J. Dairy Sci. 107:11224–11238
https://doi.org/10.3168/jds.2024-24952


# Identification of nonsense variants in the genomes of 15 Murciano-Granadina bucks and analysis of their segregation in parent-offspring trios

**Ke Wang,[1,2,3] María Gracia Luigi-Sierra,[1] Anna Castelló,[1,3] Taina Figueiredo-Cardoso,[1,3] Anna Mercadé,[3] Amparo Martínez,[4] Juan Vicente Delgado,[4] Javier Fernández Álvarez,[4] Antonia Noce,[1] Mingjing Wang,[1] Jordi Jordana,[3] and Marcel Amills[1,3]***

[1]Centre de Recerca Agrigenòmica (CRAG), CSIC-IRTA-UAB-UB, Campus Universitat Autònoma de Barcelona, Bellaterra 08193, Spain
[2]Chinese Academy of Tropical Agricultural Sciences, Zhanjiang Experimental Station, Zhanjiang, Guangdong 524000, China
[3]Departament de Ciència Animal i dels Aliments, Facultat de Veterinària, Universitat Autònoma de Barcelona, Bellaterra 08193, Spain
[4]Departamento de Genética, Universidad de Córdoba, Córdoba 14071, Spain

## ABSTRACT

Nonsense variants can inactivate gene function by causing the synthesis of truncated proteins or by inducing nonsense mediated decay of messenger RNAs. The occurrence of such variants in the genomes of livestock species is modulated by multiple demographic and selective factors. Even though nonsense variants can have causal effects on embryo lethality, abortions, and disease, their genomic distribution and segregation in domestic goats have not been characterized in depth yet. In this work, we have sequenced the genomes of 15 Murciano-Granadina bucks with an average coverage of 32.92× ± 1.45×. Bioinformatic analysis revealed 947 nonsense variants consistently detected with SnpEff and Ensembl-VEP. These variants were especially abundant in the 3′ end of the protein-coding regions. Genes related to olfactory perception, ATPase activity coupled to transmembrane movement of substances, defense to virus, hormonal response, and sensory perception of taste were particularly enriched in nonsense variants. Seventeen nonsense variants expected to have harmful effects on fitness were genotyped in parent-offspring trios. We observed that several nonsense variants predicted to be lethal based on mouse knockout data did not have such effect, a finding that could be explained by the existence of multiple mechanisms counteracting lethality. These findings demonstrate that predicting the effects of putative nonsense variants on fitness is extremely challenging. As a matter of fact, such a goal could only be achieved by generating a high-quality telomere-to-telomere goat reference genome combined with carefully curated annotation and functional testing of promising candidate variants.

**Key words:** knockout, stop gain variant, protein truncation, cost of domestication hypothesis

## INTRODUCTION

There is evidence that deleterious variants accumulated in the genomes of domestic animals and plants at a higher rate than in their wild ancestors, a circumstance that imposed a cost on their fitness (Moyers et al., 2018). In populations experiencing bottlenecks, such as the domestic ones, purifying selection is less able to remove variants with adverse effects on fitness, genetic drift being the key evolutionary force deciding their fate (Moyers et al., 2018). Inbreeding, linkage with variants under artificial selection, and fast population growth combined with long-distance migration also contributed to increase the frequencies of deleterious variants in domestic populations (Moyers et al., 2018). Notably, the linkage between deleterious variants and those with causal effects on phenotypes of interest might impose limits on the response of such traits to artificial selection (Moyers et al., 2018).

Among deleterious variation, nonsense variants are of particular importance because they can inactivate protein function by inducing either the premature cessation of translation or nonsense mediated decay of the corresponding transcript (MacArthur and Tyler-Smith 2010). Variants at canonical splice sites and missense substitutions can also have adverse effects on gene function, but in general these are less drastic than those of nonsense variants (Abramowicz and Gos 2018; González-Prendes et al., 2023). Indeed, nonsense variants are among the most frequent causes of genetic disease, representing ~11.5% of human inherited diseases and 20% of disease-associated single-nucleotide substitutions residing in gene coding regions (Mort et al., 2008). In addition, premature stop codons might also have negative effects on

Received March 25, 2024.
Accepted July 31, 2024.
*Corresponding author: marcel.amills@uab.cat

The list of standard abbreviations for JDS is available at adsa.org/jds-abbreviations-24. Nonstandard abbreviations are available in the Notes.

reproduction by impairing embryonic development and inducing abortions (Derks et al., 2019).

The potential consequences of nonsense variants have often been predicted based on phenotypic data gathered in knockout mice (Majzoub and Muglia 1996). However, there is increasing evidence that phenotypes associated with gene inactivation in mice are not always recapitulated in humans homozygous for alleles inducing the truncation of the corresponding protein (Narasimhan et al., 2016). This could be due to multiple mechanisms, such as gene redundancy (El-Brolosy and Stainier 2017), alternative splicing (Narasimhan et al., 2016), ribosome readthrough (Li and Zhang 2019), and full linkage with suppressor variants (Matsui et al., 2017), to mention a few. In this work, we wanted to characterize the genomic landscape of nonsense variants in goats by sequencing the genomes of 15 Murciano-Granadina bucks. Moreover, we aimed to select 20 nonsense variant candidates to have harmful effects on viability, according to information collected in the Mouse Genome Informatics (**MGI**; Baldarelli et al., 2021) and Online Mendelian Inheritance in Man (**OMIM**; Hamosh et al., 2005) database annotations, and to genotype them in buck-dam-offspring trios to investigate if any of them displays an altered pattern of transmission.

## MATERIALS AND METHODS

### Ethics Statement

The collection of blood is a routine procedure carried out by trained veterinarians working for the National Association of Murciano-Granadina Goat Breeders (CAPRIGRAN), so it does not require a permission from the Committee on Ethics in Animal and Human Experimentation of the Universitat Autònoma de Barcelona. Permission to extract ear samples was granted by the Review Committee for the Use of Animal Subjects of Northwest A&F University (contract number NWAFAC1008).

### Animal Material

We have extracted blood samples from 6 Murciano-Granadina bucks, 284 dams mated to the 6 bucks, and 302 offspring produced by such matings. These 6 bucks were selected because they had ~40 to 50 offspring available. Moreover, blood was also retrieved from 9 additional Murciano-Granadina bucks as part of an experiment unrelated to the current study. Blood was kept in EDTA $K_3$ coated vacuum tubes and stored at −20°C before processing. Murciano-Granadina is an autochthonous Spanish breed officially created in 1975 by the cross-breeding of 2 different Murciano and Granadina goat populations. Currently, the breed has a census of 115,105 heads (2020), and its adaptability to harsh environments as well as its good milking records (mean of 586 kg/lactation; 5.1% of fat and 3.6% of protein in milk) have made it a highly appreciated breed in Spain and other countries (https://www.mapa.gob.es/es/ganaderia/temas/zootecnia/razas-ganaderas/razas/catalogo-razas/caprino/murciano-granadina/). Part of the work developed in this project was carried out in the Zhanjiang Experimental Station of the Chinese Academy of Tropical Agricultural Sciences, located in Zhanjiang, Guangdong Province. Because of that, we had access to a collection of genomic DNA preparations from 250 ear tissues from 4 native Chinese breeds (Shaanbei White Cashmere goats, n = 50; Guanzhong goats, n = 50; Leizhou black goats, n = 50; Hezhang black goats, n = 50) and Nubian goats (n = 50) raised under the same feeding and management conditions (Gu et al., 2022), in the Demonstration Base of Tropical Herbage & Livestock Circular Agriculture (Zhanjiang, China). These goats were used to carry out the Sanger sequencing (Sanger et al., 1977) of the region containing a putative nonsense variant in the *DAXX* gene.

### Sequencing of Genomic DNA from 15 Murciano-Granadina Bucks and Bioinformatic Analysis

Genomic DNA from the 15 bucks were sequenced at the Centre Nacional d'Anàlisi Genòmica (Barcelona, Spain; https://www.cnag.eu). The synthesis of paired-end multiplex libraries was performed using the KAPA HyperPrep kit (Roche, Sant Cugat, Spain), following the instructions of the manufacturer. Libraries were sequenced on a NovaSeq 6000 sequencing platform (Illumina) with a read length of 2 × 150 base pairs. Base calling and quality control analyses were carried out with the RTA v3 analysis pipeline following the guidelines of the manufacturer. Quality-checked filtered reads were mapped to the *Capra hircus* genome version ARS1 (Bickhart et al., 2017) using the BWA aligner v0.7.17 (Li and Durbin 2009). Aligned reads in Binary Alignment Map (**BAM**) files were sorted with Samtools v0.1.19 (Li et al., 2009) and duplicated reads were marked with Picard MarkDuplicates (https://broadinstitute.github.io/picard/). Joint variant calling was performed with the Genome Analysis Toolkit (McKenna et al., 2010). The resulting variants were annotated with the database ARS1.99 using the SnpEff v5.0 (Cingolani et al., 2012) and Ensembl-VEP (McLaren et al., 2016) packages. To extract the intersection of the nonsense variants detected with both methods, we used VCF-compare (https://vcftools.github.io/perl_module.html#VCF-compare). The allelic frequencies of nonsense variants were calculated with VCFtools v4.2 (Danecek et al., 2011). Loci containing multiple potentially deleterious variants or missense

variants in the same individual underwent in-depth examination through visual inspection of BAM alignments using the Integrative Genomics Viewer (https://igv.org/; Thorvaldsdóttir et al., 2013) to check if they are pseudogenes. Variants detected within reads mapping to putative pseudogenes were systematically removed. In addition, we used the multibase codon-associated variant re-annotation (**MACARON**) software (Khan et al., 2018) to identify, re-annotate, and predict the AA change resulting from multiple single-nucleotide variants within the same genetic codon. We did so to exclude nonsense variants mapping to codons containing other variants disrupting their expected protein truncation effect.

### *Functional Annotation of Genes Containing Nonsense Variants*

The functional constraint of genes containing nonsense variants was assessed based on a loss-of-function observed/expected upper bound fraction (**LOEUF**) metric derived from the Genome Aggregation Database (gnomAD v3.1) sequencing data sets. The LOEUF metric (Karczewski et al., 2020) represents a conservative estimate of the ratio of observed to expected loss-of-function variants, and it can range from 0 (gene depleted from loss-of-function variants due to a strong evolutionary constraint) to 9 (gene not depleted from loss-of-function variants because of the absence of evolutionary constraint). Moreover, Gene Ontology (**GO**) enrichment analysis and the Kyoto Encyclopedia of Genes and Genomes (**KEGG**) analysis were performed with Annotation Visualization and Integration Discovery (https://david.ncifcrf.gov/summary.jsp), gprofiler (Raudvere et al., 2019; https://biit.cs.ut.ee/gprofiler/gost), and KOBAS (Bu et al., 2021; http://kobas.cbi.pku.edu.cn/anno_iden.php). Data from the MGI database (Baldarelli et al., 2021; http://www.informatics.jax.org/) and the OMIM (Hamosh et al., 2005; https://www.omim.org/) were used to infer the potential phenotypic effects of nonsense variants.

