

Deleterious mutation load in diverse cattle breeds of the world

Sankar Subramanian (✉ ssankara@usc.edu.au)

University of the Sunshine Coast <https://orcid.org/0000-0002-2375-3254>

Research Article

Keywords: Mutation load, Deleterious SNVs, cattle breeds, inbreeding, population bottleneck, small populations

Posted Date: February 22nd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-216864/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Deleterious mutation load in diverse cattle breeds of the world

Sankar Subramanian

*GeneCology Research Centre, School of Science and Engineering, The University of the
Sunshine Coast, Moreton Bay, QLD 4502, Australia*

Running head: Deleterious mutations in diverse cattle breeds

Keywords: Mutation load, Deleterious SNVs, cattle breeds, inbreeding, population bottleneck,
small populations

Address for correspondence:

GeneCology Research Centre
School of Science and Engineering
University of the Sunshine Coast
1 Moreton Parade, Petrie QLD 4502
Australia
Phone: + 61-7-5430 2873
Fax: +61-7- 5430 2881
E-mail: ssankara@usc.edu.au

Abstract

Domestication of wild animals results in a reduction in the effective population size and this could affect the deleterious mutation load of domesticated breeds. Furthermore, artificial selection will also contribute to accumulation deleterious mutations due to the increased rate of inbreeding among these animals. The process of domestication, founder population size, and artificial selection differ between cattle breeds, which could lead to a variation in their deleterious mutation loads. We investigated this using the whole genome data from 432 animals belonging to 54 cattle breeds of the world. Our analysis revealed a negative correlation between the genomic heterozygosity and the ratio of amino acid changing diversity to silent diversity. This suggests a proportionally higher amino acid changing Single Nucleotide variants (SNVs) in breeds with low diversity. Our results also showed that breeds with low diversity had more high-frequency ($DAF > 0.51$) deleterious SNVs than those with high diversity. A reverse trend was observed for the low-frequency ($DAF \leq 0.51$) deleterious SNVs. Overall, taurine cattle breeds had more high-frequency deleterious SNVs than indicine (or taurine-indicine hybrid) breeds. However, within taurine breeds European or Northeast Asian taurines had more high-frequency deleterious SNVs than East Asian or African taurine breeds. Similarly, within indicine breeds South Asian indicines had more high-frequency deleterious SNVs than East Asian indicine breeds. All the above observed patterns were reversed for low frequency deleterious SNVs. Some of the variation in the deleterious mutation load observed between different breeds could be attributed to the population sizes of the wild progenitors before domestication. However, the variations observed within taurine and within indicine breeds could be due to the difference in the extent of inbreeding, strength of artificial selection and/or founding population size. The findings of this study imply that the rate of incidence of genetic diseases might vary between cattle breeds.

Background

Theories of population genetics predict that at low frequencies deleterious Single Nucleotide Variants (SNVs) can contribute significantly to the heterozygosity. In contrast, they are prevented from reaching to high frequencies and are eventually eliminated by purifying selection [1]. Domestication of wild plants and animals result in a population bottleneck, because only a small subset of the wild population is typically sampled to form the founder stock [2]. Artificial selection for desired traits and inbreeding also lead to further reduction in the effective population sizes during breed formation [3]. Due to these reasons, domesticated plants and animals are expected to accumulate an excess of deleterious mutations compared to their wild types. A number of previous studies investigated this by comparing the deleterious mutational loads of wild and domesticated plants and animals. The ratio (ω) of diversities of amino acid changing (nonsynonymous) and silent (or synonymous) SNVs was used as the measure of deleterious mutational load. In a previous study ω was found to be much higher in domesticated pig breeds compared to wild pigs [4]. The ω estimated for commercial white layer chickens was found to be much higher than that observed for wild African village chickens (putatively close to jungle fowl) [4]. Similarly, much higher ω were observed for domesticated breeds of horse [5], dog [6], rabbit [7] and silkworm [7] than their wild relatives. Studies on plants were also showed higher ω for cultivated crops such as rice [2, 8], soybean [9], cassava [10], and sunflower [11] compared to their wild progenitors

Theories also predict that when population size declines harmful mutations are elevated to high-frequencies by genetic drift [12]. For instance, human migration out-of-Africa resulted in a series of bottlenecks in non-African populations as they were successively subsampled along the migratory route. This resulted in a higher proportion of high-frequency and homozygous deleterious SNVs in non-Africans compared to Africans [13, 14]. Since the

process of domestication also introduces bottlenecks, a similar pattern is expected in domesticated animals. This was confirmed by a previous study that compared the exomes of dog breeds and wild wolves and found a higher proportion of homozygous deleterious SNVs in dogs than wolves [6]. Similarly, domesticated yak populations were reported to have higher number of homozygous deleterious amino acid changing SNVs than those estimated for wild yaks [15]. High frequency deleterious variants causing diseases such as retinal degeneration in European cattle breeds has been attributed to the process of domestication and artificial selection [16].

