JAST

Genome-wide association and fine-mapping analyses identify novel candidate genes affecting serum cortisol levels using imputed whole-genome sequencing data in pigs

Tae-Hee Kim[1#], Hyeon-ah Kim[2#], Jae-Bong Lee[3], Dongwon Seo[4], Yuju Lee[1], Yong Jun Kang[2], Sang-Geum Kim[2], Sang-Hyun Han[5], Seung-Hwan Lee[6], Cedric Gondro[7], Hee-Bok Park[1,8]*, In-Cheol Cho[2]*

[1]Department of Animal Resources Science, Kongju National University, Yesan 32439, Korea
[2]Subtropical Livestock Research Center, National Institute of Animal Science, Rural Development Administration, Jeju 63242, Korea
[3]Korea Zoonosis Research Institute, Jeonbuk National University, Iksan 54531, Korea
[4]TNT Research Co., Jeonju 54810, Korea
[5]Habitat Conservation Division, Korean National Park Research Institute, Korea National Park Service, Yongju 36015, Korea
[6]Division of Animal and Dairy Science, Chungnam National University, Daejeon 34134, Korea
[7]Department of Animal Science, Michigan State University, East Lansing, MI 48824, USA
[8]Resource Science Research Institute, Kongju National University, Yesan 32439, Korea

*Corresponding authors
Hee-Bok Park
Department of Animal Resources Science, Kongju National University Yesan 32439, Korea
Tel: +82-41-330-1244
E-mail: heebokpark@kongju.ac.kr

In-Cheol Cho
Subtropical Livestock Research Center, National Institute of Animal Science, Rural Development Administration, Jeju 63242, Korea
Tel: +82-64-754-5710
E-mail: choic4753@korea.kr

Abstract

In swine breeding programs, it has now become critically important to emphasize selection for resilience to external environmental stress factors that have negatively impacted the productivity of pigs, such as those due to climate change induced temperature increases, or the intensification of housing environments. Secretion of cortisol, a neurophysiological change mediated by the hypothalamic-pituitary-adrenal axis, is a central mechanism in the biological stress response. This hormone is closely related to pig robustness and health and can serve as an informative indicator of stress resistance and robustness in pigs. To identify positional candidate genes and their genetic variants influencing blood cortisol levels, we conducted genome-wide association study (GWAS), joint linkage and linkage disequilibrium (LALD) mapping and Bayesian fine-mapping analysis in an F_2 resource population generated by crossing Duroc pigs with Korean native pigs. The data used in the study included 243 F_2 animals. We utilized imputed whole-genome sequencing data for our analyses. GWAS results revealed a genome-wide significant quantitative trait locus (q-value < 0.05) located within a ~2.46 Mb region between single nucleotide polymorphisms 7:114031215 and 7:116497417 on pig chromosome 7, which accounted for 12.65% of the phenotypic variation. LALD mapping analysis was performed to narrow down the confidence interval (CI) of the quantitative trait locus which resulted in a CI of 2.39 Mb (7:114409266~116803751). Further, to identify candidate causal genes within the 2.39 Mb region, fine-mapping analysis was

ORCID
Tae-Hee Kim
https://orcid.org/0009-0009-5551-1434
Hyeon-ah Kim
https://orcid.org/0000-0002-4203-9857
Jae-Bong Lee
https://orcid.org/0000-0001-9331-8146
Dongwon Seo
https://orcid.org/0000-0003-0548-7068
Yuju Lee
https://orcid.org/0009-0008-0806-0354
Yong Jun Kang
https://orcid.org/0000-0001-8949-7831
Sang-Geum Kim
https://orcid.org/0000-0002-0245-0465
Sang-Hyun Han
https://orcid.org/0000-0002-6615-2811
Seung-Hwan Lee
https://orcid.org/0000-0003-1508-4887
Cedric Gondro
https://orcid.org/0000-0003-0666-656X
Hee-Bok Park
https://orcid.org/0000-0002-9418-1898
In-Cheol Cho
https://orcid.org/0000-0003-3459-1999

Availability of data and material
Upon reasonable request, the datasets of this study can be available from the corresponding author.

performed within the region. The fine-mapping analysis identified *SERPINA1, ITPK1, CLMN, SERPINA12,* and *PRIMA1*, in addition to *SERPINA6*, which was previously shown to be associated with blood cortisol levels. Our results identified positional candidate genes and genetic variants associated with serum cortisol concentrations that can be included in marker panels for genomic prediction to improve selection for robustness in pigs.

Keywords: Fine-mapping, Candidate gene, Serum cortisol levels, Imputed whole-genome sequence

INTRODUCTION

With the rapid technological advancement and adoption of next-generation sequencing, genotype imputation and Bayesian statistical fine-mapping approaches, it has become feasible to conduct post genome-wide association fine-mapping analysis of quantitative trait loci (QTLs), to identify some of the genetic variants causatively associated with complex quantitative traits, such as blood cortisol levels. The genome-wide association studies (GWAS) and fine-mapping complement each other in genetic research. While GWAS is an excellent approach to identify broad genomic regions associated with complex economic traits using a sparse density of DNA markers and conventional p-values to declare statistically significant associations, fine-mapping narrows these regions to specific potential causal variants. Fine-mapping employs sophisticated statistical methods that account for linkage disequilibrium (LD) structure, utilizes denser genotyping of DNA markers, and can compute the posterior probability of causality for each variant in the genomic regions of interest. As an essential post-GWAS analysis, fine-mapping identifies putative causal variants, provides biological insights, facilitates functional studies, and improves effect estimation. By offering higher resolution and more precise identification of causal variants, fine-mapping translates GWAS discoveries into biological insights [1–3].

Although GWAS were originally developed for population studies, family-based association studies, including outbred crosses, such as F_2 intercrosses, have also become popular due to the implementation of mixed linear models (MLM) in GWAS [4–6]. Genetic selection for robustness- and health-related traits is becoming an important component of swine breeding operations because the current pig production system must cope with environmental stress factors derived from the intensification of housing environments and temperature increase due to climate changes [7]. Alterations in robustness and health can cause neuro-physiological changes, including changes in blood cortisol levels. Cortisol secretion is mainly affected by the hypothalamic-pituitary-adrenal (HPA) axis, which are the key organs of biological stress response [8,9]. Hypothalamic corticotropin-releasing hormone stimulates the release of adrenocorticotropic hormone (ACTH) from the pituitary gland, which in turn triggers cortisol production in the adrenal cortex. The sensitivity of the adrenal glands to ACTH is crucial for the regulation of cortisol secretion. Cortisol bioavailability is critically modulated by corticosteroid-binding globulin (CBG). Through its high specific affinity for cortisol, CBG plays a vital role in regulating serum cortisol levels and their access to target cells. The majority of cortisol in the bloodstream is transported bound to CBG, while only a small fraction of the total serum cortisol remains unbound and biologically active [10,11].

