P-ISSN: 2618-0723
E-ISSN: 2618-0731

NAAS Rating: 5.04
www.extensionjournal.com

International Journal of Agriculture Extension and Social Development

Volume 7; SP-Issue 3; March 2024; Page No. 207-211

Received: 10-01-2024
Accepted: 18-02-2024

Indexed Journal
Peer Reviewed Journal

In-silico investigation of porcine X-chromosome reveals deleterious non-synonymous SNPS

[1, 2]Utsav Surati, [1, 2]Ravina, [1, 2]Dharamshaw CA, [1, 2]Gaurav Patel, [1, 2]Chandana Sree Chinnareddyvari, [1, 2]Pallavi Rathi, [1, 2]Rangasai Chandra Goli, [1, 2]Kiyevi G Chishi, [3]Ayushi Singh and [1, 2]Kanaka KK

[1]Animal Genetics and Breeding Division, ICAR-National Dairy Research Institute, Karnal, Haryana, India

[2]ICAR-National Bureau of Animal Genetic Resources, Karnal, Haryana, India

[3]Animal Genetics and Breeding, Indian Veterinary Research Institute, Izzatnagar Bareilly, Uttar Pradesh, India

DOI: https://doi.org/10.33545/26180723.2024.v7.i3Sc.450

Corresponding Author: Ravina

Abstract

Deleterious mutations in genomes, particularly those occurring in germ cells, can have heritable consequences. Next Generation Sequencing (NGS) technology has facilitated the discovery and prediction of novel mutations. Computational algorithms and robust bioinformatics pipelines play a crucial role in identifying genetic mutations, especially in scenarios lacking animal records and *in vivo* testing. In this study, genomic sequence data from twenty *Sus scrofa* (pig) samples obtained from the European Variation Archive repository was annotated using VEP-Ensembl. A total of 178,745 variants were processed, including 530 novel variants. Subsequently, only sex chromosomes were considered for further analysis to identify sex-linked mutations, resulting in the processing of 2300 variants, including 28 non-synonymous SNPs (nsSNPs). Using various predictive tools such as SIFT, PANTHER, Predict SNP, Polyphen-1, Polyphen-2, and SNAP, four nsSNPs were identified as deleterious, with one nsSNP found deleterious across all tools. Additionally, the I-mutant tool confirmed a decrease in protein stability for all deleterious nsSNPs. This study identified four deleterious mutations (S261F, D664N, A704, and I76N) mapped to two genes (HUWE1 and NKRF) in pigs. One mutation name D664N (HUWEI) was found to be deleterious by all tools. While most deleterious mutations in pigs have been reported in exotic breeds, limited information exists for Indian native pig breeds. The *in-silico* identification of deleterious SNPs in this study provides a foundation for further investigation into the biological consequences of these mutations in Indian pig breeds through *in vivo* studies.

Keywords: Next generation sequencing, bioinformatics, sus scrofa, deleterious mutations, *in silico* analysis.

1. Introduction

A single nucleotide polymorphism (SNP) is a source variance in a genome. SNPs account for 90% of all genetic variations in the genome and are the most prevalent and basic type of polymorphism. The nonsynonymous SNPs can decrease the protein stability and have an impact on how protein functions, which can lead to deleterious effect [1, 2]. Deleterious nsSNPs in genomes, particularly those occurring in germ cells, can have heritable consequences. Next Generation Sequencing (NGS) technology has facilitated the discovery and prediction of novel mutations. Computational algorithms and robust bioinformatics pipelines play a crucial role in identifying genetic mutations, especially in scenarios lacking animal records and *in vivo* testing.

Utilizing *in silico* tools for identifying mutational changes represents a novel and increasingly embraced approach, offering enhanced ease and reliability. These computational tools are gaining widespread adoption for pinpointing deleterious mutations across the genome in both animals and humans. They prove particularly valuable in detecting genome-wide deleterious nucleotide mutations, notably missense SNPs, and in predicting their impact on protein structure and function [3]. Previously, *in silico* analysis was instrumental in identifying deleterious mutations by scrutinizing non-synonymous alterations in key candidate genes such as BRAF, TNF-α, BARD-1, and IGF1R in humans, and SLC11A2 in animals [4-7].

While most deleterious mutations in pigs have been reported in exotic breeds, limited information exists for Indian native pig breeds. The *in-silico* identification of deleterious SNPs in this study provides a foundation for further investigation into the biological consequences of these mutations in Indian pig breeds through *in vivo* studies.

