



REVIEW ARTICLE

Sortilin and the risk of cardiovascular disease[☆]

Maria Francisca Coutinho^{a,b,c,*}, Mafalda Bourbon^d, Maria João Prata^{b,c}, Sandra Alves^a

^a Grupo de Investigação em Doenças Lisossomais de Sobrecarga, Unidade de I&D, Departamento de Genética Humana, INSA, Porto, Portugal

^b IPATIMUP, Porto, Portugal

^c Departamento de Biologia, Faculdade de Ciências, Universidade do Porto, Porto, Portugal

^d Grupo de Investigação Cardiovascular, Unidade de I&D, Departamento de Promoção da Saúde e Doenças Crónicas, INSA, Lisboa, Portugal

Received 11 January 2013; accepted 21 February 2013

Available online 20 November 2013

KEYWORDS

Genome-wide association studies;
Coronary artery disease;
Low-density lipoprotein cholesterol;
Sortilin;
Functional genetics;
Lipoprotein metabolism

Abstract Plasma low-density lipoprotein cholesterol (LDL-C) levels are a key determinant of the risk of cardiovascular disease, which is why many studies have attempted to elucidate the pathways that regulate its metabolism. Novel latest-generation sequencing techniques have identified a strong association between the 1p13 locus and the risk of cardiovascular disease caused by changes in plasma LDL-C levels. As expected for a complex phenotype, the effects of variation in this locus are only moderate. Even so, knowledge of the association is of major importance, since it has unveiled a new metabolic pathway regulating plasma cholesterol levels. Crucial to this discovery was the work of three independent teams seeking to clarify the biological basis of this association, who succeeded in proving that *SORT1*, encoding sortilin, was the gene in the 1p13 locus involved in LDL metabolism. *SORT1* was the first gene identified as determining plasma LDL levels to be mechanistically evaluated and, although the three teams used different, though appropriate, experimental methods, their results were in some ways contradictory. Here we review all the experiments that led to the identification of the new pathway connecting sortilin with plasma LDL levels and risk of myocardial infarction. The regulatory mechanism underlying this association remains unclear, but its discovery has paved the way for considering previously unsuspected therapeutic targets and approaches.

© 2013 Sociedade Portuguesa de Cardiologia. Published by Elsevier España, S.L. All rights reserved.

PALAVRAS-CHAVE

Genome wide association studies;
Doença das artérias coronárias;

Sortilina e risco de doença cardiovascular

Resumo O nível plasmático de c-LDL constitui um determinante chave para o risco de doença cardiovascular, razão pela qual muitos estudos têm procurado elucidar as vias que regulam o seu metabolismo. As novas técnicas de sequenciação de última geração permitiram identificar

[☆] Please cite this article as: Coutinho MF, Bourbon M, Prata MJ, et al. Sortilina e risco de doença cardiovascular. Rev Port Cardiol. 2013;32:793–799.

* Corresponding author.

E-mail address: francisca.coutinho@yahoo.com (M.F. Coutinho).

Colesterol
lipoproteína de baixa
densidade;
Sortilina;
Genómica funcional;
Metabolismo das
lipoproteínas

um forte sinal de associação entre o *locus* 1p13 e o risco de doença cardiovascular causada por alteração dos níveis de LDL no plasma. Como seria de esperar para um fenótipo complexo, os efeitos da variação nesse *locus* são apenas moderados, ainda assim, o conhecimento da associação foi de grande importância uma vez que conduziu à descoberta de uma nova via metabólica reguladora dos níveis de colesterol no plasma. Para tal, foram fundamentais os trabalhos efetuados por três equipas independentes, que ao procurarem esclarecer as bases biológicas da associação em causa conseguiram provar que o gene *SORT1*, codificador da sortilina, era o gene do *locus* 1p13 implicado no metabolismo das LDL. *SORT1* foi o primeiro dos genes identificados como determinantes dos níveis plasmáticos de LDL a ser alvo de avaliação mecanística e embora cada uma das equipas recorresse a metodologias experimentais diferentes, mas igualmente apropriadas face à questão em investigação, os resultados que obtiveram foram contraditórios em alguns aspetos. Neste trabalho, revemos o caminho percorrido até à descoberta da nova via que relaciona a sortilina com os níveis plasmáticos de LDL e com o risco de enfarte do miocárdio. Ainda por esclarecer permanece o mecanismo regulador dessa ligação, mas a sua descoberta sugere novos alvos terapêuticos até há bem pouco tempo desconhecidos. © 2013 Sociedade Portuguesa de Cardiologia. Publicado por Elsevier España, S.L. Todos os direitos reservados.

