# Uterus-specific transcriptional regulation underlies eggshell pigment production in Japanese quail

Short title: Molecular basis of eggshell pigment production in Japanese quail

Satoshi Ishishita[1], Shumpei Kitahara[2], Mayuko Takahashi[2], Sakura Iwasaki[2], Shoji Tatsumoto[3], Izumi Hara[4], Yoshiki Kaneko[2], Keiji Kinoshita[1], Katsushi Yamaguchi[5], Akihito Harada[6], Yasushige Ohmori[4], Yasuyuki Ohkawa[6], Yasuhiro Go[3,7], Shuji Shigenobu[5], Yoichi Matsuda[1,2,*], Takayuki Suzuki[1,2,*]

[1] Avian Bioscience Research Center, Graduate School of Bioagricultural Sciences, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8601, Japan.

[2] Laboratory of Avian Bioscience, Department of Animal Sciences, Graduate School of Bioagricultural Sciences, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8601, Japan.

[3] Cognitive Genomics Research Group, Exploratory Research Center on Life and Living Systems (ExCELLs), National Institutes of Natural Sciences, Okazaki, Aichi 444-8585, Japan.

[4] Laboratory of Animal Morphology, Department of Animal Sciences, Graduate School of Bioagricultural Sciences, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8601, Japan.

[5] Functional Genomics Facility, National Institute for Basic Biology (NIBB), Okazaki, Aichi 444-8585, Japan.

[6] Division of Transcriptomics, Medical Institute of Bioregulation, Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka, Fukuoka 812-0054, Japan.

[7] Division of Behavioral Development, Department of System Neuroscience, National Institute for Physiological Sciences, Okazaki, Aichi 444-8585, Japan.

*Corresponding authors:

Email address: suzuki.takayuki@j.mbox.nagoya-u.ac.jp (TS), yoimatsu@agr.nagoya-u.ac.jp (YM)

# Abstract

The precursor of heme, protoporphyrin IX (PPIX), accumulates abundantly in the uterus of birds, such as Japanese quail, *Coturnix japonica*, resulting in brown-speckled eggshells. The molecular basis of PPIX production in the uterus remains largely unknown. Here, we investigated the cause of low PPIX production in a classical Japanese quail mutant exhibiting white eggshells by comparing its gene expression in the uterus with that of the wild type using transcriptome analysis and performed genetic linkage mapping to identify the causative genomic region of the white eggshell phenotype. We showed that 11 genes, including the *5-aminolevulinic acid synthase 1* (*ALAS1*) and ferroxidase *hephaestin-like 1* (*HEPHL1*) genes, were specifically upregulated in the wild-type uterus and downregulated in the mutant. We mapped the 172 kb candidate genomic region on chromosome 6, which contains several genes, including a part of the *paired-like homeodomain 3* (*PITX3*), which encodes a transcription factor. *ALAS1*, *HEPHL1*, and *PITX3* were expressed in the apical cells of the luminal epithelium and lamina propria cells of the uterine mucosa of the wild-type quail, and their expression was downregulated in these cells of the mutant quail. Biochemical analysis using uterine homogenates indicated that the restricted availability of 5-aminolevulinic acid is the main cause of low PPIX production. These results suggest that uterus-specific transcriptional regulation of heme-biosynthesis-related genes is an evolutionarily acquired mechanism of eggshell pigment production in Japanese quail.


# Introduction

Avian eggshells display diverse color and color patterns, which may be a result of adaptation to habitats by birds. Eggshell colors have various functions, including avoiding predation

(camouflage) and promoting egg recognition by parents [1–3]. Elucidating the molecular and cellular basis for eggshell pigmentation is essential for understanding the evolutionary processes and mechanisms of eggshell coloration; however, it remains largely unknown [4–6]. Avian eggshell pigments are mainly composed of intermediates and/or catabolites of heme, including protoporphyrin IX (PPIX), an organic compound comprising four pyrrole rings, which is observed as brown-colored pigment on the eggshell [7–9]. PPIX is the final intermediate in the eight-step heme biosynthesis pathway (S1 Fig), wherein the insertion of iron into PPIX generates heme, a component of hemoproteins [10]. PPIX levels are low in most tissues because heme biosynthesis is tightly regulated to avoid the toxic effect of PPIX accumulation [11], and the synthesized PPIX is efficiently converted to heme in the presence of ferrous iron. Exceptionally, PPIX accumulates in the uterus—especially in the caudal part of the oviduct—of bird species that produce brown eggs, indicating that the biosynthesis pathway of heme is specifically regulated in their uteri. The Japanese quail, *Coturnix japonica*, usually lays eggs with brown-speckled color patterns generated by PPIX. In the uterus of this species, a large amount of PPIX accumulates in the apical cells of the mucosal epithelium before its secretion [12, 13]. To elucidate the molecular basis of PPIX production in the uterus of Japanese quail, we focused on a classical quail mutant exhibiting white eggshells owing to the low PPIX production in the uterus [14, 15]. In this mutant quail, the cause of low PPIX production remains unknown. We postulated that the genes involved in PPIX production should be specifically upregulated in the uterus of wild-type quail and downregulated in the mutant quail exhibiting white eggshells. Therefore, in this study, we first compared the gene expression patterns between wild-type and mutant quail uteri by transcriptome analysis. Next, we performed a genetic linkage mapping of the causative genomic region of the white eggshell phenotype and attempted to identify the candidate gene.

# Materials & methods

## General

No statistical methods were used to predetermine the sample size, and the experiments were not randomized. The investigators were not blinded to the allocation during the experiments and outcome assessment.

## Ethics statement

Animal care and all experimental procedures were approved by the Animal Experiment Committee, Graduate School of Bioagricultural Sciences, Nagoya University (approval number: 2017030238). Experiments were conducted according to the Regulations on Animal Experiments in Nagoya University.

## Animals

Commercial female quails were purchased from a local hatchery (Cyubu-kagaku-shizai, Nagoya, Japan), and female quail strains (WE, NIES-L, and NIES-Fr) [16–18] were supplied by the National BioResource Project Chicken/Quail, Nagoya University, Japan. The commercial quail and the NIES-L and NIES-Fr strains exhibited wild-type eggshells, whereas the WE strain exhibited white eggshells. All quails were unvaccinated and were maintained individually with free access to water and a commercially available diet. Most quails were maintained under a 14:10 h light/dark photoperiod; however, some were maintained under a 16:8 h light/dark photoperiod for use with quantitative PCR (qPCR) analysis. Room temperature (RT) was

maintained at approximately 25 °C. Continuously laying quails that were 16–48 weeks old were used for the animal experiments. It should be noted that a superficial layer of pigment is formed on the shell at approximately 21–22 h following the last oviposition in Japanese quail [19–21].

## mRNA sequencing

For mRNA sequencing, females of both the commercial quail (exhibiting wild-type eggshells) and the WE strain were used ($n$ = 6 for both). To obtain mucosal tissues from the isthmuses and uteri just before eggshell pigments were secreted, female quails were sacrificed 19–20 h after the last oviposition. The isthmuses and uteri were isolated from the oviducts and opened with scissors to have the lumen side up. The tissues were rinsed with phosphate-buffered saline (PBS). Subsequently, ice-cold PBS containing 40 mM dithiothreitol was used as a buffer. Mucosal surfaces were scraped from the uteri and isthmuses by rubbing softly with an interdental brush, and the separated tissues were collected by centrifugation at 400 × $g$ and 4 °C. Preliminary histological observation of the uteri and isthmuses, which were treated using this method, confirmed that the epithelia were mainly separated from the mucosal folds (S2 Fig).

For the extraction of total RNA, 50 mg of the tissue pellet was dissolved in 1 mL of TRI Reagent (Molecular Research Center, Cincinnati, OH, USA), and total RNA was purified using the RNeasy Mini Kit (Qiagen, Hilden, Germany). RNA quality was assessed using Bioanalyzer Pico Chips (Agilent Technologies, Santa Clara, CA, USA). RNA samples with integrity numbers greater than 7.6 were used for constructing libraries. We isolated poly(A)+ RNA from 1 ng of total RNA using the NEBNext Poly(A) mRNA Magnetic Isolation Module (NEB, Ipswich, MA, USA) following the manufacturer's instructions. We converted oligo(dT)-selected RNA into a cDNA library for mRNA sequencing using the NEBNext Ultra™ Directional RNA Library Prep Kit for Illumina (NEB) in accordance with the manufacturer's instructions. The library was

sequenced on an Illumina HiSeq 1500 platform by paired-end sequencing (100 bp). We trimmed the adapter sequences from the reads using Trimmomatic software v0.33 [22], and then mapped the reads to the reference genome (Accession code: GCF_001577835.1) using TopHat2 v2.1.0 [23]. Read counts per gene and fragments per kilobase of transcript per million mapped reads (FPKM) were calculated using Cuffdiff v2.2.1 [24] and Cufflinks v2.2.1 [25]. FPKM was then transformed into transcripts per kilobase million. We obtained mRNA sequencing data from various organs and tissues of a 15-month-old male quail from the Sequence Read Archive (Bio Project number: PRJNA296888), and six RNA samples were analyzed for each organ or tissue on an Illumina HiSeq 2500 platform by paired-end sequencing (100 bp). The count data of these organs and tissues were generated using the same procedure as that used for the mRNA sequencing data of the uteri and isthmuses in this study.

Before detecting differentially expressed genes (DEGs), we excluded genes from the dataset whose counts fell below the threshold (1) in any sample. Then, we performed likelihood ratio tests using the glmLRT function in EdgeR v3.20.9 [26]. In the DEG analysis, a full model for the count data was compared to the reduced model, in which the object coefficient was set to 0. *P*-values were adjusted for comparison of the expression level between the uterus and each organ or tissue using the Benjamini–Hochberg method. Genes whose counts were significantly larger (adjusted $p < 0.05$, fold change > 2) in the uterus were classified as genes that were specifically upregulated in the uterus. Hierarchical clustering was performed using the heatmap.2 function in the gplots package of R v.3.4.3, using Ward's method with the Euclidean distance metric.

