Gene Reports xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Gene Reports

journal homepage: www.elsevier.com/locate/genrep

Computational analysis unravels novel destructive single nucleotide polymorphisms in the non-synonymous region of human caveolin gene

Nitin Chitranshi [a,*], Yogita Dheer [a], Roshana Vander Wall [a], Veer Gupta [b], Mojdeh Abbasi [a], Stuart L. Graham [a,c], Vivek Gupta [a]

[a] Faculty of Medicine and Health Sciences, Macquarie University, F10A, 2 Technology Place, North Ryde, NSW 2109, Australia
[b] School of Medical Sciences, Edith Cowan University, Perth, Australia
[c] Save Sight Institute, Sydney University, Sydney, NSW 2109, Australia

ARTICLE INFO

Article history:
Received 21 June 2016
Received in revised form 15 August 2016
Accepted 17 August 2016
Available online xxxx

Keywords:
SNP
Mutation
Homology modelling
Caveolin

ABSTRACT

The caveolin (Cav) family of proteins comprises key constituents of caveolar structures in cell membranes, which are involved in receptor-independent endocytosis and cellular signalling pathways. Three isoforms of caveolin viz. cav-1, cav-2 and cav-3 encoded by *CAV1*, *CAV2* and *CAV3* genes respectively, have been reported to be expressed in various tissues. Genetic polymorphism in *CAV* has been identified as associated with the development of pathological changes in the cardiovascular system, chronic kidney disease as well as neurodegenerative diseases of the brain and retina such as Alzheimer's disease, Parkinson's disease and glaucoma. In this study, we investigate and characterise various polymorphisms associated with *CAV1*, *CAV2* and *CAV3* by using a combination of *in silico* algorithms such as SIFT, Polyphen 2.0, I-Mutant, PROVEAN, PANTHER, SNP&Go, PhD-SNP, MutPred and SNPEffects. Three-dimensional comparative modelling was performed using Phyre2 server, *ab initio* modelling, using the I-TASSER and RaptorX program. The predicted models were evaluated using Ramachandran plot to establish the accuracy of the models generated. The resulting mutant and wild type proteins obtained were energy minimized in Swiss Deep Viewer and evaluated. The study has identified two of the non-synonymous single nucleotide polymorphism (nsSNP) in *CAV3* gene that may have a damaging effect on the protein stability. The surface residues in the wild type and mutant forms highlight different accessible surface area (ASA) of amino acid residues in the corresponding proteins. Our analysis predicted that none of the known nsSNPs have a negative effect on the CAV1 and CAV2 protein structures. Phylogenetic analysis using ConSurf further identified that most of the disease-associated nsSNPs were within the conserved regions in human cav3.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

Plasma membranes contain specialised type of lipid rafts and sub-domains for uptake of substances from the environment while maintaining the cellular integrity (Lu et al., 2015). Caveolae constitute invaginations in the membrane that collect and transfer macromolecules to sub-cellular locations such as Golgi apparatus (Hayer et al., 2010; Jeong et al., 2012). Morphologically, caveolae are around 100-nm or smaller in size and form flask-shaped structures at the plasma membrane (Briand et al., 2011; Sowa, 2012). These are rich in cholesterol and sphingolipids but also contain an important protein constituents such as caveolin. Caveolae are well known to be distributed in astrocytes (brain), endothelial cells, lungs as well as striated and smooth muscles (Song et al., 1996; Hillman et al., 2001; Thompson et al., 2014). Our group has also identified that caveolins are well expressed in the retinal ganglion cells in the retina (Gupta et al., 2012b).

Three members of the caveolin (*CAV*) gene family have previously been identified. Caveolin-1 (*CAV1*) contains three exons and is highly conserved in sequence and structure across species which further can be of either α or β isoform. *CAV1* α isoform contains 32 more residues than the β isoform and results in a protein of ~3 kDa larger size. *CAV2* has three isoforms α, β and γ (Liu and Sowa, 2014; Totta et al., 2015) and its peptide sequence shows significant similarity to that of *CAV1* while contributing to caveolae structure formation by heterodimerizing with CAV1 protein (Totta et al., 2015). Caveolin-3 (*CAV3*) is predominantly expressed in muscles. Human caveolin-2 is ~38% identical and

Abbreviations: SNP, Single nucleotide polymorphism; nsSNP, non-synonymous single nucleotide polymorphism; Cav, caveolin; IGFBP-5, insulin-like growth factor binding protein-5; CKD, chronic kidney diseases; iNOS, inducible nitric oxide synthase; HIF-1α, hypoxia inducible factor-1 alpha; IFN-γ, interferon-gamma; IL-6, interleukin 6; H3K9me3, histone H3 lysine 9 trimethylation; RGCs, retinal ganglion cells; OMIM, Online Mendelian Inheritance in Man; dbSNP, database of single nucleotide polymorphism; UNIPROT, universal protein resource; PSIC, position-specific independent count; SVM, support vector machine; PROVEAN, protein variation effect analyzer; subPSEC, substitution position-specific evolutionary conservation; GO, gene ontology; 3D, three dimensional; PB, Poisson-Boltzmann; HMM, hidden Markov model.