List of abbreviations

apo	apolipoprotein
C/EBP	CCAAT/enhancer binding protein
CAD	coronary artery disease
CVD	cerebrovascular disease
GLUT4	glucose transporter 4
GM2AP	GM2 activator protein
GWAS	genome-wide association studies
HDL	high-density lipoprotein
LDL	low-density lipoprotein
LDL-C	low-density lipoprotein
LPL	lipoprotein lipase
MI	myocardial infarction
RAP	receptor-associated protein
SAP	sphingolipid activator protein
siRNA	short interfering RNA
SNP	single-nucleotide polymorphism
TC	total cholesterol
VLDL-C	very low-density lipoprotein cholesterol

Introduction

Cardiovascular disease is the leading cause of death in developed countries,¹ and is responsible for 32% of deaths recorded in Portugal, according to the National Institute of Statistics.² Coronary artery disease (CAD), in particular, represents a major clinical problem, accounting for one in five deaths in the US.^{3,4} Multiple factors contribute to the development of CAD but it is well established that one of its key determinants is plasma LDL-C level. According to estimates by the WHO, about 9 million deaths/year and more than 75 million years of life lost/year are due to hypertension or hypercholesterolemia.⁵ Overall, hypercholesterolemia is responsible for 18% of recorded events

of cerebrovascular disease (CVD), mostly non-fatal events, and 56% of ischemic heart disease.⁵ The data for Europe suggest that hypercholesterolemia may be responsible for up to 12% of disability-adjusted life years.⁵ Given the size of these numbers, many attempts have been made to elucidate the pathways that regulate LDL metabolism. It is now known that, for small groups of individuals, high cholesterol levels may be of genetic origin. There is even a Mendelian disease associated with high blood cholesterol: familial hypercholesterolemia.⁶ Most patients suffering from this condition present pathogenic mutations in the gene that codes for the LDL receptor (*LDLR*), but it has been reported that defects in the apolipoprotein (apo) B gene (*APOB*), or less commonly, in the proprotein convertase subtilisin/kexin type 9 (*PCSK9*) gene, may also be associated with this clinical phenotype.^{6,7} Mutations in any of these genes lead to either loss (*LDLR* and *APOB*) or gain (*PCSK9*) of function of its associated protein and high cardiovascular risk.

However, there are few cases in which it is possible to relate a specific gene mutation to CVD. The pathogenesis of the major forms of CVD involves behavioral, environmental and genetic factors and the genetic component is known to be highly complex, resulting from the interaction of multiple genetic determinants.⁸ There are, however, several polymorphisms in these and other genes involved in lipid metabolism that, even though presenting a smaller effect on the protein for which they code, may play a significant part in CVD risk (reviewed in ⁶).

With the advent of new sequencing technologies, the search for a deeper understanding of these mechanisms, as well as the genetic basis of other risk factors, has gained new impetus; it has become possible to screen large populations for the genetic basis for complex diseases. Ultimately, such epidemiological studies may lead to a better understanding of etiological pathways and contribute to the development of new strategies for prevention and treatment.⁹

