907.

112 Hong KW, Go MJ, Jin HS, Lim JE, Lee JY, Han BG, et al. Genetic variations in ATP2B1, CSK, ARSG and CSMD1 loci are related to blood pressure and/or hypertension in two Korean cohorts. J Hum Hypertens. 2010; 24:367-372.

113 Zhang P, Mende U. Functional role, mechanisms of regulation, and therapeutic potential of regulator of G protein signaling 2 in the heart. Trends Cardiovasc Med. 2014; 24:85-93.

114 Hofmann F, Fabritz L, Stieber J, Schmitt J, Kirchhof P, Ludwig A, et al. Ventricular HCN channels decrease the repolarization reserve in the hypertrophic heart. Cardiovasc Res. 2012; 95:317-326.

115 Larson MG, Atwood LD, Benjamin EJ, Cupples LA, D'Agostino RB Sr, Fox CS, et al. Framingham Heart Study 100K project: genome wide associations for cardiovascular disease outcomes. BMC Med Genet. 2007; 8 Suppl 1:S5.

116 Kato N, Takeuchi F, Tabara Y, Kelly TN, Go MJ, Sim X et al. Meta-analysis of genome wide association studies identifies common variants associated with blood pressure variation in east Asians. Nat Genet. 2011; 43:531-538.

117 Wain LV, Verwoert GC, O'Reilly PF, Shi G, Johnson T, Johnson AD, et al. Genome wide association study identifies six new loci influencing pulse pressure and mean arterial pressure. Nat Genet. 2011; 43:1005-1011.

INTERNET RESOURCES

http://www.interomics.eu

http://bioinformatics.illumina.com/informatics/sequencing-microarray-data-analysis/genomestudio.html

http://pngu.mgh.harvard.edu/purcell/plink

https://mathgen.stats.ox.ac.uk/genetics_software/shapeit/shapeit.html

http://www.sph.umich.edu/csg/abecasis/MACH/index.html

http://genome.sph.umich.edu/wiki/Minimac

http://www.sph.umich.edu/csg/yli/software.html

http://www.stata.com

http://genepath.med.harvard.edu/~reich/Software.htm

http://www.r-project.org

http://csg.sph.umich.edu/locuszoom

http://grch37.ensembl.org/Homo_sapiens/Tools/VEP

http://genome.ucsc.edu/

http://www.kegg.jp/

APPENDIX

***Table A1**: Clinical research centers that recruited patients for the study.*

Recruitment center	Number of subjects
Department of Medicine and Hypertension Centre University of Catania	31
Department of Clinic and Experimental Medicine, Cardiovascular Disease Unit, University of Catanzaro 'Magna Graecia'	41
Division of Internal Medicine, ASL Isernia	46
Department of Internal Medicine and Public Health, University of L'Aquila	44
Department of Internal Medicine, University Federico II, Napoli	29
Department of Medical and Surgical Sciences, University of Padova	31
Department of Clinical And Experimental Medicine, University of Padova	11
Department of Internal Medicine, University of Pisa	23
Division of Internal Medicine, ASL Reggio Emilia	8
Division of Internal Medicine and Hypertension Center, S.Giovanni Battista e della Città di Torino Hospital, University of Torino	26
Joint ASL n.1, AOU Sassari	676