* Corresponding author at: Faculty of Medicine and Health Sciences, 75, Talavera Road, Macquarie University, Sydney, NSW 2109, Australia.
E-mail address: nitinchitranshi@gmail.com (N. Chitranshi).

http://dx.doi.org/10.1016/j.genrep.2016.08.008
2452-0144/© 2016 Elsevier Inc. All rights reserved.

ARTICLE IN PRESS

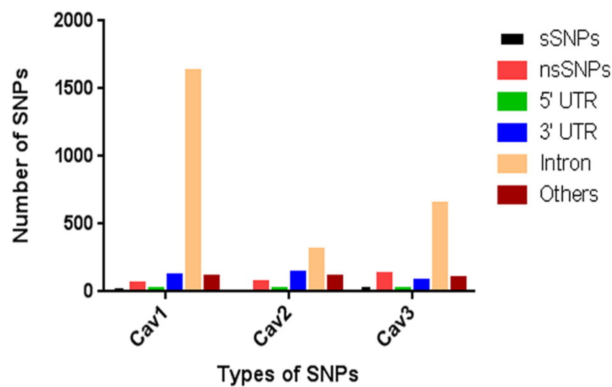

Fig. 1. Graphical representation of single nucleotide polymorphism (SNPs) in different regions denoted with different colours, synonymous SNP (black), nonsynonymous SNP (red), 5′ UTR (green), 3′ UTR (blue), introns (orange) and others (brown) for *cav* gene (based on the dbSNP database). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

~58% similar to human caveolin-1, while caveolin-3 is ~65% identical and ~85% similar to caveolin-1. In addition, a short stretch of eight amino acids FEDVIAEP has been reported that constitutes the "caveolin signature sequence," a motif that is identical between all three caveolin proteins excluding the γ isoform of CAV2.

Changes in CAV1 expression has been suggested to be involved in pulmonary hypertension and an increase in CAV1 expression in smooth muscle cells has been observed during scar tissue formation in arteries (Huang et al., 2015). Matrix protein fibronectin has been found to moderate the cell signalling pathways akin to RhoA-PI3K, Akt and ERK 1/2 for cell growth by modulation of CAV1 phosphorylation (Park et al., 2011). CAV1 has been implicated in playing a critical role in membrane trafficking and signal transduction in tissue fibrosis. It binds with insulin-like growth factor binding protein-5 (IGFBP-5) in the fibroblasts (Yamaguchi et al., 2011; Rajala et al., 2013). The changes in CAV1 expression are also associated with pathological changes in the blood-retinal barrier in diabetic retinopathy and in chronic inflammation in posterior uveitis (Klaassen et al., 2009; Hauck et al., 2010). The common variant of *CAV1* and *CAV2* at 7q31 (rs4236601) were reported to be associated with primary open angle glaucoma (POAG) (Thorleifsson et al., 2010). The clinical examination of *CAV1* single nucleotide polymorphism (SNP) rs4730751 is associated in chronic kidney diseases (CKD) (Chand et al., 2015). Kastelijn et al reported the *CAV1* SNP allele rs3807989 to be associated with the development of bronchiolitis obliterans syndrome (Kastelijn et al., 2011).

CAV2 is shown to express in bone-marrow derived macrophages (BMDMs) and regulate various inflammatory responses including expression of inducible nitric oxide synthase (iNOS), hypoxia inducible factor-1 alpha (HIF-1α), interferon-gamma (IFN-γ) and interleukin 6 (IL-6) (Maceckova et al., 2015). It also promotes induction of *Egr-1* and *JunB* genes on insulin response by decreasing histone H3 lysine 9 trimethylation (H3K9me3) (Jeong et al., 2015). Genome wide association studies have shown that *CAV2* SNP (rs2270188) interacts with dietary fat and is associated with the type 2 diabetes (Fisher et al., 2011).

