

RESEARCH ARTICLE

Open Access



Autozygosity islands and ROH patterns in Nellore lineages: evidence of selection for functionally important traits

Elisa Peripolli^{1*} , Julia Metzger², Marcos Vinícius Antunes de Lemos¹, Nedenia Bonvino Stafuzza³, Sabrina Kluska¹, Bianca Ferreira Olivieri¹, Fabieli Louise Braga Feitosa¹, Mariana Piatto Berton¹, Fernando Brito Lopes¹, Danísio Prado Munari³, Raysildo Barbosa Lôbo⁴, Cláudio de Ulhoa Magnabosco⁵, Fernando Di Croce⁶, Jason Osterstock⁶, Sue Denise⁶, Angélica Simone Cravo Pereira⁷ and Fernando Baldi¹

Abstract

Background: The aim of this study was to assess genome-wide autozygosity in a Nellore cattle population and to characterize ROH patterns and autozygosity islands that may have occurred due to selection within its lineages. It attempts also to compare estimates of inbreeding calculated from ROH (F_{ROH}), genomic relationship matrix (F_{GRM}), and pedigree-based coefficient (F_{PED}).

Results: The average number of ROH per animal was 55.15 ± 13.01 with an average size of 3.24 Mb. The Nellore genome is composed mostly by a high number of shorter segments accounting for 78% of all ROH, although the proportion of the genome covered by them was relatively small. The genome autozygosity proportion indicates moderate to high inbreeding levels for classical standards, with an average value of 7.15% (178.70 Mb). The average of F_{PED} and F_{ROH} , and their correlations (-0.05 to 0.26) were low. Estimates of correlation between F_{GRM} - F_{PED} was zero, while the correlation (-0.01 to -0.07) between F_{GRM} - F_{ROH} decreased as a function of ROH length, except for $F_{ROH > 8Mb}$ (-0.03). Overall, inbreeding coefficients were not high for the genotyped animals. Autozygosity islands were evident across the genome ($n = 62$) and their genomic location did not largely differ within lineages. Enriched terms ($p < 0.01$) associated with defense response to bacteria (GO:0042742), immune complex reaction (GO:0045647), pregnancy-associated glycoproteins genes (GO:0030163), and organism growth (GO:0040014) were described within the autozygotic islands.

Conclusions: Low F_{PED} - F_{ROH} correlation estimates indicate that F_{PED} is not the most suitable method for capturing ancient inbreeding when the pedigree does not extend back many generations and F_{ROH} should be used instead. Enriched terms ($p < 0.01$) suggest a strong selection for immune response. Non-overlapping islands within the lineages greatly explain the mechanism underlying selection for functionally important traits in Nellore cattle.

Keywords: *Bos indicus*, Indicine, Genomic inbreeding, Gene ontology

* Correspondence: elisa_peripolli@hotmail.com

¹Departamento de Zootecnia, Faculdade de Ciências Agrárias e Veterinárias, UNESP Univ Estadual Paulista Júlio de Mesquita Filho, Jaboticabal 14884-900, Brazil

Full list of author information is available at the end of the article



Background

Brazilian livestock and agriculture production have a prominent impact upon the world's food commerce. Brazilian beef production is one of the largest players in the world and produced roughly 9.56 million tons of carcass weight equivalents in 2015 [1]. The vast majority of the bovine based population reared for meat production in Brazil is composed mostly of indicine cattle (*Bos taurus indicus*). According to the Brazilian Zebu Breeders Association (ABCZ, <http://www.abcz.com.br>) such population is around 80% of the total cattle. Given the physical and physiological characteristics that they possess which greatly explain their better adaptation towards grazing systems in tropical environments [2–4], it is not surprisingly that much use of the indicine cattle has been made in these regions.

The Nellore breed has the largest number of animals (horned and polled) among the indicine cattle raised in Brazil, followed by Guzerat and Gyr. Most of Nellore importation was from India during the last century and lasted up to the seventies when the importation was banned [5]. The Nellore population in Brazil is the result of less than 7000 heads of purebred imported animals [6]. The major importation took place in 1962, when exceptional bulls were brought over the country standing out as progenitors of the main Nellore lineages [7]. Magnabosco et al. [8] reported the existence of six predominant lineages of Nellore breed (Karvadi Imp; Taj Mahal Imp; Kurupathy Imp; Golias Imp; Godhavari Imp, and Rastā Imp) that contributed to the development of the current Brazilian Nellore population. These lineages were derived from outstanding bulls named Karvadi, Taj Mahal, Kurupathy, Golias, Godhavari and Rastā which gained fame as breeders given their high rates of productive and reproductive performance [7]. Although the selection criteria used to improve the Nellore cattle among Brazilian breeding programs are closely linked and mainly associated with reproductive and carcass quality traits, there is evidence of different genetic patterns among the lineages based on the selection criterion used to improve each of them over time [9, 10]. In this manner, a question can be raised whether the genetic progress is going or not towards the same direction within the lineages raised in Brazil.

Genetic evaluations of Nellore cattle using BLUP (Best Linear Unbiased Prediction) methodology have established significant progress since the eighties, when several genetic evaluation programs started to expand in Brazil [11]. Despite the reduced number of animals imported from India, Pereira et al. [12] have reported an average inbreeding coefficient of 3% in a Nellore population, indicating that these animals have been under relative control for at least three decades. Therefore, breeding programs are always seeking for strategies to

preserve populations, and there is a growing interest in characterizing and monitoring genome-wide autozygosity to maintain the genetic diversity [13, 14], allowing a long-term conservation of genetic resources and sustainability in animal breeding programs.

Runs of homozygosity (ROH) have been widely applied to quantify individual autozygosity in livestock [15–20] given their high correlation (~ 0.7) [21]. A small number of studies have described the autozygosity in Nellore cattle and most of them do not make use of a large sample size. Karimi [22] identified region patterns with a high prevalence of ROH in taurine and indicine breeds and made use of merely 134 Nellore samples. Additionally, Zavarez et al. [19] reported the distribution of genome-wide autozygosity levels based on ROH in only 1278 Nellore cows genotyped for over 777,000 markers.

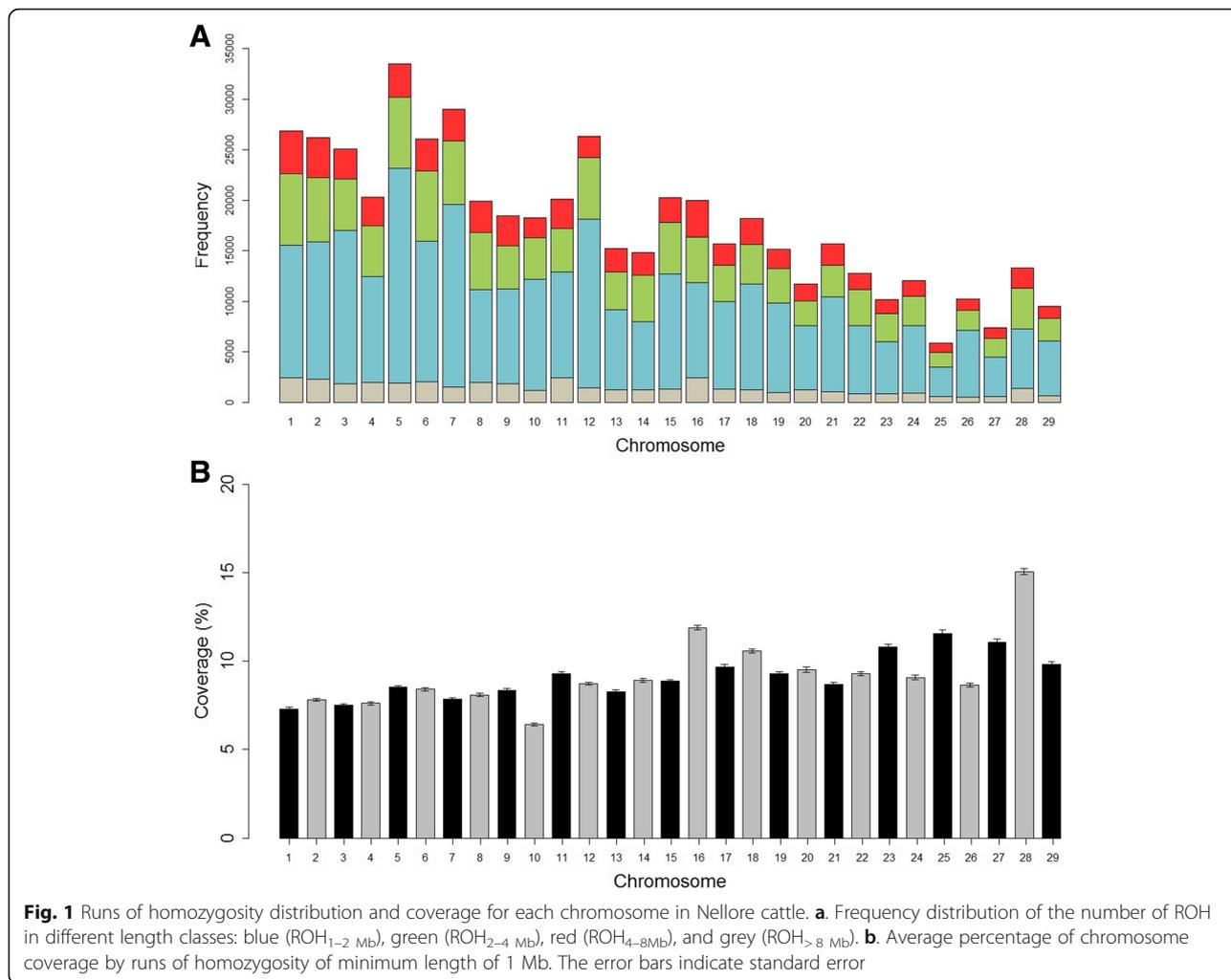
Since homozygous stretches printed on the genome may have arisen as a result of artificial selection, autozygosity based on ROH can strongly disclose the understanding of genetic selection [18]. ROH patterns are not seen to be randomly distributed across the genomes [23] and genomic regions sharing ROH patterns potentially contain alleles associated with genetic improvement in livestock [24]. The correlation of ROH and selection for productivity was first identified by Kim et al. [25]. Furthermore, ROH has been successfully utilized as a measure of inbreeding by estimating the level of autozygosity in the genome [15, 16, 25–28].