2. Materials and Methods

2.1 Data Set

Genomic sequence data from twenty *Sus scrofa* (pig) samples were obtained from the European Variation Archive repository. The data were retrieved and processed using VEP-Ensembl for annotation, encompassing a total of 178,745 variants, which included 530 novel variants.

2.2 Identification of Deleterious nsSNPs

Deleterious non-synonymous SNPs (nsSNPs) were identified to understand their impact on protein function and

phenotypic changes. The Sorting Intolerant from Tolerant (SIFT) score from the Ensembl database was utilized for predicting the deleteriousness of SNPs. SIFT assigns a tolerance index (TI) score ranging from 0.0 to 1.0, with a TI score of 0.05 or less indicating intolerance or deleteriousness [8, 9].

To validate the identified deleterious SNPs from SIFT, additional predictive tools such as PANTHER, Predict SNP, PolyPhen-1, PolyPhen-2, and SNAP were employed. Results from PolyPhen-1 and SNAP were obtained through the consensus classifier of Predict SNP.

2.3 Prediction of structural and functional effect of nsSNP

For prediction of protein stability change, I-Mutant 2.0 was used. This was a Support Vector Machine (SVM)- based web server for the automatic prediction of protein stability changes upon single-site mutations. The input protein FASTA sequence along with the residue changes were provided. I-Mutant 2.0 predicts free energy change and RI value (reliability index). If the DDG value is negative, then the mutated protein will have less stability and vice versa for high stability [10].

HOPE version 1.1.1 (https://www3.cmbi.umcn.nl/hope/) was used to recognise the structural effect of nonsynonymous change in the ANPEP protein sequence. It also provided 3D structure visualization of altered protein and superimposition of wild and mutant after providing altered protein sequence as input [11].

3. Results and Discussion

A total of 178,745 variants were processed, including 530 novel variants. Subsequently, only sex chromosomes were considered for further analysis to identify sex-linked mutations, resulting in the processing of 2300 variants, including 28 non-synonymous SNPs (nsSNPs) (Table 1). Among them, 45 percent belonged to intronic variation and 30 percent was intergenic variants. The non coding transcript variant came up to 12 percent of the total variants. The variants in upstream and downstream regions were 5 percent and 4 percent respectively. There were no synonymous variants in the coding region and 1 percent of total variants were missense in nature (Figure 1). Within the coding sequence, 64 percent were missense variants, 18 percent were synonymous. Stop retained variant was 9 percent while start loss variant constituted 2 percent of the total coding sequence variants. A 7 percent frame shift variant was also observed (Figure 2). The missense variants were analysed further.

Table 1: Details of processed variants in sex chromosome of Pig

Category	Count
Variants processed	2300
Variants filtered out	0
Novel/ existing variants	2298 (99.9)/2 (0.1)
Overlapped genes	171
Overlapped transcripts	379
Overlapped regulatory features	-

