Acta fytotechn zootechn, 25, 2022(2): 117–121
http://www.acta.fapz.uniag.sk

Review

Future perspective of NGS data for evaluation of population genetic structure in Turkish cattle

Eymen Demir[1, 2]*, Nina Moravčíková[1], Taki Karsli[2, 3], Radovan Kasarda[1]

[1]*Slovak University of Agriculture in Nitra, Faculty of Agrobiology and Food Resources, Institute of Nutrition and Genomics, Slovak Republic*
[2]*Akdeniz University, Faculty of Agriculture, Department of Animal Science, Republic of Turkey*
[3]*Eskisehir Osmangazi University, Faculty of Agriculture, Department of Animal Science, Eskisehir, Republic of Turkey*

Article Details: Received: 2022-03-07 | Accepted: 2022-04-25 | Available online: 2022-06-30

https://doi.org/10.15414/afz.2022.25.02.117-121

Developments in sequencing and SNP chip technologies have enabled scientists to obtain high-density genomic data from different livestock species, including cattle. Moreover, many bioinformatics tools are available to analyse high-density genomic data. Via these tools, several statistical approaches such as Principal Component Analysis and clusterin-based analyses could be conducted to reveal the genetic structure of cattle populations. However, revealing the genetic structure and selection signatures of Turkish cattle breeds is a new area of research, since the previous studies are limited with a few microsatellite data. On the other hand, rearing in different geographical and environmental conditions for a long period could possibly lead to more genetic variation in native Turkish cattle breeds compared to high-yielding culture breeds. These variations obviously cannot be detected by limited number of microsatellite markers, while Next Generation Sequencing is promising for further population structure studies. Hence this review aims to summarise previous studies and give a perspective of Next Generation Sequencing possibilities to reveal the population structure of Turkish cattle for further studies.

Keywords: bioinformatics, local cattle, NGS, whole-genome sequencing

---

1 Introduction

Molecular and archeologic evidence support that modern cattle reared worldwide originated from *Bos taurus* and *Bos indicus* which were domesticated in different geographic areas known as Fertile Crescent and Indus Valley, respectively (Pitt et al., 2019). Since domestication, natural and artificial selection led to the formation of modern cattle populations called "breed", which possess differences in morphology, physiology, and behaviour (Orozco-terWengel et al., 2015). Although modern cattle breeds are phenotypically distinct, they may carry similar genetic patterns across the genome due to their common ancestors and crossbreeding history (Decker et al., 2014). Additionally, intensive human-driven selection practices improving milk and beef traits may decrease genetic differentiation between cattle breeds in terms of related genomic regions (Olšanská et al., 2020). These selection practices over generations may result in fixed genomic regions which are called selection signatures (Kukučková et al., 2016; Saravanan et al., 2020). Therefore, in population genetics, it is important to reveal genetic differentiation and admixture level among cattle populations. Genetic structure based on genetic similarity and differentiation may be used for genetic improvement and conservation programs in cattle (Makina et al., 2014).

The genetic structure of cattle breeds has been assessed via different molecular techniques such as microsatellites (Demir and Balcioğlu, 2019), mtDNA (Di Lorenzo et al., 2018) and SNP Arrays (Kukučková et al., 2017). These molecular tools are useful to genetically distinguish taurine cattle breeds from indicine ones (Edea et al., 2015; van der Westhuizen et al., 2020). On the other hand, several studies highlighted that these molecular markers are limited to reveal genetic differentiation between close cattle breeds. For example, 21 microsatellite loci clustered a total of 508 individuals belonging to 11 different Indian

---

***Corresponding Author:** Eymen Demir, Akdeniz University, Faculty of Agriculture, Department of Animal Science, 07058, Antalya, Republic of Turkey, e-mail: eymendemir@akdeniz.edu.tr
ORCID: https://orcid.org/0000-0003-4507-7426

cattle breeds into six groups (Sharma et al., 2015). Similar results, in which cluster numbers were lower than studied number of breeds, were reported in sheep (Alarslan et al., 2021), goats (Nguluma et al., 2018) and chickens (Roh et al., 2020) as well due to high levels of genetic admixture. Besides, Maudet et al. (2002) reported that accuracy of microsatellite markers to assign individuals to a certain breed could be decreased even to 6.5%. These findings support the idea that genetic differentiation of cattle breeds originating from the same ancestor and common evolution history requires high-density genomic data.

