



ARTIGO DE REVISÃO

Sortilina e risco de doença cardiovascular

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

Recebido a 11 de janeiro de 2013; aceite a 21 de fevereiro de 2013

Disponível na Internet a 31 de julho de 2013

PALAVRAS-CHAVE

Genome wide
association studies;
Doença das artérias
coronárias;
Colesterol
lipoproteína de baixa
densidade;
Sortilina;
Genómica funcional;
Metabolismo das
lipoproteínas

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 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.

Abreviaturas: apo, apolipoproteína; C/EBP, CCAAT-enhancer binding protein; CT, colesterol total; DAC, doença das artérias coronárias; DALYs, disability-adjusted life years; DCI, doença cardíaca isquémica; DCV, doença cerebrovascular; EAM, enfarte agudo do miocárdio; GLUT4, glucose transport isoform 4; GM2AP, GM2 activator protein; GWAS, Genome-wide association studies; LDL, Low-density lipoproteins; HDL, High-density lipoproteins; LPL, lipoproteína lipase; RAP, receptor associated protein; SAP, sphingolipid activator protein; SNP, single nucleotide polymorphism.

* Autor para correspondência.

Correio eletrónico: francisca.coutinho@yahoo.com (M.F. Coutinho).

KEYWORDS

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

Sortilin and the risk of cardiovascular disease

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.

Introdução

As doenças cardiovasculares constituem a maior causa de mortalidade nos países desenvolvidos¹, sendo responsáveis por 32% das mortes em Portugal, segundo dados do Instituto Nacional de Estatística². A doença arterial coronária (DAC), em particular, representa um enorme problema clínico, provocando uma em cada cerca de cinco mortes nos Estados Unidos^{3,4}. Se para o desenvolvimento de DAC contribuem múltiplos fatores, está bem estabelecido que um dos determinantes chave de risco são os níveis de colesterol associado às *lipoproteínas de baixa densidade* (LDL) no plasma. Segundo estimativa da OMS, cerca de nove milhões de mortes/ano e mais de 75 milhões de anos de vida perdidos/ano podem ser atribuíveis a hipertensão ou hipercolesterolemia⁵. Globalmente, a hipercolesterolemia é responsável por 18% da doença cerebrovascular (DCV), maioritariamente eventos não fatais, e por 56% da doença cardíaca isquémica (DCI)⁵. Os dados disponíveis relativamente à Europa indicam que a hipercolesterolemia possa ser responsável por até 12% dos anos de vida perdidos ajustados à incapacidade (*Disability-Adjusted Life Years [DALYs]*)⁵. Face à ordem de grandeza destes números, muitas têm sido as tentativas no sentido de esclarecer as vias que regulam o metabolismo do colesterol associado às lipoproteínas de baixa densidade (c-LDL). Atualmente, sabe-se que em subgrupos minoritários de indivíduos o colesterol elevado pode ser de origem genética simples, resultando de doenças mendelianas como a hipercolesterolemia familiar⁶. A maioria dos doentes com esta patologia apresenta mutações patogénicas no gene que codifica o recetor das LDL (*LDLR*), mas tem-se verificado que defeitos no gene que codifica a apolipoproteína B (*APOB*) ou, mais raramente, no que codifica a proaproteína convertase subtilisina quexina tipo 9 (*PCSK9*), também estão associados ao mesmo fenótipo clínico^{6,7}. As mutações encontradas em qualquer um destes genes provocam perda (no caso do *LDLR* e *APOB*) ou ganho de função

(no caso do *PCSK9*) da proteína que codificam, alterando significativamente a sua função, sendo consideradas fator de elevado risco para o desenvolvimento de DCV.

Todavia, poucos são os casos em que se pode relacionar DCV com os efeitos drásticos de uma mutação num dado gene. Na patogénese da maioria das formas estão envolvidos fatores comportamentais, ambientais e genéticos, sabendo-se ainda que a componente genética é por si só complexa podendo implicar a interação entre múltiplos determinantes genéticos⁸. Conhecem-se já uma série de polimorfismos nos genes referidos ou noutros envolvidos no metabolismo lipídico que podem contribuir significativamente para o risco de DCV⁶, embora o efeito individual de cada variação seja pequeno.