Recently, large-scale genome-wide association studies (GWAS) have made it possible to identify a novel set

of DNA variants that influence plasma LDL-C levels. The most consistent of these associations was observed in a cluster of genes on chromosome 1p13. The clinical relevance of this novel pathway is highlighted by the 40% difference in risk of myocardial infarction (MI) between individuals homozygous for the minor (less common) and major (more common) alleles of the p13 locus on chromosome 1. The effect is comparable to that attributed to common variants of *LDLR* and *PCSK9* and greater than that described for the most common variants in *HMGCR* (the gene that codes for 3-hydroxy-3-methylglutaryl-coenzyme A reductase,¹⁰ the therapeutic target of statins, which is the class of drugs most commonly used in the treatment of hyperlipidemias). The *SORT1* gene is located in the 1p13 cluster. This gene codes for sortilin, a multi-functional protein whose biological importance is becoming clearer as it is revealed to have novel and unexpected functions. Although its functions as a receptor for various ligands were already known, given the clear association reported by different GWAS, three independent teams^{10–12} set out to elucidate the biological mechanism relating sortilin to LDL-C levels and, ultimately, to risk of CAD. To this end, they used different mechanistic approaches and, interestingly, came to different conclusions. Here we summarize each of these approaches and their main conclusions, and attempt to reconcile the apparently discrepant results.

Looking for a needle in a haystack: genome-wide association studies

Over the past few years, with the emergence and spread of third-generation sequencers, advances in sequencing and genotyping have catapulted GWAS to the forefront of population studies, with special focus on the relationship between genotype and common diseases. These studies are based on the premise that, for a large number of such diseases, the underlying hereditary variations have a minor allele frequency of more than 5%. It is then possible, through the analysis of large population samples, to identify associations of certain diseases with certain regions of the genome, both coding and non-coding. In fact, several genetic variants identified through GWAS are located in non-coding regions of the genome. Interpretation of the effects of the identified variants depends largely on the knowledge available on those regions. The possibility of genes located in the target region being responsible for the detected association is then estimated, without excluding the possibility that it may result from long-range genetic interactions or from other unknown reasons. The challenge is to understand the biological basis of the signs revealed in GWAS. Although this may be difficult, GWAS have already uncovered important genetic factors underlying a number of complex diseases. One of the most successful cases in terms of identification of single-nucleotide polymorphisms (SNPs) which are relevant to the pathogenesis of a complex disease is in fact the annotation of genes correlated with plasma lipid and lipoprotein levels, factors which have long been known to be important in pathological conditions such as dyslipidemia and MI. Over the last few years more than 100 loci have been described as associated with

genetic variation in triglyceride, LDL-C and high-density lipoprotein (HDL) cholesterol levels.^{13–20} In the cases of CAD and MI, GWAS identified a smaller number of genetic loci, some of which were also associated with changes in traditional risk factors. A comprehensive analysis of several GWAS has identified and annotated CAD-associated loci,^{14,17,21–23} by combining data from the Wellcome Trust Case Control Consortium and the German MI Family Study. It presented evidence of associations between seven chromosomal loci and CAD risk¹⁴: 1p13 (*SARS*, *CELSR2*, *PSRC1*, *MYBPHL*, *SORT1*, *PSMA5* and *SYPL2*), 1q41 (*MIA3*), 2q36 (intergenic region), 6q25.1 (*MTHFD1L*), 9p21 (*CDKN2A* and *CDKN2B*), 10q11 (intergenic region) and 15q22.33 (*SMAD3*). The immediate question that arose was whether these new loci affected already known cardiovascular risk factors. To clarify this question, Samani et al.¹⁵ investigated the association of these seven loci with a number of quantitative traits of known relevance to cardiovascular disease, and showed that only the risk locus on chromosome 1p13 was significantly associated with higher LDL-C levels. The strongest association was located in the intergenic region including the *PSRC1* and *CELSR2* genes, which code for proline/serine-rich coiled-coil 1 and cadherin EGF, respectively. The function of these proteins remains unknown, but their coding genes are located close to the gene that codes for sortilin, *SORT1*. None of these three genes, nor any of the others present in the 1p13 locus, have ever been associated with a known Mendelian disease affecting LDL-C levels.^{10,13,20}

Usual and unusual suspects: low-density lipoprotein cholesterol and sortilin

Given such a statistically significant association between the 1p13 locus and plasma LDL-C levels, the search for its explanatory mechanism became the research focus of various teams.

