

## Kalirin: a novel genetic risk factor for ischemic stroke

Tiago Krug · Helena Manso · Liliana Gouveia · João Sobral · Joana M. Xavier · Isabel Albergaria · Gisela Gaspar · Manuel Correia · Miguel Viana-Baptista · Rita Moiron Simões · Amélia Nogueira Pinto · Ricardo Taipa · Carla Ferreira · João Ramalho Fontes · Mário Rui Silva · João Paulo Gabriel · Ilda Matos · Gabriela Lopes · José M. Ferro · Astrid M. Vicente · Sofia A. Oliveira

Received: 23 March 2009 / Accepted: 17 January 2010 / Published online: 28 January 2010  
© Springer-Verlag 2010

**Abstract** Cerebrovascular and cardiovascular diseases are the leading causes of death and disability worldwide. They are complex disorders resulting from the interplay of genetic and environmental factors, and may share several susceptibility genes. Several recent studies have implicated variants of the Kalirin (*KALRN*) gene with susceptibility to cardiovascular and metabolic phenotypes, but no studies have yet been performed in stroke patients. *KALRN* is involved, among others, in the inhibition of inducible nitric oxide synthase, in the regulation of ischemic signal transduction, and in neuronal morphogenesis, plasticity, and stability. The goal of the present study was to determine whether SNPs in the *KALRN* region on 3q13, which

includes the Ropporin gene (*ROPNI*), predispose to ischemic stroke (IS) in a cohort of Portuguese patients and controls. We genotyped 34 tagging SNPs in the *KALRN* and *ROPNI* chromosomal region on 565 IS patients and 517 unrelated controls, and performed genotype imputation for 405 markers on chromosome 3. We tested the single-marker association of these SNPs with IS. One SNP (rs4499545) in the *ROPNI*–*KALRN* intergenic region and two SNPs in *KALRN* (rs17286604 and rs11712619) showed significant ( $P < 0.05$ ) allelic and genotypic (unadjusted and adjusted for hypertension, diabetes, and ever smoking) association with IS risk. Thirty-two imputed SNPs also showed an association at  $P < 0.05$ , and actual genotyping of

**Electronic supplementary material** The online version of this article (doi:10.1007/s00439-010-0790-y) contains supplementary material, which is available to authorized users.

T. Krug · H. Manso · J. Sobral · J. M. Xavier · A. M. Vicente · S. A. Oliveira  
Instituto Gulbenkian de Ciência, Oeiras, Portugal

T. Krug · J. M. Xavier · S. A. Oliveira (✉)  
Instituto de Medicina Molecular, Av. Prof. Egas Moniz,  
Edifício Egas Moniz, 1649-028 Lisbon, Portugal  
e-mail: aaoliveira@fm.ul.pt

H. Manso · J. Sobral · I. Albergaria · G. Gaspar · A. M. Vicente  
Instituto Nacional de Saúde Dr. Ricardo Jorge,  
Lisbon, Portugal

L. Gouveia · J. M. Ferro  
Serviço de Neurologia,  
Hospital de Santa Maria, Lisbon, Portugal

M. Correia · R. Taipa · G. Lopes  
Serviço de Neurologia,  
Hospital Geral de Santo António, Porto, Portugal

M. Viana-Baptista  
Serviço de Neurologia,  
Hospital Garcia de Orta, Almada, Portugal

R. M. Simões · A. N. Pinto  
Serviço de Neurologia,  
Hospital Fernando Fonseca, Amadora, Portugal

C. Ferreira · J. R. Fontes  
Serviço de Neurologia,  
Hospital São Marcos, Braga, Portugal

M. R. Silva · J. P. Gabriel  
Serviço de Neurologia,  
Hospital de São Pedro, Vila Real, Portugal

I. Matos  
Serviço de Neurologia,  
Hospital Distrital de Mirandela, Mirandela, Portugal

three of these polymorphisms (rs7620580, rs6438833, and rs11712039) validated their association. Furthermore, rs11712039 was associated with IS ( $0.001 < P < 0.01$ ) in a recent well-powered genomewide association study (Ikram et al. 2009). These studies suggest that variants in the *KALRN* gene region constitute risk factors for stroke and that *KALRN* may represent a common risk factor for vascular diseases.

## Introduction

Cerebrovascular diseases, such as stroke, and cardiovascular diseases, such as coronary artery disease (CAD) and myocardial infarction (MI), are among the top three leading causes of death and disability worldwide (Murray and Lopez 1997). They are complex disorders resulting from the interplay of genetics and environment, and they share many risk factors, including age, sex, hypertension, dyslipidemia, diabetes, obesity, smoking, and physical inactivity. These atherothrombotic diseases most likely also share common pathogenic mechanisms such as inflammation and appear to have common susceptibility loci. For instance, a locus on 9p21 has been firmly associated with vascular pathologies such as heart disease, stroke, aneurysms, and atherosclerosis (Matarin et al. 2008; Larson et al. 2007; McPherson et al. 2007; Helgadottir et al. 2007, 2008; Karvanen et al. 2009; O'Donnell et al. 2007; Ye et al. 2008).

The GENECARD study for early-onset CAD (Hauser et al. 2004), the Diabetes Heart Study (DHS) for cardiovascular disease (CVD), type 2 diabetes (T2D) and metabolic syndrome (MS) (Bowden et al. 2006), and a meta-analysis of four linkage studies for coronary heart disease (Chiodini and Lewis 2003) identified linkage peaks on chromosome 3q. Ordered subset analysis significantly increased the evidence for linkage at 3q13 (LOD = 5.10,  $P = 0.008$ ) in GENECARD families with lower-risk lipid profiles and fewer risk factors (Shah et al. 2006), and subsequent peak-wide association mapping led to the identification of twelve SNPs in the Ropporin (*ROPNI*, OMIM 611757) and *KALRN* (OMIM 604605) genes associated with early-onset CAD (Wang et al. 2007). Validation in additional datasets revealed that SNP rs9289231 in the first intron of a *KALRN* alternative transcript was associated with early-onset CAD in all white data sets examined, and the risk allele of this SNP was associated with atherosclerosis burden in human aortas (Wang et al. 2007). Additionally, follow-up of the DHS 3q linkage peak revealed that the *KALRN* polymorphism rs4234218 is associated with T2D, MS, and combined phenotype (T2D + MS + CVD + coronary calcified plaque) (Rudock et al. 2008).

Furthermore, a recent and well-powered genomewide association study (GWAS) conducted on four European

and American cohorts including 1,164 ischemic stroke patients and 18,058 controls (Ikram et al. 2009) showed that several polymorphisms in Kalirin are associated with IS ( $0.001 < P < 0.01$ ), even though below the genomewide significance level.

Kalirin is an extremely large gene with 60 exons spanning over 620 kb, characterized by multiple promoters producing developmentally regulated isoforms predominantly expressed in the brain (McPherson et al. 2004). *KALRN* encodes for a guanine nucleotide exchange factor (GEF) that activates Rho proteins, and is therefore a multifunctional protein involved, among others, in neuronal morphogenesis and secretory granule maturation (Ferraro et al. 2007; Rabiner et al. 2005). Additionally, Kalirin may have a neuroprotective role by inhibiting inducible nitric-oxide synthase (iNOS) activity (Ratovitski et al. 1999), and participates in the regulation of ischemic signal transduction (Beresevicz et al. 2008).

