Jayakodi et al. BMC Genomics (2015) 16:680
DOI 10.1186/s12864-015-1868-7

BMC Genomics

RESEARCH ARTICLE Open Access

# Genome-wide characterization of long intergenic non-coding RNAs (lincRNAs) provides new insight into viral diseases in honey bees *Apis cerana* and *Apis mellifera*

Murukarthick Jayakodi[1], Je Won Jung[2], Doori Park[2], Young-Joon Ahn[2], Sang-Choon Lee[1], Sang-Yoon Shin[3], Chanseok Shin[3], Tae-Jin Yang[1*] and Hyung Wook Kwon[2*]

## Abstract

**Background:** Long non-coding RNAs (lncRNAs) are a class of RNAs that do not encode proteins. Recently, lncRNAs have gained special attention for their roles in various biological process and diseases.

**Results:** In an attempt to identify long intergenic non-coding RNAs (lincRNAs) and their possible involvement in honey bee development and diseases, we analyzed RNA-seq datasets generated from Asian honey bee (*Apis cerana*) and western honey bee (*Apis mellifera*). We identified 2470 lincRNAs with an average length of 1011 bp from *A. cerana* and 1514 lincRNAs with an average length of 790 bp in *A. mellifera*. Comparative analysis revealed that 5 % of the total lincRNAs derived from both species are unique in each species. Our comparative digital gene expression analysis revealed a high degree of tissue-specific expression among the seven major tissues of honey bee, different from mRNA expression patterns. A total of 863 (57 %) and 464 (18 %) lincRNAs showed tissue-dependent expression in *A. mellifera* and *A. cerana*, respectively, most preferentially in ovary and fat body tissues. Importantly, we identified 11 lincRNAs that are specifically regulated upon viral infection in honey bees, and 10 of them appear to play roles during infection with various viruses.

**Conclusions:** This study provides the first comprehensive set of lincRNAs for honey bees and opens the door to discover lincRNAs associated with biological and hormone signaling pathways as well as various diseases of honey bee.

**Keywords:** *Apis cerana*, Asian honey bee, lincRNAs, RNA-seq

## Background

Advances in RNA sequencing technologies have allowed the rapid exploration of protein-coding and non-coding RNAs in both vertebrate and invertebrate genomes. Transcriptome sequencing of diverse species has revealed that much of the genome is transcribed. However, only a small portion of sequences code for proteins [1–5]. In human, only less than 2 % of the genome contains evolutionarily-conserved sequences for proteins [3, 6, 7]. Thus, many of the transcribed sequences in the genome are likely to be non-coding RNAs (ncRNAs). Non-coding RNAs include small RNAs [18–35 nucleotides (nt)], as well as longer RNAs (>200 nt) referred as long non-coding RNAs (lncRNAs), for which processing, splicing, and polyadenylation are similar to those of mRNA [8]. Based on their genomic position, lncRNAs can be classified as natural antisense transcripts, long intronic non-coding RNAs, or long intergenic non-coding RNAs (lincRNAs). Recently, lncRNAs have been found to play important roles in biological processes such as the regulation of gene expression through chromatin or histone modification [9–11], or transcriptional [12, 13] and post-transcriptional [14–16] processing. The expression of

* Correspondence: tjyang@snu.ac.kr; biomodeling@snu.ac.kr
[1]Department of Plant Science, Plant Genomics and Breeding Institute, Research Institute of Agriculture and Life Sciences, College of Agriculture and Life Sciences, Seoul National University, Seoul 151-921, Republic of Korea
[2]WCU Biomodulation Major, Department of Agricultural Biotechnology and Research Institute of Agriculture and Life Sciences, College of Agriculture and Life Sciences, Seoul National University, Seoul 151–921, Republic of Korea
Full list of author information is available at the end of the article

lncRNAs is often specific to a tissue or a particular developmental stage [17–20]. In addition, lncRNAs are associated with diseases [21] including acquired immune deficiency syndrome (AIDS) [22, 23], Alzheimer's disease [24], and cancer [25–27], necessitating their study as potential therapeutic targets. Furthermore, lncRNAs including *Xist* play a critical role in X-chromosome dosage compensation [28], genomic imprinting [29], epigenetic regulation [30], cellular pluripotency, and differentiation [31].

It is increasingly clear that lncRNAs are important regulators of diverse functions, and hence, genome-wide scans for lncRNAs are warranted to improve our understanding of cell regulatory and disease-related mechanisms. In recent years, lncRNAs have been studied via EST *in silico* mining [2, 32], whole-genome tilling array, and RNA-seq [32, 33] methods. Genome-wide lncRNA analyses have been performed in human, *Plasmodium falciparum*, mouse, zebrafish, fruit fly, worm, and yeast [34–38]. Each of the surveys in mammals has uncovered a substantial number of lncRNAs, which are often expressed at low levels in a tissue-dependent manner.

Western honey bee (*Apis mellifera*) is a key model for understanding social behavior, disease transmission, and development [39]. The genome of *A. mellifera* was revealed in 2006 [40], which paved the way for understanding regulation of behavior, immunity, and aging, and for molecular and functional genomics studies. A sister species, Asian honey bee (*Apis cerana*), is a significant pollinator in many Asian countries and its genome information was revealed recently, which enables prediction of genes and examination of evolution and comparative sociogenomics between social insects [41]. These two honey bee species have been used in medical research and for studies of neurobiology, developmental biology, behavior science and epigenomics [42–45]. Only four lncRNAs have been characterized in *A. mellifera* [46, 47], and despite their use in clinical research, no effort has been made to profile lncRNAs at the genome level in honey bees.

In the present study, we first generate a comprehensive set of lincRNAs from RNA-seq datasets in *A. cerana* and *A. mellifera*. Secondly, we identify candidate lincRNAs specifically associated with viral diseases in honey bees. Using our bioinformatics pipeline, we identified a total of 2470 lincRNAs, encoded by 2376 gene loci in the *A. cerana* genome (http://mnbldb.snu.ac.kr/; scaffold_v1) and a total of 1514 lincRNAs in the *A. mellifera* genome (www.beebase.org; Amel_4.5_scaffolds), and profiled tissue-specific lincRNA expression. Finally, we characterized the virus-specific lincRNAs in both honey bee species. Our genome-wide profiling of lincRNAs in these two sister honey bee species identifies exciting candidates for characterization of lincRNAs related to diseases as well as to hormone signaling and metabolism, and thus provides valuable information on the modulation of gene expression.

## Results

### Genome-wide identification of lincRNAs from two sister honey bee species, *A. cerana* and *A. mellifera*

To identify a comprehensive set of Asian honey bee lincRNAs, we used Illumina RNA-seq data generated for the *A. cerana* genome project [41] (for six tissues: antenna, brain, hypoharyngeal gland, gut, fat body, and venom gland) and newly generated datasets from larvae, and Sacbrood virus (SBV)-infected and non-infected honey bees (Table 1). We established a bioinformatics pipeline by modifying protocols used in various previous studies [34, 48, 49] (Fig. 1). A reference-guided assembly yielded a total of 24,529 transcripts from 18,937 gene loci. The assembled sequences were analyzed to identify putative lncRNAs, and 19,916 transcripts were selected based on nucleotide length ≥200 bp and ORF ≤ 100 amino acids (Fig. 1). We chose not to consider the protein-coding transcripts in order to increase the accuracy in identifying lncRNAs. From the filtered transcripts, we removed transcripts with overlapping Swiss-Prot protein sequences (http://www.uniprot.org/). The remaining 9373 transcripts were filtered based on coding potential evaluation, removing those with scores ≤ −1.0 using the Coding Potential Calculator (CPC) program, which is a state-of-the-art tool for assessing protein coding potential [50]. It is also necessary to remove pseudogenes and other classes of RNAs such as tRNAs, rRNAs and snRNAs to avoid misprediction. Accordingly, we established a housekeeping RNA database (see Methods) for similarity-based elimination and obtained 8715 putative long non-coding transcripts after removing housekeeping RNAs (Fig. 1). Further, transcripts derived from the mitochondrial genome were filtered by similarity searches against *A. cerana* mitochondrial protein sequences. After applying all these criteria, we identified 7376 candidate loci to encode 7969 putative lncRNAs.

