Alternative splicing contributes to the coordinated regulation of ferritin subunit levels in *Bactrocera dorsalis* (Hendel)

Xuan-Zhao Jiang, Lin Cong, Jin-Zhi Niu, Wei Dou & Jin-Jun Wang

A constant ratio of ferritin heavy chain homolog (HCH) and light chain homolog (LCH) subunits seems to be required to compose the ferritin heteropolymer protein in insects. However, the mechanism by which insect LCH genes regulate protein levels remains unclear. We report that alternative promoters and alternative splicing contribute to maintaining a constant ratio of the two subunits, BdFer1HCH and BdFer2LCH (ferritin 1 HCH and ferritin 2 LCH), in *Bactrocera dorsalis*, a notorious quarantine pest. The genes BdFer1HCH and BdFer2LCH were identified with a series of potential transcription factor binding sites and were shown to be clustered within the genome in a “head to head” fashion. Thus, we unearthed a potential post-transcriptional mechanism to regulate the levels of LCH subunits, and confirmed that the expressions of BdFer1HCH and BdFer2LCH were induced by 20-hydroecdysone, iron overload, and immune challenge.

Iron exhibits a dual nature in organisms, both as an indispensable micronutrient and a potent toxin. Organisms must intake adequate amounts of dietary iron for numerous physiological functions, such as respiration, nucleotide synthesis, steroid production and amino acid production. At the same time, organisms must avoid the potential toxic properties of iron, which can cause oxidative damage to biomolecules. To achieve iron homeostasis, organisms have evolved strategies to allow iron absorption and transport, but prevent the accumulation of redox-active iron in sensitive tissues. In insects, ferritins are iron-binding proteins, similar to the function of transferrins, and play key roles in iron transport and storage. Unlike mammalian ferritins, insect ferritins are mostly involved in the secretory pathways and circulate in the hemolymph, which is consistent with the fact that they are synthesized with a signal peptide, presumably directing the mature subunits to the secretory pathways. This kind of insect ferritin consists of two types of secretory subunits: a mammalian heavy chain homologue (HCH) that contains a ferroxidase center, and a mammalian light chain homologue (LCH) that facilitates the nucleation of the ferrihydrite iron core. Insect HCH and LCH genes are reported to be arranged “head to head” and transcribed in opposite directions. Such a structure is implicated in the coordinated regulation of ferritin at the transcriptional level, which is mediated by shared regulatory elements.

A post-transcriptional regulation mechanism mediated by the interaction of the iron responsive element (IRE) and iron regulatory protein (IRP) controls the synthesis of HCH subunits. IRE is a stem-loop structure at the 5’ untranslated region (UTR) of ferritin messages. A genomic survey of ferritin reveals that almost all insect HCH genes have an IRE, while most of the LCH genes lack an IRE, except in Lepidoptera. Recent research on *Drosophila melanogaster* ferritin (Fer1HCH and Fer2LCH) demonstrated that a constant ratio of the ferritin subunits is maintained in vivo; a protein degradation mechanism was proposed to regulate the level of the Fer2LCH subunit. However, such a post-transcriptional mechanism remains to be clarified in insects.

In addition to its role in iron homeostasis, insect ferritin may have multiple functions, such as in cell proliferation, in embryonic and post-embryonic development, as a cytotoxic protector against oxidative stress and thermal stress, and as a potential antimicrobial effector. Both subunits are critical for insect survival, and their deletion can lead to embryonic death and developmental arrest.

Given these differences from their mammalian homologues, insect ferritins are promising candidate targets for agricultural and medical pest control. *Bactrocera dorsalis* (Hendel) (Diptera: Tephritidae) is one of the most destructive and widespread polyphagous agricultural insect pests, feeding on more than 250 species of vegetables and fruits and causing severe economic losses each year. Severe resistance in *B. dorsalis* populations greatly compromises the effects of chemical insecticides, gravely threatening the effective and sustainable usage of these...
compounds. Therefore, new, effective strategies are urgently needed, and study of the ferritin of B. dorsalis may help to develop such strategies.

This study shows that the post-transcriptional mechanism, intron retention, is possibly responsible for the regulation of LCH subunit levels by inducing a premature stop codon in the LCH transcripts, which leads to the generation of truncated proteins. A classic IRE is also identified at the HCH 5' UTR of B. dorsalis, whose inclusion in the transcripts is coordinated by alternative transcription start sites (TSSs) and alternative splicing. Moreover, several potential functions for B. dorsalis ferritin are implicated by analyzing the genomic sequences and expression profiles resulting from abiotic and biotic stresses.

