

Contribution of polymorphic variation of inositol hexakisphosphate kinase 3 (*IP6K3*) gene promoter to the susceptibility to Late Onset Alzheimer's Disease.

Paolina Crocco^{2*}, Adolfo Saiardi^{1*}, Miranda S. Wilson¹, Raffaele Maletta³, Amalia C. Bruni³, Giuseppe Passarino², Giuseppina Rose^{2#}.

¹Medical Research Council Laboratory for Molecular Cell Biology, University College London, London, United Kingdom

²Department of Biology, Ecology and Earth Science, University of Calabria, Rende (CS), Italy.

³Regional Neurogenetic Centre, ASP CZ, Lamezia Terme (CZ), Italy.

*Equally contributed

#Corresponding author

Giuseppina Rose

Department of Biology, Ecology and Earth Science
Cubo 4C, University of Calabria

87036 Rende-Italy

Phone: (+ 39) 0984-492931

Fax : (+39) 09844-93601

Email : pina.rose@unical.it

Abstract

Maintenance of electric potential and synaptic transmission are energetically demanding tasks that neuronal metabolism must continually satisfy. Inability to fulfil these energy requirements leads to the development of neurodegenerative disorders, including Alzheimer's disease. A prominent feature of Alzheimer's disease is in fact neuronal glucose hypometabolism. Thus understanding the fine control of energetic metabolism might help to understand neurodegenerative disorders. Recent research has indicated that a novel class of signalling molecules, the inositol pyrophosphates, act as energy sensors. They are able to alter the balance between mitochondrial oxidative phosphorylation and glycolytic flux, ultimately affecting the cellular level of ATP. The neuronal inositol pyrophosphate synthesis relies on the activity of the neuron enriched inositol hexakisphosphate kinase 3 (IP6K3) enzyme. To verify an involvement of inositol pyrophosphate signalling in neurodegenerative disorders, we performed tagging single nucleotide polymorphism (SNP) analysis of the IP6K3 gene in patients with familial and sporadic late onset Alzheimer's disease (LOAD). Two SNPs in the 5'-flanking promoter region of the *IP6K3* gene were found to be associated with sporadic LOAD. Characterising the functionality of the two polymorphisms by luciferase assay revealed that one of them (rs28607030) affects *IP6K3* promoter activity, with the G allele showing an increased activity. As the same allele has a beneficial effect on disease risk, this may be related to upregulation of *IP6K3* expression, with a consequent increase in inositol pyrophosphate synthesis. In conclusion, we provide the first evidence for a contribution of genetic variability in the *IP6K3* gene to LOAD pathogenesis.

Keywords: IP6K3, inositol pyrophosphate, IP₇, Metabolism, SNP, Alzheimer's disease

1. Introduction

The binding of a specific hormone/ligand (a ‘first messenger’) to plasma membrane receptors results in the activation of intracellular signalling cascades, mediated by small molecules that are therefore called ‘second messengers’. The protein kinase A (PKA) activator cAMP and the calcium releasing factor Ins(1,4,5)P₃ (hereafter IP₃) represent the most classical examples of second messengers. While only two cyclic nucleotides, cAMP and cGMP, are recognized as cellular messengers, IP₃ represents the forerunner of a large and growing family of signalling molecules: the inositol phosphates [1]. The lipid form, called inositides, comprises seven members that control membrane identity as well as signalling events (for review see [2]). The cytosolic, or water soluble, inositol phosphates encompass more than 40 identified members, able to regulate virtually every aspect of cellular physiology (for review see [3, 4]).

Though IP₃ is historically and functionally the most important inositol phosphate, recently a subfamily of these signalling molecules, the inositol pyrophosphates, generated by sequential phosphorylation of IP₃ (Figure 1A) are gaining in interest and functional importance (for review see [5-7]). The inositol pyrophosphates are inositol phosphate molecules containing one or more pyrophosphate (diphospho) moiety/ies. The best characterised member of this family is synthesised from the fully phosphorylated inositol phosphate IP₆ (inositol hexakisphosphate, phytic acid or InsP₆) and is called IP₇ (diphosphoinositol pentakisphosphate or InsP₇); it possesses a pyrophosphate moiety at position five (Figure 1B) [8]. The inositol hexakisphosphate kinase enzymes, IP6Ks, are responsible for IP₇ synthesis [9] (Figure 1 A,B). Three IP6K genes are present in the human genome; *IP6K1* (*IHPK1*) and *IP6K2* (*IHPK2*, *PiUS*) are localised close to each other on the same cytogenetic band (3p21.31) on chromosome three, while *IP6K3* (*IHPK3*) is localised on chromosome six (6p21.31). The PPIP5K enzymes can further phosphorylate IP₇ to IP₈, a species containing two pyrophosphate moieties (Figure 1A).

The analysis of genetically engineered yeast, amoeba and mice has associated the absence of inositol pyrophosphates with a vast array of cell biological processes. Inositol pyrophosphates regulate telomere length [10, 11], vesicular trafficking [12], epigenetic mechanisms [13, 14], DNA recombination [15, 16], ROS signalling [17] and many other activities. The ability of inositol pyrophosphates to regulate the cellular level of ATP [18], the central molecule of intermediary metabolism, might explain the diversity of roles played by this class of molecules. These fundamental signalling roles support the notion that alteration of inositol pyrophosphate metabolism is manifested by the appearance of important human diseases such as cancer and diabetes. The generation of mouse knockout models has revealed

that *ip6k1*^{-/-} mice have altered insulin signalling and are resistant to becoming obese [19], while *ip6k2*^{-/-} mice are more susceptible to cancer [20]. Interestingly, the *ip6k3*^{-/-} mouse model displays neurological defects of motor learning and coordination [21].

New information on inositol pyrophosphates is coming from human genetic studies where polymorphisms of IP6Ks and/or their chromosomal locus have been associated with specific traits or diseases. The *IP6K1* gene is disrupted at the 3p21.31 breakpoint in a family with type 2 diabetes mellitus [22]; interestingly, a Single Nucleotide Polymorphism (SNP) at the same genetic locus has associated *IP6K1* with the pathogenesis of Crohn's disease [23]. The *IP6K2* genetic locus seems associated with variants determining human height [24]. Remarkably, a SNP at the *IP6K3* locus has also been associated with Crohn's disease in Koreans [25]. Furthermore, an *IP6K3* SNP has been associated with serum phosphorus concentration [26]. This large scale genome-wide association study of serum phosphorus concentration revealed just seven gene polymorphisms associated with serum phosphorus concentration [26], thus the association of *IP6K3* with blood phosphate level is of particular relevance. This genetic study is consistent with the demonstration that inositol pyrophosphates regulate phosphate metabolism in a lower eukaryote, the yeast *Saccharomyces cerevisiae* [27]. While IP6K3 is biochemically the least characterised of the three mammalian IP6Ks, the characterized SNP represents the strongest genetic association published so far for any IP6Ks. Therefore, we decided to investigate the genetic contribution of this locus to common human complex diseases, by analysing the influence of *IP6K3* gene variation on the susceptibility to late onset Alzheimer's disease (LOAD), the most common form of neurodegenerative disorder in the elderly. It is characterized pathologically by neuronal loss and aggregation of two proteins, A β and tau, and clinically by a progressive loss of memory and impairment of cognitive ability.