Firstly, it was important to clarify which particular genomic variant was causing this association. Due to the linkage disequilibrium (extensive and non-random relationship) between multiple SNPs at the 1p13 locus (comprising the genes *SORT1*, *PSRC1* and *CELSR2*), it was impossible to identify the causal variant solely through GWAS. In silico, in vitro and in vivo studies would be required to clarify this point, as well as the mechanisms behind this association. As these studies were being carried out, one gene began to stand out from all the others comprising this CAD risk locus: the *SORT1* gene, which codes for sortilin.^{10–12}

Sortilin belongs to the Vps10p domain receptor family, which consists of five known members. It is synthesized as a propeptide, cleaved in the Golgi apparatus by proprotein convertases, after which the protein takes its mature form, which allows proper ligand binding. Functionally, sortilin is a receptor of multiple ligands, including lipoprotein lipase (LPL),²⁴ the A-V apolipoproteins (apo A-V),²⁵ neurotensin²⁶ and receptor-associated protein (RAP).²⁷ It is also responsible for mediating Golgi-to-lysosome transport of a number of lysosomal proteins, some (but not all) enzymatic: sphingolipid activator proteins (SAPs): prosaposin and GM2 activator protein (GM2AP), acid sphingomyelinase, cathepsin H and cathepsin D.^{24,28–30} In recent years, it has been

demonstrated that sortilin is involved in a number of important biological processes such as the formation of glucose transport 4 (GLUT-4) storage vesicles in response to insulin during adipocyte differentiation.³¹ In the brain, it is part of a signaling complex that regulates cell survival.³²

The importance of these multiple properties in vivo remains unclear but it is evident that sortilin is a protein with an important biological role, deregulation of which is likely to cause severe side-effects that may go beyond its effect on plasma LDL-C levels.

An exemplary approach: mechanistic analyses

In 2010, three independent teams published results from pioneering studies which set out to clarify the biological mechanism underlying the association between the 1p13 locus and plasma LDL levels.^{10–12} Based on solid, though different, experimental approaches, all three studies indicate that the *SORT1* gene is responsible for the increased risk of CAD and/or MI. Curiously, the studies reached conclusions that were not only different but, in some cases, even opposite concerning the role of sortilin in the secretion of very low-density lipoprotein cholesterol (VLDL-C) (reviewed in^{33,34}).

The first in vitro evidence of the interaction between sortilin and LDL particles was presented by Linsel-Nitschke et al.¹¹ Through fine mapping of the 1p13 locus, the authors began by seeking the variant with the strongest signs of that association and identified the SNP rs599839, showing that the G allele was the one associated with reduced plasma LDL-C levels and lower cardiovascular disease risk. They demonstrated that individuals homozygous for the G allele showed increased expression of the *SORT1*, *CELSR2* and *PSRC1* genes in peripheral white blood cells. The strongest and most consistent association, however, was seen for *SORT1* mRNA levels. These results were confirmed in human embryonic kidney cells (HEK293) over-expressing *SORT1* that showed increased internalization of LDL-C particles, leading to lower LDL-C plasma levels.¹¹

In the same year (2010), Musunuru et al.¹⁰ presented a multifaceted approach, a *tour de force* for the follow-up of GWAS.³⁵ Based on the previous recognition that the rs646776, rs599839, rs12740374 and rs629301 SNPs from the 1p13 locus were most strongly associated with plasma LDL-C levels and on the assumption that non-coding DNA variants may alter gene expression, Musunuru et al. started by analyzing the effects of these four variants on the mRNA levels of the six genes located in that locus: *SARS*, *CELSR2*, *PSRC1*, *MYBPHL*, *SYP2* and *SORT1*. They found that in human liver the minor allele for the rs646776 SNP was associated with increased expression of the *SORT1*, *CELSR2* and *PSRC1* genes,³⁶ with the strongest association observed for *SORT1* mRNA levels and its corresponding protein, sortilin. Fine mapping of the region of interest led to the identification of the haplotypes defined by the SNPs present in 6.1 kilobases located between the *CELSR2* and *PSRC1* genes, and to the identification of the SNP rs12740374 as the one ultimately responsible for the association observed in GWAS. Bioinformatic analysis showed that, altering the wild-type sequence from GGTGCTCAAT to GTTGCTCAAT, the minor allele of this variant created a binding site for the CCAAT/enhancer