Fortunately, today, Next Generation Sequencing (NGS) technologies provide a large amount of genomic data for population structure analysis. Indeed, genetic differentiation of three rare local cattle breeds (Eastern Finncattle, Western Finncattle and Yakutian) raised in the northern Eurasian region was recently revealed by whole-genome sequencing (Weldenegodguad et al., 2019). Together with advances in sequencing technologies, rapid developments in bioinformatics tools enable scientists to analyse these high-density genomic data.

Population structure, as well as differentiation of cattle breeds, could be estimated by different statistical approaches such as Principal Component Analysis (PCA), Wright`s $F_{ST}$ index (Wright, 1965), Analysis of Molecular Variance (AMOVA) Nei`s genetic distance (Nei, 1972), Linkage Disequilibrium (LD) and clustering approaches such as STRUCTURE and Discriminant Analysis of Principal Components (DAPC). Being a mathematical procedure, PCA benefits from orthogonal transformation to reduce a set of correlated variables into a set of uncorrelated variables. The first (PC1) and second (PC2) principal component explains the highest percentage of the total variance, respectively (Fraga et al., 2016). F_{ST} index is the correlation between gametes within subpopulations proportional to randomly selected gametes in the total population, indicating that it is a parameter of genetic differentiation between breeds across the total variation (Wright, 1965). Although, Wright`s F_{ST} value is still fundamental parameter in population genetics, several analogue approaches such as G_{ST} (Nei, 1973), R_{ST} (Slatkin, 1995) and D (Jost, 2008) were proposed to analyse genetic differentiation among populations. These approaches were comprehensively reviewed elsewhere (Wang, 2012; Ma et al., 2015). AMOVA quantifies the total genetic variance across three different hierarchial levels known as among breeds, among individuals within breeds and within individuals (Excoffier et al., 1992). Nei`s genetic distance is a parameter to compare two populations in which populations coming from the same ancestors show lower genetic distance values (Nei, 1972). Today, based on Nei`s genetic distance values, several phylogenetic trees are drawn to visualise population structure in cattle. LD, caused by mutation, genetic drift and epistatic combinations, is defined as the non-random assortment of alleles at different loci (Qanbari, 2020). It is reported that admixing genetically distinct populations may create association between two unlinked loci with different allele frequencies (Qanbari, 2020). STRUCTURE is a Bayesian method to assign individuals into clusters based on their allele frequencies with variable burn-in period and Markov Chain Monte Carlo (MCMC) iterations (Pritchard et al., 2000). However, as highlighted by Kanduma et al. (2016), STRUCTURE relies on the approach in which there are Hardy-Weinberg and linkage equilibrium between all loci. These assumptions may limit using STRUCTURE approach, especially in the case of local cattle breeds with low effective population size as well as populations with high level of admixture. Alternatively, the DAPC approach was proposed by Jombart et al. (2010) to identify and describe clusters of genetically related individuals. The authors have demonstrated that compared to STRUCTURE analysis, the DAPC approach gives better results by analysing three different data sets, including simulated data, microsatellite data of worldwide human population and the data related to hemagglutinin gene sequence variation in seasonal influenza.

Recently, the population structure of native Turkish cattle breeds has been revealed by microsatellite loci. No clear genetic differentiation was detected among studied cattle breeds in previous studies due to high level of admixture. Although breeders do not conduct systematic selection practises, it is known that native Turkish cattle breeds are distinctive in terms of phenotypic and morphological traits. The main reason for differences related to phenotype, morphology and adaptive traits among native Turkish cattle breeds, which are adapted to extreme environmental conditions (diseases, forage deficiency, high or low temperature and altitude), has not been clearly addressed by autosomal makers so far. Therefore, further studies with high-density genomic data coming from SNP arrays or NGS platforms are required to obtain better genetic diversity and population structure results. In this regard, this review aims (i) to give information on the current population structure of native Turkish cattle breeds and (ii) to discuss future perspectives of NGS studies in these cattle breeds for further studies.

