Available online at www.sciencedirect.com

ScienceDirect

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

Original Article

Identification of functional SNPs in PAX3 gene and in silico analysis of damaging SNPs in relation to neural tube defect

Rashmi [a], *Royana Singh* [a,*], *A.N. Gangopadhyay* [b], *Anjali Rani* [c], *Mayank Shah* [d]

[a] *Cytogenetics Unit, Department of Anatomy, Institute of Medical Sciences, Banaras Hindu University, Varanasi, Uttar Pradesh, India*

[b] *Department of Pediatrics Surgery, Institute of Medical Sciences, Banaras Hindu University, Varanasi, Uttar Pradesh, India*

[c] *Department of Gynecology and Obstetrics, Institute of Medical Sciences, Banaras Hindu University, Varanasi, Uttar Pradesh, India*

[d] *Ruhelkhand Dental College, Bareilly, India*

ARTICLE INFO

Article history:
Received 1 October 2013
Accepted 27 January 2015
Available online 19 February 2015

Keywords:
Neural tube defect
paired box gene 3
mutation
gene

ABSTRACT

Introduction: PAX3 gene belongs to the class of transcription factor and has a significant role in neural tube development. There are a number of SNP which are associated with neural tube defect. Hence, we must sort functional SNP for a population study. To fulfill this goal data from dbSNP and literature review can be used.

Methods: In this study we analyzed the functional and structural impact of SNPs through computational prediction tool. A total of 8947 SNPs were observed from dbSNP in which SNP associated with neural tube defect having missense mutation is rs2234675. This nsSNP was found to be damaging by sequence homology based tool (Provean) and structural homology based tool (Polyphen). Modeling of wild and mutant protein structure were done using RMSD of wild and mutant protein structure were determined using Swiss PDB viewer and then the protein structure stability was determined using I-mutant 3.

Results: The nsSNP present in dbSNP i.e., rs2234675 was identified as deleterious, which lead too decrease in stability of PAX3 protein.

Discussion: A change of Thr315Lys, i.e. from polar neutral amino acid to polar basic amino acid showed a change in charge to positive and size of amino acid lead to change in structure. The modeled structure further, showed a decrease in stability. The result obtained from *insilico* study would open new prospect for association of PAX3 with neural tube defect.

Abbreviations: PAX, paired box; dbSNP, database single nucleotide polymorphism; nsSNP, non synonymous single nucleotide polymorphism.

* *Corresponding author at*: Cytogenetics Unit, Department of Anatomy, Institute of Medical Sciences, Banaras Hindu University, Varanasi, Uttar Pradesh, India.

E-mail addresses: drroyanasingh@gmail.com, singhroyana@rediffmail.com (R. Singh).

http://dx.doi.org/10.1016/j.jasi.2015.01.001

1. Introduction

Development of neural tube has been observed during 3rd week of pregnancy.[1] Any deformity in neural tube development would lead to neural tube defect. Neural tube defect is second common congenital malformation of birth defects. Neural tube defect can be classified into open neural tube defect (anencephaly, encephalocele, spina bifida) and closed neural tube defect (lipomyelomeningocele, lipomeningocele, and tethered cord). Epidemiological studies had leaded us to reveal that open neural tube defects are the most frequent in North India.[2] A study by Cherian shows an incidence of **6.57–8.21** per thousand live births in Balrampur District, UP.[2] Neural tube defect is polygenic and there is an interaction in between environmental and genetic factors determining the multifactorial nature.[3] The candidate loci for neural tube defect are on chromosome number 2, 7 and 10[4,5] however cytogenetics analysis revealed that trisomy of chromosome 13 and 18 have shown association with neural tube defect.[6] The genes associated with neural tube defect may be associated with the process of neurulation and are involved in biochemical pathways like folic acid metabolism and glucose metabolism, genes based on cellular function and developmental genes. Developmental pathway involves Wnt pathway and Hedgehog pathway. Homeobox genes are involved in vertebrate development.[7] PAX3 gene belongs to the class of transcription factor having a paired box domain (~128 amino acids long) and a unique DNA binding motif.[8,9] A homeodomain and an octapeptide chains are also present in PAX3.[10] PAX3 is expressed in commissural neuron and restricts ventral patterning.[11] Genetic association studies can be tested for variation in gene sequence for their involvement in Neural tube defect. SNP detection is the most common form of genetic variation study. A Thr315Lys variation in exon6 of PAX3 has been associated with neural tube defect have been reported in dbSNP with rs2234675.[12] A study was done to screen novel PAX3 polymorphism associated with neural tube defect having 15, novel polymorphism in which only two SNPs were in exonic region with restriction site rs12623857 and rs28945092.[13] This prompted us to focus on PAX3 gene leading to neural tube defects using bioinformatics tools to explore and extend the effect of SNPs on the stability and function of the PAX3 gene. The present study will be beneficial for understanding the role played by PAX3 gene and the genetic consequence of neural tube defect.