binding protein (C/EBP) α , increasing promoter activity and *SORT1* expression level. This was later confirmed in vitro. It should be noted that these results are in full agreement with the findings of Linsel-Nitschke's group¹¹ concerning mRNA expression levels in the liver. Finally, through studies on liver cells from mutant mice in which the gene coding for sortilin was over-expressed or inactivated, Musunuru et al. demonstrated that sortilin expression levels modulate the hepatic secretion of VLDL. The transgenic mouse chosen by this team was *Apobec1*^{-/-}, a humanized mouse in which the gene that encodes the C->U-editing enzyme APOBEC-1 is suppressed, with a lipid profile closer to that seen in humans, in whom LDL is the predominant cholesterol transporter in circulation, rather than that typical of mice. When Musunuru et al. over-expressed the *SORT1* gene in *Apobec1*^{-/-} liver cells, a 70% reduction in plasma total cholesterol (TC) and LDL-C was observed. Similarly, inactivation of *SORT1* by short-interfering RNA (siRNA) led to increases of 46% in TC and 125% in LDL.

In general, the data presented by these two teams support the findings of GWAS findings and reinforce the idea of a negative correlation between *SORT1* mRNA levels and plasma LDL-C concentrations. However, in the same year a third mechanistic study addressing this association was published and in this case, the results were not so easily reconciled either with the results from the previous studies or with the previous assumptions inferred through GWAS.

The results presented by Kjolby et al.¹² were published almost simultaneously. These authors used as a model a double knockout mouse, *Sort1*^{-/-}, *Ldlr*^{-/-}, having observed that its hepatocytes presented reductions of 30% in TC levels, ~50% in proteins containing apo B100 (VLDL and LDL), and ~60% in atherosclerotic plaque area compared to *Ldlr*^{-/-} single knockout mice. Next, they performed liver-specific *SORT1* over-expression. Briefly, they found that sortilin deficiency led to a 50% reduction in the secretion of lipoproteins, whereas over-expression resulted in a 50% increase. Together, these results indicate a positive correlation between *SORT1* expression and LDL-C levels, opposite to that observed by Musunuru et al.¹⁰

Spot the differences: analysis of the results

The question of the discrepancy between the results of these three studies has been discussed by various experts, particularly Dubé³⁴ and Tall and Ai,³³ in 2011. These authors drew attention to the methodological differences between the studies of Linsel-Nitschke et al.,¹¹ Musunuru et al.¹⁰ and Kjolby et al.,¹² and the effects that those differences may have had on their results. The three teams that set out to clarify the mechanism by which the 1p13 locus affects LDL levels and the risk of CAD opted for different experimental models, which appear to have influenced the final results. This implies that their conclusions, even though they are consistent and resulted from well-designed and consistent experiments, may not be comparable.

The first important point is the metabolic background in which each of the experiments was carried out. Linsel-Nitschke and colleagues¹¹ conducted their investigation only in humans, unlike Musunuru et al.¹⁰ and Kjolby

et al.,¹² who analyzed non-human animal models. Nevertheless, the latter teams chose mouse models with different metabolic profiles: Musunuru et al. worked with liver cells from a humanized mouse, *Apobec*^{-/-}, while Kjolby et al. studied a sortilin and LDL receptor double knockout (*Sort1*^{-/-}, *Ldlr*^{-/-}). Musunuru's mouse produced and secreted abnormally high amounts of lipoproteins, mimicking the human lipid profile, which may have artificially modified sortilin's secretory pathways and availability. Kjolby's mouse had deficient lipoprotein catabolism, created by the repression of *SORT1* expression within hepatocytes and a high-fat "western" diet. Finally, there are differences in gene regulation between the two species, man and mouse, as demonstrated by the absence of the C/EBP α binding site in mice.^{10,37} This may hinder extrapolation of mouse studies to humans with regard to sortilin.³⁴

It should be noted that, taken together, in vivo observations from studies in mouse models show that sortilin assumes complementary liver functions, depending on the metabolic milieu in which it operates, and ultimately regulates VLDL secretion. The different results reviewed here seem to suggest that sortilin regulates VLDL secretion and traffic to the lysosome when intracellular apo B-100 levels are extremely high. Conversely, at low apo B-100 expression levels, sortilin regulates the formation and secretion of VLDL.^{33,34}

Nevertheless, this putative role of sortilin in the formation and secretion of VLDL is hard to reconcile with GWAS results indicating a specific association with LDL-C but not with triglycerides, which are the major components of VLDL particles.³⁴

Taken together, even though the findings of these studies are in some ways contradictory, they also provide strong evidence of the existence of a novel regulatory pathway for lipoprotein metabolism and show that modulating this pathway could alter cardiovascular disease risk in humans. Nevertheless, there is still a long way to go before the whole process and its modulators are clearly understood.

