CORONARY ARTERY DISEASE: WHY WE SHOULD CONSIDER THE Y CHROMOSOME

Elsa Molina MSc, Elyse Michele Clarence BSc, Farah Ahmady BSc, Guat Siew Chew PhD, Fadi Joseph Charchar, PhD

School of Applied and Biomedical Sciences, Faculty of Science and Technology, Federation University, Mount Helen Campus, VIC, Australia

To whom correspondence should be addressed.

Prof. Fadi Charchar

Federation University Australia

Faculty of Science and Technology

Room YI 17, Y Building, Mount Helen Campus

P.O. Box 663, Mount Helen, VIC, 3353

Australia

Tel: +61 353 276 098

Email: f.charchar@federation.edu.au

Abstract

Coronary artery disease (CAD) is one of the leading causes of morbidity and mortality globally. Inthe last few years our understanding of the genetic and molecular mechanisms that promote CAD in individuals has increased with the advent of the genome era. This complex inflammatory disease has well-defined environmental risk factors however, in the last ten years, studies including genome-wide association studies (GWAS) have clearly demonstrated a genetic influence on CAD. Recently, studies on the human Y chromosome have also demonstrated that genetic variation within the male-specific region of the Y chromosome (MSY) could play a part in determining cardiovascular risk in men, confuming the notion that the increased risk for CAD in men cannot be fully explained through common CAD risk factors. Here, we review the literature about the pathophysiology of CAD, its potential causes and environmental risk factors known so far. Furthermore, we review the genetics of CAD, especially the latest discoveries regarding the implication of the Y chromosome, the most underexplored portion of the human genome to date, highlighting methods and difficulties arising in this research field, and discussing the importance of considering the Y chromosome into CAD research.

Introduction

Coronary Artery Disease (CAD), also known as Coronary Heart Disease (CHO) or Ischaemic Heart Disease (IHD), is the most common type of cardiovascular disease and is the major cause of morbidity and mortality in the world according to the last report of the Global Burden of Disease [1,2]. Indeed, CAD disrupts the oxygen-rich blood flow to the heart, making it the first cause of 'years of life lost' in developed countries and second in developing countries after pulmonary respiratory infections. Although life expectancy has been extended in the last decade, cardiovascular disease risk substantially increases with age, creating a heavy burden of morbidity and mortality [3,4].

CAD occurs when the arteries of the heart, which are known as the coronary arteries are damaged from plaques accumulating on the arterial wall. Over time, this buildup of plaques progressively hardens and narrows the blood vessels, a process known as atherosclerosis [5]. As a consequence, thrombosis of the vessels or stenosis can occur and lead to angina pectoris and / or myocardial infarction [6]. Atherosclerosis is a complex inflammatory disease with well-defined environmental risk factors but those risks cannot be the only explanation; it is at that point that genetics enters in the arena. Here, we will review the latest discoveries regarding the genetics of CAD and the implication of the human Y chromosome, which is too often ignored by researchers but could potentially be the key to understanding the CAD prevalence differences between men and women.

Pathogenesis of CAD and common risk factors

As CAD is a multistep and chronic disease, the build-up of plaque occurs over many years and may start in childhood [7]. Ifthe plaque ruptures, fragments stick to the site of the injury and may clump together to form blood clots, which can further narrow the arteries and worsen the angina. If a clot becomes large enough, it can mostly or completely block the artery resulting in a heart attack, stroke, or even sudden death [6].

CAD is a multifactorial and complex late-onset disease which originates from a complicated interplay of environmental and genetic factors. The environmental risk factors could influence the progression of the atherosclerotic plaque by interacting with the endothelium resulting endothelial dysfunction. The

85 latter is thought to be triggered by risk factors such as lipid disturbances (high levels of low-density
86 lipoprotein (LDL) and low level of high density lipoprotein (HDL)), hypertension, diabetes, obesity,
87 cigarette smoking, elevated plasma homocysteine concentrations, lack of physical activity, aging,
88 hereditary, and sex [7,8]. Currently the pathogenesis of CAD is not fully understood with the molecular
89 mechanisms that promote CAD in individuals affected by these environmental factors remaining
90 unclear. We know that atherosclerosis is driven by a chronic inflammatory process, elicited in part by
91 subendothelial lipoprotein retention and involving innate and adaptive immune responses [9]. Indeed,
92 lipid disturbances and other risk factors are thought to cause endothelial injury resulting in monocyte
93 adhesion and migration to the intima, as well as the release of cytokines and growth factors. These
94 include platelet-derived growth factor (PDGF) which leads to smooth muscle cells migration to the
95 intima and proliferation (Fig.A). The recruitment of activated macrophages and T cells into and within
96 the atherosclerotic lesions is guided by endothelial leukocyte adhesion molecules and chemoattractants
97 [10]. Within the intima, smooth muscle cells produce an extracellular matrix including collagen and
98 proteoglycans. LDL particles travelling in the blood and carrying cholesterol and triglycerides from the
99 liver to other body tissues get through the endothelium layer due to their size and their density, and
100 become oxidised. After migration to the sub-endothelial space, monocytes differentiate into
101 macrophages which are able to ingest oxidized-LDL, forming specialized foam cells. Macrophages are
102 not able to process the oxidized-LDL, and ultimately grow and then rupture, depositing a greater
103 amount of oxidized cholesterol into the artery wall. This triggers the recruitment of more monocytes,
104 thus increasing the inflammation and continuing the cycle. This inflammation leads to subendothelial
105 accumulation of fatty substances called atheromatous plaques [1O]. Interestingly, the pathology of
106 atherosclerosis is apparently indistinguishable and independent of the risk factor, or combination of
107 risk factors associated with disease progression. This observation suggests that the pro-atherogenic
108 pathways associated with each risk factor converge on a common molecular mechanism [11].
109 Furthermore, it has been shown that the Herpes virus infection is associated with atherosclerosis [12]
110 with cytomegalovirus infection also being a risk factor for increased arterial blood pressure, and a co-
111 factor in aortic atherosclerosis [13]. As many as 50% of patients with atherosclerosis lack currently

identified risk factors, an observation suggesting that additional factors predisposing to atherosclerosis are as yet undetected [14].

The genetics of CAD

It is now well known that heritability as risk factor for CAD should not be excluded from studies into its etiology. Indeed, CAD is a highly heritable trait, with genetic and environmental factors accounting for similar proportions of individual susceptibility [15,16]. According to the Framingham Offspring Study, the age-specific incidence of CAD is increased approximately two-fold in subjects with a family history of premature disease [17]. To date, GWAS have been able to identify more than 90 genes within various chromosomes that are involved in the pathogenesis of CAD [18-27] as summarised in **Table 1.** From the protein-coding genes in Table 1, *STRING* (software version 10.0) was used to highlight the protein-protein interactions between them **(Fig.B).** As expected, one major cluster showed up with stronger associations between the proteins APOE, APOAl, APOB, LDLR, LPA, LPL and PCSK9 which are all proteins involved in lipid metabolism. However, interestingly, 30% of the genes do not show any interactions, suggesting a field to be studied further in CAD. Moreover, among the genes found by GWAS **(Table 1),** it appears that majority of the risk loci harbor genes previously unknown to be involved in atherosclerosis. Indeed, only 15% of the identified CAD risk loci work through known risk factors, such as lipids and blood pressure, implying that key pathways leading to CAD are yet to be discovered [27].

