An *EAV-HP* Insertion in 5′ Flanking Region of *SLCO1B3* Causes Blue Eggshell in the Chicken

Zhepeng Wang, Luijia Qu, Junfeng Yao, Xiaolin Yang, Guangqi Li, Yuanyuan Zhang, Junying Li, Xiaotong Wang, Jirong Bai, Guiyun Xu, Xuemei Deng, Ning Yang, Changxin Wu

National Engineering Laboratory for Animal Breeding and Key Laboratory of Animal Genetics, Breeding, and Reproduction of the Ministry of Agriculture, China Agricultural University, Beijing, China

Abstract

The genetic determination of eggshell coloration has not been determined in birds. Here we report that the blue eggshell is caused by an *EAV-HP* insertion that promotes the expression of *SLCO1B3* gene in the uterus (shell gland) of the oviduct in chicken. In this study, the genetic map location of the blue eggshell gene was refined by linkage analysis in an *F₂* chicken population, and four candidate genes within the refined interval were subsequently tested for their expression levels in the shell gland of the uterus from blue-shelled and non-blue-shelled hens. *SLCO1B3* gene was found to be the only one expressed in the uterus of blue-shelled hens but not in that of non-blue-shelled hens. Results from a pyrosequencing analysis showed that only the allele of *SLCO1B3* from blue-shelled chickens was expressed in the uterus of heterozygous hens (*O*∗/O*N*). *SLCO1B3* gene belongs to the organic anion transporting polypeptide (OATP) family; and the OATPs, functioning as membrane transporters, have been reported for the transportation of amphipathic organic compounds, including bile salt in mammals. We subsequently resequenced the whole genomic region of *SLCO1B3* and discovered an *EAV-HP* insertion in the 5′ flanking region of *SLCO1B3*. The *EAV-HP* insertion was found closely associated with blue eggshell phenotype following complete Mendelian segregation. In situ hybridization also demonstrated that the blue eggshell is associated with ectopic expression of *SLCO1B3* in shell glands of uterus. Our finding strongly suggests that the *EAV-HP* insertion is the causative mutation for the blue eggshell phenotype. The insertion was also found in another Chinese blue-shelled breed and an American blue-shelled breed. In addition, we found that the insertion site in the blue-shelled chickens from Araucana is different from that in Chinese breeds, which implied independent integration events in the blue-shelled chickens from the two continents, providing a parallel evolutionary example at the molecular level.

Introduction

Avian eggshell coloration is the result of crypsis or mimetism and plays important roles in filtering solar radiation and strengthening the eggshell [1]. Blue eggshell color has been proposed as post-mating signals of female phenotypic quality to their mates and is related to fitness of the offspring due to the antioxidant of biliverdin, a predominant pigment for blue eggs [2,3]. Blue eggshells can be found not only in some wild birds, e.g. eastern bluebird [4], blue-footed booby [5], and pied flycatcher [6], but also in domestic birds such as Japanese quail [7], chickens [8] and ducks [9].

Brown and white are the two major eggshell colors in chickens. Protoporphyrin-IX, biliverdin, and biliverdin zinc chelate are the main pigments of the eggshell [10] and several blue egg laying breeds have been reported worldwide [11,12]. The Araucana, an indigenous breed from Chile, was the first chicken breed described to lay blue eggs [8], and has been frequently used in genetic studies of the blue eggshell phenotype. In China, Dongxiang and Lushi chickens are representative breeds laying blue eggs and show dominant inheritance as that in Araucana. However, the blue eggshell phenotype has not been fixed in these three breeds which still produce brown eggs at low frequency.

Blue eggshell color exhibits an autosomal dominant inheritance and eggs laid by homozygotes are a darker blue than those from heterozygotes (Figure 1A). In 1933, Punnett firstly reported that blue or green shell appearance of the Araucana was determined by a single genetic factor, traditionally denoted as oocyan (*O*) [8]. A series of linkage analysis involving *O* have been performed with *O* affirmatively mapped to the short arm of chromosome 1 [12–16], and closely linked to *ev1* and *P* which was identified as *SRY* (sex determining region Y)-box 5 (*SOX5*) [12,13,16,17]. In the region around *ev1*, two single nucleotide polymorphisms (SNPs) (rs15297163 and rs15297165) were found to be highly associated with the blue eggshell phenotype [18]. A 1.8 Mb genomic interval harboring the *O* gene was defined in an *F₂* resource population [19]. The localization of the *O* was further refined to the vicinity of ss244244378 by linkage and association analysis [20].
The eggshell color of birds is of wide interest, but the molecular basis remained unknown until our discovery, reported here. The blue eggshell is found not only in wild birds but also in domestic fowls. In this study, we identified that blue eggshell in chickens from different geographical regions is caused by a \(~4.2\) kb EAV-HP insertion in the 5' flanking region of SLCO1B3. The EAV-HP insertion in chicken is a derived mutation in domestic chickens. The genetic determination of blue eggshell in other birds requires further investigation. We also found that the EAV-HP insertions in the chickens from China and America were separate integration events, which presents us with a parallel molecular evolution example driven by artificial selection.

**Author Summary**

The eggshell color of birds is of wide interest, but the molecular basis remained unknown until our discovery, reported here. The blue eggshell is found not only in wild birds but also in domestic fowls. In this study, we identified that blue eggshell in chickens from different geographical regions is caused by a \(~4.2\) kb EAV-HP insertion in the 5' flanking region of SLCO1B3. The EAV-HP insertion in chicken is a derived mutation in domestic chickens. The genetic determination of blue eggshell in other birds requires further investigation. We also found that the EAV-HP insertions in the chickens from China and America were separate integration events, which presents us with a parallel molecular evolution example driven by artificial selection.