Com o advento das novas tecnologias de sequenciação, ganhou novo fôlego o esforço para alcançar uma compreensão mais profunda dos fatores genéticos subjacentes ao desenvolvimento de doenças complexas. Assistiu-se, assim, a um aumento substancial dos estudos de epidemiologia molecular e a avanços importantes no entendimento da etiologia de muitas dessas doenças, o que em última análise é fundamental para desenhar novas estratégias de prevenção e tratamento⁹.

Recentemente, através de estudos de associação envolvendo rastreios genómicos de grande escala (GWAS, *Genome Wide Association Studies*) foi possível identificar uma série nova de variantes de DNA que influenciam os níveis de LDL no plasma. Uma das associações mais consistentes foi inicialmente observada numa região do cromossoma 1p13, onde acabaria por se identificar uma variação genética comum fortemente associada ao c-LDL e enfarte de miocárdio. A importância clínica desta descoberta está patente na diferença de 40% no risco de enfarte do miocárdio verificada em indivíduos homocigóticos para os alelos maioritário (mais comum) e minoritário (menos comum) dessa variação. O efeito é comparável ao das variantes comuns dos já mencionados *LDLR* e *PCSK9* e maior do que descrito para

as variantes mais comuns de *HMGCR* (gene que codifica a 3-hidroxi-3-metilglutaril-CoA reductase)¹⁰, o gene que codifica o alvo terapêutico das estatinas, os fármacos mais usados no tratamento das hiperlipidemias. No *locus* 1p13 está localizado *SORT1*, o gene que codifica a sortilina, uma proteína multifuncional cuja relevância a nível biológico tem vindo a ser cada vez mais valorizada à medida que lhe vão sendo decifradas novas funcionalidades. Embora já fossem conhecidas as suas funções como recetor de vários ligandos, começaram a aparecer indícios claros de que a sortilina pode influenciar os níveis plasmáticos de c-LDL e, conseqüentemente, o risco de DAC. Se, inicialmente, diferentes GWAS foram fundamentais para chegar a *SORT1* como provável gene de risco para a associação c-LDL/DAC, em 2010 surgiram quase em simultâneo três estudos desenvolvidos por equipas independentes¹⁰⁻¹² que procuraram explorar o mecanismo biológico que poderia relacionar a sortilina com os níveis de c-LDL. Para tal, utilizaram abordagens mecanicistas diferentes tendo, curiosamente, chegado a conclusões também diferentes. Neste artigo, resumimos cada uma dessas abordagens e suas principais conclusões, avançando ainda com as razões que têm sido invocadas para conciliar os resultados aparentemente discrepantes.

Em busca de uma agulha num palheiro: os *Genome Wide Association Studies*

Ao longo dos últimos anos, os avanços ao nível das técnicas da biologia molecular, em grande parte proporcionados pelo aparecimento e generalização do uso da terceira geração de sequenciadores, catapultaram os GWAS para a linha da frente dos estudos populacionais, com enfoque na identificação de moduladores genéticos de doenças, em especial de doenças complexas. Os GWAS baseiam-se na premissa de que, em muitas dessas doenças, as variantes hereditárias que concorrem para explicar o desenvolvimento do processo patogénico podem ser relativamente comuns, com frequência do alelo minoritário (*minor allele frequency*, MAF) superior a 5%, o que torna possível identificar associações entre doenças e variações genómicas através de grandes rastreios populacionais. Essas variações não são necessariamente codificantes e, de facto, através de GWAS têm sido identificadas associações com uma série de variantes genéticas localizadas em zonas não codificantes do genoma. A interpretação do significado destas associações depende, em grande parte, do conhecimento fino sobre as regiões em que se encontram as variantes. No fundo, questiona-se se uma variante para a qual se detetou um sinal positivo de associação não influenciará a expressão de um gene ou se o sinal não será consequência da ocorrência de desequilíbrio de ligação com outras variações em genes localizados nessa mesma região genómica. Procura-se, então, identificar os genes com potencial para explicar a associação detetada, seguindo-se um outro desafio, que é compreender a base biológica dos sinais revelados pelos GWAS. Embora as expectativas nem sempre sejam alcançadas, este tipo de estudos tem desvendado importantes fatores genéticos que estão por detrás de uma série de doenças complexas. Um dos casos de maior sucesso é ilustrado pela identificação de *single nucleotide polymorphism* (SNP) relevantes em doenças cujos quadros