Cortisol, a crucial hormone for facilitating adaptive stress responses, ensures the body's energy supply in dynamic environmental conditions. This hormone influences multiple physiological processes, including appetite regulation, glucose metabolism, and fatty acid metabolism. Additionally, cortisol suppresses inflammation, modulates immune function, and mobilizes energy resources. It also affects energy storage. However, to be beneficial, cortisol levels must be maintained within an optimal concentration range [12]. In pigs, cortisol levels correlate with various

Lee JB, Seo D, Lee Y, Kang YJ, Kim SG, Han SH, Lee SH, Gondro C, Park HB, Cho IC.

economically important traits. Higher cortisol is associated with reduced growth and feed efficiency, and influences body weight, carcass characteristics, immune function, leanness, and meat quality [13,14]. The association between cortisol levels and various economically important traits has been well documented across species. For instance, studies in chickens and sheep have linked cortisol release to growth-related traits [15,16].

The secretion of cortisol regulated by the HPA axis is individually variable to stress responsiveness and is influenced by genetic determinants – heritability estimates of cortisol levels range between medium to high which makes it amenable to selection in breeding programs [17,18]. Hence, it is important to identify the genetic factors underlying blood cortisol levels and develop DNA markers to improve of the robustness and health of pigs, which are tightly linked to their welfare and productivity. Despite the importance of cortisol, only a limited number of GWAS and RNA seq analysis have been conducted to identify candidate genes that affect blood cortisol levels [19]. For example, previous GWA studies have identified *SERPINA6* and *SERPINA1* genes, which encode CBG, implicated in the regulation of blood cortisol levels in humans and pigs [20–22]; Crawford et al. reported strong evidence that genetic variants in the *SERPINA6/SERPINA1* locus primarily affect *SERPINA6* gene expression in the liver, which likely influences cortisol levels and its delivery to other tissues [21]. In goats, an RNA-seq experiment aimed at identifying genes and pathways associated with increased cortisol levels due to transportation stress revealed significant enrichment of genes involved in inflammation and apoptosis pathways [23].

Here, we present the results of GWAS and post-GWAS fine-mapping analyses in which we aimed to identify positional candidate genes and genetic variants affecting serum cortisol concentrations that are involved in the mechanisms of stress response in pigs. Our findings could provide insights into stress response mechanisms and targets for improving pig welfare and productivity through genomic information enabled accurate genetic evaluation.

MATERIALS AND METHODS

Animals and phenotype measurement

We generated an F_2 resource population by crossing Duroc and Korean Native Pig (KNP) pigs from Jeju Island [13–24]. There are two types of indigenous pigs in Korea: those living on the main peninsula of Korea and those living on Jeju Island. The Jeju Island indigenous pigs are particularly interesting as they have unique genetic properties that are very distinct from those of pigs raised on the Korean Peninsula since they have been largely isolated on Jeju Island for more than 1,000 years [25,26]. Hereafter, the Jeju native pig will be referred to as KNP. The coat color of KNP is black, and similarly to most indigenous breeds, their growth performance is stunted in comparison to modern commercial pig breeds. They possess however excellent meat quality attributes, such as a solid fat structure, a darker red meat color, and high levels of marbling [27–29]. In this study, nine purebred Duroc pigs were mated with five purebred KNP pigs to produce 36 F_1 animals. Subsequently, the F_1 animals were intercrossed to produce 345 F_2 animals. From the F_2 progeny comprised 31 full-sib families, high-quality serum cortisol level data from 243 F_2 animals (133 males and 110 females) were obtained as phenotypic data, and the corresponding genotypic data were included for this study. All animals were raised at the experimental farm of the National Institute of Animal Science, Rural Development Administration, Jeju, Republic of Korea. They were fed *ad libitum*, and the males were not castrated. All experimental procedures were performed in accordance with national and institutional guidelines and were approved by the Ethical Committee of the National Institute of Animal Science (No. 2020-446).

Blood samples (10 mL) were collected at 140 days of age from the jugular veins of the 243 F_2

offspring. This specific age was chosen based on the ease of blood collection, taking into account the body size of pigs at this stage of development. The 140 days of age provided a balance between the pigs being sufficiently mature for straightforward blood sampling and not being too large to handle safely and efficiently. To minimize sampling variance, the blood sampling was conducted in the morning as soon as the research farm staffs were present at the farm. Serum cortisol levels were measured using a commercially available enzyme-linked immunosorbent assay kit (Endocrine Technologies, Newark, CA, USA). Descriptive statistics are presented in Table 1. The cortisol phenotype showed significant deviation from normality and was transformed using a natural logarithm to remove skewness.

Single nucleotide polymorphism marker data

Genomic DNA was isolated from the blood samples using a standard sucrose-proteinase K method. Genotyping was conducted using the Porcine SNP 60K BeadChip technology (Illumina, San Diego, CA, USA), which contained 61,565 SNPs across the whole genome. The SNPs were filtered for a minor allele frequency < 1%, genotype call rate < 95%, and p-value of χ^2-test for Hardy–Weinberg equilibrium ≤ 0.000001. Additionally, Mendelian inconsistencies in the F_2 pedigree were assessed using the SNP marker information. The quality control procedures were performed using the PLINK v 1.90 [30]. In total, 39,463 SNP markers were retained and used for further analysis.

Whole genome sequencing and genotype imputation

To scale up the information obtained from the Porcine SNP 60K BeadChip data of the study population to whole-genome sequence (WGS) level, genotype imputation was performed. As a reference panel for imputation, we collected WGS data from 56 pigs, including 7 Duroc, 19 KNP, and 30 Duroc × KNP F_2 (DK F_2) individuals. Among them, some of the Duroc (7) and KNP (5) pigs were the parental animals of the F_2 resource population. The 30 DK F_2 pigs were sampled from the F_2 progeny of the resource population. Whole-genome sequencing was performed using the Illumina HiSeq platform (Illumina). DNA libraries were prepared according to the manufacturer's instructions and sequenced using the paired-end 150 bp sequencing protocol. The sequencing coverage for these individuals was approximately 30×. In addition, whole-genome resequencing data from 14 KNPs were obtained from the NCBI Sequence Read Archive (SRA). These sequencing data were accessed under the BioProject accession number PRJNA254936. These data were also included in the subsequent procedures for preparing the reference sequence dataset for genotype imputation.