The goal of the present study was to confirm if SNPs in the *ROPNI*–*KALRN* region predispose to ischemic stroke (IS).

## Patients and methods

### Study subjects

Five-hundred sixty-five unrelated patients with a clinical diagnosis of IS, who were under the age of 65 at stroke onset, were recruited through Neurology and Internal Medicine Departments throughout Portugal. Stroke was defined by the presence of a new focal neurological deficit, with an acute onset and with symptoms and signs persisting for more than 24 h. The stroke was confirmed in all patients by a computed tomography scan in 97% of cases and/or magnetic resonance imaging in 25% of patients. All patients were seen, and all neuroradiology tests were reviewed by study neurologists. Trauma, tumors, infection, and other causes of neurological deficit were excluded.

Data collection forms were developed for this study that included extensive clinical information such as stroke characteristics, general clinical observation, neurological symptoms and signs, complications and interventions during hospitalization, and situation at discharge. Data were also collected on relevant lifestyle aspects and previous clinical risk factors.

Five-hundred seventeen unrelated healthy individuals were included in this study as a control sample population. Since stroke is a late-onset disease, the control group was selected from a group of healthy volunteers with a higher mean age than the case group, thus minimizing the chances for mis-classification as “stroke-free”. Control individuals were verified to be free of stroke by direct interview before

recruitment, but no brain imaging studies were performed. The interview also included questions on established clinical and lifestyle risk factors for stroke. All participants were adults of Portuguese Caucasian origin.

The research protocol was approved by the Ethics Committees of participating institutions, and all participants provided informed consent.

### SNP selection and genotyping

All SNPs originally found associated with CAD in the CATHGEN initial dataset (Wang et al. 2007) were genotyped first and then additional haplotype tagging SNPs were included in the study. Genotypes of 30 European (CEU) family trios were downloaded from the HapMap Release 21/phaseII Jul06, on NCBI B35 assembly, and haplotype tagging SNPs in the *ROPNI–KALRN* region were identified in Haploview 4.0 (Barrett et al. 2005) with the following options: pairwise mode,  $r^2 > 0.8$ , and minor allele frequency (MAF)  $> 0.1$ . A total of 34 SNPs were genotyped in a 384-well format, 31 of which using Sequenom's (San Diego, USA) iPLEX assay (primer extension of multiplex products with detection by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry) following manufacturer's protocol and detected in a Sequenom MassArray K2 platform. The primer sequences are available upon request and were designed using Sequenom's MassARRAY<sup>®</sup> Assay Design 3.0. Extensive quality control was performed using eight HapMap controls of diverse ethnic affiliation, sample duplication within and across plates, non-Mendelian inheritance check in three large pedigrees, Hardy–Weinberg equilibrium (HWE) in the control group ( $P > 0.05$ ), and a minimum of 90% call rate. Genotype determinations were performed blinded to affection status. The SNPs which failed at least one quality control (rs12634530, rs12637456, and rs13075202) or that could not be genotyped (rs1444754) using the iPLEX assay were genotyped using the TaqMan allelic discrimination assays (assays-on-demand C\_\_11236577\_10, C\_\_11236574\_10, C\_\_1720000\_10, and C\_\_9532531\_10, respectively) from Applied Biosystems (Foster City, USA). Even though SNP rs13075202 was out of HWE in the control dataset ( $P < 0.05$ ), it was considered for further analysis as it was genotyped consistently in two assays (iPLEX and assay-on-demand) and passed all other quality checks. All the genotyping was performed in the Genomics Unit of the Instituto Gulbenkian de Ciência.

### Statistical analysis

An unpaired Student's *t* test and a  $\chi^2$  test were used to compare quantitative and qualitative clinical and demographic data, respectively, between cases and controls.  $\chi^2$  tests for

HWE, allelic association of SNPs with IS risk, and linkage disequilibrium (LD) plots were performed using Haploview 4.0 (Barrett et al. 2005).

The genotypic association between IS and each SNP was examined by logistic regression analyses using the SNPAssoc v.1.4-9 package (González et al. 2007) implemented in the R freeware (<http://cran.r-project.org/>). To adjust the association analyses for relevant confounding factors, hypertension, diabetes, and ever smoking were included as covariates in multivariate logistic regression with backward elimination of risk factors. The interaction *i* among these covariates in regression models was not strong ( $-0.5 < i < 0.5$ ). Odds ratios (ORs) and their associated 95% confidence intervals (CIs) were uncorrected for confounding variables in the  $\chi^2$  tests and unadjusted logistic regression, and corrected for covariates in adjusted regression models. Results were considered significant below the conventional level of 0.05. Since there is LD among some of the genotyped SNPs, we did not perform corrections for multiple testing, and uncorrected *P* values are reported.

### Imputation

Un genotyped SNPs in chromosome 3 were imputed with PLINK v1.04 (Purcell et al. 2007) using HapMap data (Release 22, 154783 SNPs in chromosome 3 with MAF greater than 0.01 and genotyping rate greater than 0.95 in the 60 CEU founders). For every imputed SNP, PLINK provides an information content metric INFO, ranging from 0 to 1 (although it can be greater than 1 occasionally). A higher INFO value generally means a better SNP imputation. All imputed SNPs with MAF in controls smaller than 0.05 and with INFO  $< 0.5$  were excluded. For SNPs that have been genotyped, PLINK calculates the concordance rate among observed and imputed genotypes.

To confirm the imputation results, three imputed SNPs (rs7620580, rs6438833, and rs11712039) significantly associated with IS were fully genotyped in our case–control biobank, and tested for association as described previously.

## Results

Table 1 summarizes the principal characteristics of the dataset used in this study. The average age at the first stroke is  $51.7 \pm 9.5$  years. Since stroke is a late-onset disease, we chose to have the control group significantly older than the case group to minimize misclassification biases. As expected, male:female ratio, hypertension, diabetes, ever smoking, and ever drinking were significantly higher in IS patients than in controls.