### *TaqMan Open Array Genotyping of 20 Nonsense Variants with Predicted Harmful Effects*

We aimed to select a set of 20 nonsense variants with likely harmful consequences to be genotyped in parent-offspring trios. The potential deleteriousness of variants was predicted according to information from the MGI database (Baldarelli et al., 2021; http://www.informatics.jax.org/) and OMIM (Hamosh et al., 2005; https://www.omim.org/) database. Although nonsense variants within 20% of the 3′ end coding region are expected to be less harmful than those outside, we did not exclude such variants from our analysis because this criterium would have been too restrictive, leading to the exclusion of many variants of potential interest. The existence of multiple mRNA isoforms might also limit the effect of a nonsense variant (due to the existence of isoforms lacking the exon carrying the nonsense variant). However, we did not consider this criterium because deleteriousness can be strongly influenced by the expression levels of each isoform, a parameter that is currently unknown. Finally, each candidate variant was manually annotated and those that do not introduce stop codons were eliminated from the data set.

The upstream and downstream flanking sequences (100 bp) of the 20 selected variants were submitted to the Custom TaqMan Assay Design Tool website (https://www5.appliedbiosystems.com/tools/cadt, Life Technologies) to check if they could be typed in a TaqMan Open Array multiplex assay platform. Selected variants were genotyped in dams (n = 284) mated with each one of the 6 bucks and their offspring (n = 302). Genotyping was performed at the Servei Veterinari de Genética Molecular of the Universitat Autònoma de Barcelona (http://sct.uab.cat/svgm/en) by using a QuantStudio 12K Flex Real-Time PCR System (Thermo Fisher Scientific). Analysis and visualization of the results were carried out in a dedicated Thermo Fisher online platform (http://apps.thermofisher.com/). The Hardy-Weinberg equilibrium test was performed for each variant by using the Science Primer platform (https://scienceprimer.com/hardy-weinberg-equilibrium-calculator).

### *Sanger Sequencing of the Region Containing a Nonsense Variant in the DAXX Gene*

The nonsense XM_018038577.1:c.1134C > G variant associated with the goat *DAXX* gene was the only one showing a significant ($P < 0.00001$) depletion of homozygous genotypes for the G variant. When designing primers to amplify the region containing this variant and blasting them against the goat genome, we became aware that the region containing this variant is duplicated. So, we designed 2 pairs of primers (Supplemental Figure S1, see Notes) amplifying the chromosome-23 region containing the XM_018038577.1:c.1134C > G variant (represented by GenBank sequence XM_018038577.1) and the duplicated region on chromosome 6 (represented by GenBank sequence XM_018049164.1), which harbors a death domain-associated protein 6-like pseudogene (*LOC108636199*). Sequences of the primers were F_5′ CCC TCG TTC GTC AGG TGT GT-3′ and R_5′-AGG GAT CTG TTG CCC CGT TC-3′ for the partial amplification of the *DAXX* gene on chromosome 23, and F_5′-AGA TGA CTA TAG GCC AGG CAT TG-3′ and R_5′-TCC AGC GAG GAT TTA GGG GT-3′ for the partial amplification of the *LOC108636199* locus on chromosome 6 (Supplemental Figure S1).

Regions flanked by these primers were amplified in a set of individuals comprising Shaanbei White Cashmere goats (n = 50), Guanzhong goats (n = 50), Leizhou black goats (n = 50), Hezhang Black goats (n = 50), and Nubian goats (n = 50). Genomic DNA was isolated from ear-tissue samples using the Genomic DNA Extraction Kit for animal tissues or cells (Solarbio, Beijing, China). The volume of the amplification reaction was 25 μL, and it contained 1.0 μL of genomic DNA (20 ng/μL), 1.0 μL of each primer (10 μ*M*), and 12.5 μL of 2× *Taq* Master mix, which included 1.5 m*M* $MgCl_2$, 200 μ*M* dNTP, and 0.625 units of *Taq* DNA polymerase (BioLinker, Shanghai, China) The thermal profile was 94°C for 5 min; followed by 35 cycles of 30 s at 94°C, 30 s at 55°C, and 1 min at 72°C, plus a final extension step at 72°C for 5 min. Sanger sequencing of amplicons corresponding to these 2 regions in 250 individuals was performed by the AuGCT Company (Beijing, China) by using an Applied Biosystems 3730XL DNA sequencer (AuGCT, Beijing, China).

## RESULTS

### Identification of Nonsense Variants and Their Genomic Distribution

The statistics obtained in the whole-genome sequencing experiment are shown in Supplemental Table S1 (see Notes). The average coverage of the 15 goat genomes was 32.92×, and the average percentage of uniquely mapped reads was 90.13%. With SnpEff (Cingolani et al., 2012) and Ensembl-VEP (McLaren et al., 2016), we detected 991 and 1,463 nonsense variants segregating in the genomes of the 15 Murciano-Granadina bucks, respectively. By using VCF-compare, we identified 947 nonsense variants consistently detected with both methods. After quality control and filtering (−max-missing 0.90 −minQ 30 −minDP 10), we retained 866 nonsense variants. From these, we filtered out 26 multinucleotide variants disrupting premature stop codons with MACARON (Khan et al., 2018). The allelic frequencies of nonsense variants were in general quite low (Figure 1A), because ~28% of them had frequencies below 0.05. Regarding the genomic distribution of nonsense variants, we observed that they are quite scattered along the protein-coding region, being particularly enriched in its 3′ end (Figure 1B).

### Functions and Biological Essentiality of Genes Containing Nonsense Variants

Genes containing nonsense variants were enriched in GO terms related to the integral component of the membrane ($P$ = 1.4E-15), the G-protein coupled receptor activity ($P$ = 6.4E-11), and the olfactory receptor activity (Table 1; $P$ = 1.5E-10). We also found a significant enrichment for the ATPase activity coupled to transmembrane movement of substances pathway ($P$ = 0.0029), as the remaining pathways, including defense response to the virus ($P$ = 0.0272), hormonal activity ($P$ = 0.0294), and sensory perception of taste ($P$ = 0.0417), were less significant. The KEGG pathway analysis also showed that, by far, olfactory transduction is the most significant pathway (P = 3.51E-29), followed by herpes simplex virus 1 infection ($P$ = 1.61E-10) and antigen processing and presentation (Table 1; $P$ = 3.34E-5). We also calculated the LOEUF scores of genes containing nonsense variants (Figure 1C). As expected, most of these genes had LOEUF scores of 5 or larger indicating that they are not under strong functional constraint, a result that is concordant with the functional enrichment in genes related to olfactory transduction.

### Genotyping of Variants Predicted to Have Harmful Effects in Parent-Offspring Trios

By using the OMIM (Hamosh et al., 2005) and MGI (Baldarelli et al., 2021) databases, we identified 23 and 51 genes, respectively, containing nonsense variants and with LOEUF scores <6. The inactivation of several of these loci is expected to have harmful consequences including, but not limited to, lethality (Table 2). Based on such information plus bibliographic references, we initially selected 20 nonsense variants mapping to 19 genes (Table 2). Manual annotations of these variants (Supplemental Figure S2, see Notes) revealed that 3 of them do not introduce stop codons (Supplemental Figure S2), so they were removed from our data set. The remaining 17 variants were genotyped with an Open Array system in buck-dam-offspring trios formed by the 284 dams mated to the 6 bucks and 302 offspring (Table 2).

After completing genotyping tasks, 4 nonsense variants in the *ACVR2B*, *RNLS*, *NPC1*, and *OTUD6B* genes did not yield valid results. As shown in Supplemental Figure S2, we observed that 12 out of the 13 successfully genotyped variants segregated in the genomes of a panel of 250 European goats (10 breeds), 249 African goats (4 breeds), and 109 Asian goats (4 breeds) that were sequenced in the VarGoats project (Denoyelle et al., 2021). Full details regarding how this whole-genome sequence data set was built can be found in Denoyelle et al. (2021). The only variant that did not segregate in any of these 608 individuals was XP_017894826.1:p.Arg204* in the *BCL2* gene (Supplemental Figure S2), and the same variant had an extremely low frequency among the genotyped Murciano-Granadina individuals (Table 2). The remaining variants segregated in the investigated 608 goats from Europe, Africa, and Asia although, in general, at a minimum allele frequency (**MAF**) lower than 10%. Ex-