Our brain has a very high ATP requirement; while this organ represents about 2% of body mass it uses 20% of oxygen consumed by a resting human [28]. Neurons have high energy requirements to maintain synaptic transmission and electric potential [29]. Several converging lines of evidence have implicated alterations in global and regional brain energetics in the pathogenesis of neurodegenerative diseases [30]. In fact, the reduction of the cerebral metabolic rate for glucose in specific brain areas is one of the striking features of Alzheimer's disease (AD) [31]. Moreover, a strong correlation between the spatial distribution of increased glycolysis, and A β plaques has been identified in the AD brain [32]. Therefore, since *IP6K3* is expressed in neurons [21, 33], and since inositol pyrophosphates have been shown to regulate ATP concentration by altering the glycolytic/mitochondrial metabolic ratio [18], we

hypothesized that *IP6K3* variants could impact on neuronal energy homeostasis and consequently on neurodegenerative LOAD processes. In order to verify this hypothesis, we performed tagging SNP analysis of the *IP6K3* gene in patients with familial and sporadic LOAD, and age-matched healthy control subjects. Two SNPs in the 5'-flanking region were found to be associated with sporadic LOAD. *In vitro* experiments were undertaken to examine whether these SNPs have an effect on *IP6K3* promoter activity.

2. Material Methods

2.1. Study population

A total of 527 individuals from Southern Italy, including 280 patients with LOAD and 247 unrelated healthy controls were considered in this study. These patients were recruited at the Regional Neurogenetics Centre (Calabria, Southern Italy). According to familiarity (at least one first-degree relative with AD), the cases were classified as affected by familial LOAD (N = 138 subjects, including in the study only one case per family) or sporadic LOAD (N = 142 subjects).

Clinical diagnosis for AD was performed through the criteria of the National Institute on Aging, and the Alzheimer's Association workgroup [34]. All patients were fully characterised from a clinical point of view and a set of physical and biochemical parameters were measured. Cognitive status was investigated through Mini Mental State Examination (MMSE) [35]. MMSE scores were adjusted for age and educational level according to procedure reported in [36].

Control subjects were recruited in the same population as AD patients in the frame of different recruitment campaigns focused on the monitoring of the quality of aging in Calabria. All subjects were carefully assessed using a rigorous clinical history evaluation and a general/neurological examination, in order to exclude the presence of any neurological disorder. To avoid population stratification effects, only subjects with at least two generations of ancestors from the Calabria region were included in this study.

2.2. Ethics statement

The study protocol was approved by the local ethics committee and conducted in accordance with the provisions of the Helsinki Declaration; informed consent for genetic screening was obtained from the study participants or, where appropriate, a relative or legal representative. After written informed consent was signed, genomic DNA was extracted from peripheral blood, according to established protocols.

2.3. SNP Selection and Genotyping Assays

A total of 17 SNPs within approximately 30 kb encompassing the entire *IP6K3* gene and its 5' and 3' flanking regions were genotyped in all subjects included in the study. The SNPs were selected by a tagging approach using the Caucasian HapMap database (www.hapmap.org) based on pairwise r^2 (≥ 0.8) among common SNPs with minor allele frequency (MAF ≥ 0.05). The rs9469578 polymorphism, associated with serum phosphorus concentration [26], was also included in the final tag SNP panel. Multiplex SNP genotyping was performed using iPlex Gold Genotyping Assay and Sequenom MassArray (Sequenom, San Diego, CA, USA) technology, following the manufacturer's instructions and as previously described [37]. SNP assays were designed using Sequenom's MassARRAY Assay Design v3.0 Software. Spectra were analyzed using MassARRAY Typer v3.4 Software (Sequenom). For quality control, 5% of the total number of samples was re-genotyped to assess the reliability of the genotype identification protocols. Concordance among duplicates was greater than 99.8% for all genotypes. For additional quality control, genotypes were excluded if Hardy-Weinberg equilibrium among controls $p < 0.05$ or call rates $< 90\%$. Ultimately, two SNPs, rs559290 and rs6457740, were excluded from analysis because they deviated from HWE ($P < 0.05$), and one, rs12203688, due to low genotyping success rates.

2.4. Bioinformatic analyses

To assess the potential function of the promoter phenotype-associated SNPs, *in silico* functional analysis was carried out using the freely available PROMO software (http://alggen.lsi.upc.es/cgi-bin/promo_v3/promo/promoinit.cgi?dirDB=TF_8.3) [38, 39]. The functionality of these SNPs (and those in strong linkage disequilibrium; $r^2 \geq 0.9$) was further explored using regulatory information from the ENCODE (<https://genome.ucsc.edu/ENCODE/>) [40] and the Roadmap Epigenome Mapping projects (<http://www.ppmroadmap.com/>; Roadmap Epigenomics et al., 2015) as implemented in HaploReg (v4.1, www.broadinstitute.org/mammals/haploreg/) [41], and RegulomeDB (www.regulomedb.org/) [42].

2.5. Reporter constructs

Two *IP6K3* promoter regions of 777 nt containing the rs10947435 (G3392A (Pos. 33718075) and rs28607030 (A2033G (Pos. 33716716)) polymorphisms were chemically

synthesized (MWG-Biotech AG) adding NheI and HindIII restriction sites at the 5' and 3' of the synthesized DNA. The SNP was localized 276 nt downstream of NheI and 500 nt upstream of HindIII site. The four plasmid synthesized were named pK31 ((A)rs28607030), pK32 ((G)rs28607030), pK33 ((A)rs10947435) and pK34 ((G)rs10947435). These constructs were digested by NheI and HindIII, and subcloned into the promoterless pGL3 Basic luciferase reporter vector (Promega) generating the following plasmids pGL3K31, pGL3K32, pGL3K33 and pGL3K34. The final constructs and presence of the appropriate SNP was confirmed by DNA sequencing.

2.6. Cell culture, transfection and luciferase assay

The cell lines HEK293T and SH-SY5Y were cultured in Dulbecco's modified Eagle's medium (DMEM, Gibco) supplemented with 10% (v/v) fetal bovine serum, and maintained in a humidified atmosphere of 5% CO₂ at 37°C. Transfection of HEK293T cells was performed using Lipofectamine 2000 (Invitrogen) on 80% confluent cells in a 24 well plate, while SH-SY5Y cells were transfected using TransFast (Promega) in a 12 well plate following the manufacturer's protocol. The pGL3-IP6K3-SNP construct (pGL3K31 to 4), 100-500 ng, was cotransfected with 1-10 ng pHRG-TK vector containing the Renilla luciferase (Promega) to correct for transfection efficiency. Twenty-four hours after transfection, cells were harvested or treated with Forskolin or the c-Myc inhibitor 10058-F4 (Sigma- Aldrich). Luciferase expression activities were assessed using the Stop & Glo kit (Promega) following manufacturer's instructions. Firefly luciferase activity was normalized to Renilla luciferase activity. Differences in luciferase activity were compared using unpaired t-test.

2.7. Statistical Analysis

For each SNP, allele and genotype frequencies were estimated by gene counting from the observed genotypes. Departure from Hardy-Weinberg equilibrium was assessed for each SNP in controls using the χ^2 test. Pairwise LD was estimated between SNPs based on r^2 statistics calculated in controls using Haploview software version 4.0 [43] (available from the Broad Institute at <http://www.broad.mit.edu/mpg/haploview/>).

Logistic regression analysis was used to examine the differences in SNP genotype frequencies between cases and controls. The association analyses were based on the estimation of the odds ratios (ORs) and their 95% confidence intervals (CIs), using sex as covariate. Additive, dominant and recessive genetic models were evaluated for each SNP. Only the dominant model was considered where the minor allele homozygote count for either cases or

controls was <5%. Akaike's Information Criterion (AIC) [44] was employed to determine the best-fitting model for each SNP.