**Table 1** General characteristics of the ischemic stroke case–control study sample

Characteristic	Controls	Cases	<i>P</i> <sup>a</sup>
<i>N</i>	517	565	
Sex ( <i>n/N</i> , % male)	238/517 (46.0)	361/565 (63.9)	<10 <sup>−4</sup>
Age-at-examination (mean ± SD, years)	63.0 ± 6.8	52.4 ± 9.3	<10 <sup>−4</sup>
Age-at-onset (mean ± SD, years)	–	51.7 ± 9.5	–
Risk factors ( <i>n/N</i> , %)			
Hypertension (>140–85 mmHg)	192/508 (37.8)	289/505 (57.2)	<10 <sup>−4</sup>
Hypercholesterolemia (>200 mg/dL)	324/515 (62.9)	328/526 (62.4)	0.853
Hypertriglycemia (>200 mg/dL)	68/437 (15.6)	42/229 (18.3)	0.359
Diabetes	57/496 (11.5)	95/538 (17.7)	0.005
Ever smoking	146/507 (28.8)	272/556 (48.9)	<10 <sup>−4</sup>
Ever drinking	218/500 (43.6)	326/558 (58.4)	<10 <sup>−4</sup>

SD standard deviation

<sup>a</sup> *P* of an unpaired Student's *t* test or a  $\chi^2$  test for quantitative and qualitative data, respectively

We first tested the association with IS of the 12 SNPs (Table 2), originally found associated with CAD in the CATHGEN initial dataset (SNPs 2, 3, 4, 7, 8, 9, 12, 14, 20, 22, 31, and 37 in Wang et al. 2007), which includes the rs4234218 (SNP 31 in Wang et al. 2007) associated with cardiovascular risk, T2D, and MS (Rudock et al. 2008). These polymorphisms are located in *ROPNI*, *ROPNI*–*KALRN* intergenic region and in the 5' region of the *KALRN* gene. In our dataset, all of these polymorphisms are also haplotype tagging SNPs, with the exception of rs7613868 and rs12637456 (SNPs 7 and 9, respectively, in Wang et al. 2007) which are in almost complete LD ( $r^2 = 0.97$ ; Fig. 1).

We found that a cluster of SNPs in low LD (pairwise  $r^2 \leq 0.56$ ) in the *ROPNI*–*KALRN* intergenic region (rs4499545, rs7613868/rs12637456, and rs12634530) was associated with IS in allelic and unadjusted genotypic tests (Fig. 1; Supplementary Table 1). The association of SNP rs4499545 remained significant after adjustment for covariates ( $P = 0.028$ , OR [95% CI] = 1.31 [1.03–1.67]). rs9289231, which has been associated with early-onset CAD and atherosclerosis burden in human aortas (Wang et al. 2007) and rs4234218 were not associated with IS in any test performed (Fig. 1).

Given the evidence for association of genetic variants in the *ROPNI*–*KALRN* region with IS, although with a different variant than previously reported to be associated with cardiovascular risk, we decided to further investigate this region by analyzing additional SNPs. We genotyped 22

additional haplotype tagging SNPs in the *ROPNI* region, the *ROPNI*–*KALRN* intergenic region, and in the 5' region of *KALRN* (Table 2). Since the *KALRN* gene is extremely large, we chose to concentrate our efforts on the *KALRN* region with the reported associations.

Two newly analyzed SNPs in the *ROPNI*–*KALRN* intergenic region LD block (rs2280422 and rs2332719) and two SNPs in *KALRN* (rs17377867 and rs6784664) were associated ( $0.003 < P < 0.042$ ) in allelic and unadjusted log-additive model tests, but these associations became marginal ( $0.05 < P < 0.15$ ) when adjusted for covariates (Fig. 1; Supplementary Table 1). Inclusion of any combination of two of these covariates had the same overall effect, but it is not always the same polymorphisms that become marginally significant, suggesting that this may be mostly a power issue and not a biologically relevant phenomenon. On the other hand, two new SNPs in *KALRN* (rs17286604 and rs11712619) demonstrated association in all tests performed (Fig. 1; Supplementary Table 1).

Table 3 presents more detailed association results for the three genotyped SNPs demonstrating an association with IS in all tests performed (rs4499545, rs17286604, and rs11712619). Similar to the rs4234218 association with cardiovascular risk, T2D, and MS (Rudock et al. 2008), the most significant results for these three SNPs were under the dominant genetic model (Table 3). These polymorphisms are in very low LD (all pairwise  $r^2 < 0.34$  and  $D' < 0.58$ ). SNP rs4499545 remains significantly associated when the other two markers are used as covariates, while the association subsides for rs11712619 when adjusted for rs17286604. These data suggest at least two independent lines of evidence for association of *KALRN* with IS.

Even though we investigated 23 haplotype tagging SNPs in the 5' end of *KALRN*, a large portion of its natural genetic variation was not assessed because it is a very large gene with a high degree of genetic diversity. We therefore performed genotype imputation for SNPs in chromosome 3 using data from HapMap as well as the genotypes observed at the 34 fully genotyped polymorphisms. We obtained imputed genotypes meeting minimum quality standards (MAF in controls  $\geq 0.05$  and SNP INFO  $\geq 0.5$ ) for 405 SNPs in and around *ROPNI* and *KALRN* (Fig. 2). As expected, the imputed SNPs with higher information content metric (SNP INFO  $> 0.8$ ) are located in the region of genotyped SNPs. As an additional quality control, the genotyped SNPs are dropped one at a time and imputed using the other observed genotypes. A concordance rate between observed and imputed genotypes can be calculated for these markers, and 90% of the genotyped SNPs had a concordance rate  $>85\%$  between imputed and observed genotypes (Fig. 2). Thirty-two of the imputed polymorphisms have an allelic association with IS risk at a more stringent *P* threshold

**Table 2** Characteristics of the investigated SNPs

SNP ID	rs number	SNP name in Wang et al. 2007	Position in chromosome 3 (bp)	Gene	Genotyping assay <sup>a</sup>	HWE <i>P</i> in controls	Allele 1:2	Frequency of allele 2 in controls
1	rs7633408		125166534	-	I	0.110	C:G	0.130
2	rs6810298	2	125177665	<i>ROPN1</i>	I	0.846	A:G	0.350
3	rs17376453	3	125179139	<i>ROPN1</i>	I	0.111	G:C	0.149
4	rs2280422		125179620	<i>ROPN1</i>	I	0.147	G:A	0.418
5	rs4499545 <sup>b</sup>	4	125188532	-	I	0.084	G:A	0.174
6	rs2332719		125195656	-	I	0.162	A:G	0.329
7	rs7613868	7	125224719	-	I	0.320	C:T	0.329
8	rs12634530	8	125227160	-	T	1.000	C:T	0.216
9	rs12637456	9	125227353	-	T	0.365	T:A	0.332
10	rs1317671		125245276	<i>AK123068</i>	I	0.096	T:C	0.304
11	rs9289231	12	125256768	<i>AK123068</i>	I	0.808	T:G	0.103
12	rs13075202	14	125304977	<i>KALRN</i>	I and T	0.036	A:G	0.302
13	rs4678085		125314057	<i>KALRN</i>	I	0.535	G:A	0.230
14	rs1950091		125383007	<i>KALRN</i>	I	0.297	G:A	0.149
15	rs1444760		125402970	<i>KALRN</i>	I	0.180	G:A	0.361
16	rs1444768	20	125406612	<i>KALRN</i>	I	0.270	A:G	0.404
17	rs1444766		125407961	<i>KALRN</i>	I	0.259	A:G	0.269
18	rs6438833		125408285	<i>KALRN</i>	I	0.199	T:A	0.079
19	rs1444754	22	125420907	<i>KALRN</i>	T	0.232	T:C	0.433
20	rs1373609		125432219	<i>KALRN</i>	I	0.537	T:C	0.483
21	rs17286604		125434907	<i>KALRN</i>	I	0.129	C:T	0.364
22	rs1158012		125440085	<i>KALRN</i>	I	0.818	G:A	0.463
23	rs4608634		125443748	<i>KALRN</i>	I	0.526	G:C	0.430
24	rs4234218	31	125443900	<i>KALRN</i>	I	0.693	G:C	0.365
25	rs4608635		125444186	<i>KALRN</i>	I	0.617	G:C	0.335
26	rs6781700		125450642	<i>KALRN</i>	I	0.349	T:C	0.078
27	rs11712039		125456276	<i>KALRN</i>	I	0.377	C:T	0.378
28	rs1444763		125457660	<i>KALRN</i>	I	0.743	A:G	0.279
29	rs11929003	37	125462011	<i>KALRN</i>	I	1.000	A:G	0.459
30	rs13064819		125468805	<i>KALRN</i>	I	0.620	G:C	0.163
31	rs11712619		125502492	<i>KALRN</i>	I	0.127	C:T	0.367
32	rs17377867		125504693	<i>KALRN</i>	I	0.366	G:A	0.110
33	rs9838361		125517439	<i>KALRN</i>	I	0.211	G:T	0.355
34	rs6784664		125524089	<i>KALRN</i>	I	0.899	C:A	0.467
35	rs7620580		125527993	<i>KALRN</i>	I	0.267	A:G	0.116
36	rs3821525		125537714	<i>KALRN</i>	I	1.000	A:G	0.271
37	rs11708466		125544276	<i>KALRN</i>	I	0.754	T:C	0.300

