

RESEARCH ARTICLE

Open Access



# A multi-breed GWAS for morphometric traits in four Beninese indigenous cattle breeds reveals loci associated with conformation, carcass and adaptive traits

Sèyi Fridaïus Ulrich Vanvanhossou<sup>1</sup>, Carsten Scheper<sup>1</sup>, Luc Hippolyte Dossa<sup>2</sup>, Tong Yin<sup>1</sup>, Kerstin Brügemann<sup>1</sup> and Sven König<sup>1\*</sup>

## Abstract

**Background:** Specific adaptive features including disease resistance and growth abilities in harsh environments are attributed to indigenous cattle breeds of Benin, but these breeds are endangered due to crossbreeding. So far, there is a lack of systematic trait recording, being the basis for breed characterizations, and for structured breeding program designs aiming on conservation. Bridging this gap, own phenotyping for morphological traits considered measurements for height at withers (HAW), sacrum height (SH), heart girth (HG), hip width (HW), body length (BL) and ear length (EL), including 449 cattle from the four indigenous Benin breeds Lagune, Somba, Borgou and Pabli. In order to utilize recent genomic tools for breed characterizations and genetic evaluations, phenotypes for novel traits were merged with high-density SNP marker data. Multi-breed genetic parameter estimations and genome-wide association studies (GWAS) for the six morphometric traits were carried out. Continuously, we aimed on inferring genomic regions and functional loci potentially associated with conformation, carcass and adaptive traits.

**Results:** SNP-based heritability estimates for the morphometric traits ranged between  $0.46 \pm 0.14$  (HG) and  $0.74 \pm 0.13$  (HW). Phenotypic and genetic correlations ranged from  $0.25 \pm 0.05$  (HW-BL) to  $0.89 \pm 0.01$  (HAW-SH), and from  $0.14 \pm 0.10$  (HW-BL) to  $0.85 \pm 0.02$  (HAW-SH), respectively. Three genome-wide and 25 chromosome-wide significant SNP positioned on different chromosomes were detected, located in very close chromosomal distance ( $\pm 25$  kb) to 15 genes (or located within the genes). The genes *PIK3R6* and *PIK3R1* showed direct functional associations with height and body size. We inferred the potential candidate genes *VEPH1*, *CNTNAP5*, *GYPC* for conformation, growth and carcass traits including body weight and body fat deposition. According to their functional annotations, detected potential candidate genes were associated with stress or immune response (genes *PTAFR*, *PBRM1*, *ADAMTS12*) and with feed efficiency (genes *MEGF11*, *SLC16A4*, *CCDC117*).

(Continued on next page)

\* Correspondence: [sven.koenig@agr.uni-giessen.de](mailto:sven.koenig@agr.uni-giessen.de)

<sup>1</sup>Institute of Animal Breeding and Genetics, Justus-Liebig-University Gießen, Gießen, Germany

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



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

(Continued from previous page)

**Conclusions:** Accurate measurements contributed to large SNP heritabilities for some morphological traits, even for a small mixed-breed sample size. Multi-breed GWAS detected different loci associated with conformation or carcass traits. The identified potential candidate genes for immune response or feed efficiency indicators reflect the evolutionary development and adaptability features of the breeds.

**Keywords:** Endangered cattle breeds, Morphometric traits, Multi-breed GWAS, SNP-based genetic parameters, Functional annotations, Potential candidate genes

## Background

In Sub-Saharan Africa, the cattle breeding sector comprises smallholder subsistence farms in extensive production systems, and only a few exclusively market-oriented productions. In consequence, there is a substantial gap with regard to organized breeding programs and routine performance testing, technical support and precise breeding policies. Against such a background, applying advanced breeding technologies combined with enhanced genomic statistical methods in order to conserve animal genetic resources remains a major challenge, as observed in many developing countries [1, 2]. First approaches to utilize genomic information in animal breeding in Africa were recently made, but based on limited resources [2, 3]. Nevertheless, the African continent is well-endowed with hundreds of indigenous, often endangered, animal genetic resources. In addition to being the source of livelihood for millions of poor farmers, these genetic resources represent specific socio-economic values, and they are unique for traits indicating adaptation to harsh environments [4–8].

High-throughput genotyping and genomic methods provide new opportunities for the genetic characterization and genetic management of African indigenous animal breeds. An elementary genetic and phenotypic evaluation is the prerequisite for future performance improvements. Such efforts might be the key to successful management and conservation of animal genetic resources in the light of policy developments, climate change, and diversifying market demands [9–11]. In this regard, studies addressing selection signatures analyses for heat tolerance, thermoregulation, tick resistance and trypanotolerance in African cattle breeds reflect the unique potential and advantage they possess [12–15]. Therefore, more efforts are needed to investigate the genetic architecture of functional and performance traits in native African breeds, in order to unravel their potential for future breeding development.

A great variety of study designs and methods for GWAS as established in recent years are powerful tools to study the genomic architecture of both qualitative and quantitative traits. Consequently, series of GWAS have been performed for performance and functional traits in various livestock species and breeds in European, American and Asian countries, e.g. [16–19]. In

Africa, however, only a few GWAS in livestock have been reported to date [20–22], due to the difficulties in collecting valid phenotypic data in the smallholder production system, as well as the lack of resources and technologies for genotyping [23].

Three main factors are decisive to design a reliable GWAS: i) accurate phenotype and genotype data, ii) sufficient sample selection and sample size, and iii) application of adequate statistical methods [24]. With regard to further challenges such as population stratification, environmental influences and the complexity of quantitative traits, statistical methods are gradually enhanced [25, 26]. The broad availability of open-source software packages implementing innovative methods such as PLINK [27] and GCTA [28] for genome-wide complex trait analyses, in combination with continuously declining genotyping costs, open the potential for pilot GWAS in unstudied and undeveloped breeds for novel traits. Nevertheless, the recording of accurate phenotypes remains one big challenge in the African livestock-breeding context. The Food and Agriculture Organization of the United Nations (FAO) guidelines highlight morphometric traits as a good starting point to initially characterize unstudied breeds phenotypically and genetically [29]. Actually, morphometric traits comprise linear measurements of animal body sizes. Morphometric traits are routinely measurable even at early ages, and are proper early indicators for animal growth, health, welfare, and longevity [30–32]. Moreover, because of the accurate measurements (mostly in cm), morphometric traits provide an objective and better assessment of body traits than subjectively conformation traits scoring [33–35].

Several GWAS associated single nucleotide polymorphism (SNP) and genes to cattle morphometric or conformation traits. Major and commonly investigated cattle body traits for GWAS were hip height (stature), height at withers, body length, hip (rump) width, chest width, and scrotal circumference [36–38]. The heritabilities generally reported for these traits were moderate to large [39–41]. Moreover, despite the fact that the genomic architecture of cattle body traits is highly polygenic, many similarities with other livestock species, human and mammals in general, were observed [42, 43]. The identified regions were mainly involved in biological functions such as regulation of fetal growth, skeletal

development, regulation of cell cycle or cell division, homeostasis, and lipid metabolism [40–44]. More interestingly, different studies identified significant overlaps in genomic architecture, and in genomic relationships between morphometric or conformation traits with cattle performance traits including body weight, carcass trait, feed intake, reproduction and health [42–46]. Indeed, phenotypic correlations between morphometric traits and animal performances such as body weight and milk offtake were very similar in various African breeds [47–49]. Furthermore, models to predict body weight from heart girth, body length or height at withers have been established [47, 48], suggesting morphometric traits as major predictors of animal performances in African livestock herds, where recording systems are poorly developed. Addressing principles of selection, Kabi et al. [50] indicated that morphometric population diversity is a result of selection for adaptive and sociocultural interests in African smallholder livestock production context. Hence, investigating genomic regions associated with morphometric traits is worthwhile in African breeds, as it allows a better understanding of animal diversity and adaptation features. In addition, it opens prospects for the effective use of morphometric traits in basic phenotype recording for any potential community-based breeding program in African smallholder livestock systems.

As discussed by many scientists [24, 51], GWAS are exploratory in their nature and further investigations based on SNP marker effects are required to get more insights into the genetic and biological basis of a trait. Hou and Zhao [52] reviewed tools and genomic features such as differential gene expressions, protein deleteriousness predictions and DNase I hypersensitive sites that could be used in understanding biological causal mechanisms and the functional relevance of identified significant SNP. In the context of limited resources, approaches that rely on documented information and public databases, such as candidate gene functional annotation and enrichment analyses of gene ontology (GO), even though non-analytical, offer new prospects for a deeper interpretation of results from GWAS [53].

Cattle in Benin are commonly kept in small herds under extensive production conditions. The indigenous cattle breeds of Benin consist of two taurine (Somba and Lagune) and two hybrid (taurine x indicine, i.e., Borgou and Pabli) breeds. These breeds were described for their adaptive potential to disease and harsh environmental conditions as well as for their importance in the livelihoods of poor farmers [5, 14, 54, 55]. However, due to their low productivity, they are increasingly threatened by indiscriminate crossbreeding with zebu animals [5, 56]. In addition, routine performance recordings and structured breeding programs have not been developed yet. The existing threats were confirmed in a genetic diversity approach considering indigenous cattle breeds from Benin, focusing on the effects of transboundary transhumance [57]. The present study builds on the dataset established in Scheper et al. [57], and combines 50 k SNP data with a basic phenotypic characterization according to FAO guidelines [29]. Given the threats surrounding the indigenous cattle breeds in Benin and in Sub-Saharan Africa in general, genetic evaluations based on marker data are an important first step to develop sustainable conservation and breeding strategies.

The aim of the present study was to estimate genetic parameters and to perform genome-wide associations for morphometric traits in four indigenous cattle breeds from different agro-ecological zones (AEZ) of Benin using medium density SNP chip data. In addition, functional annotation and gene enrichment analyses were applied to identify genes and functional loci potentially associated with morphometric traits. Finally, the indigenous Benin breeds from smallholder farms were contrasted genomically with other African livestock and exotic breeds or crossbreeds raised under improved management conditions in research stations [58, 59].

## Results

### Heritabilities, phenotypic and genetic correlations

SNP-based heritability estimates for the morphometric traits ranged between  $0.46 \pm 0.14$  (HG) and  $0.74 \pm 0.13$  (HW, EL, Table 1). Heart girth showed the largest genetic correlations ( $r_g$ ) with all other morphometric traits (0.38–0.80). Overall, genetic correlations among all

**Table 1** Estimated phenotypic and genetic correlations among morphometric traits and their heritability. Heritabilities (in bold) are on the diagonal, above the diagonal are the genetic correlations and below the diagonal are the phenotypic correlations

	HAW	SH	HG	HW	BL	EL
HAW	<b>0.72 ± 0.08</b>	0.85 ± 0.02	0.80 ± 0.28	0.46 ± 0.05	0.25 ± 0.08	0.42 ± 0.11
SH	0.89 ± 0.01	<b>0.70 ± 0.08</b>	na	0.46 ± 0.05	0.19 ± 0.09	0.42 ± 0.13
HG	0.62 ± 0.03	0.59 ± 0.03	<b>0.42 ± 0.14</b>	0.72 ± 0.15	0.38 ± 0.15	0.58 ± 0.25
HW	0.50 ± 0.04	0.49 ± 0.04	0.58 ± 0.03	<b>0.74 ± 0.13</b>	0.14 ± 0.10	0.35 ± 0.17
BL	0.37 ± 0.04	0.32 ± 0.04	0.45 ± 0.04	0.25 ± 0.05	<b>0.73 ± 0.10</b>	0.25 ± 0.10
EL	0.35 ± 0.04	0.33 ± 0.04	0.29 ± 0.04	0.25 ± 0.05	0.33 ± 0.04	<b>0.74 ± 0.10</b>

HAW height at withers, SH sacrum height, HG heart girth, HW hip width, BL body length and EL ear length

morphometric traits ranged from 0.14 (HW with BL) to 0.85 (HAW with SH). Most of the estimated genetic correlations had small standard errors below 0.11. The largest standard error ( $SE = 0.28$ ) was estimated for the genetic correlation between HG and HAW. In general, HG had larger SE for genetic correlations and for the heritability in comparison to all other traits. The phenotypic correlations ( $r_p$ ) ranged between  $0.25 \pm 0.05$  (HW with BL) and  $0.89 \pm 0.01$  (HAW with SH). Heart girth and hip width were phenotypically and genetically highly correlated ( $r_p = 0.62$ ,  $r_g = 0.72$ ). In contrast, phenotypic and genetic correlations between EL and BL ( $r_p = 0.33$ ,  $r_g = 0.25$ ) were considerably lower. In addition, BL showed the smallest phenotypic and genetic correlations with other traits, especially with HW ( $r_g = 0.14$ – $0.38$ ,  $r_p = 0.25$ – $0.45$ ).