**Results**

**Linkage analysis of Chicken blue eggshell gene**

A linkage analysis was performed in an \(F_2\) resource population segregating for the \(O\) gene to refine the location of chicken blue shell gene in the present study. Eight molecular markers in the candidate region were used for linkage analysis (Table S1). By two-point analysis, the \(O\) gene was mapped in the region between marker L4 and L5 which were the closest flanking markers to the \(O\) gene was mapped in the region between marker L4 and L5 which were the closest flanking markers to the \(O\) gene with recombination rate being both \(0.02 (\text{LOD} = 15.84)\) (Figure 2). We subsequently cloned the \(5'\) UTR end in the \(O\) gene expressed in the uterus of blue-shelled (O*LC/O*LC) Dongxiang chicken (May 2006 assembly) (Table S1) and no recombination was found between the blue eggshell phenotype and the markers within the region.

**Specific expression of SLCO1B3 in uterus of blue-shelled chicken**

Totally, four genes (SLCO1C1, SLCO1B3, LOC418198 and SLCO1A2) were found in the \(~120\) kb interval by a MapView search (http://www.ncbi.nlm.nih.gov/mapview/) (Table S2). The uterus is where pigment is secreted to eggshell. We performed expression analysis for the four candidate genes in the uterus of blue-shelled (\(n = 16\)) and brown-shelled (\(n = 16\)) Dongxiang hens by RT-PCR. We found that SLCO1B3 was the only gene expressed specifically in the uterus of blue-shelled Dongxiang chickens (Figure 1B), thus we measured its expression in the uterus in 5 blue-shelled (4 \(O^*\text{LC}/O^*\text{LC}\) and 4 \(O^*\text{LC}/O^*\text{N}\)) and 4 \(O^*\text{LC}/O^*\text{N}\) is for the blue-shell allele in the two Chinese breeds, Dongxiang and LushiChicken and \(O^*\text{N}\) denotes the non-blue-shell allele (or wild-type allele) and 4 brown-shelled (\(O^*\text{N}/O^*\text{N}\)) Dongxiang chickens, 6 blue-shelled (3 \(O^*\text{LC}/O^*\text{LC}\) and 3 \(O^*\text{LC}/O^*\text{N}\)) and 6 brown-shelled (\(O^*\text{N}/O^*\text{N}\)) Lushi chickens, and 24 chickens from three brown-shelled and one white-shelled breeds (\(O^*\text{N}/O^*\text{N}\), 6 chicken per breed) by real-time PCR. All blue-shelled chickens expressed the gene in the uterus while non-blue-shelled chickens did not (Figure 1C). In addition, expression of SLCO1B3 was 2 to 3 fold higher in homozygous blue-shell chickens than in heterozygous blue-shell Dongxiang and Lushi individuals (Figure 1C). Fluorescein labeled cDNA \textit{in situ} hybridization demonstrated that the transcripts of SLCO1B3 were only expressed in the uterus of blue-shell but not brown-shell chickens (Figure 1D). These results suggest that SLCO1B3 is the causative gene for blue eggshell in the chicken.

**Allele-specific expression of SLCO1B3**

We found a SNP (\(g.67334934\) \(G>T\)) in exon 5 of SLCO1B3 gene by sequencing the coding region and the SNP presented complete association with the blue eggshell phenotype in the Dongxiang chicken by genotyping it in Dongxiang blue-shelled and brown-shelled chickens. With six heterozygous individuals produced by mating a homozygous Dongxiang blue-shelled male with a White Leghorn female, the allelic expression of SLCO1B3 gene was demonstrated by RT-PCR analysis and pyrosequencing. More than 95% of the transcripts expressed in the uterus originated from the \(T\) allele corresponding to the blue-shell allele (Figure 1E). This means the expression of the gene is regulated by a \(\alpha\)-acting element. Surprisingly, its expression in liver is also allele specific, and \(~95\%) of the transcripts in liver come from the \(G\) allele which is non-blue-shell allele (Figure 1F).

**An EAV-HP insertion is completely associated with blue eggshell phenotype**

We sequenced the genomic region of SLCO1B3 in order to reveal the potential causative mutation of the gene with 5 blue-shelled and 5 brown-shelled Dongxiang chickens. Twenty-one SNPs evenly covering the whole genomic region (\(~24\) kb) of SLCO1B3 were taken for genotyping in 353 chickens from 3 blue-shelled breeds (Araucana, Dongxiang and Lushi) and 9 non-blue-shelled breeds. However, none of the SNPs was found to be in complete linkage disequilibrium with blue eggshell (Table 1).

We subsequently cloned the 5'UTR (GenBank accession number: JN381032) of SLCO1B3 by 5' RACE in a blue-shelled (\(O^*\text{LC}/O^*\text{LC}\) and 5 brown-shelled (\(O^*\text{N}/O^*\text{N}\)) Dongxiang chicken and an extra 24 bps were found at the beginning of 5'UTR end in blue-shelled Dongxiang chicken (Figure S1). We further sequenced 5 kb upstream of the promoter using 5 blue-shelled (\(O^*\text{LC}/O^*\text{LC}\) and 5 brown-shelled (\(O^*\text{N}/O^*\text{N}\)) Dongxiang chickens. A \(~4.2\) kb insertion adjacent to 5'UTR containing the extra 24 bps was found in the blue-shelled but not in the brown-shelled chickens. The sequence of the \(~4.2\) kb insertion (GenBank accession number: JF837512) represents an incomplete retrovirus and shows 95.8% identity with the sequence of the avian EAV-HP retrovirus (EMBL accession number: AJ238124) [22]. A typical proviral structure consists of gag, pol and env flanked by long terminal repeat (LTR), which are arranged in the order of 5'LTR-gag-pol-env-LTR3' [22]. Here, the inserted retrovirus is absent of the whole pol gene and part of gag and env (Figure 3A). The retrovirus was integrated into the blue-shelled chicken genome in an inverted orientation (Figure 3B) at Chrl: 67324641–67324642. We also found that the EAV encompassed some promoter elements by sequence analysis, indicating its expression promotion activities (Figure 3A).

A wide-range survey of the EAV-HP insertion was performed in 705 chickens from 12 worldwide breeds and the \(F_2\) resource population (Table 2) using diagnostic PCR test. The results that the EAV-HP insertion is completely associated with the blue eggshell phenotype provide strong evidence that the mutation is causative.
Independent EAV-HP insertion events in blue-shelled chickens from China and Chile

In order to elucidate whether the blue-shelled chickens from China and Chile have the same origin for the genotypic mutation, we further sequenced the EAV-HP insertion regions in a homozygous Araucana and a homozygous blue-shelled Lushi chicken. The EAV-HP insertion was found in both samples and the alignments of Araucana and Lushi to Dongxiang blue-shelled chickens.