Conclusion

By genotyping high-frequency alleles, GWAS are limited to identifying alleles which exert minimal, or even negligible, effects on the phenotype.³⁸ Furthermore, the most common alleles appear to explain only a small portion of the phenotype.³⁹ Thus, a significant proportion of the inheritance of complex phenotypes such as cardiovascular disease (CAD in particular) remains unknown, despite all the efforts in this area through numerous GWAS. This portion of heredity has been called "missing heritability" or the "dark matter" of heritability.³⁵ The supporters of GWAS argue that increasing the size of the study samples and SNP density will enable detection of alleles with very small effect sizes, revealing the portion of inheritance which remains unknown.³⁵ Nevertheless, some authors support an alternative strategy based on whole-genome direct sequencing as a way of identifying rare alleles with large effects on the phenotype.⁴⁰ Crucial to this are recent advances in DNA sequencing technology, with the development of new platforms for third-generation sequencing that enable low-cost whole genome sequencing and identification of

rare and/or new variants. It is believed that each genome has approximately 10 000 non-synonymous variants, among approximately 3.5 million SNPs. Given the size of these numbers, this kind of sequencing is expected to dominate genetic studies in coming years, a trend that can already be seen.^{41–45}

It does, however, appear that the "dark matter" of heritability is the product of complex interactions between factors of different types: genetic, genomic and epigenetic.^{39,46} Similarly, the phenotype is also the result of nonlinear and stochastic interactions between different genetic and non-genetic factors.

Nevertheless, the discovery and systematization of new genetic variants associated with a particular complex phenotype is important, particularly for the example reported here, in which these techniques have led to the discovery of a previously unknown molecular pathway.

The three reports reviewed here are exemplary approaches to the need to move from a "blind" statistical association given by GWAS to a mechanistic explanation of how a particular genetic variation can modulate a particular phenotype. In this case, GWAS results pointed to a particular starting point, the 1p13 locus,^{14,21–23} which eventually caught the attention of three independent teams who relied on different experimental approaches to unveil the basis of the statistical association, all of which identified the *SORT1* gene as the modulator of LDL-C levels and MI risk. But, while their results were in agreement concerning the relevance of *SORT1*'s role in the regulation of lipoprotein metabolism, their interpretations of the effect of its expression on plasma LDL-C levels and its underlying mechanism differed.

Linsel-Nitschke and colleagues¹¹ proposed, on the basis of their observations, that overexpression of sortilin increases the internalization of LDL, with a consequent decrease in plasma levels. Soon afterwards, through studies in human cohorts, hepatocytes and mice, Musunuru et al.¹⁰ reported an inverse relationship between sortilin expression and circulating LDL-C levels, and proposed an explanatory mechanism through transcriptional regulation (liver-specific) of the *SORT1* gene by the transcription factor C/EBP α . By contrast, Kjolby et al.¹² observed a direct relationship between *SORT1* expression and circulating LDL concentrations, suggesting this could result from increased VLDL secretion.

Several explanations have been put forward for the discrepancy between these results; the answer seems to depend on sortilin itself, which appears to be a multifaceted protein that can assume different functions depending on circumstances.

To summarize, the studies reviewed here presented strong evidence that *SORT1* is a regulator of plasma LDL-C levels, adding a significant role to the sortilin-coding gene that was unknown until recently. The cellular pathway relating sortilin to lipid metabolism is still controversial but it is surely an issue that will be further explored. A full understanding of this pathway will be crucial to assess whether sortilin is a potential target for therapeutic interventions for hypercholesterolemia or CAD (reviewed in ^{33–35}).