Global cattle breeds predominantly derive from *Bos taurus*, *Bos indicus* or from their hybrid *Bos taurus* × *Bos indicus* [17]. Therefore, the heterozygosity as well as the deleterious mutational load in different breeds could also depends on the ancestral source species/population [18]. In addition, the differences in the degree and patterns of artificial selection and the rate of inbreeding could also contribute to the deleterious mutation load of different cattle breeds [3]. The present study estimated the deleterious mutational load in various cattle breeds and investigated the potential contributions of the above-mentioned factors by analysing the whole genome data from 432 animals belonging to 54 distinct breeds of the world.

Methods

Genome data

Whole genome data from 108 cows and 314 bull was obtained from the Bovine Genome Variation Database [19]. These animals belong to 54 breeds including those from Europe (Central and Western Europe), Northeast Asia (Japan and South Korea), Africa (Western Africa and Guinea), Middle East (Iran), South Asia (India, Pakistan, and Sri Lanka) and East

Asia (China). The number of individuals in each breed varies between 1 to 45 (for details, see Supplementary Table S1). To orient the direction of mutations and to find the derived SNVs, the whole genome data of American Bison (*Bison Bison*) was used. For this purpose, the whole genome LASTZ alignment of Cow and American Bison pair was downloaded from the Ensembl genome data resource (ftp://ftp.ensembl.org/pub/release-102/maf/ensembl-compara/pairwise_alignments/btau_ars-ucd1.2.v.bbci_bison_umd1.0.lastz_net.tar.gz). Using the chromosomal positional coordinates, the corresponding Bison nucleotide was obtained for each SNVs and using this, the orientation of mutational change of the SNVs were determined.

Functional annotations

To identify amino acid changing (nonsynonymous) SNVs and silent (synonymous) SNVs the genome annotation file containing the functional consequences information was obtained from the Ensembl server (ftp://ftp.ensembl.org/pub/release-102/variation/vcf/bos_taurus/bos_taurus_incl_consequences.vcf.gz). To determine the number of synonymous and nonsynonymous sites (or positions) the CDS sequences for 15,392 unique reference genes of cow were downloaded from the GenBank reference genes database. Using the program codeml of the software PAML [20] the exact numbers of synonymous and nonsynonymous positions were calculated. In order to identify the deleterious amino acid changing SNVs the Genome Evolutionary Rate Profiling (GERP) score was used [21]. The GERP score for each chromosomal position of cow was calculated based on a whole genome alignment consisting of 111 mammalian genomes and this data was downloaded from the Ensembl server (ftp://ftp.ensembl.org/pub/release-102/bed/ensembl-compara/111_mammals.gerp_constrained_element/). Amino acid changing SNVs with a GERP score >2.0 were designated as deleterious nonsynonymous SNVs. While other higher

thresholds reduced the number of SNVs available for the analysis, use of them produced qualitatively similar results.

Data analysis

Nucleotide diversity (π) per base was estimated using the following equations [22]

$$\theta_T = \frac{n}{n-1} \sum_{i=1}^S 2p_i(1-p_i)$$

$$\pi = \frac{\theta_T}{L}$$

where p_i is the allele frequency of SNV i , S is the total number of SNVs in the whole genome or exome, n is the number of chromosomes sampled and L is the number of sites or bases in the genome, synonymous or nonsynonymous positions. The ratio (ω) was estimated as

$$\omega = \frac{\pi_N}{\pi_S}$$

where π_N and π_S are nonsynonymous and synonymous nucleotide diversities. Only biallelic SNVs are used in the analysis. To test the significance between mean estimates the Z-test was used and the Spearman rank correlation was used to estimate the strength of relationships.