The predicted lncRNAs were further filtered using the *A. cerana* genome annotation [41] to find those that were intergenic, yielding a total of 2470 lincRNAs from

**Table 1** Details of the RNA-seq data sets from *A. cerana*

| Tissue | NCBI SRA | No. of reads | Reference |
|---|---|---|---|
| *A. cerana* | | | |
| Antenna | SRR1380976 | 55,983,150 | Park et al. [41] |
| Brain | SRR1380970 | 48,168,000 | Park et al. [41] |
| Hypopharyngeal gland | SRR1380979 | 59,548,970 | Park et al. [41] |
| Gut | SRR1380984 | 52,489,846 | Park et al. [41] |
| Fat body | SRR1388774 | 124,626,606 | Park et al. [41] |
| Venom gland | SRR1406762 | 175,970,162 | Park et al. [41] |
| Larvae | SRR1653580 | 77,898,496 | This study |
| SBV control | SRR1653605 | 50,927,126 | This study |
| SBV infected (adult) | SRR1653592 | 49,363,940 | This study |

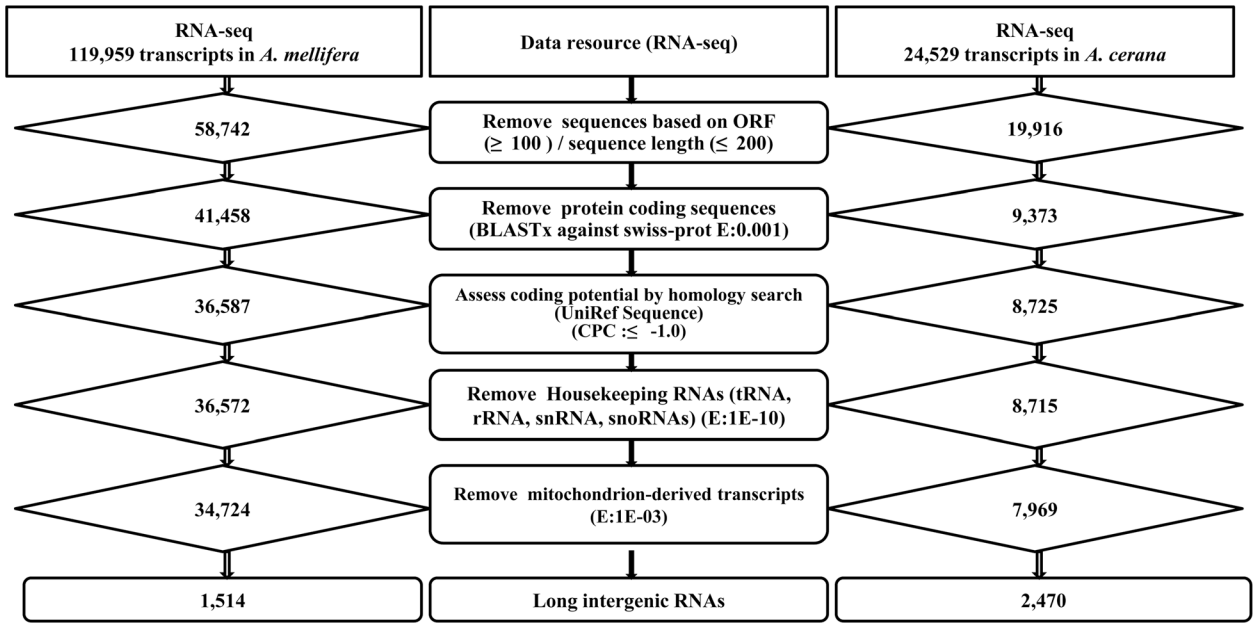


**Fig. 1** Schematic diagram of the bioinformatics pipeline for lincRNA prediction. Step-wise *in silico* filtration of lincRNAs is shown for both *A. mellifera* and *A. cerana*

2376 transcription loci (Fig. 1). From these, we selected 22 putative lincRNAs to validate their prediction and expression using RT-PCR. We used six tissues (antenna, brain, hypoharyngeal gland, gut, fat body, venom gland) for RT-PCR confirmation (Fig. 2) in *A. cerana*. The RT-PCR bands and tissue expression patterns were largely consistent with the RNA-seq data (Additional file 1: Table S1 (B). For example, we selected some lincRNAs based on expression in all analyzed tissues (Fig. 2 (Gel: AC_lincRNA.4118.1, AC_lincRNA.457.1) or specifically in 2 (Fig. 2 (Gel: AC_lincRNA.15188.1, AC_lincRNA.21243.1), 3 (Fig. 2 (Gel: AC_lincRNA.4523.1, AC_lincRNA.5679.1), or a single (Fig. 2 (Gel: AC_lincRNA.12705.1, AC_lincRNA .20163.1) tissue. Most of the lincRNAs exhibited tissue-specific expression, and many lincRNA were detected in brain, hypoharyngeal gland and fat body tissues (Fig. 2).

We also identified lincRNAs from *A. mellifera* using the above-validated method (Fig. 1). First, we retrieved a comprehensive set of 119,959 transcripts from the Transcriptome Shotgun Assembly (TSA) database, which was generated from *de novo* assemblies from RNA-seq datasets of seven tissues in the updated genome annotation of *A. mellifera* [39]. This dataset includes separate tissue transcripts from testes (10,054 transcripts), mixed antennae (14,079 transcripts), embryo (18,613 transcripts), brain & ovary (26,425 transcripts), larvae (9107 transcripts), abdomen (14,372 transcripts), and ovary (27,309 transcripts) [39]. Although these transcripts were generated from deep-transcriptome sequencing, no effort has been made to characterize lncRNAs. We obtained 13,775 putative lincRNAs that were not located at introns or overlapping with any protein coding regions in the *A. mellifera* genome according to the latest gene annotations (OGSv3.2). Since these putative lincRNAs were from separate tissue assemblies for each of the seven tissues, it was possible that the same lincRNA transcript was assembled in two or more tissues. Therefore, in order to remove redundancy, we mapped these lincRNAs in the *A. mellifera* genome sequence (Amel_4.5) and found 1514 unique genomic loci where the putative lincRNAs were clustered. LincRNAs clustered at the same locus could not be assumed to be isoforms due to different tissue assemblies, and hence we selected one lincRNA from each locus based on sequence length and consensus alignment with the genome sequence. Ultimately, we obtained a unique set of 1514 lincRNAs in *A. mellifera*, which were used for further characterization. The exon-intron boundaries for each of these lincRNAs were determined based on the genome alignment.

### Characterization of lincRNAs of *A. cerana* and *A. mellifera*

To investigate the basic features of *A. cerana* lincRNAs, we compared them with protein-coding mRNAs annotated in the Asian honey bee genome project [41]. Most (84 %) of the lincRNAs consisted of a single exon (Fig. 3b), while mRNAs had exon numbers ranging from 1 to over 16. The average length of lincRNA exons was 1232 bp, which is less than that of protein-coding exons. The majority of lincRNAs (over 85 %) were shorter than 2 kb, with very few (3 %) longer than 3 kb (Fig. 3a). The proportion of lincRNAs ranging from 1 to 2 kb was similar to