Figure 1. The eggshell color and the expression of SLC01B3 in the uterus of blue-shelled and non-blue-shelled chickens. (A) Eggshell colors of homozygous blue-shelled (O*LC/O*LC), heterozygous blue-shelled (O*LC/O*N) and brown-shelled (O*N/O*N) of Dongxiang chickens. (B) The expression analysis of the four genes which are located in the refined region between markers L4 and L5 in the uterus of Dongxiang blue-shelled hens (n = 16) and Dongxiang non-blue-shelled hens (n = 16) by RT-PCRs. DX-BS: blue-shelled Dongxiang, DX-NBS: non-blue-shelled Dongxiang. (C) Analysis of SLC01B3 expression in the uterus. Expression data were presented as fold relative to heterozygous blue-shelled Dongxiang chickens (O*LC/O*N) by the comparative C_{t} method (2^{-AC_{t}}). SLC01B3 is exclusively expressed in blue-shelled chickens, and the amount of SLC01B3 transcripts in homozygous blue-shelled chicken (O*LC/O*LC) is approximately two to three folds of that in heterozygous individuals (O*LC/O*N). DX-BS: blue-shelled Dongxiang, LS-BS: blue-shelled Lushi, DX-NBS: non-blue-shelled Dongxiang, LS-NBS: non-blue-shelled Lushi. (D) Micrographs of cDNA in situ hybridization for SLC01B3 mRNA in the uterus from blue-shelled and non-blue-shelled chickens. (E) Differential expression of SLC01B3 transcript in the uterus from blue-shelled heterozygotes using genomic DNA (gDNA) as control. The polymorphic position g. 67334934 G>T was used to monitor differential expression using pyrosequencing. T and G at this position correspond to blue-shell and non-blue-shell alleles, respectively. Due to two Ts next to the SNP at 3’ end, the peaks of T in the schema contain three Ts including one T from blue-shell allele and two Ts from non-blue-shell allele. The percent expression on the peaks for T and G are the T or G at g. 67334934 G>T. (F) Summary of the detection of differential expression in uterus and liver from six heterozygotic blue-shelled (O*LC/O*N) birds.

doi:10.1371/journal.pgen.1003183.g001
chicken showed the identity of the inserted EAV-HP being around 97%. Interestingly, the insertion sites in Araucana are different from that in the two Chinese blue-shelled chickens. The break point for EAV-HP insertion in blue-shelled Araucana is located at 23 bp upstream to that in the two Chinese breeds (Figure 3C). We sequenced the junction sites in homozygous blue-shelled chickens of Araucana (n = 5), Lushi (n = 5) and Dongxiang (n = 5) and confirmed the insertion sites in Dongxiang and Lushi are the same but different from that in Araucana.

We also typed 21 SNPs in genomic region of SLCO1B3 in multiple breeds of blue-shelled and non-blue-shelled chicken breeds. It is obvious to see that the EAV-HP insertions in blue-shelled chickens from the two continents were embedded in two distinguished different haplotypes (Table S3), which supports independent integrations accounting for the blue shelled phenotypes.

Discussion

In birds, eggshell color is a variable Mendelian trait. Colored eggshell could function as avoiding predation through either crypsis or aposematism, distinguishing from brood parasitism, reinforcing eggshell strength, regulating egg temperature, combating harmful solar radiation and sending sexually selected signal to males [2,23]. However, molecular mechanism of all kinds of eggshell color formation is poorly understood to date. Here, we demonstrate that a ~4.2 kb EAV-HP insertion at upstream of SLCO1B3 is responsible for blue eggshell phenotype in the chicken.

By linkage analysis, we fine mapped the O locus to a 120 kb region, where four candidate genes of SLCO1C1, SLCO1B3, LOC418189 and SLCO1A2 are located. These genes are all members of organic anion transporting polypeptides (OATP, gene symbol SLCO). Functionally, the OATPs serve as membrane transporters that mediate a wide range of sodium-independent transport of amphipathic organic compounds, such as some endobiotic compounds of bile salts, eicosanoids, steroids, thyroid hormones and some xenobiotic compounds of anionic oligopeptides, organic dyes, toxins, and drugs [24]. SLCO1B3 codes a membrane transporter OATP1B3 which is considered a liver-specific transporter and is highly expressed in liver where it transports a wide range of substrates including bile salts [24,25]. A genome-wide association study (GWAS) for serum bilirubin levels also showed SLCO1B3 is a plausible candidate gene responsible for changes in bilirubin levels in humans [26]. As blue egg is colored mainly by deposition of biliverdin on the eggshell and biliverdin is just one component of the bile salts, the expression of the SLCO1B3 in uterus could enhance transportation of biliverdin to eggshell. In this study, we found that SLCO1B3 is exclusively expressed in shell gland of uterus of blue-shelled chickens rather than in that of brown- or white-shelled chickens, which supports that the gene plays a pivotal role for coloration of blue eggs.

