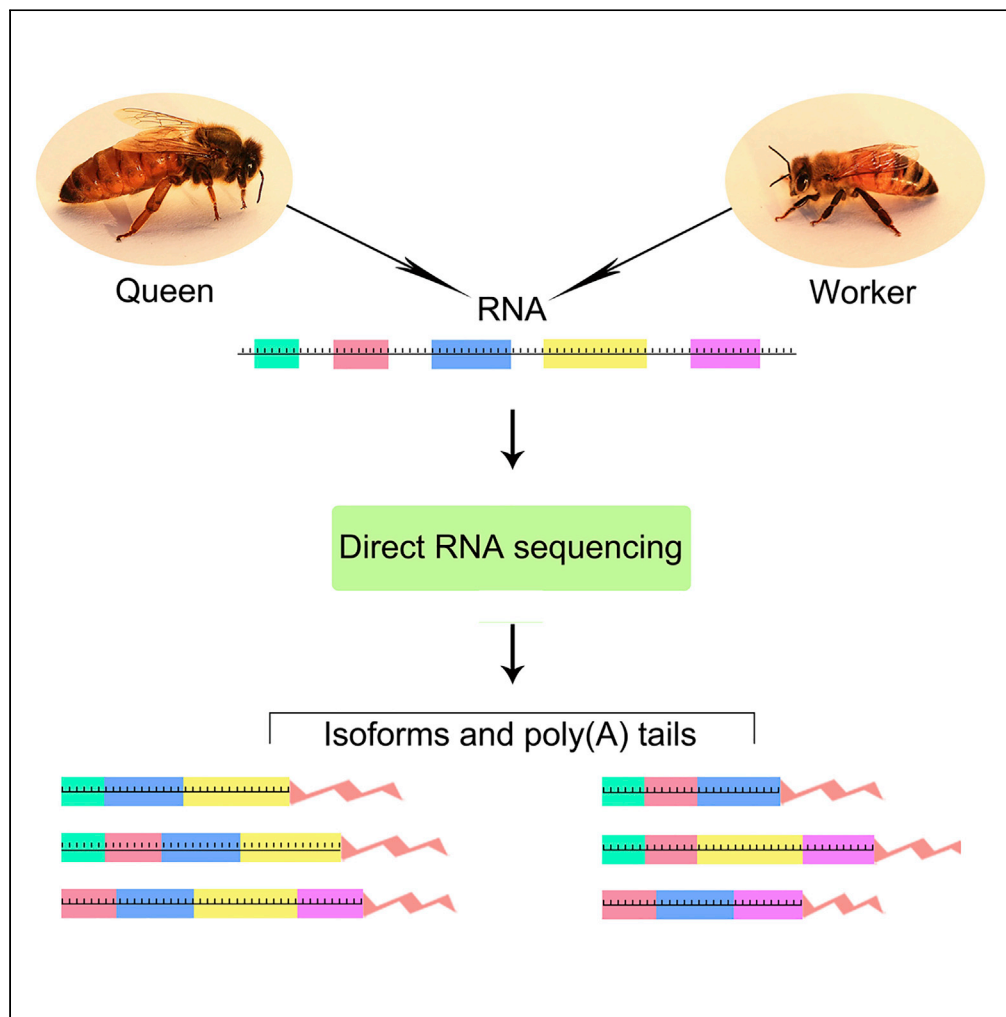


## Article

## Extent and complexity of RNA processing in honey bee queen and worker caste development



Xu Jiang He,  
Andrew B. Barron,  
Liu Yang, ..., Xiao  
Bo Wu, Wei Yu  
Yan, Zhi Jiang  
Zeng

bees1965@sina.com

**Highlights**

Honeybee caste  
differentiation has a  
complexity of RNA  
processing

Isoforms differentially  
express between queens  
and workers during larval  
development

Isoforms are formatted by  
a flexible alternative  
splicing system

Poly(A) tails are negatively  
correlated with isoform  
expression

He et al., iScience 25, 104301  
May 20, 2022 © 2022 The  
Authors.  
[https://doi.org/10.1016/  
j.isci.2022.104301](https://doi.org/10.1016/j.isci.2022.104301)

## Article

## Extent and complexity of RNA processing in honey bee queen and worker caste development

Xu Jiang He,<sup>1,2</sup> Andrew B. Barron,<sup>3</sup> Liu Yang,<sup>4</sup> Hu Chen,<sup>4</sup> Yu Zhu He,<sup>1</sup> Li Zhen Zhang,<sup>1</sup> Qiang Huang,<sup>1</sup> Zi Long Wang,<sup>1</sup> Xiao Bo Wu,<sup>1</sup> Wei Yu Yan,<sup>1</sup> and Zhi Jiang Zeng<sup>1,2,5,\*</sup>

## SUMMARY

The distinct honeybee (*Apis mellifera*) worker and queen castes have become a model for the study of genomic mechanisms of phenotypic plasticity. Here we performed a nanopore-based direct RNA sequencing with exceptionally long reads to compare the mRNA transcripts between queen and workers at three points during their larval development. We found thousands of significantly differentially expressed transcript isoforms (DEIs) between queen and worker larvae. These DEIs were formatted by a flexible splicing system. We showed that poly(A) tails participated in this caste differentiation by negatively regulating the expression of DEIs. Hundreds of isoforms uniquely expressed in either queens or workers during their larval development, and isoforms were expressed at different points in queen and worker larval development demonstrating a dynamic relationship between isoform expression and developmental mechanisms. These findings show the full complexity of RNA processing and transcript expression in honey bee phenotypic plasticity.

## INTRODUCTION

Phenotypic plasticity is a phenomenon where the same genotype showed distinct phenotypes if raised in different conditions, which is the most interesting aspect of development (Agrawal, 2001; Borges, 2005). Phenotypic plasticity is particularly dramatic in eusocial insects, and the honeybee (*Apis mellifera*) has become a valuable system for exploring this phenomenon. Honeybee queens and workers both develop from fertilized (diploid) eggs, but differences in nutrition in their larval diets determines their diverging developmental trajectories that result in different phenotypes (Asencot and Lensky, 1984; Brouwers, 1984; Maleszka, 2008; Spannhoff et al., 2011; Mao et al., 2015). The queen is the colony's only fecund female with well-developed ovaries producing around 1500 eggs per day, whereas workers are facultatively sterile females that perform all the essential tasks in the colony such as brood care, foraging, defense, construction, and cleaning (Winston, 1991).

Prior foundational studies have shown how the process of phenotypic plasticity is epigenetically determined (Chen et al., 2012; Foret et al., 2012; Cameron et al., 2013; Guo et al., 2016; He et al., 2019; Wojciechowski et al., 2018). Studies have compared the transcriptomes between honeybee queens and workers using short-length RNA sequencing (RNA-Seq) (Chen et al., 2012; Cameron et al., 2013; He et al., 2019) or cDNA microarrays (Yamazaki et al., 2006; Barchuk et al., 2007). Thousands of genes were differentially expressed between queens and workers during their development (Chen et al., 2012; Cameron et al., 2013; He et al., 2019). Consistently a few key signaling pathways have been implicated as important in queen/worker differentiation. These include target of rapamycin (TOR), the fork head box O (FoxO), Notch, Wnt, insulin/insulin-like signaling (IIS) pathways, mitogen-activated protein kinase (MAKP), Hippo, and transforming growth factor beta (TGF-beta) (Patel et al., 2007; Mutti et al., 2011a; Duncan et al., 2016; Xiao et al., 2017; Yin et al., 2018; Wang et al., 2021). In addition, key genes such as *Juvenile hormone esterase precursor (Jhe)*, *ecdysone receptor (Ecr)*, *Vitellogenin (Vg)*, and *Hexamerin 70b (Hex70b)* major royal jelly proteins (*Mrjps*) also participate in the determination of queen-worker fate (Barchuk et al., 2007; Buttstedt et al., 2013; Mello et al., 2014). Moreover, epigenetic modifications such as DNA methylation, microRNA, and alternative chromatin states have been shown to be involved in the regulation of honeybee queen-worker dimorphism (Foret et al., 2012; Ashby et al., 2016; Guo et al., 2016; Wojciechowski et al., 2018). The latest evidence showed that m6A methylation on RNA also participates in the genomic regulation leading to queen-worker dimorphism (Bataglia et al., 2021; Wang et al., 2021).

<sup>1</sup>Honeybee Research Institute, Jiangxi Agricultural University, Nanchang, Jiangxi 330045, P. R. of China

<sup>2</sup>Jiangxi Province Honeybee Biology and Beekeeping Nanchang, Jiangxi 330045, P. R. of China

<sup>3</sup>Department of Biological Sciences, Macquarie University, North Ryde, NSW 2109, Australia

<sup>4</sup>Wuhan Benagen Tech Solutions Company Limited, Wuhan, Hubei 430021, P. R. of China

<sup>5</sup>Lead contact

\*Correspondence: bees1965@sina.com

<https://doi.org/10.1016/j.isci.2022.104301>



These studies have gone a long way to establish the honeybee as a model for genomic analyses of phenotypic plasticity, but thus far, we know rather little about how RNA processing might be involved in the process. RNA processing determines the mRNA coding sequence and can lead to different isoforms of a gene. Distinct isoforms can perform different biological functions, and these can be important for animal phenotypic plasticity (Marden, 2008). Alternative transcripts produced from the same gene can differ in the position of the start site, the site of cleavage and polyadenylation, and the combination of exons spliced into the mature mRNA (Parker et al., 2020). Poly(A) tails are also very important for RNA processing. Length of the Poly(A) tail is correlated with translational efficiency and regulates transcript stability (Workman et al., 2019; Roach et al., 2020).

Alternative splicing (AS) of genes diversifies the transcriptome and increases protein coding capacity through the production of multiple distinct isoforms (Modrek and Lee, 2001; Reddy et al., 2013). AS is involved in many biological processes in plants and animals and particularly responses and adaptation to changing environments (Modrek and Lee, 2001; Reddy et al., 2013). Alternative splicing of pre-mRNA has been associated with phenotypic plasticity in insects such as the bumble bee *Bombus terrestris* and the pea aphid *Acyrtosiphon pisum* (Grantham and Brisson, 2018; Price et al., 2018). Different alternative transcripts are also involved in honeybee caste differentiation, and alternative splicing plays an important role in queen-worker differentiation (Aamodt, 2008; Mutti et al., 2011b). These studies have focused on the effects of AS on the regulation of gene expression. Thus far, none have investigated how AS determines isoform expression genome wide in honeybees.

So far, honeybee RNA-Seq studies have been based on short-length RNA sequencing, which is limited to detect full-length isoforms or alternative transcripts of a gene. The new direct RNA sequencing technology (DRS) based on the Oxford Nanopore technology (ONT) directly sequences the native full-length mRNA molecules, which offers detailed information on the mRNA and RNA modifications (Gralde et al., 2018; Harel et al., 2019; Workman et al., 2019; Parker et al., 2020; Roach et al., 2020; Zhang et al., 2020). It avoids biases from reverse transcription or amplification and yields full-length, strand-specific mRNA (Gralde et al., 2018; Harel et al., 2019; Zhang et al., 2020). This method allows a genome-wide investigation of RNA processing [e.g., alternative splicing and poly(A) tails].

Our objective in this study was to deploy this DRS technology to examine differences in RNA processing between the worker and queen honeybee castes. This allows us to investigate fundamental questions on caste-specific transcriptome patterns in honeybees: the variety of full-length isoforms and how they expressed differentially in queens and workers and the extent and complexity of the RNA processing in honeybee development.