## Double-digest restriction site-associated DNA (ddRAD) sequencing and association test

To construct the reference family for genetic linkage analysis, one male of the WE strain and one female of the NIES-L strain that exhibited wild-type eggshells were used as parents. The $F_2$ hybrids were generated by crossing an $F_1$ hybrid male with three $F_1$ hybrid females. Genomic DNA was extracted from the red blood cells of the parents and 99 $F_2$ hybrid females using the DNeasy Blood & Tissue Kit (Qiagen). The library was constructed according to a previously described method [27]. After digesting 100 ng of DNA from each sample using EcoRI and MseI (NEB), adapters containing barcode sequences were ligated to the ends of the DNA fragments. DNA fragments ranging from 300 bp to 500 bp were collected using the Pippin Prep (Sage Science, Beverly, MA, USA) and purified using AMPure XP beads (Beckman, Brea, CA, USA). The purified fragments were amplified (six cycles) with PCR primer sets containing index sequences, using Phusion High-Fidelity DNA polymerase (NEB). Library quality was validated using a 2100 Bioanalyzer with the Agilent High-Sensitivity DNA Kit (Agilent Technologies). The library was sequenced on an Illumina HiSeq 1500 platform by paired-end sequencing (100 bp) at the Functional Genomics Facility, National Institute for Basic Biology, Japan.

Sequence data were demultiplexed using their barcodes, cleaned by removing reads of low quality, and trimmed to 95 or 100 bp using Stacks v1.44 [28]. The processed reads were aligned to the reference sequence using Bowtie 2 v2.2.8 [29]. Single nucleotide polymorphisms (SNPs) were assigned by aligning paired-end reads to the reference sequence using Stacks. SNPs with a depth of coverage of less than 8× were counted as missing data. After eliminating SNPs that deviated from the Hardy–Weinberg equilibrium ($p < 0.000001$) or had call rates of less than 80%, an association test was performed using Fisher's exact test implemented in PLINK v1.90 [30]. Sequences of the adapters and primers are shown in S1 Data.

## Whole-genome resequencing

Genomic DNA was extracted from the red blood cells of the parental quails used to produce $F_1$ hybrids in genetic linkage analysis, as described above. Library construction and paired-end sequencing were performed using an Illumina HiSeq X Ten platform (150 bp) at BGI Shenzhen. After trimming the reads using Trimmomatic, the reads were mapped to the reference sequence (Accession code: GCF_001577835.1) using BWA-MEM v0.7.15-r1140 [31]. Subsequently, variants on chromosome 6 were called using Picard v1.119 (http://broadinstitute.github.io/picard/) and the Genome Analysis Toolkit [32] according to a variant call pipeline (https://github.com/gencorefacility/variant-calling-pipeline). We filtered out bases with low Phred-scaled quality scores (less than 30) and reads with low mapping quality (less than 30) using the Genome Analysis Toolkit. Genotypes with individual coverage of less than 10× or more than 70× were filtered out using SnpSift v4.3t [33], and sequence variants that were found in the candidate region were annotated using SnpEff v4.3t [33]. To narrow down the candidate region, we searched for candidate genomic intervals for informative SNP markers that could reveal the parental origins of chromosomal regions of the parent quail. To identify candidate sequence variants around the candidate genomic region, we extracted sequence variants that were homozygous in the WE strain and then filtered out those found in the NIES-L strain. We searched for structural variants around the candidate genomic region using Integrative Genomics Viewer v2.6.3 [34] and Pindel v0.2.5b9 [35]. Nucleotide sequences were determined using an ABI PRISM 3130 DNA Analyzer (Applied Biosystems, Foster City, CA, USA) after cycle-sequencing reactions using the BigDye Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems). The primer sequences are listed in S1 Table.

## qPCR analysis

To compare gene expression levels of target genes between wild-type and mutant uteri, qPCR analysis was performed using RNA samples, which were used for mRNA sequencing ($n$ = 5 for both the commercial quail and WE strain). qPCR analysis was also performed using RNA samples extracted from the uterine tissues of female quail that were not used for mRNA sequencing ($n$ = 5 for both the commercial quail and WE strain). Tissue collection and RNA extraction for these additional 10 RNA samples were conducted using the same method as that used for mRNA sequencing; however, quails were maintained under a 16:8 h light/dark photoperiod, and uterine tissues were collected 17 h after oviposition [36]. Total RNA (500 ng) was reverse transcribed into cDNA in a 10 µL reaction mixture using oligo(dT) and the ReverTra Ace qPCR RT Master Mix with a gDNA Remover (Toyobo, Osaka, Japan). PCR amplification (0.1 µL cDNA in a 10 µL reaction mixture) was performed using the Thunderbird SYBR qPCR Mix (Toyobo) and the StepOnePlus™ Real-Time PCR System (Applied Biosystems, Carlsbad, CA, USA). Melt curve analysis was performed immediately after amplification. *ACTB* was used as an endogenous control. Each experiment was performed in triplicate. For all primer sets, the amplification efficiency was 92%–112%, and the correlation coefficient was greater than 0.97. The primer sequences and PCR reaction conditions are listed in S1 Table.

## Histology

Uteri were collected from female quails ($n$ = 6 for both the commercial quail and WE strain) 19□20 h after oviposition and fixed overnight in Bouin's solution. The fixed tissues were dehydrated in a series of graded ethanol, immersed in a 1:1 ethanol/xylene solution, and subsequently embedded in paraffin. The tissues were sectioned at a thickness of 4 µm and mounted onto ovalbumin-coated glass slides. After deparaffinization, the sections were stained with Mayer's hematoxylin and eosin or left unstained.

## Transmission electron microscopy

Uterine samples for transmission electron microscopy (TEM) were collected from female quails (*n* = 1 for both the commercial quail and WE strain) 19–20 h after oviposition, fixed overnight in 2% glutaraldehyde (in 0.1 M phosphate buffer) at 4 °C, rinsed with 0.1 M phosphate buffer overnight at 4 °C, and then post-fixed in 2% osmium tetroxide (in deionized water [DW]) for 2 h at 4 °C. The fixed specimens were dehydrated in a series of graded ethanol (50%, 70%, 90%, 100%, 100%, and 100%) for 15 min at RT, immersed in propylene oxide for 30 min and then in a propylene oxide and epoxy resin mixture for 2 h at RT, and finally embedded in gelatin capsules with epoxy resin for two days at 60 °C. Ultrathin sections of 80 nm thickness were cut using an ultramicrotome with diamond knives and mounted on copper grids with a mesh size of 200 μm. The sections were stained with 2% uranyl acetate (in DW) for 15 min and lead staining solution for 5 min at RT. TEM was performed using an H-7600 TEM (Hitachi, Tokyo, Japan) operated at an accelerating voltage of 100 kV.

## *In situ* hybridization

Tissues were collected from female quails (*n* = 2 for both the commercial quail and the WE strain) at 19□20 h after ovulation, fixed with 4% PFA overnight, and dehydrated sequentially with 25% MeOH/PBT (PBS, 0.1% Tween 20), 50% MeOH/PBT, 75% MeOH/PBT, and 100% MeOH for 5 min each. After dehydration, the tissues were immersed in 100% EtOH and 100% xylene for 1 h each. Subsequently, tissues were embedded in paraffin wax 60 (Sakura, Tokyo, Japan) and 10 μm sections were prepared using a Leica RM2125 microtome (Leica, Wetzlar, Germany). Sections were dried on glass slides overnight and then hydrated sequentially with 100% xylene, 100% EtOH, 90% EtOH, 70% EtOH, and PBT for 5 min each. After the hydration process, the slides were immersed in 1 μg/mL of proteinase K in PBT for 7 min at 37 °C. Slides

were washed with PBT three times for 5 min and then fixed with 4% PFA for 20 min. The sections were hybridized with 1 μg/mL of DIG-labeled RNA in hybridization buffer (50% formamide, 5× saline-sodium citrate [SSC], 1 mg/mL total RNA, 100 μg/mL heparin, 0.1% Tween 20, 0.1% 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate [CHAPS], and 10 mM ethylenediaminetetraacetic acid [EDTA]) overnight at 68 °C. The slides were washed twice with 50% formamide in 2× SSC for 30 min. After cooling, the slides were incubated with anti-DIG antibody (1:1000; Roche, Penzberg, Germany) overnight. After the slides were washed with Tris-buffered saline with 0.1% Tween 20 three times for 5 min, the color was developed in AP buffer (100 mM Tris-HCl, 100 mM NaCl, 10 mM $MgCl_2$, NBT 4.5 μL/mL [Roche], BCIP 3.5 μL/mL [Roche], pH 9.5).

For the synthesis of RNA probes, partial sequences of cDNAs of the target genes were obtained by PCR amplification (see S1 Table for primer sequences and S2 Data for partial sequences of cDNAs) and then inserted into the pGEM-T Easy Vector System (Promega, Madison, WI, USA). RNA probes were synthesized *in vitro* using SP6 RNA polymerase and plasmid DNA digested with SphI (*ALAS1*) and ApaI (*HEPHL1* and *PITX3*).

## GO term enrichment analysis

Gene Ontology (GO) term enrichment analyses were performed using the overrepresentation test (released 2021-02-24) of the PANTHER (Protein ANalysis THrough Evolutionary Relationships) Classification System (annotation version 16.0 released 2020-12-01) [37]. The *Gallus gallus* database was used as the reference. The PANTHER GO-Slim Biological Process was used as the annotation dataset. *p*-values of Fisher's exact test were adjusted using the Benjamini–Hochberg method. GO terms were considered significantly enriched if they had an FDR of < 0.05.

## Biochemical assay for the PPIX-forming ability of the uterus

Uterine tissues were collected from female quails (*n* = 1 for both the commercial quail and WE strain) 15–20 h after oviposition and homogenized in ice-cold PBS containing 0.05% Triton-X100. After centrifugation at 20,000 × g and 4 °C for 5 min, the supernatants of each homogenate were incubated in the presence of 0, 1, or 2 mM 5-aminolevulinic acid (ALA) (Wako, Osaka, Japan) in 96-well plastic plates at 39 °C for 20 h (three reactions for each concentration of ALA were performed in 110 µL homogenates). The absorbance was measured at 405 nm using an ARVOx4 2030 Multilabel Reader (PerkinElmer, Waltham, MA, USA) because PPIX exhibits maximum light absorbance at 409 nm. Supernatants incubated without ALA supplementation were used as controls. The total protein concentrations of supernatants were determined using the Bradford method, and the mean absorbance per 1 mg/mL total protein concentration was calculated for each sample at each ALA concentration.

## Statistical analysis

R v.3.4.3 (R Core Team) and MS Excel (Microsoft Corp., Redmond, WA, USA) were used for statistical analyses. For the comparison of expression levels by qPCR analysis, Welch's two-tailed *t*-test was performed using delta *Ct* values. Welch's two-tailed *t*-test was also used to test the differences in absorbance (at 405 nm) of uterine homogenates between the wild-type and mutant quails. The alpha level of these statistical tests was set at 0.05.