#### Multi-breed GWAS for morphometric traits and functional annotation of candidate genes

Discriminant analysis of principal components (DAPC) identified four linear discriminant functions (LDF) validly representing the genetic structure in the sampled population (Additional file 1, Figure S1). The inclusion of the LDF in the PLINK GWAS resulted in sufficient correction of population stratification with desired lambda values ( $\lambda = 0.99$ – $1.09$ ). Slightly lower lambda values ( $\lambda = 0.98$ – $1.00$ ) were obtained from GWAS using GCTA (Fig. 1). The GWAS via PLINK detected a total of 28 SNP for all six morphometric traits, and the majority of these SNP were also detected via GCTA (Fig. 1). The significant SNP from PLINK were positioned within or near 15 different genes (Table 2).

The significant SNP, potential candidate genes and their functional annotation are presented in the ongoing sub-chapters for the morphometric traits, considering two traits per sub-chapter.

#### Height at withers and sacrum height

Height at withers was significantly associated with the SNP *rs109126926* on BTA1 ( $p = 1.12e-06$ ). In addition, five chromosome-wide suggestively associated SNP were detected on chromosomes 3, 17, 19 and 21 ( $p = 6.44e-06$  -  $5.04e-05$ ); see Table 2 for the exact positions). The SNP *rs110369628* and *rs109889052* on BTA19 were positioned in relative proximity to each other (7.41 Mb), within the *PIK3R6* and *SSH2* genes, respectively. The *VEPH1* gene harbored the significantly associated SNP for HAW on BTA1, while another SNP, *rs4163436*, was positioned in *CCDC117* on BTA17. No genes were annotated for the two remaining significantly associated SNP.

Only two SNP were suggestively associated with SH ( $p = 2.35e-05$  -  $3.86e-05$ ) and each was located in close distance to a gene. The SNP *rs111001850* was positionally linked to the *LYPD8* gene on BTA7, and the SNP *rs110441360* was located near the *PIK3R1* gene on BTA20.

#### Heart girth and hip width

No SNP surpassed the genome-wide significance threshold for associations with HG and HW. However, four SNP were suggestively associated with HG ( $p = 1.17e-05$  -  $4.10e-05$ ). Two of them were positioned on BTA2 and the two others were located on BTA19 and BTA22. The SNP *rs41579167* on BTA2 was positioned near the two genes *PTAFR* and *EYA3*. The gene *PBRM1* harbored the SNP *rs41637645* on BTA22.

Hip width was associated with three suggestive SNP ( $p = 2.08e-05$  -  $2.99e-05$ ). Two of the three SNP were positioned relatively near to each other (4 Mb) on BTA16. The third SNP was detected on BTA8. Only the *ABL2* gene was mapped as a potential candidate gene for HW, harboring the SNP *rs42843320* on BTA16.

#### Body length and ear length

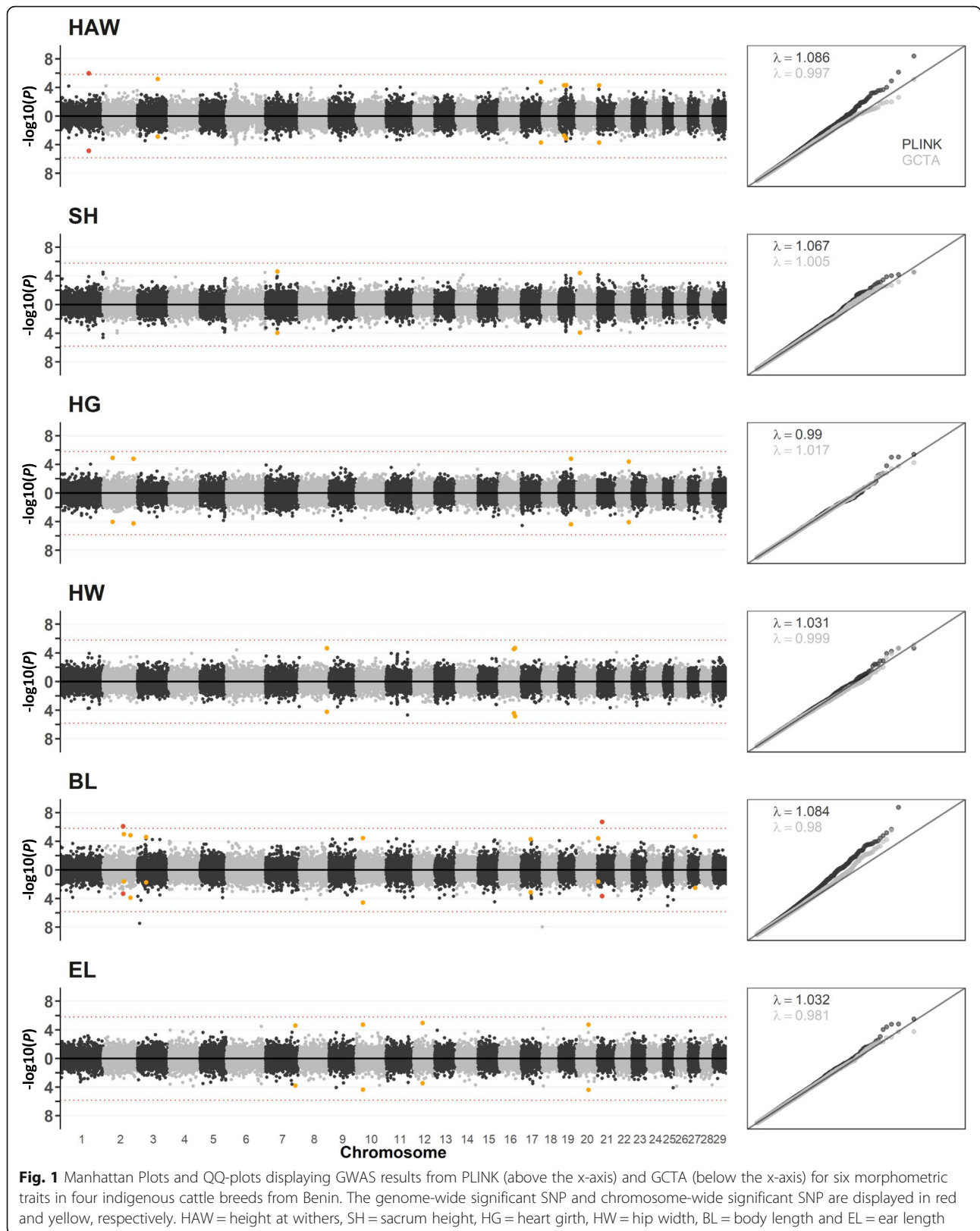
Body length presented the highest number of associated SNP with two significant and 7 suggestive SNP. BTA2 harbored one significant SNP (*rs110694334*,  $p = 8.19e-07$ ) along with two others suggestive SNP ( $p = 9.33e-06$  -  $1.38e-05$ ). Similarly, BTA21 harbored the second significant SNP (*rs41607390*,  $p = 2.03e-07$ ) and one suggestive SNP ( $p = 3.74e-05$ ), but none of them was positionally linked to a gene. The other associated SNP variants were identified on chromosomes 3, 10, 17 and 27 ( $p = 1.91e-05$  -  $4.93e-05$ ). On BTA2, the SNP *rs110694334* was positioned within the *CNTNAP5* gene, in relative proximity (2.32 Mb) to the SNP *rs109186122* positioned in the *GYPC* gene. Two additional genes, *SLC16A4* and *MEGF11*, were mapped as potential candidates for BL, harboring the SNPs *rs135705191* and *rs43616983* on BTA3 and BTA10, respectively.

Four SNP positioned on different chromosomes (7, 10, 12, 20) were suggestively associated with EL ( $p = 1.08e-05$  -  $2.54e-05$ ). The SNP *rs110608572* SNP was positioned on BTA10 within the *MEGF11* gene, which was also identified as a potential candidate gene for BL. The SNP *rs109985119* was located in the *ADAMTS12* gene on BTA20.

## Discussion

### Heritabilities, phenotypic and genetic correlations

The heritability estimates for morphometric traits obtained in the present study were larger than those commonly reported for comparable linear body traits [39–41, 60–63]. However, heritability estimates are usually higher for morphometric traits (measurements) [39–41] than for conformation traits (scores) [60, 61, 63]. Hence, quite large heritability estimates (up to 0.7) have been similarly reported in few studies based on morphometric traits [64–67]. For instance, the heritability estimate of 0.42 for HG in our study is comparable to the value of 0.43 reported for adult Brahman cattle [66]. Moreover, the heritabilities of 0.72 for HAW and of 0.70 for SH are



**Table 2** Genome-wide and chromosome-wide significant SNP and potential candidate genes associated with six morphometric traits from four indigenous breeds in Benin

SNP_rs	SNP	CHR	BP_ARS1.2	Alleles		MAF	BETA		P-value		SNP in Gene	Gene Name
				A1	A2		PLINK	GCTA	PLINK	GCTA		
Height at withers												
rs109126926 <sup>a</sup>	ARS-BFGL-NGS-57889	1	110,160,486	A	G	0.26	2.24	2.16	1.12e-06	1.44e-05	Yes	VEPH1
rs43347748	ARS-BFGL-NGS-31952	3	81,139,242	C	T	0.10	2.60	1.96	6.44e-06	1.40e-03	-	-
rs41634361	Hapmap50686-BTA-41836	17	68,037,062	G	A	0.10	2.58	2.40	1.71e-05	1.99e-04	Yes	CCDC117
rs110369628	ARS-BFGL-NGS-91812	19	20,892,297	T	C	0.02	5.58	4.55	4.82e-05	2.00e-03	Yes	SSH2
rs109889052	ARS-BFGL-NGS-81151	19	28,301,046	A	G	0.13	2.35	2.08	4.57e-05	7.94e-04	Yes	PIK3R6
rs109872376	ARS-BFGL-NGS-46597	21	7,674,101	C	G	0.44	1.62	1.54	5.04e-05	2.09e-04	-	-
Sacrum height												
rs111001850	ARS-BFGL-NGS-7310	7	42,545,291	A	G	0.45	1.53	1.50	2.35e-05	1.12e-04	No	LYPD8
rs110441360	ARS-BFGL-NGS-110086	20	11,381,378	A	G	0.48	-1.51	-1.51	3.86e-05	1.19e-04	No	PIK3R1
Heart girth												
rs110404606	ARS-BFGL-NGS-77689	2	32,163,209	C	T	0.38	2.83	2.55	1.17e-05	9.65e-05	-	-
rs41579167	BTA-49621-no-rs	2	125,244,296	G	A	0.44	-2.93	-2.75	1.55e-05	5.59e-05	No	PTAFR EYA3
rs41624005	Hapmap48676-BTA-18047	19	46,729,603	A	G	0.50	2.72	2.64	1.60e-05	4.29e-05	-	-
rs41637645	Hapmap39844-BTA-54797	22	48,208,654	C	A	0.16	-3.48	-3.36	4.10e-05	8.96e-05	Yes	PBRM1
Hip width												
rs109866742	ARS-BFGL-NGS-119529	8	103,907,956	T	C	0.02	-3.12	-3.06	2.11e-05	5.89e-05	-	-
rs41569598	BTA-39611-no-rs	16	56,426,595	C	T	0.43	-0.97	-0.97	2.99e-05	3.88e-05	-	-
rs42843320	BTB-01732320	16	60,525,984	C	T	0.22	-1.12	-1.15	2.08e-05	1.47e-05	Yes	ABL2
Body length												
rs110694334 <sup>a</sup>	ARS-BFGL-NGS-109828	2	76,610,609	C	T	0.03	-17.96	-11.94	8.19e-07	4.90e-04	yes	CNTNAP5
rs109186122	ARS-BFGL-NGS-118432	2	78,925,610	C	T	0.25	-7.18	-3.43	9.33e-06	2.48e-02	yes	GYPC
rs42301516	BTB-01145402	2	113,155,761	G	A	0.43	-5.45	-4.31	1.38e-05	1.37e-04	-	-
rs135705191	BovineHD0300010335	3	33,048,892	G	A	0.20	-6.85	-3.72	2.49e-05	1.90e-02	yes	SLC16A4
rs43616983	BTB-00409355	10	12,969,418	G	A	0.20	6.70	6.20	3.41e-05	2.94e-05	yes	MEGF11
rs42436268	BTB-01308172	17	33,487,976	G	T	0.32	5.45	4.06	4.93e-05	8.18e-04	-	-
rs41608167	BTA-96370-no-rs	21	3,408,269	T	C	0.02	-18.36	-9.53	3.74e-05	2.29e-02	-	-
rs41607390 <sup>a</sup>	Hapmap33092-BTA-51753	21	18,775,375	T	C	0.11	-10.39	-7.07	2.03e-07	2.20e-04	-	-
rs41646754	Hapmap44720-BTA-62525	27	24,546,315	G	C	0.27	-6.28	-4.04	1.91e-05	3.27e-03	-	-
Ear length												
rs109212458	ARS-BFGL-NGS-85383	7	96,266,192	T	C	0.33	0.42	0.38	2.54e-05	1.63e-04	-	-
rs110608572	ARS-BFGL-NGS-103122	10	12,919,428	A	G	0.03	1.10	1.07	1.86e-05	4.40e-05	Yes	MEGF11
rs41567897	BTA-100327-no-rs	12	41,375,572	C	G	0.17	-0.59	-0.49	1.08e-05	3.56e-04	-	-
rs109985119	ARS-BFGL-NGS-44763	20	40,060,386	T	C	0.14	0.56	0.54	1.90e-05	4.23e-05	Yes	ADAMTS12

<sup>a</sup>Genome-wide significant SNP; the remaining SNP are the chromosome-wide significant SNP

in agreement with the estimates for stature in Brahman (0.73) [65] and in Red & White cattle (0.74) [64]. These observations suggest that real and objective measurements provide a better basis for heritability estimations than subjective conformation scores.