We sampled honeybee queens and workers at different points in their larval development to assess the process of developmental canalization into the different phenotypes. Our results reveal extremely prevalent and highly complex differences in RNA processing between workers and queens. These differences changed across larval development showing a complex dynamical interaction between the epigenomic programs and the developmental pathways

## RESULTS

### Quality of direct RNA sequencing data

Twelve libraries were generated and sequenced by ONT direct RNA sequencing. The number of clean reads of each sample was  $5,416,798 \pm 1,510,975$  (Table 1). The N50 read length was  $1,297 \pm 30$  nucleotides (nt), and the read quality was  $10.63 \pm 0.13$  (Table 1), indicating a high integrity of the sequenced RNAs.  $98.13 \pm 0.79\%$  clean reads were successfully mapped to the *A.pis mellifera* reference genome (Amel HAv 3.1) (Table 1). Moreover, the total number of genes detected from DRS and RNA-seq was equal (9,592 vs 9838), and over 88.8% of them were overlapped (Figure S1). We also mapped the transcript isoforms to honey bee reference genes and showed a great quality of percentile of the gene body (5'-3') in each DRS sample (Figure S2). These results again confirmed a high and reliable integrity of the sequenced RNAs. The Pearson correlation coefficients between two biological replicates of each group were all over 0.85 (Figure S3) and the sample cluster tree also showed that two biological replicates of each group were clustered together (Figure S4), reflecting a good repeatability of biological replicates.

**Table 1. Overview of direct RNA sequencing**

| Samples              | 2Q                    | 2W                 | 4Q                  | 4W                    | 6Q                    | 6W                    | Average               |
|----------------------|-----------------------|--------------------|---------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Clean read numbers   | 6,531,006 ± 1,043,238 | 6,103,488 ± 14,688 | 6,506,243 ± 620,747 | 5,007,500 ± 1,696,509 | 3,548,493 ± 2,108,736 | 4,804,061 ± 1,564,551 | 5,416,798 ± 1,510,975 |
| Average length (nt)  | 1,030 ± 46            | 1,012 ± 15         | 1,003 ± 7           | 998 ± 42              | 1,053 ± 23            | 1,002 ± 36            | 1,016 ± 31            |
| N50 length (nt)      | 1,316 ± 57            | 1293 ± 11          | 1,286 ± 13          | 1,279 ± 52            | 1,311 ± 39            | 1,296 ± 14            | 1,297 ± 30            |
| Max read length (nt) | 20,044 ± 4,581        | 17,248 ± 987       | 24,924 ± 7,011      | 17,955 ± 2,212        | 17,548 ± 398          | 14,579 ± 1,298        | 18,716 ± 4,273        |
| Average read quality | 10.66 ± 0.07          | 10.73 ± 0.03       | 10.40 ± 0.19        | 10.64 ± 0.03          | 10.73 ± 0.06          | 10.65 ± 0.01          | 10.63 ± 0.13          |
| Alignment % identity | 98.58 ± 0.32          | 98.44 ± 0.03       | 98.43 ± 0.28        | 98.29 ± 0.13          | 97.63 ± 1.10          | 97.45 ± 1.78          | 98.13 ± 0.79          |
| Uni mapped genes     | 8,736 ± 64            | 8,658 ± 0          | 8,544 ± 159         | 8,323 ± 125           | 8,204 ± 654           | 8,592 ± 74            | 8,509 ± 286           |
| No. of isoforms      | 23,518 ± 764          | 22,896 ± 61        | 22,382 ± 899        | 20,144 ± 501          | 19,426 ± 4,156        | 20,800 ± 121          | 21,528 ± 2,040        |

The 2Q, 2W, 4Q, 4W, 6Q, and 6W are 2-day queen larvae, 2-day worker larvae, 4-day queen larvae, 4-day worker larvae, 6-day queen larvae and 6-day worker larvae, respectively. Data are represented as mean ± SD and each group has two biological replicates.

### Differences of isoform and gene expression in queen-worker comparisons

Our results showed genome-wide differentially expressed isoforms in honeybee caste differentiation. Significantly differentially expressed transcript isoforms (DEIs) and genes (DEGs) between queen and worker larvae were identified. We identified 662, 1855 and 1042 DE-Is in 2d, 4d, and 6d queen-worker larval comparisons, respectively, which were notably more than the DEGs (281, 1369, and 645, respectively, see [Figure 1A](#), detailed information see [Tables S1, S2, S3, S4, S5, and S6](#)). To compare the differences between DE-Is and DE-Gs, we firstly mapped DE-Is to reference genes (DEIGs) and showed that only 271 (34.72%), 977 (53.45%), and 410 (38.32%) genes overlapped between DEIGs and DE-Gs in 2d, 4d, and 6d queen-worker comparisons, respectively ([Figures 1B–1D](#)). This shows a considerable difference between DEIs and DEGs in queen-worker differentiation.

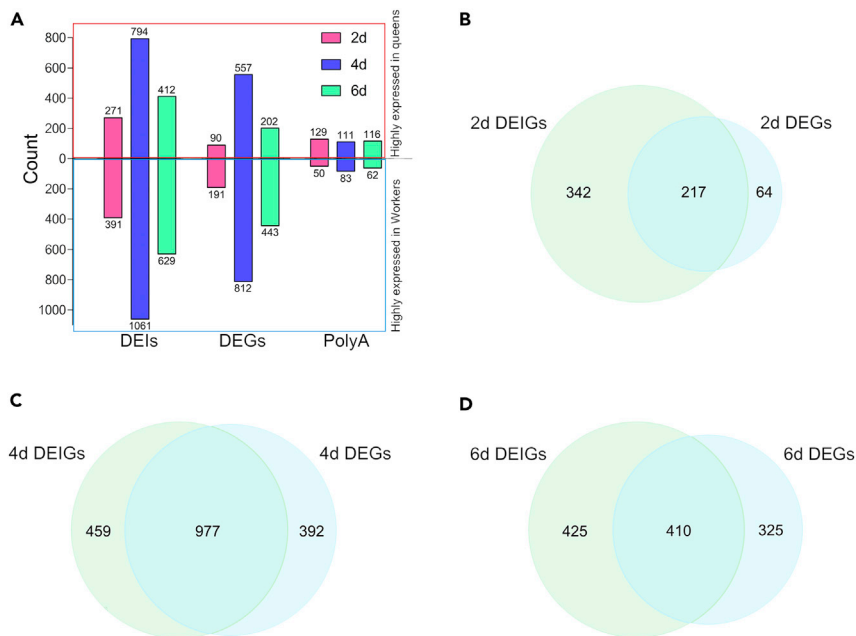
We identified DEIs (24, 54, and 48 in 2d, 4d, and 6d comparisons, respectively) that are involved in several key KEGG signaling pathways for honey bee caste differentiation, such as mTOR, Notch, FoxO, Wnt, IIS, Hippo, TGF-beta, and MAPK (see [Tables S1, S2, and S3](#)). The number of DEGs enriched in these signal pathways was 7, 34, and 33 in 2d, 4d, and 6d, respectively ([Tables S4, S5, and S6](#)). Other DEIs mapped to key genes such as *Jhe*, *Mrijps*, *Ecr*, *Vg*, and *Hex70b* (see [Tables S1, S2, and S3](#)). The number of isoforms with significantly different poly(A) length in 2d, 4d and 6d comparisons showed a negative trend as DEIs and DEGs, with more in queens ([Figure 1A](#), detailed information see [Tables S7, S8, and S9](#)) and less in workers.

DEIs and DEGs were enriched in five key KEGG signaling pathways (mTOR, Notch, FoxO, Wnt, and IIS) which have been previously shown to be participating in queen-worker differentiation. DEIs were mapped to 35 proteins in these five pathways, whereas DEG mapped onto just 17 proteins ([Figure 2A](#)). In detail, 23 proteins were uniquely mapped by DEIs, whereas only 5 were uniquely mapped by DEGs. 12 proteins mapped to both DEI and DEG ([Figure 2A](#)). We also provided detailed information about the expression of DEIs and DEGs enriched in these five key pathways ([Figure 2B](#)). Clearly, more isoforms were significantly differentially expressed in these five key pathways compared to DE-Gs, and the differentially expressed isoforms of each DEIG are presented in [Figure 2B](#).

These findings suggest that caste-specific isoform expression provided more accurate and detailed information of the real transcriptome differences between honeybee queen and worker castes than measurements of differences in gene expression.

### Uniquely expressed isoforms between queens and workers

We identified 187, 357 and 364 uniquely expressed isoforms in queen or worker larvae at 2, 4days, 4d, and 6d age, with more occurring in queens than workers ([Figure 3A](#)). The number of uniquely expressed



**Figure 1. Comparison of isoform and gene expression, poly(A) length between queen and worker larvae**

See also [Tables S1, S2, S3, S4, S5, S6, S7, S8, S9, S13, S14, S15, S16, and S17](#).

(A) Numbers of differentially expressed isoforms (DEIs) and genes (DEGs), different poly(A)-length related isoforms between queens and workers at three larval stages (day 2, day 4, and day 6). The up bars are DEIs and DEGs length highly expressed in 2d, 4d, and 6d queen larvae as well as isoforms with longer poly(A) tails in queen larvae, and the down bars are that highly expressed or with longer poly(A) tails in worker larvae.

(B) The venn diagram of DEIGs (DEI mapped genes) and DEGs from 2d queen-worker comparison.

(C) The venn diagram of DEIGs and DEGs from 4d queen-worker comparison.

(D) The venn diagram of DEIGs and DEGs from 6d queen-worker comparison.

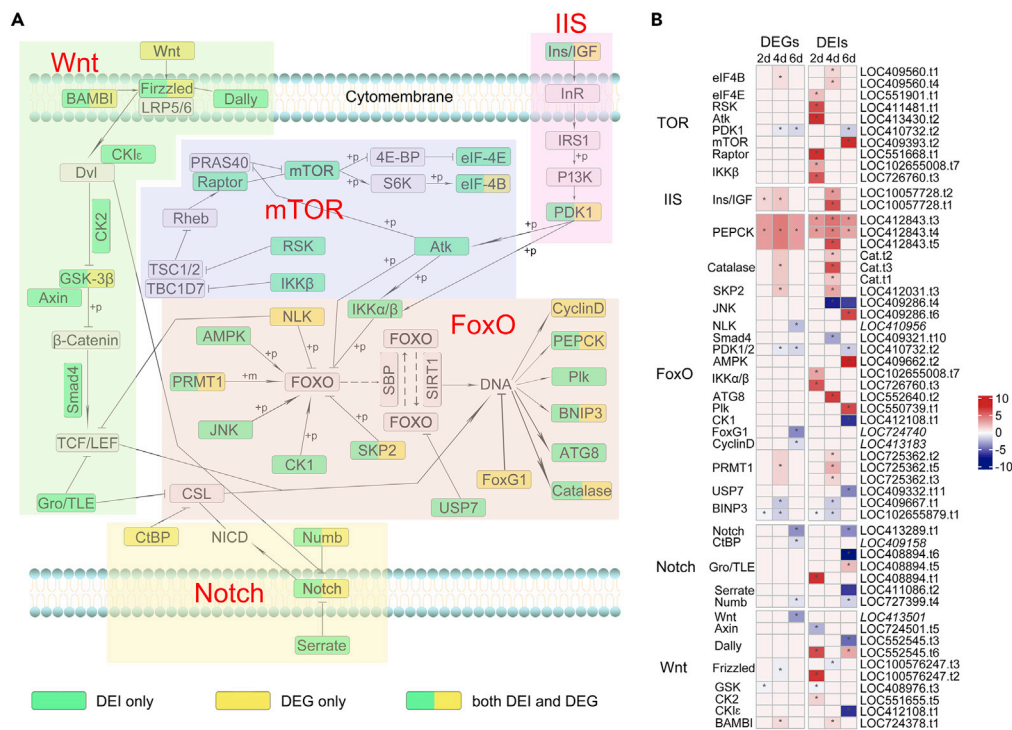
isoforms in queens reached a maximum at the 4d stage, whereas in workers it reached a maximum in the 6d sample. This could suggest that the queen developmental pathway diverges more quickly and at an early larval stage, whereas the developmental pathway of workers diverges slightly later.