We assessed whether there was evidence of interaction between *APOE-ε4* and SNPs that were found significant in the case–control analyses by including both the main effects for each variant and a product term (SNP x *APOE-ε4*) in the regression models.

Statistical significance was considered to be $p < 0.05$. As this study was exploratory, an arguably overly conservative statistical significance thresholding procedure (e.g., Bonferroni correction) was not employed. The results were analyzed using SPSS 22.0 (SPSS Inc., Chicago, IL, USA).

3. Results

3.1. Genetic Analysis

To investigate the contribution of *IP6K3* genetic variability to the risk of late-onset Alzheimer disease (LOAD), we analyzed a cohort of 280 unrelated individuals with sporadic or familial LOAD, and a similarly aged cohort of 247 cognitively normal controls. Characteristics of the samples are given in Table 1. Seventeen SNPs were analysed in these populations, 14 of which passed the quality check. Figure 2 shows the physical location of these SNPs, and the linkage disequilibrium (LD) pattern in the control sample. Table 2 reports the results of the logistic regression analysis and presents the best model for association with LOAD risk. The ORs, adjusted for the gender, showed that two polymorphisms, rs28607030 and rs10947435, both located in the 5'-flanking region of the gene, were associated with LOAD risk in the subgroup of patients with the sporadic form of the disease. For both SNPs the dominant model was the best-fitting model, but they were shown to have opposite effects on the disease risk. In fact, the carriers of at least one copy of the minor G allele of rs28607030 were at a decreased risk of LOAD, with a statistically significant OR of 0.57 (OR=0.57; 95% CI 0.36–0.90; $p = 0.011$), whereas individuals carrying at least one copy of the minor A allele of rs10947435 were at higher LOAD risk, as compared to the group of carriers (OR 1.89; 95% CI 1.15–3.12; $p = 0.010$).

We then tested for possible interactions between the *IP6K3* SNPs and *APOE* status ($\epsilon 4$ allele) but no significant interaction was observed, indicating that *IP6K3* has an independent role in LOAD, irrespective of the presence of the *APOE**4 allele. Also there was no significant association for age at onset, nor for the disease severity (measured by MMSE score).

3.2 *In silico* functional consequences of rs28607030 and rs10947435

The identification of the two polymorphisms associated with LOAD risk in the 5'-flanking region of the *IP6K3* gene suggests that they might affect *IP6K3* promoter functionality. To verify this hypothesis we first subjected the *IP6K3* promoter region to an *in silico* search for putative transcription factor binding sites. We employed the freely available PROMO software and screened the promoter using a higher stringency with a maximum matrix dissimilarity rate of 5 [38, 39]. This analysis revealed that the polymorphism rs28607030 is located at the centre of a perfectly conserved, for the G allele, c-Myc binding site CACGTG, the underline indicating the SNP polymorphism [45]. Conversely, the polymorphism rs10347435 is located at the centre of a semi-conserved CREB binding site (cAMP response element, CRE) [46] with the underlined G allele TcACGTCA in the centre and the small letter indicating a mismatch compared to the consensus palindromic CRE sequence.

Furthermore, since the observed associations might be due to different causal variants in LD with the studied ones, we annotated the two variants rs28607030 and rs10947435 and their proxies in high LD ($r^2 \geq 0.9$) for evidence of functionality using HaploReg v4.1 and RegulomeDB databases. LD patterns showed that the rs28607030 is not in LD with any other variants, while eleven SNPs, all in *IP6K3*, were identified that were proxies with the rs10947435. We found promising functional implications for these SNPs (i.e. promoter and enhancer histone marks, the presence of multiple DNase I hypersensitivity sites, and transcription factor binding sites; see Supplementary Material, Table S1 for details), indicating a regulatory potential.

3.3. *IP6K3* Promoter Activities

To verify the existence of any functional effects of the SNPs located in putative transcription factor binding sites in the promoter region of the *IP6K3* gene, four luciferase reporter constructs containing the SNP rs10947435 and rs28607030 were generated (Figure 3A). Each construct was transfected into HEK293T cells and the relative promoter activities were determined by measuring luciferase expression (Figure 3B). This analysis revealed that while the SNP rs10947435 did not affect *IP6K3* promoter fragment efficiency, the two alleles with SNP rs28607030 did show different *IP6K3* promoter efficiencies. The construct pGL3K32 rs28607030, containing the G allele with a perfect match for c-Myc binding, was three times stronger than the pGL3K31 construct (the A allele). The difference in normalised luciferase activities between constructs possessing the SNP (A) or SNP (G) was statistically significant (Figure 3B). Similar studies were performed using human neuroblastoma SH-SY5Y cells, obtaining a substantially overlapping result (Figure 3C). Using SH-SY5Y we also did not

observe any effect of the SNP rs10947435 while the SNP rs28607030, containing the G allele, was almost twice as strong as the construct containing the A allele.

The SNP rs10947435 is located in a semi-conserved cAMP response element (CRE), thus CREB binding may only occur after a strong stimulus. Therefore, we treated transfected HEK293T cells with forskolin, which increases the levels of cAMP by activating adenylyl cyclase. However, treating the cells with a wide range of forskolin concentration (1 to 50 μ M) and for 1 to 24 hours failed to reveal any effect of SNP rs10947435 on promoter activity (not shown).

Since the polymorphism rs28607030 is located at the centre of a perfectly conserved c-Myc binding site, to confirm its genuine nature we treated transfected HEK293T cells with the c-Myc inhibitor 10058-F4. This inhibitor, by destroying c-Myc-Max interaction, precludes transactivation of c-Myc target gene expression [47]. Treatment with 20 μ M of 10058-F4 for 4 hours completely inhibited the transcriptional efficiency of SNP rs28607030 (Figure 3D), confirming that the A allele of this polymorphism abolishes a functional c-Myc binding site.

4. Discussion

Inositol is essential for correct neuronal development. In mouse, inositol prevents folate-resistant neural tube defects [48]. Inositol phosphate signalling is likely responsible for correct neuronal development, since neural tube defects are observable in the mouse knockout of the inositol 1,3,4-trisphosphate 5/6-kinase (ITPK1 or PPK1) [49]. Inositol phosphate signalling is not only important during development but also for fine tuning adult brain function. In fact inositol polyphosphate multikinase (IPMK) loss of activity has been associated with the pathophysiology of Huntington's disease [50]. Here, we define IP6K3 as an important player associated with LOAD. IP6K3 synthesizes the inositol pyrophosphate IP₇, the best studied of a class of highly energetic signalling molecules acting as 'metabolic messengers' [51]. A recent review has in fact placed them 'between' signalling and metabolism [7]. The possible relevance of these molecules in the context of metabolic regulation is supported by several pieces of evidence. For example, inositol pyrophosphates regulate insulin secretion from pancreatic β -cells [52] and insulin signalling [19].

Metabolism and AD are intimately linked. The brain is one of the most energy-demanding organs in the body. It has evolved complex metabolic networks to maintain activity and energy metabolism, and any alteration in these may contribute to the neurodegenerative process. A growing body of evidence indeed supports the notion that AD represents a brain

metabolic disease. The co-existence of impaired glucose metabolism and insulin signaling in the brain has suggested that AD may represent a brain-specific form of diabetes, i.e., type 3 diabetes or an insulin-resistant brain state [53, 54]. Multiple studies have also begun to outline an intricate connection between metabolic syndrome and AD. Individuals with metabolic syndrome-associated clinical features, such as diabetes, obesity, hypertension and dyslipidemia have a higher risk of developing AD [55-59].