The raw sequencing data were processed and analyzed using a standard bioinformatics pipeline. Briefly, the quality of raw sequencing reads was assessed using FastQC (http://www.bioinformatics. babraham.ac.uk/projects/fastqc). The cleaned reads were then mapped to the *Sus scrofa* reference genome (*Sscrofa*11.1) using Burrows-Wheeler Aligner (BWA) with default parameters [31]. Duplicate reads were marked and removed using Picard MarkDuplicates (http://broadinstitute. github.io/picard). Variant calling was performed using Genome Analysis Toolkit (GATK), following the best practices for variant discovery [32]. The resulting variants were filtered based on quality scores, depth of coverage, and other standard parameters using GATK VariantFiltration to retain high-confidence genetic variants, ultimately leading to the acquisition of a total of 23,308,271 DNA markers.

Because of the substantial difference in genome coverage between the 60 K SNP data and WGS data, a two-step imputation strategy was employed to obtain WGS markers from the 60K SNP data. Initially, a subset of SNPs was extracted at regular intervals (window size of approximately 28

variants) from the WGS data obtained from the reference samples, resulting in 824,938 genetic variants. Subsequently, 1-step imputation was performed for the test sample, consisting of nine Duroc, five KNP, and 243 DK F_2 pigs. This 1-step imputation yielded a total of 578,493 DNA markers. Using these marker data from the test sample, a 2-step imputation was conducted to obtain the WGS data, resulting in a total of 15,542,014 genetic variants. Haplotyping at the whole-genome level for both the reference and test samples was performed using Beagle version 2.4.1 [33]. Beagle was strictly used for haplotype construction. Genotype imputation was then performed using Minimac4 based on chromosome-wise constructed haplotype information (https://github.com/statgen/Minimac4). An imputation accuracy was evaluated using the correlation coefficient (r^2), with a filtering criterion set at an r^2 value of 0.6 or higher. The imputed genetic variants were further filtered for a minor allele frequency < 1% and P-value of χ^2-test for Hardy-Weinberg equilibrium ≤ 0.000001. In total, 13,720,525 DNA markers were retained and used for further analyses.

Estimation of heritability and genome-wide association studies of serum cortisol levels

The efficient mixed-model association (EMMA) option of the rMVP package was used to estimate the heritability of the serum cortisol levels recorded in this study [34], and the following linear mixed model was used for the analysis:

$$y = Xb + Zu + e \quad (1)$$

where y is the vector of the log transformed serum cortisol levels; b is the vector of fixed effects, including the intercept, the effect of sex, the effect of slaughtering batch (with nine levels), and body weight at 140 days of age; u is the vector of random additive effects following a normal distribution $u \sim N(0, G\sigma_a^2)$, in which G is the genomic relationship matrix constructed using 13,720,525 DNA markers and σ_a^2 is the additive genetic variance; e is the vector of random residual effects following a normal distribution $e \sim N(0, I\sigma_e^2)$, in which I is the identity matrix and σ_e^2 is the residual variance; X and Z are the incidence matrices for b and u, respectively.

A GWAS adjusted for the familial relatedness within the F_2 intercross was performed using the MLM option of the rMVP program [34]. The following linear MLM was used to assess the association between SNP markers and the serum cortisol levels:

$$y = Xb + Z_1a + Z_2u + e \quad (2)$$

where, y is the vector of the serum cortisol levels; b is the vector of fixed effects including sex, the effect of slaughtering batch (with nine levels), and body weight at 140 days of age; a is the SNP marker effects; u is the vector of random additive effects with a distribution $u \sim N(0, G\sigma_a^2)$, where G is the genomic relationship matrix that was constructed using the 39,463 SNP markers; σ_a^2 is the additive genetic variance; e is a vector of random residuals following a distribution $e \sim N(0, I\sigma_e^2)$, in which I is the identity matrix and σ_e^2 is the residual variance. Z_1 is the incidence vector for a. X and Z_2 are the incidence matrices for b and u. The percentage of phenotypic variance explained by a marker (%Var_{SNP}) was computed as follows [35]:

$$\%Var_{SNP} = 100 \times \frac{2p(1-p)\alpha^2}{\sigma_p^2} \quad (3)$$

where *p* is the minor-allele frequency of the SNP marker; α is the additive genetic effect of the

DNA marker; σ_p^2 is the phenotypic variance for each meat quality-related trait. The p, α and σ_p^2 were estimated using the rMVP program. The q-value adjusted genome-wide suggestive (q-value < 0.10) and significant (q-value < 0.05) thresholds were used to address the multiple testing issues [36].

Joint linkage and linkage disequilibrium mapping analysis

Initial fine-mapping of QTL identified by the GWAS was performed by exploiting linkage and linkage disequilibrium (LALD) using a haplotype-based approach: 1) We used CRI-MAP version 2.503, developed by Evans and Maddox (http://www.animalgenome.org/bioinfo/tools/share/crimap), to establish a genetic linkage map of SSC7 using 899 SNP markers. These makers were selected from the Porcine SNP 60K BeadChip genotype data. 2) The founder haplotypes were then reconstructed using the DualPHASE program [37] which combines LALD information through a Hidden Markov Model setting. 3) The haplotypes then were incorporated into the following linear mixed-effect model including fixed effects (sex, slaughtering batch, and body weight at 140 days of age), random effects (i.e., the effects of the founder haplotypes and the additive polygenic effect), and random residual terms to perform high-resolution QTL mapping using QxPAK version 5.05 [38]. A 1.0-LOD drop support interval was employed to estimate the confidence interval (CI) at the location of QTL location [39].

Further Bayesian fine-mapping analyses for identifying candidate causal gene(s) for serum cortisol levels

To refine the critical region identified by the LALD analyses, a further fine-mapping approach based on the FINEMAP program was conducted [40]. The FINEMAP program uses a shotgun stochastic search algorithm that incorporates summary statistics (i.e., Z-score) from GWAS and the LD correlation structure calculated from the SNP markers in the region of interest to calculate the posterior probability of each SNP marker being a candidate causal variant. A threshold of 5% for the posterior probability was applied to select markers associated with candidate causal variants from the FINEMAP analyses [41]. A regional association plot with LD information was drawn using IntAssoPlot (https://github.com/whweve/IntAssoPlot).