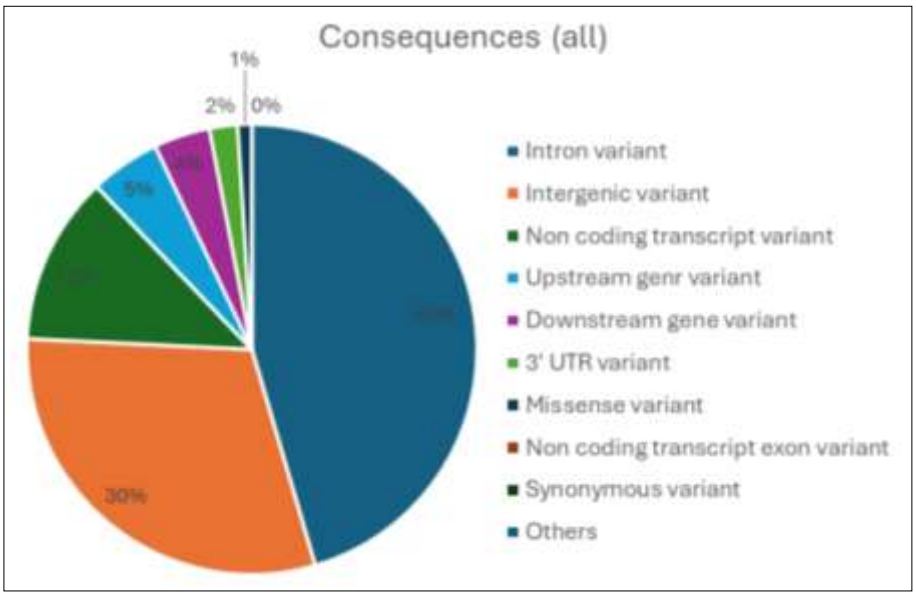

Fig 1: Details of variants in sex chromosome of Pig

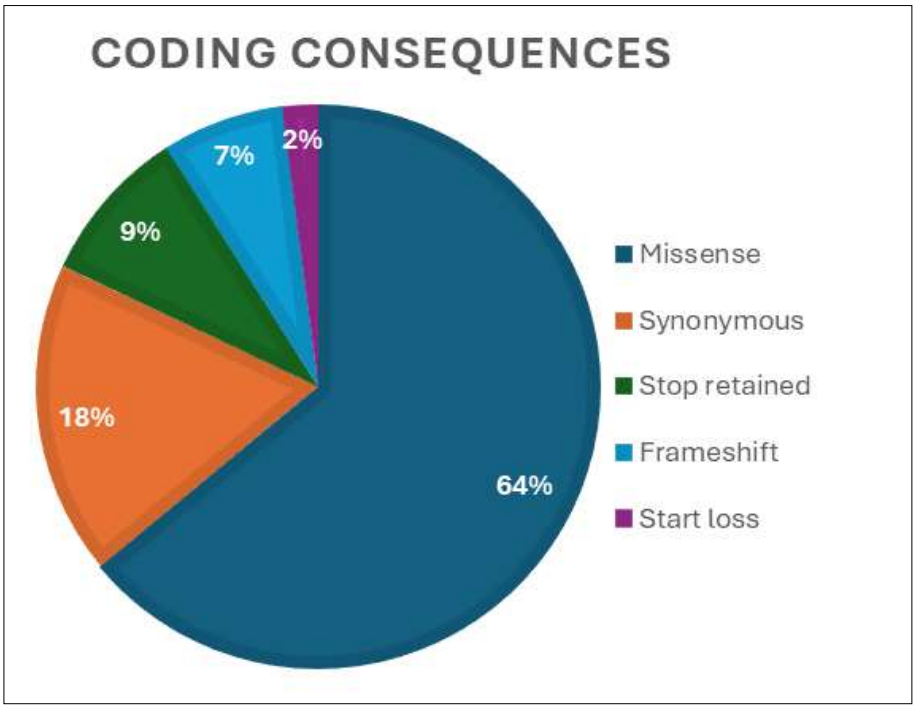

Fig 2: Details of Variants in coding regions of sex chromosome of Pigs

3.1 Identification of deleterious nsSNPs in sex chromosome of Pig

Many research efforts have looked at potential genes, including RASSF5, MECP2, CSN3, BRAF, TNF-κ, BARD-1, IGF1R, and SLC11A2, in both humans and animals in order to find harmful SNPs [12-16]. Studies have been carried out in porcine TERT gene [17] and MYH1 gene [18] using *in-silico* tools.

About 30% of false negative outcomes and 20% of false positive results are provided by the SIFT (9). To increase the accuracy of prediction by SIFT, we used other *in silico* tools. Although all *in silico* programs have certain limitations, it was imperative to use the combination of the tools to assess the deleteriousness. Reports indicate that these tools may also have a considerable error rate (19). In this study, PANTHER, Predict SNP, Polyphen-1, Polyphen-2 and SNAP were used to analyse identified missense SNPS for deleterious effects.

A total of 28 non-synonymous variants were identified on sex-chromosome (X-chromosome) of pig. Based on SIFT score obtained from VEP, 4 nsSNPs named S261F, D664N, A704T and I760N were found to be deleterious. These four nsSNPs were further analysed and three were found deleterious using PANTHER, one using Predict SNP, four using Polyphen-1, three using Polyphen-2, and one using SNAP for further confirmation. Only one nsSNP was found deleterious in common using all tools (Table 2).

Table 2: Common deleterious nsSNPs in sex chromosome of Pig

Position	Gene symbol	Variant name	PANTHER	Predict SNP	Polyphen-1	Polyphen-2	SNAP
X:10435436	-	S261F	Probably damaging	-	Deleterious	Deleterious	Neutral
X:46366236	HUWE1	D664N	Probably damaging	Deleterious	Deleterious	Deleterious	**Deleterious**
X:98011469	NKRF	A704T	possibly damaging	Neutral	Deleterious	-	Neutral
X:100773108	-	I760N	-	Neutral	Deleterious	Deleterious	Neutral

3.2 Structural and functional effects on protein

The I-mutant 2.0 tool showed a decrease in protein stability in three deleterious nsSNPs, thus adding another layer of confirmation. Furthermore, the one common SNP (D664N) from all the tools was analysed using HOPE server.