The 12 SNPs that were first genotyped are highlighted in dark gray and the three SNPs that were genotyped to validate the imputation results are highlighted in light gray

<sup>a</sup> *I* iPlex; *T* TaqMan

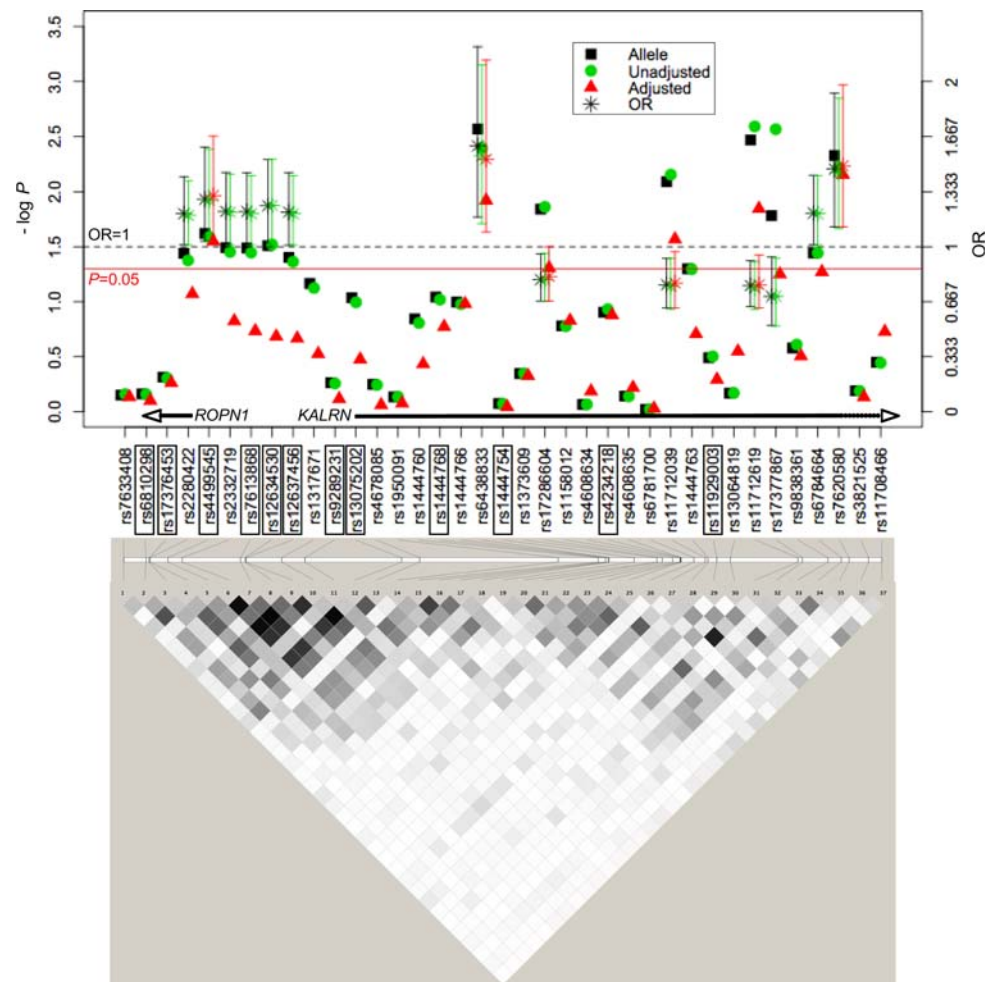
<sup>b</sup> rs4499545 is the current name for SNP rs7434266

of 0.01 (Fig. 2; Supplementary Table 2). Among the top five SNPs ( $P \leq 0.001$ ), the first one (rs6790975) has a modest SNP INFO metric (0.598), and therefore its association has to be taken with caution, but the remaining four markers (rs4678111, rs7620580, rs9820396, and rs1444770) have a very high SNP INFO metric (0.93) and are in complete LD (Fig. 2). The LD among the top 32 imputed associated SNPs (Fig. 2) also suggests the existence of several independent clusters of association with IS in the 5' region

of the *KALRN* gene, such as the rs6438833–rs1373612, rs6779809–rs11712039–rs11719349 and rs4678111–rs7620580–rs9820396–rs1444770 clusters.

To validate the imputation results, we additionally genotyped in our dataset three SNPs (rs6438833, rs11712039, and rs4678111) representing the abovementioned clusters of association (highlighted in light gray in Table 2), and tested their association with IS. The average concordance rate among observed and imputed genotypes was 92%. All

**Fig. 1** Association results and pairwise linkage disequilibrium (LD) among all genotyped polymorphisms. Allelic (black squares) as well as crude (green disks) and adjusted (red triangles) genotypic (log-additive model) association results are shown. Odds ratios (ORs) and 95% confidence intervals (CIs) are depicted with stars and bars for significantly associated ( $P < 0.05$ ) polymorphisms. In the LD plot, the white–black gradient shading within each diamond represents the magnitude of LD using the pairwise LD statistic  $r^2$  (from lower to higher values, respectively). The positions of the SNPs relative to the *ROPN1* and *KALRN* genes are indicated. Boxes surrounding SNP names indicate the 12 polymorphisms that we first genotyped



three polymorphisms were associated with IS (Table 3; Supplementary Table 1).