Regulatory mutations demonstrate an important role for phenotypic diversity which may be explained by cis-acting elements [17,27–34]. The effect of endogenous retrovirus (ERV) on hosts is extensive. It can unfavorably influence certain production traits, i.e. egg production, egg weight and body weight.
Table 1. Distribution of allelic frequencies of SNPs and EAV-HP in SLCO1B3 in several blue-shelled and non-blue-shelled breeds.

| Description     | Blue-shelled* | Brown-shelled | Red Jungle Fowl | Rhode Island Red | Dwarf | Luxi Game | Silkie | Tibetan | White Leghorn |
|-----------------|---------------|---------------|-----------------|------------------|-------|------------|--------|----------|---------------|
|                 | Dongxiang     | Lushi         | Araucana        | Redjongli         |       |            |        |          |               |
|                 | 183.1         | g.67320217A>G | 0.92 (A)        | 0.85 (A)         | 0.07 (A) | 0.00 (A)  | 0.30 (A)| 0.03 (A) | 0.46 (A)      |
|                 | g.67320798A>G | 0.92 (A)      | 0.85 (A)        | 0.07 (A)         | 0.00 (A)| 0.03 (A)  | 0.30 (A)| 0.00 (A) | 0.46 (A)      |
|                 | g.67322980C>T | 0.92 (C)      | 0.83 (C)        | 0.01 (C)         | 0.00 (C)| 0.03 (C)  | 0.23 (C)| 0.00 (C) | 0.46 (C)      |
|                 |               |               |                 |                  |       |            |        |          |               |
| EAV-HP<sup>x</sup> | 0.92 (+)<sup>x</sup> | 0.83 (+)<sup>x</sup> | 0.79 (+)<sup>x</sup> | 0.00 (<sup>x</sup>) | 0.00 (<sup>x</sup>) | 0.00 (<sup>x</sup>) | 0.00 (<sup>x</sup>) | 0.00 (<sup>x</sup>) | 0.00 (<sup>x</sup>) |
|                 | g.67325964G>C | 0.97 (G)      | 0.85 (G)        | 0.79 (G)         | 0.63 (G)| 0.65 (G)  | 0.75 (G)| 0.45 (G) | 0.43 (G)      |
|                 | g.67328567A>G | 0.97 (G)      | 0.83 (G)        | 0.00 (G)         | 0.00 (G)| 0.17 (G)  | 0.13 (G)| 0.00 (G) | 0.15 (G)      |
|                 | g.67330251C>T | 1.00 (C)      | 1.00 (C)        | 0.00 (C)         | 1.00 (C)| 0.38 (C)  | 0.48 (C)| 0.18 (C) | 0.28 (C)      |
|                 | g.67330655C>T | 0.92 (T)      | 0.83 (T)        | 0.00 (T)         | 0.00 (T)| 0.02 (T)  | 0.05 (T)| 0.00 (T) | 0.00 (T)      |
|                 | g.67332494A>G | 1.00 (A)      | 0.97 (A)        | 1.00 (A)         | 1.00 (A)| 0.43 (A)  | 0.31 (A)| 0.97 (A) | 0.50 (A)      |
|                 | g.67333488C>T | 0.92 (T)      | 0.83 (T)        | 0.00 (T)         | 0.00 (T)| 0.03 (T)  | 0.13 (T)| 0.00 (T) | 0.00 (T)      |
|                 | g.67334934G>T | 0.92 (T)      | 0.83 (T)        | 0.00 (T)         | 0.00 (T)| 0.07 (T)  | 0.08 (T)| 0.00 (T) | 0.00 (T)      |
|                 | g.67335392T>C | 0.97 (T)      | 0.85 (T)        | 0.14 (T)         | 0.63 (T)| 0.53 (T)  | 0.61 (T)| 0.78 (T)| 0.43 (T)      |
|                 | g.67336068C>T | 0.92 (T)      | 0.83 (T)        | 0.00 (T)         | 0.00 (T)| 0.03 (T)  | 0.07 (T)| 0.00 (T) | 0.00 (T)      |
|                 | g.67336530G>T | 0.92 (T)      | 0.83 (T)        | 0.00 (T)         | 0.00 (T)| 0.03 (T)  | 0.22 (T)| 0.00 (T) | 0.00 (T)      |
|                 | g.67336599C>T | 0.92 (T)      | 0.83 (T)        | 0.00 (T)         | 0.00 (T)| 0.03 (T)  | 0.21 (T)| 0.00 (T) | 0.00 (T)      |
|                 | g.67336867A>T | 0.92 (T)      | 0.83 (T)        | 0.00 (T)         | 0.00 (T)| 0.06 (T)  | 0.35 (T)| 0.00 (T) | 0.00 (T)      |
|                 | g.67337145A>G | 0.92 (G)      | 0.83 (G)        | 0.00 (G)         | 0.00 (G)| 0.03 (G)  | 0.22 (G)| 1.00 (G)| 0.00 (G)      |
|                 | g.67338442C>G | 0.92 (G)      | 0.83 (G)        | 0.00 (G)         | 0.00 (G)| 0.03 (G)  | 0.23 (G)| 0.00 (G) | 0.00 (G)      |
|                 | g.67390903C>T | 0.96 (T)      | 0.97 (T)        | 0.00 (T)         | 0.22 (T)| 0.23 (T)  | 0.24 (T)| 0.00 (T)| 0.03 (T)      |
|                 | g.6739848A>G   | 0.92 (G)      | 0.85 (G)        | 0.00 (G)         | 0.00 (G)| 0.13 (G)  | 0.32 (G)| 0.00 (G) | 0.02 (G)      |
|                 | g.67403760C>T | 0.92 (T)      | 0.83 (T)        | 0.00 (T)         | 0.00 (T)| 0.02 (T)  | 0.26 (T)| 0.17 (T)| 0.00 (T)      |
|                 | g.6742640A>G   | 0.92 (A)      | 0.85 (A)        | 0.00 (A)         | 0.00 (A)| 0.10 (A)  | 0.00 (A)| 0.00 (A)| 0.02 (A)      |

*Both homozygous and heterozygous hens in these populations lay blue-shelled eggs. Dongxiang and Lushi blue-shelled chickens are from China and Araucana chickens are from South America.

<sup>x</sup>EAV-HP is the only one mutation showing complete association with blue eggshell phenotype. The “+” sign indicates that EAV-HP insertion is positive, while the “−” sign indicates no EAV-HP.