Results

To examine the pattern of genomic variation the whole genome nucleotide diversity was estimated for 54 breeds. The X-axis in Figure 1 shows that the taurine breeds typically had low diversities (0.0004-0.0016) compared to the taurine \times indicine hybrids (0.0019-0.0028) as well as indicine (0.0020-0.0031) breeds. However, within taurine breeds there was a fourfold variation in the genome diversity and this difference was ~70% within indicine or hybrid

derived breeds. The ratio of nonsynonymous-to-synonymous diversities (ω) was then estimated for the breeds and the values were plotted against the genomic diversities. A highly significant negative correlation (Spearman $\rho = -0.92$, $P < 0.00001$) between the two variables was observed (Figure 1). This suggests that the breeds with low genomic diversity have a high proportion of amino acid changing SNVs compared to those with high diversity. Hence taurine breeds have a much higher ω (0.18 - 0.22) than those observed for the hybrid or indicine (0.16 - 0.17) breeds. However, within taurine breeds the European and Northeast Asian (Japanese and Korean) taurines have a much higher ω (0.20 – 0.22) than African and East Asian taurines (0.18 – 0.19). Hybrid breeds have slightly higher ω than those of indicine breeds (0.17 vs. 0.16).

The ω estimates indirectly revealed the nonsynonymous deleterious mutation loads in various cattle breeds. We then estimated the actual counts of derived deleterious nonsynonymous SNVs. Using a GERP threshold score of >2.0 deleterious nonsynonymous SNVs were determined. The counts of deleterious SNVs were found to vary between 680 to 870 per individual, which were plotted against the whole genome diversity of cattle breeds (Figure 2A). Although this relationship was significant ($\rho = 0.71$, $P < 0.00001$) the maximum difference between the number of deleterious SNVs between the breeds was only 20%.

To investigate the deleterious mutation load based on allele frequencies, the deleterious SNVs were separated into two groups. One group consisted of deleterious SNVs with DAF (Derived Allele Frequency) ≤ 0.51 and the other group comprising those with DAF >0.51 . The correlation of these SNVs with the genomic diversity revealed contrasting patterns. Figure 2B reveals a significant positive correlation (Spearman $\rho = 0.96$, $P < 0.00001$) between the number of low-frequency deleterious SNVs per genome and genomic diversity and a significant

negative trend (Spearman $\rho = -0.92$, $P < 0.00001$) was observed for the high-frequency deleterious SNVs.

Previous analyses revealed the difference in the mutation load among various cattle breeds. To obtain the unique patterns of the load on various breeds, they were grouped into six categories based on their geographical locations and ancestral source populations, as shown in Figure 3. The overall number of deleterious SNVs were similar across the groups. In contrast, the high and low frequency deleterious SNVs show opposite patterns. On an average taurine breeds have predominantly high-frequency deleterious SNVs and indicine (and the hybrid) breeds have the preponderance of low-frequency ones. Within taurines, Northeast Asian and European breeds have the significantly higher number of high-frequency deleterious SNVs than the East Asian breeds ($P < 0.00001$, Z test). In contrast, Northeast Asian and European breeds have significantly lower number of low-frequency deleterious SNVs than those of the East Asian ones ($P < 0.002$). On the other hand, South Asian indicine breeds have significantly higher counts of high-frequency deleterious SNVs than their East Asian counterparts ($P < 0.00001$) and this pattern was reversed for the low-frequency deleterious SNVs ($P = 0.00011$).

Discussion

The whole genome diversity estimated in this study varied significantly between the cattle breeds but are very similar those reported by a previous study [23]. The correlation between diversity and ω suggests a higher nonsynonymous mutation load in breeds with low diversity than those with high diversity. This result is similar to the correlations observed between diversity and ω estimated for various dog breeds [6], and domestic breeds of rabbit, pig, chicken [7]. Furthermore, the higher ω observed for many domestic crop varieties and animal breeds compared to their wild relatives also support our results [2, 4-11, 15, 24]. This is

because those studies showed that the diversity of the domesticated breeds or varieties were always smaller than those of their wild relatives. Analysis based on the actual number of deleterious nonsynonymous SNVs revealed a higher number of high-frequency deleterious SNVs in breeds with low diversity [6, 15]. This is due to the role of genetic drift that elevates the frequency of SNVs due to bottleneck, inbreeding and artificial selection. Previous studies on dog breeds and yaks showed a higher number of homozygous deleterious SNVs in domesticated canines or yaks than those present in their respective wild relatives [6, 15]. Since homozygous SNVs represents high frequency variants, our results are in consensus with those reported by the above-mentioned studies.