The estimates of genomic correlations for the different trait combinations ranged from 0.14 to 0.84, with

standard errors between 0.02 to 0.25. Roveglia et al. [63] reported a broad range (0.08 to 0.98) for genetic correlations among conformation traits reflecting the morphometric characteristics from the present study. The genetic correlation between HW and height traits (0.46) in the present study was smaller than the estimate of 0.75 [63]. Nevertheless, generally lower genetic

correlations (0.09 to 0.35) between these trait combinations have been previously described [64, 68]. The genetic correlation of 0.80 between HG and HAW from the present study is larger than the values that were previously reported (0.30–0.65) [40, 64, 69, 70]. The large genetic correlations between HG with the other morphometric traits are in agreement with results from a previous study [64].

The estimates of phenotypic correlations between height traits (HAW and SH) and HG (0.59 and 0.62) were slightly larger than those (0.20 to 0.50) reported in previous studies [40, 70]. However, regarding the comparisons made, it is imperative to state that our study is the only one using genomic instead of pedigree relationships. To our knowledge, this is the first study, which estimated phenotypic and genetic correlations among morphometric or conformation traits on the basis of SNP marker data. Moreover, the cited genetic parameters estimates used for the comparisons are from Holstein, Brahman, Brown Swiss and Red & White populations kept in Europe or Asia [40, 63–66]. The few genetic evaluations in the African context were made for growth traits (birth, weaning and yearling weight) in exotic breeds kept under controlled management conditions in research stations [58, 59]. In African smallholder farms, genetic parameters only have been estimated for milk production, considering exotic breeds and cross-breeds (exotic × indigenous) [71].

As a major difference to previous studies, our approach focused on multi-breed genetic parameter estimations. Meyer et al. [72] observed larger heritabilities and genetic variances in multi-breed populations, compared with estimates in purebred populations. These findings are in agreement with the moderate to high values for genetic parameters obtained in our study. However, high genetic variability also was identified in single breeds, but potentially be biased due to correlations between environments and genotypes, and due to extremely close genetic relationships [73]. Thus, accounting for population structure particularities remains a great challenge in all studies estimating genetic parameters, even within a single breed [73]. Our study attempted to avoid relatedness in the dataset by sampling animals from different herds and AEZ. Nevertheless, breed admixture was observed between and within the indigenous breeds in Benin, due to continuous indigene introgression and uncontrolled crossbreeding [57]. Thus, breed specific modelling approaches might be affected from within-breed diversity. Overall, the heritabilities, phenotypic and genetic correlations from the present multi-breed study reflect the range for single-breed estimates, indicating that most of the genetic variations in body traits is based on universal common

rather than on breed specific genetic variants [74]. Indeed, multi-breed population references enhance the potential in detecting common and conservative loci [72, 75]. In this regard, Wientjes et al. [76] focused on model comparison, and observed most accurate genetic parameters from models capturing large proportions of the genetic variance in a population. Therefore, in the context of uncontrolled crossbreeding as prevalent in indigenous cattle breeds in Africa, a multi-breed approach might be useful.

### Multi-breed GWAS for conformation traits and functional annotation of candidate genes

#### *Height at withers and sacrum height*

For the SNP *rs109126926* SNP significantly associated with HAW and the corresponding gene *VEPH1*, no direct association between the variant or the identified gene with cattle height were previously reported. However, *rs109126926* was significantly associated with recoverability from mastitis in Holstein cows [77]. *VEPH1* influenced other cattle conformation traits such as udder cleft in Holstein and rump fat thickness in Nellore cattle, reflecting its role in lipid metabolism [78, 79]. Furthermore, recent studies observed an association of *VEPH1* with residual feed intake and antibody response to parasites in cattle [80, 81].

Among the five suggestive SNP associated with HAW, *rs109889052* and *rs110369628* were located on BTA19. These two SNP have not been directly associated with cattle height, but *rs109889052* contributed to feed efficiency in cattle [82]. Moreover, *rs109889052* is located in the *PIK3R6* gene, affecting body size in sheep [31]. Likewise, *rs110369628* positioned in the *SSH2* gene, was linked to carcass traits in sheep [83]. According to the database for annotation, visualization and integrated discovery (DAVID), *SSH2* is involved in the regulation of the actin cytoskeleton pathway and in several other molecular and biological functions such as the regulation of actin polymerization or depolymerization and DNA binding. In previous studies, *SSH2* was associated with somatic cell score and semen traits in cattle [84, 85]. Bouwman et al. [42] identified several SNP on BTA19 associated with cattle stature, including *rs132693733* (19:21339030), *rs109018020* (19:23816722) and *rs137732346* (19:25980624). These markers are positioned in a distance between 0.45 Mb and 6.96 Mb from the SNP identified in our study, i.e., *rs109889052* (19:20892297) and *rs110369628* (19:28301046). Another variant *rs42741630* (19:25439551) in close proximity (2.86 Mb) to *rs110369628* (19:28301046), was detected in a recent GWAS for stature [86]. These findings suggest that this chromosomal segment on region BTA19 (at 20 Mb to 30 Mb) might represent a hotspot genomic region for height. We found no reference for the three other variants associated with HAW in the literature. Nevertheless, the SNP

*rs41634361* on BTA17 is positioned within the *CCDC117* gene, which is involved in feed intake and heat stress regulation in cattle [87, 88].

The two SNP suggestively associated with SH were not associated with cattle height in previous studies. However, the SNP *rs110441360* on BTA20 is positioned in the *PIK3R1* gene, which influenced beef fat content [89]. More interestingly, different studies described *PIK3R1* for its implication in the human short stature syndrome [90, 91], confirming the functional conservation of genes linked to stature or body size in cattle and humans [42–44]. Moreover, these findings are consistent with the functional annotation of *PIK3R1* in protein stabilization, insulin resistance, and growth hormone receptor signaling pathways. The potential effect of growth hormone receptor genes on animal conformation traits has been previously reported [92]. The second variant (*rs111001850*) associated with SH is located in the *LYPD8* gene. This gene is involved in biological processes of defense responses to gram-negative bacteria, and has been related to adaptive responses to environmental stimuli such as stress, infection and inflammation in cattle [93].

The association of two PIK regulatory subunit genes with height traits in this study is remarkable. Another *PI3K regulatory subunit* gene, i.e. *PIK3R5*, was identified in a region highly associated with body size in sheep [31]. Moreover, *PIK3R6* and *PIK3R1* are declared as potential candidates for feed intake and feed efficiency in cattle [82, 94]. Both genes are members of a metabolism pathway, specifically involved in the synthesis of PIPs at the plasma membrane. In addition, we found that *PIK3R6* and *PIK3R1* are involved in several animal immune system pathways or related biological processes, such as the B cell receptor signaling pathway, the AMPK signaling pathway and regulation of T cell differentiation. Moreover, they are members of different signal transduction pathways (G beta: gamma signaling through PI3Kgamma, signaling by SCF-KIT), which are also linked to inflammatory diseases [95, 96]. The involvement of *PIK3R6* and *PIK3R1* in signal transduction and immunity pathways reflect the associations with resistance to *Mycobacterium avium ssp. paratuberculosis* (MAP) infections in cattle [27–29].

Indigenous cattle breeds in Benin, especially the taurine breeds, are known to be resistant against diseases [14, 54, 55]. In comparison to other breeds, they are small sized, and kept in harsh environments with limited feed resources and a high disease infection risk. In this context, HAW is described as an indicator of animal adaptive attributes [29]. Hence, our findings for HAW and SH confirm that *PIK3R6* and *PIK3R1* may play an important role in the interaction of adaptation to diseases and body size variability in

African cattle, and potentially in cattle populations worldwide.

#### **Heart girth and hip width**

The four SNP suggestively associated with HG were not detected in any previous GWAS for HG. However, two of them are positionally linked to three genes influencing dairy cattle traits. First, the SNP *rs41579167* SNP is located in direct proximity to the genes *PTAFR* and *EYA3*. These two genes are involved in inflammatory responses and both are associated with immune response in cattle [97, 98]. In addition, *PTAFR* is associated with intramuscular fat deposition in Nellore cattle [99] and was identified in selective regions for production performance in different cattle breeds [100, 101]. Furthermore, in mice, *PTAFR* affected body weight by controlling feed intake and obesity [102]. According to DAVID, the *EYA3* gene is involved in cell differentiation processes and in a pathway related to DNA repair mechanisms. This gene was also identified as a potential candidate gene in a GWAS for milk production in dairy sheep [103]. The SNP *rs41579167* SNP (2:125244296) is located at 0.15 Mb distance to another variant *rs2083797338* (2:125093797), which has been identified in a meta-GWAS for cattle stature [42].

Second, the SNP *rs41637645* is located within *PBRM1*, a gene which was associated with heat stress regulation in tropical breeds [104]. *PBRM1* is also involved in the negative regulation of cell proliferation and in the RMTs methylate histone arginine pathway, while this pathway is involved in different diseases in mammals [105].

We found no direct link between the four suggestively associated SNP for HW with HW in previous studies. Nonetheless, *ABL2* is a potential candidate gene for HW, harboring *rs42843320*. *ABL2* influenced feed intake in cattle and backfat thickness in pigs [106, 107]. In addition, *ABL2* is involved in innate immune response processes through cell proliferation, migration and differentiation (according to its functional annotation in the DAVID database).

#### **Body length and ear length**

Body length was significantly associated with *rs41646754* on BTA27, and with *rs110694334*, which is located within *CNTNAP5* on BTA2. Neither the two variants, nor the identified genes, have been previously associated with BL. Nevertheless, *CNTNAP5* was detected in a chromosomal segment significantly associated with hip cross height in Brahman cattle [108]. Furthermore, *CNTNAP5* is a potential candidate gene for conformation traits in Sudanese goats and growth traits and pigs [109–111]. In addition, *CNTNAP5* was identified in a selective region for adaptation in cattle as well as in sheep [112, 113]. According to DAVID, *CNTNAP5* is an



integral component of membranes and is related to the epidermal growth factor-like protein domain.

