

RESUMO

Este projecto apresenta uma abordagem multifactorial para a identificação de novos genes que influenciem o risco de sofrer acidentes vasculares cerebrais (AVCs). Os AVCs são a principal causa de morte em Portugal e a terceira maior causa de morte no conjunto dos países desenvolvidos. São ainda mais incapacitantes do que letais e os distúrbios neurológicos persistentes e a incapacidade física que provocam têm um custo socioeconómico muito elevado. Além disso, o número de pessoas afectadas deverá aumentar com o actual envelhecimento da população. Os AVCs são um “ataque cerebral” que interrompe o fluxo de sangue e oxigénio para as células de determinadas regiões do cérebro e são uma doença complexa resultante de factores genéticos e ambientais. Os principais factores de risco conhecidos incluem a idade, hipertensão, hipercolesterolemia, diabetes, doenças cardiovasculares, consumo de tabaco e de álcool e história familiar. As formas mais comuns dos AVCs podem ser classificadas em dois grandes tipos clínicos: AVCs isquémicos (AVCI; mais frequentes), ou hemorrágicos (10-20% dos casos). A identificação de novos genes que aumentem a susceptibilidade para se sofrer AVCs pode ter um forte impacto na saúde pública. A componente genética da doença tem sido demonstrada em estudos feitos em gémeos, famílias e modelos animais, e foram encontradas mutações em diversos genes em formas mendelianas raras de AVCs. No entanto, poucos genes de susceptibilidade para as formas comuns de AVCs foram até hoje identificados e os estudos de associação realizados apresentam geralmente resultados contraditórios.

Neste projecto, para atingir os objectivos propostos no estudo dos AVCI, começaram por realizar-se estudos de associação em alguns genes candidatos. Em paralelo, foi aplicada a abordagem de convergência genómica (CG) que combina estudos de ligação em todo o genoma, análises de expressão génica e estudos de associação em casos-controlos. Esta abordagem unificada, abrangente e multidisciplinar foi pela primeira vez implementada no estudo de AVCs. A disponibilidade de novas ferramentas genéticas, moleculares e estatísticas, bem como o seu sucesso no estudo de outras doenças complexas, tornam esta nova abordagem oportuna e essencial.

Uma vez que os genes que codificam a fosfodiesterase 4D (PDE4D) e a proteína activadora de araquidonato 5-lipoxigenase (ALOX5AP) têm sido nos últimos anos controversamente implicados com o risco de sofrer AVCI, avaliou-se neste projecto a sua associação com os AVCI na nossa amostra portuguesa. A proteína PDE4D degrada o segundo mensageiro adenosina monofosfato cíclica, uma molécula de transdução de sinal chave em diferentes tipos de células inflamatórias, vasculares endoteliais e musculares lisas. A proteína ALOX5AP está envolvida nos passos iniciais da síntese de leucotrienos e é secretada por vários tipos de células inflamatórias aglomeradas em locais lesados de vasos sanguíneos. Foram genotipados 67 polimorfismos de um só nucleótido (SNPs) na extremidade 5' do gene *PDE4D* e 24 SNPs localizados no gene *ALOX5AP* e ambas regiões flangeadoras de 10kb, em 565 pacientes caucasianos portugueses e 518 controlos independentes. Os SNPs seleccionados são *tagging* SNPs do projecto HapMap ou SNPs que foram encontrados previamente associados. Testaram-se as associações dos seus alelos, genótipos e haplótipos com o risco sofrer AVCI, utilizando testes padrão do qui-quadrado (χ^2) e regressões logísticas com múltiplas variáveis para ajustar as análises de associação com o risco para outros factores conhecidos, como a