2 Current status of population structure in Turkish cattle

Turkey holds six native cattle breeds, namely South Anatolian Red (SAR), South Anatolian Yellow (SAY), Anatolian Black (AB), East Anatolian Red (EAR), Zavot (ZAV) and Turkish Grey Steppe (TGS). Showing huge diversity in morphology, physiology and behaviour, these cattle breeds are also well-adapted to the environment

Table 1 Summary of previous microsatellite studies on Turkish cattle

Studied Breed	Loci	AMOVA	FCA	Structure	UPGMA tree	NJ-tree	Reference
AB, EAR and TGS	20	+	+	+	-	-	Demir and Balcioğlu (2019)
SAR, SAY, AB, EAR and TGS	22	+	-	+	+	-	Öner et al. (2019)
All native breeds	20	-	+	+	-	+	Özşensoy et al. (2019)

and local diseases (Yilmaz et al., 2012). However, their population size has been decreasing, since farmers prefer rearing high-yielding culture breeds such as Holstein Friesian. Besides, crossbreeding practices results in genetic dilution in native cattle populations (Demir et al., 2021). A decrease in effective population size has led to awareness among scientists to reveal genetic diversity and population structure of native Turkish cattle breeds to support conservation programs. In 2019, three different studies (Table 1) were published to reveal the genetic diversity and population structure of Turkish cattle (Demir and Balcioğlu, 2019; Öner et al., 2019; Özşensoy et al., 2019). Recently, the genetic diversity and population structure of native Turkish cattle have been comprehensively reviewed by Demir et al. (2021).

Demir and Balcioğlu (2019) reported that native Turkish cattle breeds showed very clear separation from Holstein Friesian via 20 microsatellite loci. On the other hand, microsatellite markers failed to clearly separate all native Turkish cattle breeds (Demir and Balcioğlu, 2019; Öner et al., 2019; Özşensoy et al., 2019). Several explanations such as genetic admixture, sampling strategy, number of loci, number of individuals per breed may be proposed for this failure. It is obvious that there is a high level of genetic admixture between native and culture cattle breeds raised in Turkey, while some breeds (TGS and ZAV) are raised in limited geographic areas, which eliminates admixture possibilities with other native breeds. Here, we proposed that the main reason for the failure to distinguish Turkish cattle breeds clearly is due to probably used molecular technique, since these cattle breeds were screened via limited number of microsatellite markers (20–22 loci). Basically, used microsatellite markers revealed variations in a small part of the genome which does not represent the actual variations across the whole genome. It is believed that further studies are required to screen native Turkish cattle breeds in terms of the genetic structure via high-density genomic data obtained from NGS.

3 Future perspective: Proposal of NGS data for further studies

NGS, allowing for parallel sequencing of massive DNA fragments from different individuals simultaneously, have revolutionized molecular studies including whole-genome genotyping, metagenomics, epigenetics, Genome Wide Association Studies (GWAS), RNA sequencing, de novo assembling of genomes and genome-wide structural variations (Gosh et al., 2018). There are several NGS platforms such as HiSeq, Rocher 454, Illumina GA2, Ion Torrent and Miseq, whereas HiSeq is the most preferred across the world (Gosh et al., 2018). NGS platforms require library preparation in which Double Digest RADseq (ddRADseq) is one of the efficient in terms of time and economic burden (Peterson et al., 2012). DNA library of even non-model species can be prepared by ddRADseq by which total DNA is restricted by two endonuclease combinations and specific adapters are fused followed by PCR amplification (Peterson et al., 2012).

Over the past few years, NGS technologies have facilitated sequencing studies in which the genome of several farm animals including cattle, pig, sheep, and horse have been partially or completely sequenced (Bai et al., 2012). First whole-genome sequencing in *Bos taurus* conducted on Hereford (beef cattle) and Holstein (dairy cattle) has revealed approximately 22.000 genes in the bovine genome (Elsik et al., 2009; Kõks et al., 2013). Being updated via new studies, genomic data of these two breeds have been served as the reference genome for assemly of numerous whole-genome sequencing studies on native cattle breeds across the world.

Compared to SNP arrays which are currently widely used in farm animals, unbiased NGS techniques have several advantages in terms of revealing genetic diversity and population structure in cattle breeds as well as offering new opportunities for Genomic Selection (GS) and detection of variations related to production, reproduction and diseases across the entire genome (Jiang et al., 2014; Mrode et al., 2019; Zhang et al., 2020; Ghafar et al., 2021).