**Results**

**Characterization of genomic sequences and cDNAs of B. dorsalis ferritin.** BdFer1HCH (Ferritin 1 heavy chain homologue of B. dorsalis) and BdFer2LCH (Ferritin 2 light chain homologue of B. dorsalis) were clustered in the genome in a “head to head” fashion (Fig. 1 and Supplementary Fig. 1). Three TSSs, designated H-TSS1, H-TSS2, and H-TSS3, existed in the gene BdFer1HCH, among which H-TSS1 and H-TSS2 were identified by 5’ RACE, while the precise location of H-TSS3 was not determined. At least three primary transcripts were generated, which were further spliced into four types of mRNAs despite the presence of alternative polyadenylation sites (Fig. 1). An alternatively spliced exon named exon 3 was identified at the 5’ UTR between H-TSS1 and H-TSS2 (Fig. 1 and Supplementary Fig. 1). A consensus sequence, acting as IRE required for post-transcriptional regulation, was identified in exon 3. Furthermore, two potential RNA editing sites, including a U to C and a C to U, were identified in the IRE (Supplementary Fig. 2A) by sequencing multiple clones. Several potential polyadenylation sites were predicted, whereas only one for BdFer1HCH was confirmed by 3’ RACE. However, all BdFer1HCH mRNAs encoded a 204-residue polypeptide (GenBank accession number: KF499586) with a putative membrane targeting signal of 17 amino acid residues at the N-terminus (Supplementary Fig. 1). The BdFer1HCH subunit preserved all seven conserved amino acid residues forming the ferroxidase center (Supplementary Fig. 1). One putative N-linked glycosylation site (N-X-S/T) was present at its N-terminus (Supplementary Fig. 1 and 2B).

For BdFer2LCH, two types of primary transcripts were detected; they resulted from two alternative TSSs, termed L-TSS1 and L-TSS2 (Fig. 1). L-TSS1 was detected by 5’ RACE, while the exact location of L-TSS2 was not determined. No potential IRE was found at the 5’ UTR of BdFer2LCH. Of interest is that an alternatively spliced exon C, generated by intron retention, was found in the coding region (Fig. 1). Nevertheless, a premature stop codon was introduced when the alternative exon C was spliced into the transcripts. Therefore, at least three types of mRNAs were generated because of the alternative TSSs and alternative splicing. Additionally, at least two polyadenylation sites existed for BdFer2LCH as shown in Supplementary Fig. 1. The ORF of BdFer2LCH, excluding exon C (GenBank accession number: KF499587), encoded a 223-amino-acid polypeptide. The
nucleotide regions upstream of the start codons) of BdFer1HCH and BdFer2LCH to identify putative Nuclear Factor-kappaB (NF-kappaB) and Broad-Complex (BR-C) binding sites as well as BdFer2LCH lacked most of the amino acids forming the ferroxidase center (Supplementary Fig. 1 and 2B). A putative N-linked glycosylation site was present at its N-terminus. The truncated protein was composed of 39 amino acid residues.

Analyses of the upstream regions for putative transcription factor binding sites (TFBSs). We analyzed the control regions (2000-nucleotide regions upstream of the start codons) of BdFer1HCH and BdFer2LCH to identify putative Nuclear Factor-kappaB (NF-kappaB) and Broad-Complex (BR-C) binding sites as well as putative TFBSs located within the ±50-nucleotide regions around the TSSs. The results showed that no NF-kB was found within the sequences between the start codons of BdFer1HCH and BdFer2LCH. Consequently, we further analyzed the first intron sequence downstream from the start codon, revealing the presence of three NF-kappaBs (Supplementary Fig. 1). A total of 11 putative BR-C binding sites were predicted, including two BR-C Z1, four BR-C Z3, and five BR-C Z4 binding sites (Supplementary Fig. 1). The TFBSs within the ±50-nucleotide regions around the TSSs, which might be part of the core promoter, were predicted by AliBaba 2.1. Nine putative TFBSs were identified along H-TSS1, including C/EBP (CCAAT/enhancer-binding protein), NF-1 (Neurofibromatosis 1), Pit-1a (pituitary-specific positive transcription factor-1a), HNF-1 (hepaticoyte nuclear factor-1), GR (Glucocorticoid receptor), Oct-1 (octamer binding protein 1), Sp1 (specificity protein 1), and TBP (TATA-binding protein) sites (Supplementary Fig. 1). The initiator element (Inr) overlapping the H-TSS1 was 5′-TCA+1GATC-3′, in agreement with the consensus sequence 5′-PyPyA+nNT/PyPy-3′. A canonical TATA box, 5′-TATAAGG-3′ (consensus sequence: 5′-TATAAAWAW-3′) was located 27 nucleotides upstream from the H-TSS1. Eight TFBSs were identified along H-TSS2, including Sp1, CREB (cAMP response element-binding protein), Oct-1, AP-1 (activator protein 1), TBP, C/EBP, and GATA-1 (GATA-binding protein 1). The corresponding Inr sequence was 5′-TCA+1GTAG-3′, where a purine A+n replaced the pyrimidine. At position −26 to −33 nucleotides upstream from the H-TSS2 lay a non-canonical TATA box (5′-TTAATAA-3′) (Supplementary Fig. 1). Six putative TFBSs were identified around L-TSS1, including C/EBP, ICSBP (interferon consensus sequence binding protein), Oct-1, TBP, GATA-1, and HNF-3 (hepaticoyte nuclear factor 3) (Supplementary Fig. 1). A non-canonical Inr (5′-TCA+nTTGAA-3′) corresponded to L-TSS1, and the upstream sequence 5′-TATAAGG-3′ (−23 to −29 nucleotides) might be recognized as a non-canonical TATA box. Additionally, a TATA box consensus sequence (5′-TATAAAAA-3′) was present downstream at position +26 to +33 nucleotides.