Among the SNPs tested and associated with IS in this study, SNP rs11712039 has also been associated with IS ( $0.001 < P < 0.01$ ) in the Ikram et al. (2009) IS GWAS, as shown in Supplementary Table 1. rs17286604 and rs11712619, which are associated with IS in our Portuguese sample, were modestly associated in the GWAS, and SNP rs4499545 has not been tested (Supplementary Table 1).

## Discussion

In this study, we highlight for the first time an association of the Kalirin gene with risk for ischemic stroke. Given that *KALRN* has already been implicated in susceptibility to cardiovascular disorders and some of its risk factors (e.g. T2D) (Wang et al. 2007; Rudock et al. 2008; Ikram et al. 2009), we propose that Kalirin may constitute a novel genetic risk factor for vascular phenotypes.

Kalirin is an extremely complex gene generating many alternative transcripts under the control of several promot-

ers in a tissue-specific and developmentally regulated manner (McPherson et al. 2004; Johnson et al. 2000; Ma et al. 2001) encoding multidomain and multifunctional proteins (McPherson et al. 2002). A number of recent studies have demonstrated that Kalirin-7, the most abundant isoform in adult brain (Penzes et al. 2000), is involved in dendritic spine development, plasticity, and stability (Penzes and Jones 2008). Kalirin-7 gene transcripts are underexpressed in Alzheimer's disease (AD) hippocampal specimens (Youn et al. 2007a), and underexpression of Kalirin-7 increases iNOS (inducible nitric oxide (NO) synthase) activity in cultured cells and correlates with increased iNOS activity in AD hippocampus (Youn et al. 2007b). Interestingly, Kalirin is known to associate with iNOS in vitro and in vivo and to prevent iNOS dimerization, and therefore iNOS activity (Ratovitski et al. 1999). NO, a potent cell-signaling, effector, and vasodilator molecule characterized by its strong reactivity and diffusibility, is produced in a complex and tightly controlled process by three NO synthases: iNOS, neuronal synthase (nNOS), and endothelial synthase (eNOS). Cytokines, endotoxin, or other proinflammatory stimuli induce iNOS expression in

**Table 3** Detailed association results for the most consistently associated SNPs

SNP Model		Controls (%)	Cases (%)	Unadjusted		Adjusted	
				OR [95% CI]	<i>P</i>	OR [95% CI]	<i>P</i>
rs4499545							
Codominant	1/1	69.4	61.5	1.00	<b>0.022</b>	1.00	0.055
	1/2	26.4	34.2	1.46 [1.11–1.91]		1.43 [1.06–1.93]	
	2/2	4.2	4.3	1.14 [0.62–2.10]		1.33 [0.67–2.65]	
Dominant	1/1	69.4	61.5	1.00	<b>0.008</b>	1.00	<b>0.016</b>
	1/2–2/2	30.6	38.5	1.42 [1.09–1.83]		1.42 [1.07–1.89]	
Recessive	1/1–1/2	95.8	95.7	1.00	0.974	1.00	0.615
	2/2	4.2	4.3	1.01 [0.55–1.85]		1.19 [0.60–2.35]	
log-Additive	0, 1, 2			1.28 [1.03–1.59]	<b>0.026</b>	1.31 [1.03–1.67]	<b>0.028</b>
rs6438833							
Codominant	1/1	78.3	85.2	1.00	<b>0.016</b>	1.00	<b>0.042</b>
	1/2	19.0	13.7	1.51 [1.06–2.16]		1.49 [1.01–2.21]	
	2/2	2.6	1.1	2.70 [0.93–7.85]		2.58 [0.77–8.60]	
Dominant	1/1	78.3	85.2	1.00	<b>0.007</b>	1.00	<b>0.018</b>
	1/2–2/2	21.7	14.8	1.60 [1.13–2.25]		1.57 [1.08–2.28]	
Recessive	1/1–1/2	97.4	98.9	1.00	0.076	1.00	0.136
	2/2	2.6	1.1	2.52 [0.87–7.32]		2.41 [0.72–8.04]	
log-Additive	0, 1, 2			1.55 [1.14–2.10]	<b>0.004</b>	1.53 [1.09–2.13]	<b>0.012</b>
rs17286604							
Codominant	1/1	38.9	47.9	1.00	<b>0.011</b>	1.00	<b>0.028</b>
	1/2	49.5	41.5	0.68 [0.53–0.88]		0.68 [0.51–0.90]	
	2/2	11.6	10.6	0.74 [0.50–1.11]		0.80 [0.52–1.23]	
Dominant	1/1	38.9	47.9	1.00	<b>0.003</b>	1.00	<b>0.010</b>
	1/2–2/2	61.1	52.1	0.69 [0.54–0.88]		0.70 [0.54–0.92]	
Recessive	1/1–1/2	88.4	89.4	1.00	0.613	1.00	0.898
	2/2	11.6	10.6	0.91 [0.62–1.33]		0.97 [0.65–1.47]	
log-Additive	0, 1, 2			0.80 [0.67–0.96]	<b>0.014</b>	0.82 [0.67–1.00]	<b>0.049</b>
rs11712039							
Codominant	1/1	46.2	39.0	1.00	<b>0.022</b>	1.00	0.070
	1/2	45.8	48.5	0.79 [0.60–1.05]		0.84 [0.62–1.13]	
	2/2	8.0	12.5	0.54 [0.34–0.86]		0.56 [0.34–0.93]	
Dominant	1/1	46.2	39.0	1.00	<b>0.027</b>	1.00	0.093
	1/2–2/2	53.8	61.0	0.74 [0.57–0.97]		0.78 [0.59–1.04]	
Recessive	1/1–1/2	92.0	87.5	1.00	<b>0.026</b>	1.00	<b>0.045</b>
	2/2	8.0	12.5	0.61 [0.39–0.95]		0.61 [0.38–1.00]	
log-Additive	0, 1, 2			0.76 [0.62–0.93]	<b>0.007</b>	0.78 [0.63–0.97]	<b>0.027</b>
rs11712619							
Codominant	1/1	38.4	47.0	1.00	<b>0.010</b>	1.00	<b>0.050</b>
	1/2	49.8	44.5	0.73 [0.57–0.94]		0.78 [0.59–1.03]	
	2/2	11.8	8.4	0.58 [0.38–0.89]		0.60 [0.37–0.95]	
Dominant	1/1	38.4	47.0	1.00	<b>0.004</b>	1.00	<b>0.030</b>
	1/2–2/2	61.6	53.0	0.70 [0.55–0.9]		0.74 [0.57–0.97]	
Recessive	1/1–1/2	88.2	91.6	1.00	0.068	1.00	0.087
	2/2	11.8	8.4	0.69 [0.46–1.03]		0.68 [0.43–1.06]	
log-Additive	0, 1, 2			0.75 [0.62–0.91]	<b>0.003</b>	0.77 [0.63–0.95]	<b>0.014</b>