patogénicos dependem dos níveis de lípidos e de lipoproteínas no plasma, como são exemplo a dislipidemia ou o enfarte agudo do miocárdio (EAM). Ao longo dos últimos anos, foram já descritos mais de 100 *loci* associados a variação dos níveis de triglicéridos, c-LDL e colesterol HDL¹³⁻²⁰. Relativamente a DAC ou EAM, os GWAS levaram à identificação de um menor número de *loci*, alguns dos quais igualmente associados a alterações nos fatores de risco tradicionais. Um balanço conjunto dos vários GWAS que resultaram na anotação de *loci* associados a DAC^{14,17,21-23}, entre os quais se incluíam os estudos do *Welcome Trust Case Control Consortium* e do *German MI Family Study*, saldou-se no registo de indícios de associação com 7 *loci* cromossómicos¹⁴, nomeadamente localizados em 1p13.3 (compreendendo os genes *SARS*, *CELSR2*, *PSRC1*, *MYBPHL*, *SORT1*, *PSMA5* e *SYPL2*), 1q41 (gene *MIA3*), 2q36.3 (região intergénica), 6q25.1 (gene *MTHFD1L*), 9p21.3 (genes *CDKN2A* e *CDKN2B*), 10q11.21 (região intergénica) e 15q22.33 (gene *SMAD3*). A questão imediata que se colocou foi a de saber se estes novos *loci* afetavam os fatores de risco cardiovascular já conhecidos. Para a esclarecer, Samani et al.¹⁵ investigaram a relação destes sete *loci* com uma série de parâmetros mensuráveis de reconhecida relevância em termos de doença cardiovascular, tendo demonstrado que apenas o *locus* de risco de DAC no cromossoma 1p13 estava significativamente associado a níveis elevados de c-LDL, tendo detetado na região intergénica localizada entre os genes *PSRC1* e *CELSR2* o sinal de associação mais forte. Os genes *PSRC1* e *CELSR2* codificam respetivamente as proteínas *proline/serine-rich coiled-coil 1* e caderina EGF, cuja função permanece desconhecida, e estão muito próximos do gene codificador da sortilina, *SORT1*. Nenhum destes três genes ou dos outros contidos no *locus* 1p13 tinha sido conotado com alguma das doenças mendelianas conhecidas que afetam os níveis de colesterol LDL^{10,13,20}.

Os suspeitos do costume e um suspeito inesperado: lipoproteínas de baixa densidade e sortilina

Face à grande robustez estatística da associação detetada entre o *locus* 1p13 e os níveis plasmáticos de c-LDL, a busca do mecanismo biológico capaz de explicar a observação tornou-se o foco de pesquisa de diversas equipas.

Em primeiro lugar, importava conseguir uma melhor aproximação à causa genética da associação, impossível de obter através dos GWAS devido aos efeitos de *linkage disequilibrium* (relação não aleatória) entre vários SNPs do *locus* 1p13.

Para alcançar o fim em vista, seria necessário proceder à resequenciação fina da região contendo os genes *SORT1*, *PSRC1*, *CELSR2* e efetuar diversos estudos *in silico*, *in vitro* e *in vivo*. À medida que os resultados foram surgindo um gene começou a destacar-se de entre os candidatos que compõem este *locus* de risco de DAC: *SORT1*, o gene que codifica a sortilina¹⁰⁻¹².