Figure 2 | Tissue transcription profiling of Bactrocera dorsalis ferritin in larvae. RT-qPCR analysis of the relative expression levels in five larval tissues: the fat body (FB), integument (IN), midgut (MG), Malpighian tubules (MT), and trachea (TR). All values are normalized relative to the expression levels in trachea (set as 1). Data are mean 2^−ΔΔCt values of three independent biological replicates ±SEM. Differences were analyzed by one-way ANOVA. In all cases, bars denoted with different letters had significant differences (p < 0.05, Tukey HSD’s test).

Tissue and developmental transcription profiling of B. dorsalis ferritin. The primers used in RT-qPCR and RT-PCR analysis for BdFer1HCH and BdFer2LCH were designed to anneal to the common regions so that we could detect the abundances of all variants. The expression of BdFer1HCH was detectable in all tested larval and adult tissues (Fig. 2 and Fig. 3A). The highest expression was observed in the midgut, followed by the fat body, Malpighian tubules, integument, and trachea in larvae (Fig. 2). In adults, relatively high expression was also detected in the midgut and fat body, while relatively low expression was observed in the ovary (Fig. 3A). Several variants of both BdFer1HCH and BdFer2LCH were also detected in adult tissues. The product H3 was amplified using a sense primer located upstream from H-TSS2 and an antisense primer located downstream of the start codon; it represented the abundance of the variant initiating with H-TSS3. When the transcripts initiated with H-TSS3, the sequencing results indicated that intron 1, exon 2, exon 3 and intron 2 were spliced out as parts of an intron (Fig. 1). The variants initiating with H-TSS3 had relatively low expression in all nine tested adult tissues. The product H2 represented the variants initiating with H-TSS2, which were amplified using a sense primer located between H-TSS1 and H-TSS2 and an antisense primer located downstream of the start codon. Two variants were observed that initiated with H-TSS2 and were due to alternative splicing (Fig. 3A). Further sequencing analysis revealed that exon 3 was alternatively spliced. The two variants also had a wide distribution in all tested adult tissues (Fig. 3A). The midgut and ovary contained a high abundance of the variant with exon 3 included, while other tissues seemed to have a similar distribution of the two variants (Fig. 3A). The developmental expression profiling showed that BdFer1HCH had the lowest expression in eggs, the highest in larvae, and a wide distribution in all stages (Supplementary Fig. 3). The splicing variants seemed to have stage-specific expression. In eggs, only the variant including exon 3 was observed. In the larvae and both male and female adults, the variant including exon 3 was the main mRNA type, whereas in pupae the expression levels of both variants were similar (Fig. 3C).

BdFer2LCH was very similar to BdFer1HCH in its tissue and developmental expression profiling; however, the expression of BdFer2LCH dominated at the adult stage (Supplementary Fig. 3). Three variants were observed for BdFer2LCH, as illustrated by the products, BdFer2LCH and L2 (Fig. 3A). L2 was amplified using a sense primer upstream of L-TSS1 and an antisense primer downstream of the start codon to selectively target the variant initiating with L-TSS2. This variant had relatively low expression in all adult tissues except the head (Fig. 3A). The product, BdFer2LCH, which was amplified using a sense primer downstream from L-TSS1 and an antisense primer near the stop codon, had two variants, either with or without exon C as a result of alternative splicing (Fig. 3A). The variant including exon C seemed to mainly express in the ovary. To better discern the splicing variants, we amplified the product BdFer2LCH’ (Fig. 3B), which was part of BdFer2LCH but shorter, and spanned the region including exon C. Both variants were observed in all adult tissues, but the one excluding exon C dominated in all tissues (Fig. 3B). The variant
including exon C was also detected in all developmental stages, but the larvae seemed to mainly express the variant without exon C (Fig. 3C). To check whether the alternative splicing occurred when the transcript initiated with L-TSS2, the PCR product L2’ was generated using a sense primer upstream from L-TSS1 and an antisense primer downstream of exon C. Only a single product excluding exon C was amplified (Fig. 3C), which was further confirmed by sequencing multiple clones. Thus, exon C was induced only when the transcripts initiated with L-TSS1.

Iron overload dramatically altered the expression of *B. dorsalis* ferritin. The expressions of both *BdFer1HCH* and *BdFer2LCH* increased significantly 12, 24, and 36 h after the insect began feeding on an artificial diet supplemented with 1.8 mg/g FeCl₃, and the expressions of both genes peaked 24 h after feeding (Fig. 4A and B). Moreover, larvae feeding on a diet supplemented with 1.8 mg/g FeCl₃ for 24 h gained significantly less weight than those eating diets supplemented with 0 mg/g and 0.18 mg/g FeCl₃ (Fig. 5A). However, both 0.18 mg/g and 1.8 mg/g FeCl₃ could significantly up-regulate the expressions of *BdFer1HCH* and *BdFer2LCH*, while no apparent dosage-effect was observed between the 0.18 mg/g and 1.8 mg/g treatments (Fig. 5B).

