Volume : 3 | Issue : 4 | April 2014 • ISSN No 2277 - 8179

Research Paper

In Silico analysis on the effect of Alzheimer's related nsSNP (rs1049296) on the Structure and Function of Transferrin

Science

KEYWORDS : Trnasferrin, *In Silico*, SNPs, Alzheimer's disease

Waseem El-Huneidi

Department of Basic Sciences, College of Science and Health Professions, King Saud bin Abdulaziz University for Health Sciences, Riyadh, Saudi Arabia.

ABSTRACT

Genetic variations such as Single nucleotide polymorphisms (SNPs) are known to play crucial role in understanding the basis of many genetic diseases. SNPs in the transferrin (TF) gene have been associated with Alzheimer's disease decreasing the affinity of iron to TF leading to iron accumulation in brain cells. Recently, TF gene rs1049296 polymorphism was reported to be a genetic determinant of Alzheimer's disease, which is a main cause for dementia in elderly. In this report, In Silico analysis was conducted to predict the effect of rs1049296 on transferrin structure and function, and the results showed a non-deleterious effect of rs1049296 on transferrin.

Introduction

Alzheimer's disease is one of the main causes of memory loss in elderly, and till now the etiology of this diseases is not well understood (Hardy, 2006). Mutations are known to be one important factor that contribute (either independently or in association with each other) to Alzheimer's disease pathogenesis. Mutations in different genes (such as presenilin I and II genes) were reported to be associated with familial Alzheimer's disease (van Rensburg, Carstens, Potocnik, Aucamp, & Taljaard, 1993).

Mutations in transferrin (TF) gene, encoding transferrin (a major circulating glycoprotein involved in iron mobilization), were also reported to have a correlation with Alzheimer's disease, specially that C2 variants of TF gene was reported to be significantly higher in Alzheimer's disease patients. These variants could decrease the binding of iron by TF leading to iron overload, which was hypothesized to contribute in Alzheimer's development(Robson et al., 2004).

Since 1993, when the association between TF C2 and Alzheimer's was initially reported, 22 studies focused on this association. However, the data published on the association of TF gene rs1049296 polymorphism and Alzheimer's disease was refuting (Wang et al., 2013).

Recently, Wang et. al., reexamine the association between the TF gene rs1049296 polymorphism and Alzheimer's disease through a meta-analysis on data from 19 studies including 6310 Alzheimer cases and 13661 controls, and his meta-analysis showed that TF gene rs1049296 polymorphism is a genetic determinant of Alzheimer's disease (Wang et al., 2013).

In this study, In Silico methods were used to test the possible effect of the missense mutation rs1049296 on the structure and function of transferrin

Materials and methods

Data for rs1049296 and related protein

SNP information (SNP rsID, Gene ID) and protein sequence for transferrin were obtained from National Centre for Biotechnology Information (NCBI) database dbSNP (http://www.ncbi.nlm.nih.gov/projects/SNP).

SIFT

SIFT tool was used to predicts whether rs 1049296 affect transferrin function, this program was developed by Ng and Henikoff, based on sequence homolog (Ng& Henikoff, 2003). The concept on which this tool predicts upon is that protein function are interrelated. Thus any change occur in an evolutionary conserved regions is less tolerable than a change in an un-conserved region. SIFT algorithm uses a modified version of PSIBLAST (Altschul et al., 1997). Accordingly, any position with a tolerance index of less than 0.05 is predicted to be deleterious and those with tolerance index greater than w or equal to 0.05 are predicted to be tolerated.

PolyPhen-2

Polymorphism Phenotyping-2 (PolyPhen-2) tool was used to investigate the possible effect of nsSNPs on protein. The predictions obtained are rooted in physical and evolutionary comparative considerations and naïve Bayes classifier which uses eight sequence and three structure to calculate the deleterious effects of nonsynonymous variants (Adzhubei et al., 2010). If the scores of PolyPhen-2 are between 0.00 to 0.99, the prediction is to have a benign effect, and if the scores are in the range of 1.0 to 1.99 it could be possibly damaging and if the score is more than 2.0, it will be predicted as probably damaging to the protein(Adzhubei, Jordan, & Sunyaev, 2013).

I-Mutant 2.0

To predict the effect of SNP on protein stability, I-Mutant 2.0 was used. The differences in free energy (ΔΔG) between the unfolding Gibbs free energy value of the mutated protein with of the wild type (Kcal/mol) are used as a basis for the prediction by I-Mutant 2.0. A positive ΔΔG value resemble high stability of mutated protein and vice versa.(Bava, Gromiha, Uedaira, Kitajima, & Sarai, 2004)

PANTHER

Protein analysis using Protein ANalysis THrough Evolutionary Relationship (PANTHER) tool organize proteins with families and subfamilies of evolutionarily related proteins relying on function, pathway or biological process of the proteins (Mi, Muruganujan, & Thomas, 2013). To estimate whether nsSNPs have and functional influence on the protein, PANTHER uses subPSEC score, which is calculated according to revolutionary related proteins. SubPSEC score of less tha – 3 is predicted to be deleterious(Thomas et al., 2003).