A sortilina pertence à família de recetores de domínio Vps10p, constituída por cinco membros atualmente conhecidos. É sintetizada sob a forma de uma proproteína e clivada no Golgi por convertases de proproteínas, processo após o qual assume a forma madura que lhe permite a correta

ligação aos ligandos. Em termos funcionais, é um recetor de múltiplos ligandos, incluindo a lipoproteína lipase (LPL)²⁴, as apolipoproteínas A-V (apo A-V)²⁵, a neurotensina²⁶ e a proteína *receptor-associated protein* (RAP)²⁷. É ainda responsável por mediar o transporte do Golgi para o lisossoma de uma série de proteínas, enzimáticas ou não: proteínas ativadoras dos esfingolípídios (*sphingolipid activator proteins* [SAP]: prosaposina e, GM2AP, GM2 *activator protein*), esfingomielinase ácida, catepsina D e catepsina H^{24,28-30}. Demonstrou-se, ainda, que a sortilina está envolvida numa série de processos de grande relevância biológica de que é exemplo a formação de vesículas de armazenamento de *glucose transport isoform 4* (GLUT-4) na resposta à insulina durante a diferenciação dos adipócitos³¹. No cérebro, faz parte de um complexo de sinalização que regula a sobrevivência da célula³².

A relevância *in vivo* destas propriedades multifacetadas ainda permanece por esclarecer, mas cada vez mais se tem reforçado a noção de que se trata de uma proteína com importantes papéis a nível biológico e cuja desregulação parece passível de causar efeitos indesejáveis, podendo ir muito além da possível alteração dos níveis de colesterol LDL no plasma.

Uma abordagem exemplar: análise mecanística

Em 2010, três equipas independentes publicaram resultados de trabalhos pioneiros na tentativa de esclarecer o mecanismo biológico que poderia explicar a associação do *locus* 1p13 aos níveis de colesterol LDL no plasma¹⁰⁻¹². Recorrendo a sólidas abordagens experimentais distintas entre si, os três estudos apontaram para o gene *SORT1* como sendo o responsável pelo aumento do risco de DAC e/ou EAM. Pese embora esta unanimidade, curioso é o facto de os três estudos terem chegado não apenas a conclusões diferentes, mas até opostas em alguns aspetos, quanto ao papel da sortilina na secreção de colesterol das lipoproteínas de muito baixa densidade (VLDL)^{33,34}.

As primeiras evidências *in vitro* da interação entre a sortilina e partículas de c-LDL foram apresentadas por Linsel-Nitschke et al.. Através do mapeamento fino do *locus* 1p13, os autores começaram por procurar a variação com sinais mais fortes de associação, chegando ao SNP rs599839 e verificando que o alelo G era o que se associava a redução dos níveis plasmáticos de LDL e diminuição do risco de doença cardiovascular. Demonstraram depois que os indivíduos homocigóticos para o mesmo alelo apresentavam expressão aumentada dos genes *SORT1*, *CELSR2* e *PSRC1* em glóbulos brancos do sangue periférico, sendo, no entanto, nos níveis de mRNA *SORT1* que observaram o efeito mais acentuado. Estes resultados foram confirmados em células embrionárias de fígado humano (HEK293) induzidas a sobre-expressar *SORT1*, em que verificaram um aumento da internalização de partículas c-LDL, com consequente diminuição dos respetivos níveis plasmáticos¹¹.

No mesmo ano, Musunuru et al. apostaram numa abordagem multifacetada, vista como um autêntico *tour de force* exemplar para dar sentido aos resultados obtidos com os GWAS³⁵. Baseando-se no conhecimento prévio de que os SNPs rs646776, rs599839, rs12740374, e rs629301 do *locus* 1p13