## Data availability

The Illumina data generated in this study were deposited in the DDBJ Sequence Read Archive: ddRAD sequencing (Accession codes: DRA005969), mRNA sequencing (DRA006948), and whole-genome resequencing (DRA007010). All raw data are included in the Supporting

Information or deposited on GitHub (https://github.com/isst001/eggshell_color_gene). Other miscellaneous information is available from the corresponding author upon request.

# Results

## Histological analysis of the uterine mucosae of wild-type and mutant quails

We performed gross and histological examination of the uterine mucosae of wild-type and mutant quails, which lay eggs with normal and white eggshells, respectively (Fig 1A; S3 Fig). The uterine mucosa of the wild-type quail exhibited a dark brown color, and that of the mutant quail was pale brown (Fig 1B). Histological examination revealed that brown pigment granules were accumulated in the apical cells of the mucosal epithelium of the uterus in the wild-type quail before anticipated secretion (19–20 h after oviposition), whereas few brown pigment granules were observed for the mutant quail (Fig 1C, D). Instead, pink-colored, eosinophilic granules accumulated in the cytoplasm of apical cells in the mutant, suggesting that transport vesicles containing few or no PPIX are formed even in the mutant cells. Electron microscopic observation of the uterine mucosa sections also showed that transport vesicles were accumulated in the apical cells of both the wild type and mutant (S4 Fig).

**Fig 1. The *white eggshell* phenotype.** (A) Wild-type and mutant white eggs. (B) Wild-type and mutant uteri were opened so that the lumen side could be seen. (C) Schematic diagram of the uterine mucosa. Epithelium of the uterine mucosa consists of apical cells and basal cells, which are alternatingly arranged. Pigment granules containing PPIX molecules accumulate in the

luminal side of the cytoplasm of apical cells before the secretion. The lamina propria is located at the inner side of the uterine mucosa. (D) Sections of uterine mucosae obtained from wild-type and mutant quail just before pigment secretion (left and middle panels, HE staining; right panels, no staining). Ep and LP indicate epithelium and lamina propria, respectively. Brown pigments accumulated in the epithelium of the uterine mucosa of the wild-type quail, but not in that of the mutant quail (left panels). Higher-magnification images of mucosal folds (middle panels) indicate that brown pigment granules accumulated in the wild-type apical cells, but not in the mutant apical cells (black arrowheads); however, pink-colored, eosinophilic granules accumulated in the luminal side of the mutant apical cells. Eosinophilic granules also accumulated more centrally in both wild-type and mutant apical cells (white arrowheads). Sections without staining clearly demonstrate that brown pigments were accumulated in the luminal side of the apical cells (arrows) of the wild type, but not in those of the mutant. Scale bars: 1 cm in (A, B); 100 μm in (D).

## Characterization of the mutant uterus by gene expression analysis

We performed mRNA sequencing analysis using the uterine tissues of wild-type and mutant quail. The average number of mapped read pairs is shown in S1 Data. Comparison of gene expression between the wild type and mutant showed that 261 out of approximately 17,000 expressed genes were differentially expressed (adjusted $p < 0.05$, fold change < 0.5 or > 2) (Fig 2A; S3 Data). These DEGs consisted of 148 and 113 genes that showed upregulated and downregulated expression, respectively, in the mutant (Fig 2A; S3 Data). We then focused on changes in the expression levels of genes that are known to be involved in heme and iron metabolism [11] (S2 Table). Therefore, *5-aminolevulinic acid synthase 1* (*ALAS1*) [38], *hephaestin-like 1* (*HEPHL1*)

[39], and *transferrin receptor protein 1* (*TFRC*) [40] were downregulated in the mutant uterus (Fig 2A; S2 Table). The expression levels of these genes in the mutant decreased to 20% (*ALAS1*), 0.1% (*HEPHL1*), and 41% (*TFRC*) of those in the wild type (S2 Table). To identify genes that are specifically upregulated in the uterus, we compared gene expression between the uterus and the other six organs and two tissues. The results showed that 382 genes were specifically upregulated in the uterus (Fig 2B; S4 Data). Eleven of the DEGs, including *ALAS1* and *HEPHL1*, were also included in the uterus-specific upregulated genes (Fig 2C, S3 and S4 Data). The expression levels of *ALAS1* and *HEPHL1* were 4–22 times and 171–1205 times higher, respectively, in the uterus than those in other organs and tissues (Fig 2D, E; S3 Table). The results of GO term enrichment analyses of the uterus-specific upregulated genes are shown in S4 Table. It is of interest to note that many uterus-specific upregulated genes were related to GO "biological process" (GO:BP) terms, such as cellular ion homeostasis and ion transmembrane transport. This may be related to eggshell calcification in the uterus. Significant downregulation of *ALAS1* and *HEPHL1* in the mutant uterus was confirmed by qPCR analysis (Welch's two-sided *t*-test, $p < 0.01$) (S5 Fig; S5 Data).

**Fig 2. Genes that are specifically upregulated in the wild-type uterus and downregulated in the mutant uterus.** (A) Visualization of the result of differential gene expression analysis using wild-type and mutant uterine tissues. Dots indicate the expression levels of genes. Differentially expressed genes (DEGs) [adjusted $p < 0.05$, fold change ($\log_2$ scale) > 1 or < −1] are highlighted. Genes that were downregulated in the mutant and specifically upregulated in the wild type and *TFRC* are labeled. (B) Gene expression levels of the 382 genes whose expression levels were higher in the uterus than in other six organs and two tissues, visualized by log-transformed

transcripts per kilobase million (TPM) values. (C) The number of downregulated genes in the mutant uterus, uterus-specific upregulated genes in the wild type, and genes showing both expression patterns. (D, E) Comparison of expression levels of *ALAS1* (D) and *HEPHL1* (E) in the wild-type uterus with those in other wild-type organs and tissues and the mutant uterus, using mRNA sequencing data. Each bar indicates the relative value of the expression level of the gene in the wild-type uterus when the expression level in each organ or tissue and the mutant uterus is defined as 1. Values in (D, E) are not indicated by log scale.

## Genetic linkage mapping of the candidate genomic region responsible for the white eggshell phenotype

The white eggshell phenotype of the mutant quail is inherited in an autosomal recessive manner [14]. However, the causative mutation, which we refer to as *white eggshell* (*we*), remains unknown. We performed chromosomal mapping of the mutation by genetic linkage analysis using 99 $F_2$ progeny obtained by mating three pairs of $F_1$ progeny derived from crossing one mutant male (the WE strain) with one wild-type female (the NIES-L strain). SNP markers were obtained by ddRAD sequencing of the genomic DNA of the 99 $F_2$ progeny and their parents. The average number of aligned read pairs and phenotype data are shown in S1 Data. An association test using a total of 15,200 SNP markers revealed a significant association of 235 SNP markers on chromosome 6 (Fisher's exact test, adjusted $p < 0.01$) with the white eggshell phenotype (Fig 3A; S6 Data). The marker SNP201288 displayed the highest association (Fig 3A) and logarithm of the odds (LOD) score and the lowest recombination frequency (Fig 3B, C; S6 Data); its genotypes in the $F_2$ progeny were fully concordant with the eggshell color phenotype (Fig 3D). Two SNP markers, SNP201255 (which was located adjacent to SNP201288) and SNP202246,

were separated by a large genomic interval (2,230 kb) (Fig 3E). SNP202246 was located far (2,179 kb) from SNP201288 and displayed discordance between genotypes and phenotypes in eight $F_2$ progeny (Fig 3D).

**Fig 3. The candidate genomic region of the *white eggshell* (*we*) phenotype is mapped to chromosome 6.** (A) Association test using 15,200 single nucleotide polymorphisms (SNPs) markers. *x*-axis, genomic coordinates of SNPs; *y*-axis, negative logarithm of *p*-values from Fisher's exact test. Blue line indicates the Bonferroni-corrected *p*-value of 0.01. SNP201288 displayed the strongest association with the phenotype. (B) Logarithm of the odds (LOD) scores between the *we* locus and all SNP markers on chromosome 6. SNP201288 and flanking SNPs (SNP201255 and SNP202246) display high LOD scores. (C) LOD scores and recombination frequencies of three SNPs near the *we* locus. (D) Genotypes of SNPs near the *we* locus in 99 $F_2$ individuals. SNP201288 shows a complete concordance between genotypes and eggshell colors in $F_2$ progeny. Discordance between genotype and eggshell color was observed for SNP201255 in $F_2$ individual #3398 and for SNP202246 in eight $F_2$ progeny shown by brackets, including $F_2$ individual #4040. (E) Genetic and physical maps of SNP markers around the *we* locus. Genetic distances between the *we* locus and flanking markers are indicated by parentheses. Positions of SNP202482 and SNP202389 are inverted between the genetic and physical maps, which is possibly due to an assembly error in the draft genome. (F) Genotypes of Chr6-we1 and Chr-we2 are concordant and discordant, respectively, with the eggshell color, in $F_2$ individual #4040. Recombination in the interval between SNP markers is indicated by "×."

We then carried out whole-genome resequencing of the parental quail, which was used for genetic linkage analysis, to obtain sequence variant information on chromosome 6. Sequence data

of 33 and 45 GB were obtained for the father (WE strain) and mother quail (NIES-L strain), respectively (S1 Data). The average depth of coverage of chromosome 6 was approximately 16 and 23 for the father and mother quails, respectively (S1 Data). Structural variants, including copy number variation, large deletions or insertions, and inversions, could not be found around the candidate genomic region. Almost all informative SNP markers were detected in the vicinity of SNP201288 and SNP202246 (S7 Data). Among them, the genotypes of Chr6-we1 and Chr6-we2 were concordant and discordant, respectively, with the eggshell color phenotype of the $F_2$ progeny (#4040) (Figs 3E, 4A; S6 Fig; S5 Table). Furthermore, Chr6-we-a, which was detected in a 2,030 kb genomic region between Chr6-we1 and Chr6-we2, displayed a concordance between the genotypes and phenotypes of the eight $F_2$ progeny (S6 Fig; S5 Table). However, the genotype of chr6-we-a in the mutant was also observed in the quail strain NIES-Fr, which exhibited wild-type eggshells (S5 Table). Therefore, we eliminated the 2,058 kb region between Chr6-we1 and SNP202246 from the 2,230 kb candidate region. By searching sequence variants (SNPs and insertion–deletion mutations [indels]) that are homozygous in the father quail (mutant) but not found in the mother quail (wild type), 957 candidate sequence variants were detected in the remaining 172 kb region between SNP201255 and Chr6-we1 (Fig 4A, B; Table S6; S7 Data).