2. Methods

2.1. Data source

dbSNP, and literature survey were used to obtain the SNPs associated with PAX3 gene.[14] A total of 8947 SNPs were retrieved on 22nd May 2014 for PAX3 gene. In dbSNP, polymorphic site Thr315Lys in exon6 with rs2234675 is found.[12] A study was done to screen novel polymorphism during the survey in dbSNP only two SNP present in exonic region were observed. The SNPs present in intronic region and UTR region were not observed in dbSNP. SNPs present in exonic region were rs12623857 and rs28945092.[13]

2.2. Insilico analysis

Evaluation of functional significance of non synonymous SNP was done using a sequence homology tool Provean[15,16] and functional impact of nsSNPs was done using a structural homology based tool Polyphen.[17] After functional analysis modeling of nsSNPs on Protein Structures for PAX3 was done using 3we0A as template with resolution of 1.90 Å in RCSB PDB using Pcons.net.[18] and then calculation of their RMSD Difference *was done* using Swiss PDB viewer[19] and UCSF chimera 1.5.1.[20] Then the change in stability was determined because of SNP using I mutant 3.0.[21]

3. Result

3.1. SNP dataset

Polymorphism data of the PAX3 gene associated with neural tube defect investigated in this paper was retrieved from the dbSNP (Table 1) and literature review.[14] It contained about 8947 SNPs associated with PAX3 gene. Wei Lu, 2006 observed 15 novel SNPs with SNP ID rs28945096 (T/C), rs28945095 (C/T), rs28945094 (T/C), rs28945093 (A/C), rs28945092 (C/G), rs12623857 (T/C), rs28945091 (C/A), rs28945090 (G/A), rs28945089 (C/G), rs28945088 (C/G), rs28945087 (C/T), rs28945086 (T/C), rs16863657 (T/C), rs28945085 (C/G), rs28945094 (T/C), rs12623857 (T/C), and rs28945668 (C/T). Only two SNP with rs12623857 and rs28945092 in exonic region were present in dbSNP, the SNP in intronic and UTR region were not present. Hol et al; 1996 and Trembath et al; 1999 observed a polymorphic site in exon 6 from C→A with SNP ID rs2234675. The SNP observed showed Thr/Lys amino acid substitution at position 315 and is non-synonymous for neural tube defect.

3.2. Deleterious nsSNP prediction by Provean

The deleterious effect of nsSNP was identified by Provean (Table 2) which has a predefined score of −2.5. Below −2.5 the polymorphism is deleterious. The nsSNP has a predefined score of −2.601 and is observed to be deleterious.

3.2.1. Evaluation of functional impact of nsSNPs by PolyPhen server

The nsSNP rs2234675 was analyzed using Polyphen server (Fig. 1). The mutation was predicted to be possibly damaging with a score of **0.616** (sensitivity: **0.87**; specificity: **0.91**).[22–24]

Table 1 – nsSNPs and synonymous SNP identified by dbSNP.

SNP	Amino acid position	SNP type	Alleles	Amino acid change
rs12623857	43	Synonymous	$T_{[b]}/C_{[d]}$	$G_{[e]}$, $G_{[e]}$
rs28945092	52	Synonymous	$C_{[d]}/G_{[a]}$	$P_{[f]}$, $P_{[f]}$
rs2234675	315	Non Synonymous	$C_{[d]}/A_{[c]}$	$T_{[g]}$, $K_{[h]}$

[a]guanine, [b]thymine, [c] adenine, [d] cytosine [e] glycine, [f] proline, [g]threonine, [h] lysine.

Table 2 – List of SNPs that were analyzed by Provean.