Up to date, studies characterizing genome-wide autozygosity in the main Nellore lineages are incipient. Hence, this study was carried out to assess genome-wide autozygosity in a Nellore cattle population to identify and characterize ROH patterns as well as to identify autozygosity islands that may have occurred due to selection for functionally important traits in different Nellore lineages and verify whether these lineages differ or not from one another. It attempts also to compare estimates of molecular inbreeding calculated from ROH (F_{ROH}), genomic relationship matrix (F_{GRM}), and from pedigree-based coefficient (F_{PED}).

Results

Genome-wide distribution of runs of homozygosity

On individual animal basis, the average number of ROH per animal, considering the genotyped animals ($n = 9386$), was 55.15 ± 13.01 with an average size of 3.24 Mb. The longest ROH was 99.30 Mb in length (28,778 SNPs) on *Bos Taurus* autosome (BTA) 5. The number of ROH per chromosome was also greater for BTA5 (33,492 segments) (Fig. 1a) and the greatest fraction of chromosome covered with ROH was found on BTA28 (15.06% of chromosomal length within an ROH) (Fig. 1b).



ROH analysis for the different length classes for the genotyped animals ($n = 9386$) revealed that the Nellore genome is composed mostly of a high number of shorter segments (ROH_{1-2 Mb} and ROH_{2-4 Mb}), which accounted for approximately 78% of all ROH detected and roughly contributed to 43% of the cumulative ROH length (Table 1). Shorter and medium (ROH_{4-8 Mb}) ROH displayed a similar genome coverage and also a cumulative ROH length, with values varying from 20.53 to 22.88%. Despite the total length of

ROH being composed mostly of a high number of short segments, the proportion of the genome covered by them was relatively small when compared to larger ROH (ROH_{> 8 Mb}).

The most autozygous animal exhibited a ROH genome coverage encompassing 718.96 Mb of the total autosomal genome extension (UMD3.1) covered by markers (28.75% of the cattle genome), totaling 92 ROH \geq ROH_{1-2 Mb}. On average, 7.15% (178.70 Mb) of the genome was considered to be a region of homozygosity.

Table 1 Descriptive statistics of runs of homozygosity number ($nROH$) and length (in Mb) for four different length classes (ROH_{1-2 Mb}, ROH_{2-4 Mb}, ROH_{4-8 Mb}, and ROH_{> 8 Mb})

| Class | n ROH | (%) | Mean Length | Standard Deviation | Genome Coverage (%) | Cumulative ROH Length (%) |
|--------------------------|---------|-------|-------------|--------------------|---------------------|---------------------------|
| ROH _{1-2 Mb} | 285,085 | 55.07 | 1.34 | 0.27 | 1.63 | 22.88 |
| ROH _{2-4 Mb} | 123,254 | 23.81 | 2.79 | 0.56 | 1.47 | 20.53 |
| ROH _{4-8 Mb} | 68,407 | 13.21 | 5.53 | 1.11 | 1.63 | 22.59 |
| ROH _{> 8 Mb} | 40,925 | 7.91 | 13.93 | 7.18 | 2.58 | 34.00 |

Pedigree and genomic inbreeding

Descriptive statistics for F_{PED} and F_{ROH} coefficients for the genotyped animals ($n = 9386$) are presented in Table 2. The average F_{PED} and F_{ROH} were low in the studied population, and it is noteworthy that 94.20% of the genotyped animals exhibited a F_{PED} below 5%. Low correlations were observed between F_{PED} - F_{ROH} and it gradually increased as a function of ROH length (Fig. 2). No estimates of correlation were found between F_{GRM} - F_{PED} and those between F_{GRM} - F_{ROH} decreased as a function of ROH length. The inbreeding evolution (Fig. 3) demonstrates a significant ($p < 0.01$) decay in F_{GRM} and $F_{ROH > 8 \text{ Mb}}$.

F_{PED} and F_{ROH} averages for each Nellore lineage ($n = 8646$) are presented in Table 3. The highest F_{PED} ($p < 0.05$) values were observed for Karvadi, Golias, and Godhavari lineages. F_{ROH} estimates were close to F_{PED} and they did not differ ($p < 0.05$) for Karvadi and Godhavari lineages.

Autozygosity islands in Nellore lineages

Autozygosity islands were evident across the genome, and their distributions along the genome vary in length and position across chromosomes. A total of 62 regions with 100 outlying consecutive SNPs were identified for the genotyped animals ($n = 9386$) in almost all autosomes, with the exception of BTA2, BTA11, BTA18, BTA25, and BTA28 (Additional file 1). Overall, the mean length was $1.40 \pm 0.85 \text{ Mb}$, and the longest island was observed on BTA7 (107,000,000:111,700,000 bp) encompassing 4.70 Mb of length. Interestingly, BTA7 also contained the highest number of islands ($n = 8$) followed by BTA1, BTA12 and BTA20, all-encompassing five islands each.

To verify if the autozygosity islands possess genes related to environmental adaptation processes, those 62 autozygosity islands were overlapped with 9803 CNVRs strongly associated with adaptation for Nellore cattle described by Lemos et al. [29]. Only 338 CNVRs were observed within the autozygosity islands, and the overlapping regions harbored 484 genes with described functions.

When analyzing the autozygosity islands within the lineages ($n = 8646$), the Karvadi lineage showed the highest number of islands ($n = 54$), followed by Godhavari ($n = 31$), Golias ($n = 26$), Taj Mahal ($n = 18$), Akasamu ($n = 13$) and Nagpur ($n = 6$). It should be noted

that overlapping islands were observed in between the lineages (Additional files 2 and 3). Interestingly, the region on BTA7 encompassing 51,610,000 to 52,930,000 bp in length was found to be described in all lineages. Non-overlapping autozygosity islands were also observed in some lineages in specific genomic regions and were screened for gene content (Additional file 4). These regions could be an indicative of selection signatures or it may reflect inbreeding events within a lineage [26].

Functional annotation of genes

As most of autozygosity islands identified for the genotyped animals ($n = 9386$) overlapped with those described for the Nellore lineages (Additional file 5), the analysis performed using the DAVID v.6.8 [30, 31] comprised 946 genes identified for the genotyped animals (Table 4). Additional file 6 describes the set of genes involved in each GO term and KEGG pathway.

To obtain a broad functional insight into the set of genes ($n = 484$) observed within the autozygosity islands and CNVRs overlapping regions, an enrichment analysis was also performed. An enhancement of genes involved in several GO terms (four biological processes, one molecular function, and none cellular component process) was significant ($p \leq 0.01$) and one for KEGG (Additional file 7). Despite the large number of overlapping regions, and consequently, the large number of genes found in these regions, no significant GO term and KEGG pathway was found commonly associated in both studies and neither associated in some way with environmental adaptation processes.

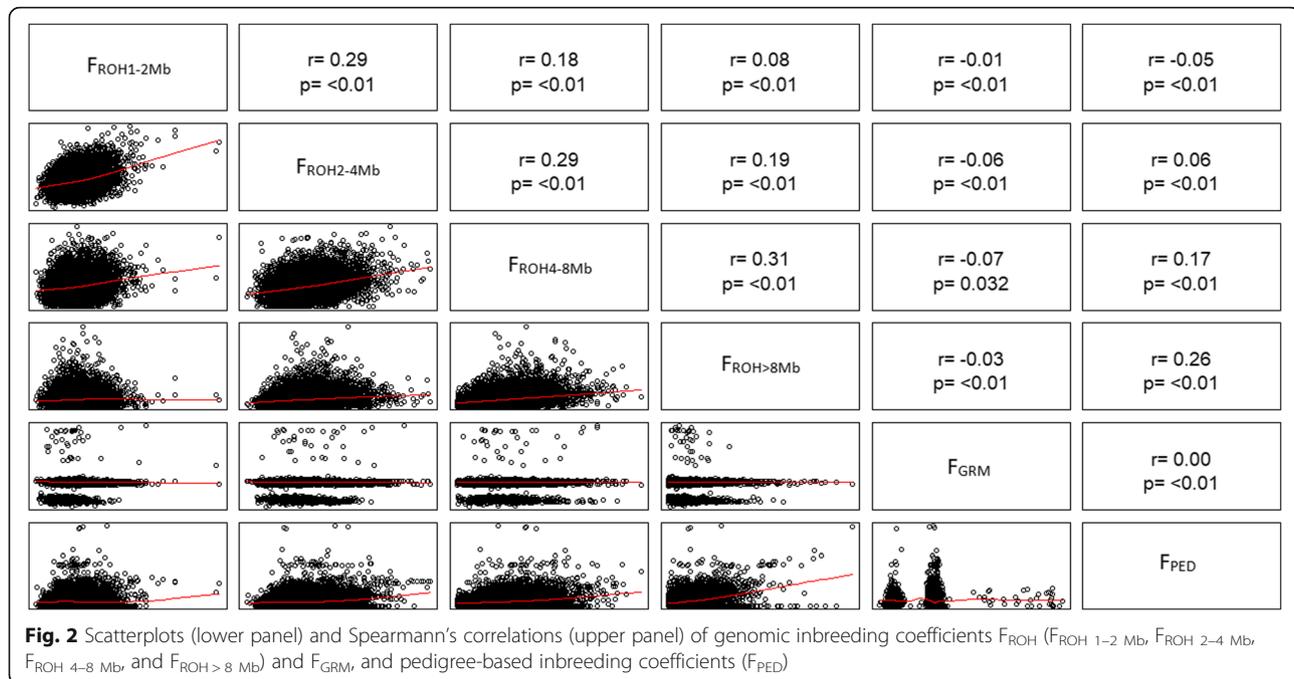
Discussion

Genome-wide distribution of runs of homozygosity

The longest ROH was described on BTA5, however, results in taurine and indicine cattle [20, 25, 32] have reported the longest on BTA8. Corroborating with the results, Peripolli et al. [20] observed the greatest number of ROH on BTA5 in indicine cattle, however, studies have described the greatest number on BTA1 [24, 32, 33]. BTA5, which presented the longest and the greater number of ROH, has been reported to harbor QTL related to weight [34, 35], reproduction [36, 37], and milk fat yield traits [37, 38] in cattle.