We further estimated the effects of iron overload on the expressions of *BdFer1HCH* and *BdFer2LCH* in the larval midgut and in the carcass without the midgut, because the midgut is the main organ responsible for iron absorption. The results showed that the expressions of *BdFer1HCH* and *BdFer2LCH* both increased in the larval midgut and carcass without midgut after feeding on the diet supplemented with FeCl₃ for 24 h (Fig. 5D). The variant excluding exon 3 was mainly detected in the carcass without midgut, and the variant...
including exon 3 was the main mRNA type in the midgut and in the carcass without midgut (Fig. 5C). Iron overload up-regulated the expression of the variant including exon 3, but had no apparent effect on the expression of the variant excluding exon 3. Similar expression levels in the midgut and carcass without midgut were observed for the variant initiated with H-TSS3, and the expression level increased in the midgut and carcass without midgut after iron overload (Fig. 5C). The transcript initiating with L-TSS2 represented by the product L2 was mainly expressed in the larval midgut. However, the expression level of the transcript initiating with L-TSS2 generally increased with an increasing dosage of iron (Fig. 5D).

**Upregulation of B. dorsalis ferritin by exogenous 20E.** BdFer1HCH and BdFer2LCH mRNAs were significantly increased after injection of 20-hydroxyecdysone (20E) at all four tested time points (Fig. 4C and D). The mRNA levels of BdFer1HCH had increased by more than 7-fold 1 h post injection. A more than 9-fold increase was observed for BdFer2LCH at 1 h post injection, and a greater increase up to 16-fold, had occurred 12 h post injection. The transcription patterns of BdFer1HCH and BdFer2LCH were very similar, but the response of BdFer2LCH was much more intense than that of BdFer1HCH.

**Repressed expression of B. dorsalis ferritin by E. coli, PGN-SA, and PGN-SB.** The antimicrobial peptide gene dipterin is sensitive to immune challenge. To ensure the effectiveness of the immune challenge, the expression patterns of the dipterin gene (GenBank accession number: KJ488999) of B. dorsalis were used as an indicator. The expression of dipterin was up-regulated by injecting sterilized saline (insect physiological saline; 150 mM NaCl, 5 mM KCl, pH 6.8) or sterilized water and further induced by E. coli, PGN-SA (peptidoglycan from Staphylococcus aureus), and PGN-EB (peptidoglycan from E. coli 0111:B4), which suggested that an immune reaction was generated in B. dorsalis (Fig. 6A and D). Additionally, dipterin was induced more by E. coli and PGN-EB than by PGN-SA at 9 h post treatment (Fig. 6A and D). Therefore, these results validate the subsequent analysis of B. dorsalis ferritin.

BdFer1HCH was significantly induced by injection of saline compared with the control group at 6 h and 9 h post treatment, while E. coli had a slightly weaker effect than saline (Fig. 6B). BdFer1HCH was also up-regulated, compared with the control group, by injection of sterile water at 9 h post treatment, but was significantly repressed by PGN-SA and PGN-EB at 9 h compared with the water group (Fig. 6E). BdFer2LCH was also induced by sterilized saline in contrast to the control group at 3 h and 9 h, and was significantly depressed by E. coli at 9 h compared with that in the saline group (Fig. 6C). The expression of BdFer2LCH was significantly decreased by PGN-SA and PGN-EB at 3 h compared with that in the water group and was sustained at a relatively low level at 6 h and 9 h post injection (Fig. 6F).

**Discussion**

HCH and LCH are characterized by post-transcriptional regulatory mechanisms that are required to facilitate the balancing of subunit
levels. An IRE/IRP system regulates the expression of the Fer1HCH subunit in *D. melanogaster*, whereas the regulatory mechanism for Fer2LCH subunit remains unknown. Additionally, studies on *D. melanogaster* demonstrated that a constant ratio of Fer1HCH and Fer2LCH subunits was required for ferritin expression, and was secured by an unknown post-transcriptional mechanism, which may be conserved within dipterans. Here, for the first time, we present a possible mechanism, named intron retention, to regulate the level of LCH subunits in *B. dorsalis*. The LCH variant including exon C, which is generated by intron retention, was detected in all developmental stages of *B. dorsalis*. The inclusion of exon C introduces a premature stop codon into the transcripts and leads to the production of truncated proteins. Nonetheless, whether such truncated proteins are functional remains to be determined.

Alternative splicing events have also been reported in LCH genes, but only at the 5' UTR. Additionally, the exclusion of exon C only occurs when the transcripts initiate with L-TSS1 but not L-TSS2, indicating coordinating regulation by transcriptional and post-transcriptional mechanisms. Thus, alternative promoters and alternative splicing cooperate to control the levels of the BdFer2LCH subunit.

