De novo transcriptome analysis shows differential expression of genes in salivary glands of edible bird’s nest producing swiftlets

Q. H. Looi1, H. Amin1, I. Ain1,2, M. Zuki1,2 and A. R. Omar1,2*

Abstract

**Background:** Edible bird’s nest (EBN), produced from solidified saliva secretions of specific swiftlet species during the breeding season, is one of the most valuable animal by-products in the world. The composition and medicinal benefits of EBN have been extensively studied, however, genomic and transcriptomic studies of the salivary glands of these birds have not been conducted.

**Results:** The study described the transcriptomes of salivary glands from three swiftlet species (28 samples) generated by RNASeq. A total of 14,835 annotated genes and 428 unmapped genes were cataloged. The current study investigated the genes and pathways that are associated with the development of salivary gland and EBN composition. Differential expression and pathway enrichment analysis indicated that the expression of CREB3L2 and several signaling pathways involved in salivary gland development, namely, the EGFR, BMP, and MAPK signaling pathways, were up-regulated in swiftlets producing white EBN (Aerodramus fuciphagus) and black EBN (Aerodramus maximus) compared with non-EBN-producing swiftlets (Apus affinis). Furthermore, MGAT, an essential gene for the biosynthesis of N-acetylneuraminic acid (sialic acid), was highly expressed in both white- and black-nest swiftlets compared to non-EBN-producing swiftlets. Interspecies comparison between Aerodramus fuciphagus and Aerodramus maximus indicated that the genes involved in N-acetylneuraminic and fatty acid synthesis were up-regulated in Aerodramus fuciphagus, while alanine and aspartate synthesis pathways were up-regulated in Aerodramus maximus. Furthermore, gender-based analysis revealed that N-glycan trimming pathway was significantly up-regulated in male Aerodramus fuciphagus from its natural habitat (cave) compared to their female counterpart.

**Conclusions:** Transcriptomic analysis of salivary glands of different swiftlet species reveal differential expressions of candidate genes that are involved in salivary gland development and in the biosynthesis of various bioactive compounds found in EBN.

**Keywords:** Aerodramus fuciphagus, Aerodramus maximus, Apus affinis, Transcriptome, Salivary gland

**Background**

Swiftlets are small, aerial, insectivorous birds, and are widely distributed; they extend from the Western Indian Ocean to the Pacific Ocean [1, 2]. The swiftlet species roost in caves or in a cave-like, man-made environments called swiftlet houses [3]. During the breeding season, both male and female swiftlets of the Aerodramus species collaborate to produce highly valuable edible bird’s nest (EBN) from saliva secretions generated from their salivary glands [4, 5]. Hence, the salivary glands of these swiftlets are relatively bigger during the breeding season compared to other seasons [6]. A study on the anatomy of swiftlet report that the sublingual glands of swiftlet enlarge tremendously during the breeding season, i.e. from 2.5 mg to 160 mg, while fledgling birds were found to have small, non-secretory salivary glands [6, 7]. Such
dramatic changes in the gland size and weight are probably associated with the proliferation and differentiation of specialized cells, presumably stem cells, which control the development of the glands during breeding season [8]. However, no significant enlargement of salivary glands was observed in non-EBN-producing swiftlets (Apus species) which produce nests that primarily comprise of dried grass [4].

EBN is a luxurious delicacy, and has been used in Traditional Chinese Medicine for hundreds of years; it is known as one of the world’s most valuable animal products [4, 7]. Nests from the white-nest swiftlet (Aerodramus fuciphagus) and the black-nest swiftlet (Aerodramus maximus) are harvested for commercial purposes. Compared with black-nests (A. maximus), white-nests from A. fuciphagus are premium grade as they comprise solely of solidified saliva secretions and contain high concentrations of N-acetylglucosamine, lactoferrin, sialic acid, amino acids, fatty acids, triacylglycerol, vitamins, minerals, and other antioxidant molecules [12]. Despite efforts to investigate the composition of EBN and its medicinal properties, little is known regarding the genes that are transcriptionally active in the salivary glands of swiftlets.

Genomic characterization of avian species has been proven to be of importance for both socio-economic and ecological reasons as many avian species are vital to agricultural and poultry industries. Hence, the genome sequences of several avian species, such as that of chickens, pigeons and turkeys, have been investigated over the years [13–15]. However, studies on genes or the genome of swiftlets are scarce. As of February 2017, there were approximately 850 sequences under Aerodramus spp. in the nucleotide database of NCBI, with Aerodramus fuciphagus being the most extensively studied species. The absence of an established swiftlet genome has hindered many fundamental studies on salivary gland development and saliva secretion during EBN construction.

Traditionally, RNA sequencing projects require tedious laboratory manipulation and are time-consuming. However, current advancements in next-generation sequencing (NGS) techniques have provided researchers with cost- and time-effective platforms for transcriptomic analyses [16–18]. In the current study, we performed de novo transcriptome sequencing of EBN-producing swiftlets (Aerodramus fuciphagus and Aerodramus maximus) and non-EBN-producing swiftlets (Apus affinis) to generate individual reference transcripts for each species. The generated reference transcripts were further analyzed for gene annotation, pathway enrichment, and differential gene expression studies. The transcriptome profiles of the salivary glands of the Aerodramus species and the Apus affinis showed us a differential expression of genes that are important in saliva secretion and the composition of EBN during nest construction.

**Methods**

**Animal ethics and sampling**

RNA sequencing libraries were derived from the salivary glands of Aerodramus fuciphagus (A. fuciphagus), Aerodramus maximus (A. maximus), and Apus affinis (Additional file 1: Table S1). A total of 16 Aerodramus fuciphagus samples (eight from a natural habitat, i.e. caves, and eight from a man-made habitat, i.e. houses), 8 Aerodramus maximus samples from a cave, and 2 Apus affinis samples were collected from various locations (Additional file 1: Table S1). Salivary gland samples of A. maximus from man-made houses were not included in this study as farmers rarely breed A. maximus due to the low value of its nests compared with A. fuciphagus. Upon necropsy, the sublingual salivary glands were removed, weighed, and then stored in RNAlater® solution (Ambion, USA) at −80 °C for RNA extraction. Only well-developed salivary glands (0.06–0.10 g) and partially developed salivary glands (0.01–0.05 g) from adult birds were considered for the RNA sequencing (Additional file 2: Table S2). This sampling strategy was adopted to obtain unbiased and comprehensive transcriptomic profiles, because during the breeding season swiftlets have both types of salivary glands.

**Species identification**

Identification of the species of the sampled swiftlets was carried out based on the sequencing of the cytochrome-b gene, as described in a previous study [1].

**RNA isolation**

Approximately, 30 mg of salivary gland tissues were individually homogenized (TissueRuptor, Qiagen) before
RNA extraction. Total RNA was extracted using the RNeasy® Plus kit (Qiagen, German) according to the manufacturer’s protocol. The extracted RNA was stored at −80 °C until further use. The integrity and purity of the extracted RNA were checked using a spectrophotometer (Eppendorf, Germany) and a bioanalyzer (Agilent, US).