In 2007, the first GWAS in relation to CAD was published, identifying what is still the most genomic susceptible locus known for CAD heritability within the intergenic non-coding region of chromosome 9p21[28]. This locus contains a long non-coding ribonucleic acid (lncRNA), referred to as antisense non-coding RNA in the INK.4 locus, commonly known as ANRIL **(Table 1),** as reviewed in [29,30]. So far, ANRIL is the most replicated marker of CAD, independent from the conventional risk factors and its expression is correlated with atherosclerotic lesions. This lncRNA is expressed in tissues and cell types affected by atherosclerosis, such as primary coronary smooth muscle cells, vascular

endothelial cells, human monocyte-derived macrophage cells and RNA extracted from carotid and arterectomy [31]. Notably, an increased expression of ANRIL transcripts was found to be directly correlated with the severity of atherosclerosis [32,33]. Subsequent studies revealed that this locus is related to a broad spectrum of vascular phenotypes, including CAD and myocardial infarction [18,34,35], coronary artery calcification [36], peripheral artery disease [37,38], and abdominal aortic aneurysm [39]. However, despite the potential importance of this lncRNA to vascular disease, the pathophysiology underlying the link between ANRIL and CAD currently remains unknown.

Taking the aforementioned studies into consideration, it has been shown that the increased risk for CAD cannot be fully explained through the conventional risk factors.

CAD **and the human** Y chromosome

The human Y chromosome is one of two sex chromosomes, also known as allosomes. Of all chromosomes in human genome, the haploid Y chromosome contains the smallest number of genes. To date, over 200 Y-linked genes have been identified [40] that encode about 27 distinct proteins [41,42]. Its major part, the male-specific region (MSY), constitutes '"95% of its length, and does not recombine with the other sex chromosome (the X chromosome) during meiosis, and is inherited as an indivisible unit from fathers to sons [41]. The fundamental biological role of the human Y chromosome is thought to impart male characteristics [43]. However, there is also data that links the Y chromosome to cardiovascular diseases. Indeed, CAD is predominately associated with males with a 3:1ratio of men to women [44,45] with males commonly developing CAD nine years earlier than women [46]. Moreover, polysomy of the Y chromosome (XYY karyotype) was linked to increased cardiovascular mortality [47], with associations found between single nucleotide polymorphisms (SNPs) of the MSY and blood pressure, circulating concentrations of total cholesterol, LDL cholesterol, proatherogenic B-phenotype of LDL cholesterol molecules, and paternal history of coronary artery disease [48-51]. Although not all studies have replicated these associations, the accumulated evidence lends support to the notion that genetic variation within the MSY could play a part in determining cardiovascular risk in men [52, 53].

Due to the haploid nature of the Y chromosome, the usual methods of analysis (such as GWAS) cannot be employed to investigate variations, and this is the reason why the Y chromosome is routinely excluded from large-scale GWAS. The Y chromosome is therefore the most underexplored portion of the human genome to date. To bypass this difficulty, Charchar *et al.* [54] performed an analysis of the Y chromosome phylogenetic tree. lbis strategy is defined by a series of biallelic SNPs which enable the MSY to be partitioned into 20 major haplogroups (non-recombining portions of DNA [55]) that descend from a common ancestor, Y-chromosomal Adam [42]. lbis study was the first to evaluate associations between main European Y chromosome lineages and coronary artery disease, as well as its underlying risk factors. Results showed that men who inherit haplogroup I (one of the most common Y chromosome types in Europe) from their male ancestors have a 50% increased risk of developing coronary artery disease compared to men with other Y chromosome haplogroups. lbis study also demonstrated that the effect of haplogroup I on CAD is not mediated by traditional cardiovascular risk factors (such as age, body-mass index (BMI), blood pressure, lipids, diabetes, smoking, alcohol consumption, socioeconomic status, or circulating concentrations of C-reactive protein) but might be mediated through a genetically programmed profile of immunity and response to inflammation [54]. lbis makes haplogroup I of the Y chromosome one of the strongest common genetic risk factors of CAD known to date.

In order to confirm these findings and identify the causative variants underlying the increased susceptibility to CAD in carriers of haplogroup I, a total of 1988 biologically unrelated men from 4 white European populations were genotyped, using 11 Y chromosome SNPs and classified into 13 of the most common European haplogroups [56]. The results of this study confirmed that haplogroup I of the Y chromosome, which has previously been linked to an increased risk of CAD, is not associated with conventional cardiovascular and metabolic risk factors in young men from the general white European population. lbis study also showed for the first time that CAD predisposing haplogroup I of the Y chromosome is associated with the downregulation of two MSY genes; ubiquitously transcribed tetratricopeptide repeat, Y-linked gene *(UTY)* and protein kinase, Y-linked, pseudogene *(PRKY)* within macrophages. The *UTY* gene encodes a protein containing tetratricopeptide repeats, involved in protein-protein interactions. lbis protein acts as an immune related minor histocompatibility antigen that may induce graft rejection of male stem cell grafts [41,57]. The dysregulated expression of this gene in

195 macrophages of subjects with haplogroup I may lead to increased risk of CAD (Fig.A). This is also
196 based on an emerging role for *UTY* in both the immune system [56], haematopoiesis [58] and
197 cardiovascular system development [59,60], which are important processes that contribute to the
198 development of CAD [60,61]. Recently published data by Wang *et al.* [62] on the role of *UTY* revealed
199 that it is essential for progression of cardiac development and that it associates with cardiovascular
200 specific transcription factors to regulate downstream target genes. Data on *Uty* mutant mice by Shpargel
201 *et al.* [63] show that *Uty* is able to regulate gene activity through demethylase independent mechanisms.
202 Furthermore, we used *GIANT*, the new human tissue-specific network webserver [64] to highlight the
203 potential tissue-specific functional interactions of *UTY* with protein-coding genes in macrophages
204 (Fig.C). According to the functional network generated the data predicts that *UTY* interacts with the
205 following genes: DDX3Y, EIFIAY, KDM5D, RPS4Yl, USP9Y, and ZFY in macrophages.
206 Interestingly, these 6 protein-coding genes are only located on the Y chromosome. These results
207 reinforce the idea that the Y chromosome should be considered in future works in relation to CAD. In
208 regards to *PRKY*, no studies have yet been published in relation to its involvement in cardiovascular
209 processes.
210
211
212 Concluding Perspectives
213 Despite a large advancement in our knowledge of CAD genes due to GWAS, studies regarding Y
214 chromosome linked-genes in relation to CAD are still sparse. The involvement and function of both
215 autosomal and sex chromosome genes in an atherosclerotic context need to be further elucidated. So far,
216 no additional studies have been published on *PRKY* and *UTY* in humans. In mice, a study linking the Y
217 chromosome, HDL-cholesterol levels, and *Uty* has been recently published [65]. This study confirmed
218 the effect of the Y chromosome on plasma HDL-cholesterol levels in mice by identifying several
219 variants associated with plasma HDL-cholesterol levels. The results notably showed that the variation
220 rs46947134 (a nonsynonymous SNP) in *Uty* was significantly associated with plasma HDL-cholesterol
221 levels, however, it is still unknown whether the G/C variants in mouse *Uty* are associated with these
222 expression levels [65].