Table 2 Number of genotyped animals (n) and descriptive statistics of the pedigree-based inbreeding coefficient (F_{PED}) and runs of homozygosity-based inbreeding coefficient (F_{ROH}) for different lengths (F_{ROH1-2} , F_{ROH2-4} , F_{ROH4-8} , and $F_{ROH > 8 \text{ Mb}}$)

| Coefficient | Mean | Median | Minimum | Maximum | Coefficient of Variation (%) | n |
|--------------------------|-------|--------|---------|---------|------------------------------|------|
| F_{PED} | 0.017 | 0.013 | 0.000 | 0.258 | 3.387 | 8502 |
| $F_{ROH1-2 \text{ Mb}}$ | 0.016 | 0.016 | 0.000 | 0.199 | 27.14 | 9387 |
| $F_{ROH2-4 \text{ Mb}}$ | 0.014 | 0.014 | 0.000 | 0.100 | 37.71 | 9352 |
| $F_{ROH4-8 \text{ Mb}}$ | 0.016 | 0.015 | 0.001 | 0.059 | 47.81 | 9281 |
| $F_{ROH > 8 \text{ Mb}}$ | 0.025 | 0.021 | 0.003 | 0.222 | 77.03 | 8836 |



Dissimilarity among animals was observed between the number of ROH and the length of the genome covered by ROH (Fig. 4). Animals exhibiting the same homozygous genome length displayed a variable number of ROH. This pattern was also described by Mészáros et al. [39], who attributed this event as a consequence of the distinct distances from the common ancestor. Therefore, when considering animals with the same homozygous genome length, we can infer that those displaying more ROH have an increased distance with the common ancestor since these segments are expected to be shorter due to repeated meiosis events that break up ROH through recombination [40].

The highest autozygosity value per animal was similar to those reported in the literature for dairy breeds [20, 24, 32, 41]. Conversely, Marras et al. [18] described that dairy breeds had a higher sum of all ROH than did beef breeds, and Purfield et al. [24] observed that dairy breeds were the most autozygous animals among several studied breeds. In addition, the autozygotic proportion of the genome described for this population seems to indicate moderate to high inbreeding levels for classical standards. Similar results were described by Marras et al. [18] for Marchigiana beef cattle (7%) and Peripolli et al. [20] for Gyr dairy cattle (7.10%). Compared to Zavarez et al. [19] study on a Nellore population whose findings showed a value of 4.58%, this sample of Nellore animals presented a higher average autosomal coverage. The high autozygosity value per animal and homozygous proportion of the genome observed for this population might be a result of the small number of imported progenitors to speed up the genetic progress and develop the

first Nellore lineages during the major importation in the sixties. Furthermore, the formation of lineages can be made by the use of consanguinity in which the same breeder is mated with its descendants along the generations aiming to fix genes related to important traits [8].

Pedigree and genomic inbreeding

F_{PED} was lower than results reported by Barbosa et al. [42] and higher than those described by Santana et al. [43], with values of 8.32% and 1.42% for inbred Nellore populations, respectively.

F_{ROH} can disclose the age of the inbreeding given the approximate correlation between the length of the ROH and the distance with the common ancestor due to recombination events over time. Therefore, calculated F_{ROH} are expected to correspond to the reference ancestral population dating 50 ($F_{ROH1-2 Mb}$), 20 ($F_{ROH2-4 Mb}$), 12.5 ($F_{ROH4-8 Mb}$), and 6 ($F_{ROH > 8 Mb}$) generations ago by considering that 1 cM equals to 1 Mb [44]. According to Zavarez et al. [19], incomplete pedigree cannot account for inbreeding caused by distant ancestors and estimates based on F_{PED} are only comparable with F_{ROH} calculated over large ROH. F_{PED} estimate was then compared with $F_{ROH > 8 Mb}$, and the genome autozygotic proportion from $F_{ROH > 8 Mb}$ exceeded F_{PED} . This variation can be attributed to the fact that the pedigree might not have been deep enough to allow F_{PED} to capture the relatedness since its average depth is close to four generations, whereas $F_{ROH > 8 Mb}$ reflects an inbreeding that occurred nearly six generations ago. Furthermore, F_{PED} does not take into account the stochastic events of recombination

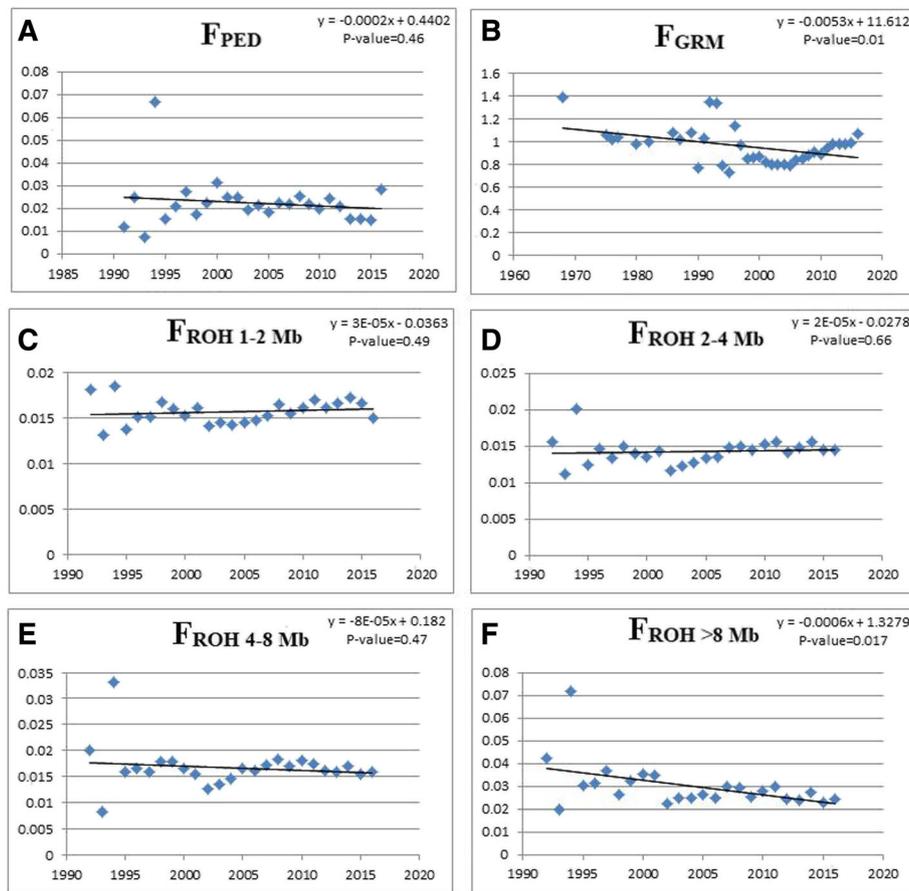


Fig. 3 Inbreeding evolution over the past 30 years for pedigree-based inbreeding (F_{PED}), genomic relationship matrix approach (F_{GRM}), and F_{ROH} ($F_{ROH1-2\ Mb}$, $F_{ROH2-4\ Mb}$, $F_{ROH4-8\ Mb}$, and $F_{ROH > 8\ Mb}$) coefficients and their respective regression equations and p-values. The X-axis represents the years and the Y-axis shows the inbreeding coefficients. Each blue dot represents the inbreeding average per year

during meiosis [26] and pedigree relatedness does not show the actual relatedness among individuals since it is estimated from statistical expectations of the probable identical by descendent (IBD) genomic proportion [45].

F_{PED} - F_{ROH} correlations were seen to be higher when longer ROH reflecting recent relatedness were included in F_{ROH} estimates. It is noticeable to highlight that most of the pedigree records did not extend back many generations, therefore, correlations with shorter ROH reflecting ancient relatedness tended to be lower and those

with longer ROH reflecting recent relatedness had a tendency to be higher [18, 46]. Additionally, several authors have reported a high correlation between F_{PED} - F_{ROH} when a deeper number of described generations are available in the pedigree [15, 16, 18, 24, 33].

No estimates of correlation between F_{GRM} - F_{PED} may be explained by considering that individuals from sub-populations for which allele frequencies diverge from the entire population may have been estimated to have high F_{GRM} [47], which may have led to biased

Table 3 Average mean (number of observations) of pedigree-based inbreeding coefficient (F_{PED}) and runs of homozygosity-based inbreeding coefficient (F_{ROH}) for different lengths (F_{ROH1-2} , F_{ROH2-4} , F_{ROH4-8} , and $F_{ROH > 8\ Mb}$) for six Nellore lineages

| Coefficient | Karvadi | Golias | Godhavari | Taj Mahal | Akasamu | Nagpur |
|-------------------|---------------------------|---------------------------|---------------------------|---------------------------|--------------------------|-------------------------|
| F_{PED}^1 | 0.020 ^a (7282) | 0.019 ^a (178) | 0.020 ^a (90) | 0.016 ^{ab} (103) | 0.011 ^b (42) | - |
| $F_{ROH1-2\ Mb}$ | 0.016 ^a (7853) | 0.014 ^c (288) | 0.015 ^{ab} (205) | 0.014 ^{bc} (149) | 0.014 ^c (79) | 0.014 ^c (50) |
| $F_{ROH2-4\ Mb}$ | 0.014 ^a (7810) | 0.012 ^b (284) | 0.014 ^a (198) | 0.012 ^b (144) | 0.011 ^b (73) | 0.012 ^b (44) |
| $F_{ROH4-8\ Mb}$ | 0.015 ^a (7664) | 0.014 ^b (266) | 0.016 ^a (185) | 0.014 ^b (136) | 0.014 ^b (70) | 0.012 ^b (40) |
| $F_{ROH > 8\ Mb}$ | 0.025 ^a (7443) | 0.022 ^{bc} (245) | 0.024 ^{ab} (171) | 0.018 ^c (130) | 0.022 ^{bc} (70) | 0.017 ^c (34) |

F_{PED} was not available for the Nagpur lineage. Means sharing a common letter within a row were not significantly different ($p < 0.05$) from one another

Table 4 Gene Ontology (GO) terms and KEGG pathways annotation analysis enriched ($P < 0.01$) based on autozygosity islands set of genes

| Terms | Genes | P-value |
|---|-------|---------|
| GO Biological Process | | |
| (GO:0042742) Defense response to bacteria | 14 | 7.07E-5 |
| (GO:0030163) Protein catabolic process | 9 | 6.33E-4 |
| (GO:0070200) Establishment of protein localization to telomere | 4 | 1.70E-3 |
| (GO:0040014) Regulation of multicellular organism growth | 6 | 2.68E-3 |
| (GO:0045647) Negative regulation of erythrocyte differentiation | 4 | 4.46E-3 |
| (GO:0030901) Midbrain development | 6 | 4.84E-3 |
| GO Molecular Function | | |
| (GO:0008289) Lipid binding | 13 | 2.07E-4 |
| (GO:0004190) Aspartic-type endopeptidase activity | 9 | 3.24E-4 |
| GO Cellular Component | | |
| (GO:0005776) Autophagosome | 8 | 3.07E-3 |
| (GO:0005634) Nucleus | 155 | 6.11E-3 |
| (GO:0005815) Microtubule organizing center | 10 | 8.36E-3 |
| (GO:0005730) Nucleolus | 41 | 8.50E-3 |
| KEGG pathway | | |
| (bta01100) Metabolic pathways | 72 | 4.21E-4 |

correlation. According to Zhang et al. [48], inbreeding coefficients based on methods using allele frequency are sensitive compared to ROH-based methods, especially for populations with divergent allele frequencies. Correlations between F_{GRM} - F_{ROH} decreased as a function of ROH length, and Zavarez et al. [19] associated it with the properties of the G matrix, which is based on individual loci, whereas F_{ROH} is based on chromosomal segments.

