Journal of Rashid Latif Medical College

Original Article

Open Access

In Silico Analysis of Structural & Functional Impact of Non-Synonymous SNPs in DC-SIGN Receptor Gene

Muhammad Zubair Mehboob, Muhammad Usman, Ubair Aziz, Naeem Mahmood Ashraf*, Muhammad Javed Iqbal, Awais Asif, Kalsoom Sughra and Sajid Mehmood
Department of Biochemistry & Molecular Biology, University of Gujrat, Punjab, Pakistan.

KEY WORDS:

- In Silico Analysis
- Structural Impact
- Functional Impact
- Non-Synonymous SNPs
- DC-SIGN Receptor
- SNP

Single Nucleotide Polymorphism

ABSTRACT

Introduction: Cd209 encodes very important trans membrane receptor DC-SIGN in DC and macrophages to recognize and bind with carbohydrate ligands present on surface of pathogen. Structural alterations of this receptor can greatly influence its functions. nsSNPs is important group of SNPs family which brings structural alteration in protein. In this study, we sorted out SNPs present in coding region of CD20 and also evaluated their impact of nsSNPs on structure of protein through computational tools. **Methodology:** Genomic data has been retrieved from 1000 Genome project and then sorted out by using computational tools including SNpEff, PolyPhen-2, SIFT, PMUT, MutPred, I-Mutant and PROVEAN. Results: Out of 177 SNPs of coding region, 30 were sorted out as nsSNPs. Among 30 nsSNPs, only 12 nsSNPs have been predicted deleterious. **Results:** All pathogenic nsSNPs have been found in Neck and C-lectin damain of receptor which is involved in recognition and binding of ligands. **Conclusion:** It can be concluded that these nsSNPs can play their role in genetic susceptibility of individuals toward infectious

Corresponding Author: Muhammad Zubair Mehbbob'Department of Biochemistry and Molecular Biology, University of Gujrat. **Email:** Zubair.mehboob@gmail.com

Introduction

CD209 gene, which is expressed in alveolar macrophages and dendritic cells, is responsible to encode DC-SIGN (Dendritic Cell-Specific ICAM3-Grabbing Non integrin)[1-3]. DC-SIGN binds with variety of endogenous and exogenous glycosylated ligands. Among exogenous ligands, there is a long list of notorious pathogenic bacteria and viruses which bind with DC-SIGN through their glycosylated moieties present at their outer surface for instance, HIV, M. tuberculosis, Streptococcus pneumonia etc [4-11].

Non-synonymous SNPs are single nucleotide substitutions which bring amino acid change in corresponding protein structure. These miss-sense variants might alter the structure, stability and function of protein which could ultimately lead to drastic phenotypic outcomes[12-15]. Therefore, discrimination between neutral and deleterious nsSNPs can help to define approprate diagnosis and can help to define approprate diagnosis and treatment strategies for disease. Recently, rumbers of research studies have been carried out to find out role of different CD209 variants in human suscepibility to different infectious diseases including HIV, Dengue, M. tuberculosis etc[16-17] and in these studies some of the variants has been found to be associated with susceptibility of human population to certain infections diseases but still there are huge number of repeated nsSNPs in CD209 which are not so far investigated for their phenotypic effect. On rational ground, taking into account the vastness of nsSNPs in CD209 it would be difficult and time consuming to analyze their impact on DC-SIGN structure through patient-control GWAS studies. In this study, by using computational tools we have tried first to sort out synonymous and non-synonymous SNPs in CD209 gene and then checked impact of each nsSNPs on protein structure.

Fig-1 Shows the work flow designed for nsSNPS effects on human DC-SIGN protein. All data of SNPs of cd209 were retrieved from ENSEMBLE v76. All mutations in cd209 were annotated by using SpEff-v4.1 based on human genome assembly GRCh37.71. There are 11 transcripts (splice variants) of cd209, out of them canonical transcript was considered for functional effect prediction.