Uniquely expressed isoforms were enriched into eight key KEGG signaling pathways such as mTOR and IIS for queen-worker differentiation as well as some key genes such as *DNA methyltransferase 3 (Dnmt3)* and *Ecr* ([Figure 3](#), details see [Tables S10, S11, and S12](#)). This indicated that there are isoforms which are likely to be involved in queen-worker differentiation that are uniquely expressed in queens or workers during their development.

### DEIs and alternative splicing

We identified 21,574 isoforms in each sample on average, and these isoforms could map to averagely 8509 honeybee reference genes ([Table 1](#)). This suggested that the honey bee genome produces notably more isoforms than genes. DRS technology allows detecting the splicing events causing an isoform. We reconstructed the complex splicing events as well as the DEIs resulting from the alternative splicing. We then define a DEI from a single gene, which is differentially spliced compared to the longest isoform of this gene as a DS-DEI. The majority of DEIs (73.26%, 67.98, and 63.53% of 2d, 4d, and 6d DEIs) were DS-DEIs, and most of them (71.13%, 66.16%, and 65.56% in 2 d, 4d, and 6d) were generated by at least two types of splicing events in a single isoform. A combination of RI, A5, and A3 in a single DEI was the most common type ([Figure 4](#)). This indicates that the formation of transcripts in honeybee development is more complex than previously known.

Of these DEIs, 9.82% (65), 5.71% (106), and 9.50% (99) of DEIs from the 2d, 4d, and 6d comparisons were produced by significantly different AS events when comparing workers and queens ([Figure S5](#), [Tables S13, S14, and S15](#)). Similarly, many of them (46.15%, 42.45, and 45.45% in 2d, 4d, and 6d



**Figure 2. Expression and enrichment of DEIs and DEGs in five KEGG pathways**

See also Tables S1, S5, S6, S16, and S17.

(A) DEIs and DEGs enriched in key steps of five KEGG pathways. Five KEGG pathways were shown with transparent boxes and the names of these pathways were red marked. Green boxes inside of transparent boxes are DEI enriched proteins; yellow boxes are DEG enriched proteins; green-yellow mixed boxes are both DEI and DEG enriched proteins; blank boxes are non DEI and DEG enriched proteins.

(B) Expression of DEIs and DEGs in the five key pathways. The left are pathway and protein names, and the right are isoform expression and names. Expression of isoforms and genes are presented with their log2 TPM values and shown with color scales. DEIs or DEGs are marked with “\*” in the middle of the boxes.

comparisons, respectively) contained at least two types of significantly different AS events (Figure S6, Tables S13, S14, and S15). Consequently, a part of DEIs were significantly differentially spliced.

### Correlation between poly(A) length and expression of DEIs and DEGs

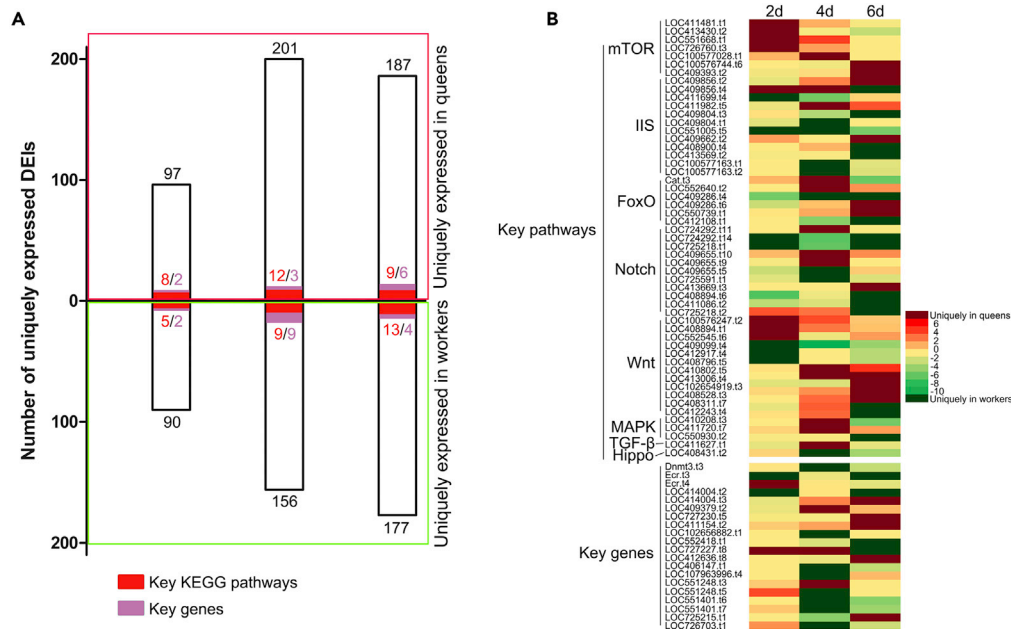
We showed a negative correlation between isoform expression and their poly(A) tail length (Figure S7). Expression of DEIs (Figure 5) and DEGs (Figure S8) was also negatively correlated with their poly(A) lengths. It is believed that poly(A) tails are involved in the regulation of isoform expression. More interestingly, the poly(A) lengths were positively correlated with log2 fold change values of DEIs, which was stronger than that of DEGs (Figures 5D–5F). This firstly showed that poly(A) lengths were more closely related to the biased expression of DEIs than DEGs.

### Correlation of isoform expression and poly(A) lengths for two key genes

Two key genes, *Jhe* and *Ecr*, which participate in queen-worker differentiation were selected as examples to show the correlation pattern of gene and isoform expression and its poly(A) length. *Jhe* had only one isoform and was negatively correlated with its poly(A) lengths (Figure 6A). The *Ecr* gene has 5 isoforms and we detected 4 isoforms (Figure 6B). Queen and worker larvae showed a more detailed difference in the expression of these 4 isoforms (see the middle of Figure 6B) than gene expression (see the left of Figure 6B). We selected the most differentially expressed isoform (*Ecr.t4*) to measure the correlation between its expression and poly(A) length. Similarly, the expression of *Ecr.t4* isoform was also negatively correlated with their poly(A) lengths (Figure 6B).

We selected two isoforms of the *Ecr* gene (*Ecr.t1* and *Ecr.t2*) for TaqMan real-time PCR confirmatory experiment. The expression of these two isoforms in TaqMan real-time PCR (Figure 7) was mainly





**Figure 3. Uniquely expressed isoforms in queen and worker larvae**

See also [Tables S10, S11, S12b, S16 and S17](#).

(A) Numbers of uniquely expressed isoforms in queen larvae (up bars) or worker larvae (down bars) at 2d, 4d and 6d stages. The red bars are numbers of uniquely expressed isoforms enriched in eight key KEGG signaling pathways (mTOR, IIS, FoxO, Notch, Wnt, MAPK, Hippo, and TGF-β). The purple bars are numbers of uniquely expressed isoforms that are key genes reported in previous studies (details see [Tables S10, S11, and S12](#)).

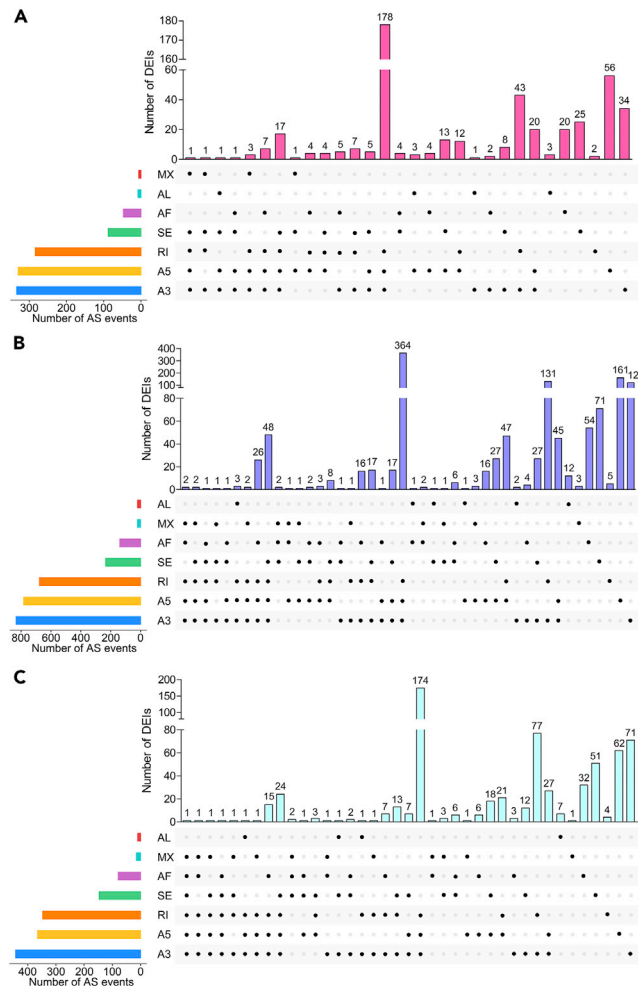
(B) The heat map of expression of uniquely expressed isoforms which are shown in red bars and green bars in (A). The red color blocks in heat map are uniquely expressed isoforms in queen larvae at 2d, 4d and 6d age, whereas the dark green blocks are uniquely expressed isoforms in worker larvae. Other color scales are the log<sub>2</sub> fold change values of isoform expression between queen and worker larvae at 2d, 4d, and 6d stages.

consistent with the DRS results, confirming a biased isoform expression in honey bee caste differentiation.

## DISCUSSION

RNA processing acts as a major driver of animal phenotypic plasticity ([Gingeras, 2007](#)), but the complexity and characteristics of RNA processing underlying this phenotypic variability is not well understood. The present study has shown the extent of RNA processing in honeybee queen-worker differentiation. Differences in isoform expression revealed the true extent of the differences in the transcriptome between queens and workers during their development, providing more detailed and accurate information on mRNA expression compared to measures of gene expression. We also showed the flexibility of alternative splicing in queen-worker differentiation. Most DEIs were formed by more than one type of alternative splicing. Poly(A) length negatively correlated with the expression of DEIs demonstrating another part of the genomic mechanisms of honey bee caste differentiation.