hipertensão, diabetes e consumo de tabaco. Nenhum dos SNPs previamente associados foram replicados na nossa amostra, e considerando o número de testes realizados, não foi encontrada nenhuma associação significativa de outras variantes destes genes na susceptibilidade para sofrer AVCs na população portuguesa. Apenas o SNP rs7442640 do gene *PDE4D* se encontra associado (p -value = 0,006) com o risco de sofrer AVCs nos testes genotípicos ajustados para covariáveis e o SNP rs4491352 a jusante do gene *ALOX5AP* apresenta uma associação moderada ($0,017 < p$ -value $< 0,025$) para os testes alélicos e genotípicos não ajustados. Realizando uma meta-análise incluindo todos os estudos publicados recentemente e as nossas amostras portuguesas e espanholas para os SNP41, SNP45, SNP56, SNP87 e SNP89 (que se encontram associados na publicação original) do gene *PDE4D*, não foram encontrados resultados de associação com o risco de sofrer AVCs. No entanto, verificou-se que o SNP rs10507391 (ou SG13S114) do gene *ALOX5AP*, que faz parte do haplótipo HapA, está associado com o risco de sofrer AVCs tal como descrito no seu estudo original, tanto na população ibérica, como na meta-análise realizada. Estes resultados sugerem que o gene *PDE4D* pode não ser o principal factor de risco para os AVCs nas populações portuguesa e espanhola, contrastando com o gene *ALOX5AP* que pode conferir um risco aumentado para se sofrerem AVCs na Península Ibérica e outras populações.

Avaliou-se também a associação do gene kalirin (*KALRN*) com os AVCs na nossa amostra portuguesa uma vez que vários estudos recentes têm implicado algumas das suas variantes com a susceptibilidade para fenótipos cardiovasculares e metabólicos, mas nenhum estudo foi ainda realizado em AVCs. As doenças cerebrovasculares e cardiovasculares são ambas complexas, resultantes da interacção de factores genéticos e ambientais, e podem partilhar diversos genes de susceptibilidade. A proteína *KALRN* está envolvida, entre outras funções, na inibição da enzima óxido nítrico sintase induzida, na regulação da transdução de sinais de isquémia, e na morfogénese neuronal, plasticidade e estabilidade. O nosso objectivo foi verificar se SNPs na região do gene *KALRN*, no cromossoma 3q13, que inclui o gene roporina (*ROPNI*), predispõem para os AVCs na nossa amostra de pacientes e controlos portugueses. Foram genotipados 34 *tagging* SNPs na região cromossómica dos genes *KALRN* e *ROPNI*, em 565 pacientes com AVCs e 517 controlos, e realizada a imputação de genótipos para 405 marcadores no cromossoma 3. Testaram-se as associações de cada SNP individualmente e dos haplótipos que constituem com os AVCs. Um SNP na região intergénica *ROPNI-KALRN* (rs4499545) e dois SNPs no gene *KALRN* (rs17286604 e rs11712619) apresentam associações alélicas e genotípicas (não ajustadas e ajustadas para hipertensão, diabetes e consumo de tabaco) significativas ($0,003 < p$ -value $< 0,049$) com o risco de sofrer AVCs. Trinta e dois SNPs imputados encontram-se também associados com valores de p -value $< 0,05$, tendo sido validada a associação de três destes polimorfismos (rs7620580, rs6438833 e rs11712039) por genotipagem. Além disso, o SNP rs11712039 foi também associado com os AVCs ($0,001 < p$ -value $< 0,01$) no estudo de associação em todo o genoma (GWAS) publicado por Ikram e seus co-autores (2009). Esses resultados sugerem que variantes no gene *KALRN* constituem factores de risco para os AVCs, podendo ser factores de risco genéticos comuns das doenças vasculares.