**Sequencing and data analysis**

The sequencing was performed on an Illumina HiSeq 2000 system (100 bp pair-end read) using Illumina TruSeq Cluster V3 flow cells, and a TruSeq SBS Kit V3 (Illumina, US) according to the manufacturer’s specifications. Each sample was sequenced individually to generate individual sequencing reads. Quality control of the RNA-Seq data was performed using FastQC. A total of 5 parameters were checked during quality control analysis, namely sequence quality scores (Phred Score), base sequence content, base and sequence GC content, sequence length distribution, sequence duplication level, and kmer content. The fastQ file adapters were removed using Cutadapt [19] and reads with a Phred score below 20 were removed using the FASTX-Toolkit. The cleaned reads were deposited in the NCBI Sequence Read Archive (SRA) under accession number SRX1640191-SRX1640203 and SRX1640246-SRX1640253. Raw reads were assembled using SOAPdenovo-Trans software with the default parameters [20]. Prior to the assembly, raw reads generated from each library were concatenated, and redundant reads were removed using Fastq-MCF software in each sampling group [21]. A total of 4 reference transcripts for *Aerodramus fuciphagus* from the cave (AFC), *Aerodramus fuciphagus* from the man-made house (AFM), *Aerodramus maximus* from the cave (AMC), and *Apus affinis* (AA) were generated. Core Eukaryotic Genes Mapping Approach (CEGMA) analysis was performed to validate the completeness and accuracy of the transcript assembly [22] (Additional file 3: Figure S1).

**Gene, functional and pathway annotation**

Raw data from the four groups were concatenated to generate a non-redundant reference assembly using the CAP3 program for gene annotation and differential expression analysis [23]. A non-redundant reference transcript was used for gene prediction by the AUGUSTUS program (University of Greifswald, Germany) [24]. Gene predictions were retrieved in a general feature file (GFF) format and used for downstream annotation and expression analyses. Functional annotation was performed using the Blast2GO tool [25]. All transcript sequences were aligned to sequences, in order of priority, using BLASTX (cut-off E-value of <0.0005), in the following databases: NR (NCBI database), InterPro database, UniProt database, Pfam database, Kyoto Encyclopedia of Genes and Genome database (KEGG), and WikiPathways. The highest ranks in the blast results were used to identify the coding region of the transcript sequences. Gene Ontology (GO) annotation was performed using Blast2Go software v2.5.0. The BLASTX cut-off value was set to 1 × 10−5. Each annotated sequence was assigned to detailed GO terms. Proteins were further classified into three main GO categories, namely molecular function, cellular component, and biological process, using Web Gene Ontology Annotation Plot (WEGO) (BGI Americas, UK) [26]. KEGG pathway analysis was conducted using the KEGG Automatic Annotation Server (KASS) available at http://www.genome.jp/tools/kaas/. GO is a hierarchical description of gene function that classifies genes based on known or predicted functions in model organisms [27]. GO terms allow a general assessment of the annotation and functional roles of an individual gene. However, GO annotation that is based on sequence similarity alone may result in an over-assignment of GO terms to genes that are functionally diverged [28]. Thus, this study investigated the large-scale pathway patterns revealed by these functional annotations.

**Gene expression analysis**

Raw reads in the fastQ format from Illumina sequencing were mapped against the reference assembly generated by the TopHat v2.0.10 software [29] to provide a uniform template to calculate gene expression levels in various groups using the cuffdiff application of the software. To compare expression analysis among samples, the BAM files from TopHat and the predicted GFF files from AUGUSTUS were used as inputs for cuffdiff v2.1.1 software, and then normalized with the FPKM (fragments per kb per million reads) method to identify the differentially expressed genes among the samples [16]. The following comparisons of gene expression profiles were made: a) EBN-producing swiftlets and non-EBN-producing swiftlets, (b) interspecies comparisons among EBN-producing swiftlets from the cave (AFC and AMC), and lastly, (c) habitat comparisons between AFC and AFM. Furthermore, analyses were also performed to determine the relationships between gender and expression profiles of EBN-producing swiftlets. To evaluate the significantly regulated pathways, the list of differentially expressed genes with a log2 fold change >2 and a p < 0.05 was keyed into the Enricher tool, and the gene set was annotated for possibly affected pathways using the KEGG, Reactome, Biocarta and PANTHER databases.

**Validation of expression profile**

A total of 21 well-developed salivary gland samples with five samples previously analyzed by NGS transcriptomic
analysis were selected for the gene expression validation study. A total of 36 genes of potential importance in salivary gland development and EBN composition were selected for validation from the total of 14,835 genes detected by NGS expression profiles using the NanoString nCounter® gene expression assay. Total RNA was extracted using the RNeasy® Plus kit (Qiagen, Germany) according to the standard protocol. The array contained four internal positive control probes and five negative controls for data normalization. Seven housekeeping genes were selected based on expression profile analysis to determine the appropriate thresholds for measuring significant hybridization signals over the background. After pre-processing and normalization, log2-transformed RNA-Seq and nCounter data from 36 genes were compared, and the concordance of the data from both technologies were evaluated.

### Results

#### Overview of the swiftlets’ transcriptome landscape

After removing the amplification adapters and the ambiguous reads, we generated a total of 17.4 million, 17.7 million, 19.6 million, and 15.1 million clean reads for AFC, AFM, AMC, and AA, respectively (Table 1). The results showed that 99% of the sequenced reads passed quality control with Phred scores of more than 20, indicating that the error rate of the sequencing process was less than 0.01%. SOAPdenovo-Trans software was used to perform de novo assembly and generated 47,596 (AFC), 55,331 (AFM), 40,330 (AMC), and 44,145 (AA) contigs; the average contig size exceeded 200 bp for all four libraries (Table 3). More than 65% of the reads could be mapped back to the reference transcripts in all four experimental groups. Furthermore, for all four groups, more than 96% of the genes were longer than 100 bp, more than 20% of the genes were 500 to 1000 bp, 11% of the genes were longer than 1000 bp, and 0.01% of the genes for AFC, AFM, and AA exceeded 10 kbp (Table 1). CEGMA analysis was used to analyze the completeness of the assembled transcript sequences. A total of 142 (57.66%) Core Eukaryotic Genes (CEGs) were identified as full-length, whereas a total of 185 (74.6%) core genes were retrieved as partial CEGs (Additional file 4: Table S3).