Thousands of DEIs were identified between queens and workers, which were notably more than DEGs ([Figure 1A](#)). Many genes were not differentially expressed, but different isoforms from these genes were significantly differentially expressed ([Figures 1B–1D, Tables S1, S2, and S3](#)), demonstrating how DRS gives a very different perception of the transcriptome differences between developing worker and queen honey bees. DEIs are enriched into more proteins of five key KEGG signaling pathways such as mTOR and IIS than DEGs ([Figure 2](#)). The mTOR signaling pathway is the central component of a conserved eukaryotic signaling pathway, regulating cell and organismal growth in response to nutrient status and has been shown to be important for determining the queen-worker developmental paths ([Colombani et al., 2003; Oldham and Hafen, 2003; Patel et al., 2007](#)). The IIS and Notch signaling pathways have been shown to be involved in the development of honeybee ovaries ([Mutti et al., 2011a; Duncan et al., 2016](#)). More DEIs were identified



**Figure 4. The alternative splicing events (AS) in DEIs of three comparisons**

See also Figure S6.

(A) DEIs of 2d comparison containing AS events. The colorful bars in the left bottom diagram are the different AS patterns and the number of AS events in DEIs. Seven different AS patterns are 3'splice site (A3), 5'splice site (A5), First exon (AF), Last exon (AL), Retained intron (RI), Skipping exon (SE), and Mutually exclusive exon (MX). The pink vertical bars in the top of the right diagram are the numbers of DEIs that were spliced by a particular combination of AS forms. The black spots in a column mean DEIs that were spliced by a combination of related AS forms. Same in (B and C).

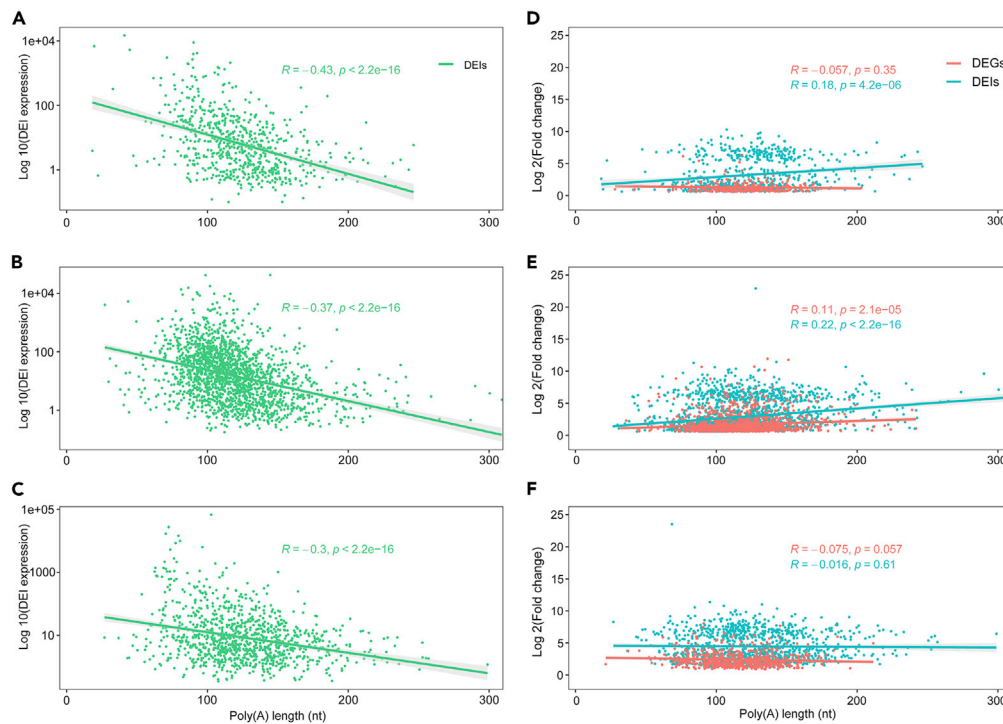
(B) DEIs of 4d comparison containing AS events.

(C) DEIs of 6d comparison containing AS events.

in these key pathways than DEGs, indicating that DEIs could provide more nuanced differences of mRNA expression between honeybee queens and workers during their development. Therefore, compared to DE-Gs that conceal a more nuanced picture of transcription differences, DE-Is show the true nature of transcriptome differences between queens and workers. Consequently, DEIs identified in the present study revealed a more informative knowledge and precise difference in the transcriptome expression of honey bee caste differentiation compared to previous RNA-Seq results (Chen et al., 2012; Cameron et al., 2013; He et al., 2019; Yin et al., 2018).

Hundreds of uniquely expressed isoforms were identified in queens or workers. Many uniquely expressed isoforms were enriched in mTOR, IIS, FoxO, Notch, Wnt, MAPK, TGF-beta, and Hippo—the eight key KEGG-signaling pathways (Figure 3). Some uniquely expressed isoforms mapped to key genes such as DNA methyltransferase 3 (*Dnmt3*) and *Ecr* genes (Figure 3). *Dnmt3* is a key driver of epigenetic global reprogramming controlling honeybee queen-worker fate (Kucharski et al., 2008). Here, we showed one isoform of *Dnmt3* (*Dnmt3.t3*) was uniquely expressed in 4 days worker larvae, suggesting that this isoform of





**Figure 5. Correlation between poly(A) length and expression of DEIs and DEGs**

See also [Figures S7 and S8](#) and [Tables S7, S8, and S9](#).

(A) Correlation between poly(A) lengths and the expression of DEIs of 2d queen-worker comparison. Expression of each DEI was the log10 TPM value. Same in (B and C).

(B) Correlation between poly(A) lengths and DEIs of 4d comparison.

(C) Correlation between poly(A) lengths and DEIs of 6d comparison.

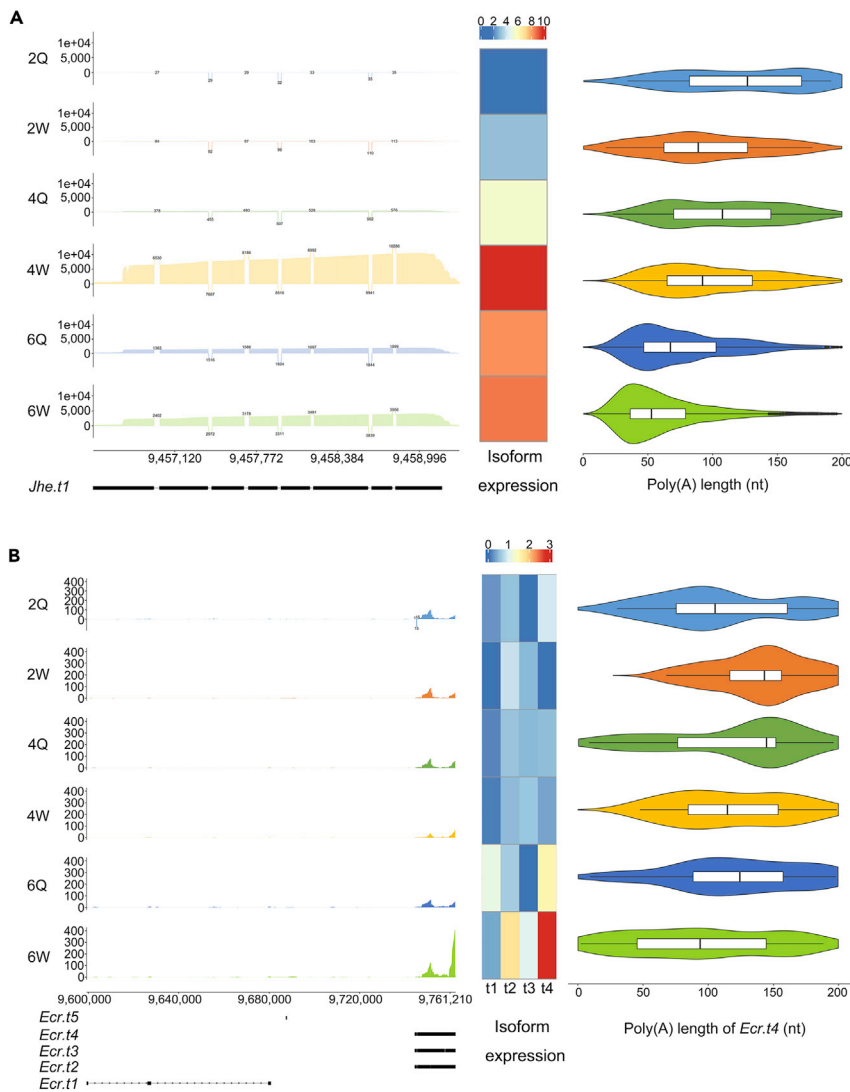
(D) Correlation between poly(A) lengths and log2 fold change values of DEIs and DEGs of 2d comparison. Same in (E and F).

(E) Correlation between poly(A) lengths and log2 fold change values of DEIs and DEGs of 4d comparison.

(F) Correlation between poly(A) lengths and log2 fold change values of DEIs and DEGs of 6d comparison.

*Dnmt3* may play an important role in the worker developmental pathway at this critical stage. For the *Ecr* gene, 2 days worker uniquely expressed *Ecr.t3*, whereas 2 days queen larvae uniquely expressed *Ecr.t4*. Different isoforms from the same gene were expressed in queens and workers, revealing the complexity of RNA processing in the honeybee queen-worker differentiation. We have documented many genes where isoform expression was not consistently biased toward either queens or workers throughout all stages of larval development. In some cases, an isoform was uniquely expressed in either queens or workers in one larval stage, but at other developmental stages the isoform occurred in both workers and queens ([Figure 3B](#)). Isoforms like *LOC409856.t4* and *LOC727227.t8* uniquely expressed in queens at 2d and 4d stages but oppositely expressed in workers at 6d stage ([Figure 3B](#)). This illustrates the dynamic relationship between RNA processing and the developmental process. A simple model might imagine different isoforms of a gene specific to either the queen or worker developmental pathway. Instead, we see isoforms featuring in both pathways but at different stages.

The number of uniquely expressed isoforms also showed a different trend in queens and workers. It reached a maximum in queens at the 4d larval stage but it peaked later in workers ([Figure 3A](#)). These different trends correlate with the known faster development of queens ([Winston, 1991](#)) and suggest the queen's developmental pathway differentiates at an earlier larvae stage than the worker developmental pathway. Queen larvae at early stage receive significantly more food than worker larvae, and the nutrient contents in royal jelly also differ from worker jelly in terms of sugar, vitamins, proteins, acids, and microRNA ([Asencot and Lensky, 1984](#); [Brouwers, 1984](#); [Maleszka, 2008](#); [Mao et al., 2015](#)). In [Figure 3B](#), most isoforms enriched in mTOR signaling pathway which is a signaling pathway in response to nutrient status ([Colombani](#)



**Figure 6. The poly(A) lengths and isoform expression of two key genes**

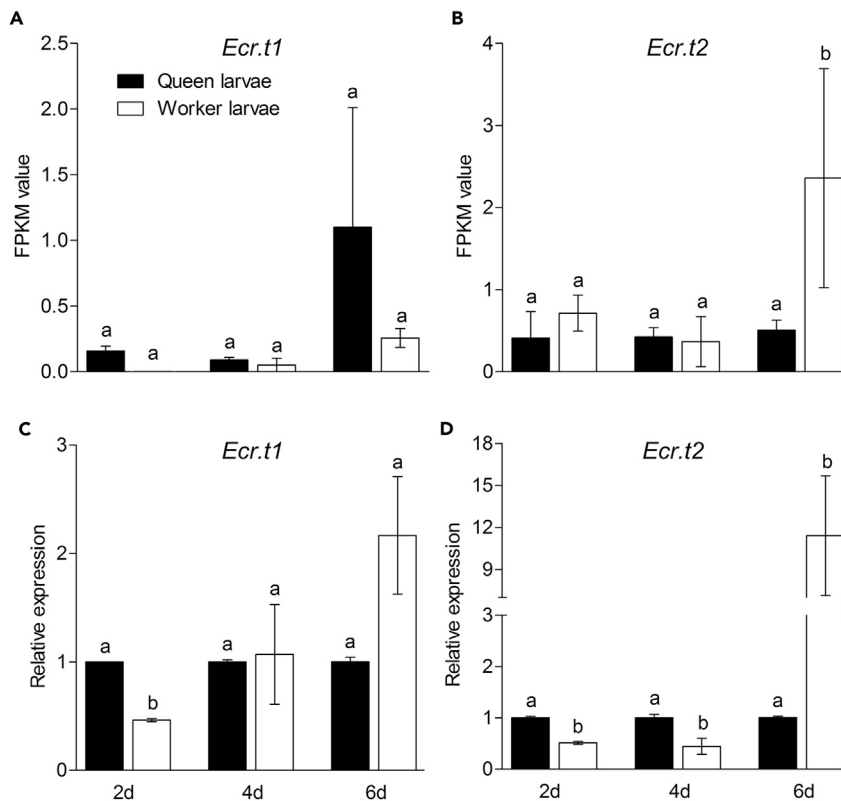
See also Tables S1, S2, S3, S16, and S17.