Thus, the regulation of inositol pyrophosphate synthesis, which have a central role in cellular metabolism and are crucial for the production and use of energy, may have a key role in the correlation between metabolism and AD. Here, we provide the first evidence for a contribution of genetic variability in the *IP6K3* gene, which codes for one of the enzymes responsible for inositol pyrophosphate synthesis (Figure 1B) in neuronal cells [21, 33], to LOAD pathogenesis. We found that carriers of the minor allele of two *IP6K3* 5'-flanking region SNPs, rs28607030-G and rs10947435-A, were at decreased (OR=0.57; 95% CI 0.36–0.90; p= 0.011) and increased (OR 1.89; 95% CI 1.15–3.12; p= 0.010) risk of LOAD, respectively. The two variants were found to affect susceptibility to sporadic LOAD while they appeared to not affect the chance of familial LOAD. This finding suggests that *IP6K3* is likely to confer a low risk, which is covered by the major risk factors involved in familial LOAD. In addition, unlike other risk factors, it does not modulate the effect of major risk factor such as Apolipoprotein E (APOE) [60].

We analysed the role of the two SNPs associated with LOAD in modulating *IP6K3* promoter activity by using luciferase reporter assays. The SNP rs10947435 did not mediate gene activation. This specific polymorphism is located within a semi-conserved CRE element, which is the binding site for the transcription factor CREB. However, not all putative CREs are bound by CREB in resting conditions, as CREB does not constitutively interact with chromatin and its binding depends on stimulation in neurons [61, 62]. Therefore, it is possible that our inability to observe a transcriptional readout for SNP rs10947435 is due to the technical limitations of the luciferase reporter assays. However, we favour a different explanation: the rs10947435 SNP is likely a marker for another linked functional polymorphism which may account for the observed effect on the risk of disease. According to Hapmap data, this SNP is in strong linkage disequilibrium (LD; $r^2 \geq 0.9$) with 11 other SNPs within *IP6K3*. Moreover, *in silico* analysis by RegulomeDB revealed that three of the SNPs in LD with rs10947435 (Table S1) have the highest level of evidence for a regulatory role; therefore, we cannot exclude the possibility that one of these markers may be the actual causal variant. Further studies are required to confirm these observations.

On the contrary, analysis of SNP rs28607030 pointed to a functional role for this polymorphism, showing greater transcriptional activity of the minor rs28607030 G allele that possesses an intact c-Myc binding site, with respect to the A allele. This SNP is not in LD with other polymorphisms; therefore, it is reasonable to speculate that rs28607030 could affect the activity of the *IP6K3* promoter. Therefore the beneficial effect of the rs28607030 G allele on disease risk could be related to upregulation of expression of the gene and a consequent increase in IP₇ production.

We are aware of the fact that our study has some weaknesses that should be addressed. First, a limitation of the study is the relatively small sample size, and therefore additional large studies are warranted to validate our findings. However, it is of note that all patients and controls were recruited in Calabria, a region from south Italy characterized, for historical and geographical reasons, by a homogeneous genetic background and therefore population stratification, which may substantially influence the results of analyses, is limited. A second limitation is the lack of proper correction for multiple testing. Since this study was exploratory, we did not control for type I errors, even though we conducted a relatively large number of statistical tests; a Bonferroni correction would have eliminated potentially important findings if applied. Despite these limitations, however, we believe that our study provides a starting point for shedding light on the potential function of inositol pyrophosphates in the nervous system, and could open a new avenue to potentially improve our understanding of LOAD pathogenesis.

Disclosure

All the authors declared no competing interests.

Acknowledgements

We are grateful to patients and families for the interest and generous participation in our research effort and the Associazione per la Ricerca Neurogenetica-ONLUS Lamezia Terme for invaluable help in assisting persons and families. We would also like to thank members of the Saiardi lab for discussion. AS is supported by the Medical Research Council (MRC) core support to the MRC/UCL Laboratory for Molecular Cell Biology University Unit (MC_UU_1201814). This work was also partially supported by the European Union's Seventh Framework Programme (FP7/2007-2011) [grant number 259679], by funds from Programma

Operativo Nazionale [01_00937] - MIUR“Modelli sperimentali biotecnologici integrati per lo sviluppo e la selezione di molecole di interesse per la salute dell'uomo”, and by the Italian Health Ministry (DGRST number 4/2760-P/I.9 ab, 2007; RFPS-2006-7-334858, 2006).

References

- [1] R.F. Irvine, M.J. Schell, Back in the water: the return of the inositol phosphates, *Nat Rev Mol Cell Biol*, 2 (2001) 327-338.
- [2] T. Balla, Phosphoinositides: tiny lipids with giant impact on cell regulation, *Physiol Rev*, 93 (2013) 1019-1137.
- [3] A.C. Resnick, A. Saiardi, Inositol polyphosphate multikinase: metabolic architect of nuclear inositides, *Front Biosci*, 13 (2008) 856-866.
- [4] S.B. Shears, S.B. Ganapathi, N.A. Gokhale, T.M. Schenk, H. Wang, J.D. Weaver, A. Zaremba, Y. Zhou, Defining signal transduction by inositol phosphates, *Subcell Biochem*, 59 (2012) 389-412.
- [5] S.B. Shears, Inositol pyrophosphates: why so many phosphates?, *Adv Biol Regul*, 57 (2015) 203-216.
- [6] S.G. Thota, R. Bhandari, The emerging roles of inositol pyrophosphates in eukaryotic cell physiology, *Journal of biosciences*, 40 (2015) 593-605.
- [7] M.S. Wilson, T.M. Livermore, A. Saiardi, Inositol pyrophosphates: between signalling and metabolism, *Biochem J*, 452 (2013) 369-379.
- [8] P. Draskovic, A. Saiardi, R. Bhandari, A. Burton, G. Ilc, M. Kovacevic, S.H. Snyder, M. Podobnik, Inositol hexakisphosphate kinase products contain diphosphate and triphosphate groups, *Chem Biol*, 15 (2008) 274-286.
- [9] A. Saiardi, H. Erdjument-Bromage, A.M. Snowman, P. Tempst, S.H. Snyder, Synthesis of diphosphoinositol pentakisphosphate by a newly identified family of higher inositol polyphosphate kinases, *Curr Biol*, 9 (1999) 1323-1326.
- [10] A. Saiardi, A.C. Resnick, A.M. Snowman, B. Wendland, S.H. Snyder, Inositol pyrophosphates regulate cell death and telomere length through phosphoinositide 3-kinase-related protein kinases, *Proc Natl Acad Sci U S A*, 102 (2005) 1911-1914.
- [11] S.J. York, B.N. Armbruster, P. Greenwell, T.D. Petes, J.D. York, Inositol diphosphate signaling regulates telomere length, *J Biol Chem*, 280 (2005) 4264-4269.
- [12] A. Saiardi, C. Sciambi, J.M. McCaffery, B. Wendland, S.H. Snyder, Inositol pyrophosphates regulate endocytic trafficking, *Proc Natl Acad Sci U S A*, 99 (2002) 14206-14211.
- [13] A. Burton, C. Azevedo, C. Andreassi, A. Riccio, A. Saiardi, Inositol pyrophosphates regulate JMJD2C-dependent histone demethylation, *Proc Natl Acad Sci U S A*, 110 (2013) 18970-18975.
- [14] J. Worley, X. Luo, A.P. Capaldi, Inositol pyrophosphates regulate cell growth and the environmental stress response by activating the HDAC Rpd3L, *Cell Rep*, 3 (2013) 1476-1482.
- [15] R.S. Jadav, M.V. Chanduri, S. Sengupta, R. Bhandari, Inositol pyrophosphate synthesis by inositol hexakisphosphate kinase 1 is required for homologous recombination repair, *J Biol Chem*, 288 (2013) 3312-3321.
- [16] H.R. Luo, A. Saiardi, H. Yu, E. Nagata, K. Ye, S.H. Snyder, Inositol pyrophosphates are required for DNA hyperrecombination in protein kinase c1 mutant yeast, *Biochemistry*, 41 (2002) 2509-2515.
- [17] S.M. Onnebo, A. Saiardi, Inositol pyrophosphates modulate hydrogen peroxide signaling, *Biochem J*, (2009).
- [18] Z. Sziogyarto, A. Garedew, C. Azevedo, A. Saiardi, Influence of inositol pyrophosphates on cellular energy dynamics, *Science*, 334 (2011) 802-805.
- [19] A. Chakraborty, M.A. Koldobskiy, N.T. Bello, M. Maxwell, J.J. Potter, K.R. Juluri, D. Maag, S. Kim, A.S. Huang, M.J. Dailey, M. Saleh, A.M. Snowman, T.H. Moran, E.