The inbreeding evolution stress out a significant ($p < 0.01$) decline in $F_{ROH > 8 \text{ Mb}}$ and it is worth highlighting that it reflects inbreeding up to six generations prior (~ 30 years). The reduction in this coefficient since the

1990's happened together with the foundation of the Nelore Brazil Breeding program in 1988 (ANCP, <http://www.ancp.org.br>). These results pointed out, that mating decisions were taken since this time by the breeders to avoid mating between relatives, decreasing the genomic inbreeding level in this population over time. The $F_{ROH 4-8 \text{ Mb}}$ reflects inbreeding up to 12.5 generations prior (~ 60 years) and the slight reduction in this coefficient since the 1960's happened together with the beginning of bull evaluation for weight gain in test stations. The results obtained for $F_{ROH1-2 \text{ Mb}}$ and $F_{ROH2-4 \text{ Mb}}$ showed that mating decisions before the major importations might have favored the increasing of inbreeding.

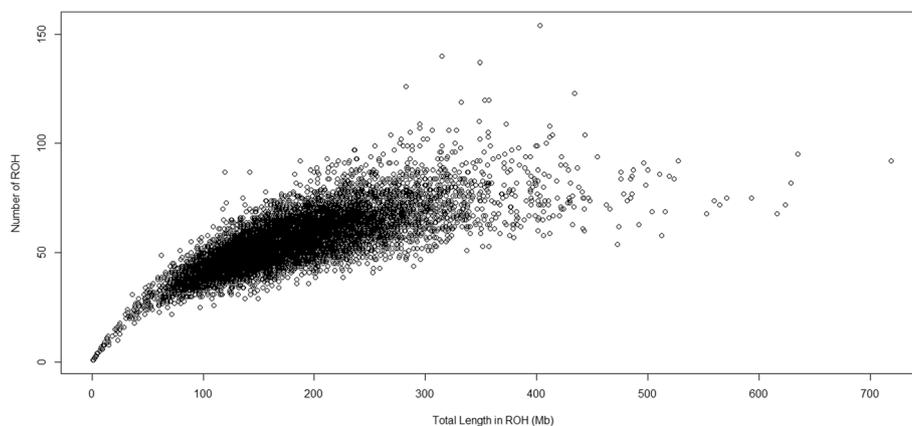


Fig. 4 Relationship between the number of runs of homozygosity (ROH) per individual and the total length of the genome covered by them. Each hollow circle stands for one animal

Inbreeding coefficients were not high for the genotyped animals with lineages records ($n = 8646$), with values around to 2%. According to Pereira [49], the lineage diversification within a breed can provide substantial gains for selection by reducing inbreeding rates and restoring the genetic variability. The use of Karvadi, Golias, and Godhavari lineages can be evidenced by the high inbreeding rates described for them when compared to other lineages. According to Oliveira et al. [7], when considering a small number of progenitors in a studied breed, the prevalence use of some ancestors can be explained by their marginal contribution in the reference population. Hence, when assessing the marginal contribution of each lineage to the ANCP Nellore cattle population, an eminent contribution of Karvadi and Godhavari lineages can be observed (10.44 and 1.48%, respectively), agreeing with F_{ROH} estimates. Lineages such as Golias, Taj Mahal, Akasamu, and Nagpur did not show an expressive marginal contribution, and interestingly, displayed lower inbreeding averages ($p < 0.05$) for $F_{ROH1-2\text{ Mb}}$, $F_{ROH2-4\text{ Mb}}$, and $F_{ROH4-8\text{ Mb}}$.

Autozygosity islands in Nellore lineages

Autozygosity islands in the genotyped animals ($n = 9386$) were seen overlapping with previous studies on several cattle breeds (Additional file 8). Within these studies, islands were not reported overlapping only with those described for Nellore cattle. Remarkably, Sölkner et al. [50] and Szmatoła et al. [41] displayed islands in common on BTA7 encompassing the same chromosomal region around 51–53 Mb, and Szmatoła et al. [41] also described islands located on the same chromosomal region on BTA7 (42–44 Mb) in Holstein, Red Polish, Simmental and Limousin cattle breeds. Sölkner et al. [50] and Gaspa et al. [51] exhibited overlapping islands around 1.3–1.9 Mb on BTA21. Overlapping islands between these studies and the current one (43,510,000:43,592,173 – BTA7; 51,574,295:52,353,000 – BTA7, and 1,360,390: 1,829,761 – BTA21) were inspected in detail. These islands are suggested to harbor targets of positive selection in cattle [52] and may be used to identify regions of the genome under selection, and to map genes that affect traits of interest [18]. Further, ROH islands were found overlapping in cattle breeds selected for different purposes, suggesting that selection pressure can also be undergoing on traits other than those specific to dairy or beef traits.

When examining in detail, the region encompassing 51–52 Mb on BTA7 harbored relevant genes for beef cattle production. Among them, we highlight the *CTNNA1* gene which has been associated with myostatin expression level in skeletal muscle of Holstein-Friesian bulls [53]. Myostatin is a key protein that plays an essential role in regulating skeletal muscle growth, and it is considered to be one of the most important factors responsible for meat

productivity traits in cattle [54]. The *MATR3* gene was also described within the overlapping region and has been related to fat deposition in cattle [55, 56]. It is also worth highlighting the *ECSCR* gene. This gene regulates insulin sensitivity and predisposition to obesity [57]. Besides, the protein encoded by this gene is primarily found in endothelial cells and blood vessels (provided by RefSeq, Jun 2014). Endothelial cells are the important players in angiogenesis, a physiological process by which new blood vessels develop from pre-existing vasculature [58]. Blood vessels dilate to dissipate heat to external environment by a process denominated vasodilation. In this regard, the *ECSCR* gene might be a key role in elucidating the better tolerance of some cattle breeds to heat stress, i.e. *Bos taurus indicus*. The increased number of blood vessels through the angiogenic process allows more blood to be dissipated, decreasing the body temperature.

Overlapping islands within the lineages ($n = 8646$) were described in this study and two reasons might have led to this result. First, the Nellore cattle sampled in Brazil is derived from the Ongole cattle imported from the Indian district of Andhra Pradesh [59]. Prior such importations, the Ongole cattle was already notorious in India due to their greater adaptation upon high temperatures, ability to carry lower burdens of cattle tick and tolerate poor feed management [60]. Therefore, these overlapping regions might reflect the acquired adaptiveness of zebu cattle in tropical environments due to natural selection over the time [61]. Second, these findings support the concept that despite having different lineages within the Nellore breed, the genetic progress of economically important traits goes toward the same direction and IBD genomic regions harboring traits of interest are being conserved over time.

The region on BTA7 described to be overlapping in all lineages (51,610,000:52930000 bp) harbored five genes (*CTNNA1*, *LRRTM2*, *SILI*, *MATR3*, and *PAIP2*). Among them, the *CTNNA1* (Catenin Alpha 1) gene has been described associated with myostatin expression level and molecular function in skeletal muscle in Holstein-Friesian bulls [53]. Furthermore, the *LRRTM2* (Leucine Rich Repeat Transmembrane Neuronal 2) gene was found related to maturation of male germ cells and male fertility [62, 63].

Non-overlapping islands within the lineages were explored for gene content and among the genes identified within the regions we can highlight those described in Table 5. Remarkably, six genes were also reported in Nellore-specific studies associated with carcass traits [64] (*PPM1*), age at first calving [65] (*NPBWRI*, *OPRKL*, and *MRPL1*), and birth weight [66] (*RPS20* and *TGS1*).

Despite having non-overlapping autozygosity islands within the lineages, several genes have been found described associated with productive and reproductive traits within the lineages. Productive related-genes were

Table 5 Gene content of non-overlapping ROH islands within the Nellore lineages highlighted according to their function

| Lineage | Gene | Function | Author |
|-----------|---|---------------------|------------------|
| Godhavari | <i>LAMB4</i> | Immune System | [91] |
| Karvadi | <i>RFX4</i> | Immune System | [92] |
| Godhavari | <i>IFRD, PPM1B, DTX4, MTMR7</i> | Productive traits | [64, 92–95] |
| Taj Mahal | <i>CAPZA2</i> | Productive traits | [96] |
| Karvadi | <i>ZBTB20, RPS20, STAC3, STAT6, RIC8B, LYPLA1, XKR4, TMEM68, TGS1</i> | Productive traits | [66, 92, 97–102] |
| Godhavari | <i>NAMPT</i> | Reproductive traits | [103, 104] |
| Godhavari | <i>PPM1B, JMJD1C</i> | Reproductive traits | [105, 106] |
| Karvadi | <i>RFX4, NPBWR1, OPRK1, MRPL15</i> | Reproductive traits | [65, 107] |
| Karvadi | <i>DRD3, ZBTB20</i> | Reproductive traits | [108, 109] |
| Karvadi | <i>CSNK1A1, TBC1D12</i> | Thermotolerance | [110, 111] |

mainly associated with average daily gain (*IFRD1*), muscle (*PPM1B* and *STAC3*), fat (*DTX4* and *XKR4*), body and birth weight (*MTMR7*, *RPS20*, and *TGS1*), meat and carcass quality traits (*MTMR7*, *CAPZA2*, *STAT6*, and *RIC8B*), and feed intake (*LYPLA1* and *TMEM68*). Reproductive related-genes largely encompassed those linked to heifer's fertility (*RFX4*), age at first calving (*NPBWR1*, *OPRK1*, and *MRPL15*), and oocyte maturation and expression (*NAMPT* and *JMJD1C*).

