

RESEARCH

Open Access



# Examining parent-of-origin effects on transcription and RNA methylation in mediating aggressive behavior in honey bees (*Apis mellifera*)

Sean T. Bresnahan<sup>1\*</sup>, Ellen Lee<sup>2,3</sup>, Lindsay Clark<sup>4,5</sup>, Rong Ma<sup>1</sup>, Juliana Rangel<sup>6</sup>, Christina M. Grozinger<sup>1</sup> and Hongmei Li-Byarlay<sup>2,7\*</sup>

## Abstract

Conflict between genes inherited from the mother (matrigenes) and the father (patrigenes) is predicted to arise during social interactions among offspring if these genes are not evenly distributed among offspring genotypes. This intragenomic conflict drives parent-specific transcription patterns in offspring resulting from parent-specific epigenetic modifications. Previous tests of the kinship theory of intragenomic conflict in honey bees (*Apis mellifera*) provided evidence in support of theoretical predictions for variation in worker reproduction, which is associated with extreme variation in morphology and behavior. However, more subtle behaviors – such as aggression – have not been extensively studied. Additionally, the canonical epigenetic mark (DNA methylation) associated with parent-specific transcription in plant and mammalian model species does not appear to play the same role as in honey bees, and thus the molecular mechanisms underlying intragenomic conflict in this species is an open area of investigation. Here, we examined the role of intragenomic conflict in shaping aggression in honey bee workers through a reciprocal cross design and Oxford Nanopore direct RNA sequencing. We attempted to probe the underlying regulatory basis of this conflict through analyses of parent-specific RNA m6A and alternative splicing patterns. We report evidence that intragenomic conflict occurs in the context of honey bee aggression, with increased paternal and maternal allele-biased transcription in aggressive compared to non-aggressive bees, and higher paternal allele-biased transcription overall. However, we found no evidence to suggest that RNA m6A or alternative splicing mediate intragenomic conflict in this species.

**Keywords** Honey bee, Intragenomic conflict, Transcriptomics, Oxford Nanopore direct RNA sequencing, Alternative splicing, Allele-biased transcription, RNA m6A

\*Correspondence:

Sean T. Bresnahan  
stb5321@psu.edu  
Hongmei Li-Byarlay  
hli-byarlay@centralstate.edu

<sup>1</sup>Huck Institutes of the Life Sciences, Pennsylvania State University, University Park, USA

<sup>2</sup>Agricultural Research and Development Program, Central State University, Wilberforce, USA

<sup>3</sup>Department of Biological Sciences, Wright State University, Dayton, USA

<sup>4</sup>HPCBio, University of Illinois at Urbana-Champaign, Champaign, USA

<sup>5</sup>Research Scientific Computing Group, Seattle Children's Research Institute, Seattle, USA

<sup>6</sup>Department of Entomology, Texas A&M University, College Station, USA

<sup>7</sup>Department of Agricultural and Life Science, Central State University, Wilberforce, USA



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Background

Conflict between genes inherited from the mother (matrigenes) and the father (patrigenes) is predicted to arise over parental investment in offspring or during social interactions among offspring, if these genes are not evenly distributed among offspring genotypes [1, 2]. According to the kinship theory of intragenomic conflict, matrigenes and patrigenes can each favor behaviors that promote their own respective fitness to the detriment of the other. This conflict drives the evolution of parent-specific transcription patterns in offspring that result from parent-specific epigenetic modifications [3]. Studies of intragenomic conflict and its epigenetic basis are therefore useful for revealing the molecular mechanisms that underlie the transgenerational inheritance of social behavior traits [4]. Given that matrigenes and patrigenes represent two transcriptional outcomes at individual genomic loci that can be assessed simultaneously within the same individual, studies of intragenomic conflict can also elucidate the regulation of gene expression more generally [5].

Honey bees (*Apis mellifera*) are an excellent model system in which to explore the mechanisms mediating intragenomic conflict, including parent-specific transcription and its epigenetic basis. The kinship theory of intragenomic conflict predicts that paternal alleles in polyandrous social insects will favor enhanced activity of reproductive traits [6]. For instance, honey bee workers typically increase their inclusive fitness by cooperating to support the queen's reproductive fitness, even at the cost of their own personal reproductive fitness. However, under queenless conditions, cooperation between workers shifts to reproductive competition, whereby workers with developed ovaries aggress their sisters and lay unfertilized eggs that develop into haploid male drones [7–11]. Similarly, virgin (unmated) honey bee queens will fight to maintain their dominance by stinging other queens detected within the hive [12–15]. Thus, the alleles of genes underlying aggressive behaviors that enhance reproductive fitness may be in conflict, with paternal alleles favoring aggression and maternal alleles favoring cooperation.

Previous studies in honey bees have found evidence for parent-of-origin related variation in reproductive morphology, physiology, transcription [8, 16–22], and aggressive behavior [14, 23], all of which are consistent with the kinship theory of intragenomic conflict [6, 24]. These studies evaluated workers from reciprocal crosses of "Africanized" honey bees (AHBs, which are derived from the subspecies *A. m. scutellata*) and European honey bees (EHBs, which are derived from subspecies from Europe, such as *A. m. ligustica*) – two strains of honey bee that vary in their genotype, physiology, and behavior [25, 26]. The activation and size of ovaries in reciprocal crosses of

AHBs and EHBs is greater in workers with AHB paternity as compared to EHB paternity, and this phenotypic difference is associated with enrichment for paternal allele-biased transcription in the ovaries, fat bodies [21], and brains [22]. Moreover, studies of worker aggression in hybrid crosses of AHBs and EHBs suggest that aggressive behaviors are increased by AHB paternity [14, 27–29]. Therefore, in this study, we tested the hypothesis that aggression (specifically focusing on aggressive stinging behavior in defense of the colony) is increased in soldier-aged workers with AHB paternity relative to workers with EHB paternity, and that this pattern is associated with an enrichment for paternal allele-biased transcription.

For most of the genes exhibiting parent-specific transcription in plants and mammals, differential DNA methylation of promoters or associated noncoding regulatory regions underlies their parent-of-origin transcription [30, 31]. Consistent with predictions of the kinship theory of intragenomic conflict [6, 24], previous studies have demonstrated that parent-specific transcription occurs in multiple tissues and developmental stages in social insects, including honey bees [17, 19, 21, 22] and bumble bees [32]. However, these transcriptional patterns are not associated with allelic DNA methylation differences in either species [32, 33]. This is perhaps unsurprising, given that differential DNA methylation does not appear to be associated with transcriptional variation in social insects specifically [34], or insects generally [35]. Thus, how parent-specific transcription is regulated in social insects remains an open area of investigation [36].

RNA methylation has recently been recognized as an important gene regulatory mechanism [37–39]. N6-methyladenosine (m6A) is the most abundant eukaryotic RNA modification, accounting for more than 80% of all RNA methylation [40]. In mammals, it plays a role in neurogenesis and morphology [41] and alters synaptic transmission by regulating transcript abundance of protein coding genes involved in neuronal signaling [42]. m6A affects RNA structure, splicing, localization, translation, stability, and metabolism [43, 44] and thus, allele-specific m6A modifications may drive differences in allele-specific transcript abundance [45]. In honey bees, recent studies revealed that chemical suppression of m6A impacts larval development and caste determination [46] and that fat body and brain tissues in workers show variation in global levels of m6A [47]. Whether there is a relationship between m6A and behavior in honey bees has yet to be investigated, however. Here, we used Oxford Nanopore direct RNA sequencing to assess (1) the relationship between parent-specific m6A and parent-specific transcription and (2) the association between each of these mechanisms with parent-of-origin effects on an aggressive behavior in soldier-aged worker honey bees from reciprocal crosses of AHBs and EHBs. We tested

the hypothesis that parent-specific m6A underlies parent-specific transcription, with higher transcript abundance of the unmethylated allele.

Additional layers of gene regulation – including alternative splicing and the activities of long intergenic non-coding RNAs (lincRNAs) – have also revealed insights into the molecular basis of honey bee social behavior [48–51]. Work in other insect and mammalian model species have revealed roles for lincRNAs in development, disease, behavior, and metabolism [52–55]. With few exceptions [56], most studies of alternative splicing and long noncoding RNAs in honey bees have used short-read sequencing or cDNA microarrays [57, 58], which are limited to detecting only full-length isoforms or alternative splicing events [59, 60]. Here, we utilized long-read sequence data to profile full-length non-coding RNA molecules to study their relevance to parent-of-origin effects on honey bee aggression. We described the alternative splicing and lincRNA profiles of soldier-aged worker honey bees, compared between aggressive and non-aggressive individuals, and assessed the relationships between alternative splicing, parent-specific m6A, and parent-specific transcript abundance. We tested the specific hypothesis that genes which show parent-specific m6A or transcript abundance also show alternative splicing.

## Results

### Parent-of-origin effects on honey bee worker aggression

In contrast with previous work [14, 27–29], we observed inconsistent parent-of-origin effects on worker aggression across our four experimental colonies (Table 1). Colony 1 showed significantly higher aggression in workers with EHB paternity than those with AHB paternity, whereas Colony 2 (from which we isolated RNA from the head and thorax of individual bees for sequencing) and Colony 3 showed higher aggression in workers with AHB paternity. In Colony 4, aggression was slightly higher in workers with AHB paternity, but this difference was not statistically significant. When data from all colonies were combined, there was no statistically significant parent-of-origin effect on aggression.