#### Functional annotation and classification of the assembled Unigenes

The generated, non-redundant, reference transcript was used for gene predictions by AUGUSTUS and a total of 14,835 genes were predicted. The functional annotation of genes was first performed using the BLASTx function in the Blast2GO tool against the non-redundant (NR) protein database from NCBI and the UniProt database with an E-value of <0.0005. Out of 14,835 high-quality genes, 14,107 genes (95.51%) had a minimum of one significant match to an existing gene model in the Blast2Go database and mapped to 7028 different protein IDs in the UniProt database. Pfam was used to further improve the functional annotation process. In total, 13,493 out of 14,835 genes (97.81%) were mapped to the known functional domain in the Pfam database using multiple sequence alignments and hidden Markov models (HMMs). However, 428 genes (2.89%) were not identified in any of the databases and were classified as hypothetical proteins (Figure 1a). Gene ontology analysis indicated that a total of 14,107 genes were categorized into 49 GO clusters based on three main categories, namely cellular components, molecular function, and biological processes. The molecular function category consisted of 11 GO groups, the largest group was that of binding (60.1%), and the cellular component category consisted of 15 GO groups; the largest group was that of cell/cell parts (90.6%). Lastly, the biological processes category consisted of 23 GO groups; the largest group was that of cellular process (77.4%) (Figure 1b). Further gene ontology enrichment analysis showed that more genes were involved in salivary gland morphogenesis and regulation in *Aerodramus* than in *Apus affinis* (Figure 1c).

| Sequencing parameters | Species and location |
|-----------------------|----------------------|
|                       | AFC                  | AFM                  | AMC                  | AA                   |
| Contig (n)            | 47,596               | 55,331               | 40,330               | 44,145               |
| Contig N50 size       | 252                  | 269                  | 256                  | 244                  |
| Total length of contig| 10,865,350           | 12,946,212           | 9,192,670            | 9,807,398            |
| Scaffold (n)          | 28,471               | 32,337               | 25,185               | 28,051               |
| Scaffold N50 size     | 1521                 | 1560                 | 1362                 | 1143                 |
| Maximum length of scaffold| 15,860            | 15,605               | 10,769               | 15,348               |
| Minimum length of scaffold| 100              | 100                  | 100                  | 100                  |
| Total length of scaffolds| 13,284,125        | 15,888,505           | 11,131,881           | 11,571,111           |
| GC percentage         | 46.72%               | 44.30%               | 47.94%               | 50.73%               |

**Table 1** Overview of the sequencing and assembly

**Gene expression analysis**

Data from EBN-producing swiftlets (AFC, AFM, and AMC) was compared to data from non-EBN-producing swiftlets (AA); interspecies comparisons of data from EBN-producing swiftlets from the cave (AFC and AMC), and comparisons of white-nest producing swiftlets from two different habitats (AFC and AFM) (Figure 2). In addition, the effects of gender on gene expression patterns were also studied for EBN-producing swiftlets. As expected, the greatest differential gene expression profile was observed between EBN-producing swiftlets (AFC, AFM, and AMC) and non-EBN-producing swiftlets (AA). Comparisons between the data from EBN-producing swiftlets and non-EBN-producing swiftlets
showed that a total of 2037 genes were significantly up-regulated and 1450 genes were significantly down-regulated. An interspecies comparison between AFC and AMC from similar habitats showed a fewer number of differentially regulated genes; only expressions of 124 genes were up-regulated and 118 genes were down-regulated. Lastly, expression profiles between AFC and AFM only identified 16 up-regulated genes and 32 down-regulated genes out of 14,835 annotated genes. Analysis of male and female samples indicated that AFM have the highest number of regulated genes, whereby a total of 180 genes were significantly up-regulated in male samples and 37 genes in female samples (Table 2). However, only 102 genes and 45 genes were significantly
Fig. 2 (See legend on next page.)
up-regulated in AFC and AMC samples, respectively (Table 2). Overall analysis of log2-fold change data suggested that 98% of the significantly regulated genes have less than 4 folds different.

**Pathway prediction and enrichment analysis**
A total of 14,835 annotated genes were exported to the Enricher tool for pathway enrichment analysis against several publicly available databases. The results indicated a total of 41, 75, 11, 69, and 34 pathways from KEGG, WIKI pathway, Reactome, Biocarta, and PANTHER databases, respectively. All genes that were significantly regulated with a q-value less than or equal to 0.05 and a log2-fold change of more than two were identified and subjected to pathways enrichment analysis.

Pathway enrichment analysis between EBN and non-EBN-producing swiftlets indicated that a total of 341 pathways were up-regulated in EBN-producing swiftlet salivary glands, with a majority of the pathways involved in mRNA processing (15.5%), protein export (14.4%), and squamous cell TarBase (9.1%), while 328 pathways were up-regulated in AA samples. The top 3 up-regulated pathways identified in AA were associated with translation factors (27.4%), mRNA processing (10.9%) and spliceosomes (8.8%). Further analysis showed a total of 152 and 120 up-regulated pathways detected among AFC and AMC samples, respectively. The majority of the up-regulation in AFC samples involved the PI3K-Akt signaling pathway (46%), ribosome biogenesis (12.5%), and cardiac muscle contraction (6.5%). Meanwhile, analysis of AMC samples showed that the top three up-regulated pathways were the Ras signaling pathway (38%), the MAPK signaling pathway (29%), and the proteasome pathway (20%). As expected, comparisons between AFM and AFC showed only 34 and 7 pathways that were significantly up-regulated in AFM and AFC samples, respectively. The majority of the detected pathways in AFM samples were associated with galactose metabolism (38%), endocytosis (15%), and oxidative phosphorylation (15%), while the top three up-regulated pathways in AFC samples were mitochondrial fatty acid oxidation (43%), RNA transport (29%), and proteoglycan (29%) (Table 3).

Gender based analysis showed that a total of 34 pathways were significantly up-regulated in female AFC samples, and the majority of these pathways were associated with DNA repair (61%), circadian clock (20%), and cell cycle (12%). Furthermore, 25 pathways were up-regulated in male samples, and 40% of the identified pathways were involved in mitochondrial protein synthesis. Analysis of AFM data did not indicate significant regulated pathways in female samples, whereas 15 pathways were significantly up-regulated in male samples. The majority of the detected pathways were associated with molecule transport (17.5%), protein metabolism (5.8%), and signal transduction (5.1%). Lastly, a total of 13 and 5 pathways were up-regulated in female and male samples from AMC, respectively. The majority of the up-regulated pathways from female samples were related to organelle biogenesis (23%), muscle contraction (17%), and axon guidance (9%), while the top three regulated pathways in male samples were vesicle-mediated transport (38%), extracellular matrix transport (29%), and signal transduction (23%) (Table 3).

A total of 24 genes of interest, that are associated with salivary gland development and EBN composition, were identified for further discussion. These genes are involved in salivary gland development, the epidermal growth factor receptor (EGFR) signaling pathway, the bone morphogenetic protein (BMP) signaling pathway, the mitogen-activated protein kinases (MAPK) signaling pathway, the fatty acid synthesis pathway, the N-glycan trimming in endoplasmic reticulum, and the N-glycan biosynthesis pathway (Table 4).