**Fig 4. Candidate sequence variants and expression analysis of genes around the 172 kb candidate region.** (A) Number of sequence variants that were homozygous in the father quail (WE strain) but not in the mother quail (NIES-L strain) around the *we* locus. White arrowheads indicate the locations of single nucleotide polymorphism (SNP) markers. (B) The frequencies of SNPs and insertion–deletion mutations (indels) within the 172 kb candidate region. (C) Eight genes and a part of *PITX3* that are harbored in the 172 kb candidate region (green arrows). (D)

Relative expression levels of 15 genes (located around the candidate genomic region) in the mutant uteri against those in wild-type uteri. NA denotes genes whose expression was not detected in the wild-type uterus. Expression levels of six genes (enclosed in squares) substantially decreased in the mutant (less than 60%–70% of those in the wild type). (E) The graphs show the results of qPCR analysis of *PITX3* using RNA samples ($n$ = 5 for both wild-type and mutant quail) that were not used for mRNA sequencing. The expression level in the wild-type uterus is shown as a relative value against that in the mutant uterus defined as 1. The expression level in each sample is shown by a circle. Double asterisks denote $p < 0.01$.

The 172 kb candidate region harbored seven protein-coding genes (*PITX3*, *GBF1*, *TMEM150A*, *NFKB2*, *PSD*, *FBXL15*, and *CUEDC2*) and two non-coding genes (*LOC107315847* and *LOC107315848*) (Fig 4C). Six missense variants were included in the 957 sequence variants (S6 Table; sheet 5 in S7 Data); however, none were predicted to have adverse effects on the functions of the protein-coding genes, including *GBF1*, *NFKB2*, *PSD*, and *CUEDC2* (S7 Table). Furthermore, although it is difficult to precisely estimate the effect of sequence variants, the remaining sequence variants were not considered to be related to functions (sheet 6 in S7 Data). These results raised the possibility that the mutation affects the expression level of the gene, rather than its function. We predicted that the recessive *we* mutation suppresses the expression of the responsible gene, because mutations that enhance gene expression levels are most likely to be dominant and not recessive. Therefore, using mRNA sequencing data, we searched for genes on chromosome 6 that were downregulated in the mutant uteri. Two genes, *PRKG1* and *C6H10orf71*, were significantly downregulated in the mutant (adjusted $p < 0.05$, fold change < 1/2) (S8 Table; sheets 3 and 4 in S3 Data). However, they do not seem to be causative genes because both are located far (>~1 Mb) from the candidate region. Thus, we focused on the genes located within the

candidate region and in the region up to 100 kb on either side (nucleotide positions 14.23–14.61 Mb), of which six genes (*PIK3AP1*, *LOC107315928* [*2-hydroxyacylsphingosine 1-beta-galactosyltransferase-like*], *PITX3*, *TMEM150A*, *PSD*, and *CUEDC2*), were substantially downregulated in the mutant (their expression was less than approximately 70% of that in the wild-type) (Fig 4D). *PITX3* encodes a transcription factor involved in the normal development, differentiation, and maintenance of cells in other tissues [41–44]. The other five genes encode proteins that may function in molecular and cellular processes other than transcriptional regulation, such as galactosylceramide biosynthesis and signal transduction [45–49]. Because *PITX3* is likely involved in a tissue-specific transcriptional regulation process, we expected that this gene could be a strong candidate. qPCR analysis using mRNA samples that were used for mRNA sequencing also showed non-significant but substantial downregulation of *PITX3* in the mutant uterus (S7 Fig; S5 Data). We observed significant downregulation of *PITX3* in the mutant uterus by qPCR analysis using additional RNA samples (Welch's two-sided *t*-test, $p < 0.01$) (Fig 4E; S5 Data). The difference between the qPCR analyses may be due to the unstable expression of *PITX3* for some reason. However, these results consistently indicate that *PITX3* is downregulated in the mutant.

Subsequently, we investigated the distribution of the transcripts of *PITX3*, *ALAS1*, and *HEPHL1* in the uterine mucosa of wild-type and mutant quail. These three genes were expressed in apical cells of the mucosal epithelium and cells of the lamina propria (LP), the latter of which are located adjacent to the basal layer of the mucosal epithelium, in the wild type (upper panels of Fig 5). In contrast, hybridization signals were very weak for all three genes in both the apical cells and LP cells in the mutant (lower panels of Fig 5). Notably, the hybridization signals of *ALAS1* were more intense in the LP cells than in the apical cells of the wild-type quail (Fig 5A).

**Fig 5 Distribution of *PITX3*, *ALAS1*, and *HEPHL1* transcripts in uterine mucosal tissues.** (A–C) Light micrographs of sections of mucosal folds of wild-type uterus (upper panels) and mutant (lower panels) uterus. *PITX*, *ALAS1*, and *HEPHL1* were expressed in the apical cells on the luminal side of the epithelia (Ep) of the wild type, and their expression was observed as bluish hybridization signals (upper panels, black arrowheads). Hybridization signals were also observed in the cells of the lamina propria (LP) (upper panels, white arrowheads). *ALAS1* was highly expressed in the cells of the LP. The hybridization signals of three genes in the luminal side of the Ep (lower panels, black arrowheads) and LP cells (lower panels, white arrowheads) of the mutant were weak compared to those of the wild type.

Finally, we examined whether the PPIX-forming ability in uterine homogenates of the mutant quail could be recovered by supplementation with a large amount of ALA. The results showed that homogenates from mutant uteri could produce the same amount of PPIX as those from wild-type uteri in both 1 mM and 2 mM concentrations of ALA (Welch's two-sided *t*-test, $P \geq 0.05$) (Fig 6A; S8 Data), indicating that the restricted availability of ALA or its precursors is the main cause of the low PPIX production.

**Fig 6 Comparison of the protoporphyrin IX (PPIX)-forming ability of uterine homogenates between the wild-type and mutant quail and a hypothetical model of PPIX production in Japanese quail uterus.** (A) Absorbance of uterine homogenates incubated in the presence of different concentrations of 5-aminolevulinic acid (ALA). Bars indicate the mean absorbance at 405 nm per 1 mg/mL total protein. Dots indicate the absorbance of each sample. Error bars

indicate standard error of the mean. (B) Hypothetical model of the uterus-specific regulation of heme biosynthesis and underlying transcriptional regulation. Transcription factors may enhance *ALAS1* and *HEPHL1* expression in the uterus. The upregulated expression of *ALAS1* may enhance ALA synthesis, which may allow uterine cells to synthesize an excessive amount of PPIX, resulting in the accumulation of PPIX by exceeding the conversion of PPIX to heme. In addition, HEPHL1, a predicted ferroxidase, may suppress the conversion of PPIX to heme by decreasing the intracellular or intramitochondrial ferrous ions, which may accelerate the accumulation of PPIX. (C) Schematic diagram of PPIX accumulation in the uterine cells.

# Discussion

Eggshell PPIX is synthesized in the uterus and accumulates in the mucosal epithelium in birds; however, where and how PPIX is synthesized in the uterus remains largely unknown. In this study, we investigated the cause of low PPIX production in a Japanese quail mutant that lays white eggshells by transcriptome and genetic linkage analyses.

Transcriptome analysis revealed that the expression levels of 11 genes were higher in the uterus than those in other organs and tissues in the wild-type quail and were downregulated in the mutant uterus. Therefore, we expected that these genes would be involved in the uterus-specific regulation of heme synthesis, which should be disrupted in the mutant quail. Among these 11 genes, *ALAS1* and *HEPHL1*, which are well-known genes related to heme and iron metabolism, were included. *ALAS1* encodes the rate-limiting enzyme of heme biosynthesis, which catalyzes ALA synthesis as the first step in heme biosynthesis. Our results indicate that high *ALAS1* expression in the uterus is required for the increased synthesis of PPIX [38,50]. The PPIX-forming ability was comparable between the wild-type and mutant uteri in the presence of large

amounts of ALA, which suggests that the restricted availability of ALA, owing to the downregulation of *ALAS1*, is the main cause of low PPIX production in the mutant quail. However, a previous biochemical study showed that the ability to synthesize PPIX was lower in mutant homogenates than in wild-type homogenates [15]. The difference in results between this study and previous studies may be due to the differences in the genetic background of the mutant and/or wild-type quail used in the experiments. A previous biochemical study using chickens also suggested that limited availability of ALA or its precursor in the uterus could be the cause of white eggshells [51]. *ALAS1* expression in the uterus was lower in a chicken strain that laid lighter brown eggshells than in one with darker brown eggshells [52]. These results and our present data suggest that low ALA synthesis, due to the downregulation of *ALAS1*, may be a genetic cause of white eggshells in both quail and chickens. However, we cannot exclude the possibility that there are additional factors that affect PPIX production other than an insufficient amount of ALA *in vivo*. *HEPHL1* encodes a copper-dependent ferroxidase that converts ferrous iron to ferric iron [39]. Thus, HEPHL1 is likely involved in PPIX accumulation by inhibiting the conversion of PPIX to heme through its ferroxidase activity. Therefore, the downregulation of *HEPHL1* is another possible cause of low PPIX production in the mutant quail. Although the remaining nine genes, except for *ALAS1* and *HEPHL1*, have not been shown to be related to heme and iron metabolism (S3 Table), these genes are also important in elucidating the molecular basis of PPIX production in the uterus.

This study demonstrated that *ALAS1* is highly expressed in the apical cells of the epithelium and the lamina propria (LP) cells of the uterine mucosa, indicating the possible involvement of these cells in PPIX synthesis through ALA synthesis. This concurs with the suggestion that the amount of PPIX produced is too large to be produced only in the apical cells

of the uterine mucosa [53]. However, gene expression analysis of each cell type in the uterine mucosa should be conducted to clarify the role of the LP and apical cells in PPIX production.

The candidate locus responsible for white eggshells was localized to a 172 kb region on chromosome 6, which contains nine genes, by genetic linkage analysis. Although many candidate mutations were found within this region, there were no sequence variants that could be expected to influence gene function. Several genes that are located up to 100 kb on either side of this 172 kb region, including *PITX3*, were substantially downregulated in the mutant. Because *PITX3* is likely involved in a tissue-specific transcriptional regulation process [41–44], it is plausible that *ALAS1* expression is regulated by PITX3. Therefore, the downregulation of *PITX3* is one possible cause of the low PPIX production in the mutant. Further investigations, such as identification of the causative mutation and transgenic rescue experiments, are required to verify this possibility. The causative mutation may affect *PITX3* expression via a change in the histone modification of enhancer chromatins. Epigenetic approaches, such as chromatin immunoprecipitation sequencing with anti-H3K27ac antibody [54] or assay for transposase-accessible chromatin using sequencing (ATAC-seq) [55], to compare active regulatory regions in uterine cells between the wild type and mutant would be useful for identifying a candidate as the uterus-specific enhancer of *PITX3*.