- Mezey, S.H. Snyder, Inositol pyrophosphates inhibit Akt signaling, thereby regulating insulin sensitivity and weight gain, *Cell*, 143 (2010) 897-910.
- [20] B.H. Morrison, J.A. Bauer, J. Hu, R.W. Grane, A.M. Ozdemir, M. Chawla-Sarkar, B. Gong, A. Almasan, D.V. Kalvakolanu, D.J. Lindner, Inositol hexakisphosphate kinase 2 sensitizes ovarian carcinoma cells to multiple cancer therapeutics, *Oncogene*, 21 (2002) 1882-1889.
- [21] C. Fu, J. Xu, R.J. Li, J.A. Crawford, A.B. Khan, T.M. Ma, J.Y. Cha, A.M. Snowman, M.V. Pletnikov, S.H. Snyder, Inositol Hexakisphosphate Kinase-3 Regulates the Morphology and Synapse Formation of Cerebellar Purkinje Cells via Spectrin/Adducin, *J Neurosci*, 35 (2015) 11056-11067.
- [22] J. Kamimura, K. Wakui, H. Kadowaki, Y. Watanabe, K. Miyake, N. Harada, M. Sakamoto, A. Kinoshita, K. Yoshiura, T. Ohta, T. Kishino, M. Ishikawa, M. Kasuga, Y. Fukushima, N. Niikawa, N. Matsumoto, The IHPK1 gene is disrupted at the 3p21.31 breakpoint of t(3;9) in a family with type 2 diabetes mellitus, *J Hum Genet*, 49 (2004) 360-365.
- [23] A.R. Morgan, D.Y. Han, W.J. Lam, A.G. Fraser, L.R. Ferguson, Association analysis of 3p21 with Crohn's disease in a New Zealand population, *Hum Immunol*, 71 (2010) 602-609.
- [24] J. Lachance, B. Vernot, C.C. Elbers, B. Ferwerda, A. Froment, J.M. Bodo, G. Lema, W. Fu, T.B. Nyambo, T.R. Rebbeck, K. Zhang, J.M. Akey, S.A. Tishkoff, Evolutionary history and adaptation from high-coverage whole-genome sequences of diverse African hunter-gatherers, *Cell*, 150 (2012) 457-469.
- [25] S.K. Yang, M. Hong, W. Zhao, Y. Jung, J. Baek, N. Tayebi, K.M. Kim, B.D. Ye, K.J. Kim, S.H. Park, I. Lee, E.J. Lee, W.H. Kim, J.H. Cheon, Y.H. Kim, B.I. Jang, H.S. Kim, J.H. Choi, J.S. Koo, J.H. Lee, S.A. Jung, Y.J. Lee, J.Y. Jang, H.D. Shin, D. Kang, H.S. Youn, J. Liu, K. Song, Genome-wide association study of Crohn's disease in Koreans revealed three new susceptibility loci and common attributes of genetic susceptibility across ethnic populations, *Gut*, 63 (2014) 80-87.
- [26] B. Kestenbaum, N.L. Glazer, A. Kottgen, J.F. Felix, S.J. Hwang, Y. Liu, K. Lohman, S.B. Kritchevsky, D.B. Hausman, A.K. Petersen, C. Gieger, J.S. Ried, T. Meitinger, T.M. Strom, H.E. Wichmann, H. Campbell, C. Hayward, I. Rudan, I.H. de Boer, B.M. Psaty, K.M. Rice, Y.D. Chen, M. Li, D.E. Arking, E. Boerwinkle, J. Coresh, Q. Yang, D. Levy, F.J. van Rooij, A. Dehghan, F. Rivadeneira, A.G. Uitterlinden, A. Hofman, C.M. van Duijn, M.G. Shlipak, W.H. Kao, J.C. Witteman, D.S. Siscovick, C.S. Fox, Common genetic variants associate with serum phosphorus concentration, *J Am Soc Nephrol*, 21 (2010) 1223-1232.
- [27] A. Lonetti, Z. Szigyarto, D. Bosch, O. Loss, C. Azevedo, A. Saiardi, Identification of an evolutionarily conserved family of inorganic polyphosphate endopolyphosphatases, *J Biol Chem*, 286 (2011) 31966-31974.
- [28] J.W. Mink, R.J. Blumenshine, D.B. Adams, Ratio of central nervous system to body metabolism in vertebrates: its constancy and functional basis, *Am J Physiol*, 241 (1981) R203-212.
- [29] J.J. Harris, R. Jolivet, D. Attwell, Synaptic energy use and supply, *Neuron*, 75 (2012) 762-777.
- [30] D. Kapogiannis, M.P. Mattson, Disrupted energy metabolism and neuronal circuit dysfunction in cognitive impairment and Alzheimer's disease, *The Lancet Neurology*, 10 (2011) 187-198.
- [31] L. Mosconi, A. Pupi, M.J. De Leon, Brain glucose hypometabolism and oxidative stress in preclinical Alzheimer's disease, *Ann N Y Acad Sci*, 1147 (2008) 180-195.