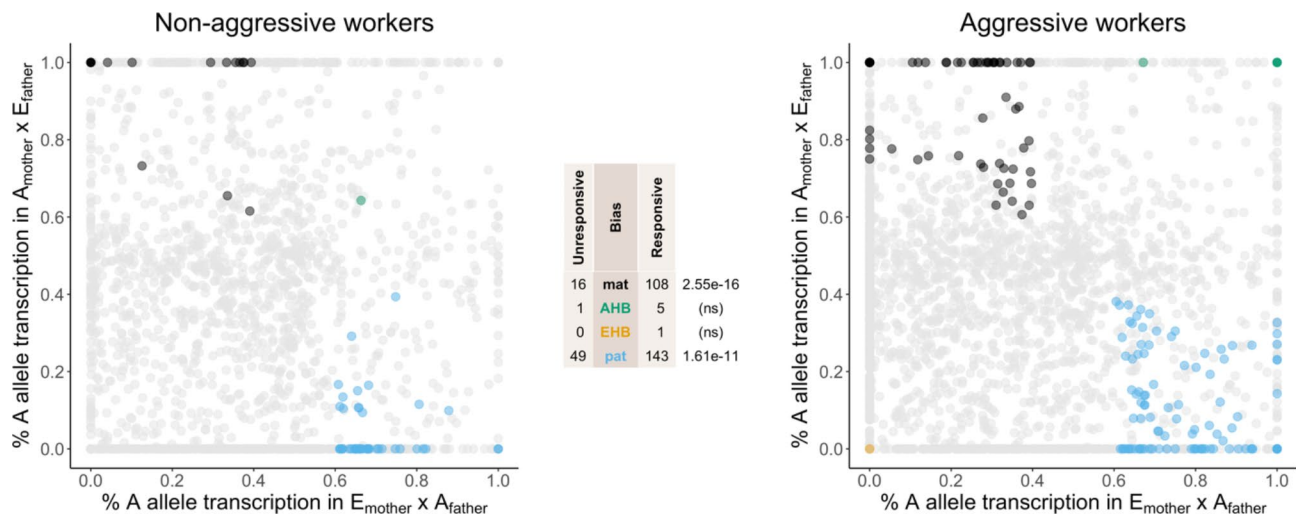
### Parent-specific allele-biased transcription associated with worker aggression

At approximately 60x genome coverage, we detected an average of 1.24 million homozygous single nucleotide polymorphisms (SNPs) per diploid queen and 2.29 million SNPs per haploid drone. We identified a total of 31,078 unique transcripts in our samples, including 3,640 lincRNAs. Of these transcripts, 3,081 (9.91%) were shared between both crosses and sufficiently varied in their sequence identity between the parents of each cross, allowing for identification of parent-of-origin reads in the workers. Specifically, we identified 35,782 SNP positions within transcripts that had at least 2 SNPs. Of the 3,081 transcripts, 2,584 were from the published annotation (NCBI *Apis mellifera* Annotation Release 104), and 497 were novel. After filtering SNP positions with low read counts in the workers, our dataset contained read counts at 18,674 positions distributed among 1,928 transcripts (approximately 9.69 SNPs per transcript) in non-aggressive workers and 33,494 positions distributed among 2,820 transcripts (approximately 11.88 SNPs per transcript) in aggressive workers.

We identified 312 transcripts that showed allelic bias in both crosses (Fig. 1), including 231 from previously annotated genes, 68 novel transcripts, and 13 lincRNAs. Some transcripts showed allele-biased transcription in both non-aggressive and aggressive workers ( $n=11$ ), whereas other transcripts showed allelic bias in only one group (non-aggressive only,  $n=55$ ; aggressive only,  $n=246$ ). In total, in non-aggressive workers, 66/1,928 (3.42%) of tested transcripts showed allelic bias, whereas in aggressive workers, 256/2,820 (9.111%) of tested transcripts showed allelic bias. See Supplementary Dataset 1 (Table S16) for a complete listing of the parent and lineage biased transcripts identified in this study. In support of our hypothesis, we found that paternal allele-biased transcripts were enriched in aggressive workers ( $\chi^2 p=1.61 * 10^{-11}$ ). Interestingly, maternal allele-biased transcripts were also enriched in aggressive workers ( $\chi^2 p=2.55 * 10^{-16}$ ), although there were fewer maternal allele-biased transcripts overall. Cytoplasm cellular compartment (GO:0005737) was overrepresented for paternal

**Table 1** Frequency of aggressive stinging observed in workers from reciprocal crosses of EHBs and AHBs.

Block	Colony	% of individuals with EHB mother observed stinging		% of individuals with AHB mother observed stinging		$\chi^2$ test
			SD		SD	<i>p</i> -value
A	1	12.62%		0.00%		$3.77 \cdot 10^{-13}$
	2	2.43%		7.89%		
B	3	1.11%		15.77%		$2.03 \cdot 10^{-8}$
	4	4.39%		5.00%		
Combined	Average		SD	Average	SD	Logistic regression <i>p</i> -value
	% of individuals with EHB mother observed stinging	5.17%	5.14%	% of individuals with AHB mother observed stinging	7.17%	



**Fig. 1** Worker aggression is associated with increased maternal and paternal allele-biased transcription. Transcript abundance of parent and lineage alleles were assessed in non-aggressive and aggressive workers from an EHB and AHB reciprocal cross. The x-axis represents, for each transcript (1,928 in non-aggressive workers and 2,820 in aggressive workers), the average proportion of AHB reads in workers with an EHB mother and AHB father ( $p_1$ ), and the y-axis represents, for each transcript, the average proportion of AHB reads in workers with an AHB mother and EHB father ( $p_2$ ). Each color represents a transcript which is significantly biased at all tested SNP positions: black is maternal, green is AHB, gold is EHB, blue is paternal, and gray is not significant. Significance was determined using the overlap between two statistical tests: a generalized linear interactive mixed model (GLIMMIX) [17, 21, 22], and a Storer-Kim test along with previously established cutoff thresholds [61] of  $p_1 < 0.4$  and  $p_2 > 0.6$  for maternal bias,  $p_1 > 0.6$  and  $p_2 < 0.4$  for paternal bias,  $p_1 < 0.4$  and  $p_2 < 0.4$  for EHB bias, and  $p_1 > 0.6$  and  $p_2 > 0.6$  for AHB bias

allele-biased transcripts in aggressive workers. No biological process terms or KEGG pathways were associated with paternal allele-biased transcripts in non-aggressive workers, or with maternal allele-biased transcripts in either non-aggressive or aggressive workers.

We found little overlap between the genes showing allele-biased transcription associated with aggression in our study and (1) genes identified in a previous study [23] as showing parent-of-origin expression in guard bees, and (2) genes within quantitative trait loci (Sting 1–3) associated with aggressive behaviors in honey bees [62]. Specifically, we identified four genes showing allele-biased transcription in both our study and Gibson et al., only two of which were biased toward the same allele (Supplementary Dataset 1, Table S18). No overlap was found between genes showing allele-biased transcription in our study and genes within the Sting 1–3 QTLs (Supplementary Dataset 1, Table S19).

#### Parent-specific RNA m6A associated with worker aggression

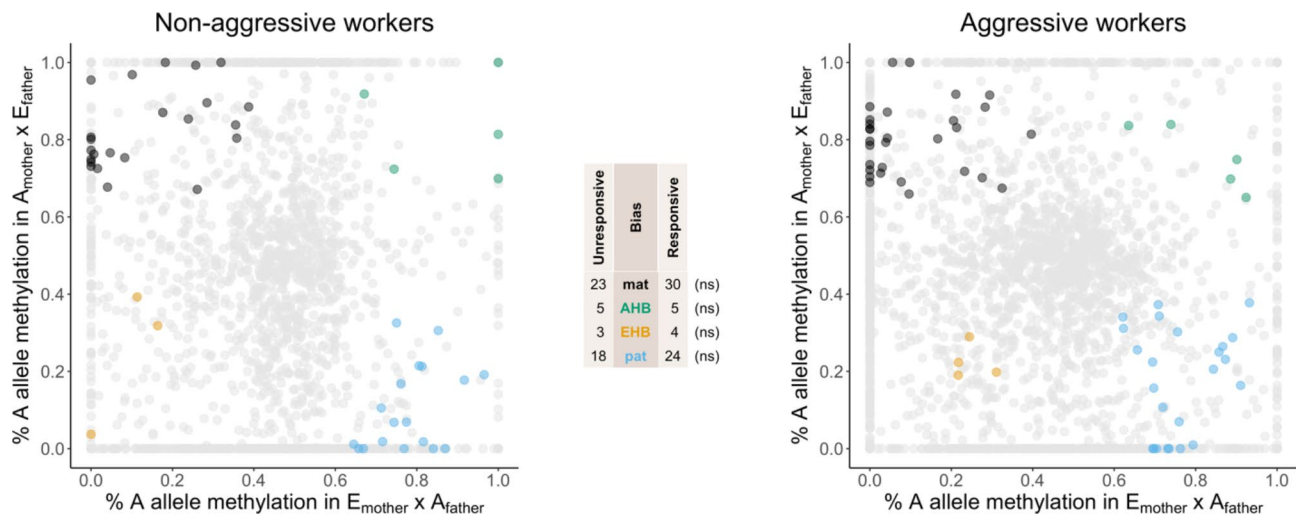
We identified 103 transcripts that showed parent or lineage biases in RNA m6A in both crosses (Fig. 2), including 87 from previously annotated genes, 15 novel transcripts, and 1 lincRNA. Some transcripts showed allele-biased m6A in both non-aggressive and aggressive workers ( $n=9$ ), whereas others showed allelic bias in only one group (non-aggressive only,  $n=40$ ; aggressive only,  $n=54$ ). See Supplementary Dataset 1 (Table S17) for a complete listing of the transcripts with parent

and lineage biases in m6A identified in this study. In contrast to transcription, there was no relationship between parent-specific m6A and worker aggression. Additionally, there were no gene ontology terms overrepresented for genes showing paternal nor maternal allele-biased m6A.