The deleterious SNVs estimated for groups of breeds showed on an average taurine cattle breeds have a higher mutational load and more high-frequency deleterious SNVs than indicine breeds. This could be due to the difference in the effective population sizes of their progenitors before domestication as suggested previously [18, 25]. Alternatively, this could also be due to the difference in the severity of the bottleneck occurred during their respective domestication. The mutation loads and the number of high-frequency deleterious SNVs also significantly vary within taurine and within indicine breeds. For instance, Northeast Asian taurine breeds have higher number of high-frequency deleterious SNVs, and lower number of low-frequency deleterious SNVs compared to those of East Asian taurines. Similarly, South Asian indicine breeds have the higher number of high-frequency deleterious SNVs and lower number of low-frequency harmful SNVs than East Asian indicines. These differences could be attributed to the practices during breed formation [3]. For example, the strength of artificial selection, extent of inbreeding and the number of founders selected could vary significantly among breeds.

Conclusion

Our study revealed a higher mutation load in cattle breeds with low genomic diversity compared to those with high diversity. The results also showed a higher number of high-frequency deleterious SNVs in the former than the latter. These results suggest that diversity, deleterious mutation load and the frequency of deleterious mutations are determined by their effective population sizes as predicted by population genetic theories. While we found higher mutational load in taurine breeds compared to indicines or hybrids, the loads did vary within taurine breeds owing to the difference in their population sizes. These results have implications in understanding the health of different cattle breeds as the mutations causing genetic diseases and their frequencies are expected to vary between them. For instance, rate of incidence of genetic diseases caused by recessive homozygous variations could potentially be higher in breeds that have small effective sizes.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12864-019-0588-8>.

Acknowledgements

This research was funded by the University of the Sunshine Coast (DVC-R&I grant).

Author's Contributions

Conceptualization, S.S.; methodology, S.S.; formal analysis, S.S.; writing, S.S.; funding acquisition, S.S. All authors have read and agreed to the published version of the manuscript.

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

All relevant results are within this paper and its additional file. The whole genome data used in this study is available at: <http://animal.nwsuaf.edu.cn/code/index.php/BosVar>

Competing interests

The authors declare that they have no competing interests.

Funding

This research was funded by the University of the Sunshine Coast (DVC-R&I grant).

List of abbreviations

SNV – Single Nucleotide Variant

DAF – Derived Allele Frequency

GERP - Genome Evolutionary Rate Profiling

LASTZ – Large-Scale genome alignment Tool

Figure Legends

Figure 1. Relationship between whole genome diversity (π) and the ratio of nonsynonymous heterozygosity to synonymous heterozygosity (ω) estimated for 54 cattle breeds. Colour codes of the data points distinguish geographical locations of the breeds and shapes denote the species/subspecies of the breeds. The correlation was highly significant (Spearman $\rho = -0.92$, $P < 0.00001$). Best fitting regression line is shown.

Figure 2. Correlation between whole genome diversity (π) and the number of derived deleterious nonsynonymous SNVs. **(A)** All derived deleterious SNVs. The relationship was significant ($\rho = 0.71$, $P < 0.00001$). **(B)** Deleterious SNVs divided into two groups based on their derived allele frequencies (DAF). The positive ($\rho = 0.96$, $P < 0.00001$) and negative ($\rho = -0.92$, $P < 0.00001$) correlations were highly significant and best fitting regression lines are shown.

Figure 3. Average number of deleterious derived nonsynonymous SNVs in cattle breeds that were grouped based on their geographical location and progenitor source population. **(A)** All deleterious SNVs **(B)** High-frequency SNVs (DAF > 0.51) **(C)** Low-frequency SNVs (DAF \leq 0.51). Error bars denotes standard error of the mean.