Based on our findings, we propose a hypothetical model of the uterus-specific regulation of heme biosynthesis and underlying transcriptional regulation, as shown in Fig 6B and C. Heme biosynthesis actively progresses with abundant ALA synthesis by the upregulation of *ALAS1*. In parallel with this, the conversion of PPIX to heme may be inhibited through a ferroxidase reaction mediated by the predicted ferroxidase HEPHL1. This inhibition of the final step of heme biosynthesis results in the accumulation of a large amount of PPIX in the uterus. Transcription

factors, such as PITX3, may be involved in the transcriptional regulation of these two genes. This model requires the production of a large amount of glycine and succinyl-CoA, which are precursors of ALA in uterine cells. However, to verify our hypothesis, the downstream targets of PITX3 need to be identified.

Many quantitative trait loci that control eggshell color in chickens have been reported [5]. In addition, correlations between the depth of brown eggshell color and the expression levels of the genes involved in heme biosynthesis have been reported in chickens [52]. However, no causative genes for eggshell brownness/whiteness in chickens have been identified. *SLCO1B3*, encoding an anion transporter, is the first and only gene that has been identified as a causative gene for eggshell color variation; the overexpression of this gene due to retrovirus insertion leads to a blue egg phenotype in chickens [56]. The findings of this study may help to identify the causative gene of white eggshells, contributing to a deeper understanding of the genetic basis of eggshell color variation.

# Conclusions

In this study, we postulated that uterus-specific transcriptional regulation plays an important role in the uterus-specific regulation of the heme biosynthesis pathway and that this specific transcriptional regulation is disrupted in the white eggshell mutant quail. Based on this postulation, we performed transcriptome analysis of wild-type and mutant uteri and genetic linkage mapping of the mutation responsible for white eggshells. The results showed that 11 genes were specifically upregulated in the uterus and downregulated in the mutant uterus. Among them, *ALAS1* and *HEPHL1* were highly expressed in the apical cells of the mucosal epithelium

and LP cells of the wild-type uterus, but downregulated in the mutant. The results from our biochemical analysis indicated that the insufficiency of ALA was mainly responsible for the low PPIX production in the mutant. Furthermore, we showed that several genes in the candidate genomic region, including the gene encoding the transcription factor PITX3, were downregulated in the mutant uterus. These results collectively suggest that tissue-specific transcriptional regulation underlies PPIX production in the uterus. Although further investigation is required to verify our model, our findings would contribute to uncovering the physiological mechanisms of PPIX production in the bird uterus and the genetic basis of eggshell color variation.

# Acknowledgments

We would like to thank Hisayo Asao and Asaka Akita (NIBB), Kayako Fukuyama (Kyushu University), Makoto Mizutani, Mikiharu Nakano, and Mitsuo Nunome (Nagoya University) for their technical assistance. The quail strains used in this study were provided by the National BioResource Project (NBRP) Chicken/Quail, supported by the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) and Japan Agency for Medical Research and Development (AMED), Japan. ddRAD sequencing was carried out under the NIBB Cooperative Research Program for Next-generation DNA Sequencing (15-833). Computations were performed on the NIG supercomputer at the National Institute of Genetics.

# References

1. Birkhead T. The most perfect thing: Inside. New York, NY: Bloomsbury Publishing; 2016.
2. Hauber, M. E. The book of eggs: A life-size guide to the eggs of six hundred of the world's bird species. Chicago, IL: University of Chicago Press; 2014.

3. Kilner RM. The evolution of egg colour and patterning in birds. Biolog. Rev. 2006;81: 383–406. Doi:10.1017/S1464793106007044

4. Samiullah S, Roberts JR, Chousalkar K. Eggshell color in brown-egg laying hens — a review. Poult Sci. 2015;94: 2566–2575. Doi:10.3382/ps/pev202

5. Wilson PB. Recent advances in avian egg science: A review. Poult Sci. 2017;96: 3747–3754. Doi:10.3382/ps/pex187

6. Wragg D, Mwacharo JM, Alcalde JA, Wang C, Han J-L, Gongora J, et al. Endogenous retrovirus EAV-HP linked to blue egg phenotype in Mapuche fowl. PloS One. 2013;8: e71393. Doi:10.1371/journal.pone.0071393

7. Kennedy GY, Vevers HG. A survey of avian eggshell pigments. Comp. Biochem. Physiol. Part B Comp. Biochem. 1976;55: 117–123. Doi:10.1016/0305-0491(76)90183-8

8. Gorchein A, Lim CK, Cassey P. Extraction and analysis of colourful eggshell pigments using HPLC and HPLC/electrospray ionization tandem mass spectrometry. Biomed. Chromatogr. 2009;23: 602–606. Doi:10.1002/bmc.1158

9. Wang XT, Zhao CJ, Li JY, Xu GY, Lian LS, Wu CX, et al. Comparison of the total amount of eggshell pigments in Dongxiang brown-shelled eggs and Dongxiang blue-shelled eggs. Poult Sci. 2009;88: 1735–1739. Doi:10.3382/ps.2008-00434

10. Heinemann IU, Jahn M, Jahn D. The biochemistry of heme biosynthesis. Arch. Biochem. Biophys. 2008;474: 238–251. Doi:10.1016/j.abb.2008.02.015

11. Sachar M, Anderson KE, Ma X. Protoporphyrin IX: the good, the bad, and the ugly. J Pharmacol Exp Ther. 2016;356: 267–275. Doi:10.1124/jpet.115.228130

12. Poole HK. A microscopic study of uterine eggshell pigment in Japanese quail. J Hered. 1967;58: 200–203. Doi:10.1093/oxfordjournals.jhered.a107586

13. Tamura T, Fujii S. Histological observations on the quail oviduct: On the secretions in the mucous epithelium of the uterus. J. Fac. Fish. Anim. Husb. Hiroshima Univ. 1966;6: 373–393. Doi:10.15027/40496

14. Poole HK. Egg shell pigmentation in Japanese quail: genetic control of the white egg trait. J Hered. 1964;55: 136–138. Doi:10.1093/oxfordjournals.jhered.a107312

15. Poole HK. Relative oöporphyrin content and porphyrin forming capacity of wild-type and white-egg Japanese quail uterine tissue. Proc Soc Exp Biol Med. 1966;122: 596–598.

16. Takahashi S, Inooka S, Mizuma Y. Selective breeding for high and low antibody responses to inactivated Newcastle disease virus in Japanese quails. Poult Sci. 1984;63: 595–599.

17. Mizutani M. The Japanese quail. In: Chang HL, Huang YC, editors. The relationship between indigenous animals and humans in APEC region. Taiwan: The Chinese Society of Animal Science; 2003. Pp. 143–163.

18. Nunome M, Nakano M, Tadano R, Kawahara-Miki R, Kono T, Takahashi S, et al. Genetic divergence in domestic Japanese quail inferred from mitochondrial DNA D-Loop and microsatellite markers. PloS One. 2017;12: e0169978. Doi:10.1371/journal.pone.0169978

19. Woodard AE, Mather FB. The timing of ovulation, movement of the ovum through the oviduct, pigmentation and shell deposition in Japanese quail (*Coturnix coturnix japonica*). Poult Sci. 1964;43: 1427–1432. Doi:10.3382/ps.0431427

20. Poole HK. Spectrophotometric identification of eggshell pigments and timing of superficial pigment deposition in the Japanese quail. Proc Soc Exp Biol Med. 1965;119: 547–551. Doi:10.3181/00379727-119-30234

21.Tanaka K, Imai T, Koga O. Superficial pigmentation of egg shell in Japanese quail, *Coturnix coturnix japonica*. J Poult Sci. 1977;14: 229–231. Doi:10.2141/jpsa.14.229

22. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30: 2114–2120. Doi:10.1093/bioinformatics/btu170

23. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 2013;14: R36. Doi:10.1186/gb-2013-14-4-r36

24. Trapnell C, Hendrickson DG, Sauvageau M, Goff L, Rinn JL, Pachter L. Differential analysis of gene regulation at transcript resolution with RNA-seq. Nat Biotechnol. 2013;31: 46–53. Doi:10.1038/nbt.2450

25. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, Baren MJ van, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol. 2010;28: 511–515. Doi:10.1038/nbt.1621

26. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26: 139–140. Doi:10.1093/bioinformatics/btp616

27. Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE. Double digest RADseq: an inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. PloS One. 2012;7: e37135. Doi:10.1371/journal.pone.0037135

28. Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA. Stacks: an analysis tool set for population genomics. Mol Ecol. 2013;22: 3124–3140. Doi:10.1111/mec.12354

29. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9: 357–359. Doi:10.1038/nmeth.1923

30. 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: 559–575. Doi:10.1086/519795

31. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25: 1754–1760. Doi: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: 1297–1303. Doi:10.1101/gr.107524.110

33. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. Fly (Austin). 2012;6: 80–92. Doi:10.4161/fly.19695

34. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. Nat Biotechnol. 2011;29: 24–26. doi:10.1038/nbt.1754

35. Ye K, Schulz MH, Long Q, Apweiler R, Ning Z. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. Bioinformatics. 2009;25: 2865–2871. doi:10.1093/bioinformatics/btp394

36. Nakao N, Yasuo S, Nishimura A, Yamamura T, Watanabe T, Anraku T, et al. Circadian clock gene regulation of steroidogenic acute regulatory protein gene expression in preovulatory ovarian follicles. Endocrinology. 2007;148: 3031–3038. Doi:10.1210/en.2007-0044

37. Mi H, Muruganujan A, Ebert D, Huang X, Thomas PD. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. Nucleic Acids Res. 2019;47: D419–D426. Doi:10.1093/nar/gky1038

38. Hunter GA, Ferreira GC. Molecular enzymology of 5-aminolevulinate synthase, the gatekeeper of heme biosynthesis. Biochim Biophys Acta. 2011;1814: 1467–1473. Doi:10.1016/j.bbapap.2010.12.015

39. Sharma P, Reichert M, Lu Y, Markello TC, Adams DR, Steinbach PJ, et al. Biallelic *HEPHL1* variants impair ferroxidase activity and cause an abnormal hair phenotype. PLOS Genetics. 2019;15: e1008143. Doi:10.1371/journal.pgen.1008143

40. Qian ZM, Li H, Sun H, Ho K. Targeted drug delivery via the transferrin receptor-mediated endocytosis pathway. Pharmacol Rev. 2002;54: 561–587. Doi:10.1124/pr.54.4.561

41. Semina EV, Ferrell RE, Mintz-Hittner HA, Bitoun P, Alward WL, Reiter RS, et al. A novel homeobox gene *PITX3* is mutated in families with autosomal-dominant cataracts and ASMD. Nat Genet. 1998;19: 167–170. Doi:10.1038/527

42. Smidt MP, Smits SM, Bouwmeester H, Hamers FPT, van der Linden AJA, Hellemons AJCGM, et al. Early developmental failure of substantia nigra dopamine neurons in mice lacking the homeodomain gene Pitx3. Development. 2004;131: 1145–1155. Doi:10.1242/dev.01022

43. L'Honoré A, Coulon V, Marcil A, Lebel M, Lafrance-Vanasse J, Gage P, et al. Sequential expression and redundancy of *Pitx2* and *Pitx3* genes during muscle development. Dev Biol. 2007;307: 421–433. Doi:10.1016/j.ydbio.2007.04.034

44. L'honoré A, Commère P-H, Ouimette J-F, Montarras D, Drouin J, Buckingham M. Redox Regulation by *Pitx2* and *Pitx3* is critical for fetal myogenesis. Developmental Cell. 2014;29: 392–405. Doi:10.1016/j.devcel.2014.04.006

45. Basu S, Schultz AM, Basu M, Roseman S. Enzymatic synthesis of galactocerebroside by a galactosyltransferase from embryonic chicken brain. J Biol Chem. 1971;246: 4272–4279.