eram os mais altamente associados ao c-LDL no plasma, e partindo do princípio que variantes de DNA não codificantes, como o caso das quatro em que se centraram, podem alterar a expressão génica, Musunuru et al. começaram por analisar os seus efeitos nos níveis de mRNA dos seis genes localizados no *locus* 1p13: *SARS*, *CELSR2*, *PSRC1*, *MYBPHL*, *SORT1* e *SYP2*. Verificaram, então, que no fígado humano o alelo minoritário de rs646776 estava associado a um aumento da expressão de três genes, *SORT1*, *CELSR2* e *PSRC1*³⁶, tendo verificado maior alteração de expressão ao nível de *SORT1*, traduzida no aumento da quantidade do seu produto proteico, a sortilina. O mapeamento fino da região genómica em estudo permitiu-lhes identificar os haplótipos definidos pelos SNPs presentes numa região de 6,1 kb localizada entre os genes *CELSR2* e *PSRC1*, e chegar ao SNP rs12740374 como sendo afinal o responsável pela associação evidenciada pelos GWAS. Os indícios começaram por surgir na sequência da análise bioinformática, indicando que ao alterar a sequência GGTGCTCAAT para GTTGCTCAAT o alelo minoritário deste SNP criava um novo local de ligação de fator de transcrição, especificamente para a proteína C/EBP (*CCAAT-enhancer binding protein*), capaz de aumentar a atividade do promotor e o nível de expressão de *SORT1*. O efeito previsto *in silico* foi depois confirmado *in vitro*. Refira-se que este resultado era consistente com o reportado pelo grupo de Linsel-Nitschke et al.¹¹ relativamente aos níveis de mRNA expressos no fígado. Finalmente, Musunuru et al.¹⁰ efetuaram estudos em células de fígado de ratinhos mutantes, em que o gene que codifica a sortilina tinha sido inativado ou sobreexpresso, demonstrando que o nível de expressão da sortilina era capaz de modular a secreção hepática das VLDL. O ratinho transgénico escolhido por esta equipa, com o gene *Apobec1* (que efetua a correção das transições C > U nos mRNAs da apo B, entre outros) suprimido, *Apobec1*^{-/-}, era um ratinho «humanizado», apresentando um perfil lipídico diferente do que é normal naquela espécie e mais próximo do que se observa em humanos, em que o c-LDL é o transportador de colesterol predominante na circulação. Quando Musunuru et al.¹⁰ sobre-expressaram o gene *Sort1* nas células hepáticas desses ratinhos, observaram uma redução 70% nas concentrações de colesterol total (CT) e de colesterol LDL. Pelo contrário, quando procederam à inativação do gene *Sort1* por siRNA, registaram um aumento de 46% dos níveis de CT e de 125% dos níveis de LDL.

De um modo geral, os dados apresentados por estas duas equipas vinham reforçar, e também dar sentido, às conclusões extraídas dos GWAS, na medida em que ao demonstrarem a existência de uma correlação negativa entre os níveis de mRNA *SORT1* e a concentração de c-LDL no plasma, sustentavam um mecanismo biológico explicativo de como variações genéticas no *locus* 1p13 podiam influenciar os níveis plasmáticos de LDL e, portanto, alterar o risco de DAC/EAM.

No entanto, ainda no mesmo ano, foi publicado um terceiro estudo mecanístico abordando a mesma questão, cujos resultados nem se conciliavam tão facilmente com os dois anteriores nem com as evidências inferidas através dos GWAS.

Referimo-nos ao trabalho de Kjolby et al.¹², que utilizaram como modelo um ratinho *knockout* duplo, *Sort1*^{-/-}, *Ldlr*^{-/-}, tendo observado que os seus hepatócitos apresentavam reduções de 30% nos níveis de CT, de ~50% nos níveis de proteínas contendo apo B100 (VLDL e LDL) e de

~60% na área das placas arteroscleróticas, quando comparados com o ratinho *knockout* simples, *Ldlr*^{-/-}. De seguida, induziram sobre-expressão do gene *Sort1* ao nível do fígado dos ratinhos, verificando, resumidamente, que enquanto a deficiência de sortilina levava a uma redução de 50% na secreção de lipoproteínas, a sua sobre-expressão se traduzia num aumento de 50% na secreção das mesmas proteínas. No conjunto, os resultados apontavam para uma correlação positiva entre a expressão *Sort1* e a concentração de c-LDL, ou seja, de sentido oposto ao observado por Musunuru et al.¹⁰