**Table 2** Number of up-regulated genes of male and female swiftlet salivary glands

| Gender of the swiftlets | AFC | AFM | AMC |
|------------------------|-----|-----|-----|
| Male*                  | 46  | 180 | 20  |
| Female                 | 56  | 37  | 25  |

*Samples from male swiftlets were used as baseline for differential expression analysis

**NanoString nCounter® Gene expression assay**
A total of 36 genes associated with salivary gland development and EBN compositions were analyzed using the NanoString nCounter® gene expression assay. All the analyzed genes showed a normal distribution of log2 expression values with a median value of 10.5-folds compared with the negative control. Out of these genes,
a total of 33 genes showed similar expression patterns as detected by NGS-based transcriptomic analysis, and three genes, namely PDHB, PRMT5 (both involved in amino acid metabolism), and HK1 (involved in sugar metabolism) showed contradicting results (Table 5). A Pearson correlation was calculated based on normalized, log2-transformed gene expression values from the nCounter data and the RNA-Seq FPKM data of the 36 genes. The mean correlation was 0.54 with a 95% bootstrap confidence interval.

Discussion
Characterization of the swiftlet transcriptome
With the advancements in sequencing technologies, the genomes of several avian species, including Peking ducks [30], passenger pigeons [15], and vultures [31], have been sequenced and characterized, whereby this inclusion has accelerated the research on functional genomics and improved our understanding of the genetic regulation of important traits. However, for some non-model organisms and minor species, it is not feasible to perform whole genome sequencing because of the expensive cost. Sequencing based on NGS technology has provided an opportunity to use a transcriptomic approach to study mechanisms that control certain traits in non-model organisms [32]. It is generally recognized that EBN is one of the most expensive animal-based products; it is often known as the “caviar of the East” [4]. Nevertheless, the transcriptomic profiles of swiftlets have not been established. The current study focused on de novo RNA sequencing of the salivary glands of both EBN-producing swiftlets and non-EBN-producing swiftlets.

The transcriptomic sequencing was completed using the Illumina Genome Analyzer system platform, HiSeq2000, with a 100 bp paired-end read. A paired-end sequencing protocol was selected for this study because it is critical for the assembly of short (<100 bp) reads and the de novo sequencing approach [33]. Paired-end sequencing allows the assembly of localized regions that are repetitive, as reads that derived from a specific localized copy of a repeat can often be inferred by the

### Table 3 Pathway enrichment analysis of up-regulated pathways in swiftlet salivary glands

| Species comparison | Number of pathways up-regulated | Top 3 regulated pathways |
|-------------------|---------------------------------|--------------------------|
| AA vs. Aerodramus | 328                             | - Translation factors (27.4%) |
| species           |                                 | - mRNA processing (10.9%) |
|                   |                                 | - Spliceosome (8.8%)      |
| Aerodramus species| 341                             | - mRNA processing (15.5%) |
|                   |                                 | - Protein export (14.4%)  |
|                   |                                 | - Squamous cell TarBase (9.1%) |
| AFC vs. AMC       | Number of pathways up-regulated  | Top 3 regulated pathway   |
|                   |                                 | - PI3K-Akt signalling pathway (46%) |
|                   |                                 | - Ribosome biogenesis in euakaryotes (12.5%) |
|                   |                                 | - Cardiac muscle contraction (6.5%) |
| AFC               | 152                             | - Ras signalling pathway (38%) |
|                   |                                 | - MAPK signalling pathway (29%) |
|                   |                                 | - Proteasome (20%)        |
| AFM               | 34                              | - Galactose metabolism (38%) |
|                   |                                 | - Endocytosis (15%)       |
|                   |                                 | - Oxidative phosphorylation (15%) |
| AFC vs. AMC       | Number of pathways up-regulated  | Top 3 regulated pathway   |
|                   |                                 | - Mitochondrial fatty acid oxidation (43%) |
|                   |                                 | - RNA transport (29%)     |
|                   |                                 | - Proteoglycan (29%)      |
| Gender comparison | Number of pathways up-regulated  | Top 3 regulated pathway   |
| AFC               | 34                              | - DNA repair (61%)        |
|                   |                                 | - Circadian clock (20%)   |
|                   |                                 | - Cell cycle (12%)        |
| Male              | 25                              | - Mitochondrial protein (40%) |
|                   |                                 | - Peptide hormone metabolism (23%) |
|                   |                                 | - Muscle contraction (15%) |
| AFM               | 0                               | - NA                     |
| Female            |                                 | - Transmembrane transport of small molecule (17.5%) |
|                   |                                 | - Protein metabolism (5.8%) |
|                   |                                 | - Signal transduction (5.1%) |
| AMC               | 13                              | - Organelle biogenesis and maintenance (23%) |
| Female            |                                 | - Muscle contraction (17%) |
|                   |                                 | - Axon guidance (9%)      |
| AMC               | 5                               | - Vesicle-mediated transport (38%) |
|                   |                                 | - Extracellular matrix transport (29%) |
|                   |                                 | - Signal transduction (23%) |