Despite these breakthroughs, the exact cause of atherosclerosis still remains unknown, and the biological mechanisms underlying the association between CAD and human Y chromosome remains to be discovered. Further studies should focus on functional characterization of the biological underpinnings of the association between haplogroup I and *UTYIPRKY* expression in order to fully elucidate the mechanisms of increased susceptibility to CAD amongst men with haplogroup I of the Y chromosome. This would help us to better understand the complex interplay between the human Y chromosome, immunity, and cardio vascular disease; and maybe discover new diagnostic markers and therapeutic targets for CAD in men in the future.

Acknowledgements

This work is supported by the Higher Degree Research (HOR) at Federation University Australia and the Collaborative Research Network (CRN) at Ballarat, VIC, Australia.

References

[1] WHO. (2015, January). *Cardiovascular diseases (CVDs), Fact sheet No 317*. Retrieved March 29, 2015, from World Health Organization: http://www.who.int/mediacentre/factsheets/fs317/en/

[2] GBD Mortality and Causes of Death Collaborators. Global, regional, and national levels of age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990--2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2015;385: 117-71.

[3] Berthold HK, Gouni-Berthold I. Lipid-lowering drug therapy in elderly patients. Curr Pharm Des 2011;17:877-93.

[4] Lakatta EG, Levy D. Arterial and cardiac aging: major shareholders in cardiovascular disease enterprises: Part I: aging arteries: a "set up" for vascular disease. Circulation 2003;107: 139-146.

[5] Sakakura K, Nakano M, Otsuka F, Ladich E, Kolodgie FD, Virmani R. Pathophysiology of atherosclerosis plaque progression. Heart Lung Circ. 2013; 22: 399-411.

[6] Roger VL, Go AS, Lloyd-Jones DM, Benjamin EJ, Berry JD, Borden WB, et al. Heart Disease and Stroke Statistics: A Report From the American Heart Association. Circulation 2012;125:e2-e220.doi: 10.1161.

[7] Heart Foundation. (n.d.). *Data and statistics.* (Heart Foundation) Retrieved April I, 2015, from HeartFoundation:http:*/lwww.hearifoundation.org.aul information:for-professionals/data-andstatistics/Pages/default.ospx.*

[8] Netter's Cardiology, 2e (Netter Clinical Science) 2nd Edition by Marschall S.Runge (Editor), George A. Stouffer (Editor), Cam Patterson (Editor), Frank H. Netter (Illustrator) Aug.IO, 2010

[9] Williams KJ, Tabas I. The Response-to-Retention Hypothesis of Early Atherogenesis. Arterioscler Thromb Vase Biol 1995;15:551-61.

[10] Kumar V, Abbas AK, Kausto N. Pathologic Basis of Disease. 7th ed. Eds. Robbins and Cotran;2005.

[II] Zeadin MG, Petlura CI, Werstuck GH. Molecular Mechanisms Linking Diabetes to the Accelerated Development of Atherosclerosis. Can J Diabetes 2013;37:345-50.

277 [12] Hsu HY, Nicholson AC, Pomerantz KB, Kaner RJ, Hajjar DP. Altered cholesterol trafficking in
278 herpesvirus-infected arterial cells. Evidence for viral protein kinase-mediated cholesterol accumulation.
279 J Biol Chem 1995;270:19630--7.
280 [13] Cheng J, Ke Q, Jin Z, Wang H, Kocher 0, Morgan JP, et al. Cytomegalovirus Infection Causes an
281 Increase of Arterial Blood Pressure. PLoS Pathog 2009;5:doi:I0.1371.
282 [14] Genco CA, Gibson FC. Infection and Atherogenesis, edited by Joseph Loscalzo. Molecular
283 Mechanisms of Atherosclerosis. Eds. Taylor & Francis; 2005.Chap.13.
284 [15] Stylianou IM, Bauer RC, Reilly MP, Rader DJ. Genetic basis ofatherosclerosis: insights from mice
285 and humans. Circ Res 2012;110:337-355.
286 [16] Kathiresan S, Srivastava D. Genetics of human cardiovascular disease. Cell 2012;148:1242-1257.
287 [17] Lloyd-Jones DM, Nam BH, D'Agostino RB, Sr, Levy D, Murabito JM, Wang TJ, et al. Parental
288 cardiovascular disease as a risk factor for cardiovascular disease in middle-aged adults a prospective
289 study of parents and offspriog. JAMA 2004;291:2204-11.
290 [18] Zdravkovic S, Wienke A, Pedersen NL, Marenberg ME, Yashin AI, De Fair U. Heritability of death
291 from coronary heart disease: a 36-year follow-up of 20 966 Swedish twins. J Intern Med 2002;252:247-
292 254.
293 [19] Clarke R, Peden JF, Hopewell JC, Kyriakou T, Goel A, Health SC, et al. (2009) Genetic variants
294 associated with Lp(a) lipoprotein level and coronary disease. N Engl J Med 2009;361:2518--2528.
295 [20] Soranzo N, Spector TD, Mangino M, Kiihnel B, Rendon A, Teumer A, et al. A genome-wide meta-
296 analysis identifies 22 loci associated with eight hematological parameters in the HaemGen consortium.
297 Nat Genet 2009;41: 1182-1190.
298 [21] Butterworth AS, Braund PS, Farrall M, Hardwick RJ, Saleheen D, Peden JF, et al. Large-scale
299 gene-centric analysis identifies novel variants for coronary artery disease. PLoS Genet
300 2011;7:e1002260.
301 [22] Peden JF, Hopewell JC, Saleheen D, Chambers JC, Hager J, Soranzo N, et al. A genome-wide
302 association study in Europeans and South Asians identifies five new loci for coronary artery disease.
303 Nat Genet 2011;43:339-44.

[23] Peden JF, Farrall M. Thirty-five common variants for coronary artery disease: the fruits of much collaborative labour. Hum Mol Genet 2011;20: 198-205.

[24] Schunkert H, Konig IR, Kathiresan S, Reilly MP, Assimes TL, Holm H, et al. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. Nat Genet 2011;43:333-338.