Only three genes which showed allele-biased transcription also showed allele-biased m6A, each of which showed bias toward the paternal allele in both transcript abundance and m6A levels. In aggressive workers, this included one of two transcript isoforms for elongation of very long chain fatty acids protein AAEL008004 (LOC724552; GB54396). In non-aggressive workers, this included the transcript for an uncharacterized gene (LOC726321) and one of four transcript isoforms for vitellogenin (LOC406088; GB13999; GB49544).

#### Differential expression

We detected transcript abundance of 19,092 genes and 3,640 lincRNAs in our samples. After filtering low-count genes, our dataset consisted of transcript levels of 14,462 genes and 2,415 lincRNAs. Of these, 8,058 (55.72%) of the genes and 451 (18.67%) of the lincRNAs were from the published annotation. When the crosses were analyzed separately, we identified one significantly differentially expressed gene (DEG) at a false discovery rate (FDR)  $< 0.05$  between the non-aggressive and aggressive bees in each cross: box A-binding factor (LOC725389; GB50932) in cross A, and myosinase 1 (LOC411978; GB54486) in cross B. However, a full model accounting for differences between crosses revealed that there were



**Fig. 2** Worker aggression is not associated with an increase in parent or lineage specific RNA m6A. Read abundance and m6A rate of parent and lineage alleles were assessed in non-aggressive and aggressive workers from an EHB and AHB reciprocal cross. The x-axis represents, for each transcript (1,928 in non-aggressive workers and 2,820 in aggressive workers), the average proportion of AHB reads that were methylated in workers with a EHB mother and AHB father, whereas the y-axis represents, for each transcript, the proportion of AHB reads that were methylated in workers with an AHB mother and EHB father. Each color represents a transcript which is significantly biased at all tested SNP positions: black is maternal, green is AHB, gold is EHB, blue is paternal, and gray is not significant. Significance was determined using an unpooled two-tailed z-test [63] along with previously established cutoff thresholds [61] of  $p1 < 0.4$  and  $p2 > 0.6$  for maternal bias,  $p1 > 0.6$  and  $p2 < 0.4$  for paternal bias,  $p1 < 0.4$  and  $p2 < 0.4$  for EHB bias, and  $p1 > 0.6$  and  $p2 > 0.6$  for AHB bias

no significant DEGs between these behavioral phenotypes in our study.

#### Detection of lincRNAs and isoform switching

In total, we detected 2,415 lincRNAs from all samples, 11 of which were differentially expressed (see Supplementary Dataset 1, Table S3). However, none of these showed parent-specific transcription or m6A biases. Additionally, we detected few genes that were significant for alternative splicing or isoform switching (see Supplementary Dataset 1, Table S2). Five genes were detected which had significant isoform switches ( $FDR < 0.1$ ) between non-aggressive and aggressive bees, with putative functional consequences: synaptic vesicle glycoprotein 2B (LOC409924; GB41065), adenylosuccinate synthetase (LOC409299; GB14705; GB43704), H(+)/Cl(-) exchange transporter 7 (LOC413069; GB51847), protein YIF1B (LOC551459; GB48886), and hypoxia up-regulated protein 1 (LOC551763; GB54222). None of these showed significant parent-specific transcription or m6A biases. Models of the genes exhibiting splicing and isoform switching are available in Supplementary Dataset 2 (Figures S21–S44).

#### Discussion

In eusocial insects like *A. mellifera*, the alleles of genes underlying aggressive behaviors that enhance reproductive fitness are predicted to be in conflict, with paternal alleles favoring aggression, and maternal alleles favoring cooperation. Here, we tested the hypothesis that an aggressive behavior (i.e., stinging) in honey bee workers

should exhibit parent-of-origin effects and should be associated with enriched paternal allele-biased transcription. Additionally, we examined whether certain gene regulatory mechanisms play a role in intragenomic conflict. Indeed, we found that both paternal and maternal alleles showed significant increases in transcription for different genes in aggressive versus non-aggressive individuals. However, there was no evidence for parent-of-origin effects on RNA m6A or alternative splicing. Moreover, there was no association between genes that showed parent-specific transcription, m6A, or alternative splicing (see Supplementary Dataset 1).

In four reciprocally crossed colonies of “Africanized” and European honey bees (AHBs and EHBs, respectively), we tested the hypothesis that aggression should be increased in workers with AHB paternity relative to workers with EHB paternity. While we observed significant parent-of-origin effects on this behavior in three of the four colonies, the results were inconsistent between them, with only colonies 2 and 3 showing significantly more aggression in workers with AHB paternity, and colony 1 not showing any stinging behavior in workers with AHB paternity. While our study contrasts with previous reports of overall higher aggression in workers with AHB paternity [14, 27–29], those studies also reported behavioral variation between colonies. Additionally, we used a mtDNA haplotype analysis [64] to identify F0 AHB queens collected from feral colonies and confirm the maternal lineage of our managed F0 EHB queens. While we can be certain that the mtDNA haplotypes of these F0 queens were indeed AHB and EHB, respectively, it is

possible that paternal alleles from the reciprocal lineage could have been present elsewhere in their genomes. Moreover, Rittschof et al. [65] demonstrated that honey bee workers which experienced higher levels of aggression within their colonies during pre-adult stages of development showed increased aggression in adulthood relative to siblings which experienced lower levels of aggression. We did not assess aggression in workers of the parental colonies and so did not control for this in our study. Therefore, our study was limited in that we were unable to disentangle the extent to which the higher level of aggression observed in workers with AHB paternity was due to inherited factors or environmental differences experienced during pre-adult stages of these bees.

Using a colony that exhibited significant differences in aggression associated with paternity, we next explored whether there were parent-specific differences in transcription associated with increased aggression. In contrast with our hypothesis that aggression should be associated with an increase in paternal allele-biased transcription, we found a significant increase in both paternal and maternal allele-biased transcription in aggressive bees relative to non-aggressive bees, although there were more paternal allele-biased than maternal allele-biased transcripts in aggressive bees. This concomitant increase in maternal allele-biased transcription could be explained by between-locus conflicts, whereby some genes have been selected to be preferentially expressed from the maternal allele to counteract the effects of genes that have been selected to be preferentially expressed from the paternal allele. This is supported by the observation in plants and mammals that imprinted genes (those that exhibit parent-specific transcription driven by parent-specific epigenetic marks) are frequently co-expressed [66]. In this context, it is hypothesized that within-locus conflicts, which are resolved by *cis*-regulatory mechanisms like DNA methylation and repressive histone marks [31], select for secondary conflicts between loci [67]. Consequently, extensions to *trans*-regulatory interactions between imprinted genes have been shown to follow naturally from the kinship theory of intragenomic conflict in both theoretical and empirical studies of imprinted gene networks [66]. However, we did not detect an enrichment for any similar gene ontology terms or KEGG pathways among the genes showing parent-specific transcription in our study.

Transgenerational epigenetic inheritance of RNA m6A has been described in brine shrimp (*Artemia franciscana*) [68] and has been hypothesized to modulate offspring metabolism through paternal environmental exposures which affect spermatogenesis [69]. Therefore, we hypothesized that m6A may underlie parent-specific transcription patterns in offspring. To our knowledge, our study is the first analysis to examine the relationship

between parent-specific RNA m6A and parent-specific transcription in a behavioral context. We tested the hypothesis that parent-specific m6A underlies parent-specific transcription, with the unmethylated allele being expressed, and the methylated allele being silenced. While we identified many transcripts that exhibited parent- and lineage-biased m6A in aggressive bees, few also showed allele-biased transcription, none which matched our expectation of higher transcript abundance of the unmethylated allele. We also examined whether parent-specific transcription is associated with alternative splicing. When examining the transcript profiles of aggressive and non-aggressive bees, we identified a handful of genes that showed differential splicing patterns. Some of the differentially spliced genes had functions that could play a role in behavioral variation, including neuronal function and stress response. There was, however, no overlap between genes that showed differential splicing and genes that showed parent-specific transcription.