Positional candidate gene analyses

A list of genes annotated within the QTL region was extracted from the NCBI database release 85 based on *Sus scrofa* 11.1 assembly (NCBI accession ID: NC_010454.4). A list of genes in each QTL region was obtained from the NCBI database. A comparative analysis with previously reported QTL locations for the trait was conducted using the Animal QTLdb [19]. The candidate causal variants identified using the FINEMAP program were annotated using the ENSEMBL pig genome database (https://asia.ensembl.org/).

RESULTS AND DISCUSSION

Descriptive summary statistics and estimated heritability of the serum cortisol trait in the DK F_2 pigs are presented in Table 1. The mean value and range of serum cortisol levels were 21.18 ng/mL and 61.6 ng/mL (2.30–63.90 ng/mL) in the F_2 population. The estimate of heritability for the cortisol traits was 0.32, indicating that a considerable contribution of genetic effects to the phenotypic variation in the trait of interests is considerable. A total of 23,129,957 genetic variants were imputed using Beagle and Minimac4. The average imputation accuracy (r^2) for the entire set of genetic variants was 0.67. A total of 7,587,919 variants had an imputation accuracy below the

Table 1. Basic statistics for the phenotypic data in the DK F$_2$ pigs

Phenotype	Total	N	*N[1]	Mean	StDev	Min	Max	h^2
Cortisol (ng/mL)	243	234	9	21.18	12.46	2.30	63.90	0.32

[1]*Number of individuals with missing phenotypes.

StDev, Standard deviation; h^2, heritability.

threshold of 0.6 (r^2). After filtering out these low-accuracy variants, the final average imputation accuracy (r^2) increased to 0.968 (Table 2). After an additional QC filtering step with PLINK for MAF and Hardy-Weinberg equilibrium deviations, a total of 13,720,525 DNA markers remained for subsequent analyses.

Genome-wide association studies

To investigate the genetic structure underlying the cortisol trait in pigs, we used an F$_2$ intercross between the Duroc and KNP pigs. Using this F$_2$ intercross population, we detected a single genome-wide significant QTL (q-value<0.05), comprising 34 DNA markers with extremely tight LD among these markers. The QTL spans a ~2.46 Mb region between 7:114031215 and rs326739326 (7:116497417), for serum cortisol levels on SSC7 (Fig. 1). However, a single most significant DNA marker associated with the phenotype of interest was not detected (Table 3). This significant QTL accounted for 12.65% of the phenotypic variance of the serum cortisol levels in the pig population (Table 3).

Table 2. Evaluation of the genotype imputation accuracy

Chromosome	Total variants[1]	Mean r^2	Pass	Pass mean r^2
1	2,144,560	0.629	1,357,945	0.963
2	1,478,956	0.653	974,485	0.962
3	1,393,829	0.591	832,824	0.960
4	1,366,608	0.701	967,277	0.971
5	1,098,829	0.656	728,091	0.965
6	1,566,762	0.652	1,033,019	0.965
7	1,323,022	0.669	893,604	0.966
8	1,515,476	0.719	1,100,848	0.970
9	1,578,875	0.607	965,412	0.968
10	1,004,923	0.725	735,704	0.971
11	978,320	0.667	658,497	0.971
12	755,615	0.636	484,594	0.968
13	1,781,238	0.706	1,268,266	0.971
14	1,358,977	0.674	922,556	0.968
15	1,263,992	0.663	844,095	0.970
16	1,034,235	0.712	743,793	0.969
17	805,850	0.653	530,034	0.968
18	679,890	0.731	500,994	0.969
Total	23,129,957	0.669	15,542,038	0.968

[1]Total_variants, total number of imputed and genotyped variants per chromosome; Mean r^2, average r^2 of all SNPs; Pass, number of variants with r^2 > 0.6; Pass mean r^2, average r^2 of variants after filtering out genotypes with r^2 < 0.6.
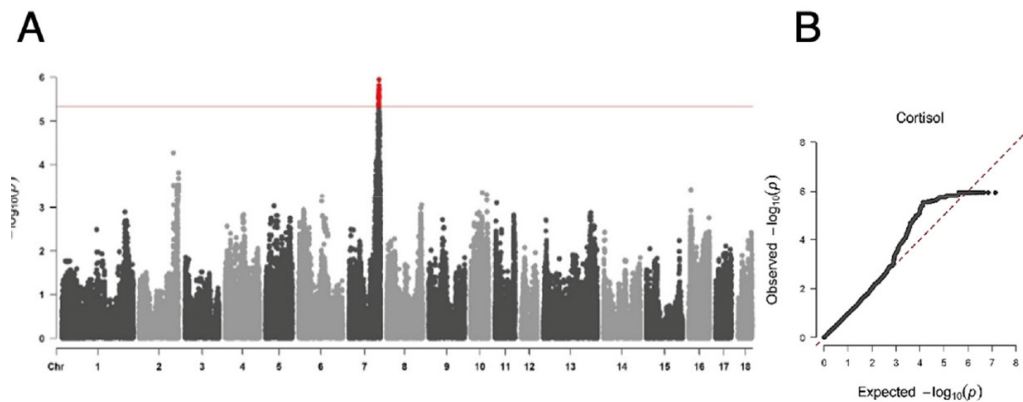
Fig. 1. Whole-genome imputed sequence association analysis for serum cortisol levels in the DK cross. The red horizontal line represents the genome-wide significant threshold (p = 4.78E-06) A. Manhattan plot B. QQ plot (Genomic inflation factor = 1.032)

Table 3. Summary of the QTL affecting serum cortisol identified by the GWAS

SSC	N_{snp}[1]	Interval (Mb)	Centered SNP	Effect	SE	%var	p-value	q-value
7	34	115575190-115584273	7:115580597	0.3216	0.0643	12.65	1.13E-06	0.0398

[1]N_{snp}, number of SNPs included in interval; Interval (Mb), range of Nsnp; %var, percentage of phenotypic variance explained by the centered SNPs; q-value, FDR-based q-value.
QTLs, quantitative trait loci; GWAS, genome-wide association studies; SSC, *Sus scrofa* chromosome; SNP, single nucleotide polymorphism; FDR, false discovery rate.

Linkage and linkage disequilibrium mapping analysis

We performed integrated LALD mapping to reduce the CI of the QTL using the 899 markers (from the Porcine SNP 60K BeadChip genotype data) located across the SSC7. The 1-LOD drop method was used to estimate the CI of the QTL in SSC7 (Fig. 2A). The CI was reduced to 2.39 Mb (7:114,409,266–116,803,751). This region overlaps with a previously reported QTL region that influences pig blood cortisol levels [42–46]. The 2.39 Mb region in SSC7 encompassed 24 annotated genes with 14,337 DNA markers in the *Sus scrofa* 11.1 genome and imputed genotype dataset (Figs. 2B and 2C).