Variant	Provean score	Prediction (cutoff = −2.5)
T315K	−2.601	Deleterious
G43G	0.000	Neutral
P52P	0.000	Neutral

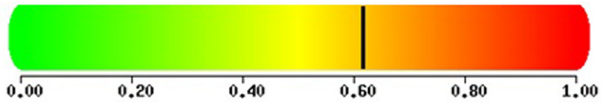

Fig. 1 – Representation of possibly damaging score using Polyphen server.

3.3. Ab-initio *modeling of mutant and wild type protein structure*

Modeling of normal PAX3 and mutant structure of nsSNP was done by Pcons.net (Table 3). Pcons.net searches for the best template 3we0A with crystallographic resolution of 1.90Å. The modeled structure has a pcons score of 0.021, Pcomb score of 0.070, ProQ2 score of 0.266 Table 4.

The native and mutant structures were aligned using Swiss PDB viewer and UCSF chimera 1. 5.3. Magic fit of Swiss PDB viewer showed the minimum deviation between two structures. To determine the similarity of wild and mutant structure superimposition was done for both the structures. The RMSD observed through Swiss PDB viewer using Magic fit was 0.94Å between the native and mutant structure (Fig. 2A). Another run was done of iterative fit leading to decrease in RMSD from 0.94Å to 0.79Å for whole structure but at polymorphic site the RMSD observed was 0.844Å without iteration but after iteration a slight decrease in RMSD was observed 0.747Å. For best structural alignment the RMSD must be zero. In this study superimposition showed a RMSD of 0.844 Å after iterative fit (Fig. 2B). The native and mutant structures observed were displayed in ribbon using UCSF chimera.

In Swiss PDB viewer after magic fit another run was done of iterative fit leading to decrease in RMSD from 0.94Å to 0.79Å for whole structure but at polymorphic site the RMSD observed was 0.844Å without iteration but after iteration a slight decrease in RMSD was observed 0.747Å (Fig. 2C).

3.4. *Prediction of protein structure stability after mutation*

The stability of protein structure after mutation was predicted using I-mutant 3 server. It predicts the protein stability using thermodynamic experimental data using Protherm.[25] The mutant structure showed a decrease in protein stability with Reliability index of 4 and Relative solvent accessible area of 56.2.

4. Discussion

A total of 8947 SNP was observed from dbSNP. However, SNP associated with neural tube defect which was non-synonymous (rs2234675) were determined using literature search. For rs2234675, the nucleotide change was in the 1325 position where ACA was replaced with AAA, *i.e.* C (Cytosine) to A (adenine) resulting in the substitution of amino acid T (threonine) to K (lysine). The change of threonine to lysine represents a missense mutation (The bases in bold are the nsSNPs). Threonine is a polar, neutral amino acid whereas lysine is polar, basic amino acid. The change of neutral amino acid to basic would be deleterious. The deleterious effect of nsSNP was identified by Provean with a score of −2.601.

The nsSNP for rs2234675 showed change of neutral polar amino acid to positively charged basic polar amino acid determining the damaging effect for the polymorphic site.

The 3D structure of nsSNP and wild type modeled structures which were subjected to determine RMSD change showed higher RMSD value, the higher the RMSD value the greater is change in wild and mutant structure. Hence, there is a chance of considerable change in the structure due to this polymorphism. The polymorphic site showed a change of threonine at 315 position of amino acid to lysine, i.e. from polar

Table 3 – Representation of modeled structure with its score.

Rank	Model	Pcons	Pcomb	ProQ2	ProQ / ProQresSum	Template	CATH	Alignment
1		0.021	0.070	0.266	0.672 / 307.370	3we0_A	N/A	

Table 4 – Predictions of protein stability change due to mutations.

Molecule model	Position of amino acid on protein molecule	Wild type	Mutant type	SVM2 Prediction effect	DDG value prediction Kcal/mol	RI: Reliability index	RSA: Relative solvent accessible area
3we0	315	Thr	Lys	Decrease	−0.71	4	56.2