(A) Expression and related poly(A) lengths of *Jhe* isoform. The left is the exon coverage of clean reads (namely gene expression) of *Jhe* gene of 6 larval groups. The *Jhe* has only one isoform (*Jhe.t1*) and its structure is present (see its exons below marked with black color). The middle heat map is the isoform expression of *Jhe.t1* (log2 TPM values) in 6 groups and presented with color scales. The right is the poly(A) length distribution of *Jhe.t1* in 6 groups. The mean poly(A) length is presented in the middle of each violinplot.

(B) Expression and related poly(A) lengths of *Ecr* isoforms. The left is the exon coverage of clean reads (namely gene expression) of *Ecr* gene. The *Ecr* gene has 4 isoforms and their expression is presented using heat map as above. The poly(A) length distribution of *Ecr.t4* in 6 groups is presented on the right side; it was the most significantly differentially expressed isoform of the gene *Ecr*. The poly(A) of *Ecr.t4* in 2dW was presented with a blank box, which means few clean data of poly(A) was detected because 2dW did not express *Ecr.t4*.

et al., 2003; Oldham and Hafen, 2003) were uniquely expressed in queens at 2d stage and its number decreased at 4d and 6d stages, which supports our speculation.

We detected  $21,528 \pm 2,040$  isoforms in each sample which could map to  $8,509 \pm 286$  honey bee reference genes (Table 1). This is consistent with a previous study that 60.3% of honeybee multi-exon genes are alternatively spliced (Li et al., 2013). The majority of DEIs, contained AS events and over 65% of them had at least two types of AS patterns in a single isoform (Figure 4). Some DEIs contained more than five different AS



**Figure 7. Verification of two isoforms of *Ecr* gene between 2d, 4d, and 6d queen and worker larvae by TaqMan real-time PCR**

See also Table S18.

(A) The *Ecr.t1* expression in three queen-worker comparisons (2 d, 4 d, and 6 d) from DRS. The expression of the isoform was the FPKM value. Data were presented as mean  $\pm$  SE. Different letters in the top of bars indicate significant difference (p-values < 0.05 and log2 Fold change values > 1.5 using DESeq2) and the same letter indicates no significant difference. Same in (B).

(B) The *Ecr.t2* expression in three queen-worker comparisons (2d, 4d, and 6d) from DRS.

(C) The *Ecr.t1* expression in three queen-worker comparisons (2d, 4d, and 6d) from TaqMan real-time PCR. Data were presented as mean  $\pm$  SE. Different letters in the top of bars indicate significant difference (p < 0.05, one-way ANOVA test) and same letter indicate no significant difference. Same in (D).

(D) The *Ecr.t2* expression in three queen-worker comparisons (2d, 4d, and 6d) from TaqMan real-time PCR.

types (Figure 4). A few DE-Is were differentially spliced between queens and workers by multiple AS patterns (Figure S3, Tables S13, S14, and S15). It means several distinct AS types together contributed to the construction of a single isoform (Figure 4), revealing a flexible system in honey bee caste differentiation.

Poly(A) tails are a regulator of translation and transcript stability (Lim et al., 2016; Tudek et al., 2018; Woo et al., 2018). Poly(A) tail length is not stable but dynamic and condition-dependent (Lim et al., 2016; Woo et al., 2018). This study showed that poly(A) tail lengths were strongly and negatively correlated with expression of isoforms in honeybees (Figure 5). The negative correlation between isoform expression and poly(A) length has been shown in the regulation of *Caenorhabditis elegans* development (Roach et al., 2020). Here we showed that poly(A) tails also participate in the regulation of isoform expression and therefore contributes to complexity of RNA processing in honey bee queen-worker differentiation. Moreover, we examined the relationship between poly(A) length and the degree of biased expression of isoforms or genes in queen-worker differentiation. Interestingly, the log2 fold change values of DEIs were positively correlated with the poly(A) lengths and their correlation were stronger than that of DEGs (Figures 5D–5F), suggesting that poly(A) tails play an important role in the RNA processing of honeybee caste differentiation. This also firstly showed that poly(A) tails had a stronger correlation with isoforms than genes in terms of insect phenotypic plasticity.

Consequently, poly(A) tails possibly contribute to the determination of honeybee queen-worker dimorphism via regulating the isoform expression.

Environmental stimulus altering patterns of gene expression has been considered as a primary molecular mechanism of phenotypic plasticity (Schlichting and Simth, 2002; Aubin-Horth and Renn, 2009). However, one gene can produce different transcripts by RNA processing, and different transcripts have distinct biological functions that are vital for phenotypic plasticity (Marden, 2008). Therefore, increasing evidence revealed that the complex RNA processing may be a principal contributor to the phenotypic complexity (Gommans et al., 2008, 2010). The present study supports this hypothesis and showed a genome-wide different isoform expression in queen-worker differentiation, which was far more complex than the gene expression (Figures 1 and 2). Especially hundreds of isoforms were uniquely expressed in queens or workers (Figure 3). These uniquely expressed isoforms cannot be easily seen in gene expression but may play an important role in the determination of honey bee caste differentiation. We also showed a flexible alternative splicing generating a variety of isoforms (Figure 4 and S6) and the polyadenylation [poly(A) tails] negatively regulating isoform expression (Figures 5 and 6). All findings demonstrated a complexity of RNA processing in honey bee phenotypic plasticity.

In summary, RNA processing plays an important role in shaping honeybee phenotypic plasticity, but investigations of gene expression by short-length RNA-Seq fail to reveal the full complexities of RNA processing (Chen et al., 2012; Cameron et al., 2013; He et al., 2019; Yin et al., 2018). By using DRS, this study has shown the extent of differential isoform expression between workers and queens, flexibility of transcript splicing, and polyadenylation. It provides a more detailed understanding of the molecular mechanisms underlying the divergence of the queen and worker developmental paths. It also contributes to our understanding of the extent and complexity of RNA processing in animal phenotypic plasticity.

### Limitations of the study

Although this study showed an extensive and complex RNA processing in honeybee caste differentiation including biased isoform expression, flexible alternative splicing, and polyadenylation, other RNA processing events like 5' capping are limited to be detected by the DRS technology but may also be involved into the honeybee caste differentiation. In addition, to extract enough amount of total RNA for DRS, we used a whole body of a 4d and 6d queen or worker larva as a sample and collected twenty 2d larvae as a mixed sample because of its small body size. This amount of RNA used in DRS was much lower than that in RNA-Seq; however, we note that the isoform expression of honey bee queen and worker larvae might differ in their distinct tissues.

### STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Insects
- METHOD DETAILS
  - Larva sampling
  - RNA preparation
  - Direct RNA sequencing
  - Preprocessing and alignments
  - DEG and DEI calculation
  - KEGG pathway enrichment analysis
  - Alternative splicing identification
  - Poly(A) length estimation and correlation with expression of genes and isoforms
  - Short-length RNA sequencing
  - TaqMan real-time PCR
- QUANTIFICATION AND STATISTICAL ANALYSIS

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.isci.2022.104301>.

## ACKNOWLEDGMENTS

We thank Mr Yi Cheng Yang and Mrs Yan Niu for help on data analysis and Mrs Xin Zhang for help on project management. This work was supported by the National Natural Science Foundation of China (32172790, 32160815, and 31872432), Major Discipline Academic and Technical Leaders Training Program of Jiangxi Province (20204BCJL23041) and the Earmarked Fund for China Agriculture Research System (CARS-44-KXJ15).

## AUTHORS CONTRIBUTION

Z.J.Z. and X.J.H. designed research. X.J.H. performed the research. X.J.H. and A.B.B. provided guidance for data. X.J.H., L.Y., and H.C. analyzed data. Y.Z.H., L.Z.Z., Q.H., Z.L.W., X.B.W., and W.Y.Y. assisted in the experiments. X.J.H., A.B.B., and Z.J.Z. wrote the paper. All authors were involved in the preparation of the final manuscript.

## DECLARATION OF INTERESTS

The authors declare no competing interest.