Among the seven suggestive SNP associated with BL, *rs43616983* on *BTA10* was associated with direct perinatal mortality in Holstein–Friesian dairy cattle [16]. Moreover, *rs43616983* is positioned within the *MEGF11* gene, which regulated daily gain and immune response to mastitis in cattle [85, 114]. Another study observed different expressions of *MEGF11* in the *musculus longissimus dorsi* of two different cattle breeds in response to low energy diets [85]. This finding is consistent with results from GWAS in pigs suggesting *MEGF11* as a potential candidate gene for feed efficiency [115]. Furthermore, *MEGF11* was significantly associated with height in Buffalo [116]. Two other SNP suggestively associated with BL, *rs109186122* and *rs43616983*, are located in the genes *GYPC* and *SLC16A4*, respectively. According to its annotation in DAVID, *GYPC* is involved in oligosaccharide binding and is reported to be associated with intramuscular fat deposition and reproduction performances in cattle [117–119]. The *SLC16A4* gene had effects on postweaning weight gain, feed efficiency and resistance to the bovine viral diarrhoea virus in cattle [120–122]. These observations are in line with the involvement of *SLC16A4* in transmembrane transport functions and glucose import processes, as well as with the impact of the SLC16 gene family on health regulation [123]. The *SLC16A4* gene is located in a genomic segment being under divergent selection in South African cattle breeds [13]. In addition, solute carrier family genes were associated with cattle body weight or conformation traits [124, 125].

None of the four SNP suggestively associated with EL are linked to EL in the literature, because of the trait relevance only for tropical production systems. Two potential candidate genes for EL are related to different traits in cattle and other mammal species. Firstly, *MEGF11*, harboring the SNP *rs110608572* on *BTA10*, overlap with our findings for BL. Secondly, *ADAMTS12*, harboring the SNP *rs109985119* on *BTA20*, was associated with body weight and supernumerary teat in cattle [126–128]. *ADAMTS12* was involved in inflammatory responses and in the regulation of the hepatocyte growth factor (HGF) receptor signaling pathway. Indeed, HGFs are known to play an important role in the stimulation of epithelial cell proliferation, motility, morphogenesis and angiogenesis [129]. The *ADAMTS12* gene was identified in pigs in a selection signature associated with genetic adaptation to high altitude [130]. These observations confirm a potential effect of *ADAMTS12* on EL variability, considered as an adaptive trait for heat tolerance in the indigenous breeds in Benin [29, 131].

### Genomic regions associated with morphometric and adaptation traits

The comparison with the literature gives convincing evidence for the validity of our results based on gene functions and associations as detected in other traits and species. The rather small number of detected SNP for the moderately to highly heritable polygenic traits is in line with our small sample size and its diversity (multi-breed), and is most likely due to a lack of power to detect variants with smaller effects. However, with regard to the functional annotations of the identified genomic regions, our study confirms the potential of multi-breed GWAS in detecting fewer variants, but more precise functional loci or causative mutation across breeds [75, 132]. In this context, the SNP associated with morphometric traits in the present study may represent novel common variants for linear body traits in African breeds, suggesting further investigations. Furthermore, the validity of the multi-breed approach may constitute to collaborative research towards better characterizations and genetic evaluations, including the animal genetic resources in Africa [2, 3, 71, 133].

Our findings reveal a close connection between genomic regions associated with morphometric traits and adaptive traits. For instance, two of the identified variants detected for morphometric traits in this study are directly linked to immune response, while several potential candidate genes have functional annotations for, e.g., signal transduction, metabolism, and immune response adaptation [134]. The comparison with previous studies confirms associations of the identified genes with selection signatures as well as immunity or resistance to diseases, feed efficiency and adaptation to harsh environments (heat stress, high altitude), addressing the main components of adaptation [4]. On the one hand, our observations are consistent with the breeding history of the indigenous breeds in Benin, which is characterized by natural and non-directional selection based on individual farmer preferences. In many African breeds, adaptive traits (disease resistance, feeding ease) and reproductive performances reflect the major breeding preferences of farmers, and they select animals according to their morphometric or conformation appearance rather than on actual recorded performance [50, 135, 136]. Hence, this might be an explanation for the observed association between morphometric and adaptive traits in our study [50, 137].

On the other hand, genomic regions associated with adaptive traits (immune response or feed efficiency) have been similarly detected in GWAS for linear body traits in more developed breeds such as Holstein and Angus [36, 62, 92, 138]. These overlaps as well as the effect of selection response for disease resistance or feed efficiency on conformation or carcass traits, and vice-versa, have been extensively discussed [138, 139]. In our view,

two hypotheses emerge from such overlap. First, the identified loci or genomic regions may simultaneously control morphometric traits and adaptive traits, indicating the pleiotropic role of many loci associated with body conformation traits such as height or body size [42]. Secondly, a high genetic correlation between morphometric and adaptive traits is due to the long history of natural selection in the studied breeds on both trait categories.

## Conclusion

Heritabilities for as well as phenotypic and genetic correlations among morphometric traits based on dense SNP marker data and a multi-breed approach were moderate to moderately high. Twenty-five SNP and fifteen genes potentially associated with the morphometric traits were detected. Comparisons with previous studies and the functional annotation of the genes revealed a clear association of loci identified in this study with conformation, growth and carcass traits in cattle or in other species. Moreover, the majority of the detected genes are associated with immune response and feed efficiency, or involved in related biological processes, suggesting a strong correlation between morphological and adaptive traits. This is in line with the evolutionary development and breeding history in these indigenous breeds mainly shaped by natural selection. Our findings suggest that accurate phenotyping (measurements) for morphometric traits combined with SNP marker data can be used for genetic evaluations, considering mixed-breed cattle populations.

## Methods

### Phenotypes and animal resources

Following the FAO guidelines [29], we recorded six morphometric traits (Table 3) on 449 animals from the four main indigenous cattle breeds in Benin (Borgou 181, Pabli 58, Lagune 150, Somba 60). The morphometric traits were chosen for their importance in cattle characterization and their association with production or adaptive traits. Sacrum height, heart girth, hip width and body length are commonly used to evaluate cattle body size and growth, and they are highly correlated with

body weight or milk yield in different cattle breeds [40, 44, 47, 48, 66, 140]. In addition, ear length and height at withers are related to adaptive traits, whereas hip width is associated with animal longevity [29, 30].

The animals were selected from a larger dataset of 462 animals as described in a previous study [57]. Thirteen animals from the larger dataset were excluded from the current study due to inconclusive genetic adherence and impact from crossbreeding [57]. The animals were sampled in small cattle herds kept under extensive management, according to their distribution across AEZ in Benin (see Additional file 2, Table S1 for animal characteristics, location and geographic coordinates of the herds). Somba and Pabli cattle were sampled in the AEZs Ouest Atacora (OA) and Cotonnière Nord (CNB). Borgou cattle were sampled in three different AEZs Cotonnière Centre (CCB), Vivrière Sud Borgou (VSB) and Cotonnière Nord (CNB), and the Lagune cattle in three other AEZs Pêcheries (P), Dépression (D) and Terre de Barre (TB). One-way analysis of variance (ANOVA) tests for the six morphometric traits revealed significant differences ( $p \leq 0.001$ ) across breeds with higher values for hybrid breeds (Borgou and Pabli). For instance for height at withers, Borgou and Pabli cattle measured  $116.5 \pm 5.67$  cm and  $111.4 \pm 7.58$ , but Somba and Lagune are smaller with  $96.37 \pm 4.97$  cm and  $92.59 \pm 7.18$ , respectively. The full description of the six morphometric traits for the four breeds are presented in Additional file 3, Figure S2.

### Genotypes, quality control and imputation

The 449 selected animals with phenotypes were genotyped using the Illumina BovineSNP50 BeadChip, and 51,278 SNP were available before quality control. Sample collection, DNA extraction and genotyping procedures are described in detail in Scheper et al. [57].

Genotyping quality control was performed using the PLINK software [26] to retain SNP with a minor allele frequency larger than 5% and a genotyping call rate of 90%, and which are in Hardy-Weinberg equilibrium ( $p \geq 10^{-6}$ ). For individual animals, a genotype call rate larger than 95% was required. After quality control, 14,518

**Table 3** Description of the six morphometric traits recorded on four indigenous cattle breeds from Benin

Morphometric traits	Description	Measuring device
Height at withers	Vertical distance from the bottom of the front foot to the highest point of the shoulder between the withers	Measuring stick
Sacrum height	Distance from the top of the bone at the base of the tail to the ground	Measuring stick
Heart girth	Circumference of the body immediately behind the shoulder blades in a vertical plane, perpendicular to the long axis of the body	Measuring tape
Hip width	Distance between the rearmost posterior points of pin bones	Wooden caliper
Body length	Horizontal distance from the point of the shoulder to the pin bone	Measuring tape
Ear length	Length on the back side of the ear from its root on the poll to the tip	Measuring tape

SNP and one animal were discarded, implying a genotype dataset with 36,760 SNP variants from 448 cattle.

Sporadic missing variants were imputed with BEAGLE [141] after remapping the SNP positions to the current reference assembly ARS1.2 and removing all markers with unknown position on ARS1.2. Genotype imputation generated a dataset of 36,720 SNP for the ongoing genomic analyses.

### Adjustment for fixed and environmental effects and genetic structure

One-way ANOVA in R was firstly applied on each morphometric trait to test the explanatory variables for significance. Given that breeds are nested within AEZ, a new variable (AEZ\_Breed), which combines the two variables AEZ and breed, was created. The factors age, sex, and AEZ\_Breed were simultaneously included in a multi-factor linear model to test their effects on the respective morphometric trait via the Type III sums of squares from ANOVA, using the Car package in R [142]. AEZ\_Breed showed significant effects on the six morphometric traits ( $p \leq 0.001$ ), while sex had significant effect only on HG, HW and BL, and age on HG, HW and HAW (Additional file 4, Table S2). In consequence, AEZ\_Breed, sex and age were considered as fixed effects in the GWAS models for all morphometric traits.

The genetic structure in the dataset was evaluated applying a DAPC, using the R package ADEGENET [143, 144]. LDF were used as covariates in addition to AEZ\_Breed, sex and age to correct for population stratification in GWAS. DAPC was chosen instead of classical principal component analysis (PCA) as it better characterizes the genetic structure of the population [145, 146] (see Additional file 4, Table S3, for significant SNP when the first four PCs were included as covariates in the GWAS model). The genotype dataset used for DAPC consisted of 25,065 SNP after pruning of the imputed dataset based on linkage disequilibrium (LD) between markers. The “--indep-pairwise” command in PLINK and defining 0.2 for the  $r^2$  threshold, was considered in this regard [27].

### Estimation of heritability, phenotypic and genetic correlation

A genomic relationship matrix between the animals was firstly generated with the -grm method in GCTA, and afterwards considered for the estimation of the genetic parameters [28, 74]. The restricted maximum likelihood (REML) method was applied for genetic parameter estimation. The respective genetic-statistical model was defined as follows:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{e} \tag{1}$$

where  $\mathbf{y}$  was a vector of morphometric traits;  $\mathbf{b}$  was a

vector of fixed effects including age, sex, and AEZ\_Breed;  $\mathbf{u}$  was a vector of polygenic effects with a variance-covariance structure of  $\mathbf{u} \sim N(0, \mathbf{G}\sigma_u^2)$ ,  $\mathbf{G}$  was the genomic relationship matrix between individuals [74],  $\sigma_u^2$  was the polygenic variance;  $\mathbf{e}$  was a vector of random residual effects with  $\mathbf{e} \sim N(0, \mathbf{I}\sigma_e^2)$ ,  $\mathbf{I}$  was an identity matrix of dimension  $n \times n$  (with  $n$ , the sample size = 449); and  $\mathbf{X}$  and  $\mathbf{Z}$  were incidence matrices for  $\mathbf{b}$  and  $\mathbf{u}$ , respectively.

Considering two conformation traits  $x$  and  $y$ , the genetic correlation ( $r_g$ ) between  $x$  and  $y$  was estimated using the “--reml-bivar  $x y$ ” option in bivariate genomic REML analyses. The phenotypic correlation ( $r_p$ ) between  $x$  and  $y$  was calculated from the bivariate genomic REML outputs using the following formula:

$$r_{p,xy} = \frac{\sigma_{u_{xy}} + \sigma_{e_{xy}}}{\sqrt{(\sigma_{u_x}^2 + \sigma_{e_x}^2) \times (\sigma_{u_y}^2 + \sigma_{e_y}^2)}} \tag{2}$$

where  $\sigma_{u_{xy}}$  and  $\sigma_{e_{xy}}$  were the genetic covariance and residual covariance between  $x$  and  $y$ , respectively; and  $\sigma_u^2$  and  $\sigma_e^2$  were the genetic variance and residual variance of  $x$  and  $y$ , respectively. The standard errors of the phenotypic correlation were calculated based on the “delta-method” function from the R-package “msm” [147].