Foi também testada a associação com o risco de sofrer AVCs do factor H inibidor do complemento (*CFH*), bem como de vários genes candidatos relacionados com mecanismos de neuroprotecção: eritropoietina (*EPO*), hemoxigenase-2 (*HO2*) e calicreína 1 (*KLK1*). Tem sido sugerido que a proteína *CFH*

desempenha um papel importante na inibição do complemento em lesões ateroscleróticas. Tem sido consistentemente associada com um risco aumentado para a degeneração macular relacionada à idade (AMD) e para os enfartes do miocárdio (MI) que compartilham vários factores de risco com os AVCs. Por outro lado, as proteínas EPO, HO2 e KLK1 são neuroprotetoras, por exemplo, em casos de hipóxia cerebral e isquémia (EPO) e contra acidente vascular cerebrais induzidos (KLK1). O polimorfismo no gene *CFH* (rs1061170) que foi previamente associado com a AMD e os MI parece estar modestamente associado (388 pacientes e 461 controlos portugueses; $0,030 < p\text{-value} < 0,035$) com o risco de sofrer AVCIs nos testes alélicos e genotípicos não ajustados, mas nenhum *tagging* SNP neste gene foi claramente associado com AVCIs no GWAS realizado por Ikram *et al.* (2009). Apesar de só ter sido estudado um SNP neste gene, estes resultados não justificam uma análise mais aprofundada. Foram genotipados 3, 3 e 5 *tagging* SNPs localizados nas regiões codificantes e regiões flanqueadoras de 10 kb dos genes *EPO*, *HO2* e *KLK1*, respectivamente, em 565 pacientes com AVCIs e 518 controlos da nossa amostra portuguesa. Não foram encontradas associações para nenhum marcador individualmente nem para haplótipos nestes genes neuroprotetores, sugerindo que eles não constituem factores de risco genéticos para AVCIs.

Para identificar novos genes de susceptibilidade para AVCIs, aplicou-se ainda a abordagem de CG proposta. Foram realizadas análises de expressão génica em células brancas mononucleares do sangue periférico de 20 casos com AVCIs e 20 controlos equilibrados para a idade e para o género, usando *Affymetrix GeneChip Human U133 Plus 2.0 arrays*, que representam 47.000 transcritos humanos e variantes. Foram identificadas várias vias metabólicas afectadas nos pacientes com AVCIs, tal como a via das moléculas de adesão celular. 16 dos genes diferencialmente expressos entre casos e controlos (para um *cut-off* de 1,2 *fold-change* e *p-values* não corrigidos $< 0,05$) encontram-se localizados em picos de ligação descritos em estudos de ligação do inteiro genoma em humanos. Todos os *tagging* SNPs desses genes e das suas regiões flanqueadoras de 10 kb (um total de 191 SNPs) foram genotipados em 565 casos com AVCIs e 520 controlos do nosso biobanco português. Foram realizados testes de associação de cada SNP individualmente e dos haplótipos que constituem. Os resultados de associação sugerem que existem variantes em 6 (*HEMGN*, *GFI1B*, *TMTC4*, *TTC7B*, *SDC4* e *TUBB1*) dos 16 genes priorizados que podem constituir factores de risco para os AVCIs na população portuguesa. Vários dos SNPs associados nestes genes também fazem parte dos haplótipos que se encontraram associados. SNPs intrónicos como os rs9582406 e rs946845 do gene *TMTC4*, e o rs2284278 do gene *SDC4*, apresentam associação ($0,015 < p\text{-value} < 0,050$) com risco de sofrer AVCIs em todos os testes realizados. Por outro lado, o SNP intrónico rs1535321 do gene *TTC7B* encontra-se associado ($p\text{-value} = 0,009$) nos testes alélicos e genotípicos não ajustados, embora não apresente associação no teste ajustado para hipertensão, diabetes e para o consumo de tabaco.