[25] Lu X, Wang L, Chen S, He L, Yang X, Shi Y, et al. Genome-wide association study in Han Chinese identifies four new susceptibility loci for coronary artery disease. Nat Genet 2012;44:890--894.

[26] Roberts R, Stewart AF. Genes and Coronary Artery Disease: Where Are We? J Am Coll Cardiol 2012;60: 1715-1721.

[27] Ruth McPherson. Genome-Wide Association Studies of Cardiovascular Disease in European and Non-European Populations. Curr Genet Med Rep 2014;2:1-12.

[28] McPherson R, Pertsemlidis A, Kavaslar N, Stewart A, Roberts R, Cox DR, et al. A common allele on chromosome 9 associated with coronary heart disease. Science 2007;316: 1488--1491.

[29] Holdt LM, Teupser D. Recent studies of the human chromosome 9p21 locus, which is associated with atherosclerosis in human populations. Arterioscler Thromb Vase Biol 2012;32:196-206.

[30] Congrains A, Kamide K, Ohishi M, Rakugi H. ANRIL: molecular mechanisms and implications in human health. Int J Mol Sci 2013; 14:1278-1292.

[31] Broadbent HM, Peden JF, Lorkowski S, Goel A, Ongen H, Green F, et al. PROCARDIS consortium. Susceptibility to coronary artery disease and diabetes is encoded by distinct, tightly linked SNPs in the ANRIL locus on chromosome 9p. Hum Mol Genet 2008;17:806-814.

[32] Holdt LM, Beutner F, Scholz M, Gielen S, Gabel G, Bergert H, et al. ANRIL expression is associated with atherosclerosis risk at chromosome 9p21. Arterioscler Thromb Vase Biol 2010;30:620--627.

[33] Schonrock N, Harvey RP, Mattick JS. Long noncoding RNAs in cardiac development and pathophysiology. Circ Res 2012;10:1349-62.

[34] Samani NJ, Erdmann J, Hall AS, Hengstenberg C, Mangino M, Mayer B, et al. WTCCC and the Cardiogenics Consortium. Genomewide association analysis of coronary artery disease. N Engl J Med 2007;357:443-453.

[35] Helgadottir A, Thorleifsson G, Manolescu A, Gretarsdottir S, Bionda! T, Jonasdottir A, et al. A common variant on chromosome 9p21 affects the risk of myocardial infarction. Science 2007;316:1491-1493.

[36] O'Donnell CJ, Kavousi M, Smith AV, Kardia SL, Feitosa MF, Hwang SJ, et al. CARDioGRAM Consortium. Genome-wide association study for coronary artery calcification with follow-up in myocardial infarction. Circulation 2011;124:2855-2864.

[37] Cluett C, McDermott MM, Guralnik J, Ferrucci L, Bandinelli S, Miljkovic I, et al. The 9p21 myocardial infarction risk allele increases risk of peripheral artery disease in older people. Circ Cardiovasc Genet 2009;2:347-353.

[38] Murabito JM, White CC, Kavousi M, Sun YV, Feitosa MF, Nambi V, et al. Association between chromosome 9p21 variants and the ankle-brachia! index identified by a meta-analysis of 21 genome-wide association studies. Circ Cardiovasc Genet 2012;5:100--112.

[39] Helgadottir A, Thorleifsson G, Magnusson KP, Gretarsdottir S, Steinthorsdottir V, Manolescu A, et al. The same sequence variant on 9p21 associates with myocardial infarction, abdominal aortic aneurysm and intracranial aneurysm. Nat Genet 2008;40:217-224.

[40] Bethesda, Maryland: National Center for Biotechnology Information. Genes and Disease. Bookshelf ID: NBK22266.

[41] Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, Cordum HS, Hillier L, Brown LG, et al. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. Nature 2003;423:825-37.

[42] Jobling MA, Tyler-Smith C. The human Y chromosome: an evolutionary marker comes of age. Nat Rev Genet 2003;4:598--612.

[43] Graves JA, Koina E, Sankovic N. How the gene content of human sex chromosomes evolved. Curr Opin Genet Dev 2006;16:219-24.

[44] Madaric J, Vulev I, Bartunek J, Mistrik A, Verhamme K, De Bruyne B, et al. Frequency of abdominal aortic aneurysm in patients >60 years of age with coronary artery disease. Am J Cardiol 2005;9:1214-6.

[45] Cardiovascular Disease: Australian Facts 2011. Cardiovascular Disease Series. Canberra, ACT, Australia Australian Institute of Health and Welfare, 2011.

[46] Yusuf S, Hawken S, Ounpuu S, Dans T, Avezwn A, Lanas F, et al. INTERHEART Study Investigators. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. Lancet 2004;364:937-52.

[47] Higgins CD, Swerdlow AJ, Schoemaker MJ, Wright AF, Jacobs PA, UK Clinical Cytogenetics Group. Mortality and cancer incidence in males with Y polysomy in Britain: a cohort study. Hum Genet 2007;121:691-96.

[48] Shanker RR, Charchar FJ, Eckert GJ, Saha C, Tu W, Dominiczak AF, et al. Studies of an association in boys of blood pressure and the Y chromosome. Am J Hypertens 2007;20:27-31.

[49] Charchar FJ, Tomaszewski M, Lacka B, Zakrzewski J, Zukowska-Szczechowska E, et al. Association of the human Y chromosome with cholesterol levels in the general population. Arterioscler Thromb Vase Biol 2004;24:308--12.

[50] Charchar FJ, Tomaszewski M, Padmanabhan S, Lacka B, Upton MN, Inglis GC, et al. The Y chromosome effect on blood pressure in two European populations. Hypertension 2002;39:353-56.

[51] Ellis JA, Stebbing M, Harrap SB. Association of the human Y chromosome with high blood pressure in the general population. Hypertension 2000;36:731-33.

[52] Charchar FJ, Tomaszewski M, Strahom P, Champagne B, Dominiczak AF. Y is there a risk to being male? Trends Endocrinol Metab 2003;14:163--68.

[53] Ely D, Underwood A, Dunphy G, Boehme S, Turner M, Milsted A. Review of the Y chromosome, Sry and hypertension. Steroids 2010;75:747-53.

[54] Charchar FJ, Bloomer LD, Barnes TA, Cowleyy MJ, Nelson CP, Wang Y, et al. Inheritance of coronary artery disease in men: an analysis of the role of the Y chromosome. Lancet 2012;379:915-922.

[55] Y Chromosome Consortium. A nomenclature system for the tree of human Y-chromosomal binary haplogroups. Genome Res 2002;12:339--348.

[56] Bloomer LD, Nelson CP, Eales J, Denniff M, Christofidou P, Debiec R, et al. Male-specific region of the Y chromosome and cardiovascular risk. Arteriorscler Thromb Vasc Biol 2013;33: 1722-1727.