### Multi-breed GWAS

GWAS were performed for the six morphometric traits using PLINK [27]. A linear regression using an additive genetic model was applied, and defined as follows:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Wg} + \mathbf{e} \tag{3}$$

where  $\mathbf{y}$  was a vector of morphometric traits;  $\mathbf{b}$  was a vector of fixed effects including Age, Sex, AEZ\_Breed, and linear discriminant functions;  $\mathbf{g}$  was a vector for the SNP effects;  $\mathbf{e}$  was a vector of random residual effects with  $\mathbf{e} \sim N(0, \mathbf{I}\sigma_e^2)$ ; and  $\mathbf{X}$ ,  $\mathbf{W}$  were incidence matrices for  $\mathbf{b}$  and  $\mathbf{g}$ , respectively.

For a verification of results from PLINK, we additionally performed GWAS applying the following model (eq. 4) in the GCTA software [28]. However, considering the complexity of the mixed linear model in GCTA [28], and the small size of our dataset, only SNP detected by PLINK are prioritized and described. In matrix notation, the mixed model is:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Wg} + \mathbf{Zu} + \mathbf{e} \tag{4}$$

where  $\mathbf{y}$ ,  $\mathbf{g}$ ,  $\mathbf{e}$  and incidence matrices  $\mathbf{X}$  and  $\mathbf{W}$  were notations as defined in Eq. 3;  $\mathbf{b}$  was a vector of fixed effects including Age, Sex, AEZ\_Breed;  $\mathbf{Z}$  and  $\mathbf{u}$  were notations as defined in Eq. 1.

The assessment of the models for population stratification based on the genomic inflation factor (lambda- $\lambda$ ) and on the quantile–quantile (Q–Q) plot. Manhattan

and Q-Q-plots plots were generated by means of the ggplot2 package in R [148].

Significantly associated SNP were detected according to the Bonferroni corrected significance threshold ( $p = 1.55 \times 10^{-06}$ ), i.e., calculated as  $p = 0.05 / m$ , with  $m = 32,185$  (the effective number of SNP). In addition, chromosome-wide Bonferroni-corrected significance thresholds ( $p_c = 0.05 / m_c$ ) with  $m_c$  denoting the effective number of SNP for each chromosome (see Additional file 5, Table S4) were used to identify suggestively associated variants. The effective numbers of SNP for the whole genome ( $m$ ) and for each chromosome ( $m_c$ ) were determined from the genetic type I error calculator (GEC) [149].

### Candidate genes and functional annotations

To identify potential candidate genes associated with the morphometric traits, the rs-accession numbers of the significant and suggestive SNP were retrieved from the Ensembl genome database (version 96), using the BioMart R package [150, 151]. Genes were mapped to identified SNP, and only those located within a window frame of  $\pm 25$  kb around each SNP were considered. In cases with more than two identified genes within the defined window frame, preference was given to the gene in which the SNP was located, or to the gene in closest SNP distance. However, if the SNP were located between two genes, both genes were selected, and the remaining genes were discarded. The gene mapping method as well as the window frame of  $\pm 25$  kb were chosen to increase precision in selection of candidate genes.

Further, functional annotation was performed on the set of identified candidate genes per trait using the DAVID (see Additional file 6, Table S5 for the extensive outputs) [53]. In addition, pathways (KEGG and reactome pathways) and the biological process GO terms for candidate genes were retrieved manually, in order to infer potential gene functions.

### Supplementary Information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12864-020-07170-0>.

**Additional file 1: Figure S1.** Scatterplots for the first four linear discriminant functions (LDF) included as covariates in GWAS model for population stratification in the four Beninese indigenous cattle breeds. The coloring represents the original breed assignment of samples.

**Additional file 2: Table S1.** Characteristics and origins of sampled animals in four Beninese indigenous cattle breeds. Animal ID, sex, breed, year of sampling and origin: farm longitude and latitude, commune, agro-ecological zones (AEZ).

**Additional file 3: Figure S2.** The variations of six morphometric traits from all (ALL) and from four respective Beninese indigenous cattle. Height at withers (HAW), sacrum height (SH), heart girth (HG), hip width (HW), body length (BL) and ear length (EL).

**Additional file 4: Table S2.** Effects of AEZ\_B, sex and age in multi-factor linear models on six morphometric traits in four Beninese indigenous cattle breeds. Analysis of variance (ANOVA) table presenting the significance of fixed effects on height at withers (HAW), sacrum height (SH),

heart girth (HG), hip width (HW), body length (BL) and ear length (EL).

**Table S3.** Genome-wide and chromosome-wide significant SNP associated with six morphometric traits from GWAS model including the first four principal components (instead of the linear discriminant functions as considered for main results).

**Additional file 5: Table S4.** Chromosome-wide significance thresholds used in multi-breed GWAS for conformation traits in four Beninese indigenous cattle breeds.

**Additional file 6: Table S5.** Functional annotation of candidate genes for six morphometric traits in four Beninese indigenous cattle breeds. Results retrieved from the database for annotation, visualization and integrated discovery (DAVID) for candidate genes associated with height at withers (HAW), sacrum height (SH), heart girth (HG), hip width (HW), body length (BL) and ear length (EL).

### Abbreviations

AEZ: Agro-ecological zone; ANOVA: Analysis of variance; BL: Body length; DAPC: Discriminant analysis of principal components; DAVID: Database for annotation, visualization and integrated discovery; EL: Ear length; FAO: Food and agriculture organization of the United Nations; GCTA: Genome-wide complex trait analysis; GEC: Genetic Type I error calculator; GWAS: Genome-wide association study; HAW: Height at withers; HG: Heart girth; HW: Hip width; kb: Kilo base pairs = 1000 base pairs; KEGG: Kyoto Encyclopedia of Genes and Genomes; LD: Linkage disequilibrium; LDF: Linear discriminant functions; Mb: Mega bases pairs = 1000 kb = 1 million base pairs; PCA: Principal component analysis; REML: Restricted maximum likelihood;  $r_g$ : Genetic correlations;  $r_p$ : Phenotypic correlation; SH: Sacrum height; SNP: Single nucleotide polymorphism

### Acknowledgements

We are thankful to all farmers who give their consent to collect phenotype data and hair samples on their animals.

### Authors' contributions

SFV, CS, LHD and SK designed the study. SFV collected, analyzed the data and drafted the manuscript. CS, TY and KB supported in statistical analyses and manuscript preparation. SK and LHD reviewed and edited the manuscript. All authors read and approved the final manuscript.

### Funding

The present study was financially supported by the VolkswagenStiftung, Hannover, Germany.

Funding was used for animal genotyping and covered all travelling costs for trait recording. Open Access funding enabled and organized by Projekt DEAL.

### Availability of data and materials

All the data supporting the results of this article are presented within the article or in the additional files. The raw phenotypic and genotypic data are stored in the cloud of the University of Giessen (<https://jluibox.uni-giessen.de>) and can be provided by the corresponding author on reasonable request.

### Ethics approval and consent to participate

The data were collected in accordance to the FAO guidelines for the characterization of animal genetic resources. Animal management, animal husbandry and trait recording followed the criteria for the assessment of animal welfare as identified and defined in by the Welfare Quality Project (WQP) [29, 152]. In Benin, there is no specific legislation for animal welfare or animal body measurements and hair sampling [153]. Hair sampling and body measurements are non-invasive methods. Hence, no approval was necessary. Before data collection, all farmers and cattle breeders gave their verbal consent through the Communal Sector for Agricultural Development (SCDA). The SCDA is the decentralized governmental institution for the management of the agricultural sector in the surveyed municipalities. In consequence, no further institutional approval was required.

### Consent for publication

Not applicable.

**Competing interests**

The authors declare that they have no competing interests.

**Author details**

<sup>1</sup>Institute of Animal Breeding and Genetics, Justus-Liebig-University Gießen, Gießen, Germany. <sup>2</sup>School of Science and Technics of Animal Production, Faculty of Agricultural Sciences, University of Abomey-Calavi, Cotonou, Benin.

