Supplementary Information for: Marker-assisted selection for heat tolerance in Criollo cattle misses slick *PRLR* carriers when only one allele is genotyped

Luiz Dione Barbosa De-Melo[1,*]
[1] Molecular Genetics Laboratory,
Federal Institute of Education, Science and Technology
of Rio de Janeiro (IFRJ), Rio de Janeiro, Brazil

* Corresponding author: luiz.melo@ifrj.edu.br

Contents

Items appear below in the order in which the main text cites them.

- **Supplementary Note on SLICK5.** SLICK5 across the panel, and what is needed before it enters a marker panel
- **Supplementary Methods.** Validation of carrier calls against the aligned reads
- **Supplementary Note on Senepol ancestry.** Genome-wide ancestry of the Senepol, and the detection limit of the estimator
- **Supplementary Note on classifier robustness.** Robustness of the coding classifier, and three tests of the expectation behind its design
- **Supplementary Table S1.** Read support for each carrier call
- **Supplementary Table S2.** Every carrier haplotype, by method
- **Supplementary Table S3.** Panel rarefaction and batch separation
- **Supplementary Table S4.** Intronic assignment by distance to the causal site
- **Supplementary Table S5.** Frequency of SLICK1 in N'Dama compatible with the observed non-detection
- **Supplementary Table S6.** Inbreeding control for tract length
- **Supplementary Figure S1.** Distribution of SLICK5 across populations
- **Supplementary Figure S2.** Intronic assignment by distance to the causal site
- **Supplementary Figure S3.** Inbreeding control: tract length at the locus against random positions

Supplementary Note on SLICK5: distribution across the panel

The position chr20:39,099,228 carries SLICK5 (`c.1396A>T`, p.Lys466Ter) in 123 phased haplotypes across 109 animals of the present panel. The change was originally described by Sanger sequencing of slick-coated Criollo cattle [1]. That study and the later survey of 14 Criollo and Criollo-derived breeds [2] both treat it as a slick allele, and the survey reports it as the most frequent slick allele in the Blanco Orejinegro. Within the Criollo breeds the present distribution agrees with that description, including nine of ten Blanco Orejinegro animals. Beyond the Criollo breeds the same change also segregates in pure European taurines (19 labelled animals) and also occurs in one N'Dama and one Ogaden zebu (Supplementary Figure S1). It never shares a haplotype with the four segregating alleles. None of the 31 haplotypes that carry SLICK1, SLICK2, SLICK3 or SLICK6 carries it, and none of the 35 carrier haplotypes present before the read-based filter did.

Published frequencies at this position vary between samples, and the present panel adds others. [3] did not detect the change in 40 Blanco Orejinegro, where it appears here in nine of ten animals. The published frequencies for the Hartón del Valle also differ, 0.24 in [1] against 0.89 in [3]. The European taurines, N'Dama and Ogaden carrying the change here come from populations outside those samples.

No phenotype accompanies the newly incorporated animals, so segregation with the slick coat cannot be tested where the change is observed. The distribution does bear on its interpretation: a truncating change carried by lineages that diverged millennia ago, most of which are not known for the slick coat, is unlikely to act on its own as a population-specific causal allele. Whether the same substitution arose more than once at this position, or whether a single ancient allele is segregating at low penetrance, these data cannot settle. SLICK5 is therefore left unassigned, pending tests against the coat in the breeds where it was described, and is excluded from the haplotype analysis.

Supplementary Methods: Validation of carrier calls against the aligned reads

Phasing returns a complete genotype matrix. Sites that were missing before phasing are filled, and a filled genotype at a rare position can create a carrier that no read supports. With carrier counts as small as those reported here, a single spurious call would change a conclusion, so every call with a local alignment was returned to the reads.

Of the twenty-seven carrier calls, the eighteen from the newly incorporated runs have local alignments and were checked; the nine from Run 8, which I hold as called genotypes, could not be. All four SLICK1 carriers fall in the second group.