Although they were not located in the same genomic regions, these autozygosity islands showed an enrichment of genes involved in cattle growth, meat and carcass quality traits, immune system, and thermotolerance functions. These findings help to reinforce the concept that the genetic progress goes towards the same direction within the lineages and different genetic patterns among the lineages based on the selection criterion used to improve each of them could not be identified in this study.

Functional annotation of genes

The analyses performed on DAVID revealed only the metabolic pathways (bta01100) KEGG pathway as significant ($p < 0.01$), while the Gene Ontology analyses showed several enriched terms for the ROH gene list. The defense response to bacteria (GO:0042742) on biological process encompasses several reactions triggered in response to the presence of a bacteria that act to protect the cell or organism. We highlighted the beta-defensin genes (*DEFB1*, *DEFB4A*, *DEFB5*, *DEFB6*, *DEFB7*, *DEFB10*, and *DEFB13*) that encode host defense peptides that are critical to protection against bacterial, viral and fungal infections, and acts as an important link between innate and adaptive immune responses [67]. In addition to their antimicrobial properties, beta-defensins have an important role in several functions including regulation of the immune response, fertility, reproduction, and embryo development [67, 68].

The negative regulation of erythrocyte differentiation (GO:0045647) on biological process is defined as any process that stops, prevents, or reduces the frequency, rate or extent of erythrocyte differentiation. Erythrocytes were described by Nelson [69] as belonging to the immune complex reaction (bacteria, complement, and antibody). In fish and chickens, erythrocytes have been shown to facilitate the clearance of pathogens by macrophages [70], and could produce specific signaling molecules such as cytokines in response to binding [71, 72].

The protein catabolic process (GO:0030163) includes chemical reactions and pathways resulting in the breakdown of mature proteins, which play an important role in the immune and inflammatory response. Khansefid et al. [73] identified the protein catabolic process enriched in genes significantly associated with residual feed intake in Angus and Holstein cattle breeds. Regarding the genes related to protein catabolic process identified in our study, most of them are pregnancy-associated glycoproteins genes (PAG) (Supplementary file 6) mapped on BTA29. Goszczynski et al. [74] identified eight genes belonging to the PAG gene family within ROH islands in Retinta cattle breed, while Szmatoła et al. [41] identified sixteen PAG genes in Holstein cattle breed. PAG glycoproteins are one major group of the proteins secreted from trophoblast cells of the placenta into the maternal blood shortly after implantation and are detectable throughout gestation [56]. These proteins have been used to monitor embryonic viability as biochemical pregnancy markers in the cow's blood or milk [75] as well as placental functions in cattle [76, 77]. Significant reductions in PAG concentrations during the late embryonic/early fetal period are associated with pregnancy failures in cattle [76, 78]. PAG proteins also play an important role in implantation, placentogenesis, fetal antigen sequestering, and fetal–maternal interactions [76, 79–81]. Modifications in circulating PAG

concentrations also were associated with several parameters linked to pregnancy loss in cattle, including parity, artificial insemination service number, milk yield, and metabolic diseases [82].

The regulation of multicellular organism growth (GO:0040014) biological process encompasses any process that modulates the frequency, rate or extent of growth of the body of an organism so that it reaches its usual body size, while the midbrain development (GO:0030901) biological process encompass the process whose specific outcome is the progression of the midbrain over time, from its formation to the mature structure.

Conclusions

This study is the first of its kind to bring out results characterizing genome-wide autozygosity in the main Nellore lineages. The average F_{PED} and F_{ROH} of different lengths were low in the studied population, however, the autozygotic proportion in the genome indicates moderate to high inbreeding levels. Low correlations between F_{PED} , F_{ROH} may be partly due to the relatively superficial depth of the pedigree, emphasizing the concept that autozygosity based on ROH should be used as an accurate estimator of ancient individual inbreeding levels [15, 24, 33, 83]. Overall, inbreeding coefficients were not high within the lineages and the findings obtained in this study suggest that lineages displaying an eminent marginal contribution in the reference population also display the highest F_{ROH} values, i.e. Karvadi and Godhavari.

Genomic regions that are selection targets tend to generate autozygosity islands and several of them have been described in the Nellore genome. Most remarkable is the clear evidence of autozygosity islands patterns within the lineages, suggesting that IBD genomic regions have been selected for the same traits over time. Autozygosity islands harbored enriched terms in which we highlight the defense response to bacteria (GO:0042742) and the negative regulation of erythrocyte differentiation (GO:0045647), which might help to better elucidate the greater adaptation of indicine cattle in host environment given its association with immune responses mechanisms. Additionally, non-overlapping autozygosity islands within the lineages were found to contain genes related to cattle growth, reproduction, and meat and carcass quality traits. The results of this study give a comprehensive insight about the autozygosity patterns in the main Nellore lineages and their potential role in explaining selection for functionally important traits in cattle. Despite having different lineages within the Nellore breed, it has clearly shown that selection is going towards the same direction and different genetic patterns could not be described.

Methods

Animals and genotyping

The animals used in this study comprise a dataset and progeny test program from the National Association of Breeders and Researchers (ANCP – Ribeirão Preto-SP, Brazil). The progeny test program headed by ANCP aims to disseminate semen of genetically superior Nellore young bulls evaluated for sexual precocity, growth, morphologic composition, feed efficiency, and carcass quality traits.

Nellore animals were genotyped with the low-density panel (CLARIFDE° Nellore 2.0) containing over 20,000 markers ($n = 7729$ animals); GGP-LD BeadChip (GeneSeek° Genomic Profiler 30 K) that contains 30,106 markers ($n = 201$ animals); Illumina BovineSNP50° Beadchip (Illumina Inc., San Diego, CA, USA) containing 54,001 markers ($n = 58$ animals); GGPi BeadChip (GeneSeek° Genomic Profiler Indicus) that contains 74,153 markers ($n = 487$ animals); and with Illumina BovineHD BeadChip (Illumina Inc., San Diego, CA, USA) containing 777,962 markers ($n = 911$ animals). Imputation was implemented using the FIMPUTE 2.2 software [84] and all genotypes were imputed to a panel containing 735,044 markers. A reference population with 963 sires and dams genotyped with the Illumina BovineHD BeadChip (Illumina Inc., San Diego, CA, USA) was used. Prior imputation, markers were edited for call rate ($< 90\%$) for the genotyped and the reference populations. SNPs unsigned to any chromosome and those assigned to sexual chromosomes were removed from the dataset. After editing, a total of 9386 animals and 735,044 SNP markers were retained for the analyses. Genotyped animals with lineages records ($n = 8646$) were categorized as follows: Karvadi Imp ($n = 7860$), Goliath Imp ($n = 290$), Godhavari Imp ($n = 210$), Taj Mahal Imp ($n = 150$), Akasamu Imp ($n = 81$), and Nagpur Imp ($n = 55$). Lineages were classified using the PEDIG package [85], which estimates the average consanguinity between a set of individuals and a reference group. The reference group encompassed founder's animals from the Nellore base population in which the Nellore lineages were derived from.

Runs of homozygosity

Individual ROH was identified using PLINK v1.90 software [86], which uses a sliding window approach to scan each individual's genotype at each marker position to detect homozygous segments [44]. The parameters and thresholds applied to define ROH were set as follows: a sliding window of 50 SNPs across the genome, a minimum number of 100 consecutive SNPs included in a ROH, a minimum ROH length of 1 Mb, a maximum gap between consecutive homozygous SNPs of 0.5 Mb, one SNP per 50 kb, and a maximum of five SNPs with

missing genotypes and up to one heterozygous genotype in a ROH. ROH were classified into four length classes: 1–2, 2–4, 4–8, and > 8 Mb, identified as $ROH_{1-2 \text{ Mb}}$, $ROH_{2-4 \text{ Mb}}$, $ROH_{4-8 \text{ Mb}}$, and $ROH_{>8 \text{ Mb}}$, respectively. ROH were performed separately for all genotyped animals ($n = 9386$) and for each Nellore lineage ($n = 8646$).

Pedigree and genomic inbreeding coefficients

Pedigree-based inbreeding coefficients (F_{PED}) were estimated using pedigree records from a dataset containing 45,917 animals born between 1934 and 2017. The pedigree dataset was provided by the National Association of Breeders and Researchers (ANCP – Ribeirão Preto-SP, Brazil). The average pedigree depth was approximately four generations, with a maximum depth value of nine. The F_{PED} was estimated for both datasets ($n = 9386$ and $n = 8646$) through the software INBUPGF90 [87]. Genomic inbreeding coefficients based on ROH (F_{ROH}) were estimated for each animal and both datasets, according to the genome autozygotic proportion described by McQuillan et al. [21]:

$$F_{ROH} = \frac{\sum_{j=1}^n L_{ROH_j}}{L_{total}}$$

where L_{ROH_j} is the length of ROH_j , and L_{total} is the total size of the autosomes covered by markers. L_{total} was taken to be 2,510,605,962 bp, based on the consensus map. For each animal, F_{ROH} ($F_{ROH_{1-2 \text{ Mb}}}$, $F_{ROH_{2-4 \text{ Mb}}}$, $F_{ROH_{4-8 \text{ Mb}}}$, and $F_{ROH_{>8 \text{ Mb}}}$) was calculated based on ROH distribution of four minimum different lengths (ROH_j): 1–2, 2–4, 4–8, and > 8 Mb, respectively. A second measure of genomic inbreeding was calculated just for the whole dataset ($n = 9386$) using the Genomic relationship matrix (G) (F_{GRM}). The G matrix was calculated according to VanRaden et al. [88] as follows:

$$G = \frac{ZZ'}{2\sum_{i=1}^n P_i(1-P_i)}$$

where Z is a genotype matrix that contains the $0-2p$ values for homozygotes, $1-2p$ for heterozygotes, and $2-2p$ for opposite homozygotes, where P_i is the reference allele frequency at locus i th. The diagonal elements of the matrix G represent the relationship of the animal with itself, thus, it was used to assess the genomic inbreeding coefficient. Spearman method was used to estimate correlations between the inbreeding measures.