Further Bayesian fine-mapping for identifying novel positional candidate gene(s) for serum cortisol levels

To further refine the critical region harbouring causative genes, we conducted fine-mapping of the 2.39 Mb region identified by the LALD analysis using the FINEMAP program. To enhance the fine-mapping accuracy for identifying putative causal genes, we first selected DNA markers located within the 24 positional candidate genes in the critical region based on the GWAS results. For each positional candidate gene, we retained only the DNA marker with the lowest p-value. Subsequently, we manually chose 297 DNA markers evenly distributed across the 2.39 Mb region. We then pruned these DNA markers using the LD pruning option in the PLINK program, resulting in 17 DNA markers evenly distributed throughout the 2.39 Mb critical region. Hence, a total of 41 DNA markers were included in the Bayesian fine-mapping analysis. No single DNA marker showed a substantial posterior probability (greater than 0.5) in the critical region (Table 4). DNA markers demonstrated low to medium levels of posterior probability, ranging from 0.066 to 0.116. This was most likely due to the modest sample size (N = 243) of the F_2 cohort. Nevertheless, our Bayesian fine-mapping approach allowed us to prioritize potential causal genes. While we acknowledge the limitations of our study, including the moderate sample size and the inability to pinpoint specific
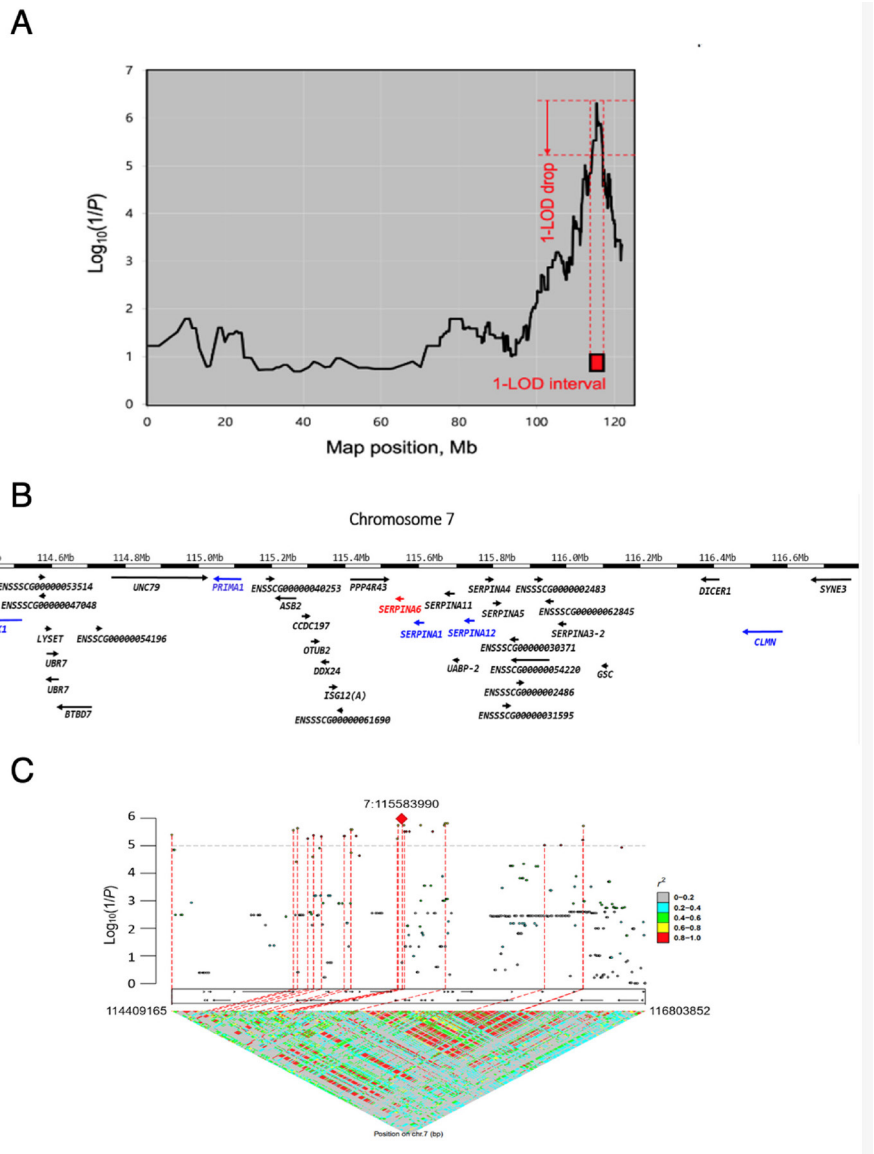
Fig. 2. Fine-mapping analysis of QTL for serum cortisol levels on SSC7. (A) LALD analyses, (B) Positional candidate genes in the 1-LOD interval (2.39 Mb) region. (C) IntAssoPlot for the 1-LOD interval (2.39 Mb) region. QTLs, quantitative trait loci; LALD, linkage and linkage disequilibrium; LOD, logarithm of odds.

causal variants, the fine-mapping approach has enabled us to create a prioritized list of potential causal candidate genes. This list provides valuable direction for subsequent studies, demonstrating the usefulness of fine-mapping even in situations where definitive causal variant identification is not achieved. For example, the most significant QTL region, explained by 34 variants in Table 3, spans only 9,083 bp. This narrow region contains just two genes: *SERPINA6* and *SERPINA1*. Based solely on the p-values obtained from the conventional GWAS, it would be challenging to determine which of these genes should be prioritized for further investigation. However, as shown in Table 4, the posterior probabilities derived from our fine-mapping analysis allow us to distinguish between *SERPINA6* and *SERPINA1*, providing a basis for prioritization. Recently, Uemoto et al. reported results of GWAS and LALD analysis on Landrace pigs [22]. While our study shares similarities with their study, employing both GWAS and LALD approaches, major differences underscore the novelty of our work. We used whole-genome imputed sequence variants instead of a conventional

Table 4. Positional candidate genes and posterior probabilities for significant variants in SSC7 Bayesian fine-mapping approach