Together with basic genetic diversity parameters, NGS techniques offer new approaches such as Copy Number Variations (CNV), the discovery of selection signatures, novel mutations and rare genetic variants (Brøndum et al., 2014; Jiang et al., 2014; Zhang et al., 2020). Particularly, detection of novel mutations and rare variants increases the accuracy of revealing total genetic diversity across the entire genome which could be integrated to conservation programs of local cattle breeds. Besides, some novel mutations and rare variants may be directly associated with environmental adaptation (Leroy et al., 2016), which allows supporting selection strategies against climate change and inherited diseases. Additionally, NGS technologies provides high-density genomic data for GS in which genetic improvement of economically important traits such as milk and beef production in cattle is faster and more accurate compared to Marker Assisted Selection (MAS). Indeed, MAS is mainly based on selection of several Quantitative Trait Loci (QTL) with limited impact on genetic improvement (van Marle-Köster et al., 2013), whereas NGS provides all genomic information, including genomic regions with both major and minor effects on economically important traits.

On the other hand, no study is available on native Turkish cattle breeds via SNP arrays and NGS technologies. Previous genetic diversity studies were mainly based on microsatellite loci in which native Turkish cattle breeds could not be clearly separated by limited genetic data which covers a small part of the genome. Fortunately, as mentioned above, obtaining high-density genomic data across the entire genome via SNP arrays and NGS technologies is now cheaper and feasible for livestock species, including cattle. In the first step, population structure in native Turkish cattle breeds may be detected by SNP arrays rather than NGS, since application and data manipulation of SNP arrays are easier and faster. Subsequently, native Turkish cattle breeds should be screened via NGS technologies that are more complex, including library preparation, assembly, and SNP calling process. It is believed that NGS data will not only clarify population structure of Turkish cattle but also allow scientists to conduct new approaches to detect genotype combinations related to environmental adaptation, production, reproduction and diseases. Additionally, GS via NGS data will facilitate selection practises in terms of accuracy and genetic gain in the future. As highlighted by Demir et al. (2021), we highly recommend that NGS technologies should be adopted by scientists to succeed these goals in native Turkish cattle breeds in the future.

Acknowledgement

The Scientific and Technological Research Council of Turkey is acknowledged for financial support of stay of Dr. Demir at Slovak University of Agriculture in Nitra (Project number: 1059B142100368). Agency for Research and Development of Slovak Republic is acknowledged for financial support under projects APVV-17-0060 and APVV-20-0161.

References

Alarslan, E. et al. (2021) Genetic identification and characterisation of some Turkish sheep. *Small Ruminant Research*, 202, 106455. https://doi.org/10.1016/j.smallrumres.2021.106455

Bai, Y. et al. (2012) Current status and future perspectives for sequencing livestock genomes. *Journal of Animal Science and Biotechnology*, 3(1), 1–6. https://doi.org/10.1186/2049-1891-3-8

Brøndum, R.F. et al. (2014) Strategies for imputation to whole genome sequence using a single or multi-breed reference population in cattle. *BMC genomics*, 15(1), 1–8. https://doi.org/10.1186/1471-2164-15-728

Decker, J.E. et al. (2014) Worldwide patterns of ancestry, divergence, and admixture in domesticated cattle. *PLoS Genetics*, 10(3), e1004254. https://doi.org/10.1371/journal.pgen.1004254

Demir, E., & Balcioğlu, M.S. (2019) Genetic diversity and population structure of four cattle breeds raised in Turkey using microsatellite markers. *Czech Journal of Animal Science*, 64(10), 411–419. https://doi.org/10.17221/62/2019-CJAS

Demir, E. et al. (2021) A comprehensive review on genetic diversity and phylogenetic relationships among native Turkish cattle breeds based on microsatellite markers. *Turkish Journal of Veterinary and Animal Sciences*, 45(1), 1–10. https://doi:10.3906/vet-2006-107

Di Lorenzo, P. et al. (2018) Mitochondrial DNA variants of Podolian cattle breeds testify for a dual maternal origin. *PLoS One*, 13(2), e0192567. https://doi.org/10.1371/journal.pone.0192567