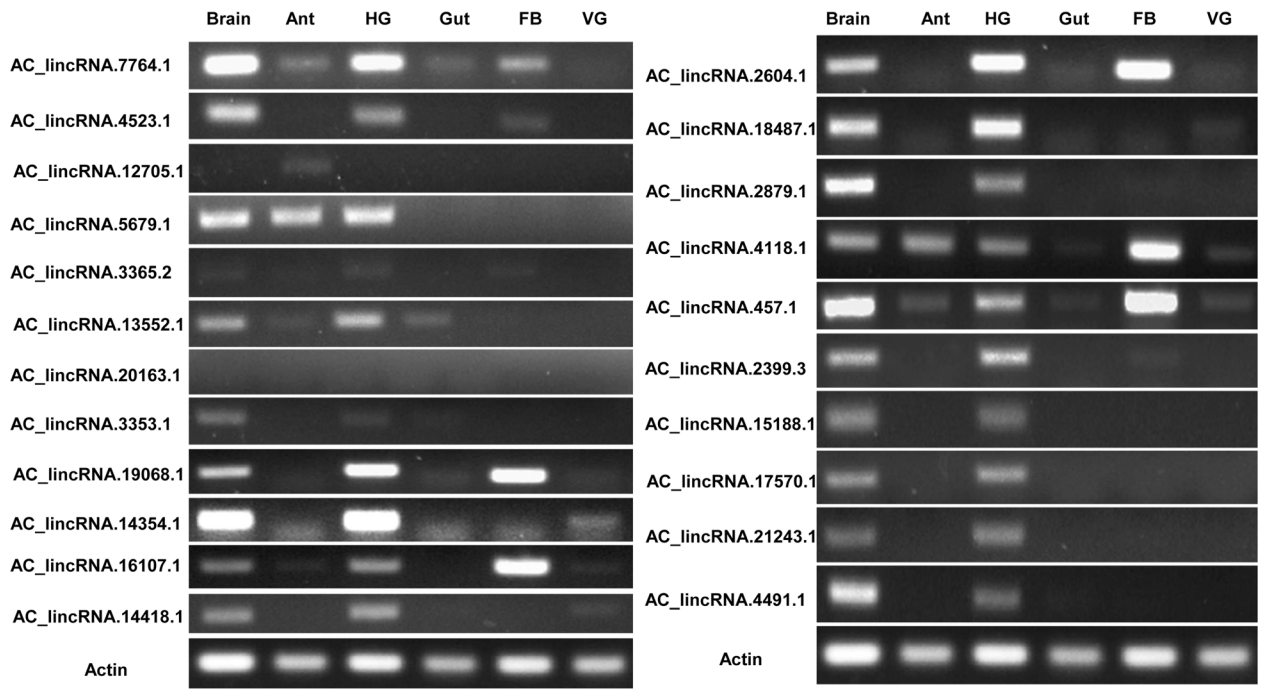


**Fig. 2** RT-PCR validation of lincRNAs in Asian honey bee (*A. cerana*) tissues. Twenty-two putative lincRNAs were selected for RT-PCR validation. Each primer set was used to perform RT-PCR on six RNA samples, including brain, antenna (Ant), hypoharyngeal gland (HG), gut, Fat body (FB), and venom gland (VG), and *actin* was used as a control to show amplification of cDNA samples

that of mRNAs. Repetitive analysis was also performed to make a global view of repeat contents in honeybee lncRNAs. We found a low amount of repetitive content in the *A. cerana* lincRNA dataset. We identified as few as 33 retroelements, of which 3 elements were LINEs and 85 were LTR elements. In addition, we identified a total of 16 DNA transposons. On the whole, simple repeats were abundant (2743) compared to the other repetitive elements in lincRNAs. When we aligned the *A. cerana* lincRNAs with those from other species using BLAST with an e-value cutoff of 1E-03, we found detectable sequence similarity to 101 (4.0 %) lincRNAs from *D. melanogaster*, 176 (7.1 %) from *C. elegans*, 65 (2.6 %) from chicken, 217 (8.7 %) from cow, and 144 (5.8 %) from human.

Similar results were obtained for *A. mellifera*, in which many of the lincRNAs (77 %) consisted of a single exon (Fig. 4b) and the average length of lincRNAs (790 bp) was shorter than that of annotated protein-coding mRNAs (1266 bp). Similar to *A. cerana*, the majority of the *A. mellifera* lincRNAs were within 2 kb of genes (Fig. 3a), and approximately 18 % were distributed within 400 to 500 bp from a gene. The *A. mellifera* lincRNAs showed similarity to fewer than 2 % of the known lncRNAs from other species and were rich in simple repeats (1156). Annotation files describing the genomic features are available as Additional file 2: Dataset S1 and Additional file 3: Dataset S2; Table 2 shows a comparison of the sequence features of lincRNAs identified in this study. All lincRNAs identified in this study and their respective annotation information can be downloaded at http://mnbldb.snu.ac.kr/data.php.

### Comparative analysis of lincRNAs between *A. cerana* and *A. mellifera*

To study the evolutionary dynamics of lincRNAs in honey bees, we aligned *A. cerana* lincRNAs to the *A. mellifera* genome and *A. mellifera* lincRNAs to the *A. cerana* genome using BLAST. The vast majority (2360; 95 %) of putative *A. cerana* lincRNAs could be aligned to the *A. mellifera* genome, with identity ranging from 74 to 100 %. Similarly, 1453 (95 %) of *A. mellifera* lincRNAs aligned to the *A. cerana* genome. The remaining unaligned lincRNAs (5 %) could be specific to each species. Next, we compared lincRNAs from these two sister species with each other and found that 299 *A. cerana* lincRNAs showed perfect matches to 263 *A. mellifera* lincRNAs. The list of 299 lincRNAs in *A. cerana* includes isoforms as identified by Cufflinks and thus includes some sets of multiple transcripts derived from the same locus in *A. cerana* that aligned to a single lincRNA in *A. mellifera*. In addition, we analyzed the exon-intron conservation between these two species based on per-base level identity. The level of identity or conservation was higher for exons as compared to introns in lincRNAs of both species (Additional file 4: Figure S1). Overall, we observed a high level of intron-exon