Marker[1]	Position	Effect	SE	[2]%var	p-value	q-value	Gene	Annotation	P.P
7:115583990	115583990	0.322	0.064	0.127	1.13E-06	0.040	SERPINA1	3' UTR variant	0.116
7:114409265	114409265	-0.289	0.061	0.103	4.02E-06	0.047	ITPK1	Intron variant	0.094
7:116490750	116490750	-0.299	0.061	0.111	1.92E-06	0.040	CLMN	Intron variant	0.091
7:115727874	115727874	0.307	0.064	0.115	3.07E-06	0.040	SERPINA12	Intron variant	0.084
7:115045439	115045439	-0.298	0.062	0.110	2.33E-06	0.040	PRIMA1	Intron variant	0.083
7:115802806	115802806	-0.302	0.061	0.113	1.55E-06	0.040	SERPINA5	Intron variant	0.077
7:115792190	115792190	-0.302	0.061	0.113	1.55E-06	0.040	SERPINA4	3' UTR variant	0.077
7:115342222	115342222	0.300	0.064	0.110	4.49E-06	0.048	DDX24	Intron variant	0.075
7:115280790	115280790	0.300	0.064	0.110	4.49E-06	0.048	CCDC197	5' UTR variant	0.075
7:115023328	115023328	-0.294	0.061	0.107	2.72E-06	0.040	UNC79	Intron variant	0.074
7:115667344	115667344	-0.298	0.061	0.110	1.82E-06	0.040	SERPINA11	Intron variant	0.074
7:115554168	115554168	-0.298	0.061	0.110	1.82E-06	0.040	SERPINA6	Open chromatin	0.073
7:115314601	115314601	-0.292	0.061	0.106	2.61E-06	0.040	OTUB2	Intron variant	0.066

[1]Marker, Marker ID; [2]%var, percentage of phenotypic variance explained by the marker; q-value, FDR based q-value; Gene, positional candidate gene; Annotation, ENSEMBL variant annotation; P.P, posterior probability of the marker to be causal.

FDR, false discovery rate.

60K SNP chip, providing comprehensive genomic coverage including rare and novel variants. In addition, we applied Bayesian fine-mapping to prioritize potential causal candidate genes, rather than presenting a conventional list of positional candidates. This approach allowed for more precise localization and prioritization of potential causal genes and variants.

In this critical region, *SERPINA6*, a gene encoding CBG, has been proposed as a putative causal gene accounting for this QTL effect [34]. A p.Arg307Gly (c.919T>C) substitution in *SERPINA6* was previously shown to increase CBG capacity and decrease CBG affinity for cortisol *in vitro* [47,48]. These findings suggest that *SERPINA6* is a positional and functional candidate gene for the QTL associated with blood cortisol levels in pigs. However, the fine-mapping results based on posterior probability ranked the top five genes as *SERPINA1*, *ITPK1*, *CLMN*, *SERPINA12*, and *PRIMA1* genes, all with higher posterior probability than *SERPINA6* (Table 4). In this study, we could not evaluate the p.Arg307Gly (c.919T>C) substitution in *SERPINA6* because the corresponding variant at position 7:115541678 was not present among the 14,337 DNA markers initially considered for the fine-mapping.

SERPINA1, which encodes alpha-1 antitrypsin, inhibits neutrophil elastase and regulates cortisol secretion by influencing CBG cleavage and subsequent reconfiguration of the reactive center loop [49]. Heterozygosity of *SERPINA1* mutations leads to alpha-1 antitrypsin deficiency and CBG cleavage, thereby increasing the free cortisol fraction [50]. Intracellular inositol triphosphate (IP$_3$) is involved in various signal transduction pathways that affect cellular metabolisms [51]. Increased IP$_3$ levels can trigger cortisol secretion by the adrenal zona fasciculate cells [52]. IP$_3$ is produced by PIP$_2$ hydrolysis, and *ITPK1* synthesizes IP$_4$, IP$_5$, and IP$_6$. Hence, *ITPK1* polymorphisms could potentially affect inositol metabolism, which may be associated with cortisol secretion via the inositol pathways. *CLMN*, which encodes a calponin-like transmembrane domain protein, regulates cell cycle exit and neurite outgrowth in murine neuroblastoma cells. Genetic variants of *CLMN* may influence synaptic function and organization, resulting in the HPA axis [53,54]. *SERPINA12*, also known as VASPIN, is an adipokine belonging to the serpin protein family. Cortisol plays an important role in adipose tissue by influencing adipokine expression, insulin sensitivity, and fatty acid metabolism [55]. *PRIMA1* encodes Proline-Rich Membrane Anchor 1 protein, which is required for anchoring of acetylcholinesterase to neuronal synapses. Acetylcholin, a product of

acetylcholinesterase, is released in response to stress, and activates the HPA axis, which is linked to cortisol production [56]. However, there is still a paucity of direct genetic evidence on the effects of *CLMN* and *PRIMA1* on blood cortisol levels.

CONCLUSION

The intensification of housing environments and the impact of climate change on current production systems have negatively impacted the breeding and growth of pigs. The cortisol level is an important indicator of stress resistance and robustness in pigs. This hormone levels in pigs could indicate environmental stress, potentially intensified by climate change. Rising temperatures and erratic weather patterns increase heat stress, alter food and water availability, and expose livestock to new pathogens, elevating cortisol and affecting animal health and productivity. Studying cortisol levels in pig populations like the KNP and its crossbreeds can provide insights into their adaptation to changing environments [57]. To elucidate the molecular genetic basis of breeding robustness, we identified QTLs affecting basal serum cortisol levels using imputed whole-genome sequencing data-based GWAS, LALD analysis, and Bayesian fine-mapping approaches. This study identified novel positional candidate genes (*SERPINA1*, *ITPK1*, *CLMN*, *SERPINA12*, and *PRIMA1*) in addition to the previously known *SERPINA6*. Our results provide a basic understanding for the development of genetic markers to improve the robustness of pigs.

References

1. Slatko BE, Gardner AF, Ausubel FM. Overview of next-generation sequencing technologies. Curr Protoc Mol Biol. 2018;122:e59. https://doi.org/10.1002/cpmb.59
2. Li Y, Willer C, Sanna S, Abecasis G. Genotype imputation. Annu Rev Genomics Hum Genet. 2009;10:387-406. https://doi.org/10.1146/annurev.genom.9.081307.164242
3. Schaid DJ, Chen W, Larson NB. From genome-wide associations to candidate causal variants by statistical fine-mapping. Nat Rev Genet. 2018;19:491-504. https://doi.org/10.1038/s41576-018-0016-z
4. Ledur MC, Navarro N, Pérez-Enciso M. Large-scale SNP genotyping in crosses between outbred lines: how useful is it? Heredity. 2010;105:173-82. https://doi.org/10.1038/hdy.2009.149
5. 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
6. Lee JB, Lim JH, Park HB. Genome-wide association studies to identify quantitative trait loci and positional candidate genes affecting meat quality-related traits in pigs. J Anim Sci Technol. 2023;65:1194-204. https://doi.org/10.5187/jast.2023.e70
7. Knap PW. Breeding robust pigs. Aust J Exp Agric. 2005;45:763-73. https://doi.org/10.1071/EA05041
8. Joseph JJ, Golden SH. Cortisol dysregulation: the bidirectional link between stress, depression, and type 2 diabetes mellitus. Ann NY Acad Sci. 2017;1391:20-34. https://doi.org/10.1111/nyas.13217
9. Ataallahi M, Nejad JG, Park KH. Selection of appropriate biomatrices for studies of chronic stress in animals: a review. J Anim Sci Technol. 2022;64:621-39. https://doi.org/10.5187/jast.2022.e38
10. Mormede P, Terenina E. Molecular genetics of the adrenocortical axis and breeding for robustness. Domest Anim Endocrinol. 2012;43:116-31. https://doi.org/10.1016/