Depth and allele counts at each focal position were taken with `samtools mpileup` at a minimum mapping quality and base quality of 20, with base alignment quality (BAQ) recalibration disabled. BAQ lowers the quality of bases it judges possibly misaligned, and at these known positions it removes some true alternative bases: with BAQ enabled, the SLICK2 call in ERR12035149 drops from four supporting reads to two. Those two are the reads that span both the SLICK2 and the SLICK6 positions while carrying the SLICK6 reference base, so the allele and its phase remain directly observed. A call was retained when at least three reads carried the alternative allele and those reads made up at least a quarter of the depth at the site. Zygosity was not reassigned from the reads: every retained call keeps its phased genotype, including ERR12035155, homozygous for SLICK2 on three reads, and ERR12035132, heterozygous for SLICK6 after phasing although six of its seven reads carry the alternative allele. Four of those eighteen calls did not meet the criterion and were removed from the assignment: two SLICK2 calls, one SLICK3 call and one SLICK6 call (Supplementary Table S1). The removed animals remain masked in the panel frequencies, which is the conservative choice, since a carrier that may be real does not then contaminate the reference.

One of the removed calls has a further consequence. At ERR12035135 the phasing placed SLICK3 and SLICK6 in *cis*, but the reads covering both positions do not support the SLICK6 call at all, so the *cis* configuration rests on an allele that is not there. Where two focal positions fall within the span of a single read, phase can be read directly rather than inferred, and I did so by walking the CIGAR (Compact Idiosyncratic Gapped Alignment Report) string of every read covering both positions. This applies to the SLICK6 and SLICK2 pair, 54 bp apart, in ERR12035149: no read carries both alternative alleles, which places them in *trans* on read evidence rather than on statistical reconstruction.

Supplementary Note on Senepol ancestry: Genome-wide ancestry of the Senepol

The Senepol was formed on St. Croix in the early twentieth century from a presumed cross between the African taurine N'Dama and the European taurine Red Poll [4,5]. Whether West African ancestry survived that founding at a measurable level is a different question from the haplotype question of the main text, and it rests on a different resource: the BovineSNP50 genotypes of the global panel, from which 753 animals and 29,869 markers survived quality control after linkage-disequilibrium pruning.

Estimated with a supervised admixture model against a lineage-matched control,

the breed carries about six percentage points more apparent African taurine ancestry than the Red Poll founder line, and a similar excess over the wider British red lineage. An unsupervised replication with sNMF returns a comparable figure.

Titration with known doses places the detection limit of the estimator between roughly 10 and 25 per cent, above the estimate itself. The quantity reported is therefore the *excess over a matched control measured with the same estimator*, not the absolute fraction. Apparent African ancestry varies from 0.00 to 8.40 per cent across 26 pure European breeds under the same supervised estimator, a span close to that detection limit, and each estimate here is given with its paired control.

An earlier analysis of 147 Senepol, with BovineSNP50 genotypes and unsupervised ADMIXTURE at K = 3, attributed 0.6 per cent of the genome to African taurine ancestry, from 0 to 4.1 per cent among animals, and found no support for an African taurine origin of the breed [4]. That estimate comes from a different sample of the breed and a different estimator. That estimate and the excess reported here both fall below the 10 to 25 per cent detection limit titrated for the estimator used here, so the two results cannot be told apart, and both are compatible with a taurine ancestry that is almost entirely European. The Senepol is not part of the phased haplotype panel, and no slick carrier haplotype derives from it.

Supplementary Note on classifier robustness: Robustness of the coding classifier

The main text reports that a single pair of adjacent exon 2 sites carries most of the evidence in the coding classifier. Three further tests bear on how much weight the classifier can take, and two of them did not support the expectation that motivated the coding-site design.