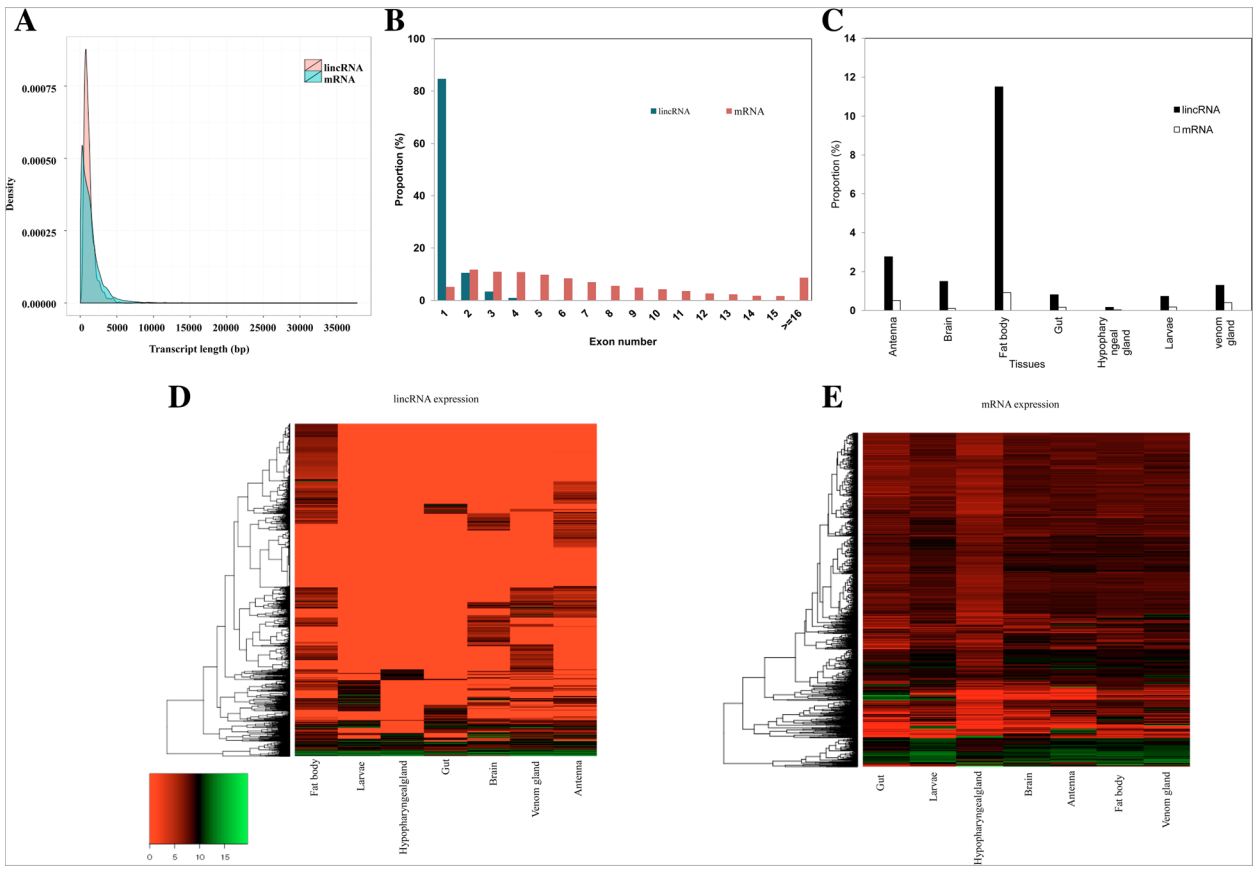


**Fig. 3** Characteristics of Asian honey bee (*A. cerana*) lincRNAs. **a** Length distribution of lincRNAs and mRNAs. **b** Distribution of exon number in lincRNAs and mRNAs. **c** Percentage of lincRNAs and mRNAs showing tissue-specific expression. **d** Heatmap showing expression levels of all lincRNAs (including the 464 tissue-specific lincRNAs, with a cutoff of a two-fold change identified by EdgeR) and **e** mRNAs based on FPKM in seven tissues (brain, antenna, hypoharyngeal gland, gut, fat body, larvae and venom gland)

identity (80–100 %) between *A. mellifera* and *A. cerana*, consistent with recent evolutionary divergence.

### Tissue expression profile

Tissue-specific expression is characteristic of lncRNAs and therefore we profiled the expression of lincRNAs in *A. cerana* (Fig. 3d). Sequencing reads from seven tissues were mapped to *A. cerana* lincRNAs separately and the FPKM (Fragments Per Kilobase of exons per Million fragments generated) score was calculated. Many lincRNAs showed low as well as moderate level expression across seven tissues (Fig. 3d). Among them, 22.3 % of the lincRNAs were expressed in at least 2 tissues, 24.6 % were expressed in 3–5 tissues, and 4.6 % were expressed in all seven tissues. Further, differential expression analysis between seven tissues was conducted using edgeR bioconductor package with $p \leq 0.001$ and fold change $\geq 2$. Overall, we found a subset of 149 differentially-expressed and 464 tissue specifically-expressed lincRNAs. Interestingly, more lincRNAs were preferentially expressed in fat body and antenna tissues than in other tissues in *A. cerana* (Fig. 3c). We also analyzed the analogous characteristics of mRNAs, finding a wide range of expression (low to high FPKM; Fig. 3e). Only 2.3 % of mRNAs showed tissue-specific expression (Fig. 3c) and over 54 % were expressed in all seven tissues.

In *A. mellifera*, the available RNA-seq reads were generated from 400- to 800-bp cDNA fragments by 454 sequencing technology [39]. Generally, long sequencing reads reduce the resolution in gene expression profiling. Therefore, we could not determine the FPKM or RPKM value for *A. mellifera* lincRNAs. However, we determined the expression or presence of each lincRNA in each tissue by comparing each separate tissue assembly with the unique lincRNA set. Approximately 863 (57 %) of the lincRNAs were detected only in one of the seven tissues. Of those, 389 (45 %) and 169 (20 %) were identified in the ovary and brain & ovary tissue assembly dataset, respectively (Fig. 4). Only 3 lincRNAs were identified in all seven tissues, whereas 384 (25 %) and 146 (9 %) lincRNAs were expressed in two and three tissues, respectively.

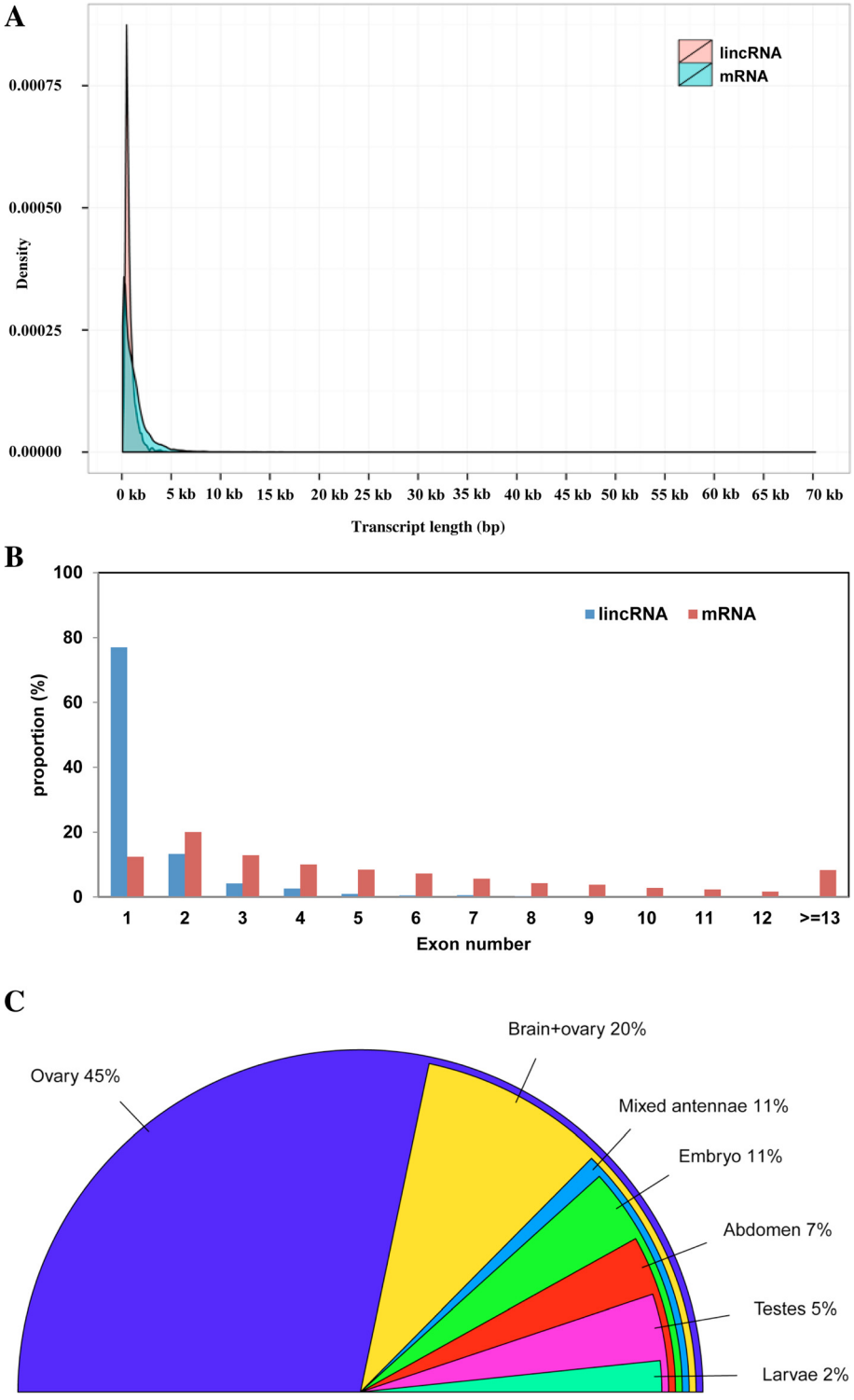


**Fig. 4** Characteristics of Western honey bee (*A. mellifera)* lincRNAs. **a** Length distribution of lincRNAs and mRNAs. **b** Distribution of exon number in lincRNAs and mRNAs. **c** Percentage of lincRNAs specifically expressed in various *A. mellifera* tissues

**Table 2** Comparison of lincRNAs identified in *A. mellifera* and *A. cerana*

| | *A. mellifera* lincRNAs | *A. cerana* lincRNAs |
|---|---|---|
| Number of lincRNAs | 1514 | 2470 |
| Total bases | 1,197,075 | 3,044,196 |
| Maximum length (bp) | 5248 | 9150 |
| Minimum length (bp) | 200 | 204 |
| Average length (bp) | 790 | 1232 |
| GC (%) | 35 | 33 |

### Candidate lincRNAs associated with honey bee viral diseases

One major aim of this study was to identify candidate lincRNAs associated with honey bee diseases. SBV disease is a major diseases of honey bee, especially *A. cerana*, in which the SBV attacks brood and adult stages of bees and thus leads to decreased life span [51]. We compared lincRNA expression between SBV-infected and non-infected (control) honey bees. First, we analyzed the RNA-seq data set derived from the SBV control and infected bees of *A. cerana*. We identified 15 lincRNAs that showed significant differential expression between SBV control and infected data using read mapping and the edgeR bioconductor package. We selected those differentially-expressed transcripts for qRT-PCR validation in both adult and larvae stages of *A. cerana*. Of 15 lincRNAs, 11 showed expression consistent with our RNA-seq results of significant differential expression between SBV control and infected honey bees. Among them, one lincRNA (ID: AC_lincRNA.3472.1) was down-regulated and the rest were up-regulated in both adult and larval SBV-infected *A. cerana*. These lincRNAs represent candidates to play specific roles in SBV replication and regulation of SBV-resistance genes.