Descubra as diferenças: análise dos resultados

A discrepância entre os resultados destes três estudos tem vindo a ser discutida entre a comunidade científica, como ilustram as publicações surgidas em 2011, de Dubé, ou de Tall e Ai^{33,34}. Nos seus comentários, estes autores chamam a atenção para as diferenças experimentais entre os trabalhos de Linsel-Nitschke et al.¹¹, Kjolby et al.¹² e Musunuru et al.¹⁰, e para como tais diferenças podem ter condicionado algum desacordo nos resultados obtidos. Tendo em vista esclarecer o mecanismo através do qual o *locus* 1p13 afetava os níveis de LDL e risco de DAC, as três equipas optaram por modelos diferentes e, embora cada uma delas tivesse recorrido a abordagens experimentais bem desenhadas e delas extraído ilações apropriadas, as diferentes conclusões podem não ser estritamente comparáveis entre si.

Primeiro, devido à diferença de *background* metabólico dos modelos animais usados nas experiências efetuadas. Linsel-Nitschke et al.¹¹ conduziram as suas investigações apenas em humanos, ao contrário das equipas de Musunuru¹⁰ e Kjolby¹², que apostaram na análise de modelos animais não-humanos, especificamente em modelos de ratinho mas, mesmo assim, os ratinhos que escolheram tinham diferentes perfis metabólicos: Musunuru et al.¹⁰ trabalharam com células hepáticas de um ratinho «humanizado», *ApoBec1*^{-/-}; Kjolby et al. estudaram um *knockout* duplo para a sortilina e para o recetor das LDL (*Sort1*^{-/-}, *Ldlr*^{-/-}). Enquanto o modelo de Musunuru¹⁰ sobre-produzia e secretava quantidades anormalmente elevadas de lipoproteínas de modo a mimetizar o perfil lipídico humano, o que pode ter alterado artificialmente as vias secretórias e a disponibilidade de sortilina, o de Kjolby¹² tinha um catabolismo deficiente em lipoproteínas, criado pela repressão da expressão no interior dos hepatócitos e acentuado pela dieta *high fat western*. Outro ponto a ter em linha de conta são as diferenças na regulação génica entre as duas espécies, homem e ratinho, como ficou demonstrado, por exemplo, pela inexistência do local C/EBP α no ratinho^{10,37}. No que se refere à sortilina, isto significa que as conclusões retiradas de observações no ratinho não podem ser linearmente extrapoladas para o homem, nem vice-versa³⁴.

Em conjunto, as observações *in vivo* dos estudos em ratinhos parecem demonstrar que a sortilina pode assumir funções hepáticas complementares dependentes do meio metabólico que, em última análise, regulam a secreção das VLDL. É de admitir que a sortilina possa regular a secreção e o tráfego das VLDL para o lisossoma quando os níveis intracelulares de apo B-100 são extremamente elevados. Pelo

contrário, em condições de menor expressão da apo B-100, a sortilina poderá regular a formação e secreção das VLDL^{33,34}.

Porém, esta possível função da sortilina na formação e secreção das VLDL não se harmoniza facilmente com os resultados obtidos com os GWAS que demonstravam uma associação específica com o c-LDL e não com os triglicérides, que são o componente principal das partículas das VLDL³⁴. Fica assim em aberto a possibilidade de a sortilina ter um papel noutras vias de regulação dos níveis lipídicos no plasma que não impliquem as VLDL.

Se os três estudos em apreço forneceram pistas que em parte são contraditórias, apresentaram também provas da existência de uma nova via regulatória para o metabolismo das lipoproteínas e da possibilidade da modulação desta via alterar o risco de DCV em humanos, não obstante haver ainda muito a percorrer para obter uma visão clara de como se processa e dos fatores que a modulam.