References

- [1] Kimura M., *The neutral theory of molecular evolution*, Oxford University Press, Oxford, 1983.
- [2] Lu J., Tang T., Tang H., Huang J., Shi S., Wu C.I., *The accumulation of deleterious mutations in rice genomes: a hypothesis on the cost of domestication*, *Trends Genet.* 22 (2006) 126-131.
- [3] Frantz L.A.F., Bradley D.G., Larson G., Orlando L., *Animal domestication in the era of ancient genomics*, *Nat Rev Genet.* 21 (2020) 449-460.
- [4] Bosse M., Megens H.J., Derks M.F.L., de Cara A.M.R., Groenen M.A.M., *Deleterious alleles in the context of domestication, inbreeding, and selection*, *Evol Appl.* 12 (2019) 6-17.
- [5] Schubert M., Jonsson H., Chang D., Der Sarkissian C., Ermini L., Ginolhac A., Albrechtsen A., Dupanloup I., Foucal A., Petersen B., Fumagalli M., Raghavan M., Seguin-Orlando A., Korneliussen T.S., Velazquez A.M., Stenderup J., Hoover C.A., Rubin C.J., Alfarhan A.H., Alquraishi S.A., Al-Rasheid K.A., MacHugh D.E., Kalbfleisch T., MacLeod J.N., Rubin E.M., Sicheritz-Ponten T., Andersson L., Hofreiter M., Marques-Bonet T., Gilbert M.T., Nielsen R., Excoffier L., Willerslev E., Shapiro B., Orlando L., *Prehistoric genomes reveal the genetic foundation and cost of horse domestication*, *Proc Natl Acad Sci U S A.* 111 (2014) E5661-5669.
- [6] Marsden C.D., Ortega-Del Vecchyo D., O'Brien D.P., Taylor J.F., Ramirez O., Vila C., Marques-Bonet T., Schnabel R.D., Wayne R.K., Lohmueller K.E., *Bottlenecks and selective sweeps during domestication have increased deleterious genetic variation in dogs*, *Proc Natl Acad Sci U S A.* 113 (2016) 152-157.
- [7] Makino T., Rubin C.J., Carneiro M., Axelsson E., Andersson L., Webster M.T., *Elevated Proportions of Deleterious Genetic Variation in Domestic Animals and Plants*, *Genome Biol Evol.* 10 (2018) 276-290.

- [8] Liu Q., Zhou Y., Morrell P.L., Gaut B.S., Deleterious Variants in Asian Rice and the Potential Cost of Domestication, *Mol Biol Evol.* 34 (2017) 908-924.
- [9] Kono T.J., Fu F., Mohammadi M., Hoffman P.J., Liu C., Stupar R.M., Smith K.P., Tiffin P., Fay J.C., Morrell P.L., The Role of Deleterious Substitutions in Crop Genomes, *Mol Biol Evol.* 33 (2016) 2307-2317.
- [10] Ramu P., Esuma W., Kawuki R., Rabbi I.Y., Egesi C., Bredeson J.V., Bart R.S., Verma J., Buckler E.S., Lu F., Cassava haplotype map highlights fixation of deleterious mutations during clonal propagation, *Nat Genet.* 49 (2017) 959-963.
- [11] Renaut S., Rieseberg L.H., The Accumulation of Deleterious Mutations as a Consequence of Domestication and Improvement in Sunflowers and Other Compositae Crops, *Mol Biol Evol.* 32 (2015) 2273-2283.
- [12] Crow J.K., Kimura M., *An Introduction to Population Genetics Theory*, Harper and Row, New York City, 1970.
- [13] Henn B.M., Botigue L.R., Peischl S., Dupanloup I., Lipatov M., Maples B.K., Martin A.R., Musharoff S., Cann H., Snyder M.P., Excoffier L., Kidd J.M., Bustamante C.D., Distance from sub-Saharan Africa predicts mutational load in diverse human genomes, *Proc Natl Acad Sci U S A.* 113 (2016) E440-449.
- [14] Subramanian S., Europeans have a higher proportion of highfrequency deleterious variants than Africans, *Hum Genet.* 135 (2016) 1-7.
- [15] Xie X., Yang Y., Ren Q., Ding X., Bao P., Yan B., Yan X., Han J., Yan P., Qiu Q., Accumulation of deleterious mutations in the domestic yak genome, *Anim Genet.* 49 (2018) 384-392.
- [16] Michot P., Chahory S., Marete A., Grohs C., Dagios D., Donzel E., Aboukadiri A., Deloche M.C., Allais-Bonnet A., Chambrial M., Barbey S., Genestout L., Boussaha M., Danchin-Burge C., Fritz S., Boichard D., Capitan A., A reverse genetic approach identifies an

ancestral frameshift mutation in RP1 causing recessive progressive retinal degeneration in European cattle breeds, *Genet Sel Evol.* 48 (2016) 56.