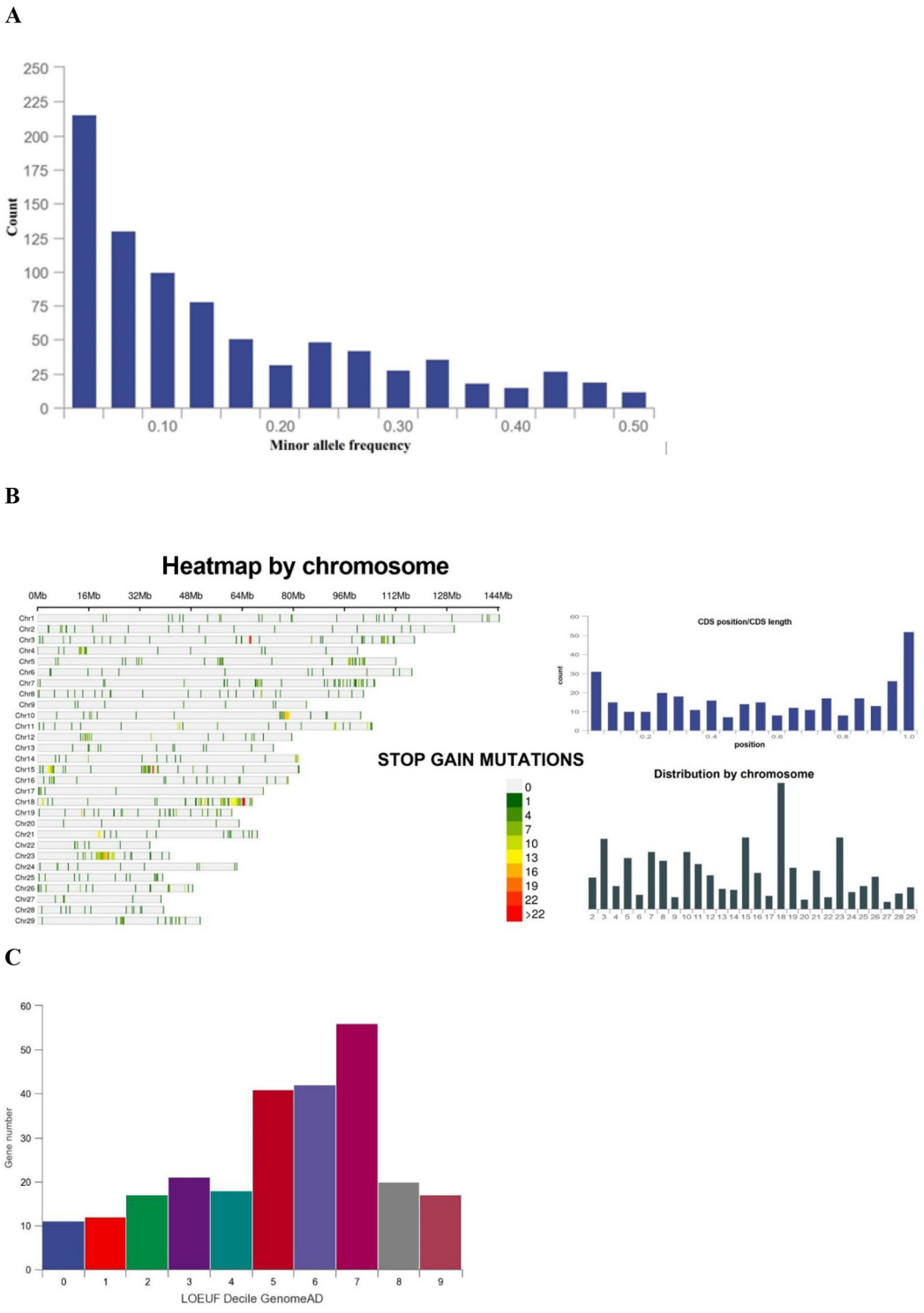


**Figure 1.** (A) Allele frequencies of nonsense variants detected in the genomes of 15 Murciano-Granadina bucks. (B) Chromosome distribution (heatmap and absolute counts per chromosome) and location along the protein-coding region of nonsense variants segregating in the genomes of 15 Murciano-Granadina bucks. (C) Absolute counts of genes containing nonsense variants that have been classified according to their LOEUF score. A LOEUF score of 0 indicates that a gene is under strong selective constraint and has an essential function, whereas a LOEUF score of 9 corresponds to a gene for which inactivation is well tolerated. CDS = coding sequence.

**Table 1.** Gene ontology terms and KEGG pathways significantly (corrected *P*-value <0.05) overrepresented among genes containing predicted nonsense variants segregating in the genomes of 15 Murciano-Granadina bucks

| Term | No. of genes | *P*-value | Corrected *P*-value[1] |
|---|---|---|---|
| Gene ontology terms | | | |
| GO:0016021~integral component of membrane | 144 | 2.67E-16 | 1.4E-15 |
| GO:0004930~G-protein coupled receptor activity | 47 | 6.05E-12 | 6.4E-11 |
| GO:0004984~olfactory receptor activity | 42 | 2.20E-11 | 1.5E-10 |
| GO:0042626~ATPase activity, coupled to transmembrane movement of substances | 7 | 0.0011707 | 0.0029473 |
| GO:0015914~phospholipid transport | 3 | 0.0120228 | 0.0148164 |
| GO:0006355~regulation of transcription, DNA-templated | 12 | 0.0125267 | 0.0176249 |
| GO:0051607~defense response to virus | 7 | 0.0127781 | 0.0272945 |
| GO:0005179~hormone activity | 6 | 0.0167250 | 0.0294567 |
| GO:0030515~snoRNA binding | 3 | 0.0168168 | 0.0294613 |
| GO:0022857~transmembrane transporter activity | 7 | 0.0188329 | 0.0322914 |
| GO:0005886~plasma membrane | 48 | 0.0238185 | 0.0341911 |
| GO:0050909~sensory perception of taste | 3 | 0.0262560 | 0.0417436 |
| GO:0005044~scavenger receptor activity | 5 | 0.0321463 | 0.0486462 |
| KEGG pathways | | | |
| Olfactory transduction | 105 | 1.56E-31 | 3.51E-29 |
| Herpes simplex virus 1 infection | 42 | 1.43E-12 | 1.61E-10 |
| Antigen processing and presentation | 13 | 4.46E-7 | 3.34E-5 |
| Autoimmune thyroid disease | 11 | 2.75E-6 | 1.55E-4 |
| Graft-versus-host disease | 9 | 1.29E-5 | 5.79E-4 |
| Allograft rejection | 9 | 2.22E-5 | 8.33E-4 |
| Type I diabetes mellitus | 9 | 5.24E-5 | 1.68E-3 |
| Viral myocarditis | 8 | 7.75E-4 | 2.00E-2 |
| Complement and coagulation cascades | 9 | 8.01E-4 | 2.00E-2 |
| ABC transporters | 8 | 1.16E-3 | 2.60E-2 |

[1]*P*-value corrected for multiple testing.

ceptions to this rule were *NANOG* NP_001272505.1:p.Gln273* in African (MAF = 0.205) and Asian (MAF = 0.1225) goats, *MAP9* XP_017917128.1:p.Arg611* in Asian goats (MAF = 0.1875), *PDE4D* XP_005694732.3:p.Cys50* in European goats (MAF = 0.357), and *DAXX* XP_017894066.1:p.Tyr378* in European (MAF = 0.16) and African (MAF = 0.115) goats.

We did not observe a significant absence of homozygous genotypes for several nonsense variants located in genes whose inactivation causes lethality in knockout mice (Table 3), such as *MECR* (XM_018057207.1:c.1152C > G, with the G-allele introducing a premature stop codon), *BRCA2* (XM_018056629.1:c.8754T > A, with the A-allele introducing a premature stop codon), and *NANOG* (NM_001285576.1:c.817C > T, with the T-allele introducing a premature stop codon). For the XM_018063707.1: c.5419C > T (with the T-allele introducing a premature stop codon) variant in the *TNRC6C* gene, which is also a lethal gene in mice, only one individual displayed a TT genotype (Table 3). When checking the annotation of these 4 variants, we verified that all of them are located within the 10% 3′ end of the protein-coding region (Supplemental Figure S2). Nonsense variants in the *PPA2*, *MAP9*, and *SLC25A26* genes also mapped to such region (Supplemental Figure S2). In Supplemental Figure S3 (see Notes), we show the repertoire of mRNA isoforms for genes with at least 2 distinct mRNA isoforms. In the case of the *MECR* and *TNRC6C* genes, the nonsense variants only affected one (*MECR*: mRNA X2, *TNRC6C*: mRNA X3) of the 6 (*MECR*) or 7 (*TNRC6C*) reported isoforms (Supplemental Figure S3). Moreover, the 2 affected isoforms showed a highly divergent sequence when compared with their counterparts. This isoform-specificity of nonsense variants was also observed for the *BCL2*, *MAP9*, *PPA2*, and *SLC25A26* genes (Supplemental Figure S3).

In addition, 6 variants in the *NBAS*, *MAP9*, *SLC25A26*, *DAXX*, and *BCL2* genes showed a complete absence of individuals homozygous for the nonsense variant (Table 3), but such depletion was only significant ($P < 0.00001$) for the XM_018038577.1: c.1134C > G variant in the *DAXX* gene. Visual inspection of *DAXX* genotype clustering with the Thermo Fisher online platform (http://apps.thermofisher.com/) showed that the cluster of heterozygotes was quite scattered (Supplemental Figure S4, see Notes), suggesting that such cluster might be the result of the overlap of CG and GG clusters.

Because the *DAXX* XM_018038577.1:c.1134C > G variant was a promising candidate variant to have lethal effects and the Open Array (Thermo Fisher) genotyping results were doubtful due to the scattered distribution of the heterozygous genotype, we decided to re-genotype this variant by Sanger sequencing. When designing primers, we became aware that the chromosome 23: 41192961–41193179 bp region containing the XM_018038577.1:c.1134C > G variant is duplicated on

**Table 2.** Location and main features of 20 predicted nonsense variants with potential harmful effects on biological viability

| CHIR[1] | Position | Gene symbol | Phenotype | AA and nucleotide replacement | |
|---|---|---|---|---|---|
| 1 | 79618036 | *RTP4* | Homozygous knockout reduces susceptibility to parasitic and viral infection[2] | XP_005675197.2: p.Arg257* | XM_005675140.3: c.769C > T |
| 2 | 11614286 | *MECR* | Homozygous variants in this gene lead to embryonic lethality due to pleiotropic effects including abnormalities in fetal placental and cardiovascular development and embryonic growth retardation[2] | XP_017912696.1: p.Tyr384* | XM_018057207.1: c.1152C > G |
| 3 | 62127712 | *MCOLN3* | Homozygotes are white with small patches of color and show severe behavioral abnormalities, poor postnatal viability and are almost infertile[2] | XP_017901176.1: p.Trp186* | XM_018045687.1: c.557G > A |
| 5 | 27520597 | *NR4A1* | NR4A1 and NR4A3 abrogation lead to rapidly lethal acute myeloid leukemia in mice[3] | **QHW05400.1: p.Arg588***[6] | **MN197544: c.1762C > T**[6] |
| 6 | 20144367 | *PPA2* | PP2A mutations may cause a spectrum of mitochondrial disease phenotypes; severe symptoms include seizures, lactic acidosis, cardiac arrhythmia, and death within days of birth[4] | XP_017904735.1: p.Arg327* | XM_018049246.1: c.979C > T |
| 11 | 82481680 | *NBAS* | The NBAS gene is recurrently mutated in pediatric hemophagocytic lymphohistiocytosis[5] | XP_017911109.1: p.Gln105* | XM_018055620.1: c.313C > T |
| 12 | 58071934 | *BRCA2* | Homozygous null mutants are embryonic lethal with abnormalities including growth retardation, neural tube defects, and mesoderm abnormalities[2] | XP_017912118.1: p.Tyr2918* | XM_018056629.1: c.8754T > A |
| 14 | 81057228 | *PLEC* | Targeted variants of this gene result in neonatal death, skin blistering, impaired myofibril integrity, reduced hemidesmosome number, and disintegration of intercalated disks in the heart[2] | **XP_017914185.1: p.Trp3754***[6] | **XM_018058696.1: c.11261G > A**[6] |
| 14 | 81058774 | *PLEC* | | **XP_017914185.1: p.Arg3239***[6] | **XM_018058696.1: c.9715A > T**[6] |
| 14 | 8887926 | *OTUD6B* | Mice homozygous for a knockout allele exhibit complete perinatal lethality, decreased fetal size, and ventricular septal defects[2] | XP_005689305.1: p.Trp10* | XM_005689248.3: c.30G > A |
| 15 | 81546465 | *NANOG* | Mice homozygous for a disruption in this gene die between E3.5 and E5.5 with abnormal embryonic and extraembryonic tissue development[2] | NP_001272505.1: p.Gln273* | NM_001285576.1: c.817C > T |
| 17 | 70604449 | *MAP9* | Mice homozygous or heterozygous for a null allele exhibit increased aneuploidy, chromosomal instability, detached centrosomes, and pericentriolar material fragmentation in mouse embryonic fibroblasts[2] | XP_017917129.1: p.Arg611* | XM_018061639.1: c.1831C > T |
| 19 | 53253007 | *TNRC6C* | Mice homozygous for a gene trap allele exhibit complete neonatal lethality with cyanosis, respiratory distress, and thickened mesenchyme in air sacs[2] | XP_017919196.1: p.Arg1807* | XM_018063707.1: c.5419C > T |
| 20 | 20004713 | *PDE4D* | Homozygotes for targeted null variants exhibit delayed growth, female infertility associated with impaired ovulation, and reduced postnatal viability[2] | XP_005694732.3: p.Cys50* | XM_005694675.3: c.150C > A |
| 22 | 11721225 | *ACVR2B* | Mice homozygous for targeted variants that inactivate the gene show abnormal lateral asymmetry and homeotic transformation of the axial skeleton, and die shortly after birth with extensive cardiac defects[2] | XP_017922111.1: p.Lys213* | XM_018066622.1: c.637A > T |