Identification and gene prospection in autozygosity islands

Autozygosity islands were defined as regions where SNPs were outliers according to boxplot distribution for each autosome (Additional files 9 and 10). A file generated by PLINK v1.90 software [86] which specifies how

many times each SNP appeared in an ROH was used and regions displaying at least 100 consecutive outlier SNPs were then classified as an autozygosity island. Raw data regarding how many times each SNP appeared in an ROH was log-transformed (\log_{10}). Autozygosity islands were identified separately for all genotyped animals ($n = 9386$) and for each Nellore lineage ($n = 8646$).

The gene content of the autozygosity islands was identified using the UMD3.1 bovine genome assembly from the Ensembl BioMart tool [89]. Database for Annotation, Visualization, and Integrated Discovery (DAVID) v6.8 tool [30, 31] was used to identify significant ($p \leq 0.01$) Gene Ontology (GO) terms and KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways using the list of genes from autozygosity islands and the *Bos taurus taurus* annotation file as background.

Autozygosity islands previously identified for the genotyped animals were overlapped with copy number variation regions (CNVRs) described for Nellore cattle by Lemos et al. [29]. Overlap analysis was carried out using the Bioconductor package *GenomicRanges* [90].

Additional files

Additional file 1: Autozygosity islands across the Nellore cattle genome. (DOCX 20 kb)

Additional file 2: Autozygosity islands within the Nellore lineages by chromosome: Karvadi (red), Golias (Black), Godhavari (Green), Taj Mahal (blue), Akasamu (purple), and Nagpur (yellow). (PDF 31 kb)

Additional file 3: Overlapping autozygosity islands within the Nellore lineages. (DOCX 28 kb)

Additional file 4: Non-overlapping autozygosity islands within the Nellore lineages. (DOCX 26 kb)

Additional file 5: Autozygosity islands within the genotyped animals (red) and those with lineages records (black). (PDF 29 kb)

Additional file 6: Gene Ontology terms and KEGG pathways annotation analysis enriched ($P < 0.01$) based on autozygosity islands set of genes identified for the genotyped animals ($n = 9386$). (DOCX 17 kb)

Additional file 7: Gene Ontology terms annotation analysis enriched ($P < 0.01$) based on copy number variation regions (CNVRs) and autozygosity islands overlapping regions set of genes identified for the genotyped animals ($n = 9386$). (DOCX 15 kb)

Additional file 8: Runs of homozygosity islands described in several cattle breeds located within those observed in the present study. (DOCX 22 kb)

Additional file 9: Outliers SNPs for the genotyped animals ($n = 9386$) according to Boxplot distribution. (PDF 227 kb)

Additional file 10: Outliers SNPs for each Nellore lineage ($n = 8646$) according to Boxplot distribution. (PDF 340 kb)

Abbreviations

ANCP: Associação Nacional de Criadores e Pesquisadores; BLUP: Best Linear Unbiased Prediction; BTA: *Bos taurus* autosome; CNV: Copy number variation; CNVRs: Copy number variation regions; DAVID: Database for Annotation, Visualization, and Integrated Discovery; F_{GRM} : Genomic relationship matrix-based estimates of inbreeding; F_{PED} : Pedigree-based estimates of inbreeding; F_{ROH} : ROH-based estimates of inbreeding; G : Genomic relationship matrix; GO: Gene Ontology; IBD: Identical by descent; KEGG: Kyoto Encyclopedia of Genes and Genomes; NCBI: National Center for Biotechnology Information;

QTL: Quantitative trait loci; ROH: Runs of homozygosity; SNP: Single Nucleotide Polymorphism

Acknowledgements

E.P is supported by São Paulo Research Foundation (FAPESP) grant #2017/24084-7.

Funding

Supported by São Paulo Research Foundation (FAPESP) grant #2011/21241-0.

Availability of data and materials

The dataset analyzed during the current study are not publicly available due to belonging to the National Association of Breeders and Researchers (ANCP).

Authors' contributions

EP, FB, and RBL conceived and designed the experiment. EP and FB carried out the data analyses. EP, JM, MVAL, NBS, SK, BFO, FLBF, MPB, FBL, DPM, CUM, FDC, JO, SD, ASCP interpreted the results and drafted the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

The DNA was extracted from semen bought from artificial insemination centers and therefore no specific ethical approval is needed (Brazil law number 11794, from October 8th, 2008, Chapter 1, Art. 3, paragraph III). All the samples were obtained with the consent of the artificial insemination centers to use for research.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Departamento de Zootecnia, Faculdade de Ciências Agrárias e Veterinárias, UNESP Univ Estadual Paulista Júlio de Mesquita Filho, Jaboticabal 14884-900, Brazil. ²Institute for Animal Breeding and Genetics, University of Veterinary Medicine Hannover, 30559 Hannover, Germany. ³Departamento de Ciências Exatas, Faculdade de Ciências Agrárias e Veterinárias, UNESP Univ Estadual Paulista Júlio de Mesquita Filho, Jaboticabal 14884-900, Brazil. ⁴Associação Nacional de Criadores e Pesquisadores (ANCP), Ribeirão Preto 14020-230, Brazil. ⁵Embrapa Cerrados, Planaltina 73310-970, Brazil. ⁶Zoetis, Kalamazoo 49007, USA. ⁷Universidade de São Paulo, Departamento de Nutrição e Produção Animal, Pirassununga 13635-900, Brazil.

Received: 9 April 2018 Accepted: 7 September 2018

Published online: 17 September 2018

References

- ABIEC. Associação Brasileira das Indústrias Exportadoras de Carnes [Internet]. 2016 [cited 2017 Jun 16]. Available from: <http://www.abiec.com.br>
- Turner JW. Genetic and biological aspects of zebu adaptability. *J Anim Sci*. 1980;50:1201–5.
- Hansen PJ. Physiological and cellular adaptations of zebu cattle to thermal stress. *Anim Reprod Sci*. 2004;82–83:349–60.
- Jonsson NN. The productivity effects of cattle tick (*Boophilus microplus*) infestation on cattle, with particular reference to *Bos indicus* cattle and their crosses. *Vet Parasitol*. 2006;137:1–10.
- Santiago AA. O Guzerá. Recife: Tropical; 1984.
- Brasil. Projeto de melhoramento genético da zebuicultura: PROZEBU: 1978-1984. Belo Horizonte: Associação Brasileira dos Criadores de Zebu; 1978.
- Oliveira JHF, Magnabosco CD, Borges AMS. Nelore: base genética e evolução seletiva no Brasil. Distrito Federal: Embrapa Cerrados; 2002.
- Magnabosco CD, Cordeiro CMT, Trovo JD, Mariante AD, Lôbo RB, Josahkian LA. Catálogo de linhagens do germoplasma zebuino: raça Nelore. Brasília: Embrapa Recursos Genéticos e Biotecnologia; 1997.
- Lôbo RB, Marcondes CR, Vozzi PA, Lima FP, Bezerra LAF, Zambianchi AR. A tecnologia da informação e a sustentabilidade da raça Nelore. 40 Reun. Anu. da Soc. Bras. Zootec. Santa Maria; 2003.
- de Nadai Bonin M, Ferraz JBS, Eler JP, da Luz Silva S, Rezende FM, Córdova D, et al. Características de carcaça e qualidade de carne em linhagens da raça Nelore Carcass and meat quality traits in lineages of Nelore breed. *Ciência Rural*. 2014;44:1860–6.
- Ferraz JBS, de FPE. Production systems - an example from Brazil. *Meat Sci*. 2010;84:238–43.
- Pereira RJ, Santana ML, Ayres DR, Bignardi AB, Menezes GRO, Silva LOC, et al. Inbreeding depression in zebu cattle traits. *J Anim Breed Genet*. 2016; 133:523–33.
- De Cara MÁR, Villanueva B, Toro MÁ, Fernández J. Using genomic tools to maintain diversity and fitness in conservation programmes. *Mol Ecol*. 2013; 22:6091–9.
- Bosse M, Megens HJ, Madsen O, Crooijmans RPMA, Ryder OA, Austerlitz F, et al. Using genome-wide measures of coancestry to maintain diversity and fitness in endangered and domestic pig populations. *Genome Res*. 2015;25:970–81.
- Ferenčaković M, Hamzić E, Gredler B, Curik I, Sölkner J. Runs of homozygosity reveal genome-wide autozygosity in the Austrian Fleckvieh cattle. *Agric Conspec Sci*. 2011;76:325–9.
- Ferenčaković M, Hamzić E, Gredler B, Solberg TR, Klemetsdal G, Curik I, et al. Estimates of autozygosity derived from runs of homozygosity: empirical evidence from selected cattle populations. *J Anim Breed Genet*. 2013;130:286–93.
- Silió L, Rodríguez MC, Fernández A, Barragán C, Benítez R, Óvilo C, et al. Measuring inbreeding and inbreeding depression on pig growth from pedigree or SNP-derived metrics. *J Anim Breed Genet*. 2013;130:349–60.
- Marras G, Gaspa G, Sorbolini S, Dimauro C, Ajmone-Marsan P, Valentini A, et al. Analysis of runs of homozygosity and their relationship with inbreeding in five cattle breeds farmed in Italy. *Anim Genet*. 2014;46:110–21.
- Zavarez LB, Utsunomiya YT, Carmo AS, Neves HHR, Carneiro R, Ferencakovic M, et al. Assessment of autozygosity in Nelore cows (*Bos indicus*) through high-density SNP genotypes. *Front Genet*. 2015;6:1–8.
- Peripolli E, Baldi F, da Silva MVGB, Irgang R, Lima ALF. Assessment of runs of homozygosity islands and estimates of genomic inbreeding in Gyr (*Bos indicus*) dairy cattle. *BMC Genomics*. 2018;19:34.
- McQuillan R, Leutenegger AL, Abdel-Rahman R, Franklin CS, Pericic M, Barac-Lauc L, et al. Runs of homozygosity in European populations. *Am J Hum Genet*. 2008;83:359–72.
- Karimi Z. Runs of homozygosity patterns in taurine and Indicine cattle breeds (master thesis). Vienna: BOKU - University of Natural Resources and Life Sciences; 2013.
- Zhang Q, Guldbrandsen B, Bosse M, Lund MS, Sahana G. Runs of homozygosity and distribution of functional variants in the cattle genome. *BMC Genomics*. 2015;16:542.
- Purfield DC, Berry DP, McParland S, Bradley DG. Runs of homozygosity and population history in cattle. *BMC Genet*. 2012;13:70.
- Kim ES, Cole JB, Huson H, Wiggans GR, Van Tassel CP, Crooker BA, et al. Effect of artificial selection on runs of homozygosity in U.S. Holstein cattle. *PLoS One*. 2013;8:e80813.
- Curik I, Ferenčaković M, Sölkner J. Inbreeding and runs of homozygosity: a possible solution to an old problem. *Livest Sci*. 2014;166:26–34.
- Kim ES, Sonstegard TS, Van Tassel CP, Wiggans G, Rothschild MF. The relationship between runs of homozygosity and inbreeding in Jersey cattle under selection. *PLoS One*. 2015;10:e0129967.
- Scraggs E, Zanella R, Wojtowicz A, Taylor JF, Gaskins CT, Reeves JJ, et al. Estimation of inbreeding and effective population size of full-blood wagyu cattle registered with the American wagyu cattle association. *J Anim Breed Genet*. 2014;131:3–10.
- Lemos MVA, Piatto Berton M, Ferreira de Camargo GM, Peripolli E, de Oliveira Silva RM, Olivieri BF, et al. Copy number variation regions in Nelore cattle: evidences of environment adaptation. *Livest Sci Elsevier BV*. 2018;207: 51–8.
- Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res*. 2009;37:1–13.