Intriguingly, like Fer1HCH, BdFer1HCH also gained an IRE in an alternatively spliced exon at the 5' UTR. However, the inclusion of exon 3 only occurs when the transcripts initiate with H-TSS2. Thus, BdFer1HCH levels are jointly regulated by alternative promoters and alternative splicing, which has not previously been reported in other insect homologues. Based on the highly similar genomic structures among ferritin genes, similar mechanisms may also exist in other insects.

The tight clustering of the secreted ferritin genes allows for their coordinated regulation at their transcriptional levels and accounted, to some extent, for their similar transcriptional profiles. Such a clustering structure is also conserved in *B. dorsalis* ferritin genes, which may explain why the tissue and developmental expression profiles of BdFer1HCH and BdFer2LCH are quite similar. The genetic structures of BdFer1HCH and BdFer2LCH are most similar to that of *D. melanogaster* ferritin. Analysis of the control region implies that BdFer1HCH and BdFer2LCH may share some regulatory elements. Despite the transcriptional similarity between BdFer1HCH and BdFer2LCH, BdFer2LCH has distinctly higher expression in female adults than in other developmental stages, unlike...
BdFer1HCH. Such variation may imply an extra role for BdFer2LCH in female adults, which remains to be further characterized.

Iron overload can dramatically alter the abundance of ferritin transcripts and alter the mRNA types. Here, we confirmed that the transcript abundance was dramatically altered by iron overload. A high dose of FeCl₃ can significantly retard the growth of *B. dorsalis* larvae, while low doses may cause small amounts of weight gain. Interestingly, although the expressions of both *BdFer1HCH* and *BdFer2LCH* were up-regulated by iron overload, no apparent dose effect was observed, which may be explained by the high storage capacity of the ferritin protein and reduced absorption of iron in the midgut. Expression was up-regulated not only in the midgut, but also in other tissues, suggesting a role for ferritins in iron transport and iron detoxification. Previously, a study on *Fer1HCH* indicated that a functional IRE was present only at the 5' UTR of certain splice variants, which were preferentially encoded under iron-limited conditions. By contrast, in *B. dorsalis*, the transcripts containing IRE dominated at all developmental stages and in all tissues, and were preferentially up-regulated by iron overload. Moreover, in the larval midgut, the variants including IRE were the main mRNA types. Hence, we can infer that *BdFer1HCH* may prefer an adjustable form, which may allow *B. dorsalis* to adjust to changing iron conditions. However, more research is needed to explain the differences among ferritins in different insect species.

Figure 6 | Effects of *Escherichia coli*, PGN-SA, and PGN-EB challenge on the mRNA levels of antimicrobial peptide gene *diptericin* (A, D), *BdFer1HCH* (B, E), and *BdFer2LCH* (C, F). (A, B, C) Relative expression levels of *diptericin*, *BdFer1HCH*, and *BdFer2LCH* after *E. coli* challenge. All values were normalized relative to the control group (set as 1). Data represent mean 2^ΔΔCt values of three independent biological replicates ± SEM. (D, E, F) Relative expression levels of *diptericin*, *BdFer1HCH*, and *BdFer2LCH* after PGN-SA and PGN-EB challenge. Results are represented as the mean 2^ΔΔCt values of three independent biological replicates ± SEM. The two-tailed, unpaired t-test was used to test significance. Asterisks indicate significant differences in relative expression. (*, p < 0.05; **, p < 0.01).
The BR-C gene is a primary response gene in the ecdysoid pathway, and encodes four BR-C isoforms (Z1–Z4) in D. melanogaster. These isoforms are involved in a number of insect physiological processes, such as metamorphosis, oogenesis, and vitellogenesis. Accordingly, the distribution of BR-C binding elements in the upstream control region is evidence that B. dorsalis ferritin may be responsive to 20E. The immediate induction of BdFer1HCH and BdFer2LCH by 20E injection corroborates our hypothesis that the expression of B. dorsalis ferritin is regulated by hormone level. Additionally, the similar expression patterns of BdFer1HCH and BdFer2LCH under 20E regulation may be attributed to the shared 20E-responsive transcription factor binding elements in the control region. The regulation of ferritin expression by 20E has never been reported, whereas ecdysoides have been demonstrated to regulate the expression of the Apis mellifera transferrin gene. Hence, ecdysoides can regulate free iron levels by regulating the expression of iron-binding proteins, which further suggests that a certain iron level is required for hormone regulation. Additionally, some components in the ecdysoidogenic and the signaling pathways require iron, so their protein activities are subject to the iron level. Accordingly, 20E is probably capable of regulating its own level and its signal transduction by controlling the iron level.

Significant transcriptional regulation of ferritins by bacterial infection has been reported in insects. However, the actual regulation mechanisms by which bacterial infection regulates ferritin remain unclear. The host recognizes the bacteria by sensing specific forms of Peptidoglycan (PGN) via peptidoglycan recognition proteins (PGRPs). Differential immune effects are stimulated by Gram-negative and Gram-positive bacteria because of the differences in their PGNs. The antimicrobial peptide (AMP) diptericin is very active against Gram-negative bacteria and is generated by the IMD pathway. The identification of putative NF-κB binding sites in the first intron sequences downstream from the start codon indicated that B. dorsalis ferritin may also participate in immune defense. Thus, we examined the transcriptional responses of ferritins to E. coli, PGN-SA, and PGN-EB.