*Continued*

**Table 2 (Continued).** Location and main features of 20 predicted nonsense variants with potential harmful effects on biological viability

| CHIR[1] | Position | Gene symbol | Phenotype | AA and nucleotide replacement | |
|---|---|---|---|---|---|
| 22 | 35072050 | *SLC25A26* | Embryos homozygous for a transposon insertion appear growth retarded and underdeveloped and die after E8.5 but before birth[2] | XP_005695813.1: p.Gln284* | XM_005695756.3: c.850C > T |
| 23 | 41193060 | *DAXX* | Mice homozygous for a targeted variant of this gene display extensive apoptosis and embryonic lethality[2] | XP_017894066.1: p.Tyr378* | XM_018038577.1: c.1134C > G |
| 24 | 33418088 | *NPC1* | Homozygotes for spontaneous and chemically induced variants may exhibit lysosomal storage of nonesterified cholesterol, neurodegeneration, ataxia, presence of foam cells, sterility, and shortened lifespan[2] | XP_017894966.1: p.Arg1264* | XM_018039477.1: c.3790C > T |
| 24 | 61676529 | *BCL2* | Homozygous null mutants show pleiotropic abnormalities including small size, increased postnatal mortality, polycystic kidneys, apoptotic involution of thymus and spleen, graying in the second hair follicle cycle, and reduced numbers of motor, sympathetic, and sensory neurons[2] | XP_017894826.1: p.Arg204* | XM_018039337.1: c.610C > T |
| 26 | 41247997 | *RNLS* | Mice homozygous for a knockout allele exhibit aggravated ischemic myocardial damage, increased heart rate, increased blood pressure, and increased serum levels of dopamine, adrenaline, and noradrenaline[2] | XP_017897022.1: p.Gln85* | XM_018041533.1: c.253C > T |

[1]CHIR = *Capra hircus* chromosome.
[2]MGI Mouse database.
[3]Mullican et al., 2007.
[4]Kennedy et al., 2016.
[5]Bi et al., 2022.
[6]Variants marked in bold are not correctly annotated and do not introduce any premature stop codon.

chromosome 6: 3909498–3909716 bp (Figure 2). In consequence, we designed primers (Supplemental Figure S1) to specifically amplify each one of the 2 copies of the duplicated region. Two amplicons of 433 bp (chromosome 23) and 222 bp (chromosome 6) were sequenced with the Sanger method. Comparison of the sequences (Figure 2A) showed that the putative XM_018038577.1:c.1134 C > G variant in the *DAXX* gene (chromosome 23) is, in reality, a dinucleotide variant mapping to the *LOC108636199* locus on chromosome 6 (XM_018049164.1:c.1270–1271 CA > GG variant). Indeed, the 2 orthologous sites in the *DAXX* gene are monomorphic. In consequence, for the *DAXX* gene only one genotype (CC) is possible for the putative XM_018038577.1:c.1134C > G variant on chromosome 23, whereas 3 genotypes (CC, CG, and GG) could eventually be found in the *LOC108636199* locus for position 1270 of the XM_018049164.1:c.1270–1271 CA > GG variant on chromosome 6 targeted in the TaqMan assay (Figure 2B). These 2 regions have very similar nucleotide sequences (Figure 2A), so when they co-amplify the resulting genotypes for the site targeted by the TaqMan assay can be exclusively CG (chromosome 23: CC + chromosome 6: CG or GG) or CC (chromosome 23: CC + chromosome 6: CC). This results in an apparent depletion of GG individuals. Indeed, no GG individuals were identified in the 250 Shaanbei White Cashmere, Guanzhong, Leizhou Black, Hezhang Black, and Nubian individuals for which the chromosome 23 region containing the nonsense variant was amplified with specific primers and sequenced with the Sanger method.

## DISCUSSION

### *The Genomic Landscape of Nonsense Variants in 15 Capra hircus Genomes*

The sequencing of 15 Murciano-Granadina bucks with good coverage (32.9×) allowed us to identify 947 nonsense variants. Sequencing of 600 bovine exomes revealed the segregation of 1,377 nonsense variants (Charlier et al., 2016). In humans, Pelak et al. (2010) sequenced 20 genomes with a coverage similar to ours and detected 554 nonsense variants. Estimates obtained by the 1000 Genomes Project Consortium (2010, 2015)

**Table 3.** Absolute frequencies[1] and departure from the Hardy-Weinberg equilibrium of 13 predicted nonsense variants successfully genotyped in dams mated with each one of the bucks and their offspring (in bold: nonsense variants never found in homozygosity)[2]

| | | | Buck[3] | | Dam | | Offspring | |
|---|---|---|---|---|---|---|---|---|
| CHIR | Gene | Variant | Frequency | HW *P*-value | Frequency | HW *P*-value | Frequency | HW *P*-value |
| 1 | *RTP4* | XP_005675197.2: p.Arg257* | 3/2/1 (8/6/1) | 0.25 (0.28) | 202/51/9 | 0.06 | 171/85/21 | 0.09 |
| 2 | *MECR* | XP_017912696.1: p.Tyr384* | 3/3/0 (10/5/0) | 0.25 (0.44) | 232/35/2 | 0.87 | 169/70/4 | 0.56 |
| 3 | *MCOLN3* | XP_017901176.1: p.Trp186* | 4/2/0 (8/6/1) | 0.44 (0.28) | 165/88/23 | 0.09 | 167/108/11 | 0.45 |
| 6 | *PPA2* | XP_017904735.1: p.Arg327* | 5/0/1 (10/3/2) | 0.69 (0.44) | 153/61/14 | 0.08 | 131/77/9 | 0.72 |
| **11** | *NBAS* | XP_017911109.1 p.Gln105* | **5/1/0 (14/1/0)** | **0.69 (0.87)** | **274/1/0** | **1.00** | **264/21/0** | **0.81** |
| 12 | *BRCA2* | XP_017912118.1: p.Tyr2918* | 4/2/0 (9/5/1) | 0.44 (0.36) | 194/70/9 | 0.69 | 199/72/13 | 0.17 |
| 15 | *NANOG* | NP_001272505.1: p.Gln273* | 5/1/0 (13/2/0) | 0.69 (0.75) | 130/101/25 | 0.71 | 148/99/12 | 0.67 |
| **17** | *MAP9* | XP_0179171289.1: p.Arg611* | **4/2/0 (11/4/0)** | **0.44 (0.54)** | **249/13/0** | **0.92** | **219/39/0** | **0.42** |
| 19 | *TNRC6C* | XP_017919196.1: p.Arg1807* | 5/1/0 (14/1/0) | 0.69 (0.87) | 254/20/1 | 0.68 | 259/26/0 | 0.72 |
| 20 | *PDE4D* | XP_005694732.3: p.Cys50* | 3/3/0 (11/4/0) | 0.25 (0.54) | 198/71/4 | 0.70 | 200/78/8 | 0.99 |
| **22** | *SLC25A26* | XP_005695813.1: p.Gln284* | **5/1/0 (13/2/0)** | **0.69 (0.75)** | **262/10/0** | **0.95** | **252/28/0** | **0.68** |
| **23** | *DAXX* | XP_017894066.1: p.Tyr378* | **5/1/0 (11/4/0)** | **0.69 (0.54)** | **103/171/0** | **0.00** | **146/138/0** | **0.00** |
| **24** | *BCL2* | XP_017894826.1: p.Arg204* | **5/1/0 (13/2/0)** | **0.69 (0.75)** | **246/1/0** | **1.00** | **237/0/0** | **1.00** |

[1]Absolute frequencies are ordered as follows: homozygous individuals for the wild-type allele or heterozygous individuals for the nonsense variant or homozygous individuals for the nonsense variant.

[2]CHIR = *Capra hircus* chromosome; HW *P*-value = statistical significance of the departure from the Hardy-Weinberg equilibrium.

[3]Data from the 6 sequenced bucks mated to dams and, between parentheses, from the 15 sequenced bucks (please see the "Animal Material" section for more details).

indicate that each individual typically differs from the reference human genome sequence by 80 to 200 nonsense variants, and in pigs such number is ~30 (Groenen et al., 2012). In general, the ability and accuracy to detect nonsense variants depend strongly on parameters, such as population size and composition, sequencing technology and coverage, variant-calling algorithms, filtering criteria, and the availability of a high-quality and well-annotated genome assembly. Usually, these parameters differ across studies making it difficult to establish meaningful comparisons between them (MacArthur et al., 2012). Another important factor influencing the abundance of deleterious variants is effective population size (Charlier et al., 2016). Indeed, the increase of effective population size leads to an increment of the average number of embryonic lethal variants harbored per individual (Charlier et al., 2016). Oliveira et al. (2016) reported that the effective number of founders of the Murciano-Granadina breed encompassed 967 individuals, whereas the most recent estimate (from 2006, same authors) of effective population size is ~100 individuals. According to simulations reported by Charlier et al. (2016), this would correspond to 0.53 ± 0.21 recessive lethals carried on average per individual.

We have observed that approximately one-fourth of the 947 nonsense variants segregating in the 15 buck genomes have frequencies below 5%. Extensive sequencing of human genomes revealed that 76% of the total variation has frequencies below 0.5%, and only 9.5% of the variants corresponded to common alleles (i.e., those with frequencies above 5%; 1000 Genomes Project Consortium, 2015). In addition, nonsense variants with harmful effects are expected to have lower frequencies than neutral alleles because they are purged by purifying selection. Our findings might be explained by the fact that we have sequenced a reduced set of individuals, making it very difficult to detect alleles with low frequencies. Indeed, the 1000 Genomes Project Consortium (2015) reported that the majority of variants identified in a single genome are common (only 1%–4% of the variants have a frequency below 0.5%), thus implying that common alleles (frequency ≥5%) are preferentially detected, over those with frequencies below 5%, when sequencing a small number of individuals.