386 [57] Mortensen BK, Rasmussen AH, Larsen ME, Larsen MV, Lund 0, Braendstrup P, et al.
387 Identification of a novel UTY-encoded minor histocompatibility antigen. Scand J Immunol
388 2012;76:141-50.
389 [58] Thieme S, Gyarfas T, Richter C, Ozhan G, Fu J, Alexopoulou D, et al. The histone demethylase
390 UTX regulates stem cell migration and hematopoiesis. Blood 2013; 121:2462-2473.
391 [59] Welstead GG, Creyghton MP, Bilodeau S, Cheng AW, Markoulaki S, Young RA, et al. X-linked
392 H3K27me3 demethylase Utx is required for embryonic development in a sex-specific manner. Proc Natl
393 Acad Sci USA 2012;109:13004--9.
394 [60] Soehnlein 0, Swirski FK.. Hypercholesterolemia links hematopoiesis with atherosclerosis. Trends
395 Endocrinol Metab 2013;24:129-136.
396 [61] Eriksson EE. Leukocyte recruitment to atherosclerotic lesions, a complex web of dynamic cellular
397 and molecular ioteractions. Curr Drug Targets Cardiovasc Haematol Disord 2003;3:309-325.
398 [62] Wang C, Lee JE, Cho YW, Xiao Y, Jin Q, Liu C, et al. UTX regulates mesoderm differentiation of
399 embryonic stem cells iodependent of H3K27 demethylase activity. Proc Natl Acad Sci U S A
400 2012;109:15324--15329.
401 [63] Shpargel KB, Sengoku T, Yokoyama S, Magnuson T. UTX and UTY demonstrate histone
402 demethylase-independent function io mouse embryonic development. PLoS Genet 2012;8:e1002964
403 doi: 10.1371.
404 [64] Greene CS, Krishnan A, Wong AK, Ricciotti E, Zelaya RA, Himmelstein DS, et al. Understandiog
405 multicellular function and disease with human tissue-specific networks. Nat Genet. 2015;47:569-76.
406 doi: 10.1038/ng.3259.
407 [65] Suto J, Satou K. Effect of the Y chromosome on plasma high-density lipoprotein-cholesterol levels
408 inY-chromosome-consomic mouse strains. BMC Res Note 2014;7:393.
409

410

411

412 Table

413

414 Table 1:Genome Wide Association Study (GWAS) genes found to be involved in CAD.

415

Chromosome	Location	Gene Name	Full Name	Gene Function
Chr 1	lp13	*SORTI*	Sortilin 1	Sorting receptor in Golgi compartment
	lpl3	*PSRCI*	Proline/serine-rich coiled-coil 1	Mitosis
	lp21	*CELSR2*	Cadherin, EGF LAG seven-pass G-type receptor 2	Cell to cell signaling during nervous system formation
	lp32	*PPAP2B*	Phosphatidic acid phosphatase tvoe 2B	Conversion of phosphatidic acid to diacylitlvcerol
	lp32	*PCSK9*	Proprotein convertase subtilisin/kexin type 9	Regulating plasma cholesterol homeostasis
	lq21	*IL6R*	Interleukin-6 receptor	Regulation of the immune response, hematoooiesis
	lq21	*AQPIO*	Aquaporin 10	Water-selective channel
	lq41	*MIA3*	Melanoma inhibitory activity family 3	Loads COL7Al at endoplasmic reticulum exit sites
	lq43	*FMN2*	Formin 2	Organization of the actin cytoskeleton and cell polarity
	lq44	*OR13GI*	Olfactory receptor 1301	Odorant receptor
Chr 2	2pl1	*VAMPS*	Vesicle-associated membrane protein 8	Autophagosome membrane fusion with lysosome
	2pl1.2	*YAMP5*	Vesicle-associated membrane orotein *5*	Myogenesis
	2p21	*ABCG5*	ATP-binding cassette sub-family G (WHITE), member 5	Selective transport of dietary cholesterol
	2p21	*ABCG8*	ATP-binding cassette sub-family G (WHITE), member 8	Stimulate the excretion of cholesterol and sterols intobile, transport of sterols back into the intestinal lumen
	2p24	*APOB*	Apolipoprotein B	Binding and internalization of LDL particles
	2ql3	*ILJFJO*	Interleukin I family, member 10 (theta)	Regulate adapted and innate immune resoonses
	2q22	*ZEB2*	Zinc finger E-box binding homeobox 2	Transcriptional inhibitor
	2q33	*WDR12*	WD repeat domain 12	Cell cycle progression, signal transduction, apoptosis
Chr 3	3q22	*MRAS*	Muscle RAS oncogene homolog	Cell growth and differentiation
Chr 4	4q22	*ABCG2*	ATP-binding cassette sub-family G (WHITE),member 2 (Junior blood III'OUD)	Xenobiotic transporter which may play a major role in multi-drug resistance
	4q31	*EDNRA*	Endothelin receptor type A	Associated with Gproteins
	4q32	*GUCYJA3*	Guanylate cyclase 1, soluble, alpha 3	Conversion of GTP to 3',5'-cyclic GMP and DvroDhosphate
	4q32.3	*PALLD*	Palladin, cytoskeletal associated orotein	Organisation the actin cytoskeleton
	4q32.3	*RPL9P16*	Ribosomal protein L9 pseudogene 16	Unknown

Chr 5	5q14.1	AP3B1	Adaptor-related protein complex 3, beta 1 subunit	Organelle biogenesis
	5q31	*SLC22A4*	Solute carrier family 22 (organic cation/Zwitterion transporter), member 4	Organic cation transporter and plasma integral membrane protein
Chr 6	6p21	*KCNK5*	Potassium channel, two pore domain subfamily K, member 5	Renal potassium channel
	6p21.3	*NFKBIL1*	Nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor-like 1	Unknown
	6p21.3	*DDX39B*	DEAD (Asp-Glu-Ala-Asp) box polypeptide 398	Splicing factor
	6p21.31	*ANKSJA*	Ankyrin repeat and sterile alpha motif domain containing IA	Controls cell migration and neurite retraction through regulation of EPHA8 receptor tyrosine kinase sil!Dalling
	6p21.33	*MCCDI*	Mitochondrial coiled-coil domain 1	Unknown
	6p21.33	*SNORD117*	Small nucleolar RNA, *CID* boc 117	Unknown
	6p21.33	*RPL15P4*	Ribosomal protein LI5 pseudogene 4	Unknown
	6p21.33	*LOC100287329*	Uncharacterized LOC100287329	Unknown
	6p24	*PHACTR1*	Phosphatase and actin regulator 1	Reorganization of actin skeleton
	6p24.1	*ADTRP*	Androgen-dependent TFPI-regulating protein	Regulates the cell expression and the activity of the inhibitor TFPI in endothelial cells (in *vitro)*
	6q22	*ROSI*	ROS proto-oncogene 1	Growth or differentiation factor recentor
	6q23.2	*TCF21*	Transcription factor 21	Epithelial-mesenchymal interactions in kidney and lung morohogenesis
	6q25	*LPA*	Lipoprotein Lp(a)	Inhibits the activity of tissue-type plasminogen activator I
	6q25.1	*MTHFDIL*	Methylenetetrahydrofolate dehydrogenase (NADP+ deoendent) 1-like	Synthesis of tetrahydrofolate (THF) in the mitochondrion
	6q25.3	*SLC22AJ*	Solute carrier family 22 (organic cation transporter), member 3	Plasma integral membrane protein
	6q26	*PLG*	Plasminogen	Dissolves fibrin in blood clots and performs as aproteolytic factor in processes such as embryonic development, tissue remodeling, tumour invasion, and inflammation
	6q26	*LPAL2*	Lipoprotein, Lp(a)-like 2, pseudogene	Similar to Lp(a) but they are candidates for nonsense-mediated decay
Chr 7	7p21.1	*HDAC9*	Histone deacetylase 9	Transcriptional regulation, cell cycle progression, and develomnental events
	7q22	*COGS*	Component of oligomeric golgi complex 5	Norm.al Golgi function
	7q22.3	*BCAP29*	B-cell receptor-associated protein 29	Transport of membrane proteins from the endoplasmic reticulum to the Golgi
	7q32.2	*ZCJHCJ*	Zinc finger, C3HC-type containing 1	Regulates the onset of cell division