HOPE

Have yOur Protein Defined (HOPE) tool was used to assess the structural impact of rs1049296 on transferrin (Venselaar, Te Beek, Kuipers, Hekkelman, & Vriend, 2010). These approaches analyzed the impact of the mutation, and the findings showed a structural impact. This report also shows equivalence to contacts such as metal, DNA, hydrogen bonds, and ionic interactions and evaluates whether a mutation impacts an essential contact, structural areas together with motifs, domains, and trans-membrane domains. These tools assure that the foremost dependable method to acquire data and facts is used, i.e., data regarding the "actual protein structure" that provide annotated information in UniProtKB that is utilized by prediction with DAS-servers(Prlic et al., 2007).

Results and Discussion

rs1049296 was submitted to different softwares, the SIFT program to check for the tolerance index (TI), PolyPhen 2 to examine if rs1049296 is functionally significant. I-Mutant 2 to test for the stability of the mutant and finally rs1049296 was submitted to PANTHER for further confirmation for SNP prediction.

As shown in Table 1, SIFT TI score was 0.32 indicating that rs1049296 is tolerable, PolyPhen2 score 0.002 representing a benign effect of rs1049296 on TF function, I-Mutant 2 ΔΔG is -1.65 , this negative ΔΔG value indicate the low stability of

Volume : 3 | Issue : 4 | April 2014 • ISSN No 2277 - 8179

the mutated TF, and finally SubPSEC score is of PANTHER is -2.54927 confirming a less deleterious effect of rs1049296 on TF. So out of the four softwares used, three of them (SIFT, PolyPhen 2 and PANTHER) predicted that rs1049296 does not have a deleterious effect on TF while one software (I-Mutant 2) predict that the mutated protein is less stable than the wild type

To find out what possible impact of rs1049296 on TF structure, How yOur Protein Explained (HOPE) tool was used. This tool was developed at the centre for molecular and bio-molecular informatics (CMBI), department of bioinformatics, Radboud university. HOPE evaluates the effect of the mutation on the following features: the location and the interaction of the mutated residue, in addition wither variants in this residue is common in other homologous proteins.

The structure of the wild type and mutated proteins are shown in figure 1 and 2 ,rs1049296 encoded a missense mutation of a Proline into a Serine at position 589, those two amino acids differ in their specific size, charge, and hydrophobicity-value. Prolines are known to be very rigid and therefore induce a special conformation in the backbone which might be required at this position. The mutation might be able to disturb this special conformation (Figure 1 a,b and c)

In addition, serine is smaller than proline, which might lead to loss of interactions, also, The hydrophobicity proline and serine differs.as a consequence, hydrophobic interactions, either in the core of the protein or on the surface, might be lost. But According to the HOPE report, the wild-type residue is not conserved at this position and the mutant residue is among the observed residue types at this position in other, homologous sequences. This would suggest that this variant is not damaging for the protein's structure and function.

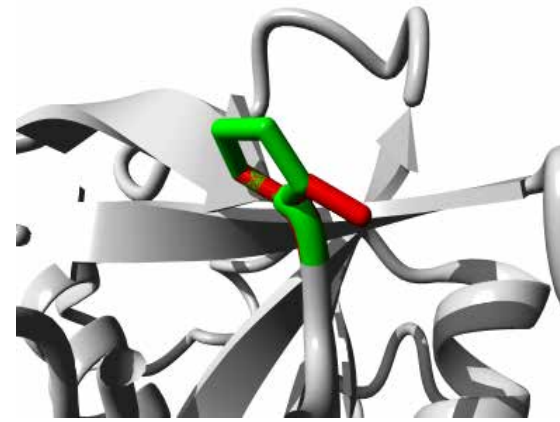
Even though, the recent meta-analysis study done by Wang et. al; (2013) concluded that rs1049296 polymorphism could be a genetic determinant of Alzheimer's disease(Wang et al., 2013), but the data were contradicting. Since 1993, when an association of TF C2 with AD was first proposed(van Rensburg et al., 1993), there have been many independent studies to further elucidate this association. Although most studies have been negative, the AlzGene meta-analysis of the allele shows a significant, although low, correlation with the Alzheimer's pathogenesis(Bertram, McQueen, Mullin, Blacker, & Tanzi, 2007) , with a similar pattern in Caucasians and east Asians. Also, a large family study supported the association(Schjeide et al., 2009). Although the results of the current study do not predict a possible deleterious effect of rs1049296 on transferrin, but it might affect the glycosylation of the protein which may affect the binding affinity of transferrin to iron.

In conclusion, the mutation encoded by rs1049296 does not show to have an apparent effect on the structure and function of the corresponding protein TF, although further dynamic simulation could be conducted to investigate whether this single nucleotide polymorphism could partially influence the function of transferrin under different chemical environmental conditions (i.e. in the presence of iron and/or glycosylation).