We did not identify any significant DEGs between aggressive and non-aggressive bees in our study. This was unexpected, as several previous studies had identified many DEGs associated with aggression in honey bees [25, 70–74], but not surprising given that our sequencing libraries were constructed from RNA pooled from multiple transcriptionally distinct tissues [75] and had lower read depth than what is typically used for statistical tests of differential expression. Additionally, genes that showed alternative splicing or isoform switching tended to be highly expressed, which suggests that our ability to identify genes with splice variants was related to our read depth. The recommended input for direct RNA sequencing reactions is 500 ng of mRNA (Direct RNA Sequencing Kit SQK-RNA002, Oxford Nanopore Technologies, Oxford, UK), and we were only able to isolate approximately 15 µg total RNA per individual bee from which approximately 150 ng mRNA was captured. Thus, the relatively low read depth in our study was due to low input to the library preparations. Moreover, our method assessed colony-level aggression, which may not clearly reflect the aggressiveness state of individual bees. We attempted to control for behavioral differences attributable to age (all bees in our colonies were age-matched) or differences in exposure to the stimuli (we disrupted each colony with noise, physical agitation, and the introduction of alarm pheromone), but individual behavior states may not be adequately described by one piece of qualitative information (stung / did not sting) recorded during one ten-minute observation period. Future studies could instead record the unstimulated response of individual bees towards an intruder in an arena assay at multiple timepoints, as described in [71], which allows for measuring quantitative information (latency for attack).

The molecular mechanisms underlying parent-specific transcription in insects remain unidentified [36], and some understanding of these mechanisms is necessary to direct investigations into the nature of their establishment and inheritance. In plants and mammals, imprinted genes, which exhibit parent-specific transcription driven by parent-specific epigenetic marks [5], are canonically regulated by DNA methylation of promoter or enhancer regions [31] that is, at least in mammals, established in primordial germ cells and remains present after fertilization [76]. In honey bees, queens mate with multiple drones [77], and evidence suggests that DNA methylation profiles of worker bees are more similar within than between patrines [78]. However, DNA methylation is not associated with parent-specific transcription in honey bees [33]. Recently, maternally inherited histone 3 (H3) methylation was demonstrated to drive the expression of “non-canonically” imprinted genes, which do not exhibit parent-specific differences in DNA methylation [31, 79, 80]. Variation in H3 methylation has been associated with transcriptional variation and caste differentiation during honey bee larval development [81], and evidence suggests that the honey bee genome is partitioned into regulatory domains with differential enrichment for H3 methylation with respect to worker ovarian plasticity [82]. Thus, future studies should examine the role of chromatin and histone post-translational modifications in mediating intragenomic conflict in honey bees.

The lack of correlation between genes that show parent-specific transcript abundance and significant differences in overall transcript abundance between behavioral phenotypes in honey bees [21, 22] is puzzling. Future studies should examine the potential regulatory linkages between these genes to determine if subtle changes in transcript abundance caused by parent-specific transcription can cascade to larger changes in transcript abundance of genes found in the same network. For example, Galbraith et al. (2021) found that genes showing differences in parent-specific transcript abundance in worker bee brains were overrepresented in gene co-expression network modules associated with metabolism, nutrition, certain cell signaling pathways (i.e., FOXO, MAPK, and HIPPO), and spliceosomes [22]. Targeted studies of the functional consequences of parent-specific transcription of individual genes in these networks would be highly informative. It is also possible that these observed biases in parent-specific transcript abundance are isolated to specific cell types [83], and our study was limited in that we sequenced material from pools of two transcriptionally distinct tissues (heads and abdomens) [75]. Therefore, future studies should utilize single tissues and/or single cell sequencing technologies for finer resolution [84, 85]. Finally, we sequenced tissue from a total of only 12 individuals (3 aggressive and 3 non-aggressive

from each of the reciprocal cross pairs), which limited our power to detect transcriptional and regulatory variation. Therefore, future studies should be performed with a higher sample size, if possible.

## Conclusions

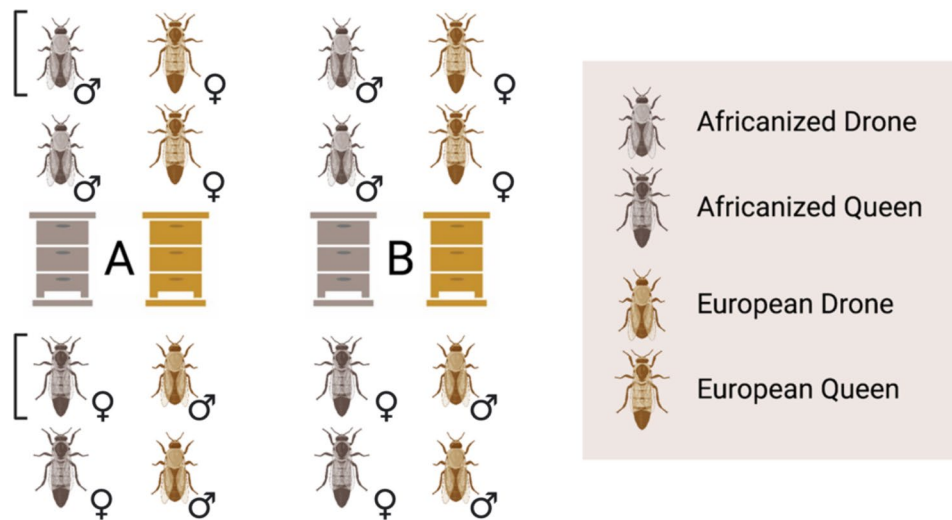
This study provides evidence for intragenomic conflict in worker honey bee aggression. When coupled with previous studies that demonstrated how intragenomic conflict shapes worker reproduction [17, 19, 21, 22], our results suggest that intragenomic conflict and parent-specific transcription may be important factors in shaping the individual variation seen within the context of social behavior in honey bees, and potentially other species [24, 32]. We examined whether RNA m6A and/or alternative splicing play a role in intragenomic conflict in honey bees and did not find evidence to suggest that either are associated with parent-specific transcription in this species. Thus, the molecular mechanisms underlying parent-specific transcription in honey bees remain unidentified [33], and insights from recent studies of genomic imprinting in plants and mammals [31, 79–82] provide direction for future investigations.

## Methods

### Biological samples

Reciprocally crossed colonies were generated in July 2019. We used two AHB colonies and two EHB colonies of *A. mellifera ligustica* stock managed at the Janice and John G. Thomas Honey Bee Facility of Texas A&M University (TAMU) in Bryan, TX. The genetic background of each colony was confirmed using a mitochondrial haplotype analysis [64]. The source colonies were separated into two blocks (“A” and “B”), with one AHB and one EHB colony assigned to each block. From each colony, queens were generated using a standard commercial practice known as “grafting” [86]. Labeled adult virgin (unmated) queens were housed in cages within a nursery colony until drones reared from those same colonies reached sexual maturity. Two queens and two drones from AHB colonies were selected from each colony and crossed with two queens and two drones from EHB colonies by instrumental insemination [87] to create two reciprocal cross pairs per block (Fig. 3). Inseminated queens were placed in their own colonies to initiate egg-laying, generating a total of four colonies with an AHB queen inseminated by a EHB drone, and four colonies with an EHB queen inseminated by an AHB drone. The entrances of these colonies were restricted with queen excluder material to prevent escape or further mating.

Approximately three weeks after the queens began laying fertilized eggs, we collected frames of sealed, emerging brood from each colony and stored them in separate containers within a standing incubator kept at constant



**Fig. 3** Breeding scheme used to generate samples for experiments. In Block A, queens and drones were derived from one Africanized (AHB) honey bee colony and one European (EHB) colony. One queen and drone from each colony was mated to create a reciprocal cross pair (an EHB queen was mated with an AHB drone, and an AHB queen was mated with an EHB drone). The worker offspring of the two crosses in a pair were placed in a common colony and used for behavioral assays and subsequently collected for molecular analysis. In Block A, workers from reciprocal cross 1 was placed in Colony 1, and reciprocal cross 2 was placed in Colony 2. A similar mating scheme was used in Block B to generate Colony 3 and Colony 4. Molecular analysis was performed on the workers generated from reciprocal cross 2 (marked with a bracket)

temperature of 35°C and 75% relative humidity. The following day, approximately five hundred age-matched newly emerged workers from each cross were collected, marked on the thorax with non-toxic paint to identify their maternal lineage of origin, and combined with an equal number of bees from the reciprocal cross colony into “nucleus” hive boxes containing food but no queen or brood. In total, we created four small colonies, each containing 500 workers from each cross in a reciprocal cross design. We also added 1,000 unmarked bees from an unrelated colony to each hive to increase the number of workers in the nucleus colony. Workers were kept in a “queenright” state for the first eight days by providing a TempQueen strip (Mann Lake, Wilkes-Barre, PA) to the colony, as done previously [88]. Once the collection of workers was completed, the queen from each colony was collected and stored at –80°C for DNA extraction and sequencing. Likewise, the bodies of the drones used for inseminating each queen were placed on dry ice and stored at –80°C immediately after semen collection for subsequent analysis.

#### Parent-of-origin effects on worker aggression

We performed an aggression assay following methods described in [89] on day 8 post-emergence, a point in adult worker bee life wherein individuals that are highly receptive to the alarm pheromone cue (soldier bees) have been demonstrated to undergo rapid induction of aggression [70]. The outer cover of each colony was repeatedly hit with a brick for 30 s to disturb the bees, and we continuously waved a swatch of black leather (approximately

10 cm x 10 cm) dosed with honey bee alarm pheromone (100 µL isopentyl-acetate; Sigma-Aldrich, St. Louis, MO) diluted at a 1:10 ratio with mineral oil (Sigma-Aldrich) in front of the colony entrance for 10 min. Colonies were set up approximately 20 m apart to prevent the spread of alarm pheromone to adjacent colonies. Worker bees that responded to the stimuli by stinging the leather swatch were collected as “aggressive” bees. Following the treatment, an equivalent number of labeled bees were collected from inside the colony as “non-aggressive” bees; these bees were likely nurse bees performing brood care duties but may have also included bees involved in food storing or comb-building. Forager bees were likely not inside the colony during the sampling period. Bees were collected immediately on dry ice, shipped to Central State University (CSU), and stored at –80°C. The proportion of bees from each cross that responded to the stimuli were calculated, and logistic regression was conducted to test for an effect of parent lineage, treating colony and block as cofactors.