Informativeness does not favour the coding core. The classifier was anchored on coding sites on the reasoning that exonic markers discriminate populations better. Measured with Rosenberg's I_{n} statistic, mean informativeness is 0.023 at coding sites against 0.020 at introns, but the difference inverts in sign once sites are matched by minor allele frequency, to between -0.0022 and -0.0026, and no comparison approaches significance ($P = 0.28$ to 0.45 across alleles). The apparent advantage was an artefact of the frequency spectrum, not a property of the class.

Physical proximity does not favour it either. The coding sites sit a median of about 1.1 kb from the causal positions, not within a few hundred base pairs, so linkage does not single them out. Assignment by distance band (Supplementary Table S4, Supplementary Figure S2) shows the predicted gradient for SLICK1 alone. For the other three alleles no band that adjudicates returns a European

panel; the Iberian panel is modal only in the bands nearest to and farthest from the causal site, where posteriors sit near chance. The bands nearest the causal site hold 12 to 18 sites and return median posteriors of 0.24 to 0.34, close to the 0.14 expected by chance with seven panels; the band beyond 20 kb returns posteriors near 0.19. Neither adjudicates, and the same criterion is applied to every band.

Panel size and batch do not explain the contrast. Rarefying all panels to the smallest and repeating the assignment 500 times keeps the modal panel for every allele in both classes, although for SLICK2 at synonymous sites and SLICK1 at non-synonymous sites in only 54 and 57 per cent of replicates (Supplementary Table S3). The SLICK2 carriers are the only ones drawn from both sequencing batches, which allows the one internal test available: the Run 8 carriers and the ERP150979 carriers return the same panel, so the assignment does not track library, depth or pipeline. The number of distinct coding configurations is the effective number of independent observations, and it is smaller than the haplotype count for every allele with more than one carrier haplotype, which Table 2 of the main text reports alongside the counts.

What these tests leave standing is narrow. The classifier is reproducible and is not an artefact of panel size or batch. It is nonetheless carried by two linked sites, and that is a property of the window, not of the sampling.

Supplementary Table S1

Supplementary Table S1. Read support for each carrier call. Depth and allele counts were taken with `samtools mpileup` at a minimum mapping and base quality of 20. A call was retained when at least three reads carried the alternative allele and those reads were at least a quarter of the depth; zygosity is that of the phased genotype.

Allele	Animal	Batch	Genotype after phasing	Reference reads	Alternative reads	Alternative fraction	Call retained
SLICK2	ERR12035115	ERP150979	`1|1`	0	7	1.00	yes
SLICK2	ERR12035117	ERP150979	`1|0`	9	5	0.36	yes
SLICK2	ERR12035119	ERP150979	`1|1`	0	15	1.00	yes
SLICK2	ERR12035125	ERP150979	`1|0`	4	2	0.33	**no**
SLICK2	ERR12035133	ERP150979	`1|0`	0	0	NA	**no**
SLICK2	ERR12035139	ERP150979	`1|0`	5	8	0.62	yes
SLICK2	ERR12035148	ERP150979	`1|0`	1	5	0.83	yes
SLICK2	ERR12035149	ERP150979	`1|0`	5	4	0.44	yes
SLICK2	ERR12035153	ERP150979	`1|1`	0	5	1.00	yes
SLICK2	ERR12035155	ERP150979	`1|1`	0	3	1.00	yes
SLICK2	ERR12035157	ERP150979	`0|1`	7	4	0.36	yes
SLICK2	ERR12035159	ERP150979	`1|1`	0	5	1.00	yes
SLICK3	ERR12035135	ERP150979	`0|1`	3	11	0.79	yes
SLICK3	ERR12035159	ERP150979	`1|0`	9	2	0.18	**no**
SLICK6	ERR12035121	ERP150979	`1|0`	7	8	0.53	yes
SLICK6	ERR12035132	ERP150979	`0|1`	1	6	0.86	yes
SLICK6	ERR12035135	ERP150979	`0|1`	16	2	0.11	**no**
SLICK6	ERR12035149	ERP150979	`0|1`	2	4	0.67	yes

Supplementary Table S2

Supplementary Table S2. Every carrier haplotype, by method. The exon 2 allele is the one at chr20:39,077,999 that carries most of the evidence in the coding classifier (Fig. 3b). Local ancestry is the call at the locus under each reference configuration; only the three-way configuration passes the validation of Table 3.