46. Chung J, Nakatsu F, Baskin JM, De Camilli P. Plasticity of PI4KIIIα interactions at the plasma membrane. EMBO Rep. 2015;16: 312–320. Doi:10.15252/embr.201439151

47. Perletti L, Talarico D, Trecca D, Ronchetti D, Fracchiolla NS, Maiolo AT, et al. Identification of a Novel Gene,PSD,Adjacent toNFKB2/lyt-10,Which Contains Sec7 and Pleckstrin-Homology Domains. Genomics. 1997;46: 251–259. Doi:10.1006/geno.1997.5022

48. Zhang P-J, Zhao J, Li H-Y, Man J-H, He K, Zhou T, et al. CUE domain containing 2 regulates degradation of progesterone receptor by ubiquitin-proteasome. EMBO J. 2007;26: 1831–1842. Doi:10.1038/sj.emboj.7601602

49. Maruoka M, Suzuki J, Kawata S, Yoshida K, Hirao N, Sato S, et al. Identification of B cell adaptor for PI3-kinase (BCAP) as an Abl interactor 1-regulated substrate of Abl kinases. FEBS Lett. 2005;579: 2986–2990. Doi:10.1016/j.febslet.2005.04.052

50. Samiullah S, Roberts J, Wu S-B. Downregulation of ALAS1 by nicarbazin treatment underlies the reduced synthesis of protoporphyrin IX in shell gland of laying hens. Sci Rep. 2017;7: 6253. Doi:10.1038/s41598-017-06527-y

51. Polin, D. Formation of porphyrin from delta-aminolevulenic acid by uterine and liver tissue from laying hens. *Proc. Soc. Exp. Biol. Med.* **94**, 276–279 (1957).

52. Li G, Chen S, Duan Z, Qu L, Xu G, Yang N. Comparison of protoporphyrin IX content and related gene expression in the tissues of chickens laying brown-shelled eggs. Poult Sci. 2013;92: 3120–3124. Doi:10.3382/ps.2013-03484

53. Baird T, Solomon SE, Tedstone DR. Localisation and 33haracterization of egg shell porphyrins in several avian species. Br Poult Sci. 1975;16: 201–208. Doi:10.1080/00071667508416177

54. Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. PNAS. 2010;107: 21931–21936. Doi:10.1073/pnas.1016071107

55. Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. Curr Protoc Mol Biol. 2015;109: 21.29.1-21.29.9. doi:10.1002/0471142727.mb2129s109

56. Li Z, Ren T, Li W, Zhou Y, Han R, Li H, et al. Association between the methylation statuses at CpG sites in the promoter region of the *SLCO1B3*, RNA expression and color change in blue eggshells in Lushi chickens. Front Genet. 2019;10. Doi:10.3389/fgene.2019.00161

# Supporting Information captions

**S1 Fig. Schematic diagram of the heme metabolism pathway.** The image in the upper panel was prepared by modifying the figure from Sachar et al. (2016) [11].

**S2 Fig. Mucosae of wild-type uteri after brushing and isthmuses before and after brushing.** (A–D) Mucosae of uteri after brushing. (B, D) Higher-magnification images of the parts shown inside frames in (A, C). The epithelium was removed from the mucosa (arrows). Arrowheads indicate remaining epithelia. (E–H) Mucosae of isthmuses before brushing (E, F) and after brushing (G, H). (F, H) Higher-magnification images of the parts shown inside frames in (E, G). Bars indicate 100 μm.

**S3 Fig. External shapes of quail oviducts.** Schematic diagram of quail oviduct containing an egg (A). The uteri of the wild-type and mutant quail were dark brown and pale brown, respectively (B). The vaginas were removed from the oviducts. Scale bars in (B) 1 cm

**S4 Fig. Transmission electron microscopy of the uterine mucosal epithelium.** (A, B) Wild-type uterus (C, D) Mutant uterus The part shown inside the frame in (A) is enlarged in (C). Nuclei of apical and basal cells are indicated by arrows and arrowheads, respectively. Transport vesicles are observed as electron-dense granules in the apical cells of both the wild-type and mutant epithelia (asterisks). Magnifications are 4000× in (A), 10000× in (B), 1500× in (C), and 7000× in (D).

**S5 Fig. Comparison of *ALAS1* and *HEPHL1* expression between wild-type and mutant uteri by qPCR.** The graphs show the results of qPCR analysis using RNA samples ($n$ = 5 for both wild-type and mutant quail) used for mRNA sequencing (A, B) and RNA samples ($n$ = 5 for both wild-type and mutant quail) that were not used for mRNA sequencing (C, D). The expression level of the wild-type uterus is shown as a relative value to that of the mutant uterus, defined as 1. The expression level in each sample is shown as a circle. Double asterisks indicate $p < 0.01$.

**S6 Fig. Genotypes of SNPs located near the candidate region in nine $F_2$ progeny that exhibited discordance between phenotypes and genotypes of SNP201255 or SNP202246.** Genotypes of SNPs 201288, Chr6-we1, and Chr6-we-a were fully concordant with the eggshell color phenotypes.

**S7 Fig. Comparison of *PITX3* expression between wild-type and mutant uteri by qPCR.** The graphs show the results of qPCR analysis using RNA samples ($n$ = 5 for both wild-type and mutant quail) used for mRNA sequencing. The expression level of the wild-type uterus is shown as a relative value to that of the mutant uterus, defined as 1. The expression level in each sample is shown as a circle. NS, not significant. Double asterisks indicate $p < 0.01$.

**S1 Table. Nucleotide sequences of primers used for genotyping, quantitative PCR, and the cloning of cDNA fragments.**

**S2 Table. Fold changes in expression of heme synthesis-related genes in mutant uteri compared to that in wild-type uteri.**

**S3 Table. Functions and relative expression levels [log2 (fold change)] of 11 genes, which were expressed at higher levels in the wild-type uterus than in other organs and tissues and**

**were downregulated in the mutant uterus. The expression level of genes in each organ or tissue was defined as 1.**

**S4 Table. Gene Ontology (GO) "biological process" (GO:BP) terms overrepresented in 382 genes that were specifically upregulated in the uterus.**

**S5 Table. Nucleotide positions and genotypes of the three informative single nucleotide polymorphism (SNP) markers.**

**S6 Table. Categories of sequence variants in the 172 kb candidate region.**

**S7 Table. Putative effects of missense variants of genes located in the 172 kb candidate region.**

**S8 Table. Fold changes of differentially expressed genes (DEGs) on chromosome 6.**

**S1 Data. Statistics of mRNA sequencing, ddRAD sequencing, and whole-genome resequencing; information on DNA samples for genetic linkage analysis; and nucleotide sequences of adaptors and primers for the construction of ddRAD sequencing libraries.**

**S2 Data. Sequence of cDNAs used for *in situ* hybridization.**

**S3 Data. Raw count data used in the comparison of gene expression between wild-type and mutant uteri and the result of differential gene expression analysis.**

**S4 Data. Raw count data used in the comparison of gene expression between uteri and other organs or tissues and the result of differential gene expression analysis.**

**S5 Data. Results from qPCR analysis of *ALAS1*, *HEPHL1*, and *PITX3*.**

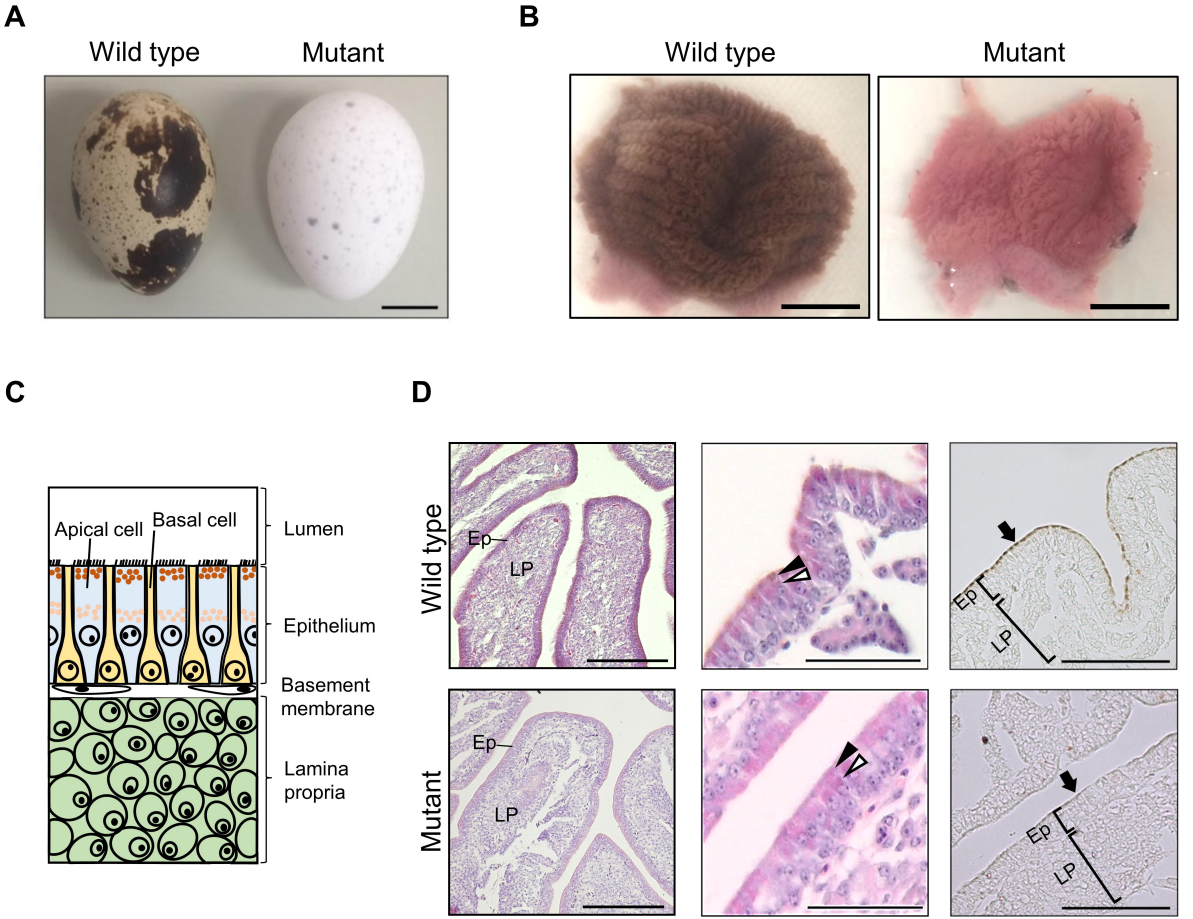
**S6 Data. Results from the genome-wide association test and LOD scores and recombination frequencies of single nucleotide polymorphisms (SNPs) on chromosome 6.**