Additionally, we examined the responses of these lincRNAs to other viral diseases in the honey bee to investigate if they are specific to SBV. DWV, a viral disease closely linked to *Varroa* mite infestation [52], causes wing deformity and premature death in adult honey bees of *A. mellifera* [52]. Hence, we investigated the expression patterns of those same 11 lincRNAs in DWV-infected and healthy uninfected *A. mellifera* honey bees using qRT-PCR. Intriguingly, we found that 10 of the lincRNAs showed a similar expression pattern in response to infection in this species, including the one down-regulated (AC_lincRNA.3472.1; Fig. 5). Furthermore, RT-PCR products for 10 lincRNAs from *A. cerana* were sequenced and found to match exactly to those lincRNAs (Additional file 5: Dataset S3). Together, these results suggest that this subset of 10 lincRNAs may play critical roles in pathogen-host interactions in honey bees. Therefore we regarded these lincRNAs as virus-specific lincRNAs in honey bee. We have submitted these virus-specific lincRNAs to GenBank (Acc. Nos.: KM889914-KM889923).

Most of these virus-specific lincRNAs contain a single exon, except AC_lincRNA.16646.1 and AC_lincRNA.3472.1, which contain two exons. The size of these lincRNAs ranged from 322 to 813 bp. In terms of digital expression, apart from the SBV datasets, AC_lincRNA.16646.1 and AC_lincRNA.4466.1 were found to have low-level expression in healthy non-infected antenna, brain, fat body and venom gland tissues, whereas AC_lincRNA.2775.1 and AC_lincRNA.231.1 showed expression only in fat body. AC_lincRNA.14547.1 and AC_lincRNA.3472.1 were uniquely found in the brain and venom gland datasets, respectively. The remaining lincRNAs were identified only in the SBV dataset. By contrast, although these lincRNAs identified in *A. cerana* were found in the genome sequence of *A. mellifera*, they were not identified in any of the *A. mellifera* tissue transcriptomes, which were assembled from healthy honey bees, suggesting that these lincRNAs were not expressed or were expressed at too low a level to be detected by RNA sequence data. These findings support the idea that the set of virus-specific lincRNAs are expressed only upon viral infection in both honey bee species. Using BLAST, we found that the level of identity of these candidate lincRNAs with those from *A. mellifera* ranged from 91 to 96 % both in exon and intron regions. This high level of identify suggests that the set of shared candidate lincRNAs might function similarly in the two species.

## Discussion

RNA sequencing (RNA-seq) is a powerful tool that enables the research community to discriminate cellular transcripts quantitatively [53]. It has been successfully employed for transcriptome profiling in various model and non-model organisms [54, 55]. Similarly, RNA-Seq has been used for lncRNA identification in both plants and mammals [56–60]. Due to numerous potential roles of lncRNAs in the genome and in organismal development [9–16], characterizing lncRNAs has become essential to understand gene regulation in eukaryotic species.

### Identification of lincRNAs in honey bees

Previously, four lncRNAs, two from the adult brain [47] and two from the larval ovaries [46], were identified in *A. mellifera*, all of which were intronic and natural antisense type. In this study, we have identified a relatively robust set of potential lincRNAs from *A. cerana* (Asian honey bee), which is a good model for behavior research due to their fascinating habits of grooming, hygiene and aggregation against predators [41], and *A. mellifera* (Western honey bee) which has served as a key model for social insects [40]. We used a total of 694 million sequencing reads from *A. cerana* and identified a set of

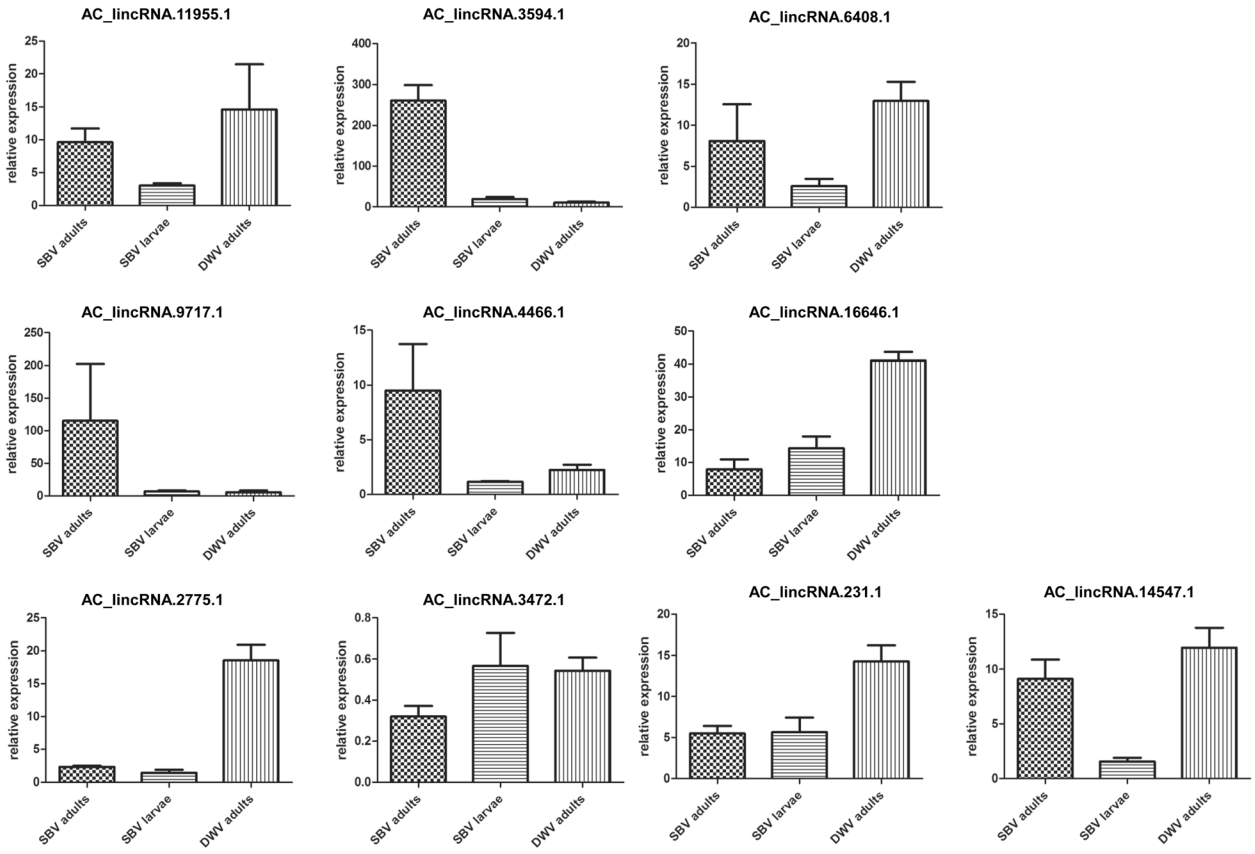


**Fig. 5** Differential expression of virus-specific lincRNAs. Relative expression of lincRNAs between uninfected (Control) and SBV- and DWV-infected bees. Quantitative analysis was carried out using StepOne plus Software V. 2.0 (Applied Biosystems). Results were normalized to a validated control gene, *actin* (data not shown), for which values were set to 1, using the ΔΔCt method for each sample [58]. SBV- and DWV-infected samples were collected from *A. cerana* and *A. mellifera*, respectively