A. Location of the XM_018049164.1:c.1270-1271 CA > GG variant in the *LOC108636199* locus. The position interrogated in the TaqMan assay is marked with a blue arrow.


B. Sanger sequencing of the region containing the position interrogated in the TaqMan assay (blue arrow) for the *LOC108636199* locus in three individuals with CC (left), CG (**center**) and GG (right) genotypes.

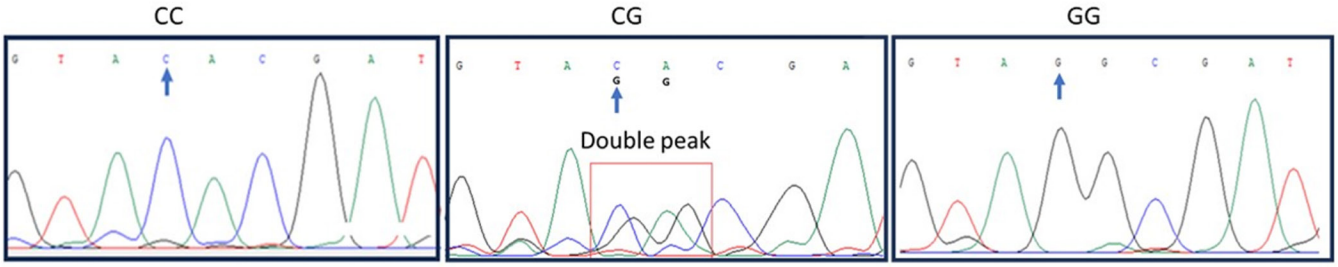


**Figure 2.** (A) Partial alignment of the chromosome 23 (*DAXX* gene) and 6 (*LOC108636199* locus) regions. The dinucleotide variant (XM_018049164.1:c.1270–1271 CA > GG) is marked with 2 red boxes. This variant is located in the *LOC108636199* locus, whereas the predicted XM_018038577.1:c.1134C > G variant is an artifact and position 1134 is, in reality, monomorphic. The position interrogated in the TaqMan assay is marked with a blue arrow. For this position, 3 genotypes (CC, CG and GG) are possible for the *LOC108636199* locus (XM_018049164.1:c.1270 C > G variant), whereas for the *DAXX* gene only one genotype (CC) can be found. If both loci are sequenced, the resulting genotypes can be only CG (23: CC + 6: CG or GG) or CC (23: CC + 6: CC). This results in an apparent complete depletion of GG individuals. (B) Electropherograms corresponding to the 3 potential genotypes for the *LOC108636199* dinucleotide variant. The position interrogated in the TaqMan assay is marked with a blue arrow.

Although nonsense variants were uniformly distributed in the coding region of goat genes, we have observed an enrichment in the 3′ end. According to MacArthur et al. (2012), nonsense variants and frameshift indels tend to accumulate toward the 3′ end of the affected genes due to a greater tolerance to truncation in such region. A substantial fraction of the nonsense variants segregating in the genomes of the 15 bucks were located in genes without essential functions (as predicted by LOEUF scores). Pathway analysis revealed that genes harboring nonsense variants were enriched in functions related to olfactory perception, such as G-protein coupled receptor activity and olfactory receptor activity. The mammalian olfactory receptor gene family constitutes the largest group of G-protein-coupled receptors (Spehr and Munger 2009), so these 2 pathways largely overlap. Because olfactory receptor genes evolve at a fast pace and do not have functions essential for organism survival, they tend to accumulate moderate-effect (e.g., missense substitutions) and high-effect (e.g., frameshift indels and nonsense variants) variants at relatively high rates (Jiménez et al., 2021). Moreover, pseudogenes are particularly abundant in the olfactory subgenomes of humans (Jiménez et al., 2021) and cattle (Lee et al., 2013), making it very difficult to discern real nonsense variants (causing protein truncation) from those variants introducing premature stop codons in pseudogenes that are not even transcribed. Pathway analysis also evidenced that nonsense variants were enriched in genes related to the viral defense response, autoimmune disease, and antigen processing, perhaps because immunity genes are generally more duplicated and less conserved than the genome average (Rausell et al., 2020)

### *Data About Knockout Mice Do Not Fully Recapitulate the Consequences of Nonsense Variants in Goats*

We genotyped 17 nonsense variants, that according to information from the OMIM and MGI databases were expected to have detrimental effects on biological viability, in buck-dam-offspring trios. Three variants in the *MECR* (XM_018057207.1:c.1152C > G, with the G-allele introducing a premature stop codon), *BRCA2* (XM_018056629.1:c.8754T > A, with the A-allele introducing a premature stop codon), and *NANOG* (NM_001285576.1:c.817C > T, with the T-allele in-

troducing a premature stop codon) genes displayed homozygous genotypes for the nonsense variant in dams as well as in their offspring, despite the fact that, according to data gathered from knockout mice, such genotypes are predicted to be lethal. For instance, the inactivation of the murine *MECR* gene leads to defects in mitochondrial respiration that are incompatible with normal embryo and placental development (Nair et al., 2017), and *BRCA2*-null embryos die before embryonic d 8.5 (Hakem et al., 1998) and *NANOG* loss of function is lethal at preimplantation stages due to the key role of this transcription factor in establishing embryonic pluripotency in the blastocyst (Sainz de Aja et al., 2019). Narasimhan et al. (2016) sequenced the exomes of 3,222 British adults of Pakistani descent and detected several individuals that carried variants predicted to be lethal in homozygosity, but that, unexpectedly, did not show any clinical sign. Even more, they identified one woman lacking the *PRMD9* gene, which in mice results in a sterility phenotype, who despite this genetic defect had children (Narasimhan et al., 2016). Indeed, dogs lack a functional *PRMD9* gene, but such loss does not have any adverse effect on reproduction indicating that another gene has assumed its function (Muñoz-Fuentes et al., 2011).

In our study, we hypothesize the existence of at least 2 mechanisms that might explain why variants with potentially lethal effects do not have such consequences. First, several of the nonsense variants (those in the *MECR*, *BRCA2*, *NANOG PPA2*, *MAP9*, and *SLC25A26* genes) investigated in the current work map to the last 10% of the coding region, and one of them (*BCL2* gene) to the last 20%. In theory, close to the end of the coding sequence has a greater tolerance for truncation (Mac Arthur et al., 2012), implying that in many instances truncation of the terminal part of the protein does not lead to its inactivation. However, this is not a universal rule. For instance, loss of the last 9 C-terminal AA of the solute carrier family 8 (sodium-calcium exchanger), member 1 (*SLC8A1*) did not lead to functional abrogation of the protein, but, in strong contrast, the deletion of the last 10 AA reduced $Na^+$ gradient-dependent $Ca^{2+}$ uptake to 35% to 39%, and additional truncation of 13 or more AA was even more damaging (Kasir et al., 1999). The second mechanism that might counteract the consequences of premature stop codons is mRNA isoform usage and specificity. As shown in Supplemental Figure S3, nonsense variants in the *MECR*, *TNRC6C*, *BCL2*, *MAP9*, *PPA2*, and *SLC25A26* genes were isoform-specific, implying that they do not necessarily abrogate protein function (unless different isoforms have also divergent functions). For instance, in the *MAP9* gene only 4 out of 6 mRNA isoforms harbor the nonsense variant, whereas in the *MECR* gene only one highly divergent mRNA isoform contains the stop gain variant (Supplemental Figure S3). In other cases (e.g., *NBAS* gene) all isoforms carry the premature stop codon (Supplemental Figure S3). It is very difficult to disentangle whether highly divergent mRNA isoforms are the product of a second copy of the gene or if they are functionally relevant because the annotation of the goat genome is still far from completion. The expression levels of mRNA isoforms are also unknown, making it difficult to evaluate the fraction of total functional mRNA levels affected by premature truncation.

An additional potential mechanism limiting the harmfulness of premature protein truncation might be gene redundancy. For instance, knockout mice for the *TP53* gene have an average time to tumor development of 4.5 mo (Donehower 1996), but such effect is not observed in elephants because they have several extra copies of this gene (Sulak et al., 2016). The adverse effects of nonsense variants can also be counteracted by suppressor variants (Matsui et al., 2017) and ribosomal readthrough (Li and Zhang 2019) that might differ across species. Finally, yet importantly, the quality of annotation, which also differs across species, is a critical factor in determining the accuracy with which loss-of-function variants are detected (MacArthur et al., 2012; Narasimhan et al., 2016).

### *A Putative Nonsense Variant in the DAXX Gene Is an Artifact Produced by the Co-amplification of a Duplicated Copy*

Six of the successfully genotyped nonsense variants showed a complete absence of homozygous genotypes but only one, located in the *DAXX* gene, showed a significant ($P < 0.00001$) depletion when testing Hardy-Weinberg equilibrium. The lack of significance of 5 of these 6 variants could be due to their very low frequencies, that make very unlikely the birth of individuals with homozygous genotypes, or to the fact that they are truly lethal. With the available information, it is very difficult to discern between these 2 alternative scenarios. As said, the nonsense variant in the *DAXX* gene was a very promising candidate to be lethal because the depletion of genotypes homozygous for the nonsense variant was highly significant ($P < 0.00001$). In addition, the loss of the *DAXX* gene results in extensive apoptosis and embryonic lethality in mice (Michaelson et al., 1999). However, manual curation of this variant revealed that it maps to a locus denominated *LOC108636199*, that contains a region perfectly duplicated in the *DAXX* gene, and the copy corresponding to the *LOC108636199* locus is the one harboring the XM_018049164.1:c.1270–1271 CA > GG dinucleotide variant. This means that when both loci co-amplify only 2 genotypes (CC and CG) can be generated. This result illustrates the remarkable complexity of correctly annotating nonsense variants (and other loss-of-function variants) and raises the possibility that many

putative variants predicted to have protein-truncating effects are, in reality, artifacts. Indeed, it is well known that loss-of-function variants are highly enriched in false positives (MacArthur et al., 2012). For instance, Narasimhan et al. (2016) detected 54 homozygous predicted loss-of-function genotypes in a cohort of Pakistani individuals, but as many as 16 of such variants were considered to be potential genome annotation errors.