Significant up-regulation of the dipterericin gene was observed post-treatment, demonstrating an acute immune reaction of B. dorsalis to E. coli, PGN-SA, and PGN-EB. Additionally, the more intense induction by E. coli and PGN-EB than by PGN-SA confirmed that the IMD pathway is activated by Gram-negative bacteria or DAP-type peptidoglycans. However, both BdFer1HCH and BdFer2LCH were downregulated by E. coli, PGN-SA, and PGN-EB exposure, as in Anopheles gambiæ. Apparently, two distinct strategies, iron-withholding and the production of reactive oxygen species (ROS), are used by different insects to defend against microorganisms by regulating the free iron levels. Iron is an essential nutrient for pathogens, and iron-withholding is an innate immune mechanism in hosts. Hence, ferritins may act as immune effectors by sequestering extra iron to limit its availability to the pathogens. Meanwhile, excessive free iron can also lead to the production of ROS, which are harmful to microorganisms; therefore, the host can increase the amount of free iron by reducing the expression levels of ferritins. This strategy may work in B. dorsalis to resist invading microorganisms, because ferritin expression was repressed by E. coli, PGN-SA, and PGN-EB. However, more evidence is needed to determine whether such defensive strategies are really used by the host insects.

This study provides an initial understanding of a regulatory mechanism used by B. dorsalis to control the LCH subunit levels through generating truncated proteins, which result from the coordination of alternative promoters and intron retention. The presence of IRE at the 5' UTR of BdFer1HCH could also be attributed to the coordination of alternative promoters and alternative splicing; the variants, including IRE, were the major mRNA type and induced by iron overload, which is contrary to the case in D. melanogaster.

Additionally, we also show that B. dorsalis ferritin could be induced by 20E and immune challenges.

**Methods**

**Insect rearing and sample preparation.** The stock colony of B. dorsalis was originally collected from Fujian Province, China. Insects were maintained under constant temperature at 27 ± 1°C, 75 ± 5% relative humidity, and a 14:10 h light:dark. Insects were reared as previously described.

Samples including eggs, 1–8-day-old larvae, 1–10-day-old pupae, and 1–19-day-old female adults were collected at 3-day intervals. In the tissue-specific experiment, 5–6-day-old larvae and 5–6-day-old adults were used for tissue isolation. The insects were dissected individually in saline solution using dissecting needles under a microscope (Olympus SZX12, Tokyo, Japan). Larval tissues, including the fat body, integument, midgut, Malpighian tubules, and trachea, and adult tissues including head (without antennae), thorax (without legs), antennae, legs, fat body, midgut, Malpighian tubules, ovary, and testis, were collected. All tissue samples were stored in RNAstore Reagent (Tiangen Biotech, Beijing, China) at –80°C until used for RNA extraction. Three replicates were used per sample.

**Iron overload, 20E injection, and immune challenge.** Three-day-old larvae with similar body sizes were starved overnight. Then, three sets of 45 individual larvae were transferred into a Petri dish (diameter: 6 cm) filled with 19 g artificial diet supplemented with FeCl3 (dissolved in 1 mL saline) at a final concentration of 1.8 mg/g or 0.18 mg/g. For the control group, 45 larvae were transferred onto a Petri dish filled with 19 g artificial diet supplemented with 1 mL saline only. Ten larvae were randomly sampled from each group after 24 h and weighed. To analyze the expression profile, five larvae were sampled from each group 12, 24, and 36 h after the beginning of feeding. The larval midgut and the remaining carcass at 24 h were dissected individually in saline and stored in RNAlater Reagent. About 5 mg of midgut were collected as one sample. The carcasses without midgut from five individuals were used as one sample.

Two milligrams of 20E (FlukaChem Ltd, Glossop, UK) were diluted into 400 μL distilled deionized (DD) water at a final concentration of 5 μg/mL. Five-day-old larvae were injected at a dose of 0.5 μg/larva using a nano-injector (Nanoject II, Drummond Scientific, Broomall, PA, USA) as the treatment group. The control group was injected with an equivalent volume of DD water. Five individual larvae were randomly sampled from each group 1, 4, 8, and 24 h post injection.

For the immune challenge, E. coli (TransZychemecentially competent cells) was pre-cultured in LB medium. Stationary-phase cells were collected by centrifugation at 6000 × g for 1 min, washed three times, and finally suspended in sterilized saline solution. Four-day-old larvae were injected with 69 μL bacteri solution (OD590 = 2.1) through the inter-segmental membrane for the E. coli group, or injected with 69 μL sterilized saline for the group. The control group was handled the same way but without injection. Five individual larvae were randomly collected from each group 3, 6, and 9 h post injection. PGN-SA (peptidoglycan from S. aureus; InvivoGen, San Diego, CA, USA) and PGN-EB (peptidoglycan from E. coli 0111:B4; InvivoGen, San Diego, CA, USA) were diluted into sterilized water at final concentrations of 0.05 mg/mL. Four-day-old larvae were injected with 2 × 69 μL PGN-SA solution, PGN-EB solution or sterilized water. The control group was handled the same way but without injection. Five individual larvae were randomly collected from each group 3, 6, and 9 h post injection.