AFC Aerodramus fuciphagus from cave
AFM Aerodramus fuciphagus from man-made house
AMC Aerodramus maximus from cave
AA Apus affinis
NA No regulated pathways with $p$-value <0.05 was identified
| Gene function          | Gene   | AFM   | AFC   | AMC   | AA    |
|------------------------|--------|-------|-------|-------|-------|
| Salivary development   | CREB3L2| 65.608| 76.666| 51.669| 0.000 |
|                        | SEC63  | 36.723| 34.793| 50.081| 5.926 |
|                        | SEC24A | 15.763| 15.414| 19.240| 19.20 |
|                        | SAR1   | 38.916| 23.855| 16.996| 6.934 |
| AA versus Aerodramus species | |       |       |       |       |
| Gene function          | Gene   | AA    | Aerodramus spp | log2 | q-value |
| EGFR signalling pathway| ARF4   | 0.683 | 48.954 | 6.162 | 0.026 |
|                        | BCAR1  | 6.407 | 51.836 | 3.016 | 0.003 |
|                        | CAV1   | 0.695 | 50.207 | 6.174 | 0.026 |
|                        | CEBPA  | 8.626 | 49.820 | 2.529 | 0.018 |
|                        | ELF3   | 6.816 | 603.607| 6.468 | 0.001 |
|                        | EPS15  | 0.198 | 10.035 | 5.657 | 0.026 |
|                        | EPS8   | 0.754 | 13.668 | 4.178 | 0.026 |
|                        | JUN    | 0.411 | 11.504 | 4.803 | 0.021 |
|                        | PLEC   | 6.047 | 31.197 | 2.367 | 0.005 |
|                        | RABBA  | 17.934| 47.435 | 1.403 | 0.046 |
|                        | REPS1  | 6.795 | 21.529 | 1.663 | 0.040 |
|                        | ZPR1   | 1.079 | 13.242 | 3.617 | 0.001 |
|                        | VAV2   | 49.553| 239.912| 2.275 | 0.011 |
|                        | STAT3  | 2.453 | 53.308 | 4.441 | 0.001 |
| N-glycan biosynthesis  | ALG2   | 6.546 | 90.860 | 3.794 | 0.001 |
|                        | MGAT1  | 4.527 | 95.917 | 4.301 | 0.004 |
|                        | RPN1   | 66.407| 552.366| 3.056 | 0.003 |
|                        | FUT8   | 13.173| 84.076 | 2.674 | 0.003 |
|                        | RPN2   | 1.069 | 71.026 | 6.053 | 0.001 |
| AMC versus AFC         | Gene   | AMC   | AFC   | log2  | q-value |
| EGFR signalling pathway| JUND   | 4.911 | 39.541| 3.009 | 0.001 |
|                        | HDAC   | 1.357 | 580.771| 8.740 | 0.001 |
|                        | BCAR   | 7.776 | 200.736| 4.690 | 0.001 |
| N-glycan biosynthesis  | MGAT1  | 20.278| 256.555| 3.661 | 0.046 |
| Alanine and Aspartate metabolism | GOT1 | 226.741| 7.0623 | −5.012| 0.000 |
|                        | PDHB   | 894.023| 135.641| −2.721| 0.000 |
| AFM versus AFC         | Gene   | AFM   | AFC   | log2  | q-value |
| MAPK signalling pathway| HMGN1  | 23,486.800| 2685.98| −3.128| 0.001 |
| BMP signalling pathway  | TOB2   | 20.463| 0.959 | −4.414| 0.001 |
| Fatty acid synthesis   | ACADVL | 10.909| 50.803 | 2.219 | 0.001 |
| Gender comparison of AFC| Gene   | Male  | Female| log2  | q-value |
| N-glycan trimming      | EDEM2  | 3.951 | 0.359 | 3.459 | 0.008 |

AFC Aerodramus fuciphagus from cave
AFM Aerodramus fuciphagus from man-made house
AMC Aerodramus maximus from cave
AA Apus affinis
placement of their mate-pair reads in unique, flanking sequences [34]. Furthermore, with short reads, the advantages of a paired-end approach are accentuated and this approach is frequently used in several short-read assemblers, including SOAPdenovo [35] that takes advantage of the deBruijn graph calculation in its reads.
assembly [36]. Using pair-end sequencing, a total of four reference transcripts were generated in this study which showed similar assembly profiles of samples from swiftlets’ salivary glands compared with previous de novo studies of other small avian species, such as the great tit (Parus major) and the songbird (Junco hyemalis) with 40,000 to 50,000 contigs and 63% to 90% of reads assembled, respectively [37, 38]. This finding indicated that the experimental design, library preparation, sequencing depth, and the number of replicates adopted by the current study were appropriate for various biological function analyses. Individual gene expression analysis, with little or no consideration on pathway effects, often reveals limited information about functional pathways, which are the key factor to a fundamental understanding of the biological systems [39]. Therefore, the analysis of the current study does not select the gene of interest based on individual gene expression profiles, but it focuses on important functional, biological pathways that are significantly up- or down-regulated to provide a comprehensive comparison between different swiftlet species. Salivary gland-specific transcriptomic profiles were examined based on gene expression patterns and pathway enrichment analysis. As expected, the results suggest that the comparison of transcriptomic profiles between non-EBN- and EBN-producing swiftlet salivary glands showed the most significant differences, while the comparison between AFC and AFM showed the least variation. We postulated that the transcriptome variations of swiftlets’ salivary glands are mainly due to dietary behavior and differences in genetics. Dietary behavior is presumed to influence the expression profiles of the salivary glands from AFC and AFM, as demonstrated in a previous study in broiler chickens that both feeding conditions and dietary manipulation significantly affect gene expression [40]. Moreover, in mice, it has been established that both genetic background and diet composition influence salivary gland gene expression [41]. Besides genetics and dietary behaviour, it seems that gender also influenced the regulation of genes in the salivary gland of EBN-producing swiftlets. Studies have reported the occurrence of sex-biased in gene expression profiles of avian and mammalian species [42]. However, the gender of the swiftlets probably has little effect on the expression profiles of the salivary glands, as 98% of the regulated genes have log2 fold changes that are equal or lesser than four.

NanoString technology was selected as an independent verification platform to determine the extent of correlation in gene expression between different salivary glands samples as the platform allows for highly multiplexed reactions and is able to generate an accurate signal capture as a digital readout [34]. Although NanoString and NGS-based analysis are robust assays for transcriptomic studies, NGS-based, RNA-sequencing detects differentially expressed transcripts by counting the sequencing library fragments that map to the exons of each gene and divide the count for each gene by a scaling factor based on the length of the exons [43]. The NanoString nCounter system is designed to detect nucleic acids based on unique probe hybridization chemistry without the involvement of any PCR amplification step. Each target molecule of interest was identified by the color code generated by the ordered fluorescent segments present on the reporter probe. A total of 33 genes from the NanoString results showed similar expression patterns when compared to their corresponding RNA-Seq. However, three genes showed contradicting patterns during the validation analysis. In addition, a low correlation of 0.54 was also detected in the expression profiles detected by both platforms. One possible reason for this variation is the sampling strategy adopted in this study, whereby in NGS analysis, both partially developed and well-developed salivary glands samples were pooled to create the reference transcript for the differential expression analysis, whereas in the NanoString system, only well-developed individual salivary gland samples were separately analyzed. Hence, the different sample batches and preparation methods might have caused the unwanted methodological bias. A previous study has also shown a poor correlation (r2) between transcriptome and NanoString analysis on formalin fixed paraffin-embedded breast tumor samples, which ranged from 0.59 to 0.72, due to variations in sample source and poor sample quality [44].

**Genes involved in salivary gland development**

Salivary glands are remarkable organs. Despite their compact size they are able to produce the required volume of saliva via the interconnected network of secretory acini and ducts through a complex, epithelial-branching, morphogenesis process. Research on fly (Drosophila melanogaster) salivary glands identified both CREBA and PAS synthetase genes as early and late markers of salivary gland development, respectively [45, 46]. A previous study on fly salivary glands showed that CREBA is a Drosophila homolog of the CREB3 family of transcription factors which includes five different proteins in mammals, namely CREB3, CREB3L1, CREB3L2, CREB3L3, and CREB3L4. Each member of the Creb3-like family has a unique expression pattern in different organs and tissues [47]. Through RNA-Seq and gene expression analysis, differential expression analysis among different swiftlet species identified that expression of the CREB3L2 gene was significantly regulated in EBN-producing swiftlets compared with AA (Fig. 3). Studies on the salivary glands D. melanogaster and Aedes aegypti revealed that CREB3L2 and the expression of several
important cell type-specific secreted component genes (SPCGs) are involved in salivary secretion and control of the saliva flow rate [45, 46]. SPCGs are genes that encode for protein machinery that targets and/or translocates proteins in the endoplasmic reticulum, regulates vesicle transport between the endoplasmic reticulum and Golgi, and cleaves the N-terminal signal sequence [47]. Several SPCGs were up-regulated in EBN-producing swiftlets compared to AA, including \textit{SEC63}, \textit{SEC24A}, and \textit{SAR1}. The \textit{SEC63} gene is an important regulatory gene involved in the translation of the endoplasmic reticulum, while both \textit{SEC24A} and \textit{SAR1} play an important role in exocytosis from the rough endoplasmic reticulum to the Golgi apparatus [48, 49].