POLYPHEN -2:

Polymorphism phenotype-2 tool was accessed by using (http://genetics. bwh.harvard.edu/pph2/) URL to predict potential effect of the amino acid substitution i.e. damaging or benign by utilizing structural and evolution characteristics. Protein sequence substituted position and amino acid were provided to server. Prediction with the score of "probably damage", "possibly damage" and "benign" are considered as "affecting the structure or function", "may or may not affecting the structure or function", and "don't affect the function of protein respectively. Polyphon score range is from 0-1. If score is near to 1, nsSPs will probably be damaging.

SIFT:

Sorting intolerant from tolerant (SIFT) web tool is accessed at (http://sift.jcvi.org) for analysis of deleterious nature of mis-sense mutation. SIFT uses the protein database by PSI-BLAST and collected functionally related protein sequences. Subsequently by performing

the homologous alignment of sequences, it finds out the probability of the existence of an amino acid at particular site. The scores <0.05 are considered as intolerated whereas scores >0.05 are took as tolerated.

PMUT:

It is accessed at (http://Mmb2.pcb.ub.es), tells about the substituted amino acid (pathological or neutral) in particular position. It works by combing the sequence alignment with structural factors by using feed-forward neural network. PMUT needs protein sequence and mutation prediction. The PMUT generated output consist of the confidence index and binary prediction of “neutral” vs “pathological”.