CAV3 is an upstream negative regulator of eNOS and forms a complex with eNOS, inhibiting its uptake in cardiac myocytes. MicroRNA-22 is shown to bind to the *CAV3* during ischemia reperfusion injury by modulating the CAV3/eNOS signalling (Chen et al., 2015b). CAV3 is also involved in immune system signalling, T-cell cytokine production and activation of inflammatory pathways (Tran et al., 2015). Various CAV3 mutations such as F97C, P104L and T78M have been found to be associated with long QT syndrome 9 (LQT9) (Vaidyanathan et al., 2013). In glaucomatous conditions CAV1 and CAV3 undergo hyper-phosphorylation in retinal ganglion cells (RGCs) under stress and bind to *PTPN11* to regulate the neurotropin signalling (Gupta et al., 2012b).

In this study using a combination of bioinformatics approaches, we investigate the impact of various known *CAV* polymorphisms on the predicted structures of these proteins to identify the deleterious changes. Our investigational study involved (i) retrieval of SNPs in caveolin gene (*CAV1*, *CAV2* and *CAV3*) from available databases, (ii) allocating the deleterious nsSNPs to their phenotypic effects, based on sequence and structure-based homology analysis, (iii) predicting the effects of specific substitutions of amino acids on secondary structures by means of solvent accessibility and stability analysis, (iv) and prediction of changes in the tertiary and domain structures due to the mutations. This study is the first in depth analysis of the caveolin gene family *in silico* and will establish a strong foundation for structure-function and population based studies in future.

2. Materials and methods

2.1. Retrieval of SNP datasets

The data on human *cav 1*, *cav 2* and *cav 3* genes were derived from web-based public repositories such as Online Mendelian Inheritance in Man (OMIM) (Amberger et al., 2015), the SNPs information (protein accession number and SNP IDs) of the *cav1*, *cav2* and *cav3* gene were retrieved from the NCBI dbSNP (database of single nucleotide polymorphism) (Sherry et al., 2001), and the protein sequence and protein structure subsequently were retrieved from Uniprot (universal protein resource) (Magrane and Consortium, 2011).

2.2. Disease-associated SNP prediction

The disease associated single nucleotide polymorphism occurring in the protein coding region were evaluated using SIFT, Polyphen, I-Mutant, PROVEAN and PANTHER tools. SIFT uses sequence homology-based approach to classify amino acid substitutions (Ng and Henikoff, 2003). The prediction score <0.05 is considered to be deleterious. The accuracy level of the SIFT program shows 88.3–90.6% specificity and 67.4–70.3% sensitivity (Leong et al., 2015), when tested with different datasets of human variants. Polyphen determines if the amino acid change is occurring at the site that is highly conserved and the variation has any deleterious effect on the protein structure (Adzhubei et al., 2010). The position-specific independent count (PSIC) score difference of 1.5 and above obtained from Polyphen server is predicted to show functional and structural impact on protein (Ramensky et al., 2002). We filtered the nsSNPs that were predicted to be deleterious and damaging using both approaches. Further, we used the program I-Mutant 2.0, support vector machine (SVM) based tool for the automatic prediction of protein stability changes upon single-point mutations (Capriotti et al., 2005b). This program was tested on a data set derived from ProTherm (Gromiha et al., 1999), which is presently the most comprehensive available database of thermodynamic experimental data of free energy changes of protein stability upon mutation under different conditions. The output files show the predicted free energy change value or sign (DDG), which is calculated from the unfolding Gibbs free energy value of the mutated protein minus the unfolding Gibbs free energy value of the native type (kcal/mol). Positive DDG values mean that the mutated protein possesses high stability and *vice versa*. PROVEAN (Protein Variation Effect Analyzer) is a sequence based predictor that estimates the effect of protein sequence variation on protein function (Choi and Chan, 2015). It is based on a clustering method where BLAST hits with more than 75% global sequence identity are clustered together and the top 30 clusters are averaged within and across clusters to generate the final PROVEAN score. A protein variant is predicted to be "deleterious" if the final score is below a certain threshold (default is -2.5), and is predicted to be "neutral" if the score is above the threshold (Leong et al., 2015). PANTHER tool estimates the likelihood of a

Please cite this article as: Chitranshi, N., et al., Computational analysis unravels novel destructive single nucleotide polymorphisms in the non-synonymous region of human caveolin g..., Gene Reports (2016), http://dx.doi.org/10.1016/j.genrep.2016.08.008

Download English Version:

https://daneshyari.com/en/article/5590027

Download Persian Version:

https://daneshyari.com/article/5590027

Daneshyari.com