<sup>+</sup>All detected blue-shelled chickens carry EAV-HP insertion and the frequency of “+” allele is not equal to 1 in blue-shelled groups because some birds are heterozygotes (+/−).
Figure 3. The EAV-HP insertion in blue-shelled chickens. (A) Schematic diagram of the relationship of a complete avian leukemia virus and the EAV-HP retrovirus. Some putative cis-regulatory elements in long terminal repeat (LTR) are underlined. The diagram was redrawn by referencing Sacco et al. [22]. (B) Arrangement of EAV-HP retrovirus and SLC01B3 in chicken genome. EAV-HP integrates into the 5' end of SLC01B3 in inverted orientation. When SLC01B3 is transcribed, an extra 24-bp sequence from the EAV-HP is also compiled into SLC01B3 transcript. The asterisk indicates the site of 24-bp sequence from EAV-HP, and arrows indicate the direction of transcription. (C) EAV-HP junction sites in blue-shelled Dongxiang, Lushi and Araucana breeds. The arrows show the EAV-HP inserted sites and the underlined sequences are the EAV-HP integration specific sequences. The sequences in red denote the inserted EAV-HP. doi:10.1371/journal.pgen.1003183.g003
induce lymphoid or erythroid leukemia and a variety of tumors [35], and cause some phenotype variants, i.e. dilute coat color mutation [36] and hairless mutation in mice [37], recessive white [38], henny-feathering mutation [39], and the sex-linked late-feathering mutation [40] in chickens and outbred wing mutation in Drosophila melanogaster [41]. ERV could alter splicing patterns of transcript to produce variants such as the recessive white mutation in the chickens [38]. ERV could also promote patterns of transcript to produce variants such as the recessive white mutation in the chickens [38]. ERV could also promote patterns of transcript to produce variants such as the recessive white mutation in the chickens [38]. ERV could also promote patterns of transcript to produce variants such as the recessive white mutation in the chickens [38]. ERV could also promote patterns of transcript to produce variants such as the recessive white mutation in the chickens [38]. ERV could also promote patterns of transcript to produce variants such as the recessive white mutation in the chickens [38]. ERV could also promote patterns of transcript to produce variants such as the recessive white mutation in the chickens [38]. ERV could also promote patterns of transcript to produce variants such as the recessive white mutation in the chickens [38].

### Table 2. EAV-HP insertion distributions in various populations.

| Breed or population        | n  | EAV-HP Insertion* |
|----------------------------|----|-------------------|
|                            | ++ | +/-               |
| Blue-shelled               |    |                   |
| Dongxiang (O*LC/-)         | 111| 100 11 0          |
| Resource family            | 5  | 5 0 0             |
| F1 (O*LC/O*N)              | 56 | 0 56 0            |
| F2 (O*LC/-)                | 121| 39 82 0           |
| Lushi (O*LC/-)             | 30 | 21 9 0            |
| Araucana (O*LA/-)          | 8  | 5 3 0             |
| Brown-shelled              |    |                   |
| Red jungle fowl (O*N/O*N)  | 31 | 0 0 31            |
| Dongxiang (O*N/O*N)        | 28 | 0 0 28            |
| Resource family            | 20 | 0 0 20            |
| F2 (O*N/O*N)               | 25 | 0 0 25            |
| Lushi (O*N/O*N)            | 30 | 0 0 30            |
| Beijing You (O*N/O*N)      | 30 | 0 0 30            |
| Silkies (O*N/O*N)          | 30 | 0 0 30            |
| Tibetan (O*N/O*N)          | 30 | 0 0 30            |
| Luxi Game (O*N/O*N)        | 30 | 0 0 30            |
| Gushi (O*N/O*N)            | 30 | 0 0 30            |
| Dwarf (O*N/O*N)            | 30 | 0 0 30            |
| Rhode Island Red (O*N/O*N) | 30 | 0 0 30            |
| White-shelled              |    |                   |
| White Leghorn (O*N/O*N)    | 30 | 0 0 30            |

*The “+” sign indicates a positive result of insertion, and the “−” denotes no insertion.

Materials and Methods

**Animals**

Two Chinese indigenous blue-shelled chicken breeds, Dongxiang and Lushi, and an American blue-shelled breed, Araucana, were used in the present study. Dongxiang chicken is from Dongxiang town, Jiangxi province of China. It is characterized by blue eggshell, single comb and black feather. Historically Dongxiang chicken is selected for blue eggshell, however, the trait has not been fixed to date. Lushi chicken is another local breed laying blue-shelled egg from Lushi town, Henan province of China. Because Lushi chicken has not been systematically bred, some appearance traits, eggshell color, as well as feather color does not show homogeneity. Araucana is an indigenous breed from Chile of South America. Besides blue-shelled egg, two distinguishing characteristics of Araucana breed are rumpless and tufts of feathers which protrude from each side of neck.

In the present study, Dongxiang chicken and Lushi chicken were collected from Jiangxi Hualv breed poultry conservation farm and Henan Sammexia Lushi chicken farm, respectively. The blood samples of Araucana were obtained from members of the Araucana Club of America. We also collected 9 non-blue-shelled chicken breeds including Red Jungle Fowl, White Leghorn, Rhode Island Red, Beijing You, Silkie, Tibetan, Luxi Game, Gushi and Dwarf (a commercial layer line in China).

A three-generation F₂ resource family was constructed by crossing homozygous blue-shelled Dongxiang (O*LC/O*LC) and white-shelled Araucana (O*LC/O*LC). The causative gene for blue eggshell was not found in the homologous region in duck (Figure S2, Primers in Table S4). Thus, the causative gene for blue eggshell in ducks may be different from that in chickens. Moreover, the genetic pattern in the chicken is also different from that for Japanese blue-eggshell quail which arise from a recessive mutation ce [7]. Because there is no record showing that the two ancestral species of domestic chickens, red jungle fowl and grey jungle fowl [33], lay blue eggs, we may conclude that the causative EAV-HP insertion for blue eggshell is a derived mutation in the domestic chicken.