Conflicts of interest

The authors have no conflicts of interest to declare.

References

- Deaton C, Froelicher E, Wuc L, et al. The global burden of cardiovascular disease. *Eur J Cardiovasc Nurs*. 2011;10:55–13.
- Anuário Estatístico de Portugal. Instituto Nacional de Estatística; 2010. Available from: http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_publicacoes&PUBLICACOESpub_boui=133813349&PUBLICACOESmodo=2 [accessed 18.08.12].
- Rosamond W, Flegal K, Friday G, et al. Heart disease and stroke statistics – 2007 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. *Circulation*. 2007;115:e69–171.
- Prospective Studies Collaboration. Blood cholesterol and vascular mortality by age, sex and blood pressure: meta-analysis of individual data from 61 prospective studies with 55,000 vascular deaths. *Lancet*. 2007;370:1829–39.
- Costa J, Borges M, Oliveira E, et al. Incidence and prevalence of hypercholesterolemia in Portugal: a systematic review. *Port J Cardiol*. 2003;22:569–77.
- Bourbon M. Genetic factors and cardiovascular disease. *Rev Port Cardiol*. 2008;27:1559–63.
- Medeiros AM, Alves AC, Francisco V, et al., Investigators of the Portuguese FH Study. Update of the Portuguese Familial Hypercholesterolaemia Study. *Atherosclerosis*. 2010;212:553–8.
- Watkins H, Farrall M. Genetic susceptibility to coronary artery disease: from promise to progress. *Nat Rev Genet*. 2006;7:163–73.
- Burton PR, Tobin MD, Hopper JL. Key concepts in genetic epidemiology. *Lancet*. 2005;366:941–51.
- Musunuru K, Strong A, Frank-Kamenetsky M, et al. From non-coding variant to phenotype via SORT1 at the 1p13 cholesterol locus. *Nature*. 2010;466:714–9.
- Linsel-Nitschke P, Samani NJ, Schunkert H. Sorting out cholesterol and coronary artery disease. *N Engl J Med*. 2010;363:2462–3.
- Kjolby M, Andersen OM, Breiderhoff T, et al. Sort1, encoded by the cardiovascular risk locus 1p13.3, is a regulator of hepatic lipoprotein export. *Cell Metab*. 2010;12:213–23.
- Kathiresan S, Melander O, Guiducci C, et al. Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nat Genet*. 2008;40:189–97.
- Samani NJ, Erdmann J, Hall AS, et al. Genomewide association analysis of coronary artery disease. *N Engl J Med*. 2007;357:443–53.
- Samani NJ, Braund PS, Erdmann J, et al. The novel genetic variant predisposing to coronary artery disease in the region of the PSRC1 and CELSR2 genes on chromosome 1 associates with serum cholesterol. *J Mol Med (Berl)*. 2008;86:1233–41.
- Willer CJ, Sanna S, Jackson AU, et al. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat Genet*. 2008;40:161–9.
- Sandhu MS, Waterworth DM, Debenham SL, et al. LDL-cholesterol concentrations: a genome-wide association study. *Lancet*. 2008;371:483–91.
- Wallace C, Newhouse SJ, Braund P, et al. Genomewide association study identifies genes for biomarkers of cardiovascular disease: serum urate and dyslipidemia. *Am J Hum Genet*. 2008;82:139–49.
- Samani NJ, Deloukas P, Erdmann J, et al. Large scale association analysis of novel genetic loci for coronary artery disease. *Arterioscler Thromb Vasc Biol*. 2009;29:774–80.
- Kathiresan S, Willer CJ, Peloso GM, et al. Cupples common variants at 30 loci contribute to polygenic dyslipidemia. *Nat Genet*. 2009;41:56–65.
- McPherson R, Pertsemlidis A, Kavaslar N, et al. A common allele on chromosome 9 associated with coronary heart disease. *Science*. 2007;316:1488–91.
- Helgadottir A, Thorleifsson G, Manolescu A, et al. A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science*. 2007;316:1491–3.