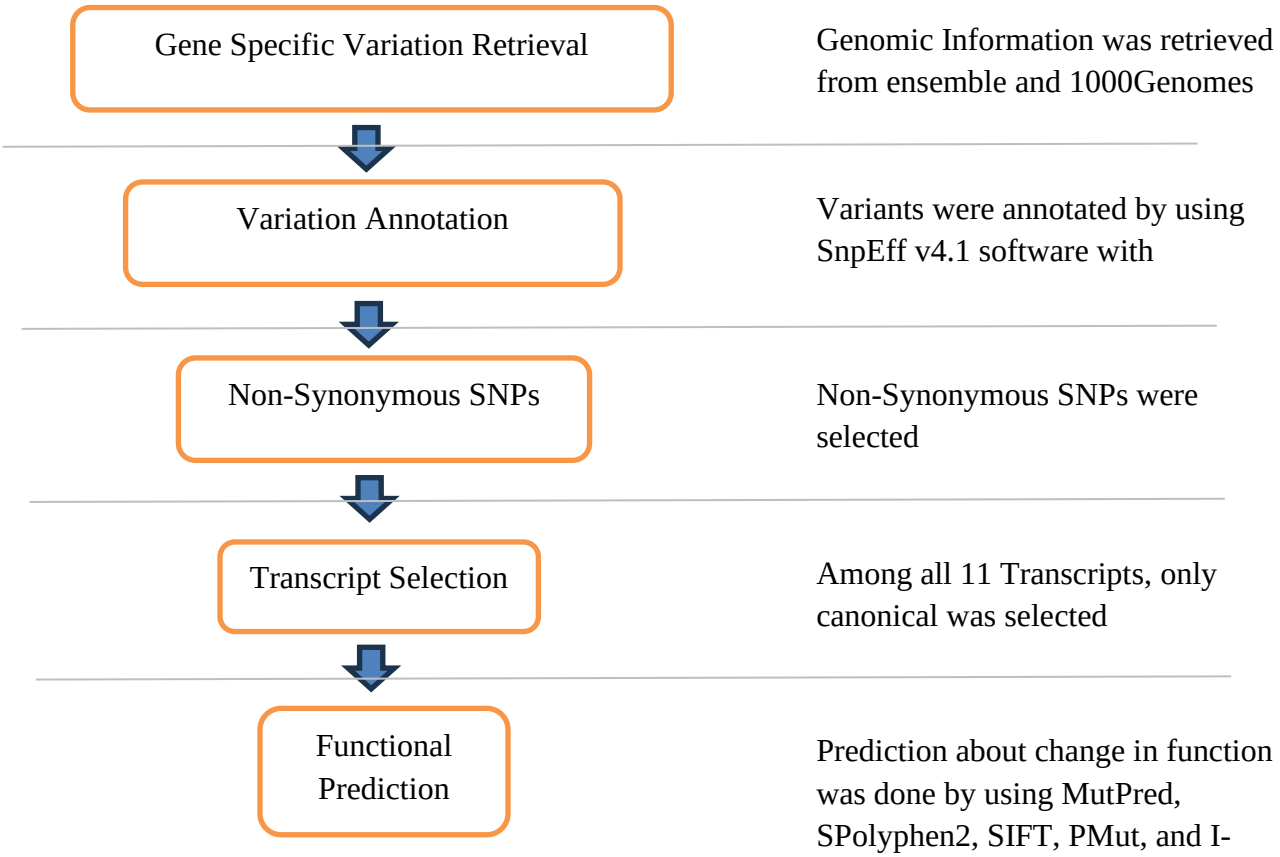


Figure 1: workflow for predicting the effect of nsSNPs on cd209 human protein

MutPred:
Mutation prediction tool was accessed by using (http://mutpred.mutdb.org) URL to predict about the changes of structural features and functional site due to amino acid substitution. MutPred builds upon the established SIFT method and a gain or loss of 14 different functional and structural properties. Protein sequence and substituted amino acid position was given to server. In MutPred results G-value range is from 0-1. If the g-value is near to 1, it shows more effect on protein function due to substituted amino acid.

I-MUTANT 3.0:
I-mutant 3.0 is a web server accessed at (), used for predicting stability of the protein in case of single change in amino acid. This software use data from ProTherm, which is a database providing experimental proved free energy change of protein stability due to change of amino acid. Protein sequence is provided along with new residue and position for obtaining the free energy change. It gives the DDG value with sign. DDG value with positive sign is the indication of mutated protein with high stability.

Results
Genomic information which was retrieved from Ensemble and 1000 Genome project has revealed presences of total 177 SNPs in CD209 gene. Sorting of these SNPs through SpEff software, by usingmGRCh37.71 chromosomal assembly has been carried out. This computation analysis has shown that among 177 SNPs only 30 were nsSNPs while rest were synonymous SNPs, having no impact on protein structure. Now these 30 nsSNPs were further investigated by using different bioinformatic tools to check their impact on protein structure.

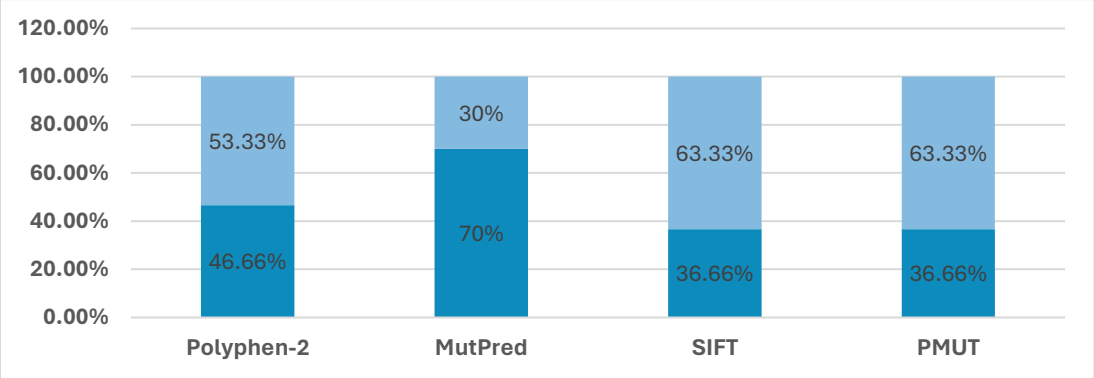
In first of sturctural characterization of 30 nsSNPs was carried out by using PolyPhen-2 which predicted 16 nsSNPs damaging for protein structure. Furthermore, results of Polyphen-2 were validated by using SIFT which predicted 19 nsSNPs. Among 19 deleterious nSSNPs, 12 damaging nsSNPs were common in PolyPhen-2 and SIFT. Moreover, PMUT and PROVEN tools were used to filter out more pathogenic nsSNPs.
Results of these softwares have validated the outputs of PolyPhen-2 and SIFT. Summary of all these results is presented in table-2 (Data not shown).
In next step, effect of each 30 nsSNPs on structual stability of protein was calculated by using I-Mutant software package. Result of this analysis has revealed that 24 nsSNPs among 30 might reduce the stabilty of protein. Results of this analysis is presented in table-2(Data not shown). The functional aspect of each nsSNPs was evluvated by using MutPred software. Functional changes, which have been predicted by MutPred include loss of glycosylation, Gain of helix, loss of sheet etc. The result of this functional annotation has been shown in table-2(Data not shown).
Since all the tools, which were used in this study are based on different alogorithum that's why every tools has predicted different number and type of possibly damaging and benign nsSNPs among 30 nsSNPs as shown in Fig-2. By carefully analyzing results of all computational tools, 12 nsSNPs are further sorted out which are common in all and their results are summarized in Table-3(Data not shown).