Received: November 10, 2021

Revised: March 12, 2022

Accepted: April 21, 2022

Published: May 20, 2022

## REFERENCES

- Aamodt, R.M. (2008). The caste- and age-specific expression signature of honeybee heat shock genes shows an alternative splicing-dependent regulation of Hsp90. *Mech. Ageing Dev.* 129, 632–637. <https://doi.org/10.1016/j.mad.2008.07.002>.
- Agrawal, A.A. (2001). Phenotypic plasticity in the interactions and evolution of species. *Science* 294, 321–326. <https://doi.org/10.1126/science.1060701>.
- Asencot, M., and Lensky, Y. (1984). Juvenile hormone induction of 'queenliness' on female honeybee (*Apis mellifera* L.) larvae reared on worker jelly and on stored royal jelly. *Comp. Biochem. Phys. B* 78, 109–117. [https://doi.org/10.1016/0305-0491\(84\)90153-6](https://doi.org/10.1016/0305-0491(84)90153-6).
- Ashby, R., Forêt, S., Searle, I., and Maleszka, R. (2016). MicroRNAs in honey bee caste determination. *Sci. Rep.* 6, 18794. <https://doi.org/10.1038/srep18794>.
- Aubin-Horth, N., and Renn, S.C.P. (2009). Genomic reaction norms: using integrative biology to understand molecular mechanisms of phenotypic plasticity. *Mol. Ecol.* 18, 3763–3780. <https://doi.org/10.1111/j.1365-294x.2009.04313.x>.
- Barchuk, A.R., Cristino, A.S., Kucharski, R., Costa, L.F., Simões, Z.L., and Maleszka, R. (2007). Molecular determinants of caste differentiation in the highly eusocial honeybee *Apis mellifera*. *BMC Dev. Biol.* 7, 70. <https://doi.org/10.1186/1471-213x-7-70>.
- Bataglia, L., Simões, Z.L.P., and Nunes, F.M.F. (2021). Active genic machinery for epigenetic RNA modifications in bees. *Insect Mol. Biol.* 30, 566–579. <https://doi.org/10.1111/imb.12726>.
- Borges, R.M. (2005). Do plants and animals differ in phenotypic plasticity? *J. Biosci.* 30, 41–50. <https://doi.org/10.1007/bf02705149>.
- Brouwers, E.V.M. (1984). Glucose/fructose ratio in the food of honeybee larvae during caste differentiation. *J. Apicult Res.* 23, 94–101. <https://doi.org/10.1080/00218839.1984.11100616>.
- Buttstedt, A., Moritz, R.F., and Erler, S. (2013). More than royal food-Major royal jelly protein genes in sexuals and workers of the honeybee *Apis mellifera*. *Front. Zool.* 10, 72. <https://doi.org/10.1186/1742-9994-10-72>.
- Cameron, R.C., Duncan, E.J., and Dearden, P.K. (2013). Biased gene expression in early honeybee larval development. *BMC Genomics* 14, 903. <https://doi.org/10.1186/1471-2164-14-903>.
- Chen, X., Hu, Y., Zheng, H., Cao, L., Niu, D., Yu, D., Sun, Y., Hu, S., and Hu, F. (2012). Transcriptome comparison between honey bee queen- and worker-destined larvae. *Insect Biochem. Mol. Biol.* 42, 665–673. <https://doi.org/10.1016/j.ibmb.2012.05.004>.
- Colombani, J., Raisin, S., Pantalacci, S., Radimerski, T., Montagne, J., and Léopold, P. (2003). A nutrient sensor mechanism controls *Drosophila* growth. *Cell* 114, 739–749. [https://doi.org/10.1016/s0092-8674\(03\)00713-x](https://doi.org/10.1016/s0092-8674(03)00713-x).
- De Coster, W., D'Hert, S., Schultz, D.T., Cruts, M., and Van Broeckhoven, C. (2018). NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics* 34, 2666–2669. <https://doi.org/10.1093/bioinformatics/bty149>.
- Duncan, E.J., Hyink, O., and Dearden, P.K. (2016). Notch signalling mediates reproductive constraint in the adult worker honeybee. *Nat. Commun.* 7, 12427. <https://doi.org/10.1038/ncomms12427>.
- Fesenko, I., Glushkevich, A., and Knyazev, A. (2020). Direct RNA sequencing dataset of SMG1 KO mutant *Physcomitrella* (*Physcomitrium patens*). *Data Brief* 33, 106602. <https://doi.org/10.1016/j.dib.2020.106602>.
- Foret, S., Kucharski, R., Pellegrini, M., Feng, S., Jacobsen, S.E., Robinson, G.E., and Maleszka, R. (2012). DNA methylation dynamics, metabolic fluxes, gene splicing, and alternative phenotypes in honey bees. *Proc. Nat. Acad. Sci. U S A* 109, 4968–4973. <https://doi.org/10.1073/pnas.1202392109>.
- Garalde, D.R., Snell, E.A., Jachimowicz, D., Sipos, B., Lloyd, J.H., Bruce, M., Pantic, N., Admassu, T., James, P., Warland, A., et al. (2018). Highly parallel direct RNA sequencing on an array of nanopores. *Nat. Methods* 15, 201–206. <https://doi.org/10.1038/nmeth.4577>.
- Gingeras, T.R. (2007). Origin of phenotypes: genes and transcripts. *Genome Res.* 17, 682–690. <https://doi.org/10.1101/gr.6525007>.
- Gommans, W.M., Mullen, S.P., and Maas, S. (2009). RNA editing: a driving force for adaptive evolution? *Bioessays* 31, 1137–1145. <https://doi.org/10.1002/bies.200900045>.
- Gommans, W.M., Dupuis, D.E., McCane, J.E., Tatalias, N.E., and Maas, S. (2008). Diversifying exon code through A-to-I RNA editing. In *DNA RNA editing*, H. Smith, ed. (Wiley & Sons, Inc), pp. 3–30.

- Grantham, M.E., and Brisson, J.A. (2018). Extensive differential splicing underlies phenotypically plastic aphid morphs. *Mol. Biol. Evol.* 35, 1934–1946. <https://doi.org/10.1093/molbev/msy095>.
- Guo, X., Su, S., Geir, S., Li, W., Li, Z., Zhang, S., Chen, S., and Chen, R. (2016). Differential expression of miRNAs related to caste differentiation in the honey bee, *Apis mellifera*. *Apidologie* 47, 495–508. <https://doi.org/10.1007/s13592-015-0389-3>.
- Harel, N., Meir, M., Gophna, U., and Stern, A. (2019). Direct sequencing of RNA with MinION Nanopore: detecting mutations based on associations. *Nucleic Acids Res.* 47, e148. <https://doi.org/10.1093/nar/gkz907>.
- He, X.J., Jiang, W.J., Zhou, M., Barron, A.B., and Zeng, Z.J. (2019). A comparison of honeybee (*Apis mellifera*) queen, worker and drone larvae by RNA-Seq. *Insect Sci.* 26, 499–509. <https://doi.org/10.1111/1744-7917.12557>.
- Kucharski, R., Maleszka, J., Foret, S., and Maleszka, R. (2008). Nutritional control of reproductive status in honeybees via DNA methylation. *Science* 319, 1827–1830. <https://doi.org/10.1126/science.1153069>.
- Li, H. (2018). Minimap2: versatile pairwise alignment for nucleotide sequences. *Bioinformatics* 34, 1–6.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>.
- Li, Y., Li-Byarlay, H., Burns, P., Borodovsky, M., Robinson, G.E., and Ma, J. (2013). TrueSight: a new algorithm for splice junction detection using RNA-seq. *Nucleic Acids Res.* 41, e51. <https://doi.org/10.1093/nar/gks1311>.
- Lim, J., Lee, M., Son, A., Chang, H., and Kim, V.N. (2016). mTAIL-seq reveals dynamic poly(A) tail regulation in oocyte-to-embryo development. *Gene Dev.* 30, 1671–1682. <https://doi.org/10.1101/gad.284802.116>.
- Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550. <https://doi.org/10.1186/s13059-014-0550-8>.
- Maleszka, R. (2008). Epigenetic integration of environmental and genomic signals in honey bees: the critical interplay of nutritional, brain and reproductive networks. *Epigenetics* 3, 188–192. <https://doi.org/10.4161/epi.3.4.6697>.
- Mao, W., Schuler, M.A., and Berenbaum, M.R. (2015). A dietary phytochemical alters caste-associated gene expression in honey bees. *Sci. Adv.* 1, e1500795. <https://doi.org/10.1126/sciadv.1500795>.
- Marden, J.H. (2008). Quantitative and evolutionary biology of alternative splicing: how changing the mix of alternative transcripts affects phenotypic plasticity and reaction norms. *Heredity* 100, 111–120. <https://doi.org/10.1038/sj.hdy.6800904>.
- Mello, T.R.P., Aleixo, A.C., Pinheiro, D.G., Nunes, F.M.F., Bitondi, M.M.G., Hartfelder, K., Barchuk, A.R., and Simões, Z.L.P. (2014). Developmental regulation of ecdysone receptor (EcR) and EcR-controlled gene expression during pharate-adult development of honeybees (*Apis mellifera*). *Front. Genet.* 5, 445. <https://doi.org/10.3389/fgene.2014.00445>.
- Modrek, B., and Lee, C. (2002). A genomic view of alternative splicing. *Nat. Genet.* 30, 13–19. <https://doi.org/10.1038/ng102-13>.
- Mutti, N.S., Dolezal, A.G., Wolschin, F., Mutti, J.S., Gill, K.S., and Amdam, G.V. (2011a). IRS and TOR nutrient-signaling pathways act via juvenile hormone to influence honey bee caste fate. *J. Exp. Biol.* 214, 3977–3984. <https://doi.org/10.1242/jeb.061499>.
- Mutti, N.S., Wang, Y., Kaftanoglu, O., Amdam, G.V., and Moritz, R. (2011b). Honey bee PTEN-description, developmental knockdown, and tissue-specific expression of splice-variants correlated with alternative social phenotypes. *PLoS One* 6, e22195. <https://doi.org/10.1371/journal.pone.0022195>.
- Oldham, S., and Hafen, E. (2003). Insulin/IGF and target of rapamycin signaling: a TOR de force in growth control. *Trends Cell Biol.* 13, 79–85. [https://doi.org/10.1016/s0962-8924\(02\)00042-9](https://doi.org/10.1016/s0962-8924(02)00042-9).
- Pan, Q.Z., Wu, X.B., Guan, C., and Zeng, Z.J. (2013). A New method of queen rearing without grafting larvae. *Am. Bee J.* 12, 1279–1280.
- Parker, M.T., Knop, K., Sherwood, A.V., Church, N.J., Mackinnon, K., Gould, P.D., Hall, A.J., Barton, G.J., and Simpson, G.G. (2020). Nanopore direct RNA sequencing maps the complexity of Arabidopsis mRNA processing and m6A modification. *eLife* 9, e49658. <https://doi.org/10.7554/eLife.49658>.
- Patel, A., Fondrk, M.K., Kaftanoglu, O., Emore, C., Hunt, G., Frederick, K., and Amdam, G.V. (2007). The making of a queen: TOR pathway is a key player in diphenic caste development. *PLoS One* 2, e509. <https://doi.org/10.1371/journal.pone.0000509>.
- Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* 14, 417–419. <https://doi.org/10.1038/nmeth.4197>.
- Pertea, G., and Pertea, M. (2020). GFF utilities: GffRead and GffCompare. *F1000 Res.* 9, 304. <https://doi.org/10.12688/f1000research.23297.2>.
- Pertea, M., Pertea, G.M., Antonescu, C.M., Chang, T.C., Mendell, J.T., and Salzberg, S.L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33, 290–295. <https://doi.org/10.1038/nbt.3122>.
- Price, J., Harrison, M.C., Hammond, R.L., Adams, S., Gutierrez-Marcos, J.F., and Mallon, E.B. (2018). Alternative splicing associated with phenotypic plasticity in the bumble bee *Bombus terrestris*. *Mol. Ecol.* 27, 1036–1043. <https://doi.org/10.1111/mec.14495>.
- Reddy, A.S., Marquez, Y., Kalyna, M., and Barta, A. (2013). Complexity of the alternative splicing landscape in plants. *Plant Cell* 25, 3657–3683. <https://doi.org/10.1105/tpc.113.117523>.
- Roach, N.P., Sadowski, N., Alessi, A.F., Timp, W., Taylor, J., and Kim, J.K. (2020). The full-length transcriptome of *C. elegans* using direct RNA sequencing. *Genome Res.* 30, 299–312. <https://doi.org/10.1101/gr.251314.119>.
- Schlichting, C.D., and Smith, H. (2002). Phenotypic plasticity: linking molecular mechanisms with evolutionary outcomes. *Evol. Ecol.* 16, 189–211. <https://doi.org/10.1023/a:1019624425971>.
- Spannhoff, A., Kim, Y.K., Raynal, N.J.M., Gharibyan, V., Su, M.B., Zhou, Y.Y., Li, J., Castellano, S., Sbardella, G., Bedford, M.T., et al. (2011). Histone deacetylase inhibitor activity in royal jelly might facilitate caste switching in bees. *EMBO Rep.* 12, 238–243. <https://doi.org/10.1038/embor.2011.9>.
- Tang, A.D., Soulette, C.M., Baren, M., Hart, K., Hrabeta-Robinson, E., Wu, C.J., and Books, A.N. (2020). Full-length transcript characterization of SF3B1 mutation in chronic lymphocytic leukemia reveals downregulation of retained introns. *Nat. Commun.* 11, 1438.
- Trincado, J.L., Entizne, J.C., Hysenaj, G., Singh, B., Skalic, M., Elliott, D.J., and Eyra, E. (2018). SUPPA2: fast, accurate, and uncertainty-aware differential splicing analysis across multiple conditions. *Genome Biol.* 19, 40. <https://doi.org/10.1186/s13059-018-1417-1>.
- Tudek, A., Lloret-Llinares, M., and Jensen, T.H. (2018). The multitasking polyA tail: nuclear RNA maturation, degradation and export. *Phil. Trans. Biol. Sci.* 373, 20180169. <https://doi.org/10.1098/rstb.2018.0169>.
- Wang, J.R., Holt, J., McMillan, L., and Jones, C.D. (2018). FMLRC: hybrid long read error correction using an FM-index. *BMC Bioinformatics* 19, 50. <https://doi.org/10.1186/s12859-018-2051-3>.
- Wang, L., Wang, S., and Li, W. (2012). RSeQC: quality control of RNA-seq experiments. *Bioinformatics* 28, 2184–2185. <https://doi.org/10.1093/bioinformatics/bts356>.
- Wang, M., Xiao, Y., Li, Y., Wang, X., Qi, S., Wang, Y., Zhao, L., Wang, K., Peng, W., Luo, G.Z., et al. (2021). RNA m6A modification functions in larval development and caste differentiation in honeybee (*Apis mellifera*). *Cell Rep.* 34, 108580. <https://doi.org/10.1016/j.celrep.2020.108580>.
- Winston, M. (1991). *The Biology of the Honey Bee* (Harvard University Press).
- Wojciechowski, M., Lowe, R., Maleszka, J., Conn, D., Maleszka, R., and Hurd, P.J. (2018). Phenotypically distinct female castes in honey bees are defined by alternative chromatin states during larval development. *Genome Res.* 28, 1532–1542. <https://doi.org/10.1101/gr.236497.118>.
- Woo, Y.M., Kwak, Y., Namkoong, S., Kristjansson, K., Lee, S.H., Lee, J.H., and Kwak, H. (2018). TED-Seq identifies the dynamics of Poly(A) length during ER stress. *Cell Rep.* 24, 3630–3641.e7. <https://doi.org/10.1016/j.celrep.2018.08.084>.