Received: 4 June 2020 Accepted: 20 October 2020

Published online: 11 November 2020

**References**

- Rewe TO, Herold P, Kahi AK, Zárate AV. Breeding indigenous cattle genetic resources for beef production in sub-Saharan Africa. *Outlook Agric.* 2009;38:317–26. <https://doi.org/10.5367/00000009790422205>.
- Ducrocq V, Laloe D, Swaminathan M, Rognon X, Tixier-Boichard M, Zerjal T. Genomics for ruminants in developing countries: from principles to practice. *Front Genet.* 2018;9:251. <https://doi.org/10.3389/fgene.2018.00251>.
- Marshall K, Gibson JP, Mwai O, Mwacharo JM, Haile A, Getachew T, et al. Livestock genomics for developing countries – African examples in practice. *Front Genet.* 2019. <https://doi.org/10.3389/fgene.2019.00297>.
- Mirkena T, Duguma G, Haile A, Tibbo M, Okeyo AM, Wurzinger M, Sölkner J. Genetics of adaptation in domestic farm animals: a review. *Livest Sci.* 2010;132:1–12. <https://doi.org/10.1016/j.livsci.2010.05.003>.
- Dossa LH, Vanvanhossou FUS. The indigenous Somba cattle of the hilly Atacora region in north-West Benin: threats and opportunities for its sustainable use. *Trop Anim Health Prod.* 2016;48:349–59. <https://doi.org/10.1007/s11250-015-0958-5>.
- Mwambene PL, Katule AM, Chenyambuga SW, Mwakilembe PAA. Fipa cattle in the southwestern highlands of Tanzania: socio-economic roles, traditional management practices and production constraints. *Anim Genet Resour/Resour Génét Anim/Recur Génét Anim.* 2012;51:1–14. <https://doi.org/10.1017/S2078633612000112>.
- Mapiye C, Chimonyo M, Muchenje V, Dzama K, Marufu MC, Raats JG. Potential for value-addition of Nguni cattle products in the communal areas of South Africa: a review. *Afr J Agric Res.* 2007;2:488–95.
- Yaro M, Munyard KA, Stear MJ, Groth DM. Combatting African animal Trypanosomiasis (AAT) in livestock: the potential role of trypanotolerance. *Vet Parasitol.* 2016;225:43–52. <https://doi.org/10.1016/j.vetpar.2016.05.003>.
- Windig JJ, Engelsma KA. Perspectives of genomics for genetic conservation of livestock. *Conserv Genet.* 2010;11:635–41. <https://doi.org/10.1007/s10592-009-0007-x>.
- Joost S, Bruford MW, Consortium TG-R. Editorial: advances in farm animal genomic resources. *Front Genet.* 2015;6:333. <https://doi.org/10.3389/fgene.2015.00333>.
- Toro MA, Villanueva B, Fernández J. Genomics applied to management strategies in conservation programmes. *Livest Sci.* 2014;166:48–53. <https://doi.org/10.1016/j.livsci.2014.04.020>.
- Bahbahani H, Afana A, Wragg D. Genomic signatures of adaptive introgression and environmental adaptation in the Sheko cattle of southwest Ethiopia. *PLoS One.* 2013:e0202479. <https://doi.org/10.1371/journal.pone.0202479> Public Library of Science.
- Makina SO, Muchadeyi FC, van Marle-Köster E, Taylor JF, Makgahlela ML, Maiwashe A. Genome-wide scan for selection signatures in six cattle breeds in South Africa. *Genet Sel Evol.* 2015;47:1–14. <https://doi.org/10.1186/s12711-015-0173-x>.
- Dayo G-K, Thevenon S, Berthier D, Moazami-Goudarzi K, Denis C, Cuny G, et al. Detection of selection signatures within candidate regions underlying trypanotolerance in outbred cattle populations. *Mol Ecol.* 2009;18:1801–13. <https://doi.org/10.1111/j.1365-294X.2009.04141.x>.
- Taye M, Kim J, Yoon SH, Lee W, Hanotte O, Dessie T, et al. Whole genome scan reveals the genetic signature of African Ankole cattle breed and potential for higher quality beef. *BMC Genet.* 2017;18:1–14. <https://doi.org/10.1186/s12863-016-0467-1>.
- Purfield DC, Bradley DG, Kearney JF, Berry DP. Genome-wide association study for calving traits in Holstein–Friesian dairy cattle. *Animal.* 2018;224–35. <https://doi.org/10.1017/S175173111300195X> Cambridge University Press.
- Abreu Silva BC, Eler JP, Santana ML, Mattos EC, Menezes IR, Ferraz JBS. Genetic association between mature weight and early growth and heifer pregnancy traits in Nelore cattle. *Livest Sci.* 2018;211:61–5. <https://doi.org/10.1016/j.livsci.2018.03.003>.
- Hay EH, Roberts A. Genome-wide association study for carcass traits in a composite beef cattle breed. *Livest Sci.* 2018;213:35–43. <https://doi.org/10.1016/j.livsci.2018.04.018>.
- Tenghe AMM, Bouwman AC, Berglund B, Strandberg E, de Koning DJ, Veerkamp RF. Genome-wide association study for endocrine fertility traits using single nucleotide polymorphism arrays and sequence variants in dairy cattle. *J Dairy Sci.* 2016;99:5470–85. <https://doi.org/10.3168/jds.2015-10533>.
- Benavides MV, Sonstegard TS, Kemp S, Mugambi JM, Gibson JP, Baker RL, et al. Identification of novel loci associated with gastrointestinal parasite resistance in a red Maasai x Dorper backcross population. *PLoS One.* 2015;10:e0122797. <https://doi.org/10.1371/journal.pone.0122797>.
- Dong X, Li J, Zhang Y, Deng X, Wu C. P3020 the potential relationship between comb color and egg production revealed by GWAS in blue-shelled chicken. *J Anim Sci.* 2016;94:61–2. <https://doi.org/10.2527/jas2016.94supplement461a>.
- Mapholi NO, Maiwashe A, Matika O, Riggio V, Bishop SC, MacNeil MD, et al. Genome-wide association study of tick resistance in south African Nguni cattle. *Ticks Tick Borne Dis.* 2016;7:487–97. <https://doi.org/10.1016/j.ttbdis.2016.02.005>.
- Martin AR, Teferra S, Möller M, Hoal EG, Daly MJ. The critical needs and challenges for genetic architecture studies in Africa. *Curr Opin Genet Dev.* 2018;53:113–20. <https://doi.org/10.1016/j.gde.2018.08.005>.
- Cantor RM, Lange K, Sinsheimer JS. Prioritizing GWAS results: a review of statistical methods and recommendations for their application. *Am J Hum Genet.* 2010;86:6–22. <https://doi.org/10.1016/j.ajhg.2009.11.017>.
- Marchini J, Donnelly P, Cardon LR. Genome-wide strategies for detecting multiple loci that influence complex diseases. *Nat Genet.* 2005;37:413–7. <https://doi.org/10.1038/ng1537>.
- Gondro C, van der Werf J, Hayes B. Genome-wide association studies and genomic prediction. Totowa: Humana Press; 2013.
- 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.
- Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet.* 2011;88:76–82. <https://doi.org/10.1016/j.ajhg.2010.11.011>.
- FAO. Phenotypic characterization of animal genetic resources. Rome: Food and Agriculture Organization of the United Nations; 2012.
- Vollema AR, van der Beek S, Harbers AGF, de Jong G. Genetic evaluation for longevity of Dutch dairy bulls. *J Dairy Sci.* 2000;83:2629–39. [https://doi.org/10.3168/jds.S0022-0302\(00\)75156-3](https://doi.org/10.3168/jds.S0022-0302(00)75156-3).
- Kominakis A, Hager-Theodorides AL, Zoidis E, Saridaki A, Antonakos G, Tsiamis G. Combined GWAS and ‘guilt by association’-based prioritization analysis identifies functional candidate genes for body size in sheep. *Genet Sel Evol.* 2012:1–16. <https://doi.org/10.1186/s12711-017-0316-3> BioMed Central.
- Dadpasand M, Miraei-Ashtiani SR, Moradi Shahrebabak M, Vaez TR. Impact of conformation traits on functional longevity of Holstein cattle of Iran assessed by a Weibull proportional hazards model. *Livest Sci.* 2008;118:204–11. <https://doi.org/10.1016/j.livsci.2008.01.024>.
- Abo-Ismael MK, Brito LF, Miller SP, Sargolzaei M, Grossi DA, Moore SS, et al. Genome-wide association studies and genomic prediction of breeding values for calving performance and body conformation traits in Holstein cattle. *Genet Sel Evol.* 2017;49:82. <https://doi.org/10.1186/s12711-017-0356-8>.
- ICAR. ICAR guidelines for conformation recording of dairy cattle, beef cattle and dairy goats; 2018.
- Alphonso C, Akpa G, Mukasa C, Rekwot P, Barje P. Genetic evaluation of linear udder and body conformation traits in Bunaji cows. *Anim Res Int.* 2011;8:1366–74. <https://doi.org/10.4314/ari.v8i11>.
- Cole JB, Wiggans GR, Ma L, Sonstegard TS, Lawlor TJ, Crooker BA, et al. Genome-wide association analysis of thirty one production, health, reproduction and body conformation traits in contemporary U.S. Holstein cows. *BMC Genomics.* 2011;12:408. <https://doi.org/10.1186/1471-2164-12-408>.
- Terakado APN, Costa RB, de Camargo GMF, Irano N, Bresolin T, Takada L, et al. Genome-wide association study for growth traits in Nelore cattle. *Animal.* 2018;12:1358–62. <https://doi.org/10.1017/S1751731117003068>.
- Bolormaa S, Hayes BJ, Savin K, Hawken R, Barendse W, Arthur PF, et al. Genome-wide association studies for feedlot and growth traits in cattle. *J Anim Sci.* 2011;99:1684–97. <https://doi.org/10.2527/jas.2010-3079>.
- Munim T, Oikawa T, Ibi T, Kunieda T. Genetic relationship of body measurement traits at early age with carcass traits in Japanese black cattle. *Anim Sci J.* 2013;84:206–12. <https://doi.org/10.1111/asj.12005>.

40. Zhang X, Chu Q, Guo G, Dong G, Li X, Zhang Q, et al. Genome-wide association studies identified multiple genetic loci for body size at four growth stages in Chinese Holstein cattle. *PLoS One*. 2017;12:e0175971. <https://doi.org/10.1371/journal.pone.0175971> .
41. An B, Xia J, Chang T, Wang X, Xu L, Zhang L, et al. Genome-wide association study reveals candidate genes associated with body measurement traits in Chinese Wagyu beef cattle. *Anim Genet*. 2019;50:386–90. <https://doi.org/10.1111/age.12805> .
42. Bouwman AC, Daetwyler HD, Chamberlain AJ, Ponce CH, Sargolzaei M, Schenkel FS, et al. Meta-analysis of genome-wide association studies for cattle stature identifies common genes that regulate body size in mammals. *Nat Genet*. 2018;50:362–7. <https://doi.org/10.1038/s41588-018-0056-5> .
43. Pryce JE, Hayes BJ, Bolormaa S, Goddard ME. Polymorphic regions affecting human height also control stature in cattle. *Genetics*. 2011;187:981–4. <https://doi.org/10.1534/genetics.110.123943> .
44. Setoguchi K, Watanabe T, Weikard R, Albrecht E, Kühn C, Kinoshita A, et al. The SNP c.1326TG in the non-SMC condensin I complex, subunit G (NCAPG) gene encoding a p.Ile442Met variant is associated with an increase in body frame size at puberty in cattle. *Anim Genet*. 2011;42:650–5. <https://doi.org/10.1111/j.1365-2052.2011.02196.x> .
45. Al-Mamun HA, Kwan P, Clark SA, Ferdosi MH, Tellam R, Gondro C. Genome-wide association study of body weight in Australian merino sheep reveals an orthologous region on OAR6 to human and bovine genomic regions affecting height and weight. *Genet Sel Evol*. 2015;47:66. <https://doi.org/10.1186/s12711-015-0142-4> .
46. Doyle JL, Berry DP, Veerkamp RF, Carthy TR, Walsh SW, Evans RD, Purfield DC. Genomic regions associated with skeletal type traits in beef and dairy cattle are common to regions associated with carcass traits, feed intake and calving difficulty. *Front Genet*. 2020;11:20. <https://doi.org/10.3389/fgene.2020.00020> .
47. Lukuyu MN, Gibson JP, Savage DB, Duncan AJ, Mujibi FDN, Okeyo AM. Use of body linear measurements to estimate liveweight of crossbred dairy cattle in smallholder farms in Kenya. *Springerplus*. 2016;5:63. <https://doi.org/10.1186/s40064-016-1698-3> .
48. Vanvanhossou SFU, Diogo RVC, Dossa LH. Estimation of live bodyweight from linear body measurements and body condition score in the west African Savannah shorthorn cattle in north-West Benin. *Cogent Food Agric*. 2018;4:135. <https://doi.org/10.1080/23311932.2018.1549767> .
49. Misganaw G, Wuletaw Z, Ayalew W. Relationships between conformation traits and milk off-take of indigenous cattle breeds in North-Western Ethiopia. *Anim Genet Resour*. 2013;53:27–32. <https://doi.org/10.1017/S2078633613000283> .
50. Kabi F, Masembe C, Negri N, Muwanika V. Patterns of indigenous female cattle morphometric traits variations in Uganda: evidence for farmers' selection to enhance agro-ecological fitness. *Anim Genet Resour/Resour Génét Anim/Recur Génét Anim*. 2015;56:79–90. <https://doi.org/10.1017/S2078633614000551> .
51. Visscher PM, Wray NR, Zhang Q, Sklar P, McCarthy MI, Brown AM, Yang J. 10 years of GWAS discovery: biology, function, and translation. *Am J Hum Genet*. 2017;101:5–22. <https://doi.org/10.1016/j.ajhg.2017.06.005> .
52. Hou L, Zhao H. A review of post-GWAS prioritization approaches. *Front Genet*. 2013;4:280. <https://doi.org/10.3389/fgene.2013.00280> .
53. Dennis G, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, Lempicki RA. DAVI D: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol*. 2003;1–11. <https://doi.org/10.1186/gb-2003-4-9-r60> BioMed Central.
54. Berthier D, Brenière SF, Bras-Gonçalves R, Lemesre J-L, Jamonneau V, Solano P, et al. Tolerance to Trypanosomatids: a threat, or a key for disease elimination? *Trends Parasitol*. 2016;32:157–68. <https://doi.org/10.1016/j.pt.2015.11.001> .
55. Flori L, Thevenon S, Dayo G-K, Senou M, Sylla S, Berthier D, et al. Adaptive admixture in the west African bovine hybrid zone: insight from the Borgou population. *Mol Ecol*. 2014;23:3241–57. <https://doi.org/10.1111/mec.12816> .
56. Houessou SO, Dossa LH, Diogo RVC, Houinato M, Buerkert A, Schlecht E. Change and continuity in traditional cattle farming systems of west African coast countries: a case study from Benin. *Agric Syst*. 2019;168:112–22. <https://doi.org/10.1016/j.agry.2018.11.003> .
57. Scheper C, Bohlouli M, Brügemann K, Weimann C, Vanvanhossou SFU, König S, Dossa LH. The role of agro-ecological factors and transboundary transhumance in shaping the genetic diversity in four indigenous cattle populations of Benin. *J Anim Breed Genet*. 2020. <https://doi.org/10.1111/jbg.12495> .
58. van Niekerk M, Naser FWC. Genetic parameters for growth traits in south African Limousin cattle. *SA J An Sci*. 2007. <https://doi.org/10.4314/sajas.v36i5.4075> .
59. Norris D, Banga C, Benyi K, Sithole BC. Estimation of genetic parameters and variance components for growth traits of Nguni cattle in Limpopo Province, South Africa. *Trop Anim Health Prod*. 2004;36:801–6. <https://doi.org/10.1023/b:trop.0000045966.59590.96> .
60. Kadarmideen HN, Wegmann S. Genetic parameters for body condition score and its relationship with type and production traits in Swiss Holsteins. *J Dairy Sci*. 2003;86:3685–93. [https://doi.org/10.3168/jds.S0022-0302\(03\)73974-5](https://doi.org/10.3168/jds.S0022-0302(03)73974-5) .
61. Pryce JE, Coffey MP, Brotherstone S. The genetic relationship between calving interval, body condition score and linear type and management traits in registered Holsteins. *J Dairy Sci*. 2000;83:2664–71. [https://doi.org/10.3168/jds.S0022-0302\(00\)75160-5](https://doi.org/10.3168/jds.S0022-0302(00)75160-5) .
62. Schrooten C, Bovenhuis H, Coppieters W, van Arendonk JAM. Whole genome scan to detect quantitative trait loci for conformation and functional traits in dairy cattle. *J Dairy Sci*. 2000;83:795–806. [https://doi.org/10.3168/jds.S0022-0302\(00\)74942-3](https://doi.org/10.3168/jds.S0022-0302(00)74942-3) .
63. Roveglia C, Niero G, Bobbo T, Penasa M, Finocchiaro R, Visentin G, et al. Genetic parameters for linear type traits including locomotion in Italian Jersey cattle breed. *Livest Sci*. 2019;229:131–6. <https://doi.org/10.1016/j.livsci.2019.09.023> .
64. de Haas Y, Janss LLG, Kadarmideen HN. Genetic and phenotypic parameters for conformation and yield traits in three Swiss dairy cattle breeds. *J Anim Breed Genet*. 2007;124:12–9. <https://doi.org/10.1111/j.1439-0388.2007.00630.x> .
65. Vargas CA, Elzo MA, Chase CC, Olson TA. Genetic parameters and relationships between hip height and weight in Brahman cattle. *J Anim Sci*. 2000;78:3045–52. <https://doi.org/10.2527/2000.78123045x> .
66. Kamprasert N, Duijvesteijn N, van der Werf JHJ. Estimation of genetic parameters for BW and body measurements in Brahman cattle. *Animal*. 2019;13:1576–82. <https://doi.org/10.1017/S1751731118003348> .
67. Bolormaa S, Pryce JE, Reverter A, Zhang Y, Barendse W, Kemper K, et al. A multi-trait, meta-analysis for detecting pleiotropic polymorphisms for stature, fatness and reproduction in beef cattle. *PLoS Genet*. 2014. <https://doi.org/10.1371/journal.pgen.1004198> .
68. Biscarini F, Biffani S, Canavesi F. Genetic analysis of type traits for the Italian Jersey breed. *Interbull Bull*. 2003;80–3.
69. Koenen EPC, Groen AF. Genetic evaluation of body weight of lactating Holstein heifers using body measurements and conformation traits. *J Dairy Sci*. 1998;81:1709–13. [https://doi.org/10.3168/jds.S0022-0302\(98\)75738-8](https://doi.org/10.3168/jds.S0022-0302(98)75738-8) .
70. Hagger C, Hofer A. Phenotypic and genetic relationships with height, heart girth and milk yield in the Swiss Braunvieh and Simmental breeds. *Livest Prod Sci*. 1991;28:265–71. [https://doi.org/10.1016/0301-6226\(91\)90147-1](https://doi.org/10.1016/0301-6226(91)90147-1) .
71. Ojango JMK, Mrode R, Rege JEO, Mujibi D, Strucken EM, Gibson J, Mwai O. Genetic evaluation of test-day milk yields from smallholder dairy production systems in Kenya using genomic relationships. *J Dairy Sci*. 2019;102:5266–78. <https://doi.org/10.3168/jds.2018-15807> .
72. Meyer K, Carrick MJ, Donnelly BJ. Genetic parameters for growth traits of Australian beef cattle from a multibreed selection experiment. *J Anim Sci*. 1993;71:2614–22. <https://doi.org/10.2527/1993.71102614x> .
73. Browning SR, Browning BL. Population structure can inflate SNP-based heritability estimates. *Am J Hum Genet*. 2011;89:191–3; author reply 193–5. <https://doi.org/10.1016/j.ajhg.2011.05.025> .
74. Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, et al. Common SNPs explain a large proportion of the heritability for human height. *Nat Genet*. 2010;42:565–9. <https://doi.org/10.1038/ng.608> .
75. Raven L-A, Cocks BG, Hayes BJ. Multibreed genome wide association can improve precision of mapping causative variants underlying milk production in dairy cattle. *BMC Genomics*. 2014;15:62. <https://doi.org/10.1186/1471-2164-15-62> .
76. Wientjes YCJ, Calus MPL, Goddard ME, Hayes BJ. Impact of QTL properties on the accuracy of multi-breed genomic prediction. *Genet Sel Evol*. 2015;47:42. <https://doi.org/10.1186/s12711-015-0124-6> .
77. Welderufael BG, Løvendahl P, de Koning D-J, Janss LLG, Fikse WF. Genome-Wide Association Study for Susceptibility to and Recoverability From Mastitis in Danish Holstein Cows. *Front Genet*. 2018;9:141. <https://doi.org/10.3389/fgene.2018.00141> .
78. Santana MHA, Ventura RV, Utsunomiya YT, Neves HHR, Alexandre PA, Junior GO, et al. A genomewide association mapping study using ultrasound-