Differences in the size, affinity towards water, flexibility and preferred secondary structure have been observed between wild and mutant residues which also led to structural and functional effects in the new mutant protein. Project HOPE identified the deleterious effect of point mutation on protein structure. This analysis revealed the effect of D664N (HUWE1) mutation; that is the mutation is damaging to the protein. A difference in charge between the wild-type and mutant amino acid (Table 3) was also identified. The net charge of amino acid sequence has a dynamic effect on interaction between other proteins as well as cell membrane phospholipids [20]. surface charge can also influence protein localisation based on the motifs present on it [21]. A change in the net charge due to the SNP may deleteriously affect the normal functioning of the coded protein

HUWE1 is a ubiquitin ligase-containing HECT domain linked to the spermatogenesis, oncogenesis and neurogenesis. Mouse models have demonstrated the critical role HUWE1 gene plays in the N-Myc pathway-mediated regulation of neurogenesis in the cerebral cortex. It is now widely acknowledged that X-linked intellectual impairment (XLID) is connected to HUWE1 mutations or rearrangements [22]. Genetic variants of HUWE1 gene were found to cause severe intellectual disabilities, disfigured

limbs, absence of speech, dysmorphic features and craniosynostosis in several human patients [23-27]. Mouse embryo study shows that HUWE1 knockdown could lead to increased apoptosis in blastocysts [28]. Thus, a deleterious SNP may lead to cognitive as well as reproductive impairment of the affected pigs.

Table 3: Change in protein stability due to mutation

Position	Allele	Gene symbol	Variant name	ddg	I Mutant 2.0
X:10435436	A	-	S261F	0.21	Increase Stabilit
X:46366236	T	HUWE1	D664N	-1.1	Decrease Stability
X:98011469	T	NKRF	A704T	-1.1	Decrease Stability
X:100773108	A	-	I760N	-0.21	Decrease Stability

4. Conclusion

This study highlights the utility of computational algorithms and bioinformatics pipelines in identifying deleterious mutations, particularly in scenarios lacking extensive animal records. The identified deleterious SNPs in pigs, mapped to genes HUWE1 and NKRF, provide valuable insights into potential genetic risks. Further research, particularly in Indian native pig breeds, is warranted to elucidate the biological consequences of these mutations through *in vivo* studies.

5. Acknowledgements

The authors acknowledge the support of Co-authers and Institute.

6. Conflict of Interest

The authors declare no conflict of interest.