- The Wellcome Trust Case Control Consortium. Genomewide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*. 2007;447:661–78.
- Nielsen MS, Jacobsen C, Olivecrona G, et al. Sortilin/neurotensin receptor-3 binds and mediates degradation of lipoprotein lipase. *J Biol Chem*. 1999;274:8832–6.
- Nilsson SK, Christensen S, Raarup MK, et al. Endocytosis of apolipoprotein A-V by members of the low density lipoprotein receptor and the VPS10p domain receptor families. *J Biol Chem*. 2008;283:25920–7.
- Mazella J, Zsürger N, Navarro V, et al. The 100-kDa neurotensin receptor is gp95/sortilin, a non-G-protein-coupled receptor. *J Biol Chem*. 1998;273:26273–6.
- Petersen CM, Nielsen MS, Nykjaer A, et al. Molecular identification of a novel candidate sorting receptor purified from human brain by receptor-associated protein affinity chromatography. *J Biol Chem*. 1997;272:3599–605.
- Lefrançois S, Zeng J, Hassan AJ, et al. The lysosomal trafficking of sphingolipid activator proteins (SAPs) is mediated by sortilin. *EMBO J*. 2003;22:6430–7.
- Ni X, Canuel M, Morales CR. The sorting and trafficking of lysosomal proteins. *Histol Histopathol*. 2006;21:899–913.
- Canuel M, Korkidakis A, Konnyu K. Sortilin mediates the lysosomal targeting of cathepsins D and H. *Biochem Biophys Res Commun*. 2008;373:292–7.
- Shi J, Kandror KV. Sortilin is essential and sufficient for the formation of Glut4 storage vesicles in 3T3-L1 adipocytes. *Dev Cell*. 2005;9:99–108.
- Vaegter CB, Jansen P, Fjorback AW, et al. Sortilin associates with Trk receptors to enhance anterograde transport and neurotrophin signaling. *Nat Neurosci*. 2011;14:54–61.
- Tall AR, Ai D. Sorting out sortilin. *Circ Res*. 2011;108:158–60.
- Dubé JB, Johansen CT, Hegele RA. Sortilin: an unusual suspect in cholesterol metabolism: from GWAS identification to in vivo biochemical analyses, sortilin has been identified as a novel mediator of human lipoprotein metabolism. *Bioessays*. 2011;33:430.
- Marian AJ. Genome-wide association studies complemented with mechanistic biological studies identify sortilin 1 as a novel regulator of cholesterol trafficking. *Curr Atheroscler Rep*. 2011;13:190–2.
- Schadt EE, Molony C, Chudin E, et al. Mapping the genetic architecture of gene expression in human liver. *PLoS Biol*. 2008;6:e107.
- Russell JC, Proctor SD. Small animal models of cardiovascular disease: tools for the study of the roles of metabolic syndrome, dyslipidemia, and atherosclerosis. *Cardiovasc Pathol*. 2006;15:318–30.
- Marian AJ. Nature's genetic gradients and the clinical phenotype. *Circ Cardiovasc Genet*. 2009;2:537–9.
- Eichler EE, Flint J, Gibson G, et al. Missing heritability and strategies for finding the underlying causes of complex disease. *Nat Rev Genet*. 2010;11:446–50.
- Cirulli ET, Goldstein DB. Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nat Rev Genet*. 2010;11:415–25.
- Wang J, Wang W, Li R, et al. The diploid genome sequence of an Asian individual. *Nature*. 2008;456:60–5.

42. Levy S, Sutton G, Ng PC, et al. The diploid genome sequence of an individual human. *PLoS Biol.* 2007;5:e254.
43. Wheeler DA, Srinivasan M, Egholm M, et al. The complete genome of an individual by massively parallel DNA sequencing. *Nature.* 2008;452:872–6.
44. Ng PC, Levy S, Huang J, et al. Genetic variation in an individual human exome. *PLoS Genet.* 2008;15:e1000160.
45. Ng SB, Turner EH, Robertson PD, et al. Targeted capture and massively parallel sequencing of 12 human exomes. *Nature.* 2009;461:272–6.
46. Kapranov P, Laurent GS. Genomic «dark matter»: implications for understanding human disease mechanisms, diagnostics, and cures. *Front Genet.* 2012;3:95.