Chr 8	8p22	*LPL*	Lipoprotein lipase	Triglyceride hydrolase and ligand/bridging factor for receptor-mediated lipoprotein uptake
	8q24	*TRIBl*	Tribbles pseudokinase 1	Interacts with MAPK kinases and ree:ulates activation of MAP kinases
Chr 9	9p21	*CDKN2B-AS1 (ANRJL)*	CDKN2B antisense RNA 1	RNA molecule leading to epigenetic silencinst
	9q33	*TNC*	Tenascin C	Encodes an extracellular matrix protein
	9q34	*ABO*	ABO blood group (transferase A, alpha 1-3-N-acetylgalactosaminyltransferase ; transferase B, alpha 1-3-galactosyltransferase)	Protein basis for blood grouping
	9q34	*AGPAT1*	1-acylglycerol-3-phosphate 0-acyltransferase 2 (lysophosphatidic acid acyltransferase, beta)	Converts lysophosphatidic acid to phosphaditidic acid
	9q34	*EGFL7*	EGF-like-domain, multiple 7	Codes for a secreted endothelial cell protein that contains two epidermal growth factor-like domains
Chr 10	lOpll	*KIAA1462*	KIAA1462	Cell adhesion
	lOqll	*CXCL12*	Chemokine (C-X-C motif) ligand 12	Embryogenesis, immune surveillance, inflammation resnonse, tissue homeostasis
	10q23	*IFIT6P*	Interferon-induced protein with tetratricopeptide repeats 6	Unknown
	10q23	*LIPA*	Lipase A, lysosomal acid, cholesterol esterase	In the lysosome tocatalyze the hydrolysis of cholesteryl esters and triglycerides
	10q24	*CYP17Al*	Cytochrome P450, family 17, subfamily A, polypeptide 1	Produces progestins, mineralocorticoids, glucocorticoids, androgens, and estrogens
	10q24	*CNNM2*	Cyclin and CBS domain divalent metal cation transport mediator 2	Magnesium homeostasis. Mutations are associated with renal h. -esemia
Chr 11	llq22	*PDGFD*	Platelet derived growth factor D	Cell proliferation, cell migration, survival and chemotaxis. Involved in wound healing. Induces macrophage recruitment, increased interstitial pressure, and blood vessel maturation during angiogenesis
	llq23	*A.POA.l*	Apolipoprotein A-I	Reverse transport of cholesterol from tissues to the liver
	llq23	*ZNF259*	Zinc finger protein ZPRl	Signaling molecule that communicates proliferative growth signals from the cytoplasm to the nucleus
Chr 12	12pl3	*PRHI*	Proline-rich protein Haem subfamily 1	Provide protective and reparative environment for dental enamel
	12pl3	*PRR4*	Proline rich 4 (lacrimal)	Involved inprotective functions in the eye
	12p13	*TAS1R50*	Taste receptor, type 2, member 50	Mediate the perception of bitterness through a G protein-coupled second messenlrel' nathway
	12q24	*ALDH2*	Aldehyde dehydrogenase 2	Encodes a mitochondrial isoform