Allele	Animal	Chromosome	Exon 2 allele	Coding classifier (synonymous / non-synonymous)	Local ancestry: 3-way	4-way	5-way
SLICK1	SAMN09213442	0	ALT	N'Dama / N'Dama	European taurine	N'Dama	N'Dama
SLICK1	SAMN09213442	1	ALT	N'Dama / N'Dama	European taurine	N'Dama	N'Dama
SLICK1	SAMN09213443	1	ALT	N'Dama / N'Dama	European taurine	N'Dama	Asian taurine
SLICK1	SAMN08166129	0	REF	Iberian / Continental European	European taurine	N'Dama	Indicine
SLICK1	SAMN08166130	1	ALT	N'Dama / N'Dama	European taurine	N'Dama	Asian taurine
SLICK3	ERR12035135	1	REF	Iberian / Iberian	European taurine	Continental European	Continental European

Allele	Animal	Chromosome	Exon 2 allele	Coding classifier (synonymous / non-synonymous)	Local ancestry: 3-way	4-way	5-way
SLICK6	ERR12035121	0	REF	Iberian / Continental European	European taurine	Continental European	Continental European
SLICK6	ERR12035132	1	REF	Iberian / Iberian	European taurine	Iberian	Iberian
SLICK6	ERR12035149	1	REF	Iberian / Iberian	European taurine	Iberian	Continental European
SLICK2	SAMN08166127	1	REF	Continental European / Continental European	European taurine	Indicine	Continental European
SLICK2	SAMN08166128	0	REF	Continental European / Continental European	European taurine	Indicine	Continental European

Allele	Animal	Chromosome	Exon 2 allele	Coding classifier (synonymous / non-synonymous)	Local ancestry: 3-way	4-way	5-way
SLICK2	SAMN08166128	1	REF	Continental European / Continental European	European taurine	Indicine	Continental European
SLICK2	SAMN08166129	1	REF	Continental European / Continental European	European taurine	Continental European	Continental European
SLICK2	SAMN08166132	1	REF	Continental European / Continental European	European taurine	Indicine	Continental European
SLICK2	SAMN08166133	0	REF	Continental European / Continental European	European taurine	Indicine	Continental European

Allele	Animal	Chromosome	Exon 2 allele	Coding classifier (synonymous / non-synonymous)	Local ancestry: 3-way	4-way	5-way
SLICK2	SAMN08166133	1	REF	Continental European / Continental European	European taurine	Indicine	Continental European
SLICK2	ERR12035115	0	REF	Continental European / Continental European	European taurine	Indicine	Continental European
SLICK2	ERR12035115	1	REF	Continental European / Continental European	European taurine	Indicine	Indicine
SLICK2	ERR12035117	0	REF	Continental European / Continental European	European taurine	N'Dama	N'Dama

Allele	Animal	Chromosome	Exon 2 allele	Coding classifier (synonymous / non-synonymous)	Local ancestry: 3-way	4-way	5-way
SLICK2	ERR12035119	0	REF	Continental European / Continental European	European taurine	Indicine	Continental European
SLICK2	ERR12035119	1	REF	Continental European / Continental European	European taurine	Indicine	Indicine
SLICK2	ERR12035139	0	REF	Continental European / Continental European	European taurine	Indicine	Continental European
SLICK2	ERR12035148	0	REF	Iberian / Continental European	European taurine	Iberian	Iberian

Allele	Animal	Chromosome	Exon 2 allele	Coding classifier (synonymous / non-synonymous)	Local ancestry: 3-way	4-way	5-way
SLICK2	ERR12035149	0	REF	Continental European / Continental European	European taurine	Indicine	Continental European
SLICK2	ERR12035153	0	REF	Continental European / Continental European	European taurine	Indicine	Continental European
SLICK2	ERR12035153	1	REF	Continental European / Continental European	European taurine	Indicine	Continental European
SLICK2	ERR12035155	0	REF	Iberian / Iberian	European taurine	Indicine	Indicine
SLICK2	ERR12035155	1	REF	Continental European / Continental European	European taurine	Continental European	Continental European