- [32] A.G. Vlassenko, S.N. Vaishnavi, L. Couture, D. Sacco, B.J. Shannon, R.H. Mach, J.C. Morris, M.E. Raichle, M.A. Mintun, Spatial correlation between brain aerobic glycolysis and amyloid-beta (A β) deposition, *Proc Natl Acad Sci U S A*, 107 (2010) 17763-17767.
- [33] A. Saiardi, E. Nagata, H.R. Luo, A.M. Snowman, S.H. Snyder, Identification and characterization of a novel inositol hexakisphosphate kinase, *J Biol Chem*, 276 (2001) 39179-39185.
- [34] G.M. McKhann, D.S. Knopman, H. Chertkow, B.T. Hyman, C.R. Jack, Jr., C.H. Kawas, W.E. Klunk, W.J. Koroshetz, J.J. Manly, R. Mayeux, R.C. Mohs, J.C. Morris, M.N. Rossor, P. Scheltens, M.C. Carrillo, B. Thies, S. Weintraub, C.H. Phelps, The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease, *Alzheimers Dement*, 7 (2011) 263-269.
- [35] M.F. Folstein, S.E. Folstein, P.R. McHugh, "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician, *J Psychiatr Res*, 12 (1975) 189-198.
- [36] E. Magni, G. Binetti, A. Bianchetti, R. Rozzini, M. Trabucchi, Mini-Mental State Examination: a normative study in Italian elderly population, *Eur J Neurol*, 3 (1996) 198-202.
- [37] P. Crocco, A. Montesanto, G. Passarino, G. Rose, Polymorphisms Falling Within Putative miRNA Target Sites in the 3'UTR Region of SIRT2 and DRD2 Genes Are Correlated With Human Longevity, *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, (2015).
- [38] D. Farre, R. Roset, M. Huerta, J.E. Adsuara, L. Rosello, M.M. Alba, X. Messeguer, Identification of patterns in biological sequences at the ALGGEN server: PROMO and MALGEN, *Nucleic Acids Res*, 31 (2003) 3651-3653.
- [39] X. Messeguer, R. Escudero, D. Farre, O. Nunez, J. Martinez, M.M. Alba, PROMO: detection of known transcription regulatory elements using species-tailored searches, *Bioinformatics*, 18 (2002) 333-334.
- [40] E.P. Consortium, An integrated encyclopedia of DNA elements in the human genome, *Nature*, 489 (2012) 57-74.
- [41] L.D. Ward, M. Kellis, HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease, *Nucleic Acids Res*, 44 (2016) D877-881.
- [42] A.P. Boyle, E.L. Hong, M. Hariharan, Y. Cheng, M.A. Schaub, M. Kasowski, K.J. Karczewski, J. Park, B.C. Hitz, S. Weng, J.M. Cherry, M. Snyder, Annotation of functional variation in personal genomes using RegulomeDB, *Genome Res*, 22 (2012) 1790-1797.
- [43] J.C. Barrett, B. Fry, J. Maller, M.J. Daly, Haploview: analysis and visualization of LD and haplotype maps, *Bioinformatics*, 21 (2005) 263-265.
- [44] H. Akaike, A new look at the statistical model identification, *IEEE Transactions on Automatic Control*, 19 (1974) 716-723.
- [45] S. Adhikary, M. Eilers, Transcriptional regulation and transformation by Myc proteins, *Nat Rev Mol Cell Biol*, 6 (2005) 635-645.
- [46] A.J. Shaywitz, M.E. Greenberg, CREB: a stimulus-induced transcription factor activated by a diverse array of extracellular signals, *Annu Rev Biochem*, 68 (1999) 821-861.
- [47] M.J. Huang, Y.C. Cheng, C.R. Liu, S. Lin, H.E. Liu, A small-molecule c-Myc inhibitor, 10058-F4, induces cell-cycle arrest, apoptosis, and myeloid differentiation of human acute myeloid leukemia, *Exp Hematol*, 34 (2006) 1480-1489.
- [48] N.D. Greene, A.J. Copp, Inositol prevents folate-resistant neural tube defects in the mouse, *Nat Med*, 3 (1997) 60-66.

- [49] M.P. Wilson, C. Hugge, M. Bielinska, P. Nicholas, P.W. Majerus, D.B. Wilson, Neural tube defects in mice with reduced levels of inositol 1,3,4-trisphosphate 5/6-kinase, *Proc Natl Acad Sci U S A*, 106 (2009) 9831-9835.
- [50] I. Ahmed, J.I. Sbodio, M.M. Harraz, R. Tyagi, J.C. Grima, L.K. Albacarys, M.E. Hubbi, R. Xu, S. Kim, B.D. Paul, S.H. Snyder, Huntington's disease: Neural dysfunction linked to inositol polyphosphate multikinase, *Proc Natl Acad Sci U S A*, 112 (2015) 9751-9756.
- [51] S.B. Shears, Diphosphoinositol polyphosphates: metabolic messengers?, *Mol Pharmacol*, 76 (2009) 236-252.
- [52] C. Illies, J. Gromada, R. Fiume, B. Leibiger, J. Yu, K. Juhl, S.N. Yang, D.K. Barma, J.R. Falck, A. Saiardi, C.J. Barker, P.O. Berggren, Requirement of inositol pyrophosphates for full exocytotic capacity in pancreatic beta cells, *Science*, 318 (2007) 1299-1302.
- [53] S.M. de la Monte, J.R. Wands, Alzheimer's disease is type 3 diabetes-evidence reviewed, *J Diabetes Sci Technol*, 2 (2008) 1101-1113.
- [54] V. Frisardi, V. Solfrizzi, C. Capurso, B.P. Imbimbo, G. Vendemiale, D. Seripa, A. Pilotto, F. Panza, Is insulin resistant brain state a central feature of the metabolic-cognitive syndrome?, *J Alzheimers Dis*, 21 (2010) 57-63.
- [55] S. Craft, The role of metabolic disorders in Alzheimer disease and vascular dementia: two roads converged, *Arch Neurol*, 66 (2009) 300-305.
- [56] H. Cai, W.N. Cong, S. Ji, S. Rothman, S. Maudsley, B. Martin, Metabolic dysfunction in Alzheimer's disease and related neurodegenerative disorders, *Curr Alzheimer Res*, 9 (2012) 5-17.
- [57] B. Kim, E.L. Feldman, Insulin resistance as a key link for the increased risk of cognitive impairment in the metabolic syndrome, *Exp Mol Med*, 47 (2015) e149.
- [58] Z. Chen, C. Zhong, Decoding Alzheimer's disease from perturbed cerebral glucose metabolism: implications for diagnostic and therapeutic strategies, *Prog Neurobiol*, 108 (2013) 21-43.
- [59] S.C. Correia, R.X. Santos, C. Carvalho, S. Cardoso, E. Candeias, M.S. Santos, C.R. Oliveira, P.I. Moreira, Insulin signaling, glucose metabolism and mitochondria: major players in Alzheimer's disease and diabetes interrelation, *Brain Res*, 1441 (2012) 64-78.
- [60] J.T. Yu, L. Tan, J. Hardy, Apolipoprotein E in Alzheimer's disease: an update, *Annu Rev Neurosci*, 37 (2014) 79-100.
- [61] T.K. Kim, M. Hemberg, J.M. Gray, A.M. Costa, D.M. Bear, J. Wu, D.A. Harmin, M. Laptewicz, K. Barbara-Haley, S. Kuersten, E. Markenscoff-Papadimitriou, D. Kuhl, H. Bito, P.F. Worley, G. Kreiman, M.E. Greenberg, Widespread transcription at neuronal activity-regulated enhancers, *Nature*, 465 (2010) 182-187.
- [62] A. Riccio, R.S. Alvania, B.E. Lonze, N. Ramanan, T. Kim, Y. Huang, T.M. Dawson, S.H. Snyder, D.D. Ginty, A nitric oxide signaling pathway controls CREB-mediated gene expression in neurons, *Mol Cell*, 21 (2006) 283-294.

Figure Legends

Figure 1. Schematic of the inositol pyrophosphate biosynthetic pathway.

The biochemical pathway of inositol pyrophosphate synthesis is illustrated in (A). The membrane phosphoinositide PI(4,5)P₂ (PIP₂) is hydrolysed by the action of Phospholipase C (PLC) to I(1,4,5)P₃ (IP₃). This is converted to IP₄ by the IP₃-3Kinase or the inositol polyphosphate multikinase (IPMK), which also converts IP₄ to IP₅. This is subsequently metabolised by the inositol pentakisphosphate 2-kinase (IP₅-2K). The fully phosphorylated IP₆ is further acted on by inositol hexakisphosphate kinase (IP6K) to IP₇ that can be further phosphorylated to IP₈ by the PPIP5Ks. The chemical reaction catalysed by the mammalian IP6K3 enzyme, using ATP as phosphate donor and IP₆, generating IP₇ pyrophosphorylated at position 5 is depicted in (B).