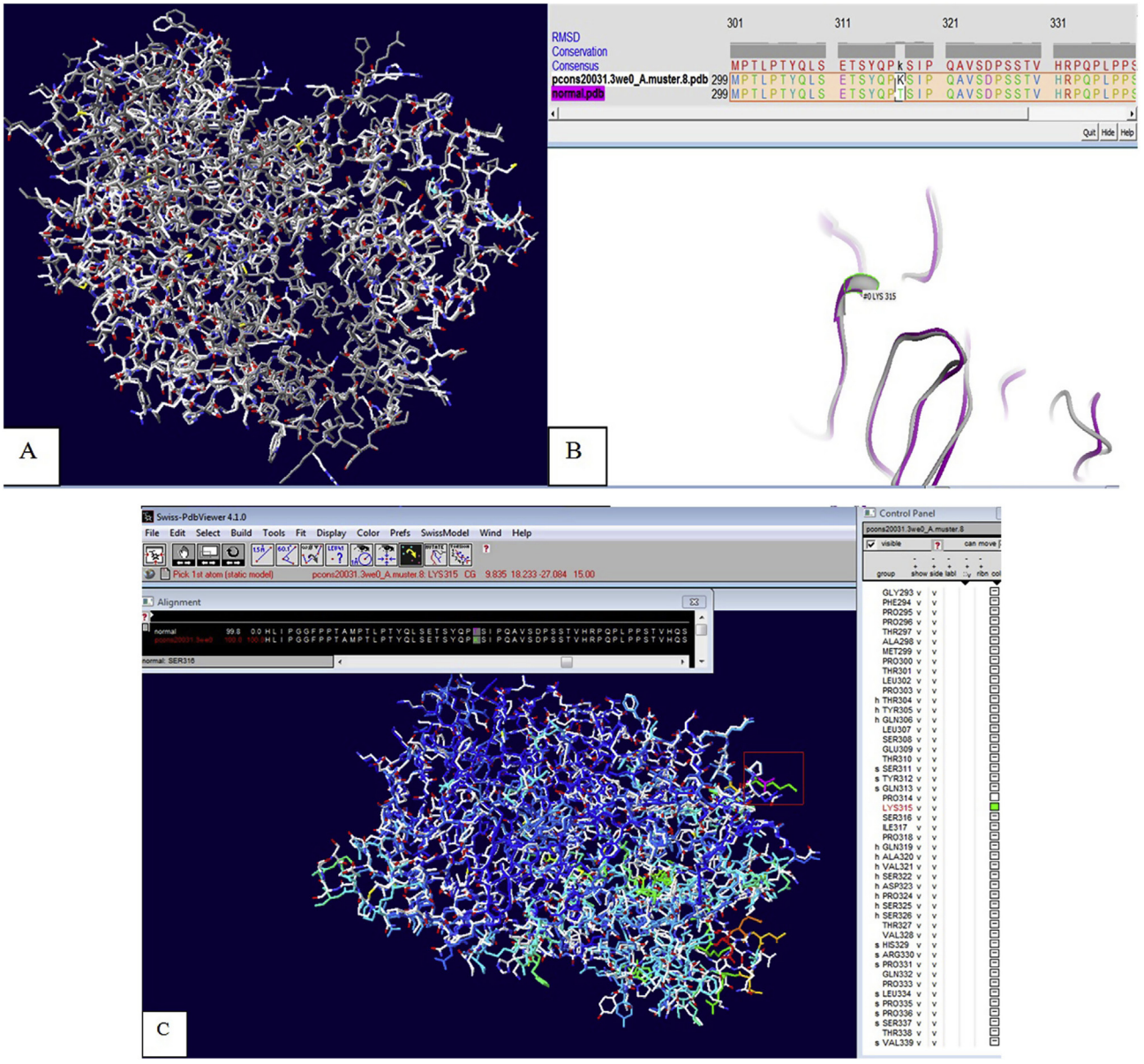

Fig. 2 – A. Superimposed structure of modeled native protein with modeled mutant protein for SNP rs2234675 with RMSD 0.94 Å. B. Representation of Lys315 on superimposition using chimera. C. Superimposition of native and mutant T315K after iterative fit. Red Box represents the superimposed structure with Thr (Pink) 315 Lys (Green) with RMSD 0.747 Å.

neutral amino acid to polar basic amino acid having positive charge. Hence a change in charge to positive and size of amino acid lead to change in structure. The modeled structure further, showed a decrease in stability.

5. Conclusion

Pax3 gene plays an important role in neural crest cell migration and development. The nsSNP observed in PAX3 gene were subjected to check for change at sequence and structural level. The nsSNP with rs2234675 seems to deregulate the function of PAX3 gene leading to improper development of neural tube leading to neural tube defect.

Conflicts of interest

All authors have none to declare.