Allele	Animal	Chromosome	Exon 2 allele	Coding classifier (synonymous / non-synonymous)	Local ancestry: 3-way	4-way	5-way
SLICK2	ERR12035157	1	REF	Continental European / Iberian	European taurine	Indicine	Indicine
SLICK2	ERR12035159	0	REF	Iberian / Iberian	European taurine	Continental European	Continental European
SLICK2	ERR12035159	1	REF	Continental European / Continental European	European taurine	Indicine	Continental European

Supplementary Table S3

Supplementary Table S3. Panel rarefaction and batch separation. Rarefaction draws every panel down to the size of the smallest and repeats the assignment 500 times; the three most frequent modal panels are shown. SLICK2 is the only allele with carriers in both sequencing batches, and so the only internal test of batch effect available.

Allele	Class	Panels rarefied to the smallest (500 replicates)	Run 8 carriers	ERP150979 carriers
SLICK1	synonymous	N'Dama 94%, Ogaden 4%, Boran 2%	N'Dama 4, Iberian 1	no carriers
SLICK1	non-synonymous	N'Dama 57%, Ankole 33%, Boran 9%	N'Dama 4, Continental European 1	no carriers
SLICK3	synonymous	Iberian 83%, Continental European 16%, Criollo 1%	no carriers	Iberian 1
SLICK3	non-synonymous	Iberian 94%, Criollo 3%, Continental European 3%	no carriers	Iberian 1
SLICK6	synonymous	Iberian 84%, Continental European 15%, Criollo 1%	no carriers	Iberian 3
SLICK6	non-synonymous	Iberian 95%, Criollo 3%, Continental European 2%	no carriers	Continental European 1, Iberian 2
SLICK2	synonymous	Continental European 54%, Criollo 29%, N'Dama 8%	Continental European 7	Continental European 12, Iberian 3

Allele	Class	Panels rarefied to the smallest (500 replicates)	Run 8 carriers	ERP150979 carriers
SLICK2	non-synonymous	Continental European 91%, Iberian 5%, Criollo 4%	Continental European 7	Continental European 12, Iberian 3

Supplementary Table S4

Supplementary Table S4. Intronic assignment by distance to the causal site. The chance level with seven panels is 0.14. A band is recorded as adjudicating only where the median posterior reaches 0.50, and the criterion is applied identically to every band.

Allele	Distance band	Sites	Modal panel	Median posterior	Adjudicates?
SLICK1	<1 kb	12	Iberian 4/5	0.24	no, near chance
SLICK1	1-5 kb	216	N'Dama 4/5	0.74	yes
SLICK1	5-10 kb	370	Iberian 2/5	1.00	yes
SLICK1	10-20 kb	721	N'Dama 4/5	0.99	yes
SLICK1	>20 kb	41	Iberian 5/5	0.19	no, near chance
SLICK3	<1 kb	13	Iberian 1/1	0.24	no, near chance
SLICK3	1-5 kb	213	Criollo 1/1	1.00	yes
SLICK3	5-10 kb	372	Criollo 1/1	1.00	yes
SLICK3	10-20 kb	720	Criollo 1/1	1.00	yes
SLICK3	>20 kb	42	Iberian 1/1	0.19	no, near chance
SLICK6	<1 kb	14	Iberian 3/3	0.24	no, near chance