j.domaniend.2012.05.002

11. Verbeeten KC, Ahmet HA. The role of corticosteroid-binding globulin in the evaluation of adrenal insufficiency. J Pediatr Endocrinol Metab. 2018;31:107-15. https://doi.org/10.1515/jpem-2017-0270

12. Moisan MP, Castanon N. Emerging role of corticosteroid-binding globulin in glucocorticoid-driven metabolic disorders. Front Endocrinol. 2016;7:160. https://doi.org/10.3389/fendo.2016.00160

13. Foury A, Devillers N, Sanchez MP, Griffon H, Le Roy P, Mormède P. Stress hormones, carcass composition and meat quality in Large White × Duroc pigs. Meat Sci. 2005;69:703-7. https://doi.org/10.1016/j.meatsci.2004.11.002

14. Li J, Kim IH. Effects of levan-type fructan supplementation on growth performance, digestibility, blood profile, fecal microbiota, and immune responses after lipopolysaccharide challenge in growing pigs. J Anim Sci. 2013;91:5336-43. https://doi.org/10.2527/jas.2013-6665

15. Yang L, Wang X, He T, Xiong F, Chen X, Chen X et al. Association of residual feed intake with growth performance, carcass traits, meat quality, and blood variables in native chickens. J Anim Sci. 2020;98:skaa121. https://doi.org/10.1093/jas/skaa121

16. Hewagalamulage SD, Lee TK, Clarke IJ, Henry BA. Stress, cortisol, and obesity: a role for cortisol responsiveness in identifying individuals prone to obesity. Domest Anim Endocrinol. 2016;56 Suppl:S112-20. https://doi.org/10.1016/j.domaniend.2016.03.004

17. Bartels M, Van den Berg M, Sluyter F, Boomsma DI, de Geus EJC. Heritability of cortisol levels: review and simultaneous analysis of twin studies. Psychoneuroendocrinology. 2003;28:121-37. https://doi.org/10.1016/S0306-4530(02)00003-3

18. Larzul C, Terenina E, Foury A, Billon Y, Louveau I, Merlot E, et al. The cortisol response to ACTH in pigs, heritability and influence of corticosteroid-binding globulin. Animal. 2015;9:1929-34. https://doi.org/10.1017/S1751731115001767

19. Hu ZL, Park CA, Reecy JM. Bringing the Animal QTLdb and CorrDB into the future: meeting new challenges and providing updated services. Nucleic Acids Res. 2022;50:D956-61. https://doi.org/10.1093/nar/gkab1116

20. Bolton JL. Genome wide association identifies common variants at the *SERPINA6/SERPINA1* locus influencing plasma cortisol and corticosteroid binding globulin. PLOS Genet. 2014;10:e1004474. https://doi.org/10.1371/journal.pgen.1004474

21. Crawford AA, Bankier S, Altmaier E, Barnes CLK, Clark DW, Ermel R, et al. Variation in the *SERPINA6/SERPINA1* locus alters morning plasma cortisol, hepatic corticosteroid binding globulin expression, gene expression in peripheral tissues, and risk of cardiovascular disease. J Hum Genet. 2021;66:625-36. https://doi.org/10.1038/s10038-020-00895-6

22. Uemoto Y, Ichinoseki K, Matsumoto T, Oka N, Takamori H, Kadowaki H, et al. Genome-wide association studies for production, respiratory disease, and immune-related traits in Landrace pigs. Sci Rep. 2021;11:15823. https://doi.org/10.1038/s41598-021-95339-2

23. Naldurtiker A, Batchu P, Kouakou B, Terrill TH, McCommon GW, Kannan G. Differential gene expression analysis using RNA-seq in the blood of goats exposed to transportation stress. Sci Rep. 2023;13:1984. https://doi.org/10.1038/s41598-023-29224-5

24. Cho IC, Park HB, Ahn JS, Han SH, Lee JB, Lim HT, et al. A functional regulatory variant of MYH3 influences muscle fiber-type composition and intramuscular fat content in pigs. PLOS Genet. 2019;15:e1008279. https://doi.org/10.1371/journal.pgen.1008279

25. Cho IC, Han SH, Fang M, Lee SS, Ko MS, Lee H, et al. The robust phylogeny of Korean wild boar (*Sus scrofa* coreanus) using partial D-loop sequence of mtDNA. Mol Cells. 2009;28:423-

30. https://doi.org/10.1007/s10059-009-0139-3

26. Lee E, Jang JC, Oh SH. The current status of Korean native pig production. J Anim Sci Technol. 2023;65:1169-79. https://doi.org/10.5187/jast.2023.e120

27. Kim DH, Seong PN, Cho SH, Kim JH, Lee JM, Jo C, et al. Fatty acid composition and meat quality traits of organically reared Korean native black pigs. Livest Sci. 2009;120:96-102. https://doi.org/10.1016/j.livsci.2008.05.004

28. Ko KB, Kim GD, Kang DG, Kim YH, Yang ID, Ryu YC. Comparison of pork quality and muscle fiber characteristics between Jeju Black pig and domesticated pig breeds. J Anim Sci Technol. 2013;55:467-73. https://doi.org/10.5187/JAST.2013.55.5.467

29. Kim GW, Kim HY. Physicochemical properties of M. longissimus dorsi of Korean native pigs. J Anim Sci Technol. 2018;60:6. https://doi.org/10.1186/s40781-018-0163-y

30. Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience. 2015;4:7. https://doi.org/10.1186/s13742-015-0047-8

31. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics. 2009;25:1754-60. https://doi.org/10.1093/bioinformatics/btp324

32. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20:1297-303. https://doi.org/10.1101/gr.107524.110

33. Browning SR, Browning BL. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. Am J Hum Genet. 2007;81:1084-97. https://doi.org/10.1086/521987