***Table A2**: Association results for LVMh[2.7].*

SNP	chr	position (bp)	alleles	risk allele freq	gene	location	beta	SE	p-value
rs7207593	17	66037405	T/C	0.02	KPNA2	intron	-13.8	3.52	8.94E-05
rs78633628	16	78971380	C/T	0.01	WWOX	intron	-10.95	2.46	8.66E-06
rs4966230	15	28969573	A/G	0.04	WHAMMP2	intron	-7.79	1.56	5.66E-07
rs2486612	9	99730155	A/G	0.04	HIATL2	intron	-6.99	1.47	1.93E-06
rs9620166	22	23076572	T/C	0.04	IGLV3-17	upstream	-6.4	1.59	5.79E-05
rs17252060	5	53321543	G/A	0.03	ARL15	intron	-5.85	1.4	2.89E-05
rs114996389	3	99379688	T/G	0.03	COL8A1	intron	-5.16	1.3	7.02E-05
rs111739640	1	242430997	C/T	0.06	PLD5	intron	-4.41	1.09	5.49E-05
rs9274407	6	32632832	A/T	0.06	HLA-DQB1-AS1	exon (missense)	-4.27	1.09	8.63E-05
rs74840030	3	111901099	C/T	0.06	SLC9C1	exon (missense)	-3.63	0.86	2.63E-05
rs73853324	3	111887788	G/A	0.07	SLC9C1	exon (missense)	-3.59	0.85	2.54E-05
rs76400391	3	111915627	C/T	0.07	SLC9C1	intron	-3.58	0.84	2.09E-05
rs59688356	14	58224134	T/C	0.06	SLC35F4	intron	-3.43	0.87	8.00E-05
rs10167952	2	29809505	G/C	0.12	ALK	intron	-2.8	0.67	3.17E-05
rs56096309	10	17877539	A/C	0.2	MRC1	intron	-2.65	0.59	6.56E-06
rs17068332	8	3833181	T/C	0.12	CSMD1	intron	-2.63	0.63	3.05E-05
rs13023211	2	164504320	G/A	0.14	FIGN	intron	-2.55	0.63	4.85E-05
rs6007872	22	48897704	A/G	0.13	FAM19A5	intron	-2.5	0.63	8.14E-05
rs9867121	3	114631548	A/C	0.18	ZBTB20	intron	-2.38	0.57	2.46E-05
rs9442871	6	73648263	T/C	0.16	KCNQ5	intron	-2.38	0.59	5.21E-05
rs77732888	15	49088198	A/G	0.16	CEP152	intron	-2.31	0.58	7.67E-05
rs35422477	8	38242712	G/A	0.2	WHSC1L1	upstream	-2.08	0.52	7.14E-05
rs2288696	8	38286225	A/G	0.19	FGFR1	intron	-2.06	0.53	8.79E-05
rs7959200	12	9330417	A/C	0.24	PZP	intron	-2.04	0.51	5.46E-05
rs4731960	7	133206438	G/A	0.28	EXOC4	intron	-1.97	0.49	6.45E-05
rs682856	3	174415647	C/T	0.47	NAALADL2	intron	-1.76	0.42	2.29E-05
rs10863888	1	211502769	A/G	0.41	TRAF5	intron	-1.71	0.42	4.23E-05
rs12403576	1	211439571	A/G	0.41	RCOR3	intron	-1.67	0.42	6.42E-05
rs3002258	1	211617397	T/G	0.42	ARPC3P2	downstream	-1.64	0.42	8.73E-05
rs6579841	5	150622225	G/A	0.42	GM2A	intron	-1.63	0.42	9.70E-05
rs5930	19	11224265	A/G	0.42	LDLR	downstream	-1.63	0.42	9.08E-05
rs784369	1	211558096	A/G	0.42	LINC00467	intron	-1.62	0.42	9.91E-05

SNP	chr	position (bp)	alleles	risk allele freq	gene	location	beta	SE	p-value
rs7137607	12	23778584	C/A	0.45	SOX5	intron	-1.61	0.41	9.63E-05
rs6700747	1	211431105	A/C	0.44	RCOR3	upstream	-1.6	0.41	9.79E-05
rs9646862	2	220233328	C/T	0.55	DNPEP	downstream	-1.65	0.42	8.02E-05
rs12198266	6	83743782	T/C	0.55	UBE3D	intron	-1.66	0.42	8.28E-05
rs2993503	1	3028987	G/A	0.5	PRDM16	intron	-1.67	0.43	9.30E-05
rs10437924	12	52951208	G/A	0.63	KRT71	upstream	-1.7	0.43	8.01E-05
rs6590636	11	100047729	C/A	0.56	CNTN5	intron	-1.71	0.42	4.30E-05
rs1790650	16	57511069	G/A	0.62	DOK4	intron	-1.71	0.44	9.76E-05
rs9284436	19	607108	T/C	0.57	HCN2	intron	-1.75	0.44	5.89E-05
rs7504272	18	19172917	C/T	0.39	ESCO1	downstream	-1.76	0.45	8.58E-05
rs291785	18	19015831	T/C	0.38	GREB1L	intron	-1.77	0.43	4.26E-05
rs6887266	5	13837218	G/A	0.38	DNAH5	intron	-1.77	0.45	7.45E-05
rs56202747	11	89700849	A/G	0.31	TRIM64	upstream	-1.78	0.45	8.68E-05
rs114443041	1	24281021	G/A	0.54	CNR2	intron	-1.86	0.39	1.60E-06
rs6946807	7	146608784	G/A	0.64	CNTNAP2	intron	-1.88	0.45	2.83E-05
rs3825448	13	25145058	C/T	0.75	PSPC1P2	downstream	-1.93	0.49	8.42E-05
rs10166469	2	29829167	C/T	0.76	ALK	intron	-1.94	0.5	9.49E-05
rs115299983	21	15199811	G/A	0.69	CNN2P7	nc transcript	-1.95	0.45	1.55E-05
rs7302899	12	118302322	G/A	0.45	KSR2	intron	-1.95	0.41	2.24E-06
rs7642531	3	23842814	A/C	0.77	UBE2E1-AS1	downstream	-2.05	0.52	7.71E-05
rs35910954	15	92581551	A/T	0.79	SLCO3A1	intron	-2.06	0.53	8.88E-05
rs7894754	10	73074310	C/T	0.81	SLC29A3	upstream	-2.09	0.53	8.01E-05
rs1149050	1	31200342	A/G	0.78	MATN1-AS1	downstream	-2.11	0.54	9.81E-05
rs55803951	9	7134281	C/T	0.77	KDM4C	intron	-2.22	0.52	2.07E-05
rs1198591	1	98541836	G/C	0.81	NFU1P2	upstream	-2.23	0.55	4.27E-05
rs12369523	12	118309008	G/A	0.43	KSR2	intron	-2.25	0.43	1.37E-07
rs35996865	18	18692344	T/G	0.7	ROCK1	upstream	-2.33	0.46	3.17E-07
rs1501127	4	16554662	T/C	0.81	LDB2	intron	-2.38	0.52	5.14E-06
rs1599951	3	28359788	A/G	0.87	CMC1	intron	-2.5	0.62	5.48E-05
rs55704326	3	28534629	A/T	0.86	ZCWPW2	intron	-2.69	0.69	9.77E-05
rs153452	5	150612619	A/G	0.89	GM2A	intron	-2.75	0.68	5.65E-05
rs9544846	13	79121654	A/G	0.91	RNF219-AS1	intron	-2.83	0.7	5.67E-05
rs10970978	9	32442256	A/G	0.88	ACO1	intron	-2.93	0.67	1.02E-05
rs17665445	3	29347554	A/G	0.88	RBMS3	intron	-2.99	0.68	9.70E-06