31. Huang DW, Sherman BT, Li L. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc.* 2009;4:44–57.
32. Mastrangelo S, Tolone M, Di Gerlando R, Fontanesi L, Sardina MT, Portolano B. Genomic inbreeding estimation in small populations: evaluation of runs of homozygosity in three local dairy cattle breeds. *Animal.* 2016;10:746–54.
33. Gurgul A, Szmatoła T, Topolski P, Jasielczuk I, Żukowski K, Bugno-Poniewierska M. The use of runs of homozygosity for estimation of recent inbreeding in Holstein cattle. *J Appl Genet.* 2016;57:527–30.
34. McClure MC, Morsci NS, Schnabel RD, Kim JW, Yao P, Rolf MM, et al. A genome scan for quantitative trait loci influencing carcass, post-natal growth and reproductive traits in commercial Angus cattle. *Anim Genet.* 2010;41:597–607.
35. Li C, Basarab J, Snelling WM, Benkel B, Murdoch B, Moore SS, et al. The identification of common haplotypes on bovine chromosome 5 within commercial lines of *Bos taurus* and their associations with growth traits. *J Anim Sci.* 2002;80:1187–94.
36. Kirkpatrick BW, Byla BM, Gregory KE. Mapping quantitative trait loci for bovine ovulation rate. *Mamm Genome.* 2000;11:136–9.
37. Lien S, Karlsen A, Klemetsdal G, Våge DI, Olsaker I, Klungland H, et al. A primary screen of the bovine genome for quantitative trait loci affecting twinning rate. *Mamm Genome.* 2000;11:877–82.
38. Wiener P, Maclean I, Williams JL, Woolliams JA. Testing for the presence of previously identified QTL for milk production traits in new populations. *Anim Genet.* 2000;31:385–95.
39. Mészáros G, Boison AS, Pérez O'Brien AM, Ferenčaković M, Curik I, da Silva MVGB, et al. Genomic analysis for managing small and endangered populations: a case study in Tyrol Grey cattle. *Front Genet.* 2015;6:173.
40. Kirin M, McQuillan R, Franklin CS, Campbell H, Mckeigue PM, Wilson JF. Genomic runs of homozygosity record population history and consanguinity. *PLoS One.* 2010;5:e13996.
41. Szmatoła T, Gurgul A, Ropka-molik K, Jasielczuk I, Tomasz Z, Bugno-Poniewierska M. Characteristics of runs of homozygosity in selected cattle breeds maintained in Poland. *Livest Sci.* 2016;188:72–80.
42. Barbosa ACB, Malhado CHM, Carneiro PLS, Muniz LMS, Ambrosini DP, Carrillo JA, et al. Population structure of Nelore cattle in northeastern Brazil. *Rev Bras Zootec.* 2013;42:639–44.
43. Santana ML, Oliveira PS, Pedrosa VB, Eler JP, Groeneveld E, Ferraz JBS. Effect of inbreeding on growth and reproductive traits of Nelore cattle in Brazil. *Livest Sci.* 2010;131:212–7.
44. Howrigan DP, Simonson MA, Keller MC. Detecting autozygosity through runs of homozygosity: a comparison of three autozygosity detection algorithms. *BMC Genomics.* 2011;12:460.
45. Visscher PM, Medland SE, Ferreira MAR, Morley KI, Zhu G, Cornes BK, et al. Assumption-free estimation of heritability from genome-wide identity-by-descent sharing between full siblings. *PLoS Genet.* 2006;2:0316–25.
46. Saura M, Fernández A, Varona L, Fernández AI, de Cara MÁR, Barraán C, et al. Detecting inbreeding depression for reproductive traits in Iberian pigs using genome-wide data. *Genet Sel Evol.* 2015;47:1.
47. Pryce JE, Haile-Mariam M, Goddard ME, Hayes BJ. Identification of genomic regions associated with inbreeding depression in Holstein and Jersey dairy cattle. *Genet Sel Evol.* 2014;46:71.
48. Zhang Q, Calus MPL, Gulbrandsen B, Lund MS, Sahana G. Estimation of inbreeding using pedigree, 50k SNP chip genotypes and full sequence data in three cattle breeds. *BMC Genet.* 2015;16:88.
49. Pereira JCC. Melhoramento genético aplicado à produção animal. Belo Horizonte: Fundação de Estudo e Pesquisa em Medicina Veterinária e Zootecnia; 1996.
50. Sölkner J, Karimi Z, Pérez O'Brien AM, Mészáros G, Eaglen S, Boison SA, et al. Extremely non-uniform: patterns of runs of homozygosity in bovine populations. 10th world Congr. Genet. Appl. To Livest. Prod. Vancouver; 2014.
51. Gaspa G, Marras G, Sorbolini S, Marsan PA, Williams JL, Valentini A, et al. Genome-wide homozygosity in Italian Holstein cattle using HD SNP panel. 10th world Congr. Genet. Appl. To Livest. Prod. Vancouver; 2014.
52. Pemberton TJ, Absher D, Feldman MW, Myers RM, Rosenberg NA, Li JZ. Genomic patterns of homozygosity in worldwide human populations. *Am J Hum Genet.* 2012;91:275–92.
53. Sadkowska T, Jank M, Zwierzchowski L, Siadkowska E, Oprzdek J, Motyl T. Gene expression profiling in skeletal muscle of Holstein-Friesian bulls with single-nucleotide polymorphism in the myostatin gene 5'-flanking region. *J Appl Genet.* 2008;49:237–50.
54. McPherron AC, Lawler AM, Lee S-J. Regulation of skeletal muscle mass in mice by a new TGF- β superfamily member. *Nature.* 1997;387:83–90.
55. Lehnert SA, Reverter A, Byrne KA, Wang Y, Nattrass GS, Hudson NJ, et al. Gene expression studies of developing bovine longissimus muscle from two different beef cattle breeds. *BMC Dev Biol.* 2007;7:95.
56. D'Andre CH, Wallace P, Shen X, Nie Q, Yang G, Zhang X. Genes related to economically important traits in beef cattle. *Asian J Anim Sci.* 2011;5:34–45.
57. Akakabe Y, Koide M, Kitamura Y, Matsuo K, Ueyama T, Matoba S, et al. Ecsr regulates insulin sensitivity and predisposition to obesity by modulating endothelial cell functions. *Nat. Commun. Nat Publ Group.* 2013;4: 2389.
58. Risau W. Mechanisms of angiogenesis. *Nature.* 1997;386:671–4.
59. Felius M. Cattle breeds: an Encyclopaedia. First. Doetinchem: Misset; 1995.
60. Karthickeyan SMK, Kumarasamy P, Sivaselvam SN, Saravanan R, Thangaraju P. Analysis of microsatellite markers in Ongole breed of cattle. *Indian J Biotechnol.* 2008;7:113–6.
61. Sanders JO. History and development of zebu cattle in the United States. *J Anim Sci New Orleans.* 1980;50:1188–200.
62. Delbes G, Yanagiya A, Sonenberg N, Robaire B. PABP interacting protein 2 (Paip2) is a major translational regulator involved in the maturation of male germ cells and male fertility. *Biol Reprod Oxford University Press.* 2009;81:167.
63. Guan D, Luo N, Tan X, Zhao Z, Huang Y, Na R, et al. Scanning of selection signature provides a glimpse into important economic traits in goats (*Capra hircus*). *Sci Rep Nat Publ Group.* 2016;6:36372.
64. Silva-Vignato B, Coutinho LL, Cesar ASM, Poleti MD, Regitano LCA, Balieiro JCC. Comparative muscle transcriptome associated with carcass traits of Nelore cattle. *BMC Genomics.* 2017;18:506.
65. Mota RR, Guimarães SEF, Fortes MRS, Hayes B, Silva FF, Verardo LL, et al. Genome-wide association study and annotating candidate gene networks affecting age at first calving in Nelore cattle. *J Anim Breed Genet.* 2017;134:484–92.
66. Utsunomiya YT, do Carmo AS, Carvalheiro R, Neves HH, Matos MC, Zavarez LB, et al. Genome-wide association study for birth weight in Nelore cattle points to previously described orthologous genes affecting human and bovine height. *BMC Genet.* 2013;14:52.
67. Meade KG, Cormican P, Narciandi F, Lloyd A, O'Farrelly C. Bovine β -defensin gene family: opportunities to improve animal health? *Physiol Genomics.* 2014;46:17–28.
68. Dorin JR, Barratt CLR. Importance of β -defensins in sperm function. *Mol Hum Reprod.* 2014;20:821–6.
69. Nelson RA. The immune-adherence phenomenon; an immunologically specific reaction between microorganisms and erythrocytes leading to enhanced phagocytosis. *Science.* 1953;118:733–7.
70. Passantino L, Altamura M, Cianciotta A, Patruno R, Tafaro A, Jirillo E, et al. Fish immunology. I. Binding and engulfment of candida albicans by erythrocytes of rainbow trout (*Salmo gairdneri* Richardson). *Immunopharmacol Immunotoxicol.* 2002;24:665–78.
71. Passantino L, Massaro MA, Jirillo F, Di Modugno D, Ribaud MR, Di Modugno G, et al. Antigenically activated avian erythrocytes release cytokine-like factors: a conserved phylogenetic function discovered in fish. *Immunopharmacol Immunotoxicol.* 2007;29:141–52.
72. Passantino L, Altamura M, Cianciotta A, Jirillo F, Ribaud MR, Jirillo E, et al. Maturation of fish erythrocytes coincides with changes in their morphology, enhanced ability to interact with *Candida albicans* and release of cytokine-like factors active upon autologous macrophages. *Immunopharmacol Immunotoxicol.* 2004;26:573–85.
73. Khansefid M, Millen CA, Chen Y, Pryce JE, Chamberlain AJ, Vander Jagt CJ, et al. Gene expression analysis of blood, liver, and muscle in cattle divergently selected for high and low residual feed intake. *J Anim Sci.* 2017; 95:4764–75.
74. Goszczynski D, Molina A, Terán E, Morales-Durand H, Ross P, Cheng H, et al. Runs of homozygosity in a selected cattle population with extremely inbred bulls: descriptive and functional analyses revealed highly variable patterns. *PLoS One.* 2018;13:e0200069.
75. Commun L, Velek K, Barbry J-B, Pun S, Rice A, Mestek A, et al. Detection of pregnancy-associated glycoproteins in milk and blood as a test for early pregnancy in dairy cows. *J Vet Diagnostic Investig.* 2016;28:207–13.
76. Pohler KG, Geary TW, Johnson CL, Atkins JA, Jinks EM, Busch DC, et al. Circulating bovine pregnancy associated glycoproteins are associated with late embryonic/fetal survival but not ovulatory follicle size in suckled beef cows1. *J Anim Sci.* 2013;91:4158–67.