Figure 2: Location and linkage disequilibrium map of 14 SNPs in *IP6K3*.

The relative physical position of each SNP is given in the upper diagram. Exons are represented by black boxes; non-coding exons by gray boxes; intronic and 5' and 3' regions are represented by solid lines. The dbSNP reference numbers are indicated below each SNP. The pairwise linkage disequilibrium coefficients r^2 for the control participants were calculated using Haploview.

Figure 3. Effect of rs28607030 and rs10947435 SNP on IP6K3 promoter activity.

Schematic, to scale, of IP6K3 promoter region (A) with the sections containing the SNPs rs28607030 and rs10947435, used for luciferase assays, shown as boxes. The first transcribed exon (Exon 1) is positioned more than 11 Kb from the first protein coding exon containing the ATG translation start site. Luciferase reporter constructs were transfected into HEK293T cells (B) and SH-SY5Y cells (C). Luciferase activity was assayed 24 h post-transfection. The rs28607030 polymorphism affected promoter strength while the rs10947435 polymorphism did not. Transfected HEK293T cells were treated 24 h post-transfection for 4 h with the c-Myc inhibitor 100058-F4 at the indicated doses (D).

The arbitrary units of firefly luciferase activity were normalized to renilla luciferase activity. Data are presented as average +/- standard deviation for experiments performed in triplicate.

$p < 0.01$ (**); $p < 0.001$ (***) ; not significant (ns). Representative results are shown; very similar results were obtained from three or two independent experiments for HEK293T and SH-SY5Y cells respectively .

Table 1. Baseline characteristics of the study participants.

	LOAD		Controls (N=247)
	Familial (N=138)	Sporadic (N=142)	
Age (mean \pm SD)	77.8 \pm 4.9	77.8 \pm 5.3	73.5 \pm 8.8
Males (%)	37.0	37.0	54.0
Age onset (mean \pm SD)	73.6 \pm 5.39	73.9 \pm 4.97	-
MMSE* (mean \pm SD)	14.35 \pm 6.53	13.86 \pm 7.11	23.84 \pm 3.78
APOE- ϵ 4 carriers (%)	45.26	37.90	8.33

MMSE scores were adjusted for educational level and age at inclusion in the present study according to procedure reported in [36].

Table 2. Logistic regression analysis for the association between *IP6K3* genotypes and the risk of LOAD

SNP	Minor allele	Frequency			Familial vs Controls			Sporadic vs controls		
		Controls (N=247)	Familial (N=138)	Sporadic (N=142)	OR	95% CI	p ^{Model}	OR	95% CI	p ^{Model}
rs4713668	T	45%	41%	48%	0.77	0.49-1.19	0.24 ^D	1.27	0.79-2.06	0.32 ^D
rs34252064	A	5%	5%	4%	1.31	0.60-2.86	0.50 ^D	1.17*	0.50-2.75	0.72 ^D
rs12661400	T	22%	17%	16%	0.83*	0.53-1.32	0.43 ^D	0.66*	0.40-1.07	0.09 ^D
rs622917	T	50%	48%	47%	0.81	0.50-1.32	0.40 ^D	0.86	0.50-1.46	0.57 ^R
rs623813	T	13%	16%	15%	2.23	0.72-6.88	0.16 ^R	1.30	0.78-2.16	0.32 ^D
rs545787	A	31%	29%	29%	0.74	0.32-1.70	0.47 ^R	0.89	0.56-1.41	0.61 ^D
rs9469578	T	9%	7%	9%	0.67*	0.35-1.27	0.22 ^D	0.79*	0.42-1.51	0.47 ^D
rs791904	A	15%	16%	16%	1.15	0.71-1.88	0.57 ^D	1.17	0.70-1.94	0.56 ^D
rs16869463	T	9%	8%	8%	0.78*	0.43-1.43	0.42 ^D	0.77	0.40-1.47	0.42 ^D
rs78748610	C	8%	7%	6%	0.88*	0.47-1.65	0.69 ^D	0.76*	0.38-1.49	0.41 ^D
rs28607030	G	37%	35%	29%	0.82	0.53-1.27	0.38 ^D	0.57	0.36-0.90	0.011^D
rs9469583	T	50%	50%	44%	0.87	0.53-1.44	0.60 ^D	0.65	0.36-1.16	0.14 ^R
rs10947435	A	39%	43%	47%	1.28	0.81-2.03	0.29 ^D	1.90	1.15-3.14	0.010^D
rs4713675	T	45%	47%	52%	1.41	0.83-2.40	0.21 ^R	1.45	0.85-2.48	0.17 ^D

OR, odds ratio; CI, 95% confidence interval. OR adjusted for sex. Significant results in bold.

P^{Model} is the p-value of the best-fit genetic model. The choice of each genetic model was based on AIC value. D is dominant model (LOAD risk in heterozygotes or minor allele homozygotes relative to common allele homozygotes) and R is recessive model (LOAD risk in minor allele homozygotes relative to common allele homozygotes or heterozygotes).

* For these SNPs only the dominant model was considered since the rare homozygous genotype was less than 3%.