China and Chile are two countries reported for having indigenous blue-shelled chicken breeds. Araucana from Chile, Dongxiang and Lushi from China got the blue eggshell phenotype and were all bred for several hundred years. Analysis with mtDNA showed both Indo-European and Asian origins of Chilean and Pacific chickens and blue/green-shell trait in the Araucana did not originated from ancient pacific/pre-Columbian chickens [43]. It is noted in the present study that though all these blue-shelled chickens had the EAV-HP insertion, the EAV-HP insertion into two different genomic sites in the 5' flanking region of SLC01B3 in the blue-shelled chickens from the two countries (Figure 3C) and the EAV-HP insertion in blue-shelled Araucana embedded in a haplotype which is distinctly different from the corresponding haplotype from blue-shelled Lushi and Dongxiang chickens (Table S3). Here, we provide unambiguous evidences that the genetic basis of blue shell phenotype in Araucana is different from that in Chinese blue-shelled breeds, indicating independent origins of the trait in different continents. Due to the blue eggshell mutation having been artificially selected for consumption and variable eggshell color types for human requirements, the separated insertion events present us another parallel evolution case at the molecular level under adaptive selection by humans.
males, which has been verified by a test cross, and brown-shelled Dongxiang (O*N/O*N) females. All F₁ hens laid blue shelled eggs and individual egg color phenotypes were recorded for all 146 F₂ hens.

Six heterozygous (O*LC/O*N) hens were produced by mating one homozygous Dongxiang blue-shelled (O*LC/O*LC) male and one White Leghorn (O*N/O*N) female, and the progeny were used for pyrosequencing analysis.

All animal research was approved by Beijing Administration Committee of Laboratory Animals under the leadership of the Beijing Association for Science and Technology, the approve ID is SYXX (Beijing) 2007–0023. DNA was extracted from blood using standard phenol/chloroform method. RNA was extracted from the liver and uterus. All the tissue samples for RNA isolation were collected at 3 to 5 hours before the next expected oviposition.

Linkage analysis
The F₂ resource family was used for linkage analysis. A set of 8 markers covering the region anchored by GWAS and SOX5, eJ were used in the linkage analysis (Figure 2). Marker L1 and L4–L7 were adopted from previous reports [18,20] and L2, L3 and L8 were mined from the chicken genome assembly (Build 2.1) at http://genome.ucsc.edu/cgi-bin/hgGateway. Fifteen SNP markers (L9–L23) between L4 and L5 were added to narrow the mapping region. Primers and genotyping methods for all markers were present in Table S1. CRI-MAP 2.4 was used for linkage analysis [44]. The TWO-POINT option was used to calculate the recombination fractions between loci as well as corresponding LOD-scores. The CHROMPIC option was used to find unlikely double recombinants.

Expression analysis
Total RNA was extracted from the uterus using Trizol reagent (TianGen, Dalian, China), followed by synthesis of cDNA from 2 μg of RNA using M-MLV reverse transcriptase (Promega, CA, USA). The cDNA probe 5'-AATCTGGGCTGAACGCATCT-3' was labeled by 6-FAM and were synthesized from mRNA of SLCO1B3 (GenBank accession No. G.67334934) by Boxing Bio-engineering Limited Company (Boxing, Guangdong, China). The in situ hybridization was then carried out according to the instruction of the FISH Detection Kit (Boxing, Guangdong, China). Imaging was performed using a fluorescence microscope equipped with vision software.

Resequencing of SLCO1B3
Twenty-four kilobases fragment (GenBank accession No. JN020139) covering the whole SLCO1B3 was resequenced using a panel of ten birds from 5 blue-shelled (O*LC/O*LC) and 5 brown-shelled (O*N/O*N) Dongxiang chickens. Seventeen primer pairs were designed to generate overlapping PCR amplicons ranging from approximately 800 bp to 2000 bp in size were listed in Table S5. The PCR amplifications were performed in a total volume of 50 μL containing 5 μL of 10×Taq polymerase buffer, 10 mmol of each deoxynucleotide triphosphate (dNTP), 20 pmol of each primer, 2.5 U Taq DNA polymerase (HT-biotech, Beijing, China), and 50 ng genomic DNA. All purified PCR products were directly sequenced in both directions using the same primers. The sequences were assembled and analyzed for polymorphisms using the ChromasPro 1.5 or BLAST program in UCSC (http://genome.ucsc.edu/cgi-bin/hgBlat?command=start). Related to the calibrator was calculated as described in Livak et al. [45].

Pyrosequencing
Tissues (the uterus and liver) were collected from the six blue-eggshell heterozygotes. Total RNA was extracted from the uterus and liver with trizol (Tiangen, Dalian, China). The RNA quality was controlled using NanoVue plus spectrophotometer (GE Healthcare, USA). The first-strand cDNA synthesis used M-MLV (Promega, CA, USA) with the 18 hexamers. A fragment containing the SNP (g. 67334934 G>T) in exon 5 was amplified with forward (CCTGTGGCGGAGATTGGT) and reverse (TCCCTTAGAAGAATTGTAAGATA) primers. The relative expression of the two allele (O*LC or O*N) transcripts in heterozygotes was scored by analyzing the SNP (g. 67334934 G>T) by pyrosequencing. A pyro-seq primer (CGTGAAGAGAATGGTG) was used as the sequencing primer and all steps were performed according to manufacturer’s protocol. All samples were analyzed in triplicates.

Fluorescent in situ hybridization
The uterus from a 60-week-old egg laying blue-shelled hen and a non-blue-shelled hen were collected and fixed in 4% paraformaldehyde in phosphate buffered saline (PBS) for 24 hours at room temperature. Fixed uterus was embedded in irrigation solution PBS for six hours to eliminate 4% paraformaldehyde. Then slides were dehydrated in increasing concentrations of ethanol (50%, 70%, 80%, 90%, 95% for 1.5 hours each and 100% for 2 hours) followed by transparentizing in two clearing agents respectively of xylene for 15 minutes each. After transparentizing, the slides were pretreated by the mixture of xylene and low melting paraffin for 30 minutes then were directly transferred into pure melting paraffin (36°C) twice for 3 hours each.