### *Implications in the Context of the Cost of Domestication Hypothesis*

As previously explained, the cost of domestication hypothesis assumes that domestication involved an increased accumulation of deleterious variants due to the occurrence of serial bottlenecks that decreased the efficiency of purifying selection. This hypothesis has been experimentally tested in several animal species (Schubert et al., 2014; Marsden et al., 2016; Makino et al., 2018). A key limitation of these studies is that deleteriousness has been mostly assessed with in silico tools, making predictions exclusively based on evolutionary constraint (Schubert et al., 2014; Marsden et al., 2016) or genomic and functional annotations (Makino et al., 2018; Jagannathan et al., 2019). According to our results, such predictions should be considered with caution because the genomic and functional annotation of domestic species is still incomplete, leading, on some occasions, to flawed predictions. For instance, Mármol-Sánchez et al. (2020) identified pigs homozygous for a putative nonsense variant in the porcine arginine succinate synthase (*ASS1*) gene that was expected to abrogate the urea cycle and cause perinatal death. However, close inspection of this variant revealed that, in fact, it is located in the 5′UTR of the *ASS1* gene (Mármol-Sánchez et al., 2020). Results obtained by us for the goat *DAXX* gene, with an apparent significant depletion of individuals homozygous for the nonsense variant that happened to be an artifact, are also a good example of the limitations of the current methods to predict and assess deleteriousness. Thus, the generation of a credible set of deleterious variants, in goats or any other species, to test the "cost of domestication" hypothesis is extremely challenging, and it can only be achieved by generating a high-quality telomere-to-telomere goat reference genome combined with carefully curated annotation and functional testing of promising candidate variants (Bosse et al., 2019), rather than relying on results obtained in knockout mouse experiments.

## NOTES

This research was funded by grant PID2019-105805RB-I00 funded by MCIN/AEI/10.13039/501100011033 (Madrid, Spain), and by grant PID2022-136834OB-I00 funded by MCIIN/AEI/10.13039/501100011033 (Madrid, Spain) and "ERDF A way of making Europe." We also acknowledge the support of the CERCA Programme/Generalitat de Catalunya and of the Spanish Ministerio de Ciència e Innovación for the Center of Excellence Severo Ochoa 2020–2023 (CEX2019–000902-S) grant awarded to the Centre for Research in Agricultural Genomics (CRAG, Bellaterra, Spain). María Luigi-Sierra was funded with a PhD fellowship Formación de Personal Investigador (BES-2017-079709) awarded by the Ministerio de Economía y Competitividad (Madrid, Spain). This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement no. 945043 through the postdoctoral fellowship awarded to Antonia Noce by the AGenT H2020-MSCA-COFUND-2019 program. Antonia Noce was also awarded with a Beatriu de Pinós postdoctoral fellowship (2020 BP 00101) funded by the Secretaria d'Universitats i Recerca del Departament d'Empresa i Coneixement de la Generalitat de Catalunya (Barcelona, Spain). Taina Figueiredo-Cardoso was awarded by the Universitat Autònoma de Barcelona a María Zambrano postdoctoral fellowship funded by the European Union-NextGeneration EU and managed under the scope of a program organized by the Ministerio de Universidades (Madrid, Spain) to promote the requalification and international mobility in the Spanish university system (Royal Decree 289/2021). Many thanks to CAPRIGRAN (https://caprigran.com/) for carrying out phenotype recording and blood sample collection from Murciano-Granadina goats. We are also grateful to the VarGoats Consortium (https://www.goatgenome.org/vargoats.html) for giving us access to their data set with the goal of analyzing the segregation of a set of variants in several goat breeds. The raw fastq files have been uploaded to SRA, and they can be accessed through BioProject *PRJNA1113629* at the National Center for Biotechnology Information (NCBI) Sequence Read Archive. Supplemental material for this article is available at https://doi.org/10.6084/m9.figshare.24580870). The collection of blood is a routine procedure carried out by trained veterinarians working for the National Association of Murciano-Granadina Goat Breeders (CAPRIGRAN), so it did not require permission from the Committee on Ethics in Animal and Human Experimentation of the Universitat Autònoma de Barcelona. Permission to extract ear samples was granted by the Review Committee for the Use of Animal Subjects of Northwest A&F University (contract number NWAFAC1008). The authors have not stated any conflicts of interest.

**Nonstandard abbreviations used:** BAM = Binary Alignment Map; CDS = coding sequence; GO = Gene

ontology; KEGG = Kyoto Encyclopedia of Genes and Genomes; LOEUF = the loss-of-function observed/expected upper bound fraction; MACARON = multibase codon-associated variant re-annotation; MAF = major allele frequency; MGI = Mouse Genome Informatics; OMIM: Online Mendelian Inheritance in Man.

## REFERENCES


1000 Genomes Project Consortium. 2012. An integrated map of genetic variation from 1,092 human genomes. Nature 491:56–65. https://doi.org/10.1038/nature11632.

1000 Genomes Project Consortium. 2015. A global reference for human genetic variation. Nature 526:68–74. https://doi.org/10.1038/nature15393.

Abramowicz, A., and M. Gos. 2018. Splicing mutations in human genetic disorders: Examples, detection, and confirmation. J. Appl. Genet. 59:253–268. https://doi.org/10.1007/s13353-018-0444-7.

Baldarelli, R. M., C. M. Smith, J. H. Finger, T. F. Hayamizu, I. J. McCright, J. Xu, D. R. Shaw, J. S. Beal, O. Blodgett, J. Campbell, L. E. Corbani, P. J. Frost, S. C. Giannatto, D. B. Miers, J. A. Kadin, J. E. Richardson, and M. Ringwald. 2021. The mouse gene expression database (GXD): 2021 update. Nucleic Acids Res. 49(D1):D924–D931. https://doi.org/10.1093/nar/gkaa914.

Bi, X., Q. Zhang, L. Chen, D. Liu, Y. Li, X. Zhao, Y. Zhang, L. Zhang, J. Liu, C. Wu, Z. Li, Y. Zhao, H. Ma, G. Huang, X. Liu, Q. F. Wang, and R. Zhang. 2022. NBAS, a gene involved in cytotoxic degranulation, is recurrently mutated in pediatric hemophagocytic lymphohistiocytosis. J. Hematol. Oncol. 15:101. https://doi.org/10.1186/s13045-022-01318-z.

Bickhart, D. M., B. D. Rosen, S. Koren, B. L. Sayre, A. R. Hastie, S. Chan, J. Lee, E. T. Lam, I. Liachko, S. T. Sullivan, J. N. Burton, H. J. Huson, J. C. Nystrom, C. M. Kelley, J. L. Hutchison, Y. Zhou, J. Sun, A. Crisà, F. A. Ponce de León, J. C. Schwartz, J. A. Hammond, G. C. Waldbieser, S. G. Schroeder, G. E. Liu, M. J. Dunham, J. Shendure, T. S. Sonstegard, A. M. Phillippy, C. P. van Tassell, and T. P. Smith. 2017. Single-molecule sequencing and chromatin conformation capture enable de novo reference assembly of the domestic goat genome. Nat. Genet. 49:643–650. https://doi.org/10.1038/ng.3802.

Bosse, M., H. J. Megens, M. F. L. Derks, Á. M. R. de Cara, and M. A. M. Groenen. 2019. Deleterious alleles in the context of domestication, inbreeding, and selection. Evol. Appl. 12:6–17. https://doi.org/10.1111/eva.12691.

Bu, D., H. Luo, P. Huo, Z. Wang, S. Zhang, Z. He, Y. Wu, L. Zhao, J. Liu, J. Guo, S. Fang, W. Cao, L. Yi, Y. Zhao, and L. Kong. 2021. KOBAS-i: Intelligent prioritization and exploratory visualization of biological functions for gene enrichment analysis. Nucleic Acids Res. 49(W1):W317–W325. https://doi.org/10.1093/nar/gkab447.

Charlier, C., W. Li, C. Harland, M. Littlejohn, W. Coppieters, F. Creagh, S. Davis, T. Druet, P. Faux, F. Guillaume, L. Karim, M. Keehan, N. K. Kadri, N. Tamma, R. Spelman, and M. Georges. 2016. NGS-based reverse genetic screen for common embryonic lethal mutations compromising fertility in livestock. Genome Res. 26:1333–1341. https://doi.org/10.1101/gr.207076.116.

Cingolani, P., A. Platts, L. L. Wang, M. Coon, T. Nguyen, L. Wang, S. J. Land, X. Lu, and D. M. Ruden. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; *iso*-2; *iso*-3. Fly 6:80–92. https://doi.org/10.4161/fly.19695.

Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, and R. Durbin.1000 Genomes Project Analysis Group. 2011. The variant call format and VCFtools. Bioinformatics 27:2156–2158. https://doi.org/10.1093/bioinformatics/btr330.

Denoyelle, L., E. Talouarn, P. Bardou, L. Colli, A. Alberti, C. Danchin, M. Del Corvo, S. Engelen, C. Orvain, I. Palhière, R. Rupp, J. Sarry, M. Salavati, M. Amills, E. Clark, P. Crepaldi, T. Faraut, C. W. Masiga, F. Pompanon, B. D. Rosen, A. Stella, C. P. Van Tassell, G. Tosser-Klopp, J. Kijas, B. Guldbrandtsen, J. Kantanen, D. Duby, P. Martin, C. Danchin, D. Duclos, D. Allain, R. Arquet, N. Mandonnet, M. Naves, I. Palhière, R. Rupp, C. A. B. R. I. C. O. O. P. breeders, F. Pompanon, H. R. Rezaei, S. Carolan, M. Foran, A. Stella, P. Ajmone-Marsan, L. Colli, A. Crisà, D. Marletta, P. Crepaldi, M. Ottino, E. Randi, B. Benjelloun, H. Lenstra, M. Moaeen-ud-Din, J. Reecy, F. Goyache, I. Alvarez, M. Amills, A. Sànchez, J. Capote, J. Jordana, A. Pons, A. Martínez, A. Molina, B. Rosen, C. Visser, C. Drögemüller, G. Luikart, C. W. Masiga, D. F. Mujibi, H. A. Mruttu, T. Gondwe, J. Sikosana, M. T. Da Gloria, and O. Nash. 2021. VarGoats project: A dataset of 1159 whole-genome sequences to dissect *Capra hircus* global diversity. Genet. Sel. Evol. 53:86. https://doi.org/10.1186/s12711-021-00659-6.

Derks, M. F. L., A. B. Gjuvsland, M. Bosse, M. S. Lopes, M. van Son, B. Harlizius, B. F. Tan, H. Hamland, E. Grindflek, M. A. M. Groenen, and H. J. Megens. 2019. Loss of function mutations in essential genes cause embryonic lethality in pigs. PLoS Genet. 15:e1008055. https://doi.org/10.1371/journal.pgen.1008055.

Donehower, L. A. 1996. The p53-deficient mouse: A model for basic and applied cancer studies. Semin. Cancer Biol. 7:269–278. https://doi.org/10.1006/scbi.1996.0035.

El-Brolosy, M. A., and D. Y. R. Stainier. 2017. Genetic compensation: A phenomenon in search of mechanisms. PLoS Genet. 13:e1006780. https://doi.org/10.1371/journal.pgen.1006780.

González-Prendes, R., M. F. L. Derks, M. A. M. Groenen, R. Quintanilla, and M. Amills. 2023. Assessing the relationship between the in silico predicted consequences of 97 missense mutations mapping to 68 genes related to lipid metabolism and their association with porcine fatness traits. Genomics 115:110589. https://doi.org/10.1016/j.ygeno.2023.110589.