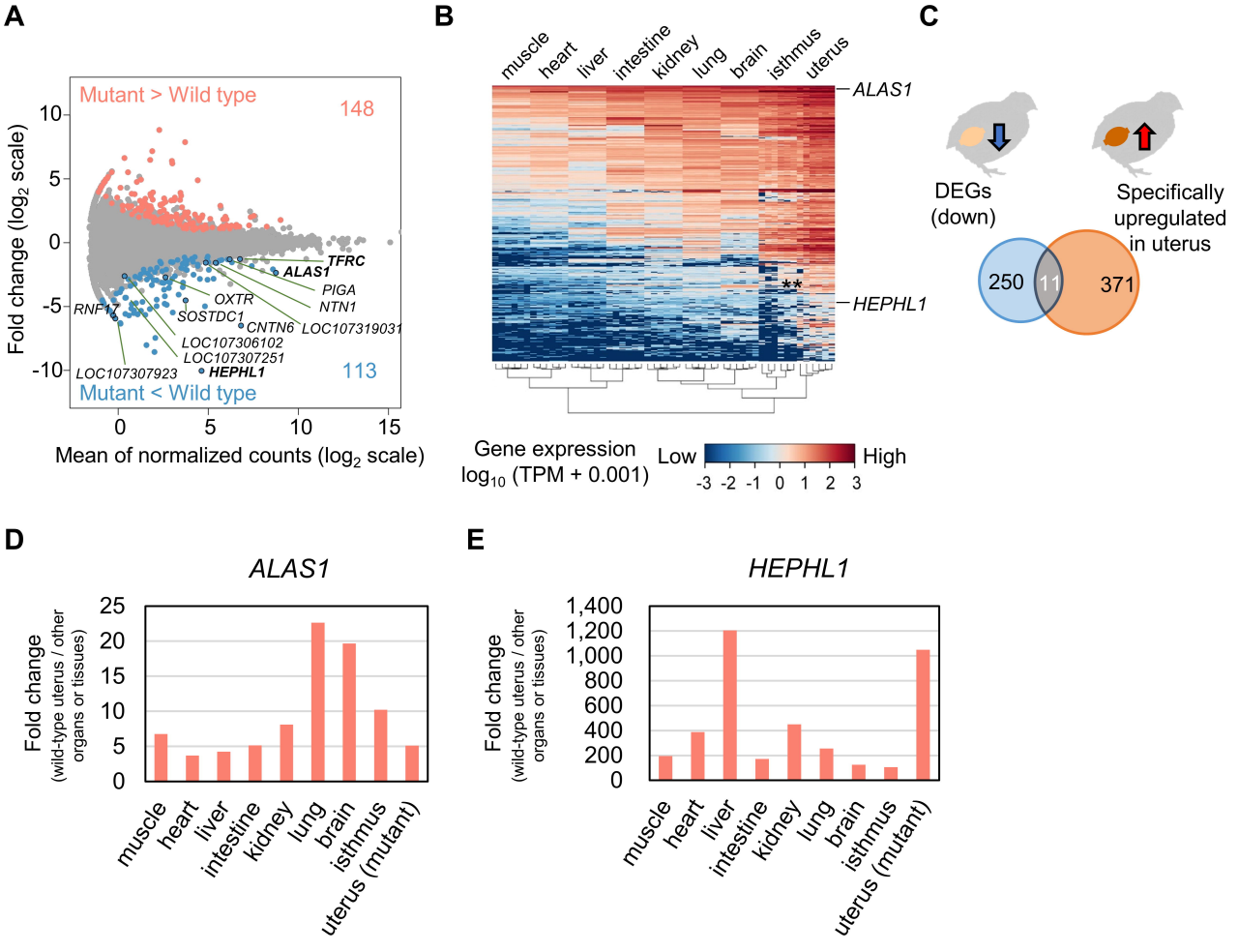
**S7 Data. Sequence variants around the candidate region and annotation of candidate sequence variants within the candidate region.**

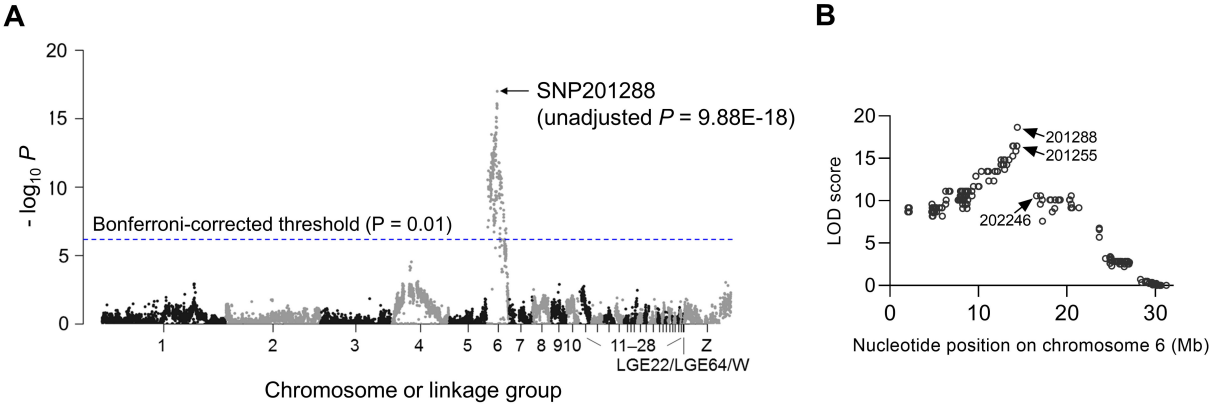
**S8 Data. Results from biochemical analysis.**

# Author contributions

Conceptualization, S.Ishishita, K.K., Y.M., S.S., Y.G., K.Y., and S.T.; Formal Analysis, S.Ishishita, M.T., and S.K.; Funding acquisition, Y.M.; Investigation, S.Ishishita, M.T., S.K., S.Iwasaki, S.T., K.Y., I.H., Y.Ohmori, T.S., and A.H.; Methodology, S.Ishishita and S.K.; Project administration; S.Ishishita; Resources, K.K., T.S., Y.M., S.S., Y.G., and Y.Ohkawa; Software, S.Ishishita, S.T., A.H., and K.Y.; Supervision, Y.M. and T.S.; Validation, S.Ishishita, S.T., K.Y, S.S., and Y.G.; Visualization, S.Ishishita., M.T., S.K., Y.K.; Writing – original draft, S.Ishishita; Writing – review & editing, S.Ishishita, T.S., and Y.M.

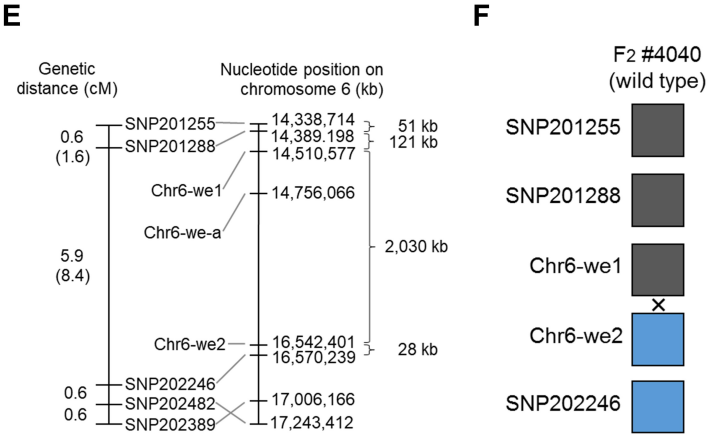

A
Wild type
Mutant
B
Wild type
Mutant
C
Apical cell
Basal cell
Lumen
Epithelium
Basement membrane
Lamina propria
D
Wild type
Mutant
Ep
LP