Workman, R.E., Tang, A.D., Tang, P.S., Jain, M., Tyson, J.R., Razaghi, R., Zuzarte, P.C., Gilpatrick, T., Payne, A., Quick, J., et al. (2019). Nanopore native RNA sequencing of a human poly(A) transcriptome. *Nat. Methods* 16, 1297–1305. <https://doi.org/10.1038/s41592-019-0617-2>.

Chen, X., Liu, Z., Wang, H., Xiao, C., Ma, C., Chen, C., Chao, C., Lu, Q., Qian, L., Shi, W., and Guo, H. (2017). Integration of lncRNA-miRNA-mRNA reveals novel insights into oviposition regulation in honey bees. *PeerJ* 5, e3881. <https://doi.org/10.7717/peerj.3881>.

Xie, C., Mao, X., Huang, J., Ding, Y., Wu, J., Dong, S., Kong, L., Gao, G., Li, C.Y., and Wei, L. (2011). Kobas 2.0: a web server for annotation and

identification of enriched pathways and diseases. *Nucleic Acids Res.* 39, W316–W322. <https://doi.org/10.1093/nar/gkr483>.

Yamazaki, Y., Shirai, K., Paul, R.K., Fujiyuki, T., Wakamoto, A., Takeuchi, H., and Kubo, T. (2006). Differential expression of HR38 in the mushroom bodies of the honeybee brain depends on the caste and division of labor. *FEBS Lett.* 580, 2667–2670. <https://doi.org/10.1016/j.febslet.2006.04.016>.

Yin, L., Wang, K., Niu, L., Zhang, H., Chen, Y., Ji, T., and Chen, G. (2018). Uncovering the changing gene expression profile of honeybee (*Apis mellifera*) worker larvae Transplanted to queen cells. *Front. Genet.*

9, 416. <https://doi.org/10.3389/fgene.2018.00416>.

Zhang, S., Li, R., Zhang, L., Chen, S., Xie, M., Yang, L., Xia, Y., Foyer, C.H., Zhao, Z., and Lam, H.M. (2020). New insights into *Arabidopsis* transcriptome complexity revealed by direct sequencing of native RNAs. *Nucleic Acids Res.* 48, 7700–7711. <https://doi.org/10.1093/nar/gkaa588>.

Zhang, X., Li, G., Jiang, H., Li, L., Ma, J., Li, H., and Chen, J. (2019). Full-length transcriptome analysis of *Litopenaeus vannamei* reveals transcript variants involved in the innate immune system. *Fish Shellfish Immun.* 87, 346–359. <https://doi.org/10.1016/j.fsi.2019.01.023>.

## STAR★METHODS

### KEY RESOURCES TABLE

| REAGENT or RESOURCE  | SOURCE   | IDENTIFIER              |
|--|--|-------------------------|
| <b>Biological samples</b>                                  |  |                         |
| Honeybee ( <i>Apis mellifera</i> ) queen and worker larvae | This paper   | N/A                     |
| <b>Critical commercial assays</b>                          |  |                         |
| Total RNA Kit I (For DRS)                                  | Omega Bio-tek (USA)  | Cat#:R6834              |
| NEBNext Poly(A) mRNA Magnetic Isolation Module             | NEB (USA)  | Cat#:E7490S             |
| Oxford Nanopore DRS protocol                               | Oxford Nanopore Technologies (UK)                                    | Cat#:SQK-RNA002         |
| Qubit RNA HS Assay Kit                                     | Thermo Fisher Scientific (USA)                                       | Cat#: Q32852            |
| Reverse Transcriptase Adapter                              | Oxford Nanopore Technologies (UK)                                    | Cat#:SQK-RNA002         |
| dNTP solution  | New England Biolabs (USA)  | Cat#:NEB N0447          |
| Quick Ligation Module                                      | New England Biolabs (USA)  | Cat#: E6056             |
| Super-Script III Reverse Transcriptase                     | Thermo Fisher Scientific (USA)                                       | Cat#: 18080044          |
| RNA Adapter Mix  | Oxford Nanopore Technologies (UK)                                    | Cat#:SQK-RNA002         |
| TRIzol reagent (For RNA-Seq)                               | Tiagen (China)   | Cat#: DP424             |
| oligo (dT) magnetic beads                                  | Vazyme (China)   | Cat#: N401-2            |
| RNA Reverse Transcriptase                                  | Thermo Fisher Scientific (USA)                                       | Cat#: 18080044          |
| DNA polymerase I system                                    | New England Biolabs (USA)  | Cat#: M0209S            |
| AMPure XP beads  | Beckman Coulter (USA)  | Cat#:A63882             |
| Taq 2X Master Mix  | New England Biolabs (USA)  | Cat#:M0287              |
| Primer-Script RT reagent Kit                               | TaKaRa (Japan)   | Cat#:RR037A             |
| AceQ universal U+ probe master mix V2                      | Vazyme (China)   | Cat#:Q513-01            |
| 50×TAE buffer solution                                     | Solarbio (China)   | Cat#:T1060              |
| 6×DNA Loading Buffer                                       | TransGen (China)   | Cat#:GH101-01           |
| 50 bp DNA Ladder   | Tiagen (China)   | Cat#:MD108              |
| Gsafe Red plus nucleic acid dye                            | Glpbio (USA)   | Cat#:GK20002            |
| <b>Deposited data</b>                                      |  |                         |
| Direct RNA sequencing data                                 | NCBI database  | BioProject: PRJNA748829 |
| RNA-Seq data   | NCBI database  | BioProject: PRJNA748829 |
| <b>Experimental models: Organisms/strains</b>              |  |                         |
| Honey bees ( <i>Apis mellifera</i> )                       | Honeybee Research institute, Jiangxi Agricultural university (China) | N/A                     |
| <b>Software and algorithms</b>                             |  |                         |
| Guppy  | <a href="#">Fesenko et al. (2020)</a>                                | version 3.2.6           |
| Nanofilt   | <a href="#">Coster et al. (2018)</a>                                 | version 2.7.1           |
| Fclmr2   | <a href="#">Wang et al. (2018)</a>                                   | version 0.1.2           |
| Minimap2   | <a href="#">Li (2018)</a>  | version 2.17-r941       |
| Samtools   | <a href="#">Li et al. (2009)</a>                                     | version 1.10            |
| Flair collapse   | <a href="#">Tang et al. (2020)</a>                                   | version 1.5.0           |
| Stringtie  | <a href="#">Pertea et al. (2015)</a>                                 | version 2.1.4           |
| Gffcompare   | <a href="#">Pertea and Pertea (2020)</a>                             | version 0.12.1          |
| Salmon   | <a href="#">Patro et al. (2017)</a>                                  | version 1.4.0           |

(Continued on next page)

| <b>Continued</b>               |  |                   |
|--------------------------------|--|-------------------|
| <b>REAGENT or RESOURCE</b>     | <b>SOURCE</b>                          | <b>IDENTIFIER</b> |
| DESeq2                         | <a href="#">Love et al. (2014)</a>     | version 1.26.0    |
| KOBAS                          | <a href="#">Xie et al. (2011)</a>      | version 2.0       |
| SUPPA2                         | <a href="#">Trincado et al. (2018)</a> | version 2.3       |
| Nanopolish                     | <a href="#">Coster et al. (2018)</a>   | version 0.13.2    |
| R package                      | R Core Team                            | version 4.0.3     |
| Statview                       | SAS institute Inc                      | version 5.0.0.0   |
| PrimerPremier 5                | Premier (Canada)                       | version 5.0       |
| <b>Others</b>                  |  |                   |
| Queen cell frame and excluder  | <a href="#">Pan et al., (2013)</a>     | N/A               |
| Worker cell frame and excluder | <a href="#">Pan et al., (2013)</a>     | N/A               |

## RESOURCE AVAILABILITY

### Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Prof. Zhi Jiang Zeng ([bees1965@sina.com](mailto:bees1965@sina.com)).

### Materials availability

This study did not generate new unique reagents.

### Data and code availability

The DRS sequence data for all 12 libraries and the fastq sequences of RNA-Seq of 18 libraries have been deposited in the NCBI database under the accession: NCBI Bioproject PRJNA748829. All data reported in this paper will be shared by the [lead contact](#) upon request. This paper does not report original code. Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

## EXPERIMENTAL MODEL AND SUBJECT DETAILS

### Insects

Three healthy honeybee colonies (*Apis mellifera*) each with a mated queen were used for this study. Each colony had ten frames with approximately 35,000 bees. These colonies were maintained at the Honeybee Research Institute, Jiangxi Agricultural University, Nanchang, China, according to the standard beekeeping techniques.