SNP	chr	position (bp)	alleles	risk allele freq	gene	location	beta	SE	p-value
rs153458	5	150607972	C/A	0.89	CCDC69	upstream	-3.01	0.7	1.59E-05
rs73226977	3	139928004	G/T	0.91	CLSTN2	intron	-3.04	0.75	5.09E-05
rs113551213	12	12513870	G/A	0.92	LOH12CR1	intron	-3.31	0.85	9.70E-05
rs7856710	9	15521362	G/A	0.94	RN7SL98P	downstream	-3.57	0.89	6.19E-05
rs62571027	9	15515899	G/A	0.94	PSIP1	upstream	-3.59	0.89	5.70E-05
rs116435336	2	238563732	C/T	0.95	LRRFIP1	intron	-3.66	0.94	9.15E-05
rs34077724	16	81739784	G/T	0.95	CMIP	intron	-3.87	0.96	5.81E-05
rs74870432	11	84932409	T/C	0.98	DLG2	intron	-5.36	1.31	4.14E-05
rs182304522	11	85216321	G/A	0.98	RNU6-1292P	upstream	-5.39	1.29	3.14E-05
rs17487808	4	166024258	A/G	0.98	TMEM192	intron	-5.51	1.4	8.44E-05
rs79910493	12	102843754	C/T	0.98	IGF1	intron	-5.94	1.46	4.60E-05
rs7933179	11	85419919	C/T	0.97	SYTL2	intron	-5.94	1.29	4.05E-06
rs76472108	1	241304791	G/C	0.98	RGS7	intron	-6.07	1.54	7.71E-05
rs148974495	7	1577796	G/A	0.98	TMEM184A	downstream	-6.32	1.58	6.06E-05
rs183544012	3	53747902	C/T	0.98	CACNA1D	intron	-6.46	1.57	3.93E-05
rs72766128	9	115417463	C/T	0.98	KIAA1958	intron	-6.71	1.64	4.14E-05
rs150455300	4	79007274	C/T	0.98	FRAS1	intron	-7.47	1.92	9.69E-05
rs76156580	8	4474130	C/A	0.99	CSMD1	intron	-8.53	2.16	7.98E-05
rs148372781	3	124555852	G/A	0.99	ITGB5	intron	-8.64	2.2	8.50E-05

$LVMh^{2.7}$ association was evaluated using a linear regression analysis under an additive model, adjusted for ancestry PCs, sex, SBP, DBP, serum creatinine, uNa24h, sPRA, BMI, HR, and echocardiography operator. To retrieve information about SNPs and their genomic context (the nearest gene) we used the hg19 assembly (National Center for Biotechnology Information 37). SNP indicates single nucleotide polymorphism; chr, chromosome; bp, base pair; and SE, standard error.