The cDNA probe 5’-AACCTGAGCTGAACGCATCT-3’ was used by 6-FAM and were synthesized from mRNA of SLCO1B3 (XM416418.2) by Boxing Bio-engineering Limited Company (Boxing, Guangdong, China). The 5’ and 3’ untranslated regions (UTR) of SLCO1B3 gene, RACE experiments were performed on 2 µg of total RNA extracted from the uterus of a homozygous blue-shelled (O*LC/O*LC) and a brown-shelled (O*N/O*N) Dongxiang chicken...
using 5′ and 3′-Full RACE Kit (Takara, Dalian, China), according to
the manufacturer’s instructions. 5′ and 3′ UTR of SLCO1B3
gene transcripts were amplified by nested PCR with gene specific
(Table S4) and adaptor primers (Table S4) for the first and second
amplifications of 5′ and 3′ UTR respectively. First and second
PCR amplifications were carried out in a 50 μL reaction volume
containing 20 pmol of each primer, 5 μL of 10× LA PCR buffer
(Mg2+ plus), 2.5 U of LA Taq (Takara, Dalian, China), 20 mM of
each dNTP and 1–2 μL of cDNA or 1st PCR product. RACE
products were cloned to pMD-18 vector (Takara, Dalian, China),
and then sequenced in both directions.

Long-range PCR
A long-range PCR amplification with 1B3_5F & 5R primer pair
(Table S5) was performed in volumes of 30 μL containing 5 μL of
10× LA PCR buffer (Mg2+ plus), 2.5 U of LA Taq (Takara, Dalian,
China), 20 mM of each dNTP, 20 pmol of each primer and
50 ng genomic DNA. The PCR condition was as follow: 94°C
for 3 min followed by 33 cycles of 94°C for 30 s, 58°C for 30 s,
72°C for 5 min, and a final extension at 72°C for 10 min. The
PCR product was completely sequenced using the other three
72°C for 3 min followed by 35 cycles of 94°C for 30 s, 57°C for 30 s and 54°C for 40 s, respectively, 72°C for 45 s, and a
final extension at 72°C for 10 min. The PCR products were
sequenced bidirectionally using the PCR primers.

Supporting Information

Figure S1 The sequencing results of 5′ RACE for SLCO1B3 in
blue-shelled Dongxiang chicken. Sequences showed in blue color
are newly obtained 5′ UTR of SLCO1B3 which has been
submitted to GenBank with accession No. JN381032. The
underlined sequences are transcription from EAV-HP insertion.
(TIF)

Figure S2 Expression analysis for SLCO1B3 in white-shelled and
blue-shelled ducks by RT-PCR.
(TIF)

Figure S3 Diagnostic genotyping test of EAV-HP insertion. (A)
primers information for diagnostic genotyping test of EAV-HP
insertion. (B) results for diagnostic genotyping test of EAV-HP
insertion in blue-shelled and non-blue-shelled Dongxiang chick-
ens. Single 425 bp band represents homozygous blue-shelled
chickens, single 340 bp band corresponds to non-blue-shelled
chickens, and two amplifications are heterozygous chickens.
(TIF)

Table S1 Information of markers used in linkage analysis.
(DOCX)

Table S2 Two point linkage analysis with markers in L4–L5
interval. (DOCX)

Table S3 Haplotype frequency analysis of all the breeds or
populations. (DOCX)

Table S4 Primers used in expression analysis, 5′ and 3′ RACE
of candidate genes. (DOCX)

Table S5 Primer sequences used in resequencing of SLCO1B3
and EAV-HP. (DOCX)

Table S6 Information of SNPs in SLCO1B3 gene. (DOCX)

Acknowledgments
We are grateful to Rosalyn Upson, XQ Zhang, XT Kang, JS Xu for
providing samples of the Araucana, Red jungle fowl, Lushi, and
Dongxiang chickens, and to L Andersson from Uppsala University, PB
Siegel from Virginia Tech, Zhi-Liang Hu from Iowa State University, and
Ning Li from China Agricultural University for their helpful comments.

Author Contributions
Conceived and designed the experiments: NY XD LQ CW. Performed the
experiments: ZW JY. Analyzed the data: ZW JY. Contributed reagents/
materials/analysis tools: XY GL YZ JI XW JB GX CW. Wrote the paper:
NY LQ XD ZW.

References
1. Underwood TJ, Sealy SG (2002) Adaptive significance of egg coloration. In:
Deeming DC, editors. Avian Incubation: Behaviour, Environment and
Evolution. Oxford University Press. pp. 280–298.