REFERENCES

1. Kaufman BA. *Neural Tube Defects*. vol. 51. Elsevier Inc; 2004: 389–419.
2. Anil Cherian SS, Bullock Robyn Kbn, Antony Ashok Chk. Incidence of neural tube defects in the least-developed area of India: a population-based study. *Lancet*. 2005;366: 930–931.

3. Detrait ER, George TM, Etchevers HC, Gilbert JR, Vekemans M, Speer MC. Human neural tube defects: developmental biology, epidemiology, and genetics. *Neurotoxicology Teratol*. 2005;27:515–524.
4. Rampersaud EBA, Enterline DS, George TM, et al. Whole genome wide linkage screen for neural tube defects reveals regions of interest on chromosomes 7 and 10. *J Med Genet*. 2005;42:940–946.
5. Stamm DS, Siegel DG, Mehltretter L, et al. Refinement of 2q and 7p loci in a large multiplex NTD family. *Birth Defects Res A Clin Mol Teratol*. 2008;82:441–452.
6. Goetzinger KR, Stamilio DM, Dicke JM, Macones GA, Odibo AO. Evaluating the incidence and likelihood ratios for chromosomal abnormalities in fetuses with common central nervous system malformations. *Am J Obstet Gynecol*. 2008;199:285–286.
7. Boyles AL, Hammock P, Speer MC. Candidate gene analysis in human neural tube defects. *Am J Med Genet C Semin Med Genet*. 2005;135C:9–23.
8. Walther C, Guenet JL, Simon D, et al. Pax: a murine multigene family of paired box-containing genes. *Genomics*. 1991;11:424–434.
9. Noll M. Evolution and role of Pax genes. *Curr Opin Genet Dev*. 1993;3:595–605.
10. DA U. PAX proteins and fables of their reconstruction. *Crit Rev Eukaryot Gene Expr*. 2012;22:161–177.
11. Mansouri A, Gruss P. Pax3 and Pax7 are expressed in commissural neurons and restrict ventral neuronal identity in the spinal cord. *Mech Dev*. 1998;78:171–178.
12. Hol FA, Geurds MP, Chatkupt S, et al. PAX genes and human neural tube defects: an amino acid substitution in PAX1 in a patient with spina bifida. *J Med Genet*. 1996;33:655–660.
13. Lu WZH, Wen S, Laurent C, Shaw GM, Lammer EJ, Finnell RH. Screening for novel PAX3 polymorphisms and Risks of spina bifida. *Birth Defects Res (Part A)*. 2007;79:45–49.
14. Smigielski EM, Sirotkin K, Ward M, Sherry ST. dbSNP: a database of single nucleotide polymorphisms. *Nucleic Acids Res*. 2000;28:352–355.
15. Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the functional effect of amino acid substitutions and indels. *PLoS One*. 2012;7:e46688.
16. Choi Y. Fast computation of Pairwise sequence alignment scores between a protein and a set of single-locus variants of another protein. In: *Proceedings of the ACM Conference on Bioinformatics, Computational Biology and Biomedicine (BCB '12) ACM*. 2012; 414–417.
17. Johnson MMH, Chen C. Screening for deleterious non-synonymous single-nucleotide polymorphisms in genes involved in steroid hormone metabolism and response. *Cancer Epidemiol Biomarkers Prev*. 2005;4:1326–1329.
18. Wallner BEA. Identification of correct regions in protein models using structural, alignment and consensus information. *Protein Sci*. 2006;15:900–913.
19. Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*. 1997;18:2714–2723.
20. Pettersen EFG, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera–a visualization system for exploratory research and analysis. *J Comput Chem*. 2004;13:1605–1612.
21. Capriotti EFP, Casadio R. I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. *Nucleic Acids Res*. 2005;33:W306–W310.
22. Adzhubei IJD, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr Protoc Hum Genet*. 2013 [Chapter 7(Unit7):20].
23. Ramensky VBP, Sunyaev S. Human non-synonymous SNPs: server and survey. *Nucleic Acids Res*. 2002;30: 3894–3900.
24. Sunyaev SRE, Rodchenkov IV, Eisenhaber B, Tumanyan VG, Kuznetsov EN. PSIC: profile extraction from sequence alignments with position-specific counts of independent observations. *Protein Eng*. 1999;12:387–394.
25. Bava KAG, Uedaira H, Kitajima K, Sarai A. ProTherm, version 4.0: thermodynamic database for proteins and mutants. *Nucleic Acids Res*. 2004;32:D120–D121.