In addition to \textit{CREB3L2} and SPCGs, EBN-producing swiftlets also showed higher expression of the family of epidermal growth factor receptor (EGFR) genes compared with non-EBN-producing swiftlets. The EGFR family plays a vital role in mammalian development because it is involved in cell proliferation, migration, differentiation, survival, and apoptosis [50, 51] and it participates in the development and morphogenesis of salivary glands [52]. Multiple members of the epidermal growth factor (EGF) family of genes are expressed during salivary gland development, namely: \textit{EGFR} (\textit{ERBB1}), \textit{ERBB2}, \textit{ERBB3}, transforming growth factor-alpha (\textit{TGF-a}), heparin binding EGF (\textit{HBEGF}), and neuregulin (\textit{NRG}) [53]. Furthermore, mutations in \textit{EGFR}, \textit{ERBB3} and \textit{NRG1 type III} genes are also often associated with a reduction in salivary gland branching [54–56]. The relatively low expression of \textit{CREB3L2}, SPCGs, and EGF receptor family genes in AA compared with EBN-producing swiftlets might restrict the development of the salivary glands in non-EBN-producing swiftlets and further hinder their ability to produce EBN.

The bone morphogenetic protein (BMP) signaling pathway and the mitogen-activated protein kinase (MAPK) signaling pathway were up-regulated in AFM compared to AFC (Table 4). BMPs are members of the transforming growth factor-beta (\textit{TGF-\beta}) superfamily and are important during embryonic patterning and in the development of many different organ systems [57]. The BMP signaling pathway is closely associated with cell migration, proliferation, differentiation, and apoptosis. The BMP signaling pathway is important for the activation of SMAD-dependent (canonical) pathways as well as SMAD-independent (non-canonical) signaling pathways (\textit{MAPK, PI3K/AKT, PKC, Rho-GTPases}), which are usually associated with cell proliferation and survival [58, 59]. Furthermore, \textit{BMP7}-deficient mice have aberrant salivary glands morphogenesis with fewer end buds compared to wild-type controls [59–61], which suggests the involvement of BMP signaling pathway in acinar formation during salivary gland development. A previous study also suggested that multiple signaling pathways are tightly regulated in a spatiotemporal fashion during salivary gland development and morphogenesis [61].

![Fig. 3 CREB3L2 gene and cell type-specific secreted components genes (SPCGs). Radar draft representation of CREB3L2 gene and cell type-specific secreted components genes (SPCGs) expression in various swiftlet groups; NEBN: Non-edible bird's nest swift, AFM: A. fuciphagus from man-made house, AFC: A. fuciphagus from cave, AMC: A. maximus from cave.](image)
Genes involved in the composition of EBN
In Traditional Chinese Medicine (TCM), EBN is believed to promote the wellbeing of several organs and body systems, in addition to improving the immune system, strengthening bones, boosting metabolism, as well as having high antioxidant, anti-inflammatory, anti-aging, and anti-viral properties [4]. Various studies have characterized the composition and the therapeutic properties of EBN [10–12, 62–64]. However, the composition of EBN varies according to nest type, geographical factors, and harvesting period [62]. The fatty acid composition of EBN is highly correlated to environmental factors, such as the dietary pattern of the swiftlet and the availability of food at a specific location [63]. Swiftlet salivary glands are unique tissues with functions beyond the digestion of food and maintenance of the health of the oral cavity. Similarly, in certain animals for example arthropods, the salivary glands are equipped with cells that produce threads for the formation of webs, while in poisonous snakes, the salivary glands produce venomous saliva [61]. Among the major components isolated from EBN, EGF, glycoprotein, and sialic acid are used as indicators for the grading of EBN [63]. The results of the current study showed that MGA1, a gene essential for the biosynthesis of hybrid and complex N-acetyllneuraminic acid, is highly expressed in EBN-producing swiftlets compared to AA. Further analysis indicated that AFC samples have the highest expression of this gene, followed by AFM, AMC, and lastly AA samples (Table 4). This finding is in conformity with that of a previous study that reported that non-EBN (grass-nest) contained a similar composition of essential monosaccharides and EGF that is similar to EBN, but at a lower concentration [64]. The current study also suggested that white EBN produced by A. fuciphagus have the highest sialic acid and EGF contents, followed by black EBN produced by A. maximus, and lastly, grass-nest produced by non-EBN-producing swiftlets. Furthermore, gender-based analysis from the current study too indicated that EDEM2, a gene that is involved in trimming and biosynthesis of glycoproteins from the endoplasmic reticulum lumen [65], was significantly up-regulated in male AFC compared to female samples. A previous study on swiftlet nest building suggested that male swiftlets from man-made house contribute more to nest building compared to female swiftlets [66]. The current study also postulated that such phenomenon is generally because spermato-genesis is demanding less energy compared to oogenesis. Nevertheless, a future study on sex-biased gene expression profiling of swiftlet species from different habitats will be valuable to provide a comprehensive understanding of swiftlet nesting behaviour.

Generally, the quality of EBN is determined by the size, shape, origin, and color of the nest [9]. Currently, white EBN is considered the highest grade EBN and is commercialized at a premium price compared to black EBN [2]. One possible reason for this is that white EBN produced by A. fuciphagus contains higher concentrations of sialic acid and EGF [10, 62–64]. Nevertheless, the potential of black EBN should not be overlooked as our transcriptome analysis indicated that the alanine and aspartate metabolic pathways were up-regulated in salivary gland samples from AMC compared with AFC as the gene expression of both pyruvate dehydrogenase (PDHB) and glutamic-oxaloacetic transaminase (GOT1) was higher in samples from AMC compared to AFC (Table 2). PDHB is an important cofactor in amino acid biosynthesis and GOT1 is a pyridoxal phosphate-dependent enzyme that plays a role in amino acid metabolism and the urea and tricarboxylic acid cycles [67]. A further in-depth study should be carried out to determine the composition and the medicinal values of black EBN.

Conclusions
In conclusion, the transcriptomic profiling of salivary glands of different swiftlet species reveal differential expressions of candidate genes that are involved in salivary gland development and in the biosynthesis of various bioactive compounds found in EBN. Further functional characterization of these genes can provide insights into the mechanisms that regulate the production of saliva and the identification of markers that are unique to the origins of EBN.