2. Moreno J, Osorno JL (2003) Avian egg colour and sexual selection: Does
eggshell pigmentation reflect female condition and genetic quality? Ecol Lett 6:
803–806.
3. Morales J, Sanz JJ, Moreno J (2006) Egg colour reflects the amount of yolk maternal antibodies and fledging success in a songbird. Bird Lett 2: 334–336.
4. Siefferman L, Navara KJ, Hill GE (2006) Egg coloration is correlated with female condition in eastern bluebirds (Sialia sialis). Behav Ecol Sociobiol 59: 651–658.
5. Morales J, Torres R, Velando A (2010) Parental conflict and blue egg coloration in a seabird. Naturwissenschaften 97: 173–180.
6. Moreno J, Morales J, Lobato E, Merino S, Tomas G, et al. (2005) Evidence for the signaling function of egg color in the pied flycatcher Ficedula hypoleuca. Behav Ecol 16: 931–937.
7. Ito S, Tsuzuki M, Komori M, Mizutani M (1993) Celadon: An eggshell color mutation in Japanese quail. J Hered 84: 145–147.
8. Pannett RC (1933) Genetic study in poultry-IX. The blue egg. Genetics 27: 465–470.
9. Wang CT, Wan TC, Pan CM, Chen YH (1997) Comparisons of physiological-chemical properties and alkalization process between greenish and whitish eggs of duck (Anas platyrhynchos). Acta Agron Chem 23: 265–272.
10. Lang MR, Wells JW (1987) A review of eggshell pigmentation. World Poult Sci J 43: 238–245.
11. Kennedy GY, Vevers HG (1973) Eggshell pigments of the Araucano fowl. Comp Biochem Physiol B 44: 11–23.
12. Brittner JH, Hunt PB (1935) Linkage of Pebb Com and Blue Egg in the Fowl. Science 90: 488.
13. Bitgood JG, Shoffner RN, Otis JS, Brides WE (1980) Mapping of the genes for pea comb, blue egg, barring, silver, and blood groups A, E, H, and F in the domestic fowl. Poult Sci 59: 1696–1699.
14. Bitgood JG, Otis JS, Shoffner RN (1983) Refined linkage value for comb and blue Egg: lack of effect of pea comb, blue egg, and naked on age at first egg in the domestic fowl. Poult Sci 62: 235–238.
15. Bitgood JG, Brides RW, Brides WE (2000) Further tests for genetic linkages of three morphological traits, three blood groups, and break points of two chromosome translocations on chromosome one in the chicken. Poult Sci 79: 295–295.
16. Barlow JR, Jones CP, Smith EJ (1996) Linkage analysis of endogenous viral element 1, blue eggshell, and pea comb loci in chickens. J Hered 87: 67–70.
17. Wright D, Beije H, Meadow JR, Bed’om B, Gourichon D, et al. (2009) Copy number variation in intron 1 of SOX5 causes the Pea-comb phenotype in chickens. PLoS Genet 5: e1000312.
18. Zhao R, Liu ZZ, Xu GY, Yang N (2007) Analysis of SNP markers for chicken blue-shelled gene using PCR-SSCP. Chin J Agric Biotech 4: 53–56.
19. Wang XT, Bai JR, Zhao CJ, Zhang H, Bao HG, et al. (2010) Localization of the gonochorism sequence interval for the blue eggshell gene using an F2 resource population of Dongxiang chickens. Brit Poult Sci 51: 507–509.
20. Wang ZP, Wang XT, Liu RF, Li JY, Deng XM (2010) Refined localization of the O gene for blue egg phenotype on chicken chromosome 1. J Anim Vet Adv 9: 2947–2950.
21. Wragg D, Mwacharo JM, Alcalde JA, Hocking PM, Hanotte O (2012) Analysis of genome-wide structure, diversity and fine mapping of Mendelian traits in traditional and village chickens. Heredity 109: 6–18.
22. Sacco MA, Flannery DM, Howes K, Venugopal K (2000) Venugopal, Avian endogenous retrovirus EAV-HP shares regions of identity with avian leukosis virus subgroup J and the avian retrotransposon ART-CH. J Virol 74: 1296–1306.
23. Reynolds SJ, Martin GR, Casery P (2009) Is sexual selection blurring the functional significance of eggshell coloration hypotheses? Annu Rev Env 78: 209–215.
24. Popovic M, Zaja R, Smital T (2010) Organic anion transporting polypeptides (OATP) in zebrafish (Danio rerio): Physiogenetic analysis and tissue distribution. Comp Biochem Physiol A 155: 327–335.
25. Hagenbuch B, Gui C (2008) Xenobiotic transporters of the human organic anion transporting polypeptides (OATP) family. Xenobiotica 38: 778–801.
26. Sanna S, Buccheri F, Maschio A, McArtic PF, Usala G, et al. (2009) Common variants in the SLC20A3 locus are associated with bilirubin levels and unconjugated hyperbilirubinemia. Hum Mol Genet 18: 2711–2718.
27. Gunnarsson U, Kerje S, Bed’om B, Sahliqvist AS, Ekkwall O, et al. (2010) The Dark brown plumage color in chickens is caused by an 8.3-kb deletion upstream of SOX10. Pigment Cell Melano J Res 24: 260–274.
28. Paillhois E, Viajar B, Chafiaux S, Served N, Taouret S, et al. (2001) A 11.7-kb deletion triggers intersexuality and poledness in goats. Nat Genet 29: 453–458.
29. Giuffra E, Tarustom A, Marldlund S, Bengcan-Rudloff E, Chardon P, et al. (2002) A large duplication associated with dominant white color in pigs originated by homologous recombination between LINE elements flanking KIT. Mamm Genome 13: 569–577.
30. Van Laere AS, Nguyen M, Braamschweig M, Nezer C, Collette C, et al. (2003) A regulatory mutation in IGFI2 causes a major QTL effect on muscle growth in the pig. Nature 425: 832–836.
31. Clop A, Marocq F, Takeda H, Pirotin D, Tordoir X, et al. (2006) A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep. Nat Genet 38: 813–818.
32. Karlsson EK, Baranowska I, Wade CM, Salmon Hiltz Jr, Zody MC, et al. (2007) Efficient mapping of mendelian traits in dogs through genome-wide association. Nat Genet 39: 1321–1328.
33. Eriksson J, Larson G, Gunnarsson U, Bed’om B, Tixier-Boichard M, et al. (2008) Identification of the yellow skin gene shows a hybrid origin of the domestic chicken. PLoS Genet 4: e1000110.
34. Roosungren Peilberg G, Godokvo A, Sundström E, Curik I, Lennartsson J, et al. (2000) A cis-acting regulatory mutation causes premature hair greying and susceptibility to melanoma in the horse. Nat Genet 40: 1004–1009.
35. Payne LN (1998) Retrovirus-induced disease in poultry. Poult Sci 77: 1204–1212.
36. Jenkins NA, Copeland NG, Taylor BA, Lee BK (1981) Dilute (d) coat colour mutation of DBA/2J mice is associated with the site of integration of an ectropic MuLV genome. Nature 293: 370–374.
37. Stove JP, Fennor S, Greenoak GE, Moran C, Coffin JM (1988) Role of endogenous retroviruses as mutagens: the hairless mutation of mice. Cell 54: 383–391.
38. Chang CM, Couville JL, Coquerelle G, Gourichon D, Oulmouden A, et al. (2006) Identification of the yellow skin gene reveals a hybrid origin of the domestic chicken. PLoS Genet 4: e1000010.
39. Bhatia DN, Jafri WS, Mehta KN, Abrego ZA, Gozzi K, et al. (2004) Mouse hair growth from the body surface to the corneum stratum depends on the hair follicle stem cell niche. Dev Biol 272: 273–286.
40. Bacon LD, Crittenden LB, Havenstein GB (1988) Association of the blue-shelled gene using PCR-SSCP. Chin J Agric Biotech 4: 53–56.
41. Stoye JP, Fenner S, Greensko GE, Moran C, Coffin JM (1988) Role of endogenous retroviruses as mutagens: the hairless mutation of mice. Cell 54: 383–391.