Groenen, M. A. M., A. L. Archibald, H. Uenishi, C. K. Tuggle, Y. Takeuchi, M. F. Rothschild, C. Rogel-Gaillard, C. Park, D. Milan, H.-J. Megens, S. Li, D. M. Larkin, H. Kim, L. A. F. Frantz, M. Caccamo, H. Ahn, B. L. Aken, A. Anselmo, C. Anthon, L. Auvil, B. Badaoui, C. W. Beattie, C. Bendixen, D. Berman, F. Blecha, J. Blomberg, L. Bolund, M. Bosse, S. Botti, Z. Bujie, M. Bystrom, B. Capitanu, D. Carvalho-Silva, P. Chardon, C. Chen, R. Cheng, S.-H. Choi, W. Chow, R. C. Clark, C. Clee, R. P. M. A. Crooijmans, H. D. Dawson, P. Dehais, F. De Sapio, B. Dibbits, N. Drou, Z.-Q. Du, K. Eversole, J. Fadista, S. Fairley, T. Faraut, G. J. Faulkner, K. E. Fowler, M. Fredholm, E. Fritz, J. G. R. Gilbert, E. Giuffra, J. Gorodkin, D. K. Griffin, J. L. Harrow, A. Hayward, K. Howe, Z.-L. Hu, S. J. Humphray, T. Hunt, H. Hornshøj, J.-T. Jeon, P. Jern, M. Jones, J. Jurka, H. Kanamori, R. Kapetanovic, J. Kim, J.-H. Kim, K.-W. Kim, T.-H. Kim, G. Larson, K. Lee, K.-T. Lee, R. Leggett, H. A. Lewin, Y. Li, W. Liu, J. E. Loveland, Y. Lu, J. K. Lunney, J. Ma, O. Madsen, K. Mann, L. Matthews, S. McLaren, T. Morozumi, M. P. Murtaugh, J. Narayan, D. Truong Nguyen, P. Ni, S.-J. Oh, S. Onteru, F. Panitz, E.-W. Park, H.-S. Park, G. Pascal, Y. Paudel, M. Perez-Enciso, R. Ramirez-Gonzalez, J. M. Reecy, S. Rodriguez-Zas, G. A. Rohrer, L. Rund, Y. Sang, K. Schachtschneider, J. G. Schraiber, J. Schwartz, L. Scobie, C. Scott, S. Searle, B. Servin, B. R. Southey, G. Sperber, P. Stadler, J. V. Sweedler, H. Tafer, B. Thomsen, R. Wali, J. Wang, J. Wang, S. White, X. Xu, M. Yerle, G. Zhang, J. Zhang, J. Zhang, S. Zhao, J. Rogers, C. Churcher, and L. B. Schook. 2012. Analyses of pig genomes provide insight into porcine demography and evolution. Nature 491:393–398. https://doi.org/10.1038/nature11622.

Gu, L., Q. He, W. Xia, G. Rong, D. Wang, M. Li, F. Ji, W. Sun, T. Cao, H. Zhou, and T. Xu. 2022. Integrated analysis of lncRNA and gene expression in *longissimus dorsi* muscle at two developmental stages of Hainan black goats. PLoS One 17:e0276004. https://doi.org/10.1371/journal.pone.0276004.

Hakem, R., J. L. D. L. Pompa, and T. W. Mak. 1998. Developmental studies of Brca1 and Brca2 knock-out mice. J. Mammary Gland Biol. Neoplasia 3:431–445. https://doi.org/10.1023/A:1018792200700.

Hamosh, A., A. F. Scott, J. S. Amberger, C. A. Bocchini, and V. A. McKusick. 2005. Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. Nucleic Acids Res. 33:D514–D517. https://doi.org/10.1093/nar/gki033.

Jagannathan, V., C. Drögemüller, and T. Leeb.Dog Biomedical Variant Database Consortium. 2019. A comprehensive biomedical variant catalogue based on whole genome sequences of 582 dogs and

eight wolves. Anim. Genet. 50:695–704. https://doi.org/10.1111/age.12834.

Jiménez, R. C., N. Casajuana-Martin, A. García-Recio, L. Alcántara, L. Pardo, M. Campillo, and A. Gonzalez. 2021. The mutational landscape of human olfactory G protein-coupled receptors. BMC Biol. 19:21. https://doi.org/10.1186/s12915-021-00962-0.

Karczewski, K. J., L. C. Francioli, G. Tiao, B. B. Cummings, J. Alföldi, Q. Wang, R. L. Collins, K. M. Laricchia, A. Ganna, D. P. Birnbaum, L. D. Gauthier, H. Brand, M. Solomonson, N. A. Watts, D. Rhodes, M. Singer-Berk, E. M. England, E. G. Seaby, J. A. Kosmicki, R. K. Walters, K. Tashman, Y. Farjoun, E. Banks, T. Poterba, A. Wang, C. Seed, N. Whiffin, J. X. Chong, K. E. Samocha, E. Pierce-Hoffman, Z. Zappala, A. H. O'Donnell-Luria, E. V. Minikel, B. Weisburd, M. Lek, J. S. Ware, C. Vittal, I. M. Armean, L. Bergelson, K. Cibulskis, K. M. Connolly, M. Covarrubias, S. Donnelly, S. Ferriera, S. Gabriel, J. Gentry, N. Gupta, T. Jeandet, D. Kaplan, C. Llanwarne, R. Munshi, S. Novod, N. Petrillo, D. Roazen, V. Ruano-Rubio, A. Saltzman, M. Schleicher, J. Soto, K. Tibbetts, C. Tolonen, G. Wade, M. E. Talkowski, C. A. Aguilar Salinas, T. Ahmad, C. M. Albert, D. Ardissino, G. Atzmon, J. Barnard, L. Beaugerie, E. J. Benjamin, M. Boehnke, L. L. Bonnycastle, E. P. Bottinger, D. W. Bowden, M. J. Bown, J. C. Chambers, J. C. Chan, D. Chasman, J. Cho, M. K. Chung, B. Cohen, A. Correa, D. Dabelea, M. J. Daly, D. Darbar, R. Duggirala, J. Dupuis, P. T. Ellinor, R. Elosua, J. Erdmann, T. Esko, M. Färkkilä, J. Florez, A. Franke, G. Getz, B. Glaser, S. J. Glatt, D. Goldstein, C. Gonzalez, L. Groop, C. Haiman, C. Hanis, M. Harms, M. Hiltunen, M. M. Holi, C. M. Hultman, M. Kallela, J. Kaprio, S. Kathiresan, B.-J. Kim, Y. J. Kim, G. Kirov, J. Kooner, S. Koskinen, H. M. Krumholz, S. Kugathasan, S. H. Kwak, M. Laakso, T. Lehtimäki, R. J. F. Loos, S. A. Lubitz, R. C. W. Ma, D. G. MacArthur, J. Marrugat, K. M. Mattila, S. McCarroll, M. I. McCarthy, D. McGovern, R. McPherson, J. B. Meigs, O. Melander, A. Metspalu, B. M. Neale, P. M. Nilsson, M. C. O'Donovan, D. Ongur, L. Orozco, M. J. Owen, C. N. A. Palmer, A. Palotie, K. S. Park, C. Pato, A. E. Pulver, N. Rahman, A. M. Remes, J. D. Rioux, S. Ripatti, D. M. Roden, D. Saleheen, V. Salomaa, N. J. Samani, J. Scharf, H. Schunkert, M. B. Shoemaker, P. Sklar, H. Soininen, H. Sokol, T. Spector, P. F. Sullivan, J. Suvisaari, E. S. Tai, Y. Y. Teo, T. Tiinamaija, M. Tsuang, D. Turner, T. Tusie-Luna, E. Vartiainen, M. P. Vawter, J. S. Ware, H. Watkins, R. K. Weersma, M. Wessman, J. G. Wilson, R. J. Xavier, B. M. Neale, M. J. Daly, and D. G. MacArthur. 2020. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 581:434–443. https://doi.org/10.1038/s41586-020-2308-7.

Kasir, J., X. Ren, I. Furman, and H. Rahamimoff. 1999. Truncation of the C terminus of the rat brain $Na^{+}$-$Ca^{2+}$ exchanger RBE-1 (NCX1.4) impairs surface expression of the protein. J. Biol. Chem. 274:24873–24880. https://doi.org/10.1074/jbc.274.35.24873.

Kennedy, H., T. B. Haack, V. Hartill, L. Mataković, E. R. Baumgartner, H. Potter, R. Mackay, C. L. Alston, S. O'Sullivan, R. McFarland, G. Connolly, C. Gannon, R. King, S. Mead, I. Crozier, W. Chan, C. M. Florkowski, M. Sage, T. Höfken, B. Alhaddad, L. S. Kremer, R. Kopajtich, R. G. Feichtinger, W. Sperl, R. J. Rodenburg, J. C. Minet, A. Dobbie, T. M. Strom, T. Meitinger, P. M. George, C. A. Johnson, R. W. Taylor, H. Prokisch, K. Doudney, and J. A. Mayr. 2016. Sudden cardiac death due to deficiency of the mitochondrial inorganic pyrophosphatase PPA2. Am. J. Hum. Genet. 99:674–682. https://doi.org/10.1016/j.ajhg.2016.06.027.

Khan, W., G. Varma Saripella, T. Ludwig, T. Cuppens, F. Thibord, E. Génin, J. F. Deleuze, and D. A. Trégouët. 2018. MACARON: A python framework to identify and re-annotate multi-base affected codons in whole genome/exome sequence data. Bioinformatics 34:3396–3398. https://doi.org/10.1093/bioinformatics/bty382.

Lee, K., D. T. Nguyen, M. Choi, S. Y. Cha, J. H. Kim, H. Dadi, H. G. Seo, K. Seo, T. Chun, and C. Park. 2013. Analysis of cattle olfactory subgenome: The first detail study on the characteristics of the complete olfactory receptor repertoire of a ruminant. BMC Genomics 14:596. https://doi.org/10.1186/1471-2164-14-596.

Li, C., and J. Zhang. 2019. Stop-codon read-through arises largely from molecular errors and is generally nonadaptive. PLoS Genet. 15:e1008141. https://doi.org/10.1371/journal.pgen.1008141.

Li, H., and R. Durbin. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754–1760. https://doi.org/10.1093/bioinformatics/btp324.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, and R. Durbin.1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–2079. https://doi.org/10.1093/bioinformatics/btp352.