34. Yin L, Zhang H, Tang Z, Xu J, Yin D, Zhang Z, et al. rMVP: a memory-efficient, visualization-enhanced, and parallel-accelerated tool for genome-wide association study. Genom Proteom Bioinf. 2021;19:619-28. https://doi.org/10.1016/j.gpb.2020.10.007

35. Falconer DS, Mackay TFC. Introduction to quantitative genetics. 4th ed. Essex: Longman; 1996.

36. Storey JD, Tibshirani R. Statistical significance for genomewide studies. Proc Natl Acad Sci USA. 2003;100:9440-5. https://doi.org/10.1073/pnas.1530509100

37. Druet T, Georges M. A hidden Markov model combining linkage and linkage disequilibrium information for haplotype reconstruction and quantitative trait locus fine mapping. Genetics. 2010;184:789-98. https://doi.org/10.1534/genetics.109.108431

38. Pérez-Enciso M, Misztal I. Qxpak.5: old mixed model solutions for new genomics problems. BMC Bioinform. 2011;12:202. https://doi.org/10.1186/1471-2105-12-202

39. Lander ES, Botstein D. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics. 1989;121:185-99. https://doi.org/10.1093/genetics/121.1.185

40. Benner C, Spencer CCA, Havulinna AS, Salomaa V, Ripatti S, Pirinen M. FINEMAP: efficient variable selection using summary data from genome-wide association studies. Bioinformatics. 2016;32:1493-501. https://doi.org/10.1093/bioinformatics/btw018

41. Mahajan A, Taliun D, Thurner M, Robertson NR, Torres JM, Rayner NW, et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. Nat Genet. 2018;50:1505-13. https://doi.org/10.1038/s41588-018-0241-6

42. Désautés C, Bidanel JP, Milan D, Iannuccelli N, Amigues Y, Bourgeois F, et al. Genetic linkage mapping of quantitative trait loci for behavioral and neuroendocrine stress response traits in pigs. J Anim Sci. 2002;80:2276-85. https://doi.org/10.1093/ansci/80.9.2276

43. Okamura T, Onodera W, Tayama T, Kadowaki H, Kojima-Shibata C, Suzuki E, et al. A genome-wide scan for quantitative trait loci affecting respiratory disease and immune capacity in Landrace pigs. Anim Genet. 2012;43:721-9. https://doi.org/10.1111/j.1365-

2052.2012.02359.x

44. Murani E, Reyer H, Ponsuksili S, Fritschka S, Wimmers K. A substitution in the ligand binding domain of the porcine glucocorticoid receptor affects activity of the adrenal gland. PLOS ONE. 2012;7:e45518. https://doi.org/10.1371/journal.pone.0045518

45. Görres A, Ponsuksili S, Wimmers K, Muráni E. Analysis of non-synonymous SNPs of the porcine *SERPINA6* gene as potential causal variants for a QTL affecting plasma cortisol levels on SSC7. Anim Genet. 2015;46:239-46. https://doi.org/10.1111/age.12276

46. Uemoto Y, Ichinoseki K, Matsumoto T, Oka N, Takamori H, Kadowaki H, et al. Genome-wide association studies for production, respiratory disease, and immune-related traits in Landrace pigs. Sci Rep. 2021;11:15823. https://doi.org/10.1038/s41598-021-95339-2

47. Guyonnet-Dupérat V, Geverink N, Plastow GS, Evans G, Ousova O, Croisetière C, et al. Functional implication of an Arg307Gly substitution in corticosteroid-binding globulin, a candidate gene for a quantitative trait locus associated with cortisol variability and obesity in pig. Genetics. 2006;173:2143-9. https://doi.org/10.1534/genetics.105.053983

48. Esteve A, Ojeda A, Huang LS, Folch JM, Pérez-Enciso M. Nucleotide variability of the porcine *SERPINA6* gene and the origin of a putative causal mutation associated with meat quality. Anim Genet. 2011;42:235-41. https://doi.org/10.1111/j.1365-2052.2010.02138.x

49. Henley D, Lightman S, Carrell R. Cortisol and CBG — getting cortisol to the right place at the right time. Pharmacol Ther. 2016;166:128-35. https://doi.org/10.1016/j.pharmthera.2016.06.020

50. Boyle LD, Nixon M, Underhill CM, Hill LA, Homer NZM, Andrew R, et al. SUN-221 subclinical alpha-1 antitrypsin deficiency is associated with increased free cortisol fraction in plasma and altered glucocorticoid delivery to tissues. J Endocr Soc. 2020;4:SUN-221. https://doi.org/10.1210/jendso/bvaa046.1849

51. Berridge MJ, Irvine RF. Inositol trisphosphate, a novel second messenger in cellular signal transduction. Nature. 1984;312:315-21. https://doi.org/10.1038/312315a0

52. Seccia TM, Caroccia B, Gomez-Sanchez EP, Gomez-Sanchez CE, Rossi GP. The biology of normal zona glomerulosa and aldosterone-producing adenoma: pathological implications. Endocr Rev. 2018;39:1029-56. https://doi.org/10.1210/er.2018-00060

53. Marzinke MA, Clagett-Dame M. The all-trans retinoic acid (atRA)-regulated gene Calmin (Clmn) regulates cell cycle exit and neurite outgrowth in murine neuroblastoma (Neuro2a) cells. Exp Cell Res. 2012;318:85-93. https://doi.org/10.1016/j.yexcr.2011.10.002

54. Inoue W, Bains JS. Beyond inhibition: GABA synapses tune the neuroendocrine stress axis. BioEssays. 2014;36:561-9. https://doi.org/10.1002/bies.201300178

55. Stimson RH, Andersson J, Andrew R, Redhead DN, Karpe F, Hayes PC, et al. Cortisol release from adipose tissue by 11β-hydroxysteroid dehydrogenase type 1 in humans. Diabetes. 2009;58:46-53. https://doi.org/10.2337/db08-0969

56. Paul S, Jeon WK, Bizon JL, Han JS. Interaction of basal forebrain cholinergic neurons with the glucocorticoid system in stress regulation and cognitive impairment. Front Aging Neurosci. 2015;7:43. https://doi.org/10.3389/fnagi.2015.00043

57. Podder M, Bera S, Naskar S, Sahu D, Mukherjee J, Patra AK. Physiological, blood-biochemical and behavioural changes of Ghoongroo pigs in seasonal heat stress of a hot-humid tropical environment. Int J Biometeorol. 2022;66:1349-64. https://doi.org/10.1007/s00484-022-02281-7