7. References

1. Barroso I, Gurnell M, Crowley VE, Agostini M, Schwabe JW, Soos MA, *et al*. Dominant negative mutations in human PPARγ associated with severe insulin resistance, diabetes mellitus and hypertension. Nature. 1999 Dec;402(6764):880-883.
2. Chasman D, Adams RM. Predicting the functional consequences of non-synonymous single nucleotide polymorphisms: structure-based assessment of amino acid variation. Journal of Molecular Biology. 2001 Mar 23;307(2):683-706.
3. Nailwal M, Chauhan JB. Analysis of consequences of non-synonymous SNPs of USP9Y gene in human using bioinformatics tools. Meta Gene. 2017;12:13-17. doi: 10.1016/j.mgene.2016.12.011
4. Alshatwi AA, Hasan TN, Syed NA, Shafi G, Grace BL. Identification of functional SNPs in BARD1 gene and *in silico* analysis of damaging SNPs: based on data procured from dbSNP database. PLoS ONE. 2012;7(10):e43939. doi: 10.1371/journal.pone.0043939
5. Dabhi B, Mistry KN. *In silico* analysis of single nucleotide polymorphism (SNP) in human TNF-α gene. Meta Gene. 2014;2:586-595. doi: 10.1016/j.mgene.2014.07.005
6. Hussain MR, Shaik NA, Al-Aama JY, Asfour HZ, Khan FS, Masoodi TA, *et al*. *In silico* analysis of Single Nucleotide Polymorphisms (SNPs) in human BRAF gene. Gene. 2012 Oct 25;508(2):188-96.
7. Patel SM, Koringa PG, Reddy BB, Nathani NM, Joshi CG. *In silico* analysis of consequences of non-synonymous SNPs of Slc11a2 gene in Indian bovines. Genomics Data. 2015;5:72-79. doi: 10.1016/j.gdata.2015.05.015
8. Ng PC, Henikoff S. Predicting deleterious amino acid substitutions. Genome Research. 2001 May 1;11(5):863-74.
9. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Research. 2003 Jul 1;31(13):3812-3814.
10. Capriotti E, Fariselli P, Casadio R. I-Mutant2. 0: predicting stability changes upon mutation from the protein sequence or structure. Nucleic Acids Research. 2005 Jul 1;33(suppl_2):W306-10.
11. Venselaar H, Te Beek TA, Kuipers RK, Hekkelman ML, Vriend G. Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. BMC Bioinformatics. 2010 Dec;11(1):1-0.
12. Hossain MS, Roy AS, Islam MS. *In silico* analysis predicting effects of deleterious SNPs of human RASSF5 gene on its structure and functions. Sci Rep. 2020;10(1):1-14. doi: 10.1038/s41598-020-71457-1
13. Desai M, Chauhan J. *In silico* analysis of nsSNPs in human methyl CpG binding protein 2. Meta Gene. 2016;10:1-7. doi: 10.1016/j.mgene.2016.09.004
14. Patel JB, Chauhan JB. Computational analysis of non-synonymous single nucleotide polymorphism in the bovine cattle kappa-casein (CSN3) gene. Meta Gene. 2018;15:1-9. doi: 10.1016/j.mgene.2017.10.002
15. Hussain MR, Shaik NA, Al-Aama JY, Asfour HZ, Khan FS, Masoodi TA, *et al*. *In silico* analysis of Single Nucleotide Polymorphisms (SNPs) in human BRAF gene. Gene. 2012 Oct 25;508(2):188-96.
16. Dabhi B, Mistry KN. *In silico* analysis of single nucleotide polymorphism (SNP) in human TNF-α gene. Meta Gene. 2014;2:586-595. doi: 10.1016/j.mgene.2014.07.005
17. Peka M, Balatsky V, Saienko A, Tsereniuk O. Bioinformatic analysis of the effect of SNPs in the pig TERT gene on the structural and functional characteristics of the enzyme to develop new genetic markers of productivity traits. BMC Genomics. 2023 Aug 25;24(1):487.
18. Ghosh M, Sodhi SS, Sharma N, Mongre RK, Kim N, Singh AK, *et al*. An integrated *in silico* approach for functional and structural impact of non-synonymous SNPs in the MYH1 gene in Jeju Native Pigs. BMC Genetics. 2016 Dec;17:1-6.
19. Ng PC, Henikoff S. Predicting the effects of amino acid substitutions on protein function. Annual Review of Genomics and Human Genetics. 2006;7:61-80. doi: 10.1146/annurev.genom.7.080505.115630
20. Silvius JR, Bhagatji P, Leventis R, Terrone D. K-ras4B and prenylated proteins lacking "second signals" associate dynamically with cellular membranes. Molecular Biology of the Cell. 2006 Jan;17(1):192-202.
21. Heo WD, Inoue T, Park WS, Kim ML, Park BO, *et al*. PI (3, 4, 5) P3 and PI (4, 5) P2 lipids target proteins with polybasic clusters to the plasma membrane. Science. 2006 Dec 1;314(5804):1458-1461.

22. Froyen G, Corbett M, Vandewalle J, Jarvela I, Lawrence O, *et al*. Submicroscopic duplications of the hydroxysteroid dehydrogenase HSD17B10 and the E3 ubiquitin ligase HUWE1 are associated with mental retardation. The American Journal of Human Genetics. 2008 Feb 8;82(2):432-443.
23. Nava C, Lamari F, Héron D, Mignot C, Rastetter A, Keren B, *et al*. Analysis of the chromosome X exome in patients with autism spectrum disorders identified novel candidate genes, including TMLHE. Translational Psychiatry. 2012 Oct;2(10):e179-.
24. Hu H, Haas SA, Chelly J, Van Esch H, Raynaud M, de Brouwer AP, *et al*. X-exome sequencing of 405 unresolved families identifies seven novel intellectual disability genes. Molecular Psychiatry. 2016 Jan;21(1):133-148.
25. Taylor JC, Martin HC, Lise S, Broxholme J, Cazier JB, Rimmer A, *et al*. Factors influencing success of clinical genome sequencing across a broad spectrum of disorders. Nature Genetics. 2015 Jul;47(7):717.
26. Brooks SS, Wisniewski K, Brown WT. New X-linked mental retardation (XLMR) syndrome with distinct facial appearance and growth retardation. American Journal of Medical Genetics. 1994;51:586-590.
27. Friez MJ, Brooks SS, Stevenson RE, Field M, Basehore MJ, Adès LC, *et al*. HUWE1 mutations in Juberg-Marsidi and Brooks syndromes: the results of an X-chromosome exome sequencing study. BMJ Open. 2016 Apr 1;6(4):e009537.
28. Chen LJ, Xu WM, Yang M, Wang K, Chen Y, Huang XJ, *et al*. HUWE1 plays important role in mouse preimplantation embryo development and the dysregulation is associated with poor embryo development in humans. Scientific Reports. 2016 Nov 30;6(1):37928.