[17] Pitt D., Sevane N., Nicolazzi E.L., MacHugh D.E., Park S.D.E., Colli L., Martinez R., Bruford M.W., Orozco-terWengel P., Domestication of cattle: Two or three events?, *Evol Appl.* 12 (2019) 123-136.

[18] Bovine HapMap C., Gibbs R.A., Taylor J.F., Van Tassell C.P., Barendse W., Eversole K.A., Gill C.A., Green R.D., Hamernik D.L., Kappes S.M., Lien S., Matukumalli L.K., McEwan J.C., Nazareth L.V., Schnabel R.D., Weinstock G.M., Wheeler D.A., Ajmone-Marsan P., Boettcher P.J., Caetano A.R., Garcia J.F., Hanotte O., Mariani P., Skow L.C., Sonstegard T.S., Williams J.L., Diallo B., Hailemariam L., Martinez M.L., Morris C.A., Silva L.O., Spelman R.J., Mulatu W., Zhao K., Abbey C.A., Agaba M., Araujo F.R., Bunch R.J., Burton J., Gorni C., Olivier H., Harrison B.E., Luff B., Machado M.A., Mwakaya J., Plastow G., Sim W., Smith T., Thomas M.B., Valentini A., Williams P., Womack J., Woolliams J.A., Liu Y., Qin X., Worley K.C., Gao C., Jiang H., Moore S.S., Ren Y., Song X.Z., Bustamante C.D., Hernandez R.D., Muzny D.M., Patil S., San Lucas A., Fu Q., Kent M.P., Vega R., Matukumalli A., McWilliam S., Sclep G., Bryc K., Choi J., Gao H., Grefenstette J.J., Murdoch B., Stella A., Villa-Angulo R., Wright M., Aerts J., Jann O., Negrini R., Goddard M.E., Hayes B.J., Bradley D.G., Barbosa da Silva M., Lau L.P., Liu G.E., Lynn D.J., Panzitta F., Dodds K.G., Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds, *Science.* 324 (2009) 528-532.

[19] Chen N., Fu W., Zhao J., Shen J., Chen Q., Zheng Z., Chen H., Sonstegard T.S., Lei C., Jiang Y., BGVD: An Integrated Database for Bovine Sequencing Variations and Selective Signatures, *Genomics Proteomics Bioinformatics.* 18 (2020) 186-193.

[20] Yang Z., PAML 4: phylogenetic analysis by maximum likelihood, *Mol Biol Evol.* 24 (2007) 1586-1591.

- [21] Cooper G.M., Stone E.A., Asimenos G., Program N.C.S., Green E.D., Batzoglou S., Sidow A., Distribution and intensity of constraint in mammalian genomic sequence, *Genome Res.* 15 (2005) 901-913.
- [22] Tajima F., Statistical method for testing the neutral mutation hypothesis by DNA polymorphism, *Genetics.* 123 (1989) 585-595.
- [23] Chen N., Cai Y., Chen Q., Li R., Wang K., Huang Y., Hu S., Huang S., Zhang H., Zheng Z., Song W., Ma Z., Ma Y., Dang R., Zhang Z., Xu L., Jia Y., Liu S., Yue X., Deng W., Zhang X., Sun Z., Lan X., Han J., Chen H., Bradley D.G., Jiang Y., Lei C., Whole-genome resequencing reveals world-wide ancestry and adaptive introgression events of domesticated cattle in East Asia, *Nat Commun.* 9 (2018) 2337.
- [24] Koenig D., Jimenez-Gomez J.M., Kimura S., Fulop D., Chitwood D.H., Headland L.R., Kumar R., Covington M.F., Devisetty U.K., Tat A.V., Tohge T., Bolger A., Schneeberger K., Ossowski S., Lanz C., Xiong G., Taylor-Teeple M., Brady S.M., Pauly M., Weigel D., Usadel B., Fernie A.R., Peng J., Sinha N.R., Maloof J.N., Comparative transcriptomics reveals patterns of selection in domesticated and wild tomato, *Proc Natl Acad Sci U S A.* 110 (2013) E2655-2662.
- [25] Bollongino R., Burger J., Powell A., Mashkour M., Vigne J.D., Thomas M.G., Modern taurine cattle descended from small number of near-eastern founders, *Mol Biol Evol.* 29 (2012) 2101-2104.