Para a replicação destes resultados, os SNPs que foram encontrados associados com $p\text{-value} < 0,05$ em pelo menos um dos testes realizados, assim como alguns SNPs que definem os haplótipos associados, foram genotipados numa amostra espanhola. Um total de 570 casos caucasianos com AVCIs e 390 controlos foram incluídos e foram realizados testes de associação dos seus alelos, genótipos e haplótipos com o risco de sofrer AVCIs, utilizando também testes de χ^2 e análises de regressão logística de múltiplas variáveis para ajustar as análises feitas para a hipertensão, diabetes, dislipidemia e para o consumo de tabaco. As mesmas

análises foram realizadas para os subtipos de AVCs aterotrombótico, cardioembólico e lacunar. Encontraram-se algumas associações significativas com o risco de sofrer AVCIs nos genes *TMTC4*, *TTC7B* e *SDC4*, no entanto, o único SNP replicado na amostra de casos-controlos espanhola foi o rs9582406 do gene *TMTC4*, significativamente associado nos testes não ajustados (p-value = 0,019). Este SNP está também associado com o risco para os subtipos aterotrombótico e lacunar de AVCs para os testes alélicos e genotípicos ($0,011 < \text{p-value} < 0,049$). Para os genes *TTC7B* e *SDC4*, os SNPs e os haplótipos associados na amostra espanhola não foram os mesmos que na portuguesa. O SNP rs6073708 do gene *SDC4* foi associado para todos os testes realizados ($0,027 < \text{p-value} < 0,029$). Nenhum dos SNPs estudados foi claramente replicado GWAS publicado recentemente por Ikram e seus colaboradores (2009).

Globalmente, os resultados obtidos sugerem que os genes *HEMGN* e *GFI1B* (que não foram replicados na amostra espanhola) podem constituir factores de risco para AVCIs na população portuguesa, sendo importante o alargamento da amostra portuguesa para validação dos resultados. Da mesma forma, o gene *TUBB1*, que não foi genotipado na amostra espanhola por problemas técnicos, pode constituir um factor de risco para AVCIs na população portuguesa, mas deve ser ainda estudado noutras amostras para validar e compreender o seu papel. Por outro lado, o gene *TMTC4* pode constituir um factor de risco para os AVCIs na população ibérica, e os genes *TTC7B* e *SDC4*, com a heterogeneidade observada nos seus resultados de associação entre diferentes amostras, podem ser factores de risco para ACVIs, sendo provável que as suas verdadeiras variantes de susceptibilidade não tenham sido estudadas ainda. O gene *TTC7B* foi um dos mais associados para as principais doenças cardiovasculares no GWAS do projecto *Framingham Heart Study 100K* (Larson et al. 2007). Para este gene, vários SNPs e haplótipos na região intrão 5 - intrão 6 associados nas amostras portuguesa e espanhola, encontram-se modestamente associados no GWAS publicado por Ikram e seus colaboradores (2009). Existem portanto várias linhas de evidência independentes que apoiam a importância do gene *TTC7B* na susceptibilidade dos AVCs, embora futuros trabalhos sejam necessários para identificar variantes de risco exatas e para elucidar o seu potencial patogénico.

Neste projecto, dado o grande número de SNPs testados, nenhum dos resultados significativos obtidos sobrevive à correcção para múltiplos testes. No entanto, é geralmente aceite que a replicação em múltiplas amostras independentes continua a ser a melhor estratégia para os estudos de associação, mesmo para associações modestas. Se os resultados apresentados puderem ser confirmados noutras amostras independentes e um estudo mais completo de outras possíveis variações genéticas nos loci de interesse forem realizados, pensamos que estudos funcionais dos genes e das suas variantes de susceptibilidade, permitirão uma melhoria significativa do nosso conhecimento sobre AVCs. Pode recorrer-se à sequenciação de nova geração para identificar com precisão as verdadeiras variantes de susceptibilidade, ou variantes raras que os estudos de associação não tenham poder para detectar.

É sugerido que a identificação dos determinantes genéticos dos AVCs utilizando diferentes estratégias e populações, e a sua análise de uma forma integrada, como a realizada neste projecto, é a forma mais completa para o seu estudo, a fim de melhorar o conhecimento da doença.