**Table 3** continued

SNP Model		Controls (%)	Cases (%)	Unadjusted		Adjusted	
				OR [95% CI]	<i>P</i>	OR [95% CI]	<i>P</i>
rs7620580							
Codominant	1/1	70.4	78.6	1.00	<b>0.019</b>	1.00	<b>0.020</b>
	1/2	26.7	19.5	1.52 [1.11–2.09]		1.58 [1.13–2.23]	
	2/2	2.9	1.9	1.72 [0.72–4.14]		1.72 [0.68–4.35]	
Dominant	1/1	70.4	78.6	1.00	<b>0.005</b>	1.00	<b>0.005</b>
	1/2–2/2	29.6	21.4	1.54 [1.14–2.09]		1.60 [1.15–2.22]	
Recessive	1/1–1/2	97.1	98.1	1.00	0.317	1.00	0.355
	2/2	2.9	1.9	1.56 [0.65–3.73]		1.54 [0.61–3.87]	
log-Additive	0, 1, 2			1.45 [1.11–1.90]	<b>0.006</b>	1.49 [1.12–1.98]	<b>0.007</b>

Unadjusted (without covariates) and adjusted (for hypertension, diabetes, and ever smoking) association testings were performed using different models (codominant, dominant, recessive, and log-additive). Significant *P* values are shown in bold

OR odds ratio, CI confidence interval

virtually all tissues, while nNOS and eNOS are constitutively expressed primarily in neural tissue and endothelium, respectively, which are the principal tissues involved in stroke. Several published case–control studies describe an association of genetic variants in *eNOS* with ischemic stroke risk (Domingues-Montanari et al. 2008), and we have found an association of nNOS with IS (personal communication). Kalirin-7 is involved in the regulation of ischemic signal transduction (Beresewicz et al. 2008), and may play a neuroprotective role during inflammation of the central nervous system by inhibiting iNOS activity (Ratovitski et al. 1999). It would therefore be interesting to assess the role played by genetic variation in Kalirin in stroke severity and recovery, as well as epistatic effects between *KALRN* and NO synthases in ischemic stroke risk.

Several independent Kalirin polymorphisms have now been associated in previous studies (Wang et al. 2007; Rudock et al. 2008; Ikram et al. 2009) and in the present report with vascular phenotypes, and therefore it is not clear which is(are) the exact susceptibility variant(s). Kalirin appears to be a pleiotropic protein, and it is possible that different variants in the same gene cause related phenotypes. Alternatively, the causal polymorphism might not have been directly studied yet, as suggested by our imputation approach. Given that we had genotyped a relatively high number of haplotype tagging SNPs in the *KALRN* gene region, we were able to impute hundreds of neighboring polymorphisms with high confidence and identified many SNPs with stronger association than those directly genotyped. Genomewide imputation is currently performed routinely in whole-genome association studies to increase power, but “local” imputation may also be a powerful approach to detect novel associations when there is a dense map of observed genotypes. We validated this local imputation approach in our dataset through direct genotyping.

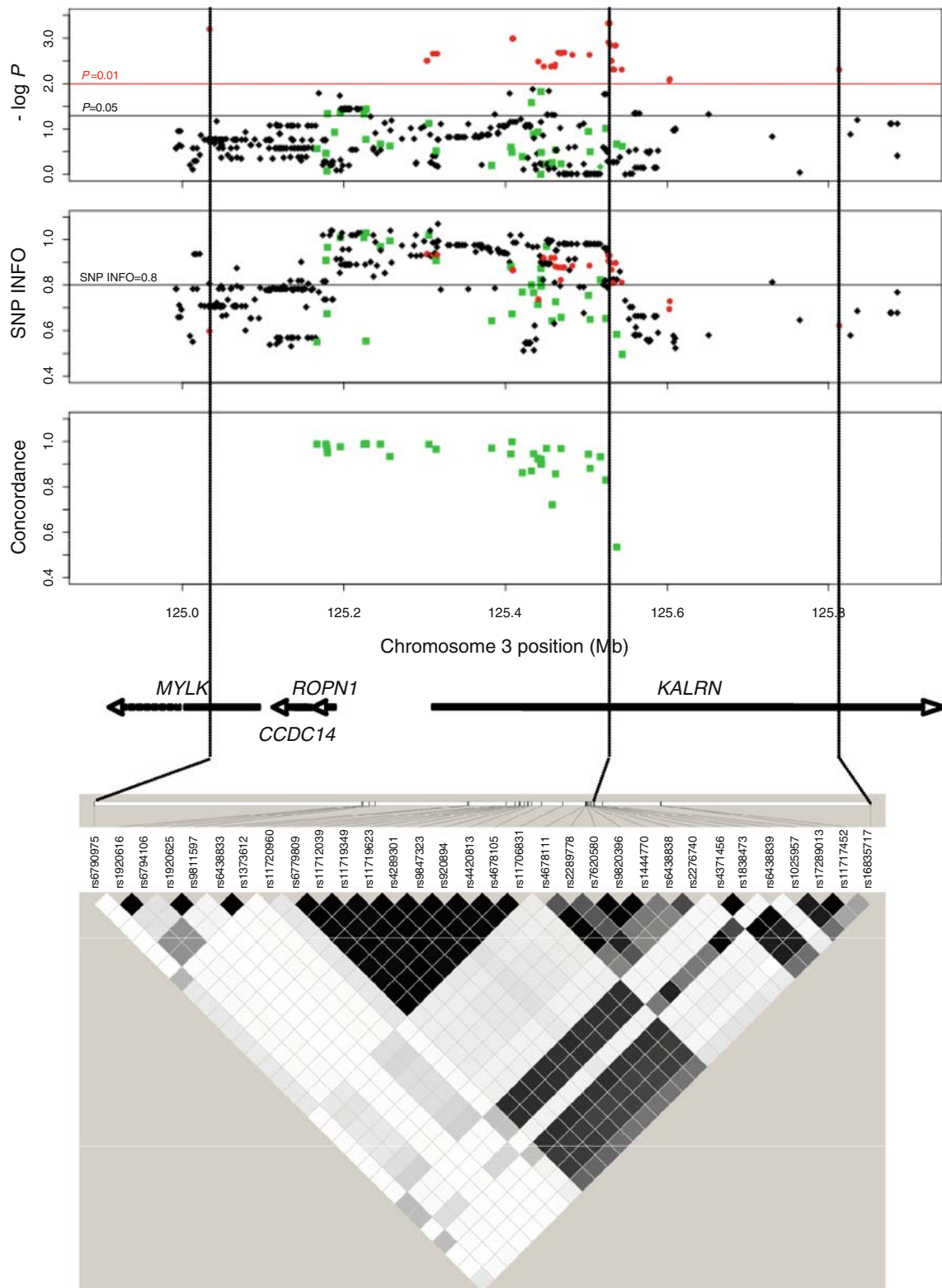
Another possibility is that Kalirin demonstrates an even stronger association with a pathogenic mechanism (e.g. inflammation) or subtype (e.g. small vessel) of vascular disease that has not yet been tested directly or corrected for in the statistical analyses. This hypothesis is supported by the significant increase in the evidence for linkage at 3q13 (LOD = 5.10, *P* = 0.008) in GENECARD families with lower-risk lipid profiles and fewer known risk factors (Shah et al. 2006). We did not test the association of Kalirin with IS subtypes since a much larger sample size would be required to have a representative number of individuals in each subtype and genotype category.