Conclusão

Como os GWAS envolvem a genotipagem de variantes comuns em termos populacionais, têm a limitação de só terem capacidade para identificar alelos cujo efeito num fenótipo associado a um quadro patogénico é mínimo, senão negligenciável³⁸. Ou seja, permitem identificar alelos relativamente comuns mas que explicam apenas uma pequena porção da variação genética associada a um fenótipo³⁹. Assim, mesmo depois de todos os esforços de investimento em GWAS, uma parte significativa da hereditariedade das doenças complexas, como é o caso das doenças cardiovasculares e das DAC em particular, permanece desconhecida. Esta porção da hereditariedade foi denominada «a hereditariedade perdida» ou «a matéria negra da hereditariedade»³⁵. Os grandes defensores dos GWAS argumentam que aumentando o tamanho das amostras e reforçando substancialmente a densidade dos SNPs a estudar, será possível ir desvendando a porção da hereditariedade que se mantém desconhecida³⁵. Porém, há quem defenda uma estratégia diferente que consiste na sequenciação direta de todo o genoma tendo em vista a identificação de alelos raros que exerçam grandes efeitos no fenótipo⁴⁰. As portas para o recurso generalizado a esta estratégia começam agora a abrir-se com o aparecimento e aperfeiçoamento de novas plataformas de sequenciação de terceira geração que permitem a sequenciação do genoma completo a um custo cada vez mais acessível. Estima-se que o genoma de cada indivíduo contenha aproximadamente 10 000 variantes não sinónimas entre cerca de 3,5 milhões de SNP. Tendo em mente a ordem de grandeza destes valores, é compreensível que este tipo de sequenciação venha a dominar os estudos genéticos nos próximos anos, tendência que já hoje se verifica⁴¹⁻⁴⁵.

Mas a descoberta de novas variantes genéticas associadas a doenças complexas pode não ser suficiente para descodificar a «matéria negra da hereditariedade», dado que esta pode ser o produto de interações complexas entre fatores genéticos, genómicos e epigenéticos^{39,46}. E a um outro nível, também o fenótipo pode ser consequência de interações não lineares entre diversos fatores de natureza genética e ambiental.

Apesar disso, é muito importante ir enriquecendo o catálogo das variantes genéticas associadas a um determinado fenótipo complexo, até porque a descoberta de novas variantes pode abrir caminho para chegar a via moleculares até então insuspeitas de influenciar a doença, como aconteceu com o exemplo que temos vindo a referir.

Os três *reports* aqui revistos representam abordagens exemplares e paradigmáticas de como se pode partir de uma associação estatística «cega» fornecida pelos GWAS, até encontrar uma explicação mecanística, mais ou menos esclarecida, do modo como uma variação genética pode modular um certo fenótipo. Neste caso, os GWAS orientaram para um ponto de partida particular, o *locus* 1p13^{14,21-23}, que acabaria por despertar a atenção de três equipas independentes, levando-as a enveredar por diferentes abordagens experimentais para compreender em que se fundamentava a associação estatística, mas todas apostando no gene *SORT1* como o modulador chave dos níveis de c-LDL e risco de EAM. Se os resultados foram concordantes na indicação de que *SORT1* tinha um papel importante na regulação do metabolismo lipoproteico, já foram menos consonantes as interpretações sobre a direção do efeito da expressão do gene nos níveis plasmáticos do c-LDL, bem como sobre o mecanismo subjacente¹⁰⁻¹².

Face às suas observações, Linsel-Nitschke et al.¹¹ propuseram que a sobre-expressão da sortilina aumentava a internalização de LDL, com conseqüente diminuição dos respectivos níveis plasmáticos¹¹. Pouco depois, Musunuru et al.¹⁰ reportaram uma relação inversa entre a expressão da sortilina e a concentração de LDL circulante, avançando com um mecanismo explicativo que passava pela regulação transcrricional, específica ao nível do fígado, do gene *SORT1* por fatores de transcrição C/EBP α , em que a sobre-expressão da *SORT1* reduziria a produção/secreção de VLDL¹⁰. Pelo contrário, Kjolby et al.¹² observaram uma relação direta entre a expressão da *Sort1* e a concentração de LDL circulante, sugerindo que tal poderia resultar de secreção aumentada das VLDL¹¹.

Entre as várias explicações que foram já apresentadas para justificar a discrepância entre resultados, a resposta parece em parte residir na própria sortilina, que emerge cada vez mais como uma proteína multifacetada podendo assumir diferentes funções consoante as circunstâncias.