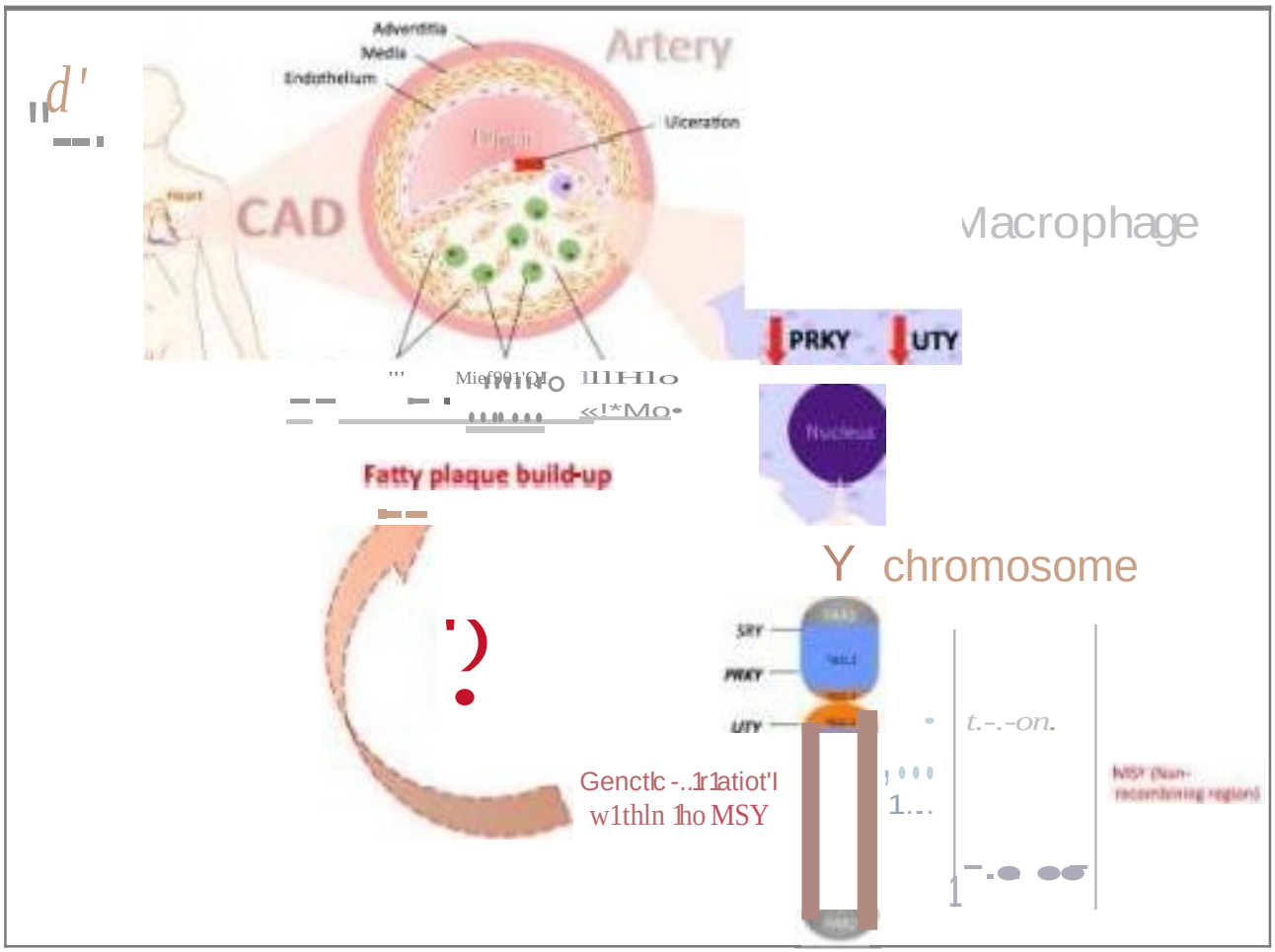
	12q24	*BIUP*	BRCA 1associated protein	Regulates nuclear targeting by retaining proteins with anuclear localization simal in the cytoplasm.
	12q24	*HNFJA*	HNFl homeobox A	Transcription factor
	12q24	*SHZB3*	SH2B adapter protein 3	Negative regulator of cytokine signaling. Plays a critical role in hematoooiesis
Chr 13	13ql2	*FLTl*	Vascular endothelial growth factor receptor 1	Embryonic vasculature development, angiogenesis regulation, cell survival and migration, macrophage function, chemotaxis
	13q34	*COL4Al*	Collagen alpha-I OV) chain	Inhibits angiogenesis
	13q34	*COL4AZ*	Collagen alpha-2 OV) chain	Inhibits angiogenesis, tumour growth, proliferation and migration of endothelial cells, reduces mitochondrial membrane potential, induces apoptosis
Chr 14	14q24-31	*CALMl*	Calmodulin 1 (phospharylase kinase, delta)	Regulates centrosom.e cycle and progression through cytokinesis
	14q32	*HHIPLl*	HHIP -like 1	Carbohydrate metabolic process
Chr 15	15q22	*SMAD3*	SMAD family member 3	Transforms growth factor-beta, involved in the regulation of carcino2e11esis
	15q25	*ADAMTS7*	ADAM metallopeptidase with thrombospondin type 1motif, 7	Degradation of COMP
	15q26	*FURJN*	FURIN (paired basic amino acid cleaving enzyme)	Codes for atype I membrane bound orotease
Chr 16	16q23	*CDH13*	Cadherin 13	Negative regulator of axon growth during neural differentiation. Protects vascular endothelial cells from apoptosis due to oxidative stress
Chr 17	17pll	*PEMT*	Phosphatidylethanolamine N-methyltransferase	Converts phosphatidylethanolamine to phosphatidylcholine by seauential methylation in the liver
	17pll	*RAil*	Retinoic acid-induced protein 1	Transcriptional regulator of circadian clock comoonents
	17pll	*RASDl*	Dexamethasonindiced Ras-related protein 1	Alterations in cell morphology, growth and cell-extracellular matrix interactions
	17pl3	*CLUH*	clustered mitochondria (cluA/CLUl) homolog	Regulates transport or translation of transcripts close to mitochondria
	17p13	*SMG6*	SMG6 nonsense mediated mRNA decay factor	Replication and maintenance of chromosome ends. Telomere regulation. Nonsense-mediated mRNA decay
	17q21	*GIP*	Gastrin inhibitory polypeptide	Potent stimulator of insulin secretion, poor inhibitor of gastric acid secretion
	17q21	*HAPl*	Huntingtin-associated protein 1	Codes for a protein that interacts with huntingtin, two cytoskeletal proteins, and a hepatocyte growth fact<r-regulated tyrosine kinase substrate
	17q21	*UBE2Z*	Ubiquitin-conjugating enzyme E2Z	Signalling pathways and apoptosis

Chr 19	19p13	*LDLR.*	Low-density lipoprotein receptor	Binds and transports LDL
	19pl3	*ZNF627*	Zinc finger protein 627	Transcriptional regulation
	19p13	*SMARCA4*	SWI/SNF related, matrix associated, actin dependent regulator of chromatin. subfamily a, member 4	Binds to BRCAI and regulates the expression of the tumorigenic protein CD44
	19q13	*APOE*	Apolipoprotein E	Mediates the binding, internalization, and catabolism of lipoprotein particles. Serves as a ligand for the LDL and apo-E recentors in henatic tissues
	19ql3	*HNRNPULI*	Heterogeneous nuclear nlxmucleonrotein U-like 1	Involved in nucleocytoplasmic RNA transoort
Chr 21	21q22	*KCNE2*	Potassium voltage-gated channel subfamily E member 2	Modulates the gating kinetics and enhances stability of the potassium channel complex
Chr 22	22ql2	*SEZ6L*	Seizure related 6 homolog (mouse)-like	Endoplasmic reticulum functions in neurons

434 -.....