#### Sample preparation for whole genome sequencing and Oxford nanopore direct RNA-seq

The queen and drone parents from each cross of one reciprocal cross pair in Block A (signified by the brackets in Fig. 3) were selected for whole genome sequencing (WGS). The thorax of each bee was removed, and the flight muscle tissue was dissected on a platform surrounded by dry ice to keep the tissue frozen until the DNA isolation procedure. Genomic DNA was isolated using a DNeasy Blood & Tissue Kit (Qiagen,



Germantown, MD) and treated with 1  $\mu$ L RNase A (NEB, Ipswich, MA). Samples were submitted to the Penn State Genomics Core for quantitation with a Qubit Fluorometer (Life Technologies, Carlsbad, CA) and quality control assessment by Genomic DNA ScreenTape analysis on a TapeStation 4150 (Agilent, Santa Clara, CA). WGS libraries were prepared using an Illumina DNA prep kit with unique 5' and 3' indexes, pooled, and sequenced on a NextSeq 2000 P3 flow cell to generate an average of 45 million 150 bp paired-end reads per sample (Supplementary Dataset 1, Table S1). WGS reads for this study are deposited in the NCBI Sequence Read Archive under SRA project accession PRJNA732718.

Three aggressive bees and three non-aggressive bees were randomly selected from each of the same crosses as above for sequencing. Heads and abdomens were dissected from each bee, and total RNA was isolated using an RNA Miniprep extraction kit (Zymo Research, Irvine, CA). Poly(A) RNA was isolated from approximately 15  $\mu$ g total RNA using the NEBNext Poly(A) mRNA Magnetic Isolation Module (E7490, NEB, Ipswich, MA, USA) according to the manufacturer's protocol with one adjustment: 60  $\mu$ L magnetic oligo d(T)25 beads were used for the initial binding to accommodate the increased total RNA input. Immediately following isolation, library preparation from poly(A) RNA was performed using the Direct RNA Sequencing Kit (SQK-RNA002, Oxford Nanopore Technologies, Oxford, UK) and the manufacturer's protocol (version DRS\_9080\_v2\_revM\_14Aug2019). One Nanopore flow cell (R9, Oxford Nanopore Technologies, Oxford, UK) was used for each sample. Flow cell check was performed to confirm 800 active pores immediately prior to using each flow cell, following the manufacturer's protocol (Flow Cell Check protocol version PQE\_1004\_v1\_revAC\_06Jan2016). Flow cell priming was performed according to the manufacturer's protocol for direct RNA sequencing. Data were acquired in the minKNOW UI version 4.1.22 using the default settings. Flow cells were run to depletion of the RNA input for approximately 25 h. Nanopore direct-seq reads for this study have been deposited the NCBI Sequence Read Archive in FAST5 format under SRA project accession PRJNA940795.

#### Generation of parental transcriptomes

WGS reads were adapter trimmed with fastp [90], then aligned to the RefSeq *Apis mellifera* HAv3.1 genome assembly [91] with BWA-MEM [92] using the default settings. Variants were detected using freebayes [93] to account for differences in ploidy between queens and drones. Heterozygous variants, those with 0/0 genotype (indicating a quality below 1), and variants with QUAL < 30, were filtered using a custom R script utilizing the VariantAnnotation (v1.18.5) package [94], which

is provided in the GitHub repository associated with this manuscript [95]. FAST5 files from all direct RNA-seq samples were processed with Guppy (v5.0.1.6) using the rna\_r9.4.1\_70bps\_hac.cfg configuration (Oxford Nanopore Technologies) for base calling on a GPU. FASTQ files labeled "pass" were concatenated together to generate one FASTQ file per sample. Minimap (v2.21) [96] was used for alignment to the HAv3.1 genome assembly [91] in splice-aware mode with a kmer length of 14 using the RefSeq gene annotation (NCBI *Apis mellifera* Annotation Release 104) to guide alignment. SAMtools (v1.12) [97] was used to sort and index the resulting BAM files. Flair (v1.5) [98] was then used to convert BAM files to BED format, correct intron-exon junctions based on the gene annotation, and call isoforms across all samples. These isoforms included novel transcripts and genes in addition to published transcripts and genes. The filtered variants, along with the transcriptome output by Flair, were then used to generate parent-specific transcriptomes using a custom R script [95]. For each cross, the two parental transcriptomes were combined, with identical transcripts being labeled "both" and transcripts differing between parents being labeled with the parent ID. As expected, the majority of mitochondrial transcripts mapped to the maternal alleles in each cross (Supplementary Dataset 2, Figures S12-S13).

#### Detection of allelic transcription and m6A differences

FAST5 files for all direct-seq samples were processed by Guppy (v3.1.5) (Oxford Nanopore Technologies) and the resulting FASTQ files labeled "pass" were concatenated together to generate one FASTQ per sample. Minimap 2 was used for alignment to the cross transcriptomes generated above. SAMtools was used to sort and index the resulting BAM files. We then used EpiNano [99] to calculate the read coverage and m6A probability at each site, combined with custom R scripts [95] to subset the EpiNano output to the A sites at the center of RRACH motifs. Read coverage and m6A probability at each position within published and novel transcripts, separated by allele, were then combined to generate read count and m6A probability matrices in R. Transcripts with  $n < 2$  positions were filtered from the datasets prior to analyses of allele-biased transcription and m6A.

We utilized the Storer Kim binomial exact test of two proportions [61, 100] and a generalized linear mixed effects model with interaction terms (GLIMMIX) to identify genes that exhibited parent-of-origin allele-biased transcription following methods described in [17, 21, 22], and the two-tailed unpaired z-test to identify genes that exhibited parent-of-origin allele-biased m6A following methods described in [63]. A Storer-Kim test was conducted for each transcript for each SNP position to test for significant differences in the proportion

of maternal and paternal read counts among the samples. Test results were then corrected to control for the FDR [101] and aggregated by transcript. For each transcript, all positions were required to exhibit the same direction of parent- or lineage-specific bias at previously established thresholds for  $p1$  (the proportion of reads mapping to the AHB allele in offspring from the EHB queen and AHB drone cross) and  $p2$  (the proportion of reads mapping to the AHB allele in offspring from the AHB queen and EHB drone cross) for the transcript to be reported as exhibiting bias. This design allows for differentiating between parent-of-origin bias (for example, a transcript exhibiting paternal bias has >60% of reads aligning to the paternal allele in both crosses) and lineage-of-origin bias (for example, a transcript exhibiting AHB bias has >60% of reads aligning to the paternal allele in individuals from the EHB mother x AHB father cross and >60% of reads aligning to the maternal allele in individuals from the AHB mother x EHB father cross – i.e., > 60% of reads align to the nucleotide variant present in the AHB-lineage parent in both crosses) [61]. Additionally, a GLIMMIX [17, 21, 22] was fit for each transcript, separately, to test for an effect of parent, lineage, and their interaction. Test results were FDR corrected, and only transcripts with a significant effect of parent or lineage, but not their interaction, were considered to exhibit bias. Only genes considered to exhibit bias in both tests were reported as exhibiting bias. A two-tailed unpaired  $z$ -test of two proportions was conducted for each transcript for each SNP position to test for differences between maternal and paternal m6A probability among the samples [63].  $Z$ -test results were then FDR corrected and aggregated by transcript. For each transcript, all positions were required to exhibit the same direction of parent- or lineage-specific bias as described above. Our reciprocal cross design and statistical framework allow for differentiating between random monoallelic expression (no consistent bias toward the paternal or maternal allele across samples [102, 103]), lineage-dependent expression (consistent bias toward either the EHB or AHB allele across samples) and parent-of-origin expression (consistent bias toward either the maternal or paternal allele across samples) [61].

#### Detection of lincRNAs

Transcripts generated by Flair were defined as “intergenic” if they had no overlap with any annotated protein-coding, rRNA, miRNA, snRNA, snoRNA, tRNA, or guide RNA gene. Transcripts were retained for further analysis if they were non-mitochondrial, intergenic, non-viral, and were at least 200 nucleotides in length [104]. The findORFs function in the ORFhunterR [105] R package was used to identify putative open reading frames (ORFs) and their lengths. CPC2 [106] was used

to identify transcripts with coding potential. Additionally, BLASTX [107] was used to align transcripts to the Uniprot and Swiss-Prot combined database [108] under default parameters and a required cutoff E-value of 0.001. The transcripts were searched against annotated rRNAs, snRNAs, snoRNAs, miRNAs, tRNAs, and gRNAs, in the NCBI RefSeq Amel HAv3.1 assembly using discontinuous megablast (from BLAST+v2.10.1). Transcripts were considered long intergenic non-coding RNAs (lincRNAs) if they met all the above criteria, did not appear to be protein-coding by any of the three above methods, and did not align to other types of non-coding RNA using discontinuous megablast.