Table A3 The loci of SNPs with less clear or direct role in cardiac hypertrophy, found in our study and their functional role in experimental setting.

Lead SNPs	Locus gene	Gene description *(official gene name is reported in Italic)*
rs76472108	RGS7	*Regulator Of G-Protein Signaling 7.* Member of G-protein signaling regulator family that fine-tunes G protein-coupled receptor-induced signaling. Changes in the RGS protein expression and/or function in the heart often lead to pathophysiological changes and are associated with cardiac disease in animals and humans, including hypertrophy, fibrosis development, heart failure, and arrhythmias.[a]
rs17068332 rs76156580	CSMD1	*CUB and Sushi multiple domains protein 1.* Transmembrane protein belonging to the vacuolar-protein-sorting-13 family. CSMD1 gene was associated with peripheral arterial disease and metabolic syndrome.[b,c] Moreover CSMD1 gene was associated with increased risk of hypertension in two case-control studies in Korean cohorts[d].
rs6590636	CNTN5	*Contactin 5.* Member of the immunoglobulin superfamily, which mediates cell surface interactions during nervous system development and has been associated with atrial fibrillation and heart failure.[e]
rs13023211	FIGN	*Fidgetin.* Member of a family of ATPases associated with diverse cellular activities. Variants at this locus have been associated with blood pressure regulation and with pulse pressure and mean arterial pressure.[f,g]
rs9284436	HCN2	*Hyperpolarization Activated Cyclic Nucleotide-Gated Potassium Channel 2*. Voltage-gated potassium channel that contributes to spontaneous rhythmic activity in both heart and brain. HCN2 and HCN4 are the predominant HCN transcripts in ventricular cells under basal conditions. Several studies showed an increase in the mRNA levels of these two genes following the induction of hypertrophy, but other studies are in contrast with these results.[h]

a. Zhang P, Mende U. Functional role, mechanisms of regulation, and therapeutic potential of regulator of G protein signaling 2 in the heart. Trends Cardiovasc Med 2014; 24:85-93.

b. Koriyama H, Nakagami H, Katsuya T, Sugimoto K, Yamashita H, Takami Y, et al. Identification of evidence suggestive of an association with peripheral arterial disease at the OSBPL10 locus by genome-wide investigation in the Japanese population. J Atheroscler Thromb 2010; 17:1054-1062.

c. Nock NL, Wang X, Thompson CL, Song Y, Baechle D, Raska P, et al. Defining genetic determinants of the Metabolic Syndrome in the Framingham Heart Study using association and structural equation modelling methods. BMC Proc 2009; 3(suppl 7):S50.

d. Hong KW, Go MJ, Jin HS, Lim JE, Lee JY, Han BG,et al. Genetic variations in ATP2B1, CSK, ARSG and CSMD1 loci are related to blood pressure and/or hypertension in two Korean cohorts. J Hum Hypertens 2010; 24:367-372.

e. Larson MG, Atwood LD, Benjamin EJ, Cupples LA, D'Agostino RB Sr, Fox CS, et al. Framingham Heart Study 100K project: genome-wide associations for cardiovascular disease outcomes. BMC Med Genet 2007; 8(suppl 1):S5.

f. Kato N, Takeuchi F, Tabara Y, Kelly TN, Go MJ, Sim X, et al. Meta-analysis of genome-wide association studies identifies common variants associated with blood pressure variation in east Asians. Nat Genet 2011; 43:531-538.

g. Ehret GB, Munroe PB, Rice KM, Bochud M, Johnson AD, Chasman DI, et al. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. Nature 2011; 478:103-109.

h. Hofmann F, Fabritz L, Stieber J, Schmitt J, Kirchhof P, Ludwig A, et al. Ventricular HCN channels decrease the repolarization reserve in the hypertrophic heart. Cardiovasc Res 2012; 95:317-326.

FUNDING

The European Union (grant FP7-HEALTH-F4-2007-201550, HYPERGENES) and InterOmics (PB05 MIUR-CNR Italian Flagship Project) provided financial support for genotyping. The “Associazione per lo sviluppo della ricerca sull'ipertensione arteriosa e sulle malattie cardiovascolari - ONLUS” supports the Coordinating Centre in Sassari for the collection of phenotypes.