Additional files

Additional file 1: Table S1. Sampling locations. (DOCX 23 kb)
Additional file 2: Table S2. List of salivary gland samples selected for NGS-based transcriptome profiling. (DOCX 26 kb)
Additional file 3: Figure S1. Workflow of transcriptome of swiftlets salivary glands based on NGS pipelines and downstream bioinformatics analyses. (DOCX 556 kb)
Additional file 4: Table S3. CEGMA analysis readings of assembled combine reference transcript. (DOCX 23 kb)

Abbreviations
BMP: Bone Morphogenetic Protein; CEGMA: Core Eukaryotic Genes Mapping Approach; EBN: Edible-bird’s Nest; EGF: Epidermal Growth Factor Receptor; GO: Gene Ontology; HMMs: Hidden Markov Models; KASS: Automatic Annotation Server; KEGG: Kyoto Encyclopedia of Genes and Genome; MAPK: Mitogen-activated Protein Kinases; NGS: Next-generation Sequencing; NF: Non-redundant; SRA: Sequence Read Archive; WEGO: Web Gene Ontology Annotation Plot

Acknowledgements
We gratefully acknowledge the assistance and support provided by the Sabah Biodiversity Centre, the Sabah Wildlife Department and the Department of Veterinary Service, Malaysia during the sampling of swiftlets. We also would like to thank Sengenics Malaysia for providing the next-generation sequencing and bioinformatics analysis services.
Funding
This work was fully by the Scifund Grant No: 5,450,668 under Ministry of Science, Technology & Innovation, Malaysia and Swiftlet Co Grant No: 6,371,400 under Ministry of Agriculture and Agro-Based Industry, Malaysia. The funding body has no role in the design of the study and collection, analysis, and interpretation of data, or in writing the manuscript.

Availability of data and materials
Sequencing reads with phred score more than 20 from Illumina HiSeq 2000 system were deposited in the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) under accession number: SRX1640191-SRX1640203 and SRX1640246-SRX1640253.

Authors’ contributions
LOH, AI, ZM, and OAR conceived the research. Animal study, organ collection, RNA extraction and sample handling were conducted by LOH and supervised by ZM and OAR. Sequencing data acquisition, data management and data analysis and interpretation were performed by LOH with assistance from AH in pathway enrichment analysis and overseen by OAR. The manuscript was written by LOH and revised by AI, ZM and OAR. The manuscript was finalized and submitted by LOH. All authors reviewed and approved the submitted manuscript.

Competing interests
The authors declare that they have no competing interests. Although we have used commercial software in this study supplied by the manufacturers, they had no influence on the experiments performed or on the results reported here.

Consent for publication
Not applicable.

Ethics approval
Before the sampling phase of the study, the approval for handling and sampling of swiftlet samples were obtained from Institutional Animal Care and Use Committee (IAUCUC), Faculty of Veterinary Medicine, UPM, with the reference number UPM/IAUCUC/AUP-R044/2014. In additional, authorization for sampling of swiftlets from their natural habitats at Gomantong Cave, Sandakan, Malaysia was obtained by Sabah Biodiversity Council, with the reference number JKM/BMS.1000–2/2(113).

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

Received: 15 September 2016 Accepted: 13 June 2017
Published online: 03 July 2017