Population stratification was a frequent confounding factor in previous association studies. It can be corrected for when genomewide data is available, but is harder to address in candidate gene studies when no ancestry-informative markers are known. Population genetic studies with very large numbers of autosomal markers have never been conducted in self-described Portuguese Caucasians. Additionally, there are no major historical events that lead us to suspect that a significant proportion of Portuguese Caucasian individuals are of mixed backgrounds. Therefore, we do not expect that an underlying population substructure in our sample accounts for the observed associations.

The type of correction for multiple testing that must be employed is subject to controversy, but our positive association findings would not survive the most conservative Bonferroni correction. However, replication in multiple independent datasets remains the gold-standard of association studies, even for modest associations. Our study highlights the importance of replicating associations that fall below the genomewide significance threshold. These variants may account for part of the missing heritability.

We believe that the Kalirin association is promising because of the multiple independent lines of evidence for





**Fig. 2** Association results of imputed SNPs. The *top plot* displays the negative logarithm of the  $P$  for the allelic association test of all 405 imputed SNPs in chromosome 3. The *second plot* shows the SNP information content metric SNP INFO, and the *third plot* displays the concordance rate of observed and imputed genotypes (for genotyped SNPs only). In all plots, the 32 imputed SNPs with  $P < 0.01$  are repre-

sented with *red spheres*, and the 34 SNPs that have been genotyped are represented with *green squares*. Known genes in this region are indicated (*MYLK* myosin light chain kinase isoform; *CCDC14* coiled-coil domain containing 14). The LD plot at the *bottom* shows all the pairwise correlation coefficients ( $r^2$ ) among the 32 imputed SNPs with  $P < 0.01$

association, namely the association of several uncorrelated SNPs, the replication of the most associated SNP results on the Ikram et al. (2009), GWAS published study, and the previous reports of association with cardiometabolic syndrome (Wang et al. 2007; Rudock et al. 2008). Additional replication studies in other independent datasets with stroke and vascular pathologies are needed to further confirm Kalirin's role in vascular disease susceptibility. Deep sequencing may have to be used to precisely identify the causal genetic variant.

**Acknowledgments** We are deeply grateful to all study participants and to the genotyping unit at the Instituto Gulbenkian de Ciência. This work was supported in part by the Marie Curie International Reintegration Grant 513760 (SAO), the Marie Curie Intra-European Fellowship 024563 (SAO), the grant PTDC/SAU-GMG/64426/2006 from the Portuguese Fundação para a Ciência e a Tecnologia (FCT), and by fellowships from FCT (TK, HM), from the Portuguese Instituto do Emprego e Formação Profissional (TK, JX) and from Fundação AstraZeneca/Faculdade de Medicina de Lisboa (LG).

## References

- Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21:263–265
- Beresevicz M, Kowalczyk JE, Zablocka B (2008) Kalirin-7, a protein enriched in postsynaptic density, is involved in ischemic signal transduction. *Neurochem Res* 33:1789–1794
- Bowden DW, Rudock M, Ziegler J, Lehtinen AB, Xu J, Wagenknecht LE, Herrington D, Rich SS, Freedman BI, Carr JJ, Langefeld CD (2006) Coincident linkage of type 2 diabetes, metabolic syndrome, and measures of cardiovascular disease in a genome scan of the diabetes heart study. *Diabetes* 55:1985–1994
- Chiodini BD, Lewis CM (2003) Meta-analysis of 4 coronary heart disease genome-wide linkage studies confirms a susceptibility locus on chromosome 3q. *Arterioscler Thromb Vasc Biol* 23:1863–1868
- Domingues-Montanari S, Mendioroz M, del Rio-Espinola A, Fernández-Cadenas I, Montaner J (2008) Genetics of stroke: a review of recent advances. *Expert Rev Mol Diagn* 8:495–513
- Ferraro F, Ma XM, Sobota JA, Eipper BA, Mains RE (2007) Kalirin/Trio Rho guanine nucleotide exchange factors regulate a novel step in secretory granule maturation. *Mol Biol Cell* 18:4813–4825
- González JR, Armengol L, Solé X, Guinó E, Mercader JM, Estivill X, Moreno V (2007) SNPAssoc: an R package to perform whole genome association studies. *Bioinformatics* 23:644–645
- Hauser ER, Crossman DC, Granger CB, Haines JL, Jones CJ, Mooser V, McAdam B, Winkelmann BR, Wiseman AH, Muhlestein JB, Bartel AG, Dennis CA, Dowdy E, Estabrooks S, Eggleston K, Francis S, Roche K, Cleverger PW, Huang L, Pedersen B, Shah S, Schmidt S, Haynes C, West S, Asper D, Booze M, Sharma S, Sundseth S, Middleton L, Roses AD, Hauser MA, Vance JM, Pericak-Vance MA, Kraus WE (2004) A genomewide scan for early-onset coronary artery disease in 438 families: the GENE-CARD study. *Am J Hum Genet* 75:436–447
- Helgadóttir A, Thorleifsson G, Manolescu A, Gretarsdóttir S, Blondal T, Jonasdóttir A, Jonasdóttir A, Sigurdsson A, Baker A, Palsson A, Masson G, Gudbjartsson DF, Magnusson KP, Andersen K, Levey AI, Backman VM, Matthiassdóttir S, Jonsdóttir T, Palsson S, Einarsdóttir H, Gunnarsdóttir S, Gylfason A, Vaccarino V, Hooper WC, Reilly MP, Granger CB, Austin H, Rader DJ, Shah SH, Quyyumi AA, Gulcher JR, Thorgeirsson G, Thorsteinsdóttir U, Kong A, Stefansson K (2007) A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science* 316:1491–1493
- Helgadóttir A, Thorleifsson G, Magnusson KP, Grétarsdóttir S, Steinthorsdóttir V, Manolescu A, Jones GT, Rinkel GJ, Blankensteijn JD, Ronkainen A, Jääskeläinen JE, Kyo Y, Lenk GM, Sakalihan N, Kostulas K, Gottsäter A, Flex A, Stefansson H, Hansen T, Andersen G, Weinsheimer S, Borch-Johnsen K, Jorgensen T, Shah SH, Quyyumi AA, Granger CB, Reilly MP, Austin H, Levey AI, Vaccarino V, Palsdóttir E, Walters GB, Jonsdóttir T, Snorraddóttir S, Magnúsdóttir D, Gudmundsson G, Ferrell RE, Sveinbjornsdóttir S, Hernesniemi J, Niemelä M, Limet R, Andersen K, Sigurdsson G, Benediktsson R, Verhoeven EL, Teijink JA, Grobbee DE, Rader DJ, Collier DA, Pedersen O, Pola R, Hillert J, Lindblad B, Valdimarsson EM, Magnadóttir HB, Wijmenga C, Tromp G, Baas AF, Ruigrok YM, van Rijn AM, Kuivaniemi H, Powell JT, Matthiasson SE, Gulcher JR, Thorgeirsson G, Kong A, Thorsteinsdóttir U, Stefansson K (2008) The same sequence variant on 9p21 associates with myocardial infarction, abdominal aortic aneurysm and intracranial aneurysm. *Nat Genet* 40:217–224
- Ikram MA, Seshadri S, Bis JC, Fornage M, DeStefano AL, Aulchenko YS, Debette S, Lumley T, Folsom AR, van den Herik EG, Bos MJ, Beiser A, Cushman M, Launer LJ, Shahar E, Struchalin M, Du Y, Glazer NL, Rosamond WD, Rivadeneira F, Kelly-Hayes M, Lopez OL, Coresh J, Hofman A, DeCarli C, Heckbert SR, Koudstaal PJ, Yang Q, Smith NL, Kase CS, Rice K, Haritunians T, Roks G, de Kort PLM, Taylor KD, de Lau LM, Oostra BA, Uitterlinden AG, Rotter JJ, Boerwinkle E, Psaty BM, Mosley TH, van Duijn CM, Breteler MMB, Longstreth WT, Wolf PA (2009) Genomewide association studies of stroke. *N Engl J Med* 360:1718–1728
- Johnson RC, Penzes P, Eipper BA, Mains RE (2000) Isoforms of kalirin, a neuronal Dbl family member, generated through use of different 5'- and 3'-ends along with an internal translational initiation site. *J Biol Chem* 275:19324–19333
- Karvanen J, Silander K, Kee F, Tiret L, Salomaa V, Kuulasmaa K, Wilklund PG, Virtamo J, Saarela O, Perret C, Perola M, Peltonen L, Cambien F, Erdmann J, Samani NJ, Schunkert H, Evans A, Project MORGAM (2009) The impact of newly identified loci on coronary heart disease, stroke and total mortality in the MORGAM prospective cohorts. *Genet Epidemiol* 33:237–246
- Larson MG, Atwood LD, Benjamin EJ, Cupples LA, D'Agostino RB Sr, Fox CS, Govindaraju DR, Guo CY, Heard-Costa NL, Hwang SJ, Murabito JM, Newton-Cheh C, O'Donnell CJ, Seshadri S, Vasan RS, Wang TJ, Wolf PA, Levy D (2007) Framingham Heart Study 100 K project: genome-wide associations for cardiovascular disease outcomes. *BMC Med Genet* 8:S5
- Ma XM, Johnson RC, Mains RE, Eipper BA (2001) Expression of kalirin, a neuronal GDP/GTP exchange factor of the trio family, in the central nervous system of the adult rat. *J Comp Neurol* 429:388–402
- Matarin M, Brown WM, Singleton A, Hardy JA, Meschia JF, for the ISGS investigators (2008) Whole genome analyses suggest ischemic stroke and heart disease share an association with polymorphisms on chromosome 9p21. *Stroke* 39:1586–1589
- McPherson CE, Eipper BA, Mains RE (2002) Genomic organization and differential expression of Kalirin isoforms. *Gene* 284:41–51
- McPherson CE, Eipper BA, Mains RE (2004) Kalirin expression is regulated by multiple promoters. *J Mol Neurosci* 22:51–62
- McPherson R, Pertsemliadis A, Kavaslar N, Stewart A, Roberts R, Cox DR, Hinds DA, Pennacchio LA, Tybjaerg-Hansen A, Folsom AR, Boerwinkle E, Hobbs HH, Cohen JC (2007) A common allele on chromosome 9 associated with coronary heart disease. *Science* 316:1488–1491