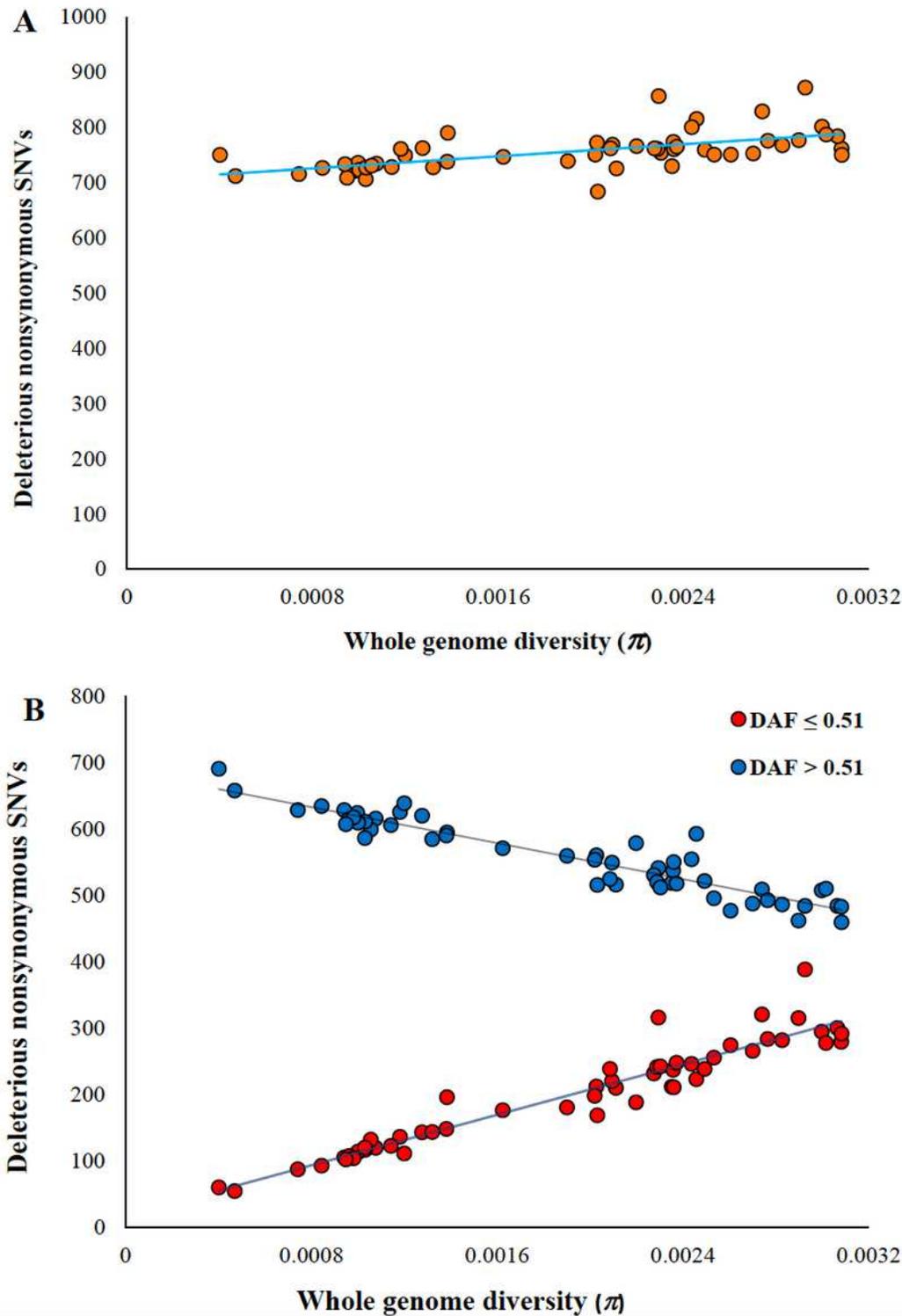


Figure 2

Correlation between whole genome diversity (π) and the number of derived deleterious nonsynonymous SNVs. (A) All derived deleterious SNVs. The relationship was significant ($\rho = 0.71$, $P < 0.00001$). (B) Deleterious SNVs divided into two groups based on their derived allele frequencies (DAF). The positive ($\rho = 0.96$, $P < 0.00001$) and negative ($\rho = -0.92$, $P < 0.00001$) correlations were highly significant and best fitting regression lines are shown.