#### Detection of isoform switching

Flair was used to quantify the abundance of each transcript in each sample. Counts were then imported into IsoformSwitchAnalyzeR (v1.18.0) [109] in R to test for differential isoform transcription between aggressive and non-aggressive bees and between the crosses with DEXseq [110]. Transcripts detected to exhibit isoform switching at  $FDR < 0.1$  were retained for further analysis. ORFs were then predicted *de novo* from the transcript sequences using IsoformSwitchAnalyzeR. Additionally, the transcript sequences were analyzed by CPC2 to predict coding potential, overriding ORFs called by IsoformSwitchAnalyzeR if the transcript was determined to be non-coding at a cutoff of 0.7. Hmmscan [111] was then used to compare predicted ORFs to the Pfam database [112] to predict protein domains. SignalP (v5.0) [113] was used to predict signal peptides within predicted ORFs. NetSurfP2 [114] was used to predict intrinsically disordered regions within predicted ORFs. Finally, alternative splicing was then analyzed within IsoformSwitchAnalyzeR, and isoform switches were determined to have “consequences” if there were changes in coding potential, domains, signal peptides, intrinsically disordered regions, or intron retention.

#### Differential expression analyses

Flair was used for each cross to quantify transcript abundances for each sample for each gene. Counts were then summed within genes regardless of parent ID or transcript ID, and samples from both crosses were combined into one gene count matrix in R. Counts were then TMM normalized using the edgeR package [115], and genes with  $n < 8$  counts per million in at least  $n = 3$  samples were filtered. After filtering, TMM normalization was performed again on the raw counts to calculate normalized  $\log_2$ -based count per million values (logCPM) using the edgeR package with a prior count of 2 to stabilize fold-changes of lowly transcribed genes. Multidimensional scaling using the limma package [116] was used to assess the largest effects on gene composition among the

samples, using the top 5,000 most variable genes between each pair of samples. Differential expression analysis was then performed using the limma-trend method to test for differences in transcript abundance between aggressive and non-aggressive bees and between the crosses. Test results were FDR corrected, and only genes with  $FDR < 0.05$  were considered to exhibit differential expression.

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-023-09411-4>.

Supplementary Material 1

Supplementary Material 2

### Acknowledgements

We thank E. T. Ash and Elizabeth (Liz) Walsh for helping to generate the original colonies used in this study, Tonya Shepherd for helping with sample collections, David Galbraith for feedback on our statistical methods, and members of the Li-Byarlay, Grozinger, and Rangel labs for feedback on our manuscript. Additionally, we thank two anonymous reviewers for their thoughtful feedback which we used to improve our manuscript.

### Authors' contributions

S.T.B., C.M.G., R.M., J.R., and H.L.-B. designed the study. S.T.B. and R.M. collected samples. S.T.B. performed DNA extractions, and E.L. performed RNA extractions. S.T.B. and L.C. conducted bioinformatic and statistical analyses. S.T.B. prepared the initial drafts of the manuscript, and all authors contributed to the writing and editing of the final manuscript.

### Funding

This work was supported by the National Science Foundation (Grant # RIA-1900793 to H.L.-B., and MCB-0950896 to C.M.G.) and the National Science Foundation Graduate Research Fellowship Program (Grant # DGE1255832 to S.T.B.).

### Data Availability

All analysis scripts for this study and detailed markdowns describing their use are available at <https://github.com/sbresnahan/allele-specific-transcription-and-m6A> [95]. WGS reads for this study are deposited in the NCBI Sequence Read Archive under SRA accession PRJNA732718, and Nanopore direct-seq reads are available in FAST5 format under SRA accession PRJNA940795.

### Declarations

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

Received: 23 February 2023 / Accepted: 27 May 2023

Published online: 12 June 2023

### References

- Haig D. The Kinship Theory of genomic imprinting. *Annu Rev Ecol Syst.* 2000;31(1):9–32.
- Haig D. Intragenomic conflict and the evolution of eusociality. *J Theor Biol.* 1992;156(3):401–3.
- Gardner A, Úbeda F. The meaning of intragenomic conflict. *Nat Ecol Evol.* 2017;1(12):1807–15.
- Reik W, Walter J. Genomic imprinting: parental influence on the genome. *Nat Rev Genet.* 2001;2(1):21–32.
- Ferguson-Smith AC. Genomic imprinting: the emergence of an epigenetic paradigm. *Nat Rev Genet.* 2011;12(8):565–75.
- Queller DC. Theory of genomic imprinting conflict in social insects. *BMC Evol Biol.* 2003;3(1):15.
- Sakagami ShF. Occurrence of an aggressive behaviour in queenless hives, with considerations on the social organization of honeybee. *Insectes Sociaux.* 1954;1(4):331–43.
- Galbraith DA, Wang Y, Amdam GV, Page RE, Grozinger CM. Reproductive physiology mediates honey bee (*Apis mellifera*) worker responses to social cues. *Behav Ecol Sociobiol.* 2015;69(9):1511–8.
- Malika O, Shnieor S, Katzav-Gozansky T, Hefetz A. Aggressive reproductive competition among hopelessly queenless honeybee workers triggered by pheromone signaling. *Naturwissenschaften.* 2008;95(6):553–9.
- Backx AG, Guzmán-Novoa E, Thompson GJ. Factors affecting ovary activation in honey bee workers: a meta-analysis. *Insectes Sociaux.* 2012;59(3):381–8.
- Hoover SER, Keeling CI, Winston ML, Slessor KN. The effect of queen pheromones on worker honey bee ovary development. *Naturwissenschaften.* 2003;90(10):477–80.
- Tarpy DR, Fletcher DJC. Spraying<sup>®</sup> Behavior during Queen Competition in Honey Bees. *J Insect Behav.* 2003;16:425–37.
- Pflugfelder J, Koeniger N. Fight between virgin queens (*Apis mellifera*) is initiated by contact to the dorsal abdominal surface. *Apidologie.* 2003;34(3):249–56.
- Guzmán-Novoa E, Hunt GJ, Page RE, Uribe-Rubio JL, Prieto-Merlos D, Becerra-Guzmán F. Paternal Effects on the defensive behavior of honeybees. *J Hered.* 2005;96(4):376–80.
- Gilley D. The behavior of Honey Bees (*Apis mellifera ligustica*) during queen duels. *Ethology.* 2001;107(7):601–22.
- Oldroyd BP, Allsopp MH, Roth KM, Remnant EJ, Drewell RA, Beekman M. A parent-of-origin effect on honeybee worker ovary size. *Proc R Soc B Biol Sci.* 2014;281(1775):20132388.
- Kocher SD, Tsuruda JM, Gibson JD, Emore CM, Arechavala-Velasco ME, Queller DC et al. A search for parent-of-Origin Effects on Honey Bee Gene expression. *G3 Bethesda Md.* 2015;5(8):1657–62.
- Reid RJ, Remnant EJ, Allsopp MH, Beekman M, Oldroyd BP. Paternal effects on *Apis mellifera capensis* worker ovary size. *Apidologie.* 2017;48(5):660–5.
- Smith NMA, Yagound B, Remnant EJ, Foster CSP, Buchmann G, Allsopp MH, et al. Paternally-biased gene expression follows kin-selected predictions in female honey bee embryos. *Mol Ecol.* 2020;29(8):1523–33.
- Linksvayer TA, Rueppell O, Siegel A, Kaftanoglu O, Page RE, Amdam GV. The genetic basis of transgressive ovary size in Honeybee Workers. *Genetics.* 2009;183(2):693–707.
- Galbraith DA, Kocher SD, Glenn T, Albert I, Hunt GJ, Strassmann JE, et al. Testing the kinship theory of intragenomic conflict in honey bees (*Apis mellifera*). *Proc Natl Acad Sci.* 2016;113(4):1020–5.
- Galbraith DA, Ma R, Grozinger CM. Tissue-specific transcription patterns support the kinship theory of intragenomic conflict in honey bees (*Apis mellifera*). *Mol Ecol.* 2021;30(4):1029–41.
- Gibson JD, Arechavala-Velasco ME, Tsuruda JM, Hunt GJ. Biased allele expression and aggression in hybrid honeybees may be influenced by inappropriate nuclear-cytoplasmic signaling. *Front Genet.* 2015;6:343.
- Pegoraro M, Marshall H, Lonsdale ZN, Mallon EB. Do social insects support Haig's kin theory for the evolution of genomic imprinting? *Epigenetics.* 2017;12(9):725–42.
- Alaux C, Sinha S, Hasadsri L, Hunt GJ, Guzmán-Novoa E, DeGrandi-Hoffman G, et al. Honey bee aggression supports a link between gene regulation and behavioral evolution. *Proc Natl Acad Sci.* 2009;106(36):15400–5.
- Harrison JF, Hall HG. African-european honeybee hybrids have low nonintermediate metabolic capacities. *Nature.* 1993;363(6426):258–60.
- Gonçalves LS, Stort AC. Genetics of defensive behavior II. The "African" honey bee. Boulder, CO: Westview Press; 1991. 329–56.
- Guzmán-Novoa E, Page RE. Backcrossing Africanized Honey Bee Queens to European Drones reduces colony defensive behavior. *Ann Entomol Soc Am.* 1993;86(3):352–5.
- Schneider SS, DeGrandi-Hoffman G. Queen replacement in african and european honey bee colonies with and without afterwarms. *Insectes Sociaux.* 2008;55(1):79–85.