Allele	Distance band	Sites	Modal panel	Median posterior	Adjudicates?
SLICK6	1-5 kb	211	Criollo 3/3	1.00	yes
SLICK6	5-10 kb	373	Criollo 3/3	0.66	yes
SLICK6	10-20 kb	718	Criollo 3/3	1.00	yes
SLICK6	>20 kb	44	Iberian 3/3	0.19	no, near chance
SLICK2	<1 kb	18	Ankole 17/22	0.34	no, near chance
SLICK2	1-5 kb	209	N'Dama 18/22	0.89	yes
SLICK2	5-10 kb	373	Ankole 9/22, N'Dama 9/22 (tie)	1.00	yes
SLICK2	10-20 kb	715	N'Dama 11/22	0.97	yes
SLICK2	>20 kb	45	Iberian 21/22	0.19	no, near chance

Supplementary Table S5

Supplementary Table S5. Frequency of SLICK1 in N'Dama compatible with observing no copies, from $(1-p)^n \geq 0.05$. Bounds are computed under the one-sided binomial sampling model at 95% confidence, where n is the number of clean alleles called and p the population frequency compatible with observing no copies. Fifteen N'Dama in the incorporated panel are called `0/0` for SLICK1 (twelve for SLICK2, which excludes frequencies above 11.7% at 24 alleles); the remainder of both panels are no-calls.

N'Dama sampled	Alleles	Excludes frequencies above
15 *(the present panel, SLICK1)*	30	**9.50%**
20	40	7.22%
50	100	2.95%
100	200	1.49%

Supplementary Table S6

Supplementary Table S6. Inbreeding control for tract length (full distributions in Supplementary Figure S3). Random positions were drawn from the same 10 Mb window on BTA20 (*Bos taurus* autosome 20), each at least 1 Mb from the receptor. The baseline is the median block extent over 150 such positions, capturing the animal's own genomic homozygosity; the ratio compares the tract observed at the locus with that baseline, and the empirical P value is obtained from the same 150 draws. The SLICK1 row rests on a single homozygous Costeño con Cuernos animal and the SLICK2 row on four haplotype copies from two homozygous animals, so both rows are anchored by very small numbers of carriers.

Allele	Homozygotes	Tract at locus	Baseline (median of 150 positions)	Ratio	Empirical P
SLICK1	1	5,177.6 kb	5,026.2 kb	1.0	0.21
SLICK2	2	**155.7 kb**	**13.6 kb**	**11.5**	**0.020**