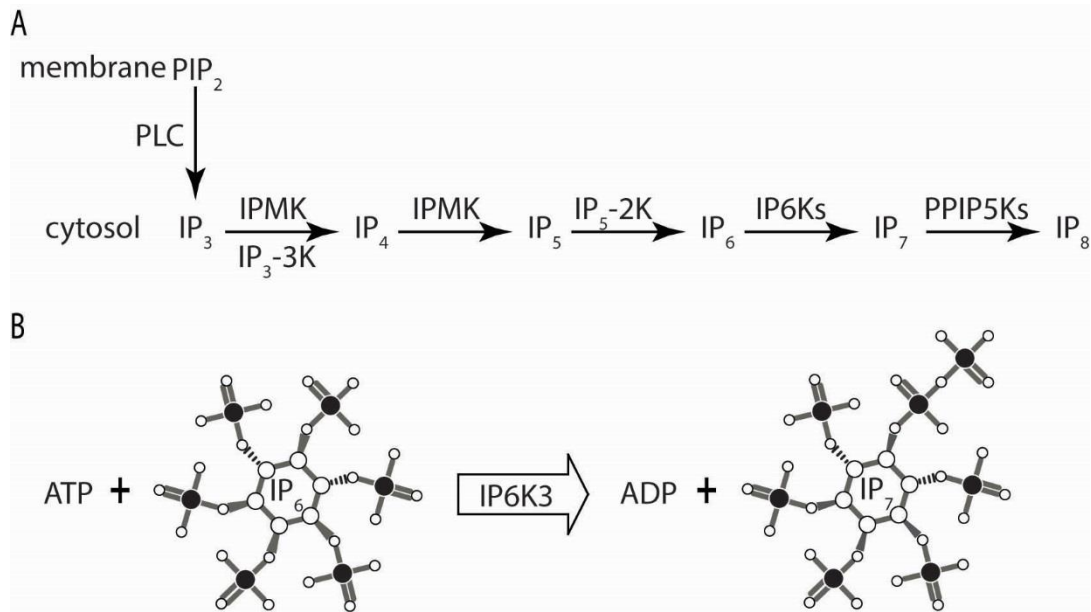


Figure 1

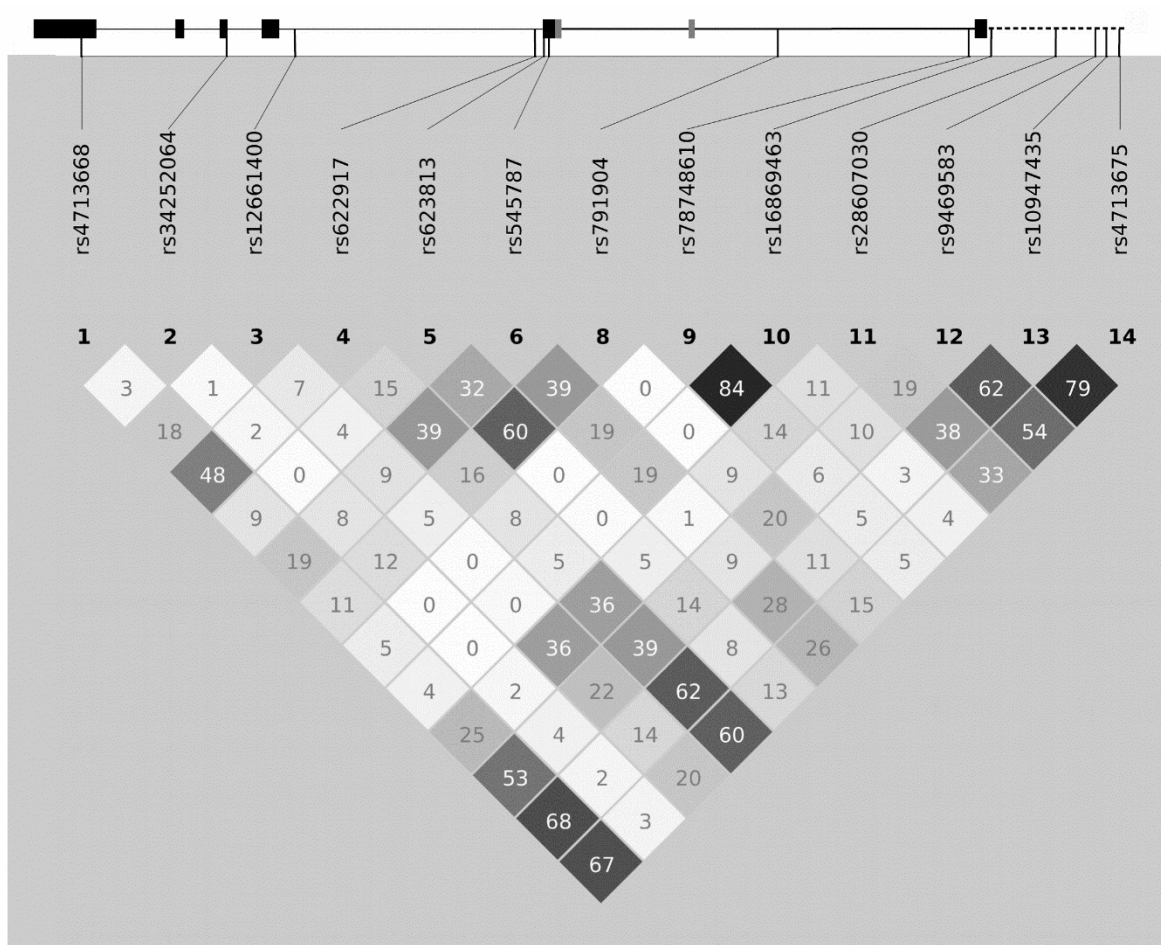


Figure 2

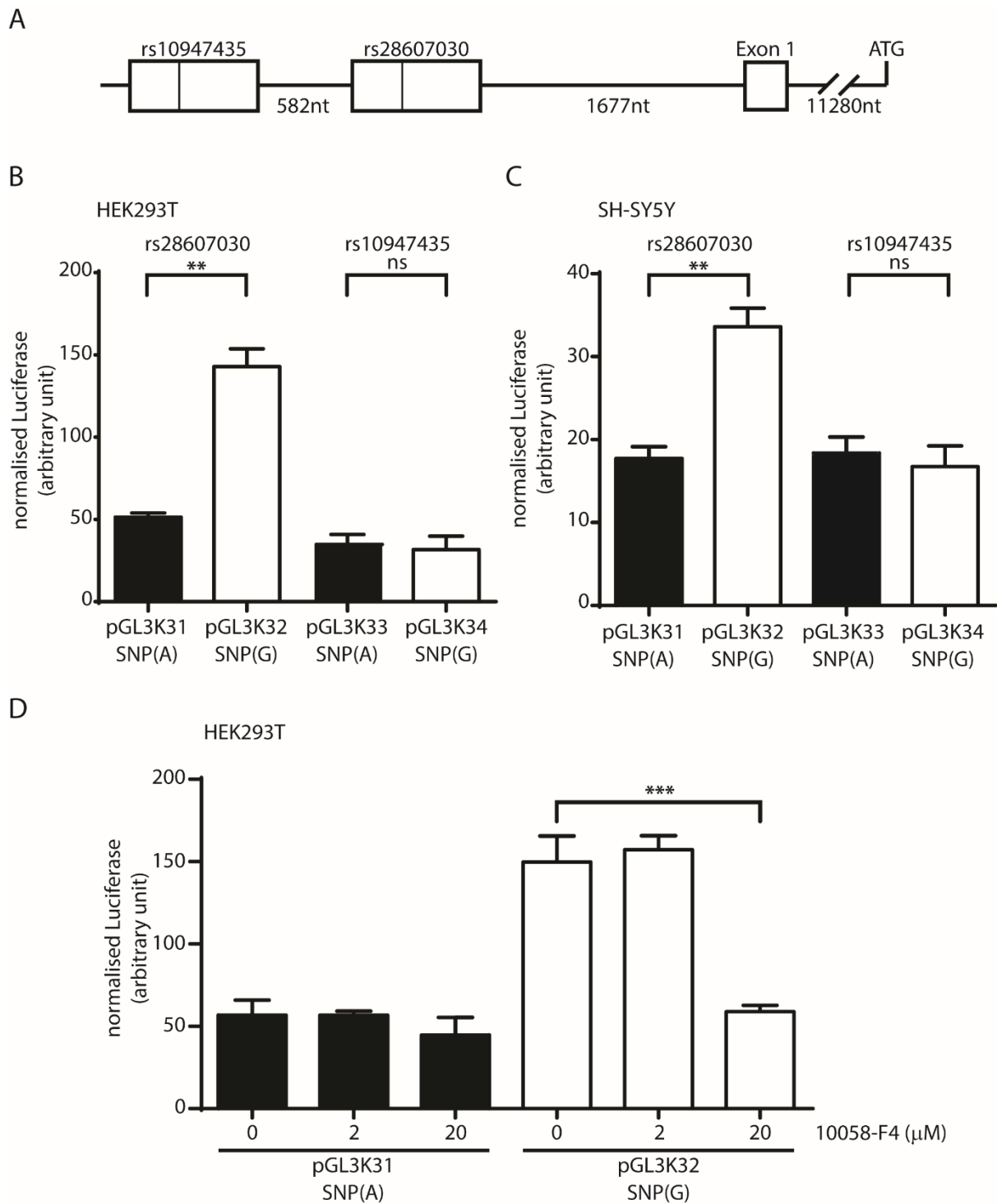


Figure 3