Em síntese, os estudos aqui revistos proporcionaram uma série de provas consistentes de que *SORT1* é um elemento regulador dos níveis de c-LDL no plasma, o que veio acrescentar ao gene que codifica a sortilina uma importante função de que não se desconfiava até há bem pouco tempo. De momento, é ainda controversa a via celular que relaciona a sortilina com o metabolismo lipídico, mas certamente que nos próximos anos a questão vai continuar a ser explorada, sendo fundamental para avaliar se a sortilina pode representar um alvo potencial para abordagens terapêuticas direcionadas para hipercolesterolemia ou DCV³³⁻³⁵.

Responsabilidades éticas

Proteção de pessoas e animais. Os autores declaram que para esta investigação não se realizaram experiências em seres humanos e/ou animais.

Confidencialidade dos dados. Os autores declaram que não aparecem dados de pacientes neste artigo.

Direito à privacidade e consentimento escrito. Os autores declaram que não aparecem dados de pacientes neste artigo.

Conflito de interesses

Os autores declaram não haver conflito de interesses.

Bibliografia

1. Deaton C, Froelicher E, Wuc L, et al. The global burden of cardiovascular disease. *Eur J Cardiovasc Nurs*. 2011;10:55-13.
2. Anuário Estatístico de Portugal (2010). Instituto Nacional de Estatística. [consultado 18 Ago 2012]. Disponível em http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_publicacoes&PUBLICACOESpub_boui=133813349&PUBLICACOESmodo=2
3. 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.
4. 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.
5. Costa J, Borges M, Oliveira E, et al. Incidence and prevalence of hypercholesterolemia in Portugal: A systematic review. Part I *Rev Port Cardiol*. 2003;22:569-77.
6. Bourbon M. Genetic factors and cardiovascular disease. *Rev Port Cardiol*. 2008;27:1559-63.
7. 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.
8. Watkins H, Farrall M. Genetic susceptibility to coronary artery disease: From promise to progress. *Nat Rev Genet*. 2006;7:163-73.
9. Burton PR, Tobin MD, Hopper JL. Key concepts in genetic epidemiology. *Lancet*. 2005;366:941-51.
10. 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.
11. Linsel-Nitschke P, Samani NJ, Schunkert H. Sorting out cholesterol and coronary artery disease. *N Engl J Med*. 2010;363:2462-3.
12. 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.
13. 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.
14. Samani NJ, Erdmann J, Hall AS, et al. Genomewide association analysis of coronary artery disease. *N Engl J Med*. 2007;357:443-53.
15. 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.
16. 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.
17. Sandhu MS, Waterworth DM, Debenham SL, et al. LDL-cholesterol concentrations: A genome-wide association study. *Lancet*. 2008;371:483-91.

18. 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.
19. 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.
20. Kathiresan S, Willer CJ, Peloso GM, et al. Cupples Common variants at 30 loci contribute to polygenic dyslipidemia. *Nat Genet.* 2009;41:56–65.
21. McPherson R, Pertsemlidis A, Kavasar N, et al. A common allele on chromosome 9 associated with coronary heart disease. *Science.* 2007;316:1488–91.
22. 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.
23. 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.
24. 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.
25. 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.
26. 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.
27. 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.
28. 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.
29. Ni X, Canuel M, Morales CR. The sorting and trafficking of lysosomal proteins. *Histol Histopathol.* 2006;21:899–913.
30. Canuel M, Korkidakis A, Konnyu K. Sortilin mediates the lysosomal targeting of cathepsins D and H. *Biochem Biophys Res Commun.* 2008;373:292–7.
31. 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.
32. 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.
33. Tall AR, Ai D. Sorting out sortilin. *Circ Res.* 2011;108:158–60.
34. 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.
35. 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.
36. Schadt EE, Molony C, Chudin E, et al. Mapping the genetic architecture of gene expression in human liver. *PLoS Biol.* 2008;6:e107.
37. 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.
38. Marian AJ. Nature's genetic gradients and the clinical phenotype. *Circ Cardiovasc Genet.* 2009;2:537–9.
39. 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.
40. Cirulli ET, Goldstein DB. Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nat Rev Genet.* 2010;11:415–25.
41. 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.