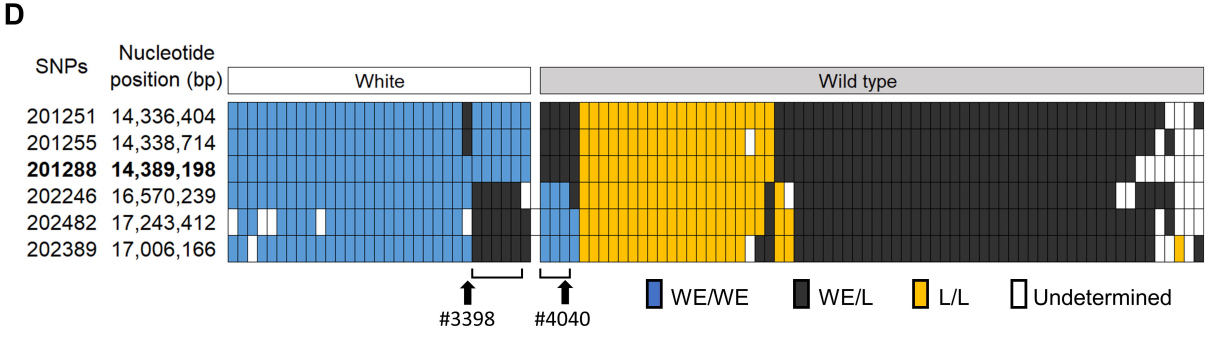

A
Mutant > Wild type
148
113
Mutant < Wild type
Fold change (log2 scale)
Mean of normalized counts (log2 scale)
TFRC
ALAS1
PIGA
NTN1
OXTR
SOSTDC1
CNTN6
LOC107319031
RNF17
LOC107306102
LOC107307251
LOC107307923
HEPHL1
B
muscle
heart
liver
intestine
kidney
lung
brain
isthmus
uterus
ALAS1
HEPHL1
**
Gene expression
log10 (TPM + 0.001)
Low
High
C
DEGs
(down)
Specifically
upregulated
in uterus
250
11
371
D
ALAS1
Fold change
(wild-type uterus / other
organs or tissues)
muscle
heart
liver
intestine
kidney
lung
brain
isthmus
uterus (mutant)
E
HEPHL1
Fold change
(wild-type uterus / other
organs or tissues)
muscle
heart
liver
intestine
kidney
lung
brain
isthmus
uterus (mutant)

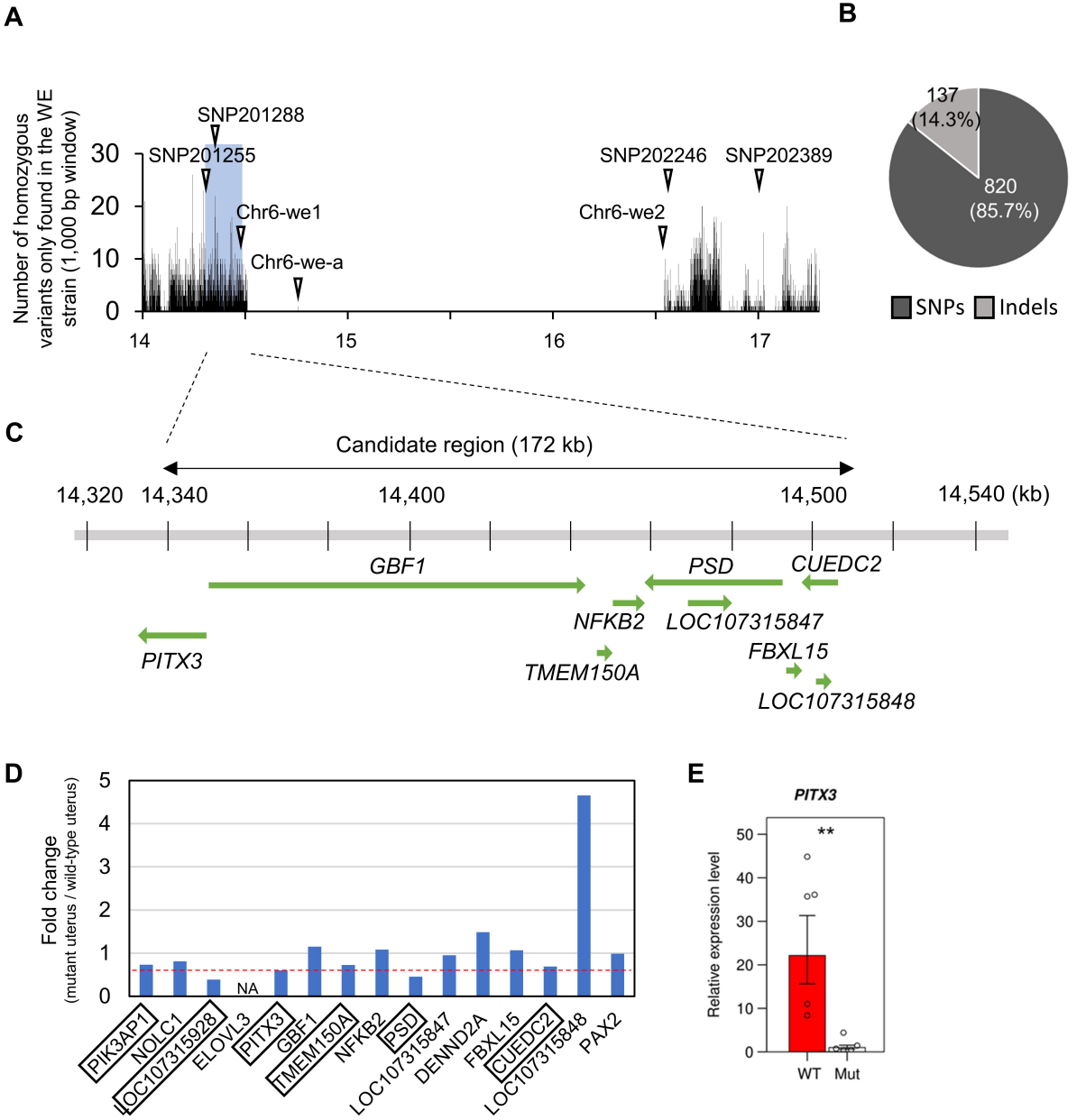


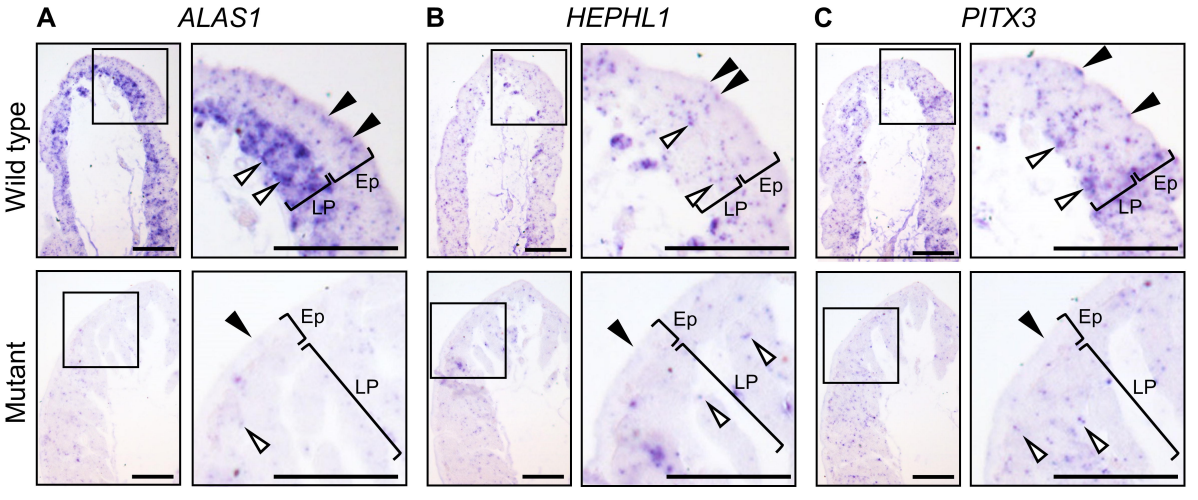

C

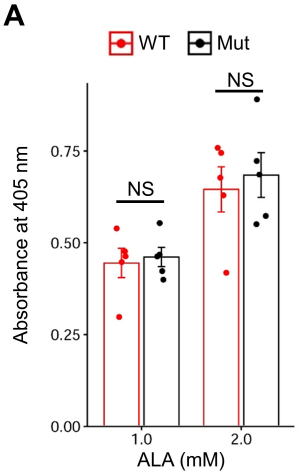
| SNP | Nucleotide position (bp) | LOD score | Recombination frequency |
|---|---|---|---|
| 201255 | 14,339,714 | 16.44 | 0.02 |
| 201288 | 14,389,198 | 18.66 | 0 |
| 202246 | 16,570,239 | 10.59 | 0.08 |

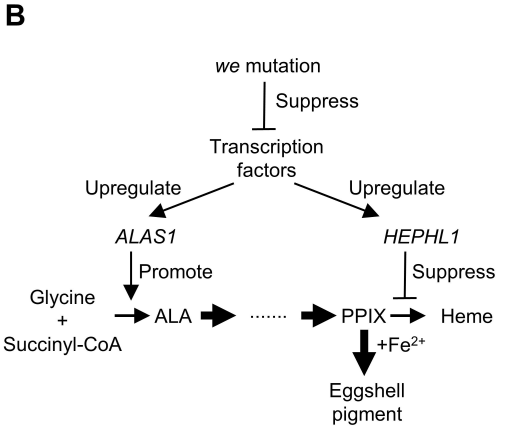


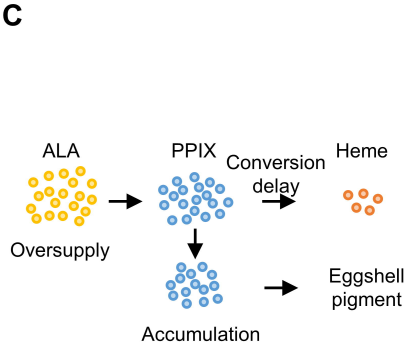




A
Number of homozygous variants only found in the WE strain (1,000 bp window)
30
20
10
0
SNP201288
SNP201255
Chr6-we1
Chr6-we-a
SNP202246
SNP202389
Chr6-we2
14
15
16
17
B
137 (14.3%)
820 (85.7%)
SNPs
Indels
C
Candidate region (172 kb)
14,320
14,340
14,400
14,500
14,540 (kb)
GBF1
PSD
CUEDC2
NFKB2
LOC107315847
PITX3
FBXL15
TMEM150A
LOC107315848
D
Fold change
(mutant uterus / wild-type uterus)
5
4
3
2
1
0
NA
PIK3AP1
NOLC1
LOC107315928
ELOVL3
PITX3
GBF1
TMEM150A
NFKB2
PSD
LOC107315847
DENND2A
FBXL15
CUEDC2
LOC107315848
PAX2
E
PITX3
Relative expression level
50
40
30
20
10
0
**
WT
Mut


A
ALAS1
B
HEPHL1
C
PITX3
Wild type
Mutant
Ep
LP


A
WT
Mut
NS
NS
Absorbance at 405 nm
0.75
0.50
0.25
0.00
1.0
2.0
ALA (mM)


B
we mutation
Suppress
Transcription factors
Upregulate
Upregulate
ALAS1
HEPHL1
Promote
Suppress
Glycine + Succinyl-CoA
ALA
PPIX
Heme
+Fe2+
Eggshell pigment


C
ALA
Oversupply
PPIX
Conversion delay
Heme
Accumulation
Eggshell pigment