- scanned information identifies potential genomic regions and candidate genes affecting carcass traits in Nellore cattle. *J Anim Breed Genet.* 2015: 420–7. <https://doi.org/10.1111/jbg.12167> John Wiley & Sons, Ltd (10.1111).
79. Jiang J, Cole JB, Da Y, VanRaden PM, Ma L, et al. bioRxiv. 2018:428227. <https://doi.org/10.1101/428227> Cold Spring Harbor Laboratory.
  80. Twomey AJ, Berry DP, Evans RD, Doherty ML, Graham DA, Purfield DC. Genome-wide association study of endo-parasite phenotypes using imputed whole-genome sequence data in dairy and beef cattle. *Genet Sel Evol.* 2019;51:15. <https://doi.org/10.1186/s12711-019-0457-7>.
  81. de Las Heras-Saldana S, Clark SA, Duijvesteijn N, Gondro C, van der Werf JHJ, Chen Y. Combining information from genome-wide association and multi-tissue gene expression studies to elucidate factors underlying genetic variation for residual feed intake in Australian Angus cattle. *BMC Genomics.* 2019;20:939. <https://doi.org/10.1186/s12864-019-6270-4>.
  82. Serão NVL, González-Peña D, Beever JE, Faulkner DB, Southey BR, Rodríguez-Zas SL. Single nucleotide polymorphisms and haplotypes associated with feed efficiency in beef cattle. *BMC Genet.* 2014;1–20. <https://doi.org/10.1186/1471-2156-14-94> BioMed Central.
  83. Daetwyler HD, Swan AA, van der Werf JHJ, Hayes BJ. Accuracy of pedigree and genomic predictions of carcass and novel meat quality traits in multi-breed sheep data assessed by cross-validation. *Genet Sel Evol: BioMed Central.* 2012;1–11. <https://doi.org/10.1186/1297-9686-44-33>.
  84. Sarakul M, Elzo MA, Koonawootrittriron S, Suwanasopee T, Jattawa D, Laodim T. Characterization of biological pathways associated with semen traits in the Thai multibreed dairy population. *Anim Reprod Sci.* 2018;197: 324–34. <https://doi.org/10.1016/j.anireprosci.2018.09.002>.
  85. Strillacci MG, Frigo E, Schiavini F, Samoré AB, Canavesi F, Vevey M, et al. Genome-wide association study for somatic cell score in Valdostana Red Pied cattle breed using pooled DNA. *BMC Genet.* 2014;1–10. <https://doi.org/10.1186/s12863-014-0106-7> BioMed Central.
  86. Yan Z, Wang Z, Zhang Q, Yue S, Yin B, Jiang Y, Shi K. Identification of whole-genome significant single nucleotide polymorphisms in candidate genes associated with body conformation traits in Chinese Holstein cattle. *Anim Genet.* 2020;51:141–6. <https://doi.org/10.1111/age.12865>.
  87. Lindholm-Perry AK, Butler AR, Kern RJ, Hill R, Kuehn LA, Wells JE, et al. Differential gene expression in the duodenum, jejunum and ileum among crossbred beef steers with divergent gain and feed intake phenotypes. *Anim Genet.* 2014;408–27. <https://doi.org/10.1111/age.12440> John Wiley & Sons, Ltd (10.1111).
  88. Kubik R. Genomic investigation of beta agonist supplementation and heat stress in livestock species. Public Access Theses and Dissertations from the College of Education and Human Sciences. 2018.
  89. Ahlberg CM, Schiermiester LN, Howard TJ, Calkins CR, Spangler ML. Genome wide association study of cholesterol and poly- and monounsaturated fatty acids, protein, and mineral content of beef from crossbred cattle. *Meat Sci.* 2014;98:804–14. <https://doi.org/10.1016/j.meatsci.2014.07.030>.
  90. Schroeder C, Riess A, Bonin M, Bauer P, Riess O, Döbler-Neumann M, et al. PIK3R1 mutations in SHORT syndrome. *Clin Genet.* 2014;86:292–4. <https://doi.org/10.1111/cge.12263>.
  91. Dymont DA, Smith AC, Alcantara D, Schwartzentruber JA, Basel-Vanagaite L, Curry CJ, et al. Mutations in PIK3R1 cause SHORT syndrome. *Am J Hum Genet.* 2013;93:158–66. <https://doi.org/10.1016/j.ajhg.2013.06.005>.
  92. Ashwell MS, Heyen DW, Weller JL, Ron M, Sonstegard TS, van Tassel CP, Lewin HA. Detection of quantitative trait loci influencing conformation traits and calving ease in Holstein-Friesian cattle. *J Dairy Sci.* 2005;88:4111–9. [https://doi.org/10.3168/jds.S0022-0302\(05\)73095-2](https://doi.org/10.3168/jds.S0022-0302(05)73095-2).
  93. Pošćić N, Montanari T, D'Andrea M, Licastro D, Pilla F, Ajmone-Marsan P, et al. Breed and adaptive response modulate bovine peripheral blood cells' transcriptome. *J Anim Sci Biotechnol.* 2017;1–16. <https://doi.org/10.1186/s40104-017-0143-y> BioMed Central.
  94. Fitzsimons C, Kenny DA, Waters SM, Earley B, McGee M. Effects of phenotypic residual feed intake on response to a glucose tolerance test and gene expression in the insulin signaling pathway in longissimus dorsi in beef cattle. *J Anim Sci.* 2011. p. 4616–4631. doi:<https://doi.org/10.2527/jas.2014-7699>. Namia.
  95. Hawken RJ, Zhang YD, Fortes MRS, Collis E, Barris WC, Corbet NJ, et al. Genome-wide association studies of female reproduction in tropically adapted beef cattle. *J Anim Sci.* 2012;90:1398–410. <https://doi.org/10.2527/jas.2011-4410>.
  96. Reber L, Da Silva CA, Frossard N. Stem cell factor and its receptor c-kit as targets for inflammatory diseases. *Eur J Pharmacol.* 2006;533:327–40. <https://doi.org/10.1016/j.ejphar.2005.12.067>.
  97. Xerxa E. Gene expression analysis of whole blood from preclinical and clinical cattle infected with atypical bovine spongiform encephalopathy. PhD thesis, International School for Advanced Studies. Italy: SISSA; 2016. <http://hdl.handle.net/20.500.11767/3586>.
  98. Trovato A, Panelli S, Strozzi F, Cambulli C, Barbieri I, Martinelli N, et al. Expression of genes involved in the T cell signalling pathway in circulating immune cells of cattle 24 months following oral challenge with bovine Amyloidotic spongiform encephalopathy (BASE). *Vet Res.* 2015;11:1–9. <https://doi.org/10.1186/s12917-015-0412-y>.
  99. Cesar ASM, Regitano LCA, Koltes JE, Fritz-Waters ER, Gasparin G, Mourão GB, et al. RNA sequencing analysis identifies retinoic acid pathway genes differentially expressed in animals with extreme intramuscular fat GEBVs in Nellore steers: Asas; 2014.
  100. Somavilla AL, Sonstegard TS, Higa RH, Rosa AN, Siqueira F, Silva LOC, et al. A genome-wide scan for selection signatures in Nellore cattle. *Anim Genet.* 2014: 771–81. <https://doi.org/10.1111/age.12210> John Wiley & Sons, Ltd (10.1111).
  101. Campos BM, do Carmo AS, da Silva, Ribeiro TB, Verardo LL, de Simoni Gouveia JJ, Mendes Malhado CH, et al. Identification of artificial selection signatures in Caracu breed lines selected for milk production and meat production. *Livest Sci.* 2017;206:82–7. <https://doi.org/10.1016/j.livsci.2017.10.014>.
  102. Li W, McIntyre TM. Platelet-activating factor receptor affects food intake and body weight. *Genes Dis.* 2015;2:255–60. <https://doi.org/10.1016/j.gendis.2015.06.002>.
  103. Suter A. Comparison of genome wide association studies for milk production traits in Valle del Belice dairy sheep; 2018.
  104. Fortes MRS, Reverter A, Nagaraj SH, Zhang Y, Jonsson NN, Barris W, et al. A single nucleotide polymorphism-derived regulatory gene network underlying puberty in 2 tropical breeds of beef cattle. *J Anim Sci.* 2011: 1669–83. <https://doi.org/10.2527/jas.2010-3681> Narnia.
  105. Bedford MT, Clarke SG. Protein arginine methylation in mammals: who, what, and why. *Mol Cell.* 2009;33:1–13. <https://doi.org/10.1016/j.molcel.2008.12.013>.
  106. Velasquez AV. Evaluation of internal and external markers for estimating dry matter intake and digestibility in cattle 2017: Universidade de São Paulo; Biblioteca Digital de Teses e Dissertações da Universidade de São Paulo. <https://doi.org/10.11606/T.10.2017.tde-17052017-152938>.
  107. Lee Y-S, Shin D. Genome-wide association studies associated with Backfat thickness in Landrace and Yorkshire pigs. *Genomics Inform.* 2018;16:59–64. <https://doi.org/10.5808/GI.2018.16.3.59>.
  108. Chen Q, Huang B, Zhan J, Wang J, Qu K, Zhang F, et al. Whole-genome analyses identify loci and selective signals associated with body size in cattle. *J Anim Sci.* 2020. <https://doi.org/10.1093/jas/skaa068>.
  109. Puig-Oliveras A, Ballester M, Corominas J, Revilla M, Estellé J, Fernández AI, et al. A Co-Association Network Analysis of the Genetic Determination of Pig Conformation, Growth and Fatness. *PLoS One.* 2013:e114862. <https://doi.org/10.1371/journal.pone.0114862> Public Library of Science.
  110. Rohrer GA, Nonneman DJ, Wiedmann RT, Schneider JF. A study of vertebra number in pigs confirms the association of vertnin and reveals additional QTL. *BMC Genet.* 2015;16:129. <https://doi.org/10.1186/s12863-015-0286-9>.
  111. Rahmatalla SA, Arends D, Reissmann M, Wimmers K, Reyer H, Brockmann GA. Genome-wide association study of body morphological traits in Sudanese goats. *Anim Genet.* 2014;478–82. <https://doi.org/10.1111/age.12686> John Wiley & Sons, Ltd (10.1111).
  112. Terefe MT. Identification of adaptive signatures in the cattle genome. PhD thesis, Seoul National University; 2018.
  113. de Simoni Gouveia JJ, Paiva SR, McManus CM, Caetano AR, Kijas JW, Facó O, et al. Genome-wide search for signatures of selection in three major Brazilian locally adapted sheep breeds. *Livest Sci.* 2017;197:36–45. <https://doi.org/10.1016/j.livsci.2017.01.006>.
  114. Olivieri BF, Mercadante MEZ, Cyrillo JND, Branco RH, Bonilha SFM, de Albuquerque LG, et al. Genomic Regions Associated with Feed Efficiency Indicator Traits in an Experimental Nellore Cattle Population. *PLoS One.* 2016;11:e0164390. <https://doi.org/10.1371/journal.pone.0164390>.
  115. Ding R, Yang M, Wang X, Quan J, Zhuang Z, Zhou S, et al. Genetic architecture of feeding behavior and feed efficiency in a Duroc pig population. *Front Genet.* 2018;9:220. <https://doi.org/10.3389/fgene.2018.00220>.
  116. Gonzalez Guzman JL, Lázaro SF, do Nascimento AV, de Abreu Santos DJ, Cardoso DF, Becker Scalez DC, et al. Genome-wide association study applied to type traits related to milk yield in water buffaloes (*Bubalus bubalis*). *J Dairy Sci.* 2020;103:1642–50. <https://doi.org/10.3168/jds.2019-16499>.