43S

436 llpNA

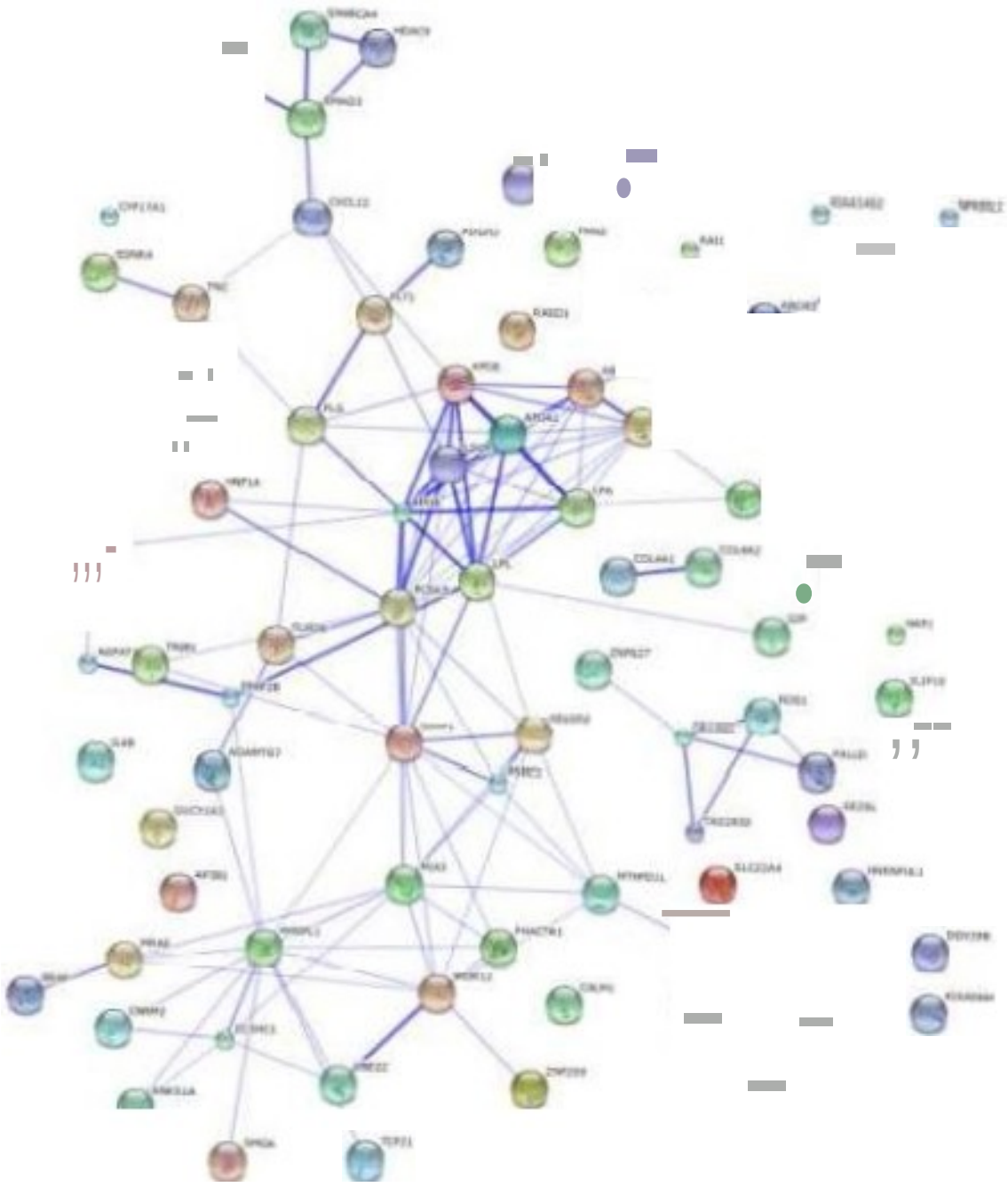


437

438

441

444 **Figure B**



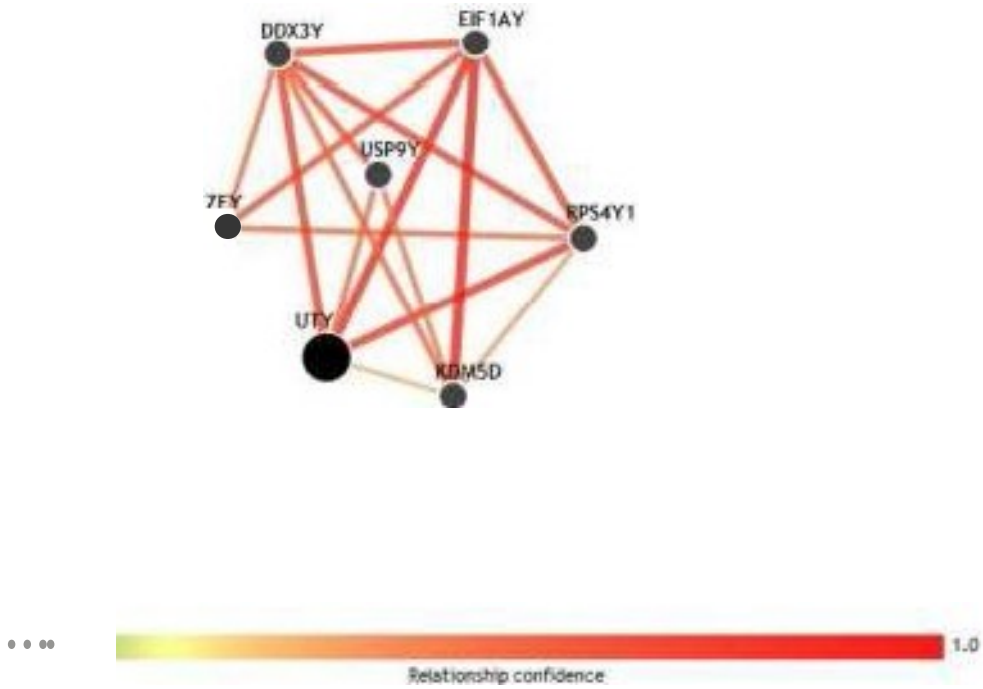

DDX3Y
EIF1AY
USP9Y
7FY
RPS4Y1
UTY
KDM5D
1.0
Relationship confidence

Figure Legends

Figure A: Schematic diagram of hypothetical links between a genetic variation within the MSY and the fatty plaque build-up in CAD.

Independent of traditional risk factors, a genetic variation within the male-specific region of the human Y chromosome (MSY) results in a downregulation of two genes: ubiquitously transcribed tetratricopeptide repeat, Y-linked gene (UTY) and protein kinase, Y-linked, pseudogene *(PRKY)* in macrophages of men with haplogroup I.1bistriggers an endothelial dysfunction resulting macrophages migrating to the intima, and the release of cytokines and growth factors which further leads to smooth muscle cells migrating to the intima and proliferating. Also, LDL particles travelling through the blood pass through the endothelium and become oxidized. Then, macrophages absorb the oxidised-LDL, which forms specialized foam cells, which grow and then rupture, depositing a greater amount of oxidized-LDL into the artery wall.

SRY: sex-determining region of the Y chromosome.

Regions of the human Y chromosome: AZFa, azoospermia factor a; AZFb, azoospermia factor b; AZFc, azoospermia factor c; PARI, pseudo-autosomal region I; PAR2, pseudo-autosomal region 2.

Figure B: Protein-protein interaction network generated from GWAS protein-coding genes involved in CAD.

This network was generated using STRING *(Search Too/for the Retrieval of Interacting Genes/Protein)* database version 10.0 (http://string.embl.de) and represents the protein-protein interactions from the 86 GWAS protein-coding genes found to be involved in CAD (Table 1). The interactions include direct (physical) and indirect (functional) associations derived from genomic context, high-throughput experiments, co-expression, and literature mining. Stronger associations are represented by thicker lines.

Figure C: UTY functional predicted interaction partners network in the macrophages.

This network was generated using GIANT *(Genome-scale Integrated Analysis of gene Networks in Tissues)* webserver (http://giant.minceton.edu/) and represents the predicted 6 protein-coding genes most tightly connected to *UTY* in macrophages. Edge thickness correspond to edge strength.