Discussion

DC-SIGN, encoded by CD209 structurally composed of three different domains; a neck domain, C-type lectin carbohydrate recognition domain and N-terminal trans-membrane domain. Neck domain and C-type lectin domain are part of extra-cellular region of DC-SIGN receptor where they are responsible to recognize pathogens and bind with them through carbohydrate ligands present on surface of pathogens. Any variations in these two regions can affect the capability DC-SIGN to efficiently recognize pathogen and bind with specific ligand. In this study among 30 nsSPs, 12 nsSNPs has been sorted out depending upon their potential deleterious impact on protein structure. Deleterious nsSNPs which are found in neck domain include S233C, R221W, T2201, S210F, R198W while T3141, A283T, R275W, A357T, G317E, R312H and V293L have been observed in C-lectin domain. Most of SNPs which have been validated through wet lab are present in promoter region of CD20916.17 gene that's why no literature has been found regarding nsSNPs which have been figured out in this study. nsSNPs which have been predicted deleterious by all algorithms in C-lectin domain include T3141, A283T, R275W. Similarly in Neck domain, only S233C is predicted to be pathogenic by all.

It can be concluded from above mentioned data that these changes will affect binding affinities of this receptor with pathogens. Moreover, there is need to evaluate impact of these nsSNPs through GWAS to find their association with particular infectious diseases.

Figure 2: Percentage of possible benign and damaging nsSNPs Predicted by different tools.



Conclusion

This study highlights the critical role of DC-SIGN, encoded by the CD209 gene, in pathogen recognition and binding through its neck and C-type lectin carbohydrate recognition domains. Among the 30 non-synonymous single nucleotide polymorphisms (nsSNPs) analyzed, 12 were identified as potentially deleterious to the protein's structure. Specifically, the nsSNPs S233C, R221W, T2201, S210F, and R198W in the neck domain, and T3141, A283T, R275W, A357T, G317E, R312H, and V293L in the C-type lectin domain, were found to potentially impair the binding efficiency of DC-SIGN with pathogens. Notably, T3141, A283T, and R275W in the C-type lectin domain, and S233C in the neck domain, were consistently predicted to be pathogenic by all algorithms used. These findings suggest that variations in these domains may significantly impact the receptor's binding affinities, potentially altering susceptibility to infectious diseases. Further evaluation through genome-wide association studies (GWAS) is essential to determine the clinical relevance of these nsSNPs and their association with specific infectious diseases

Declaration Of Interest: The authors declare no conflict of interest.

Author's Contribution:

Z.M: Created concept and design of the research, prepared initial draft, collected data, interpreted the results and generated discussion and conclusion.

U.A.: Data collection and Proof reading.