MacArthur, D. G., S. Balasubramanian, A. Frankish, N. Huang, J. Morris, K. Walter, L. Jostins, L. Habegger, J. K. Pickrell, S. B. Montgomery, C. A. Albers, Z. D. Zhang, D. F. Conrad, G. Lunter, H. Zheng, Q. Ayub, M. A. DePristo, E. Banks, M. Hu, R. E. Handsaker, J. A. Rosenfeld, M. Fromer, M. Jin, X. Mu, E. Khurana, K. Ye, M. Kay, G. I. Saunders, M. M. Suner, T. Hunt, I. H. Barnes, C. Amid, D. R. Carvalho-Silva, A. H. Bignell, C. Snow, B. Yngvadottir, S. Bumpstead, D. N. Cooper, Y. Xue, I. G. Romero, J. Wang, Y. Li, R. A. Gibbs, S. A. McCarroll, E. T. Dermitzakis, J. K. Pritchard, J. C. Barrett, J. Harrow, M. E. Hurles, M. B. Gerstein, and C. Tyler-Smith. 2012. A systematic survey of loss-of-function variants in human protein-coding genes. Science 335:823–828. https://doi.org/10.1126/science.1215040.

MacArthur, D. G., and C. Tyler-Smith. 2010. Loss-of-function variants in the genomes of healthy humans. Hum. Mol. Genet. 19(R2):R125–R130. https://doi.org/10.1093/hmg/ddq365.

Majzoub, J. A., and L. J. Muglia. 1996. Knockout mice. N. Engl. J. Med. 334:904–906. https://doi.org/10.1056/NEJM199604043341407.

Makino, T., C. J. Rubin, M. Carneiro, E. Axelsson, L. Andersson, and M. T. Webster. 2018. Elevated proportions of deleterious genetic variation in domestic animals and plants. Genome Biol. Evol. 10:276–290. https://doi.org/10.1093/gbe/evy004.

Mármol-Sánchez, E., M. G. Luigi-Sierra, R. Quintanilla, and M. Amills. 2020. Detection of homozygous genotypes for a putatively lethal recessive mutation in the porcine argininosuccinate synthase 1 (*ASS1*) gene. Anim. Genet. 51:106–110. https://doi.org/10.1111/age.12877.

Marsden, C. D., D. Ortega-Del Vecchyo, D. P. O'Brien, J. F. Taylor, O. Ramirez, C. Vilà, T. Marques-Bonet, R. D. Schnabel, R. K. Wayne, and K. E. Lohmueller. 2016. Bottlenecks and selective sweeps during domestication have increased deleterious genetic variation in dogs. Proc. Natl. Acad. Sci. USA 113:152–157. https://doi.org/10.1073/pnas.1512501113.

Matsui, T., J. T. Lee, and I. M. Ehrenreich. 2017. Genetic suppression: Extending our knowledge from lab experiments to natural populations. BioEssays 39:1700023 https://doi.org/10.1002/bies.201700023.

McKenna, A., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, and M. A. DePristo. 2010. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20:1297–1303. https://doi.org/10.1101/gr.107524.110.

McLaren, W., L. Gil, S. E. Hunt, H. S. Riat, G. R. Ritchie, A. Thormann, P. Flicek, and F. Cunningham. 2016. The Ensembl variant effect predictor. Genome Biol. 17:122. https://doi.org/10.1186/s13059-016-0974-4.

Michaelson, J. S., D. Bader, F. Kuo, C. Kozak, and P. Leder. 1999. Loss of Daxx, a promiscuously interacting protein, results in extensive apoptosis in early mouse development. Genes Dev. 13:1918–1923. https://doi.org/10.1101/gad.13.15.1918.

Mort, M., D. Ivanov, D. N. Cooper, and N. A. Chuzhanova. 2008. A meta-analysis of nonsense variants causing human genetic disease. Hum. Mutat. 29:1037–1047. https://doi.org/10.1002/humu.20763.

Moyers, B. T., P. L. Morrell, and J. K. McKay. 2018. Genetic costs of domestication and improvement. J. Hered. 109:103–116. https://doi.org/10.1093/jhered/esx069.

Mullican, S. E., S. Zhang, M. Konopleva, V. Ruvolo, M. Andreeff, J. Milbrandt, and O. M. Conneely. 2007. Abrogation of nuclear receptors Nr4a3 and Nr4a1 leads to development of acute myeloid leukemia. Nat. Med. 13:730–735. https://doi.org/10.1038/nm1579.

Muñoz-Fuentes, V., A. Di Rienzo, and C. Vilà. 2011. *Prdm9*, a major determinant of meiotic recombination hotspots, is not functional in dogs and their wild relatives, wolves and coyotes. PLoS One 6:e25498. https://doi.org/10.1371/journal.pone.0025498.

Nair, R. R., J. M. Kerätär, K. J. Autio, A. J. Masud, M. A. J. Finnilä, H. I. Autio-Harmainen, I. J. Miinalainen, P. A. Nieminen, J. K. Hiltunen, and A. J. Kastaniotis. 2017. Genetic modifications of *Mecr* reveal a role for mitochondrial 2-enoyl-CoA/ACP reductase in placental development in mice. Hum. Mol. Genet. 26:2104–2117. https://doi.org/10.1093/hmg/ddx105.

Narasimhan, V. M., K. A. Hunt, D. Mason, C. L. Baker, K. J. Karczewski, M. R. Barnes, A. H. Barnett, C. Bates, S. Bellary, N. A. Bockett, K. Giorda, C. J. Griffiths, H. Hemingway, Z. Jia, M. A. Kelly, H. A. Khawaja, M. Lek, S. McCarthy, R. McEachan, A. O'Donnell-Luria, K. Paigen, C. A. Parisinos, E. Sheridan, L. Southgate, L. Tee, M. Thomas, Y. Xue, M. Schnall-Levin, P. M. Petkov, C. Tyler-Smith, E. R. Maher, R. C. Trembath, D. G. MacArthur, J. Wright, R. Durbin, and D. A. van Heel. 2016. Health and population effects of rare gene knockouts in adult humans with related parents. Science 352:474–477. https://doi.org/10.1126/science.aac8624.

Oliveira, R. R., L. H. A. Brasil, J. V. Delgado, J. Peguezuelos, J. M. León, D. G. P. Guedes, J. K. G. Arandas, and M. N. Ribeiro. 2016. Genetic diversity and population structure of the Spanish Murciano–Granadina goat breed according to pedigree data. Small Rumin. Res. 144:170–175. https://doi.org/10.1016/j.smallrumres.2016.09.014.

Pelak, K., K. V. Shianna, D. Ge, J. M. Maia, M. Zhu, J. P. Smith, E. T. Cirulli, J. Fellay, S. P. Dickson, C. E. Gumbs, E. L. Heinzen, A. C. Need, E. K. Ruzzo, A. Singh, C. R. Campbell, L. K. Hong, K. A. Lornsen, A. M. McKenzie, N. L. Sobreira, J. E. Hoover-Fong, J. D. Milner, R. Ottman, B. F. Haynes, J. J. Goedert, and D. B. Goldstein. 2010. The characterization of twenty sequenced human genomes. PLoS Genet. 6:e1001111. https://doi.org/10.1371/journal.pgen.1001111.

Raudvere, U., L. Kolberg, I. Kuzmin, T. Arak, P. Adler, H. Peterson, and J. Vilo. 2019. g:Profiler: A web server for functional enrichment analysis and conversions of gene lists (2019 update). Nucleic Acids Res. 47(W1):W191–W198. https://doi.org/10.1093/nar/gkz369.

Rausell, A., Y. Luo, M. Lopez, Y. Seeleuthner, F. Rapaport, A. Favier, P. D. Stenson, D. N. Cooper, E. Patin, J. L. Casanova, L. Quintana-Murci, and L. Abel. 2020. Common homozygosity for predicted loss-of-function variants reveals both redundant and advantageous effects of dispensable human genes. Proc. Natl. Acad. Sci. USA 117:13626–13636. https://doi.org/10.1073/pnas.1917993117.

Robinson, J. T., H. Thorvaldsdóttir, A. M. Wenger, A. Zehir, and J. P. Mesirov. 2017. Variant review with the integrative genomics viewer. Cancer Res. 77:e31–e34. https://doi.org/10.1158/0008-5472.CAN-17-0337.

Sainz de Aja, J., S. Menchero, I. Rollan, A. Barral, M. Tiana, W. Jawaid, I. Cossio, A. Alvarez, G. Carreño-Tarragona, C. Badia-Careaga, J. Nichols, B. Göttgens, J. Isern, and M. Manzanares. 2019. The pluripotency factor NANOG controls primitive hematopoiesis and directly regulates *Tal1*. EMBO J. 38:e99122. https://doi.org/10.15252/embj.201899122.

Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463–5467. https://doi.org/10.1073/pnas.74.12.5463.

Schubert, M., H. Jónsson, D. Chang, C. Der Sarkissian, L. Ermini, A. Ginolhac, A. Albrechtsen, I. Dupanloup, A. Foucal, B. Petersen, M. Fumagalli, M. Raghavan, A. Seguin-Orlando, T. S. Korneliussen, A. M. Velazquez, J. Stenderup, C. A. Hoover, C. J. Rubin, A. H. Alfarhan, S. A. Alquraishi, K. A. Al-Rasheid, D. E. MacHugh, T. Kalbfleisch, J. N. MacLeod, E. M. Rubin, T. Sicheritz-Ponten, L. Andersson, M. Hofreiter, T. Marques-Bonet, M. T. Gilbert, R. Nielsen, L. Excoffier, E. Willerslev, B. Shapiro, and L. Orlando. 2014. Prehistoric genomes reveal the genetic foundation and cost of horse domestication. Proc. Natl. Acad. Sci. USA 111:E5661–E5669. https://doi.org/10.1073/pnas.1416991111.

Spehr, M., and S. D. Munger. 2009. Olfactory receptors: G protein-coupled receptors and beyond. J. Neurochem. 109:1570–1583. https://doi.org/10.1111/j.1471-4159.2009.06085.x.

Sulak, M., L. Fong, K. Mika, S. Chigurupati, L. Yon, N. P. Mongan, R. D. Emes, and V. J. Lynch. 2016. *TP53* copy number expansion is associated with the evolution of increased body size and an enhanced DNA damage response in elephants. eLife 5:e11994. https://doi.org/10.7554/eLife.11994.

Thorvaldsdóttir, H., J. T. Robinson, and J. P. Mesirov. 2013. Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. Brief. Bioinform. 14:178–192. https://doi.org/10.1093/bib/bbs017.

## ORCIDS

Ke Wang https://orcid.org/0000-0003-3400-3608
María Gracia Luigi-Sierra https://orcid.org/0000-0002-6414-0550
Anna Castelló https://orcid.org/0000-0001-8497-6251
Juan Vicente Delgado https://orcid.org/0000-0003-1657-8838
Javier Fernández Álvarez https://orcid.org/0000-0003-0450-5869
Antonia Noce https://orcid.org/0000-0003-1385-3187
Mingjing Wang https://orcid.org/0000-0002-2443-6840
Jordi Jordana https://orcid.org/0000-0002-7789-2989
Marcel Amills https://orcid.org/0000-0002-8999-0770