Edea, Z. et al. (2015) Genome-wide genetic diversity, population structure and admixture analysis in African and Asian cattle breeds. *Animal*, 9(2), 218–226. https://doi.org/10.1017/S1751731114002560

Elsik, C.G. et al. (2009) The genome sequence of taurine cattle: a window to ruminant biology and evolution. *Science*, 324(5926), 522–528. https://doi.org/10.1126/science.1169588

Excoffier, L. et al. (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, 131(2), 479–491. https://doi.org/10.1093/genetics/131.2.479

Fraga, A.B. et al. (2016) Multivariate analysis to evaluate genetic groups and production traits of crossbred Holstein × Zebu cows. *Tropical Animal Health and Production*, 48(3), 533–538. https://doi.org/10.1007/s11250-015-0985-2

Ghafar, A. et al. (2021). Targeted next-generation sequencing and informatics as an effective tool to establish the composition of bovine piroplasm populations in endemic regions. *Microorganisms*, 9(1), 21. https://doi.org/10.3390/microorganisms9010021

Ghosh, M. et al. (2018). Transformation of animal genomics by next-generation sequencing technologies: a decade of challenges and their impact on genetic architecture. *Critical Reviews in Biotechnology*, 38(8), 1157–1175. https://doi.org/10.1080/07388551.2018.1451819

Jiang, L. et al. (2014). Targeted resequencing of GWAS loci reveals novel genetic variants for milk production traits. *BMC Genomics*, 15(1), 1–9. https://doi.org/10.1186/1471-2164-15-1105

Jombart, T. et al. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*, 11, 1–15. https://doi.org/10.1186/1471-2156-11-94

Jost, L. (2009). D vs. GST: response to Heller and Siegismund (2009) and Ryman and Leimar (2009). *Molecular Ecology*, 18, 2088–2091. https://doi.org/10.1111/j.1365-294X.2009.04186.x

Kanduma, E.G. et al. (2016). Multi-locus genotyping reveals absence of genetic structure in field populations of the brown ear tick (*Rhipicephalus appendiculatus*) in Kenya. *Ticks and Tick-borne Diseases*, 7, 26–35. https://doi.org/10.1016/j.ttbdis.2015.08.001

Kõks, S. et al. (2013). Sequencing and annotated analysis of the Holstein cow genome. *Mammalian Genome*, 24(7), 309–321. https://doi.org/10.1007/s00335-013-9464-0

Kukučková, V. et al. (2017). Genomic characterisation of Pinzgau cattle: genetic conservation and breeding perspectives. *Conservation Genetics*, 18(4), 893–910. https://doi.org/10.1007/s10592-017-0935-9

Kukučková, V., Moravčíková, N., & Kasarda, R. (2016). Evaluating signatures of selection through variation in linkage disequilibrium among different cattle breeds. *Acta Fytotechnica et Zootechnica*, 19(5). http://dx.doi.org/10.15414/afz.2016.19.si.105-110

Leroy, G. et al. (2016). Rare phenotypes in domestic animals: unique resources for multiple applications. *Animal Genetics*, 47(2), 141–153. https://doi.org/10.1111/age.12393

Ma, L. et al. (2015). Statistical measures of genetic differentiation of populations: Rationales, history and current states. *Current Zoology*, 61, 886–897. https://doi.org/10.1093/czoolo/61.5.886

Makina, S.O. et al. (2014) Genetic diversity and population structure among six cattle breeds in South Africa using a whole genome SNP panel. *Frontiers in Genetics*, 5, 333. https://doi.org/10.3389/fgene.2014.00333

Maudet, C. et al. (2002) Genetic diversity and assignment tests among seven French cattle breeds based on microsatellite DNA analysis. *Journal of Animal Science*, 80(4), 942–950. https://doi.org/10.2527/2002.804942x

Mrode, R. et al. (2019). Genomic selection and use of molecular tools in breeding programs for indigenous and crossbred cattle in developing countries: Current status and future prospects. *Frontiers in Genetics*, 9, 694. https://doi.org/10.3389/fgene.2018.00694

Nei, M. (1972) Genetic distance between populations. *The American Naturalist*, 106(949), 283–292. https://doi.org/10.1086/282771

Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences*, 70, 3321–3323. https://doi.org/10.1073/pnas.70.12.3321

Nguluma, A.S. et al. (2018) Assessment of genetic variation among four populations of Small East African goats using microsatellite markers. *South African Journal of Animal Science*, 48(1), 117–127. https://doi.org/10.4314/sajas.v48i1.14

Olšanská, B. et al. (2020) Genome-wide characterisation of regions under intense selection based on runs of homozygosity in Charolais cattle. *Acta Fytotechnica et Zootechnica*, 23(5), 350–355. https://doi.org/10.15414/afz.2020.23.mi-fpap.350-355

Öner, Y. et al. (2019) Genetic diversity and population structure of Turkish native cattle breeds. *South African Journal of Animal Science*, 49(4), 628–635. https://doi.org/10.4314/sajas.v49i4.4

Orozco-terWengel, P. et al. (2015) Revisiting demographic processes in cattle with genome-wide population genetic analysis. *Frontiers in genetics*, 6(1), 191. https://doi.org/10.3389/fgene.2015.00191

Özşensoy, Y. et al. (2019) Phylogenetic relationships of native Turkish cattle breeds using microsatellite markers. *Turkish Journal of Veterinary and Animal Sciences*, 43(1), 23–29. https://doi.org/10.3906/vet-1805-10

Peterson, B.K. et al. (2012) Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PloS one*, 7(5), e37135. https://doi.org/10.1371/journal.pone.0037135

Pitt, D. et al. (2019) Domestication of cattle: Two or three events? *Evolutionary applications*, 12(1), 123–136. https://doi.org/10.1111/eva.12674

Pritchard, J.K. (2000) Inference of population structure using multilocus genotype data. *Genetics*, 155(2), 945–959. https://doi.org/10.1534/genetics.116.195164

Qanbari, S. (2020) On the extent of linkage disequilibrium in the genome of farm animals. *Frontiers in Genetics*, 10, 1304. https://doi.org/10.3389/fgene.2019.01304

Roh, H.J. et al. (2020). Estimating genetic diversity and population structure of 22 chicken breeds in Asia using microsatellite markers. *Asian-Australasian Journal of Animal Sciences*, 33(12), 1896. https://doi.org/10.5713/ajas.19.0958

Saravanan, K.A. et al. (2020) Selection signatures in livestock genome: A review of concepts, approaches and applications. *Livestock Science*, 241, 104257. https://doi.org/10.1016/j.livsci.2020.104257

Sharma, R. et al. (2015) Genetic diversity and relationship of Indian cattle inferred from microsatellite and mitochondrial DNA markers. *BMC genetics*, 16(1), 1–12. https://doi.org/10.1186/s12863-015-0221-0

Slatkin, M. (1995). A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, 139, 457–462. https://doi.org/10.1093/genetics/139.1.457

van der Westhuizen, L. et al. (2020) Genetic variability and relationships in nine South African cattle breeds using microsatellite markers. *Tropical Animal Health and Production*, 52(1), 177–184. https://doi.org/10.1007/s11250-019-02003-z

van Marle-Köster, E. et al. (2013). A review of genomic selection-Implications for the South African beef and dairy cattle industries. *South African Journal of Animal Science*, 43(1), 1–17. https://doi.org/10.4314/sajas.v43i1.1

Wang, J. (2012). On the measurements of genetic differentiation among populations. *Genetics Research*, 94, 275–289. https://doi.org/10.1017/S0016672312000481

Weldenegodguad, M. et al. (2019). Whole-genome sequencing of three native cattle breeds originating from the northernmost cattle farming regions. *Frontiers in Genetics*, 9, 728. https://doi.org/10.3389/fgene.2018.00728

Yilmaz, O. et al. (2012) The domestic livestock resources of Turkey: cattle local breeds and types and their conservation status. *Animal Genetic Resources*, 50, 65–73. https://doi.org/10.1017/S2078633612000033

Zhang, Y. (2020) Population structure, and selection signatures underlying high-altitude adaptation inferred from genome-wide copy number variations in Chinese indigenous cattle. *Frontiers in genetics*, 10, 1404. https://doi.org/10.3389/fgene.2019.01404