References

1. Soilleux EJ, Barten R, Trowsdale J (2000). DC-SIGN; a related gene, DC-SIGNR; and CD23 form a cluster on 19p13. J Immunol 165: 2937-2942.

2. Tailleux L, Pham-Thi N, Bergeron-Lafaurie A, Hermann JL, Charles P, et al. (2005) DC-SIGN induction in alveolar macrophages defines privileged target host cells for mycobacteria in patients with tuberculosis. PLoS Med 2: e381.

3. Tailleux L, Schwartz O, Hermann JL, Pivert E, Jackson M, et al. (2003) DC-SIGN is the major Mycobacterium tuberculosis receptor on human dendritic cells. J Exp Med 197: 121-127.

4.Gordon S (2002) Pattern recognition receptors: doubling up for the innate immune response. Cell
111:927-930.

5. Alvarez CP, Lasala F, Carrillo J, Muniz O, Corbi AL, et al. (2002) C-type lectins DC-SIGN and L-SIGN mediate cellular entry by Ebola virus in cis and in trans. J Virol 76: 6841-6844.

6. Geitenbeek TB, Kwon DS, Torensma R, van Vliet SJ, van Duijnhoven GC, et al. (2000) DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells. Cell 100:
587-597.

7. Geitenbeek TB, Van Vliet SJ, Koppel EA, Sanchez-Hernandez M, Vandenbroucke-Grauls CM, et al. (2003) Mycobacteria target DC-SIGN to suppress dendritic cell function. J Exp Med 197:7-17.

8. Lozach PY, Amara A, Bartosch B, Virelizier JL, Arenzana-Seisdedos F, et al. (2004) C-type lectins L-SIGN and DC-SIGN capture and transmit infectious hepatitis C virus pseudotype particles. J Biol Chem 279: 32035-32045

9. Pohlmann S, Soilleux EJ, Baribaud F, Leslie GJ, Morris LS, et al. (2001) DC-SIGNR, a DC-SIGN homologue expressed in endothelial cells, binds to human and simian immunodeficiency viruses and activates infection in trans. Proc Nati Acad Sci. 98: 2670-2675.

10. Pohlmann S, Zhang J, Baribaud F, Chen Z, Leslie GJ, et al. (2003) Hepatitis C virus glycoproteins interact with DC-SIGN and DC-SIGNR. J Virol 77:
4070-4080.

11. Tassaneetrithep B, Burgess TH, Granelli-Piperno A, Trumpheller C, Finke J, et al. (2003) DC-SIGN (CD209) mediates dengue virus infection of human dendritic cells. J Exp Med 197: 823-829.

12. C. M. Yates and M. J. E. Sternberg. (2013). The effects of non-synonymous single nucleotide polymorphisms (nsSNPs) on proteinprotein interactions. J. Mol Bio. 425 (21): 3949-3969.

13. P. C. Ng and S. Henikoff. (2006) Predicting the effects of amino acid substitutions on protein function. Ann Rev Gen Hum Genet. 7:61-80.

14. M. Gonzalez-Castej on, F. Mar' in, C. Soler-Rivas,G. Reglero, F. Visioli, and A. Rodriguez-Casado. (2011). Functional non-synonymous polymorphisms prediction methods: current approaches and future developments. C Med Chem. 18(33): 5095-5103.

15. A. Rodriguez-Casado (2012). In silico investigation of functional nsSNPs an approach to rational drug design. Res Rep in Med Chem. 2: 31-42.

16. Barreiro LB, Neyrolles O, Babb CL, Tailleux L, Quach H, et al. (2006) Promoter variation in the DC-SIGN-encoding gene CD209 is associated with tuberculosis. PLoS Med 3: e20.

17. Sakuntabhai A, Turbpaiboon C, Casademont I, ChuansumritA, Lownoo T, et al. (2005) A variant in the CD209 promoter is associated with severity of dengue disease. Nat Genet 37: 507-513.