30. Batista RA, Köhler C. Genomic imprinting in plants—revisiting existing models. *Genes Dev.* 2020;34(1–2):24–36.
31. Hanna CW, Kelsey G. Features and mechanisms of canonical and noncanonical genomic imprinting. *Genes Dev.* 2021;35(11–12):821–34.
32. Marshall H, Jones ARC, Lonsdale ZN, Mallon EB. Bumblebee Workers Show Differences in Allele-Specific DNA Methylation and Allele-Specific Expression. O'Neill R, editor. *Genome Biol Evol.* 2020;12(8):1471–81.
33. Wu X, Galbraith DA, Chatterjee P, Jeong H, Grozinger CM, Yi SV. Lineage and parent-of-origin Effects in DNA methylation of Honey Bees (*Apis mellifera*) revealed by reciprocal crosses and whole-genome bisulfite sequencing. *Genome Biol Evol.* 2020;12(8):1482–92.
34. Duncan EJ, Cunningham CB, Dearden PK. Phenotypic plasticity: what has DNA methylation got to do with it? *Insects.* 2022;13(2):110.
35. Glastad KM, Hunt BG, Goodisman MAD. Epigenetics in insects: genome regulation and the generation of phenotypic diversity. *Annu Rev Entomol.* 2019;64(1):185–203.
36. Oldroyd BP, Yagound B. Parent-of-origin effects, allele-specific expression, genomic imprinting and paternal manipulation in social insects. *Philos Trans R Soc B Biol Sci.* 2021;376:20200425.
37. Dai Z, Ramesh V, Locasale JW. The evolving metabolic landscape of chromatin biology and epigenetics. *Nat Rev Genet.* 2020;21(12):737–53.
38. Wang S, Lv W, Li T, Zhang S, Wang H, Li X, et al. Dynamic regulation and functions of mRNA m6A modification. *Cancer Cell Int.* 2022;22(1):48.
39. Wei JW, Huang K, Yang C, Kang CS. Non-coding RNAs as regulators in epigenetics. *Oncol Rep.* 2017;37(1):3–9.
40. Barbieri I, Kouzarides T. Role of RNA modifications in cancer. *Nat Rev Cancer.* 2020;20(6):303–22.
41. Chen J, Zhang YC, Huang C, Shen H, Sun B, Cheng X, et al. m6A regulates neurogenesis and neuronal development by modulating histone methyltransferase Ezh2. *Genomics Proteom Bioinf.* 2019;17(2):154–68.
42. Merkurjev D, Hong WT, Iida K, Oomoto I, Goldie BJ, Yamaguti H, et al. Synaptic N6-methyladenosine (m6A) epitranscriptome reveals functional partitioning of localized transcripts. *Nat Neurosci.* 2018;21(7):1004–14.
43. Nainar S, Marshall PR, Tyler CR, Spitale RC, Bredy TW. Evolving insights into RNA modifications and their functional diversity in the brain. *Nat Neurosci.* 2016;19(10):1292–8.
44. Satterlee JS, Basanta-Sanchez M, Blanco S, Li JB, Meyer K, Pollock J, et al. Novel RNA modifications in the nervous system: form and function. *J Neurosci.* 2014;34(46):15170–7.
45. Xu Z, Shi X, Bao M, Song X, Zhang Y, Wang H, et al. Transcriptome-wide analysis of RNA m6A methylation and gene expression changes among two Arabidopsis Ecotypes and their reciprocal hybrids. *Front Plant Sci.* 2021;12:685189.
46. Wang M, Xiao Y, Li Y, Wang X, Qi S, Wang Y, et al. RNA m6A modification functions in Larval Development and Caste differentiation in Honeybee (*Apis mellifera*). *Cell Rep.* 2021;34(1):108580.
47. Bataglia L, Simões Z, Nunes F. Active genic machinery for epigenetic RNA modifications in bees. *Insect Mol Biol.* 2021;30(6):566–79.
48. Li-Byarlay H, Li Y, Stroud H, Feng S, Newman TC, Kaneda M, et al. RNA interference knockdown of *DNA methyl-transferase 3* affects gene alternative splicing in the honey bee. *Proc Natl Acad Sci.* 2013;110(31):12750–5.
49. Foret S, Kucharski R, Pellegrini M, Feng S, Jacobsen SE, Robinson GE, et al. DNA methylation dynamics, metabolic fluxes, gene splicing, and alternative phenotypes in honey bees. *Proc Natl Acad Sci.* 2012;109(13):4968–73.
50. Jarosch A, Stolle E, Crewe RM, Moritz RFA. Alternative splicing of a single transcription factor drives selfish reproductive behavior in honeybee workers (*Apis mellifera*). *Proc Natl Acad Sci.* 2011;108(37):15282–7.
51. Jayakodi M, Jung JW, Park D, Ahn YJ, Lee SC, Shin SY, et al. Genome-wide characterization of long intergenic non-coding RNAs (lincRNAs) provides new insight into viral diseases in honey bees *Apis cerana* and *Apis mellifera*. *BMC Genomics.* 2015;16(1):680.
52. Choudhary C, Sharma S, Meghwanshi KK, Patel S, Mehta P, Shukla N, et al. Long Non-Coding RNAs in Insects Animals. 2021;11(4):1118.
53. Wei CW, Luo T, Zou SS, Wu AS. The role of long noncoding RNAs in Central Nervous System and neurodegenerative Diseases. *Front Behav Neurosci.* 2018;12:175.
54. Tang J, Yu Y, Yang W. Long noncoding RNA and its contribution to autism spectrum disorders. *CNS Neurosci Ther.* 2017;23(8):645–56.
55. Rogoyski O, Pueyo J, Couso J, Newbury S. Functions of long non-coding RNAs in human disease and their conservation in *Drosophila* development. *Biochem Soc Trans.* 2017;45(4):895–904.
56. He XJ, Barron AB, Yang L, Chen H, He YZ, Zhang LZ, et al. Extent and complexity of RNA processing in honey bee queen and worker caste development. *iScience.* 2022;25(5):104301.
57. Aamodt Randi M. The caste- and age-specific expression signature of honeybee heat shock genes shows an alternative splicing-dependent regulation of Hsp90. *Mech Ageing Dev.* 2008;129(11):632–7.
58. Mutti NS, Wang Y, Kaftanoglu O, Amdam GV, Honey Bee PTEN –. Description. Developmental Knockdown, and Tissue-Specific Expression of Splice-Variants Correlated with Alternative Social Phenotypes. Moritz RFA, editor. *PLoS ONE.* 2011;6(7):e22195.
59. Huang KK, Huang J, Wu JKL, Lee M, Tay ST, Kumar V, et al. Long-read transcriptome sequencing reveals abundant promoter diversity in distinct molecular subtypes of gastric cancer. *Genome Biol.* 2021;22(1):44.
60. Lagarde J, Uszczyńska-Ratajczak B, Carbonell S, Pérez-Lluch S, Abad A, Davis C, et al. High-throughput annotation of full-length long noncoding RNAs with capture long-read sequencing. *Nat Genet.* 2017;49(12):1731–40.
61. Wang X, Clark AG. Using next-generation RNA sequencing to identify imprinted genes. *Heredity.* 2014;113(2):156–66.
62. Calfee E, Agra MN, Palacio MA, Ramírez SR, Coop G. Selection and hybridization shaped the rapid spread of African honey bee ancestry in the Americas. *PLoS Genet.* 2020;16(10):e1009038.
63. Pratanwanich PN, Yao F, Chen Y, Koh CWQ, Wan YK, Hendra C, et al. Identification of differential RNA modifications from nanopore direct RNA sequencing with xPore. *Nat Biotechnol.* 2021;39(11):1394–402.
64. Meixner MD, Pinto MA, Bouga M, Kryger P, Ivanova E, Fuchs S. Standard methods for characterising subspecies and ecotypes of *Apis mellifera*. *J Apic Res.* 2013;52(4):1–28.
65. Rittschof CC, Coombs CB, Frazier M, Grozinger CM, Robinson GE. Early-life experience affects honey bee aggression and resilience to immune challenge. *Sci Rep.* 2015;5(1):15572.
66. Patten MM, Cowley M, Oakey RJ, Feil R. Regulatory links between imprinted genes: evolutionary predictions and consequences. *Proc R Soc B Biol Sci.* 2016;283(1824):20152760.
67. Haig D. Conflicting messages: imprinting and internal communication. In: d'Ettore P, Hughes DP, editors. *Sociobiology of communication.* Oxford, UK: Oxford University Press; 2008.
68. Roy S, Kumar V, Bossier P, Norouzitallab P, Vanrompoy D. Phloroglucinol Treatment induces transgenerational epigenetic inherited resistance against *Vibrio* infections and thermal stress in a brine shrimp (*Artemia franciscana*) Model. *Front Immunol.* 2019;10:2745.
69. Zhang Y, Shi J, Rassoulzadegan M, Tuorto F, Chen Q. Sperm RNA code programmes the metabolic health of offspring. *Nat Rev Endocrinol.* 2019;15(8):489–98.
70. Alaux C, Robinson GE. Alarm pheromone induces Immediate–Early Gene expression and slow behavioral response in Honey Bees. *J Chem Ecol.* 2007;33(7):1346–50.
71. Shpigler HY, Saul MC, Murdoch EE, Cash-Ahmed AC, Seward CH, Sloofman L, et al. Behavioral, transcriptomic and epigenetic responses to social challenge in honey bees: biological embedding of social experience in honey bees. *Genes Brain Behav.* 2017;16(6):579–91.
72. Shpigler HY, Saul MC, Murdoch EE, Corona F, Cash-Ahmed AC, Seward CH, et al. Honey bee neurogenomic responses to affiliative and agonistic social interactions. *Genes Brain Behav.* 2019;18(1):e12509.
73. Rittschof CC, Bukhari SA, Sloofman LG, Troy JM, Caetano-Anollés D, Cash-Ahmed A, et al. Neuromolecular responses to social challenge: common mechanisms across mouse, stickleback fish, and honey bee. *Proc Natl Acad Sci.* 2014;111(50):17929–34.
74. Rittschof CC, Rubin BER, Palmer JH. The transcriptomic signature of low aggression in honey bees resembles a response to infection. *BMC Genomics.* 2019;20(1):1029.
75. Margotta J, Mancinelli G, Benito A, Ammons A, Roberts S, Elekonich M. Effects of Flight on Gene expression and aging in the Honey Bee Brain and Flight muscle. *Insects.* 2012;4(1):9–30.
76. Zeng Y, Chen T. DNA methylation reprogramming during mammalian development. *Genes.* 2019;10(4):257.
77. Laidlaw HH, Page RE. Polyandry in honey bees (*Apis mellifera*): sperm utilization and intracolony genetic relationships. *Genetics.* 1984;108:985–97.
78. Yagound B, Remnant EJ, Buchmann G, Oldroyd BP. Intergenerational transfer of DNA methylation marks in the honey bee. *Proc Natl Acad Sci.* 2020;117(51):32519–27.