SCIENTIFIC PRODUCTS

1. Frau F, Zaninello R, Salvi E, Ortu MF, Braga D, Velayutham D, Argiolas G, Fresu G, Troffa C, Bulla E, Bulla P, Pitzoi S, Piras DA, Glorioso V, Chittani M, Bernini G, Bardini M, Fallo F, Malatino L, Stancanelli B, Regolisti G, Ferri C, Desideri G, Scioli GA, Galletti F, Sciacqua A, Perticone F, Degli Esposti E, Sturani A, Semplicini A, Veglio F, Mulatero P, Williams TA, Lanzani C, Hiltunen TP, Kontula K, Boerwinkle E, Turner ST, Manunta P, Barlassina C, Cusi D, Glorioso N. Genome-wide association study identifies CAMKID variants involved in blood pressure response to losartan: the SOPHIA study. Pharmacogenomics. 2014 Sep;15(13):1643-52.

2. Chittani M and Zaninello R, Lanzani C, Frau F, Ortu MF, Salvi E, Fresu G, Citterio L, Braga D, Piras DA, Carpini SD, Velayutham D, Simonini M, Argiolas G, Pozzoli S, Troffa C, Glorioso V, Kontula KK, Hiltunen TP, Donner KM, Turner ST, Boerwinkle E, Chapman AB, Padmanabhan S, Dominiczak AF, Melander O, Johnson JA, Cooper-Dehoff RM, Gong Y, Rivera NV, Condorelli G, Trimarco B, Manunta P, Cusi D, Glorioso N,

Barlassina C. TET2 and CSMD1 genes affect SBP response to hydrochlorothiazide in never-treated essential hypertensives. J Hypertens. 2015; 33:1301-9.

3. Chittani M and Glorioso N, Ferrandu P, Mureddu V, Argiolas G, Salvi E, M. Francesca Ortu, Zaninello R, Braga D, Barcella M, Lupoli S, Conti C, Troffa C, Glorioso V, Fresu G, Piras DA, Masala W, Cusi D, Barlassina C. Multiple gene variants are associated with left ventricular mass in untreated essential hypertension. In preparation.
4. Rizzi F, Conti C, Dogliotti E, Terranegra A, Salvi E, Braga D, Ricca F, Lupoli S, Mingione A, Pivari F, Brasacchio C, Barcella M, Chittani M, D'Avila F, Turiel M, Lazzaroni M, Soldati L, Cusi D, Barlassina C. PON1 gene variants, polyphenols and markers of cardiovascular disease in the ATHENA study. Accepted by The Journal of Translational Medicine.
5. Justice AE, Winkler TW, Feitosa MF, Graff M, Fisher VA, Young K, Barata L, Deng X , Czajkowski J, Hadley D, Ngwa JS, Ahluwalia TS, Chu AY, Heard-Costa NL, Lim E , Perez J, Eicher JD, Kutalik Z, Xue L, Mahajan A, Renström F, Wu J, Qi Q, Ahmad S, Alfred T, Amin T, Bielak LF, Böger C, Cadby G, Chittani M, et al. Genome-Wide Meta-Analysis Accounting for Smoking Behavior Identifies Novel Loci 2 for Obesity Traits. Submitted to Nature Genetics.
6. Salvi E, Wang Z, Rizzi F, Gong Y, Padmanabhan S, Hiltunen TP, Lanzani C, Zaninello R, Chittani M, Bailey KR, Barcella M, Chapman AB, Johnson JA, Dominiczak AF, Kontula KK, Manunta P, Glorioso N, Barlassina C, Cusi D, Boerwinkle E, Cooper-DeHoff RM, Turner ST. Genome-wide and Gene-based Meta-Analyses to identify loci influencing antihypertensive response to hydrochlorothiazide. In preparation.

ACKNOLEDGMENTS

I thank Prof. Mario Clerici for the opportunity to perform my PhD work at University of Milan. I express my deep gratitude to my tutor, Dr. Cristina Barlassina for her support and guidance throughout my research work. I am very thankful to Prof. Nicola Glorioso and his research group for giving me the opportunity to collaborate with them to this research project.

I would also like to thanks Dr. Erika Salvi and Dr. Francesca Frau for their friendly help during my work. I express my sincere thanks to all my colleagues of the Genomics and Bioinformatics platform at Health Science Department, Dr. Federica Rizzi, Dr. Sara Lupoli, Dr. Francesca D'Avila, Dr. Matteo Barcella, Dr. Daniele Braga, Dr. Costanza Conti, and Dr. Federico Tagliaferri.

Sincerely thanks to my parents and my sisters: so far, yet so close. Finally sincerely thanks to my family, the most exciting "project" I have realized during my three years of PhD course.