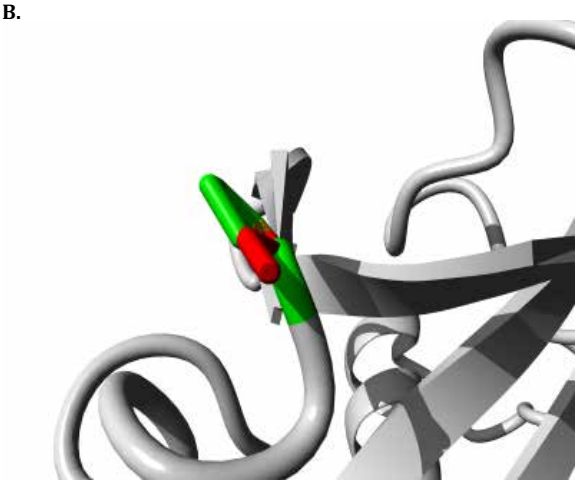
Table 1: Results of rs1049296 from SIFT, PolyPhen 2, PANTHER and I-Mutant 2

Gene	rs ID	Amino acid position	SIFT	PolyPhen 2	PANTHER	I-Mutant 2
TF	rs 1049296	P589S	0.32	0.002	-2.54927	-1.65

A.



B.



C.

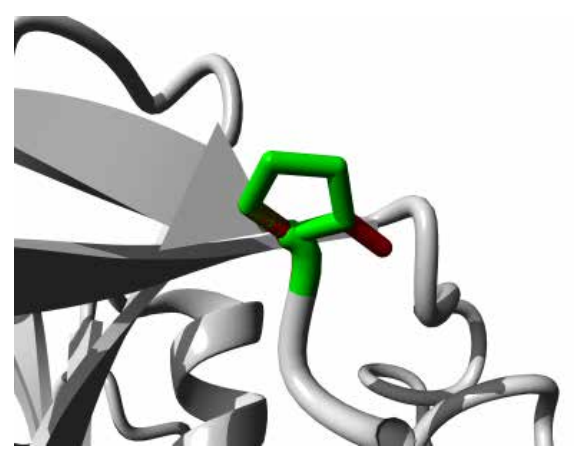


Figure 1 a. Close-up of the mutation. The protein is coloured grey, the side chains of both the wild-type and the mutant residue are shown and coloured green and red respectively. b. and c seen from a slightly different angles.

REFERENCE

Adzhubei, I., Jordan, D. M., & Sunyaev, S. R. (2013). Predicting functional effect of human missense mutations using PolyPhen-2. Curr Protoc Hum Genet, Chapter 7, Unit7 20. | Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., . . . Sunyaev, S. R. (2010). A method and server for predicting damaging missense mutations. Nat Methods, 7(4), 248-249. | Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res, 25(17), 3389-3402. | Bava, K. A., Gromiha, M. M., Uedaira, H., Kitajima, K., & Sarai, A. (2004). ProTherm, version 4.0: thermodynamic database for proteins and mutants. Nucleic Acids Res, 32(Database issue), D120-121. | Bertram, L., McQueen, M. B., Mullin, K., Blacker, D., & Tanzi, R. E. (2007). Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. Nat Genet, 39(1), 17-23. | Hardy, J. (2006). A hundred years of Alzheimer's disease research. Neuron, 52(1), 3-13. | Mi, H., Muruganujan, A., & Thomas, P. D. (2013). PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. Nucleic Acids Res, 41(Database issue), D377-386. | Ng, P. C., & Henikoff, S. (2003). SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res, 31(13), 3812-3814. | Prlic, A., Down, T. A., Kulesha, E., Finn, R. D., Kahari, A., & Hubbard, T. J. (2007). Integrating sequence and structural biology with DAS. BMC Bioinformatics, 8, 333. | Robson, K. J., Lehmann, D. J., Wimhurst, V. L., Livesey, K. J., Combrinck, M., Merryweather-Clarke, A. T., . . . Smith, A. D. (2004). Synergy between the C2 allele of transferrin and the C282Y allele of the haemochromatosis gene (HFE) as risk factors for developing Alzheimer's disease. J Med Genet, 41(4), 261-265. | Schjeide, B. M., McQueen, M. B., Mullin, K., DiVito, J., Hogan, M. F., Parkinson, M., . . . Bertram, L. (2009). Assessment of Alzheimer's disease case-control associations using family-based methods. Neurogenetics, 10(1), 19-25. | Thomas, P. D., Campbell, M. J., Kejariwal, A., Mi, H., Karlak, B., Daverman, R., . . . Narechania, A. (2003). PANTHER: a library of protein families and subfamilies indexed by function. [Comparative Study]. Genome Res, 13(9), 2129-2141. | van Rensburg, S. J., Carstens, M. E., Potocnik, F. C., Aucamp, A. K., & Taljaard, J. J. (1993). Increased frequency of the transferrin C2 subtype in Alzheimer's disease. Neuroreport, 4(11), 1269-1271. | Venselaar, H., Te Beek, T. A., Kuipers, R. K., Hekkelman, M. L., & Vriend, G. (2010). Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. BMC Bioinformatics, 11, 548. | Wang, Y., Xu, S., Liu, Z., Lai, C., Xie, Z., Zhao, C., . . . Bi, J. Z. (2013). Meta-analysis on the association between the TF gene rs1049296 and AD. Can J Neurol Sci, 40(5), 691-697. |