lincRNAs using our bioinformatics protocol. In general, transcriptome assembly produces many partial sequences and some antisense transcripts that can be aligned only at intronic regions of a protein coding gene rather than overlapping onto that gene's exons due to partial assembly. It is possible that such transcripts could be wrongly annotated as intronic-type lncRNAs as they might not contain potential ORFs or protein-coding domains. Therefore, intergenic type lncRNAs are most appropriately and reliably predicted from RNA-seq datasets obtained using poly-A primers and non-strand-specific methods, and hence, our lincRNA datasets did not include non-polyadenylated, antisense, nor intronic-type lncRNAs. In our current data, the distance between identified lincRNAs and UTR regions of the neighboring genes varies from over 200 bp to 1 kb. It is also possible that some of the lincRNAs, for which we might have only partial data at present, span a UTR region, especially if the honeybee genomes have poor UTR annotation. This possibility can be addressed in the future with more comprehensive transcriptome assembly and genome annotation from very large-scale RNA-seq data. We have made available the lincRNA features in GTF file format, which can be used in genome browsers and should enhance the honey bee genome annotation. LncRNAs show more plasticity than protein-coding genes and thus lack of conservation in general [61, 62]. Not surprisingly, our lincRNAs also exhibited rather less similarity to those of other species and shared more similarity with those of the sister species.

## Expression profiling

Similar to the results of other lncRNA studies, we identified a subset of tissue-dependent lincRNAs with almost no expression elsewhere, suggesting that the expression of these lincRNAs is tightly controlled in a tissue/development-specific manner. Approximately 11 % of the total lincRNAs were uniquely expressed in fat body tissue in *A. cerana*. Fat bodies, in which energy is stored and released according to the demands of the insect, play an important role in honey bee health [63]. In addition, major metabolic and hormone signaling pathways reside in fat bodies [64]. Consistent with our results, lncRNAs have been implicated in the development

of insulin resistance and tissue-specific regulation of metabolism [65]. Therefore, we speculate that lincRNAs in fat bodies may be involved in various metabolic pathways, including hormone biosynthesis, in *A. cerana*. Ovary has also been reported as a preferential tissue for lncRNA expression [66], and we identified many lincRNAs specific to *A. mellifera* ovary tissue. We found more tissue-specific lincRNAs in *A. mellifera* than in *A. cerana*, which could be due to unequal sequencing depth and biased transcriptome assembly in the various tissues.

### Implications for disease-specific lincRNAs of the honey bee

To date, 22 viruses have been reported to infect honey bees [52], and many of these have also been reported to be associated with *Varroa* parasitism [52]. Pathogens are proposed to be the major contributors to honey bee mortality [52]. Previously, Peng et al. [67] discovered that lncRNAs exhibit unique expression in response to viral infection [68] in mice. Here, we identified candidate lincRNAs associated with viral diseases of honey bee. Among the virus-specific lincRNAs, one lincRNA was expressed more highly in healthy adult bees of both *A. mellifera* and *A. cerana*, whereas the ten others were all up-regulated in both SBV- and DWV-infected bees. This demonstrates that the virus infections modulate the expression levels of many lincRNAs. Since we identified 11 lincRNAs as differentially expressed in SBV-infected honey bee, our findings support the idea that lncRNAs can be regulators in determining the outcome of infection as demonstrated by Peng et al. [66]. Ten lincRNAs were also observed as responding to two different viral diseases, suggesting that a subset of lncRNAs plays critical roles during viral infection in general. Determining what specific roles lincRNAs play in virus-host interactions awaits further research.

## Conclusion

Emerging reports have suggested that lncRNAs play important functional roles in disease, development and various biological processes in eukaryotes. In this study, we have provided a comprehensive set of lincRNAs in honey bees. We identified more than 1000 lincRNAs in *A. cerana* and *A. mellifera*, which were likely to exhibit tissue-specific expression patterns, as validated by expression profiling and RT-PCR analysis. The lincRNAs were less conserved than protein-coding mRNAs and contained low repeat content. Finally, we identified lincRNAs associated with SBV and DWV diseases and confirmed their differential expression by qRT-PCR. This study thus provides the first comprehensive genome-wide analysis of honey bee lincRNAs and paves the way for identification of lincRNAs associated with general development, biological and hormone signaling pathways and disease resistance.

## Methods

### RNA isolation and next generation sequencing

Bees of *A. cerana* were taken from 3 different colonies at an apiary located at the College of Agriculture and Life Sciences, Seoul National University (SNU), Seoul, Korea during the summer season. While we conducted this experiment, the queen bee was not changed genetically. Worker bees were captured and directly placed in liquid $N_2$, and stored at −80 °C. Tissues were dissected in cold RNase-free PBS (pH = 7.4). Total RNA was isolated from *A. cerana* larvae and SBV-infected and non-infected honey bees using a QIAGEN RNeasy Mini kit (Qiagen, CA, US) according to the manufacturer's protocol. The complementary DNA was prepared for each tissue using the Illumina mRNA sequencing kit (Illumina, CA, US) and the Clontech SMART cDNA Library Construction Kit (Invitrogen). Libraries were sequenced using Illumina HiSeq2000.

### Pipeline for identifying lincRNAs

1) RNA-seq data were aligned to the *A. cerana* reference genome using the spliced aligner Tophat [46] (with –no-discordant, –-no-mixed parameters) and then assembled using Cufflinks [47] (with parameters -u –library-type fr-unstranded).

2) The assembled transcripts were initially filtered based on size and ORF. An in-house perl script was used to select transcripts that were ≥200 bp in length and ORFs of ≤100 amino acids.

3) Then, transcripts were compared to the Swiss-Prot protein database to eliminate protein coding transcripts using BlastX with an E-value cutoff of 1E-03.

4) To calculate the coding potential, the coding potential calculator (CPC) [48] was utilized with default parameters. Transcripts with non-coding scores were considered as lncRNAs.

5) To eliminate housekeeping non-coding RNAs (transfer (t) RNAs, small nuclear (sn) RNAs and small nucleolar (sno) RNAs), a housekeeping RNA database was made, and putative lncRNAs were aligned to the database with an E-value cutoff of 1E-10. This database contained tRNA sequences from the genomic tRNA database (http://gtrnadb.ucsc.edu/), rRNAs from the silva database (http://www.arb-silva.de/no_cache/download/archive/current/Exports/), and other ncRNAs (snRNAs, snoRNAs, 7SL/SRP) downloaded from NONCODE (http://noncode.org/).

6) Transcripts derived from the mitochondrial genome were removed by alignment against the mitochondrial protein sequences of *A. cerana* and *A. mellifera* downloaded from NCBI (GenBank accession GQ162109 and NC_001566).

7) Finally, gene annotation information for both *A. cerana* (http://mnbldb.snu.ac.kr/) [41] and *A. mellifera*

(OGSv3.2: http://hymenopteragenome.org/beebase/) were retrieved from their genome databases. An in-house python script was developed to identify lncRNAs located between two genes, and those lncRNAs were regarded as lincRNAs in this study.

### Expression and alignment of lincRNAs

Exon-intron alignment between the two honey bee sister species were performed using BLAST with over 40 % alignments and an E-value threshold of $e^{-10}$. Tissue specificity was determined based on the condition that the FPKM should be >1 in the specific tissue and zero in the rest of the tissues. This condition was set based RT-PCR results from randomly selected lincRNAs in *A. cerana*.