Each of the above-mentioned treatments was performed in triplicate. Three replicates were collected for each sample. All samples were stored at –80°C until used for total RNA extraction.

**RNA isolation and cDNA synthesis.** Total RNA from tissues (5 mg) was isolated using RNaseasy® Plus Micro Kit (Qiagen GmbH, Hilden, Germany) according to the manufacturer’s instructions. Total RNA from other samples was extracted using TRIzol® reagent (Invitrogen, Carlsbad, CA, USA), with an additional digestion step using RQI RNAase-Free DNase (Promega, Madison, WI, USA) to digest the genomic DNA. The quality, purity, and concentration of RNA were measured at 260 nm and 280 nm using a NanoVue UV–Vis spectrophotometer (Bio-sciences, Uppsala, Sweden), and RNA integrity was checked by 1.0% agarose gel electrophoresis. For first strand cDNA synthesis, 500 ng of total RNA was reverse transcribed using the PrimerScript 1st Strand cDNA Synthesis Kit (Takara, Dalian, China) with random hexamer and oligo (dT) primers, according to the manufacturer’s instructions.

cDNA and genomic sequence cloning. Based on the transcriptomic data of the B. dorsalis larval midgut, the ORFs were identified. To determine the UTRs, rapid amplification of cDNA ends (RACE) was performed using the SMARTer RACE cDNA Amplification Kit (Clontech, Mountain View, CA, USA) according to the manufacturer’s instructions. To determine the genetic structure, B. dorsalis genomic DNA was extracted from a female adult using a Tissue/Cell gDNA Mini Kit (Watson Biotechnologies, Shanghai, China) following the manufacturer’s protocols. All PCRs were carried out in a C1000TM Thermal Cycler (Bio-Rad, Hercules, CA, USA) in 25 μL reaction mixtures containing 1 μL template cDNA/genomic DNA, 2.5 μL 10 × PCR buffer (Mg2+ free), 2.5 μL Mg2+ (25 mM), 2.0 μL dNTPs (2.5 mM each), 1 μL of each primer (10 μM), and 0.25 μL rTaq (5 U/μL) (TaKaRa). PCR was carried out as follows: one cycle at 94°C for 1 min; 30 cycles at 94°C for 30 s, 60°C for 30 s, and 72°C for 1 min; one cycle of 72°C for 5 min and holding at 16°C. All primers are
listed in Supplementary Table 1. All PCR products were cloned into the pGEM-T Easy vector (Promega) and transformed into Trans5™ chemically competent cells (Transgen, Beijing, China). Three clones for each PCR product were randomly selected for sequencing (BGI, Shenzhen, China).

Sequence analysis. The SignalP 4.1 server was used to predict the signal peptide. N-Glycosylation sites were predicted by the NetGlyc 1.0 server. To identify putative TFBs in the control regions, we used the program Alababi 2.1 provided by TESS (Transcription Element Search System) (http://www.gene-regulation.com/pub/programs/alabab2/index.html) and an online server TFSEARCH [Searching Transcription Factor Binding Sites (ver. 1.3)] (http://www.cbrc.jp/research/db/TFSEARCH.html).

Reverse transcriptase polymerase chain reactions (RT-PCR) analysis. RT-PCR was performed to indicate the relative expression levels of the variants. PCR conditions and procedures were as described above. Primers are listed in Supplementary Table 2. Each PCR product (10 μl) was amplified by 2.0% agarose gel electrophoresis with GelView II (Solarbio Co., Ltd., Shanghai, China) staining. The gel was imaged using a Gel Doc XR System (Bio-Rad) and analyzed using Quantity One software (version 4.6.7; Bio-Rad). All electrophoresis images were refined using Adobe Photoshop CS5 (Adobe Systems, San Jose, CA, USA). Three replicates were performed for each trial.

Real-time quantitative polymerase chain reactions (RT-qPCR). The efficiency and determination coefficient (R²) of PCR amplification were calculated from the slope of the standard curve derived from five serial dilutions of one cDNA sample. Primers are listed in Supplementary Table 3. Each reaction was performed within the StepOne Plus Real-Time PCR System (Applied Biosystems, Carlsbad, CA, USA) in a 20 μl volume containing 1 μl cDNA template, 10 μl iQ SYBR Green Supermix (Bio-Rad), 1 μl of each forward primer (10 μM), and 6.8 μl of nuclease-free water. The reaction conditions were one cycle at 95 °C for 2 min followed by 40 cycles at 95 °C for 15 s and 60 °C for 30 s. Melting curve analysis from 60 °C to 95 °C was applied to all reactions to ensure specificity and consistency of all generated products.