## METHOD DETAILS

### Larva sampling

The mother queen was caged onto a plastic worker-cell frame developed by [Pan et al. \(2013\)](#) to lay eggs into worker cells for 6 h. Afterwards, half of these eggs were removed into commercial plastic queen cells to rear queens before hatching. Queen and worker larvae at 48 h (2 days), 96hrs (4 days) and 144 h (6 days) after hatching were harvested and were flash-frozen in liquid nitrogen. In the DRS experiment, for the 2 days queen larvae (2Q) or 2 days worker larvae (2W), larvae were very small, therefore we collected twenty 2days queen larvae and mixed them into one sample. The 4d and 6d larvae were large enough, and one sample needed only one larva. Each group had two replicates from two colonies, therefore in total 12 samples were collected. Larval samples for short-length RNA-Seq were sampled the same as DRS, but with twice the number of larvae in each sample. We had three replicates for each of these 6 groups: 18 samples in total. Within a colony queens and workers developed from a same mother queen and eggs were laid at the same time. All sampled eggs were reared in a same colony. We compared queen and worker larvae sampled at the same time point: 2Q vs 2W, 4Qvs 4W, and 6Q vs 6W.

### RNA preparation

Total RNA of 2d, 4d and 6d queen and worker larvae was extracted in accordance with the standard protocol of the Total RNA Kit I (R6834, Omega Bio-tek), respectively. Each sample contained 30 µg total RNA for DRS. RNA was cleaned and concentrated by NEBNext Poly(A) mRNA Magnetic Isolation Module (E7490S) according to the manufacturer's instructions. From each sample we extracted 1 µg of total RNA for the RNA purity and concentration measurement using Nanodrop (Thermo Fisher Scientific). RNA integrity was checked using an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA). The RNA samples with RNA Integrity Number (RIN) >8.5 and  $2.0 < OD_{260/280} < 2.2$  (Zhang et al., 2019) were used for DRS library preparation.

### Direct RNA sequencing

Prepared RNA of each sample was used for a DRS library preparation using the Oxford Nanopore DRS protocol (SQK-RNA002, Oxford Nanopore Technologies). For reversed connector connection, 9 µL prepared RNA, 3 µL NEBNext Quick Ligation Reaction Buffer (NEB), 1 µL RT Adapter (RTA) (SQK-RNA002) and 2 µL T4 DNA Ligase (NEB) were mixed together and incubated under 25°C for 10 min. Afterwards, 8 µL 5× first-strand buffer (NEB), 2 µL 10 mM dNTPs (NEB), 9 µL Nuclease-free water, 4 µL 0.1 M DTT (Thermo Fisher) and 2 µL Super-Script III Reverse Transcriptase (Thermo Fisher Scientific, 18,080,044) were added into the above 15 µL reaction system and incubated under 50°C for 50 min then 70°C for 10 min. Reverse-transcribed mRNA was purified with 1.8× Agencourt RNAClean XP beads and washed with 23 µL Nuclease-free water, and subsequently sequencing adapters were added using 8 µL NEBNext Quick Ligation Reaction Buffer, 6 µL RNA Adapter (RMX) and 3 µL T4 DNA Ligase. The mix was purified and washed again as above, and 75 µL RRB (SQK-RNA002) were added with 35 µL Nuclease-free water. This final reaction system was loaded into Nanopore R9.4 sequencing micro-array and sequenced for 48–72 h using PromethION sequencer (Oxford Nanopore Technologies). This was performed by Wuhan Benagen Tech Solutions Company Limited.

### Preprocessing and alignments

To assess the read quality, raw reads were base called using Guppy software (version 3.2.6) (Fesenko et al., 2020) and adapter sequences were trimmed by Guppy as well. Low quality reads (Q-value < 7) and short-length reads (<50 nt) were filtered by Nanofilt (version 2.7.1) (Coster et al., 2018). The remaining clean reads were corrected using the Fclmr2 (version 0.1.2) with our short-length RNA-Seq data (Wang et al., 2018). Afterwards, clean reads from each library were aligned to the genome of honeybees *Apis mellifera* (Amel HAv3.1) using Minimap2 (version 2.17-r941) (Li, 2018). The alignment ratio of clean reads to the honeybee reference genes were calculated using Samtools (version 1.10) (Li et al., 2009) and showed in Table 1. After clearing redundant sequences using Flair collapse (version 1.5.0) (Tang et al., 2020) and Stringtie (version 2.1.4) (Pertea et al., 2015), the clean reads from each library were merged by Gffcompare (version 0.12.1) (Pertea and Pertea, 2020) to obtain the final isoforms. To test the RNA integrity of DRS samples, raw reads were used to calculate its distribution density on the gene body of honey bee genome by RSeQC software (4.0.0 version, Wang et al., 2012), and results were shown in Figure S2.

### DEG and DEI calculation

Isoform expression was annotated to the honeybee reference genes and quantified using the selective-alignment-base model in Salmon software (version 1.4.0) (Patro et al., 2017). The raw sequence and annotation information of each isoform refer to Tables S16 and S17 respectively. Gene expression was measured according to standard methods as in He et al. (He et al., 2019) using DRS raw long reads rather than RNA-Seq data, and then Significantly differentially expressed genes (DEGs) between queen and worker larvae were identified as those with p values <0.05 and log2 Fold change values >1.5 using DESeq2 (version 1.26.0) (Love et al., 2014). Since DRS provides the expression of each isoform, therefore significantly differentially expressed isoforms (DEIs) between queen and worker larvae were identified using the same analysis (p values <0.05 and log2 Fold change values >1.5 using DESeq2). To compare the expression level of genes/isoforms, we computed TPM values (transcripts per kilobase of exon model per million mapped reads) based on reads counts as expression. Since it could not compare the DE-Gs and DE-Is directly, therefore, all DEIs were referenced to their related genes to obtain differentially expressed isoforms of genes (DEIGs). Sequences of the DE-Gs and DEIGs were blasted to the Swiss-Prot database, non-redundant protein (Nr) and non-redundant nucleotide sequence (Nt) database with a cut-off E-value of  $10^{-5}$ .

### KEGG pathway enrichment analysis

Five KEGG pathways [TOR, FoxO, Notch, Wnt and IIS] which have been repeatedly associated with honeybee queen-worker differentiation (Patel et al., 2007; Mutti et al., 2011a; Duncan et al., 2016; Xiao et al., 2017; Yin et al., 2018; Wang et al., 2021) were selected and combined into a pathway net according to KEGG pathway database and a previous study (Wang et al., 2021). The statistical enrichment of DEIGs and DEGs in these KEGG pathways were analyzed using a hypergeometric test (Q-value < 0.05) using the KOBAS 2.0 software (Xie et al., 2011). Results were shown in Figure 3A. Expression of DE-Is and genes in these five pathways were also presented in Figure 3B.

### Alternative splicing identification

The AS events of each comparison were analyzed using SUPPA2 (version 2.3) (Trincado et al., 2018). AS events were divided into seven patterns including 3'splice site (A3), 5'splice site (A5), First exon (AF), Last exon (AL), Retained intron (RI), Skipping exon (SE) and Mutually exclusive exon (MX). The PSI values of each AS events in DE-Is were calculated after filtering percent spliced in index (PSI) values with 'not available' (NA) or '0' (Trincado et al., 2018). Significant differences in AS events were identified as its dPSI-value >0 and p value <0.05.

### Poly(A) length estimation and correlation with expression of genes and isoforms

Poly(A) lengths of each read were calculated by the Nanopolish software (version 0.13.2) (Coster et al., 2018). The 3' poly(A) tail lengths of each read were directly identified using the ionic current signals (Workman et al., 2019). Low-quality poly(A) tails was filtered by the "qc\_tag" software in Nanopore sequencing platform, and only the PASS reads were used for further analysis. A previous study showed that poly(A) lengths were negatively correlated with isoform expression, therefore, the median of poly(A) length of each isoform was used for correlation and regression analysis with their TPM values (results were shown in Figure S5) in R package (version 4.0.3). We also evaluated the correlation between poly(A) tail lengths and DEIs or DEGs.

For this we used median of poly(A) length and TPM or log2Fold change values of DEIs or DEGs for correlation and regression analysis as above. Two key genes (*Jhe* and *Ecr*) for the determination of queen-worker differentiation were selected as an example to show the correlation between expression of isoforms and poly(A) lengths. Results were shown in Figure 6.

### Short-length RNA sequencing

Total RNA of each sample was extracted using TRIzol reagent (Tiangen, Beijing). The concentration and integrity of RNA were assessed using Qubit 3.0 Fluorometer and Agilent 2100 bioanalyzer. Each group of queen and worker larvae had 3 biological replicates.

Total mRNA of each sample was enriched using oligo (dT) magnetic beads. First-strand cDNA was synthesized by random hexamers, and the second-strand cDNA was synthesized in DNA polymerase I system using dNTPs and RNaseH. Afterwards, cDNA was subjected to end repair and performed with poly(A) tail and ligation sequencing adapter. 200–350 bp cDNA were selected by AMPure XP beads, and then were PCR amplified and purified by AMPure XP beads. Library quality was evaluated using Agilent 2100 bioanalyzer and qRT-PCR. Totally 18 libraries were sequenced by an Illumina NovaSeq 6000 platform. This was performed by Wuhan Benagen Tech Solutions Company Limited also.

### TaqMan real-time PCR

The honeybee sample collection method (2d, 4d and 6d queen and worker larvae) was same to that of RNA-Seq, and all samples were from the same colonies of DRS and RNA-Seq. Each larval group had 3 biological replicates, and each biological replicate had 4 technical replicates in the TaqMan real-time PCR. Total RNA of each sample was extracted using TRIzol reagent (Tiangen, Beijing). The cDNA of each sample was synthesized from the total RNA using the Primer-Script RT reagent Kit (TaKaRa) according to the manufacturer's instructions.

Two isoforms of honeybee *Ecr* gene (*Ecr.t1* and *Ecr.t2*) were selected for the TaqMan real time PCR verification experiment, and the honeybee GAPDH gene was selected as the house keeping gene. Since the sequences of *Ecr.t1* and *Ecr.t2* were highly coincident, the TaqMan probes were designed spanning two

isoform-specific exons in order to specifically detect these isoforms. The primers were designed that its products fully covered the probe detecting sequence. The fluorophore and quenching group of probes were 5'-FAM (green, excitation wavelength: 494 nm, emission wavelength: 522) and BHQ-1 respectively. The primers and probes were designed using Primerpremier 5 (version 5.0) and produced by Shanghai GenePharma Co., Ltd (Shanghai, China).

CFX Connect™ real-time PCR (Bio-Rad, USA) was used for this study. Each 20 µL reaction tube contained 10 µL AceQ Universal U+ Probe Master Mix V2, 3 µL cDNA, 0.4 µL Primer Set (10 µM), 0.4 µL probe (10 µM) and 6.2 µL RNase Free ddH<sub>2</sub>O. The cycling conditions were: 3 min incubation period at 95°C followed by 40 cycles of PCR, each cycle consisting of 12 s at 94°C, 30 s at 59°C and 30 s at 72°C. A cycle threshold (Ct) was calculated by determining the point at which the fluorescence exceeded a threshold limit.

### QUANTIFICATION AND STATISTICAL ANALYSIS

For the data analysis of TaqMan real-time PCR, relative expression of these two isoforms (*Ecr.t1* and *Ecr.t2*) were calculated using  $2^{-\Delta\Delta C_t}$  comparative Ct method and were transformed by taking their square root to be normally distributed. Data were analyzed by one-way ANOVA test in Statview package (version 5.0.0.0, SAS institute inc). The p value < 0.05 was considered as significantly different. All groups (2Q, 2W, 4Q, 4W, 6Q and 6W) had three biological replicates from three different honeybee colonies, and each biological replicate had four technical replicates.