77. Pohler KG, Pereira MHC, Lopes FR, Lawrence JC, Keisler DH, Smith MF, et al. Circulating concentrations of bovine pregnancy-associated glycoproteins and late embryonic mortality in lactating dairy herds. *J Dairy Sci.* 2016;99:1584–94.
78. Pohler KG, Peres RFG, Green JA, Graff H, Martins T, Vasconcelos JLM, et al. Use of bovine pregnancy-associated glycoproteins to predict late embryonic mortality in postpartum Nelore beef cows. *Theriogenology.* 2016;85:1652–9.
79. Wooding FB, Roberts RM, Green JA. Light and electron microscope immunocytochemical studies of the distribution of pregnancy associated glycoproteins (PAGs) throughout pregnancy in the cow: possible functional implications. *Placenta.* 2005;26:807–27.
80. Breukelman SP, Perényi Z, Taverne MAM, Jonker H, van der Weijden GC, Vos PLAM, et al. Characterisation of pregnancy losses after embryo transfer by measuring plasma progesterone and bovine pregnancy-associated glycoprotein-1 concentrations. *Vet J.* 2012;194:71–6.
81. Wallace RM, Pohler KG, Smith MF, Green JA. Placental PAGs: gene origins, expression patterns, and use as markers of pregnancy. *Reproduction.* 2015;149:R115–26.
82. Mercadante PM, Ribeiro ES, Risco C, Ealy AD. Associations between pregnancy-associated glycoproteins and pregnancy outcomes, milk yield, parity, and clinical diseases in high-producing dairy cows. *J Dairy Sci.* 2016;99:3031–40.
83. Bjelland DW, Weigel K a, Vukasinovic N, Nkrumah JD. Evaluation of inbreeding depression in Holstein cattle using whole-genome SNP markers and alternative measures of genomic inbreeding. *J Dairy Sci Elsevier.* 2013;96:4697–706.
84. Sargolzaei M, Chesnais JP, Schenkel FS. A new approach for efficient genotype imputation using information from relatives. *BMC Genomics.* 2014;15:478.
85. Boichard D. PEDIG : A Fortran package for pedigree analysis suited for large populations. 7th world Congr. Genet. Appl. To Livest. Prod. Montpellier; 2002.
86. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet.* 2007;81:559–75.
87. Aguilar I, Misztal I. Technical note: recursive algorithm for inbreeding coefficients assuming nonzero inbreeding of unknown parents. *J Dairy Sci.* 2008;91:1669–72.
88. VanRaden PM, Olson KM, Wiggans GR, Cole JB, Tooker ME. Genomic inbreeding and relationships among Holsteins, Jerseys, and Brown Swiss. *J Dairy Sci.* 2011;94:5673–82.
89. Haider S, Ballester B, Smedley D, Zhang J, Rice P, Kasprzyk A. BioMart central portal - Unified access to biological data. *Nucleic Acids Res.* 2009;37:23–7.
90. Lawrence M, Huber W, Pagès H, Aboyoun P, Carlson M, Gentleman R, et al. Software for computing and annotating genomic ranges. *PLoS Comput Biol.* 2013;9:e1003118.
91. van Hulzen KJE, Schopen GCB, van Arendonk JAM, Nielen M, Koets AP, Schrooten C, et al. Genome-wide association study to identify chromosomal regions associated with antibody response to *Mycobacterium avium* subspecies paratuberculosis in milk of Dutch Holstein-Friesians. *J Dairy Sci Elsevier.* 2012;95:2740–8.
92. Taye M, Lee W, Jeon S, Yoon J, Dessie T, Hanotte O, et al. Exploring evidence of positive selection signatures in cattle breeds selected for different traits. *Mamm Genome.* 2017;28:528–41.
93. Sorbolini S, Bongiorno S, Cellesi M, Gaspa G, Dimauro C, Valentini A, et al. Genome wide association study on beef production traits in Marchigiana cattle breed. *J Anim Breed Genet.* 2017;134:43–8.
94. Guo Y, Zhang X, Huang W, Miao X. Identification and characterization of differentially expressed miRNAs in subcutaneous adipose between wagu and Holstein cattle. *Sci Rep.* 2017;7:44026.
95. Borowska A, Reyer H, Wimmers K, Varley PF, Szwaczkowski T. Detection of pig genome regions determining production traits using an information theory approach. *Livest Sci.* 2017;205:31–5.
96. Schellander K. Identifying genes associated with quantitative traits in pigs: integrating quantitative and molecular approaches for meat quality. *Ital J Anim Sci.* 2009;8:19–25.
97. Utsunomiya YT, Pérez O'Brien AM, Sonstegard TS, Van Tassell CP, do Carmo AS, Mészáros G, et al. Detecting loci under recent positive selection in dairy and beef cattle by combining different genome-wide scan methods. *PLoS One.* 2013;8:e64280.
98. Cong X, Doering J, Mazala DAG, Chin ER, Grange RW, Jiang H. The SH3 and cysteine-rich domain 3 (Stac3) gene is important to growth, fiber composition, and calcium release from the sarcoplasmic reticulum in postnatal skeletal muscle. *Skelet Muscle.* 2016;6:17.
99. Rincon G, Farber EA, Farber CR, Nkrumah JD, Medrano JF. Polymorphisms in the STAT6 gene and their association with carcass traits in feedlot cattle. *Anim Genet.* 2009;40:878–82.
100. Kawaguchi F, Kigoshi H, Nakajima A, Matsumoto Y, Uemoto Y, Fukushima M, et al. Pool-based genome-wide association study identified novel candidate regions on BTA9 and 14 for oleic acid percentage in Japanese black cattle. *Anim Sci J.* 2018;8:1060–6.
101. Lindholm-Perry AK, Kuehn LA, TPL S, Ferrell CL, Jenkins TG, Freetly HC, et al. A region on BTA14 that includes the positional candidate genes LYPLA1, XKR4 and TMEM68 is associated with feed intake and growth phenotypes in cattle. *Anim Genet.* 2012;43:216–9.
102. Porto Neto LR, Bunch RJ, Harrison BE, Barendse W. Variation in the XKR4 gene was significantly associated with subcutaneous rump fat thickness in indicine and composite cattle. *Anim Genet.* 2012;43:785–9.
103. Brisard D, Desmarchais A, Touzé J-L, Lardic L, Freret S, Elis S, et al. Alteration of energy metabolism gene expression in cumulus cells affects oocyte maturation via MOS-mitogen-activated protein kinase pathway in dairy cows with an unfavorable "Fertile" haplotype of one female fertility quantitative trait locus. *Theriogenology.* 2014;81:599–612.
104. Reverchon M, Rame C, Bunel A, Chen W, Froment P, Dupont J. VISFATIN (NAMPT) improves in vitro IGF1-induced steroidogenesis and IGF1 receptor signaling through SIRT1 in bovine granulosa cells. *Biol Reprod.* 2016;94:54.
105. Höglund JK, Sahana G, Gulbrandsen B, Lund MS. Validation of associations for female fertility traits in Nordic Holstein, Nordic Red and Jersey dairy cattle. *BMC Genet BioMed Central.* 2014;15:8.
106. Li CH, Gao Y, Wang S, Xu FF, Dai LS, Jiang H, et al. Expression pattern of JMJD1C in oocytes and its impact on early embryonic development. *Genet Mol Res.* 2015;14:18249–58.
107. Gao Y, Gautier M, Ding X, Zhang H, Wang Y, Wang X, et al. Species composition and environmental adaptation of indigenous Chinese cattle. *Sci Rep.* 2017;7:16196.
108. dos Santos FC, Peixoto MGCD, Fonseca PA de S, Pires M de FÁ, Ventura RV, Rosse I da C, et al. Identification of candidate genes for reactivity in Guzerat (*Bos indicus*) cattle: a genome-wide association study. *Davoli R, editor. PLoS One Public Libr Sci;* 2017;12:e0169163.
109. Garza-Brenner E, Sifuentes-Rincón AM, Randel RD, Paredes-Sánchez FA, Parra-Bracamonte GM, Arellano Vera W, et al. Association of SNPs in dopamine and serotonin pathway genes and their interacting genes with temperament traits in Charolais cows. *J Appl Genet.* 2017;58:363–71.
110. Taye M, Lee W, Caetano-Anolles K, Dessie T, Hanotte O, Mwai OA, et al. Whole genome detection of signature of positive selection in African cattle reveals selection for thermotolerance. *Anim Sci J.* 2017;88:1889–901.
111. Kim E-S, Elbeltagy AR, Aboul-Naga AM, Rischkowsky B, Sayre B, Mwacharo JM, et al. Multiple genomic signatures of selection in goats and sheep indigenous to a hot arid environment. *Heredity.* 2016;116:255–64.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