### Validation of putative lincRNAs by RT-PCR

RT-PCR was conducted for 22 lincRNAs in 6 tissues (antenna, brain, hypoharyngeal gland, gut, fat body, venom gland) in *A. cerana*. The primer pairs were selected using Primer 3, and the primer sequences are presented in Additional file 1: Table S1 (A). PCR was conducted using 2X Premix-MG Taq (Macrogen, Cat No. MP018S) following the manufacturer's instructions under the following conditions: pre-denaturation step at 95 °C for 3 min; 30 amplification cycles of denaturation at 95 °C for 30 s, annealing at 50 °C for 30 s, and elongation at 72 °C for 30 s; followed by a final elongation step at 72 °C for 5 min. Electrophoresis was conducted using 1.2 % agarose gels.

### Quantitative real-time PCR assay

Expression of selected differentially expressed lincRNAs between SBV control and infected bee samples was analyzed through qRT-PCR. RNA was isolated from the SBV-infected and healthy adult as well as larval bees of *A. cerana* using the RNeasy kit (Qiagen) according to manufacturer's instructions. Similarly, RNA was isolated from DWV control and infected bees of *A. mellifera*. cDNA was prepared from 500 ng RNA using Superscript III (Invitrogen, USA). The PCR amplification was carried out using SYBR Green PCR Master Mix (Applied biosystems, UK). Primer sequences are presented in Additional file 6: Table S2. The data presented here represent three independent biological and technical replicates.

### Availability of supporting data

The RNA-seq data generated in this study are available in the NCBI Sequence Read Archives (SRA) with accessions SRR1653580, SRR1653605, and SRR1653592 for larvae, SBV-non-infected (control), and SBV-infected adult bees of *A. cerana*. The candidate virus-specific lincRNAs are available at GenBank (KM889914-KM889923). All the assembled lincRNAs and their analysis results are available at http://mnbldb.snu.ac.kr/data.php.

## Additional files

**Additional file 1: Table S1.** (A) Primers used for RT-PCR validation and (B) FPKM values of seven tissues of *A. cerana* used in this study. (XLS 328 kb)

**Additional file 2: Dataset S1.** Annotation of *A. cerana* lincRNAs. (GTF 1130 kb)

**Additional file 3: Dataset S2.** Annotation of *A. mellifera* lincRNAs. (GTF 403 kb)

**Additional file 4: Figure S1.** Intron-exon conservation of lincRNAs between *A. mellifera* and *A. ceana*. Y-axis (density) represents the total lincRNAs and X-axis indicates their identity. (TIFF 1637 kb)

**Additional file 5: Dataset S3.** Sequenced PCR products of virus-specific lincRNAs. (TXT 1 kb)

**Additional file 6: Table S2.** Primers used in qRT-PCR analysis. (XLS 28 kb)

#### Abbreviations

lincRNA: Long intergenic non-coding RNA; SBV: Sacbrood virus; DWV: Deformed wing virus; lncRNA: Long non-coding RNA; bp: Basepair; kb: Kilobase; FPKM: Fragments per kilobase of exon per million fragments; PCR: Polymerase chain reaction; OGS: Official gene set; ORF: Open reading frame.

#### Competing interests

The authors declare that they have no competing interests.

#### Authors' contributions

MJ, DP, TJY, and HWK conceived and designed this project. MJ and SCL carried out the bioinformatics analysis. DP, SS, and JW performed all the experimental parts. HWK participated in its design and coordination. MJ, DP, SCL, CSS, TJY, YJA, and HWK wrote the manuscript. All authors read and approved the final manuscript.

#### Acknowledgements

This work was supported by the World Class University (WCU) program (R31-10056) through the Korea Science and Engineering Foundation funded by the National Research Foundation of Korea. Also, this work was supported by Cooperative Research Program for Agriculture Science & Technology Development (Project No. PJ010487) from the Rural Development Administration (RDA) of the Republic of Korea to HWK.

#### Author details

[1]Department of Plant Science, Plant Genomics and Breeding Institute, Research Institute of Agriculture and Life Sciences, College of Agriculture and Life Sciences, Seoul National University, Seoul 151-921, Republic of Korea. [2]WCU Biomodulation Major, Department of Agricultural Biotechnology and Research Institute of Agriculture and Life Sciences, College of Agriculture and Life Sciences, Seoul National University, Seoul 151–921, Republic of Korea. [3]Department of Agricultural Biotechnology and Research Institute of Agriculture and Life Sciences, College of Agriculture and Life Sciences, Seoul National University, Seoul 151-921, Republic of Korea.

Received: 24 January 2015 Accepted: 19 August 2015
Published online: 04 September 2015

#### References

1. Okazaki Y, Furuno M, Kasukawa T, Adachi J, Bono H, Kondo S, et al. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. Nature. 2002;420:563–73.
2. Maeda N, Kasukawa T, Oyama R, Gough J, Frith M, Engstrom PG, et al. Transcript annotation in FANTOM3: mouse gene catalog based on physical cDNAs. PLoS Genet. 2006;2:e62.
3. Kapranov P, Willingham AT, Gingeras TR. Genome-wide transcription and the implications for genomic organization. Nat Rev Genet. 2007;8:413–23.
4. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, et al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res. 2012;22:1775–89.
5. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012;489:57–74.

6. Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, et al. Landscape of transcription in human cells. Nature. 2012;489:101–8.
7. Bertone P, Stolc V, Royce TE, Rozowsky JS, Urban AE, Zhu X, et al. Global identification of human transcribed sequences with genome tiling arrays. Science. 2004;306:2242–6.
8. Wierzbicki AT. The role of long non-coding RNA in transcriptional gene silencing. Curr Opin Plant Biol. 2012;15:517–22.
9. Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell. 2007;129:1311–23.
10. Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ, et al. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature. 2010;464:1071–6.
11. Tsai MC, Manor O, Wan Y, Mosammaparast N, Wang JK, Lan F, et al. Long noncoding RNA as modular scaffold of histone modification complexes. Science. 2010;329:689–93.
12. Martianov I, Ramadass A, Serra Barros A, Chow N, Akoulitchev A. Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript. Nature. 2007;445:666–70.
13. Wang X, Arai S, Song X, Reichart D, Du K, Pascual G, et al. Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. Nature. 2008;454:126–30.
14. Tripathi V, Ellis JD, Shen Z, Song DY, Pan Q, Watt AT, et al. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. Mol Cell. 2010;39:925–38.
15. Mercer TR, Dinger ME, Mattick JS. Long non-coding RNAs: insights into functions. Nat Rev Genet. 2009;10:155–9.
16. Wilusz JE, Sunwoo H, Spector DL. Long noncoding RNAs: functional surprises from the RNA world. Genes Dev. 2009;23:1494–504.
17. Yang X, Gao L, Guo X, Shi X, Wu H, Song F, et al. A network based method for analysis of lncRNA-disease associations and prediction of lncRNAs implicated in diseases. PLoS One. 2014;9, e87797.
18. Clark BS, Blackshaw S. Long non-coding RNA-dependent transcriptional regulation in neuronal development and disease. Front Genet. 2014;5:164.
19. Paralkar VR, Mishra T, Luan J, Yao Y, Kossenkov AV, Anderson SM, et al. Lineage and species-specific long noncoding RNAs during erythro-megakaryocytic development. Blood. 2014;123:1927–37.
20. Li J, Chen Z, Tian L, Zhou C, He MY, Gao Y, et al. LncRNA profile study reveals a three-lncRNA signature associated with the survival of patients with oesophageal squamous cell carcinoma. Gut. 2014.
21. Chen G, Wang Z, Wang D, Qiu C, Liu M, Chen X, et al. LncRNADisease: a database for long-non-coding RNA-associated diseases. Nucleic Acids Res. 2013;41:D983–6.
22. Ammosova T, Yedavalli VR, Niu X, Jerebtsova M, Van Eynde A, Beullens M, et al. Expression of a protein phosphatase 1 inhibitor, cdNIPP1, increases CDK9 threonine 186 phosphorylation and inhibits HIV-1 transcription. J Biol Chem. 2011;286:3798–804.
23. Saayman S, Ackley A, Turner AM, Famiglietti M, Bosque A, Clemson M, et al. An HIV-encoded antisense long noncoding RNA epigenetically regulates viral transcription. Mol Ther. 2014;22:1164–75.
24. Faghihi MA, Modarresi F, Khalil AM, Wood DE, Sahagan BG, Morgan TE, et al. Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of beta-secretase. Nat Med. 2008;14:723–30.
25. Calin GA, Liu CG, Ferracin M, Hyslop T, Spizzo R, Sevignani C, et al. Ultraconserved regions encoding ncRNAs are altered in human leukemias and carcinomas. Cancer Cell. 2007;12:215–29.
26. Chen H, Xu J, Hong J, Tang R, Zhang X, Fang JY. Long noncoding RNA profiles identify five distinct molecular subtypes of colorectal cancer with clinical relevance. Mol Oncol. 2014;8(8):1393–403.
27. Crea F, Watahiki A, Quagliata L, Xue H, Pikor L, Parolia A, et al. Identification of a long non-coding RNA as a novel biomarker and potential therapeutic target for metastatic prostate cancer. Oncotarget. 2014;5:764–74.
28. Chow JC, Yen Z, Ziesche SM, Brown CJ. Silencing of the mammalian X chromosome. Annu Rev Genomics Hum Genet. 2005;6:69–92.
29. Sleutels F, Zwart R, Barlow DP. The non-coding Air RNA is required for silencing autosomal imprinted genes. Nature. 2002;415:810–3.
30. Khalil AM, Guttman M, Huarte M, Garber M, Raj A, Rivea Morales D, et al. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. Proc Natl Acad Sci U S A. 2009;106:11667–72.
31. Dinger ME, Amaral PP, Mercer TR, Pang KC, Bruce SJ, Gardiner BB, et al. Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. Genome Res. 2008;18:1433–45.
32. Li L, Wang X, Stolc V, Li X, Zhang D, Su N, et al. Genome-wide transcription analyses in rice using tiling microarrays. Nat Genet. 2006;38:124–9.
33. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods. 2008;5:621–8.
34. Liao Q, Shen J, Liu J, Sun X, Zhao G, Chang Y, et al. Genome-wide identification and functional annotation of Plasmodium falciparum long noncoding RNAs from RNA-seq data. Parasitol Res. 2014;113:1269–81.
35. Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature. 2009;458:223–7.
36. Young RS, Marques AC, Tibbit C, Haerty W, Bassett AR, Liu JL, et al. Identification and properties of 1,119 candidate lincRNA loci in the Drosophila melanogaster genome. Genome Biol Evol. 2012;4:427–42.
37. Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev. 2011;25:1915–27.
38. Guttman M, Garber M, Levin JZ, Donaghey J, Robinson J, Adiconis X, et al. Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. Nat Biotechnol. 2010;28:503–10.
39. Elsik CG, Worley KC, Bennett AK, Beye M, Camara F, Childers CP, et al. Finding the missing honey bee genes: lessons learned from a genome upgrade. BMC Genomics. 2014;15:86.
40. Honeybee Genome Sequencing Consortium. Insights into social insects from the genome of the honeybee Apis mellifera. Nature. 2006;443:931–49.
41. Park D, Jung JW, Choi BS, Jayakodi M, Lee J, Lim J, et al. Uncovering the novel characteristics of Asian honey bee, Apis cerana, by whole genome sequencing. BMC Genomics. 2014.
42. Galizia GC, Eisenhardt D, Giurfa M. Honeybee neurobiology and behavior:a tribute to randolf menzel. Springer Netherlands: Dordrecht, Netherlands; 2012.
43. Begna D, Han B, Feng M, Fang Y, Li J. Differential expressions of nuclear proteomes between honeybee (Apis mellifera L.) Queen and worker larvae: a deep insight into caste pathway decisions. J Proteome Res. 2012;11:1317–29.
44. Zayed A, Robinson GE. Understanding the relationship between brain gene expression and social behavior: lessons from the honey bee. Annu Rev Genet. 2012;46:591–615.
45. 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 U S A. 2012;109:4968–73.
46. Humann FC, Tiberio GJ, Hartfelder K. Sequence and expression characteristics of long noncoding RNAs in honey bee caste development–potential novel regulators for transgressive ovary size. PLoS One. 2013;8, e78915.
47. Sawata M, Yoshino D, Takeuchi H, Kamikouchi A, Ohashi K, Kubo T. Identification and punctate nuclear localization of a novel noncoding RNA, Ks-1, from the honeybee brain. RNA. 2002;8:772–85.
48. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics. 2009;25:1105–11.
49. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol. 2010;28:511–5.
50. Kong L, Zhang Y, Ye ZQ, Liu XQ, Zhao SQ, Wei L, et al. CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. Nucleic Acids Res. 2007;35:W345–9.
51. Choe SE, Nguyen TT, Hyun BH, Noh JH, Lee HS, Lee CH, et al. Genetic and phylogenetic analysis of South Korean sacbrood virus isolates from infected honey bees (Apis cerana). Vet Microbiol. 2012;157:32–40.
52. Mondet F, de Miranda JR, Kretzschmar A, Le Conte Y, Mercer AR. On the front line: quantitative virus dynamics in honeybee (Apis mellifera L.) colonies along a New expansion front of the parasite varroa destructor. PLoS Pathog. 2014;10:e1004323.
53. Berretta J, Morillon A. Pervasive transcription constitutes a new level of eukaryotic genome regulation. EMBO Rep. 2009;10:973–82.
54. Garg R, Patel RK, Tyagi AK, Jain M. De novo assembly of chickpea transcriptome using short reads for gene discovery and marker identification. DNA Res. 2011;18:53–63.

55. Li D, Deng Z, Qin B, Liu X, Men Z. De novo assembly and characterization of bark transcriptome using Illumina sequencing and development of EST-SSR markers in rubber tree (Hevea brasiliensis Muell. Arg.). BMC Genomics. 2012;13:192.
56. Lv J, Cui W, Liu H, He H, Xiu Y, Guo J, et al. Identification and characterization of long non-coding RNAs related to mouse embryonic brain development from available transcriptomic data. PLoS One. 2013;8, e71152.
57. Pauli A, Valen E, Lin MF, Garber M, Vastenhouw NL, Levin JZ, et al. Systematic identification of long noncoding RNAs expressed during zebrafish embryogenesis. Genome Res. 2012;22:577–91.
58. Nam JW, Bartel DP. Long noncoding RNAs in C. elegans. Genome Res. 2012;22:2529–40.
59. Zhou ZY, Li AM, Adeola AC, Liu YH, Irwin DM, Xie HB, et al. Genome-wide identification of long intergenic noncoding RNA genes and their potential association with domestication in pigs. Genome Biol Evol. 2014;6:1387–92.
60. Li L, Eichten SR, Shimizu R, Petsch K, Yeh CT, Wu W, et al. Genome-wide discovery and characterization of maize long non-coding RNAs. Genome Biol. 2014;15:R40.
61. Johnsson P, Lipovich L, Grander D, Morris KV. Evolutionary conservation of long non-coding RNAs; sequence, structure, function. Biochim Biophys Acta. 1840;2014:1063–71.
62. Ponting CP, Oliver PL, Reik W. Evolution and functions of long noncoding RNAs. Cell. 2009;136:629–41.
63. Arrese EL, Soulages JL. Insect fat body: energy, metabolism, and regulation. Annu Rev Entomol. 2010;55:207.
64. Wang Y, Brent CS, Fennern E, Amdam GV. Gustatory perception and fat body energy metabolism are jointly affected by vitellogenin and juvenile hormone in honey bees. PLoS Genet. 2012;8, e1002779.
65. Kornfeld JW, Bruning JC. Regulation of metabolism by long, non-coding RNAs. Front Genet. 2014;5:57.
66. Necsulea A, Soumillon M, Warnefors M, Liechti A, Daish T, Zeller U, et al. The evolution of lncRNA repertoires and expression patterns in tetrapods. Nature. 2014;505:635–40.
67. Peng X, Gralinski L, Armour CD, Ferris MT, Thomas MJ, Proll S, et al. Unique signatures of long noncoding RNA expression in response to virus infection and altered innate immune signaling. MBio. 2010;1.
68. Zhang Q, Jeang KT. Long non-coding RNAs (lncRNAs) and viral infections. Biomed Pharmacother. 2013;3:34–42.