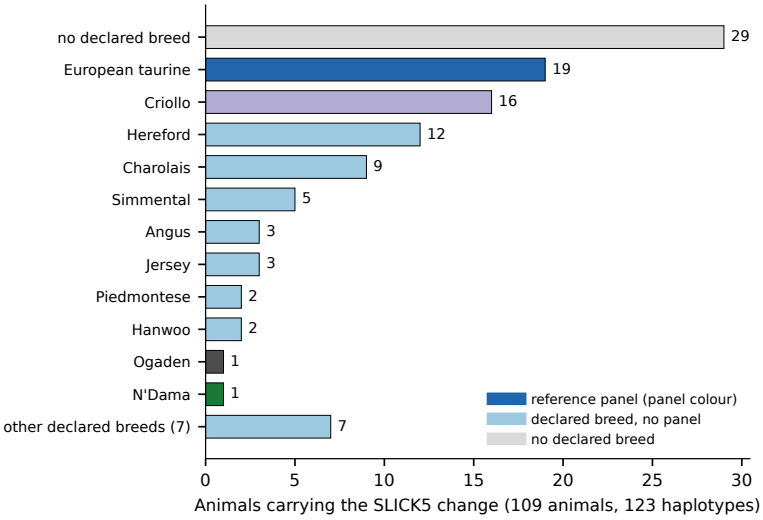
Supplementary Figure S1

Supplementary Figure S2

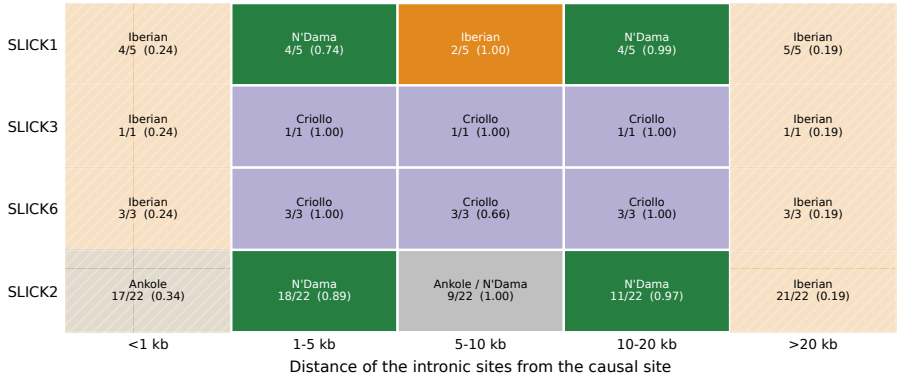
Supplementary Figure S3

References cited in this Supporting Information

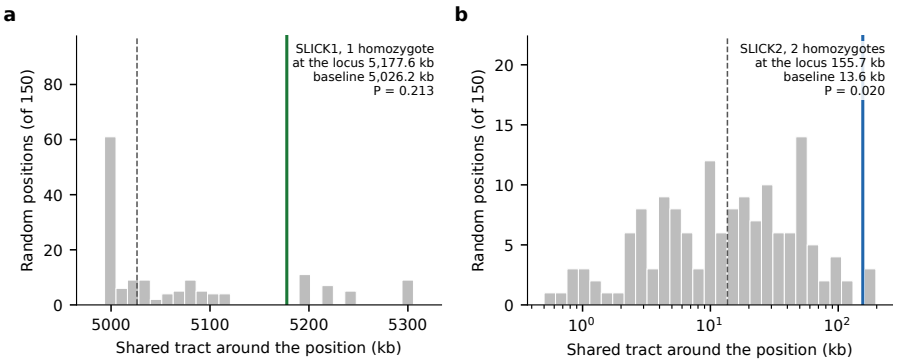
1. Flórez Murillo, J. M. *et al.* Three novel nonsense mutations of prolactin receptor found in heat-tolerant Bos taurus breeds of the Caribbean Basin. *Animal Genetics* **52**, 132–134 (2021).
2. Flórez Murillo, J. M. *et al.* Allele frequencies of PRLR truncating variants in heat-tolerant taurine cattle native to Latin America. *Animal Genetics* **55**, 910–913 (2024).
3. Gebeyehu, S. *et al.* Analysis of SLICK allele in African taurine and Zebu cattle breeds. *Animal Genetics* **56**, e13499 (2025).



Supplementary Figure S1. Animals carrying SLICK5 (chr20:39,099,228, c.1396A>T) in the phased panel, by reference panel or, for animals outside the panels, by the breed declared in the BioSample record: 109 animals and 123 haplotypes. Breeds with a single carrier animal are pooled.



Supplementary Figure S2. Intronic assignment by distance to the causal site, the data of Supplementary Table S4. Each cell gives the modal panel, its count among the carrier haplotypes and, in brackets, the median posterior. Cells with a median posterior below 0.50, near the chance level of 0.14 for seven panels, are pale and hatched; those bands do not adjudicate. Tied panels are shown together.



Supplementary Figure S3. Inbreeding control for tract length, the data of Supplementary Table S6. Histograms give the median shared tract around each of 150 random positions in the same 10 Mb window of BTA20, each at least 1 Mb from *PRLR*, in the Run 8 homozygotes for (**a**) SLICK1 (one animal) and (**b**) SLICK2 (two animals). The dashed line is the median over those positions and the coloured line the tract at the locus.

4. Flori, L. *et al.* A quasi-exclusive European ancestry in the Senepol tropical cattle breed highlights the importance of the slick locus in tropical adaptation. *PLoS ONE* **7**, e36133 (2012).
5. Huson, H. J. *et al.* Genome-wide association study and ancestral origins of the slick-hair coat in tropically adapted cattle. *Frontiers in Genetics* **5**, 101 (2014).