- Murray CJ, Lopez AD (1997) Alternative projections of mortality and disability by cause 1990–2020: Global Burden of Disease Study. *Lancet* 349:1498–1504
- O'Donnell CJ, Cupples LA, D'Agostino RB, Fox CS, Hoffmann U, Hwang SJ, Ingellson E, Liu C, Murabito JM, Polak JF, Wolf PA, Demissie S (2007) Genome-wide association study for subclinical atherosclerosis in major arterial territories in the NHLBI's Framingham Heart Study. *BMC Med Genet* 8:S4
- Penzes P, Jones KA (2008) Dendritic spine dynamics—a key role for kalirin-7. *Trends Neurosci* 31:419–427
- Penzes P, Johnson RC, Alam MR, Kambampati V, Mains RE, Eipper BA (2000) An isoform of kalirin, a brain-specific GDP/GTP exchange factor, is enriched in the postsynaptic density fraction. *J Biol Chem* 275:6395–6403
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC (2007) PLINK: a toolset for whole-genome association and population-based linkage analysis. *Am J Hum Genet* 81:559–575
- Rabiner CA, Mains RE, Eipper BA (2005) Kalirin: a dual Rho guanine nucleotide exchange factor that is so much more than the sum of its many parts. *Neuroscientist* 11:148–160
- Ratovitski EA, Alam MR, Quick RA, McMillan A, Bao C, Kozlovsky C, Hand TA, Johnson RC, Mains RE, Eipper BA, Lowenstein CJ (1999) Kalirin inhibition of inducible nitric-oxide synthase. *J Biol Chem* 274:993–999
- Rudock ME, Ziegler JT, Lehtinen AB, Freedman BI, Carr JJ, Langefeld CD, Bowden DW (2008) Analysis of Kalirin polymorphisms with cardiovascular risk, type 2 diabetes, metabolic syndrome in the Diabetes Heart Study. In: Meeting of the American Society of Human Genetics. Abstract book, abstract 2018, p 388
- Shah SH, Kraus WE, Crossman DC, Granger CB, Haines JL, Jones CJ, Mooser V, Huang L, Haynes C, Dowdy E, Vega GL, Grundy SM, Vance JM, Hauser ER (2006) Serum lipids in the GENECARD study of coronary artery disease identify quantitative trait loci and phenotypic subsets on chromosomes 3q and 5q. *Ann Hum Genet* 70:738–748
- Wang L, Hauser ER, Shah SH, Pericak-Vance MA, Haynes C, Crosslin D, Harris M, Nelson S, Hale AB, Granger CB, Haines JL, Jones CJ, Crossman D, Seo D, Gregory SG, Kraus WE, Goldschmidt-Clermont PJ, Vance JM (2007) Peakwide mapping on chromosome 3q13 identifies the kalirin gene as a novel candidate gene for coronary artery disease. *Am J Hum Genet* 80:650–663
- Ye S, Willeit J, Kronenberg F, Xu Q, Kiechl S (2008) Association of genetic variation on chromosome 9p21 with susceptibility and progression of atherosclerosis: a population-based, prospective study. *J Am Coll Cardiol* 52:378–384
- Youn H, Jeoung M, Koo Y, Ji H, Markesbery WR, Ji I, Ji TH (2007a) Kalirin is under-expressed in Alzheimer's disease hippocampus. *J Alzheimers Dis* 11:385–397
- Youn H, Ji I, Ji HP, Markesbery WR, Ji TH (2007b) Under-expression of Kalirin-7 Increases iNOS activity in cultured cells and correlates to elevated iNOS activity in Alzheimer's disease hippocampus. *J Alzheimers Dis* 12:271–281