79. Zhang M, Xie S, Dong X, Zhao X, Zeng B, Chen J, et al. Genome-wide high resolution parental-specific DNA and histone methylation maps uncover patterns of imprinting regulation in maize. *Genome Res.* 2014;24(1):167–76.
80. Moreno-Romero J, Jiang H, Santos-González J, Köhler C. Parental epigenetic asymmetry of PRC 2-mediated histone modifications in the *Arabidopsis* endosperm. *EMBO J.* 2016;35(12):1298–311.
81. Wojciechowski M, Lowe R, Maleszka J, Conn D, Maleszka R, Hurd PJ. Phenotypically distinct female castes in honey bees are defined by alternative chromatin states during larval development. *Genome Res.* 2018;28(10):1532–42.
82. Duncan EJ, Leask MP, Dearden PK. Genome Architecture facilitates phenotypic plasticity in the Honeybee (*Apis mellifera*). *Mol Biol Evol.* 2020;37(7):1964–78.
83. Liu NMP, Bousounis H, Spurr P, Alomran L, Ibeawuchi N. Estimating the allele-specific expression of SNVs from 10x Genomics single-cell RNA-Sequencing data. *Genes.* 2020;11(3):240.
84. Traniello IM, Bukhari SA, Kevill J, Ahmed AC, Hamilton AR, Naeger NL, et al. Meta-analysis of honey bee neurogenomic response links deformed wing virus type A to precocious behavioral maturation. *Sci Rep.* 2020;10(1):3101.
85. Zhang W, Wang L, Zhao Y, Wang Y, Chen C, Hu Y, et al. Single-cell transcriptomic analysis of honeybee brains identifies vitellogenin as caste differentiation-related factor. *iScience.* 2022;25(7):104643.
86. Connor L. Queen rearing essentials. 1st ed. Wicwas Press, LLC; 2009.
87. Cobey SW, Tarpay DR, Woyke J. Standard methods for instrumental insemination of *Apis mellifera* queens. *J Apic Res.* 2013;52(4):1–18.
88. Pankiw T, Winston ML, Slessor KN. Variation in worker response to honey bee (*Apis mellifera* L) queen mandibular pheromone (Hymenoptera: Apidae). *J Insect Behav.* 1994;7(1):1–15.
89. Rittschof CC, Robinson GE. Manipulation of colony environment modulates honey bee aggression and brain gene expression: Bee aggression. *Genes Brain Behav.* 2013;12(8):802–11.
90. Chen S, Zhou Y, Chen Y, Gu J. Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics.* 2018;34(17):i884–90.
91. Wallberg A, Bunikis I, Pettersson OV, Mosbech MB, Childers AK, Evans JD, et al. A hybrid de novo genome assembly of the honeybee, *Apis mellifera*, with chromosome-length scaffolds. *BMC Genomics.* 2019;20(1):275.
92. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM [Internet]. arXiv; 2013. Available from: <http://arxiv.org/abs/1303.3997>.
93. Garrison E, Marth G. Haplotype-based variant detection from short-read sequencing [Internet]. arXiv; 2012. Available from: <http://arxiv.org/abs/1207.3907>.
94. Obenchain V, Lawrence M, Carey V, Gogarten S, Shannon P, Morgan M. VariantAnnotation: a Bioconductor package for exploration and annotation of genetic variants. *Bioinformatics.* 2014;30(14):2076–8.
95. Bresnahan S. sbresnahan/allele-specific-transcription-and-m6A: Initial release [Internet]. Zenodo; 2023. Available from: <https://doi.org/10.5281/zenodo.7607957>.
96. Li H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics.* 2018;34(18):3094–100.
97. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. *GigaScience.* 2021;10(2):giab008.
98. Tang AD, Soulette CM, van Baren MJ, Hart K, Hrabeta-Robinson E, Wu CJ, et al. Full-length transcript characterization of SF3B1 mutation in chronic lymphocytic leukemia reveals downregulation of retained introns. *Nat Commun.* 2020;11(1):1438.
99. Liu H, Begik O, Novoa EM, EpiNano. Detection of m6A RNA modifications using Oxford Nanopore Direct RNA sequencing. In: McMahon M, editor. RNA modifications. *Methods in Molecular Biology.* Volume 2298. New York, NY: Springer US; 2021. pp. 31–52.
100. Storer BE, Kim C. Exact Properties of some exact test statistics for comparing two binomial proportions. *J Am Stat Assoc.* 1990;85(409):146–55.
101. Benjamini Y, Hochberg Y. Controlling the false Discovery rate: a practical and powerful Approach to multiple testing. *J R Stat Soc Ser B Methodol.* 1995;57(1):289–300.
102. Gimelbrant A, Hutchinson JN, Thompson BR, Chess A. Widespread monoallelic expression on human autosomes. *Science.* 2007;318(5853):1136–40.
103. Metsalu T, Viltrop T, Tiirats A, Rajashekar B, Reimann E, et al. Using RNA sequencing for identifying gene imprinting and random monoallelic expression in human placenta. *Epigenetics.* 2014;9(10):1397–409.
104. Deniz E, Erman B. Long noncoding RNA (lincRNA), a new paradigm in gene expression control. *Funct Integr Genomics.* 2017;17(2–3):135–43.
105. Grinev VV, Yatskou MM, Skakun VV, Chepeleva MK, Nazarov PV. ORFhunteR: an accurate approach to the automatic identification and annotation of open reading frames in human mRNA molecules. *Softw Impacts.* 2022;12:100268.
106. Kang YJ, Yang DC, Kong L, Hou M, Meng YQ, Wei L, et al. CPC2: a fast and accurate coding potential calculator based on sequence intrinsic features. *Nucleic Acids Res.* 2017;45(W1):W12–6.
107. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. *BMC Bioinformatics.* 2009;10(1):421.
108. Boutet E, Lieberherr D, Tognolli M, Schneider M, Bairoch A. UniProtKB/Swiss-Prot. In: Edwards D, editor. *Plant Bioinformatics.* Totowa, NJ: Humana Press; 2007. pp. 89–112.
109. Vitting-Seerup K, Sandelin A. IsoformSwitchAnalyzerR: analysis of changes in genome-wide patterns of alternative splicing and its functional consequences. *Berger B, editor. Bioinformatics.* 2019;35(21):4469–71.
110. Anders S, Reyes A, Huber W. Detecting differential usage of exons from RNA-seq data. *Genome Res.* 2012;22(10):2008–17.
111. Wheeler TJ, Eddy SR. nhmmer: DNA homology search with profile HMMs. *Bioinformatics.* 2013;29(19):2487–9.
112. Paysan-Lafosse T, Blum M, Chuguransky S, Grego T, Pinto BL, Salazar GA, et al. InterPro in 2022. *Nucleic Acids Res.* 2023;51(D1):D418–27.
113. Teufel F, Almagro Armenteros JJ, Johansen AR, Gislason MH, Pihl SI, Tsirigos KD, et al. SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nat Biotechnol.* 2022;40(7):1023–5.
114. Høie MH, Kiehl EN, Petersen B, Nielsen M, Winther O, Nielsen H, et al. NetSurfP-3.0: accurate and fast prediction of protein structural features by protein language models and deep learning. *Nucleic Acids Res.* 2022;50(W1):W510–5.
115. McCarthy DJ, Chen Y, Smyth GK. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* 2012;40(10):4288–97.
116. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 2015;43(7):e47–7.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.