117. Fernández JC, Pérez JE, Herrera N, Martínez R, Bejarano D, Rocha JF. Genomic association study for age at first calving and calving interval in Romosinuano and Costeño con Cuernos cattle. *Genet Mol Res*. 2019;18:1–13.
118. Hansen PJ, Dobbs KB, Denicol AC. Programming of the preimplantation embryo by the embryokine colony stimulating factor 2. *Anim Reprod Sci*. 2014;149:59–66. <https://doi.org/10.1016/j.anireprosci.2014.05.017>.
119. Cesar ASM. Identification of genes associated with intramuscular fat deposition and composition in Nellore breed 2014: Universidade de São Paulo; Biblioteca Digital de Teses e Dissertações da Universidade de São Paulo. <https://doi.org/10.11606/T.11.2014.tde-12082014-103102>.
120. Campos GS, Sollero BP, Reimann FA, Junqueira VS, Cardoso LL, Yokoo MJ, et al. Tag-SNP selection using Bayesian genomewide association study for growth traits in Hereford and Braford cattle. *J Anim Breed Genet*. 2019. <https://doi.org/10.1111/jbg.12458>.
121. Graham C, Gatherer I, Haslam I, Glanville M, Simmons NL. Expression and localization of monocarboxylate transporters and sodium/proton exchangers in bovine rumen epithelium. *Am J Physiol Regul Integr Comp Physiol*. 2007;292:R997–1007. <https://doi.org/10.1152/ajpregu.00343.2006>.
122. Nilson SM. Comparative analyses of Transcriptome data from beef cattle persistently infected with bovine viral diarrhoea virus. *Theses Diss Anim Sci*. 2016;127:79. <http://digitalcommons.unl.edu/animalscidiss/127>.
123. Halestrap AP. The SLC16 gene family - structure, role and regulation in health and disease. *Mol Asp Med*. 2013;34:337–49. <https://doi.org/10.1016/j.mam.2012.05.003>.
124. Wu X, Fang M, Liu L, Wang S, Liu J, Ding X, et al. Genome wide association studies for body conformation traits in the Chinese Holstein cattle population. *BMC Genomics*. 2013;14:897. <https://doi.org/10.1186/1471-2164-14-897>.
125. Sugimoto M, Watanabe T, Sugimoto Y. The molecular effects of a polymorphism in the 5'UTR of solute carrier family 44, member 5 that is associated with birth weight in Holsteins. *PLoS One*. 2012;7:e41267. <https://doi.org/10.1371/journal.pone.0041267>.
126. Edea Z, Jeoung YH, Shin S-S, Ku J, Seo S, Kim I-H, et al. Genome-wide association study of carcass weight in commercial Hanwoo cattle. *Asian Australas J Anim Sci*. 2018;31:327–34. <https://doi.org/10.5713/ajas.17.0276>.
127. Li Y, Gao Y, Kim Y-S, Iqbal A, Kim J-J. A whole genome association study to detect additive and dominant single nucleotide polymorphisms for growth and carcass traits in Korean native cattle, Hanwoo. *Asian Australas J Anim Sci*. 2017;30:8–19. <https://doi.org/10.5713/ajas.16.0170>.
128. Joerg H, Meili C, Ruprecht O, Bangerter E, Burren A, Bigler A. A genome-wide association study reveals a QTL influencing caudal supernumerary teats in Holstein cattle. *Anim Genet*. 2014;871–3. <https://doi.org/10.1111/age.12215> John Wiley & Sons, Ltd (10.1111).
129. Nakamura T, Mizuno S. The discovery of hepatocyte growth factor (HGF) and its significance for cell biology, life sciences and clinical medicine. *Proc Jpn Acad Ser B Phys Biol Sci*. 2010;86:588–610. <https://doi.org/10.2183/pjab.86.588>.
130. Ai H, Huang L, Ren J. Genetic diversity, linkage disequilibrium and selection signatures in Chinese and Western pigs revealed by genome-wide SNP markers. *PLoS One*. 2013;8:e56001. <https://doi.org/10.1371/journal.pone.0056001>.
131. Barendse W. Climate adaptation of tropical cattle. *Annu Rev Anim Biosci*. 2017;5:133–50. <https://doi.org/10.1146/annurev-animal-022516-022921>.
132. Yin T, Jaeger M, Scheper C, Grodkowski G, Sakowski T, Klopčič M, et al. Multi-breed genome-wide association studies across countries for electronically recorded behavior traits in local dual-purpose cows. *PLoS One*. 2019;14:e0221973. <https://doi.org/10.1371/journal.pone.0221973>.
133. Mrode R, Tarekegn GM, Mwacharo JM, Djikeng A. Invited review: genomic selection for small ruminants in developed countries: how applicable for the rest of the world? *Animal*. 2018;12:1333–40. <https://doi.org/10.1017/S1751731117003688>.
134. Wang MD, Dzama K, Rees DJG, Muchadeyi FC. Tropically adapted cattle of Africa: perspectives on potential role of copy number variations. *Anim Genet*. 2016;47:154–64. <https://doi.org/10.1111/age.12391>.
135. Kugonza DR, Nabasiye M, Hanotte O, Mpairwe D, Okeyo AM. Pastoralists' indigenous selection criteria and other breeding practices of the long-horned Ankole cattle in Uganda. *Trop Anim Health Prod*. 2012;44:557–65. <https://doi.org/10.1007/s11250-011-9935-9>.
136. Kamuanga M, Tano K, Pokou K, Jabbar M, d'Iteren GM. Farmers' preferences of cattle breeds, their market values and prospects for improvement in West Africa: a summary review. *OAU/STRC*; 2001.
137. Ndumu DB, Baumung R, Wurzing M, Drucker AG, Okeyo AM, Semambo D, Sölkner J. Performance and fitness traits versus phenotypic appearance in the African Ankole longhorn cattle: a novel approach to identify selection criteria for indigenous breeds. *Livest Sci*. 2008;113:234–42. <https://doi.org/10.1016/j.livsci.2007.04.004>.
138. Crowley JJ, Evans RD, Mc Hugh N, Pabiou T, Kenny DA, McGee M, et al. Genetic associations between feed efficiency measured in a performance test station and performance of growing cattle in commercial beef herds. *J Anim Sci*. 2011;89:3382–93. <https://doi.org/10.2527/jas.2011-3836>.
139. Arthur PF, Herd RM, Wilkins JF, Archer JA. Maternal productivity of Angus cows divergently selected for post-weaning residual feed intake. *Aust J Exp Agric*. 2005;45:985. <https://doi.org/10.1071/EA05052>.
140. Bardakçoglu HE, Sekkin S, Toplu HDO. Relationship between some teat and body measurement of Holstein cows and sub-clinical mastitis and milk yield. *J Anim Vet Adv*. 2011;10:1735–7.
141. Browning BL, Zhou Y, Browning SR. A one-penny imputed genome from next-generation reference panels. *Am J Hum Genet*. 2018;103:338–48. <https://doi.org/10.1016/j.ajhg.2018.07.015>.
142. Fox J, Weisberg S. *An R companion to applied regression*. Thousand Oaks: Sage; 2019.
143. Jombart T. Adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*. 2008;24:1403–5. <https://doi.org/10.1093/bioinformatics/btn129>.
144. Jombart T, Ahmed I. Adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics*. 2011;27:3070–1. <https://doi.org/10.1093/bioinformatics/btr521>.
145. Jombart T, Devillard S, Balloux F. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet*. 2010;11:1–15. <https://doi.org/10.1186/1471-2156-11-94>.
146. Hayes B. Overview of statistical methods for genome-wide association studies (GWAS). *Methods Mol Biol*. 2013;1019:149–69. [https://doi.org/10.1007/978-1-62703-447-0\\_6](https://doi.org/10.1007/978-1-62703-447-0_6).
147. Jackson C. Multi-state modelling with R: the msm package. Cambridge. 2019. p. 1–57.
148. Wickham H. *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag; 2016.
149. Li M-X, Yeung JMY, Cherny SS, Sham PC. Evaluating the effective numbers of independent tests and significant *p*-value thresholds in commercial genotyping arrays and public imputation reference datasets. *Hum Genet*. 2012;131:747–56. <https://doi.org/10.1007/s00439-011-1118-2>.
150. Durinck S, Moreau Y, Kasprzyk A, Davis S, de Moor B, Brazma A, Huber W. BioMart and bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics*. 2005;21:3439–40. <https://doi.org/10.1093/bioinformatics/bti525>.
151. Durinck S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc*. 2009;1184–91. <https://doi.org/10.1038/nprot.2009.97> Nature Publishing Group.
152. Vapnek J, Chapman M. *Legislative and regulatory options for animal welfare*. Rome: Food and Agriculture Organization of the United Nations; 2011.
153. Gautier P, Escobar SL. Rapport de la mission de suivi de l'évaluation PVS des services vétérinaires du Bénin. Paris: OIE - World Organisation for Animal Health; 2013.

## Publisher's Note

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

**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](https://biomedcentral.com/submissions)