Validation of stable reference genes under iron overload and bacterial challenge. Validation of candidate reference genes (Supplementary Table 4) under iron overload and immune challenge conditions. The primers were derived from a previous study (Supplementary Table 5). All data were analyzed by the tool RefFinder (http://www.leonxie.com/referencegene.php), which integrated the current major computational programs including geNorm4, NormFinder4, BestKeeper4, and the comparative Delta Ct method4.

Data analysis. Relative expression levels were determined using the 2−ΔΔCt method5. The relative expression levels at various developmental stages were normalized to that of 18S ribosomal RNA6, α-tubulin was used to normalize the data of tissue expression7 and expression under 20% treatment8. Expression differences were determined using one-way analysis of variance (ANOVA) or two-tailed, unpaired t-test within SPSS 20.0 (IBM Corp, Chicago, IL, USA). All pairwise comparisons were considered to be significant when p < 0.05. All experiments were conducted in triplicate, and data for a given variable were analyzed at the same time.

1. Andrews, N. C. Forging a field: the golden age of iron biology. Blood 112, 219–230 (2008).
2. Hower, V. et al. A general map of iron metabolism and tissue-specific subnetworks. Mol Biosyst 5, 422–443 (2009).
3. Kell, D. B. Iron behaving badly: inappropriate iron chelation as a major contributor to the aetiology of vascular and other progressive inflammatory and degenerative diseases. BMC Med Genomics 2, 2 (2009).
4. Weinberg, E. D. & Miklósy, J. Iron withholding: A defense against disease. J. Alzheimer Dis. 13, 451–463 (2008).
5. Nichol, H., Law, J. H. & Winzeler, J. I. Iron metabolism in insects. Adv. Insect Mol Biol 14, 263–270 (2005).
6. Serrano, A., et al. Heme catabolism in invertebrates. Curr. Opin. Chem. Biol. 14, 414–420 (2010).
7. Yang, L. et al. Ferritin: functional analysis of a conserved iron storage protein in the whitefly Bemisia tabaci. Insect Biochem Mol Biol 43, 352–365 (2013).
42. Vandesompele, J. et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 3, research0034 (2002).

43. Andersen, C. L., Jensen, J. L. & Ørntoft, T. F. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res* 64, 5245–5250 (2004).

44. Pfaffl, M. W., Tichopad, A., Prgomet, C. & Neuvians, T. P. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper–Excel-based tool using pair-wise correlations. *Biotech Lett* 26, 509–515 (2004).

45. Silver, N., Best, S., Jiang, J. & Thein, S. L. Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR. *BMC Mol Biol* 7, 33 (2006).

46. Livak, K. J. & Schmittgen, T. D. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2^(-ΔΔCt) Method. *Methods* 25, 402–408 (2001).

47. Shen, G. M., Wang, X. N., Dou, W. & Wang, J. J. Biochemical and molecular characterisation of acetylcholinesterase in four field populations of *Bactrocera dorsalis* (Hendel)(Diptera: Tephritidae). *Pest Manag Sci* 68, 1553–1563 (2012).

48. Shen, G. M., Jiang, H. B., Wang, X. N. & Wang, J. J. Evaluation of endogenous references for gene expression profiling in different tissues of the oriental fruit fly *Bactrocera dorsalis* (Diptera: Tephritidae). *BMC Mol Biol* 11, 76 (2010).

49. Yang, W. J., Xu, K. K., Cong, L. & Wang, J. J. Identification, mRNA Expression, and functional analysis of Chitin Synthase I gene and its two alternative splicing variants in oriental fruit fly, *Bactrocera dorsalis*. *Int J Biol Sci* 9, 331 (2013).

**Acknowledgments**

We thank Dr Ningjia He (Southwest University, Chongqing, China) for her valuable suggestions for this manuscript. This work was supported in part the Special Fund for Agro-scientific Research in the Public Interest (201203038), the Natural Science Foundation of Chongqing (CSTC, 2013jB0176), the earmarked fund for Modern Agro-industry (Citrus) Technology Research System, and the Fundamental Research Funds for the Central Universities (XDJK2014D025, XDJK2013A017) of China.

**Author contributions**

J.J.W. and X.Z.J. designed this project and experiments. X.Z.J. and L.C. performed the experiments. X.Z.J., J.Z.N. and W.D. completed the figures and tables. X.Z.J. and J.J.W. analyzed the data and wrote the manuscript.

**Additional information**

Supplementary information accompanies this paper at http://www.nature.com/scientificreports

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Jiang, X.-Z., Cong, L., Niu, J.-Z., Dou, W. & Wang, J.-J. Alternative splicing contributes to the coordinated regulation of ferritin subunit levels in *Bactrocera dorsalis* (Hendel). *Sci. Rep.* 4, 4806; DOI:10.1038/srep04806 (2014).

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License. The images in this article are included in the article’s Creative Commons license, unless indicated otherwise in the image credit; if the image is not included under the Creative Commons license, users will need to obtain permission from the license holder in order to reproduce the image. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-nd/3.0/