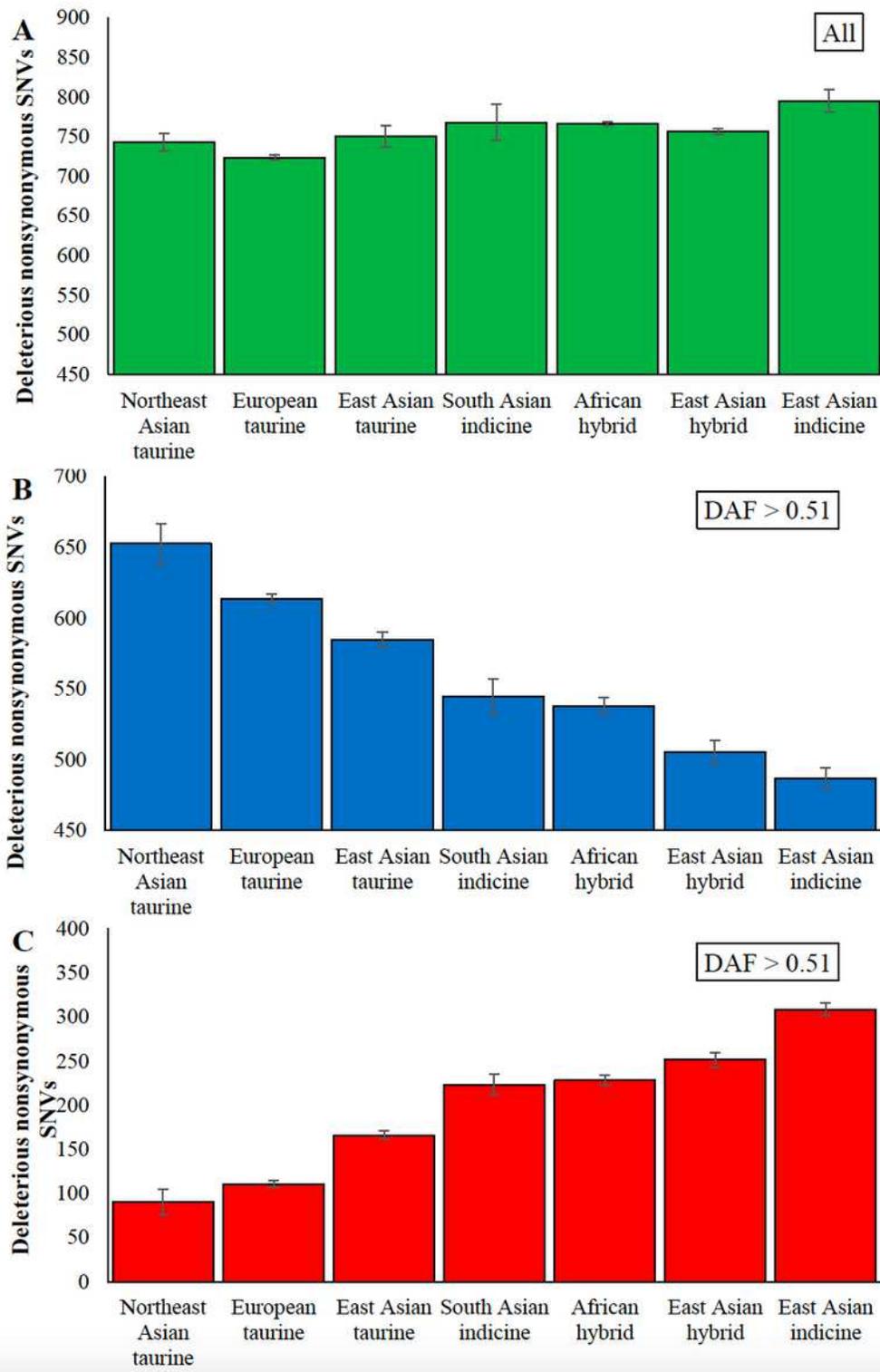


Figure 3

Average number of deleterious derived nonsynonymous SNVs in cattle breeds that were grouped based on their geographical location and progenitor source population. (A) All deleterious SNVs (B) High-frequency SNVs (DAF > 0.51) (C) Low-frequency SNVs (DAF ≤ 0.51). Error bars denotes standard error of the mean.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementaryinformation.pdf](#)