References
1. Thomassen HA, Wiersma AT, de Bakker MA, de Knijff P, Hetebrij E, Povel GDE. A new phylogeny of swiftlets (Aves: Apodidae) based on cytochrome-b DNA. Mol Phylogenet Evol. 2003;20(1):86–93.
2. Hobbs LL. Problems in the harvest of edible birds’ nests in Sarawak and Borneo. Aust J Basic Appl Sci. 2012;6(1):301–5.
3. Ma F, Liu D. Sketch of the edible bird’s nest and its important bioactivities. Food Res Int. 2012;48(2):559–67.
4. Norhayati MK, Azman O, Nazainoon WM. Preliminary study of the nutritional content of Malaysian edible bird’s nest. Malays J Nutr. 2010;16(3):389–96.
5. Yida Z, Imam MU, Ismail M. In vitro bioassay and antioxidant properties of edible bird’s nest following simulated human gastro-intestinal digestion. BMC Complement Altern Med. 2014;14(1):468.
6. Hillier LW, Miller W, Bimey E, Warren W, Hardison RC, Ponting CP, et al. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature. 2004;432(7018):695–716.
7. Reed KM, Chaves LD, Hall MK, Knutson TP, Harry DE. A comparative genetic map of the turkey genome. Cytogeten Genome Res. 2005;111(2):118–27.
8. Kan XZ, Li XF, Zhang LQ, Chen L, Qian CJ, Zhang XW, et al. Characterization of the complete mitochondrial genome of the rock pigeon, Columba Livia (Columbiformes: Columbidae). Genet Mol Res. 2010;9:1254–9.
9. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nature Rev. Genet. 2009;10(1):57–63.
10. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature Biotechnol. 2011;29(6):544–52.
11. Glenn TC. Field guide to next-generation DNA sequencers. Mol Ecol Resour. 2011;11(5):759–69.
12. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBNet Journal. 2011;17(1):10.
13. Xie Y, Wu G, Tang J, Luo R, Patterson J, Liu S, et al. SOAPdenovo-trans: de novo transcriptome assembly with short RNA-Seq data. Bioinformatics. 2014;30(12):1660–6.
14. Aromanis E. Command-line tools for processing biological sequencing data, ea-utils. Expression analysis. Durham. Available online at: https://expressionanalysis.github.io/ea-utils/. 2011. Accessed 17 Jul 2015.
15. Parra G, Bradnam K, Korf I. CEGAV: a pipeline to accurately annotate core genes in eukaryotic genomes. Bioinformatics. 2007;23(9):1061–7.
16. Huang X, Madan A. CAP3: a DNA sequence assembly program. Genome Res. 1999;9(9):688–77.
17. Sommerfeld D, Lingner T, Stanke M, Morgenstern B, Richter H. AUGUSTUS at MedGRID: adaption of a bioinformatics application to grid computing for efficient genome analysis. Future Gener Comp Sys. 2009;25(3):337–45.
18. Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics. 2005;21(18):3674–6.
19. Ye J, Fang L, Zhang H, Zhang Y, Chen J, Zhang Z, et al. WEGO: a web tool for plotting GO annotations. Nucleic Acids Res. 2006;34Suppl 3:W378–79.
20. Bottstein D, Cherry JM, Ashburner M, Ball CA, Blake JA, Butler H, et al. Gene ontology: tool for the unification of biology. Nature Genet. 2000;25(1):25–9.
21. Devos D, Valencia A. Practical limits of function prediction. Proteins: Struct Funct Bioinf. 2000;41(1):98–107.
22. Trapnell C, Pachter L, Salberg KL. TopHat: discovering splice junctions with RNA-Seq Bioinformatics. 2009;25(9):105–11.
23. Rao M, Morisson M, Faraut T, Bardes S, Feve K, Labarte E, et al. A duck RH panel and its potential for assisting NGS genome assembly. BMC Genomics. 2012;13:513.
24. Chung O, Jin S, Cho YS, Lim J, Kim H, Jho S, et al. The first whole genome and transcriptome of the cinereous vulture reveals adaptation in the gastric and immune defense systems and possible convergent evolution between the old and new world vultures. Genome Biol. 2015;16(1):1–11.
25. Meyer E, Agyamova GV, Wang S, Buchanan-Carter J, Abrego D, Colbourne JK, et al. Sequencing and de novo analysis of a coral larval transcriptome using 454 GSFlx. BMC Genomics. 2009;10(1):29.
26. Chapman MA, Lawrence MS, Keats JJ, Cibulskis K, Schinzel AC, et al. Initial genome sequencing and analysis of multiple myeloma. Nature. 2011;471(7339):467–72.
27. Chaisson MJ, Brinza D, Pevzner PA. De novo fragment assembly with short mate-paired reads: does the read length matter? Genome Res. 2009;19(2):336–46.
28. Looi CK, Cranbrook GGH, Zoologist GB, Zoologist GB. Swiftlets of Borneo: the old and new world swiftlets (Aves: Apodidae)? A test using cytochrome b mitochondrial DNA. Mol Phylogenet Evol. 2003;29(1):86–93.
29. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics. 2009;25(9):105–11.
30. Aromanis E. Command-line tools for processing biological sequencing data, ea-utils. Expression analysis. Durham. Available online at: https://expressionanalysis.github.io/ea-utils/. 2011. Accessed 17 Jul 2015.
31. Parra G, Bradnam K, Korf I. CEGAV: a pipeline to accurately annotate core genes in eukaryotic genomes. Bioinformatics. 2007;23(9):1061–7.
32. Sommerfeld D, Lingner T, Stanke M, Morgenstern B, Richter H. AUGUSTUS at MedGRID: adaption of a bioinformatics application to grid computing for efficient genome analysis. Future Gener Comp Sys. 2009;25(3):337–45.
33. Conesa A, Gótz S, García-Gómez JM, Terol J, Talón M, Robles M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics. 2005;21(18):3674–6.
34. Ye J, Fang L, Zhang H, Zhang Y, Chen J, Zhang Z, et al. WEGO: a web tool for plotting GO annotations. Nucleic Acids Res. 2006;34Suppl 3:W378–79.
35. Bottstein D, Cherry JM, Ashburner M, Ball CA, Blake JA, Butler H, et al. Gene ontology: tool for the unification of biology. Nature Genet. 2000;25(1):25–9.
36. Devos D, Valencia A. Practical limits of function prediction. Proteins: Struct Funct Bioinf. 2000;41(1):98–107.
37. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq Bioinformatics. 2009;25(9):105–11.
38. Rao M, Morisson M, Faraut T, Bardes S, Feve K, Labarte E, et al. A duck RH panel and its potential for assisting NGS genome assembly. BMC Genomics. 2012;13:513.
39. Chung O, Jin S, Cho YS, Lim J, Kim H, Jho S, et al. The first whole genome and transcriptome of the cinereous vulture reveals adaptation in the gastric and immune defense systems and possible convergent evolution between the old and new world vultures. Genome Biol. 2015;16(1):1–11.
40. Meyer E, Agyamova GV, Wang S, Buchanan-Carter J, Abrego D, Colbourne JK, et al. Sequencing and de novo analysis of a coral larval transcriptome using 454 GSFlx. BMC Genomics. 2009;10(1):29.
36. Pevzner PA, Tang H, Waterman MS. An Eulerian path approach to DNA fragment assembly. Proc Natl Acad Sci. 2001;98(17):9748–53.

37. Santure AW, Gratten J, Mossman JA, Sheldon BC, Slate J. Characterisation of the transcriptome of a wild great tit Parus major population by next generation sequencing. BMC Genomics. 2011;12(1).

38. Peterson MR, Whittaker DJ, Ambreth S, Sureshchandra S, Buechlein A, Podcheli A, et al. De novo transcriptome sequencing in a songbird, the dark-eyed junco (Junco Hyemalis): genomic tools for an ecological model system. BMC Genomics. 2012;13(1).

39. Tomofr J, Lu I, Kepler TB. Pathway level analysis of gene expression using singular value decomposition. BMC Bioinformatics. 2005;6(1).

40. Richards MP, Poch SM, Coon CN, Rosebrook RW, Ashwell CM, McMurtry JP. Feed restriction significantly alters lipogenic gene expression in broiler breeder chickens. J Nutr. 2003;133(3):707–15.

41. Simon J, DiCarlo LM, Kruger C, Johnson WD, Kappen C, Richards BK. Gene expression in salivary glands: effects of diet and mouse chromosome 17 locus regulating macronutrient intake. Physiol Rep. 2015;3(2):e12311.

42. Ellegren H, Parsch J. The evolution of sex-biased genes and sex-biased gene expression. Nature Rev Genet. 2007;8(9):689–98.

43. Quil MA, Smith M, Coupland P, Otto TD, Harris SR, Connor TR, et al. A tale of three next generation sequencing platforms: comparison of ion torrent, Pacific biosciences and Illumina MiSeq sequencers. BMC Genomics. 2012;13(1).

44. Norton N, Perez EA, Asmann YW, Carr JM, Necela BM, Kachergus JM, et al. Analysis of gene expression and copy number variation in breast tumors using both sequencing and hybridization-based platforms. Cancer Res. 2013;73:2035.

45. Andrew DJ, Henderson KD, Seihaib P. Salivary gland development in Drosophilia melanogaster. Mech Dev. 2000;92(1):5–17.

46. Nguyen MP, Whittaker DJ, Le C, Sun L, Annan Z, Clemons A, et al. Functional genetic characterization of salivary gland development in Aedes Aegypti. EvoDevo. 2013;4(1).

47. Fox RM, Hanlon CD, Fox RM, Hanlon CD, Jordan AJ, Poch ME. The Crebic/Crebic-like transcription factors are major and direct regulators of secretory capacity. J Cell Biol. 2010;191(3):479–92.

48. Murakami T, Saito A, Hino SI, Kondo S, Kanemoto S, Chihara K, et al. Signalling mediated by the endoplasmic reticulum stress transducer OASIS is involved in bone formation. Nat Cell Biol. 2009;11(10):1205–11.

49. Saito A, Hino SI, Murakami T, Kanemoto S, Kondo S, Saitoh M, et al. Regulation of endoplasmic reticulum stress response by a BBF2H7-mediated Sec23a pathway is essential for chondrogenesis. Nat Cell Biol. 2008;10(11):1197–204.

50. Wieduwilt MJ, Moasser MM. The epidermal growth factor receptor family: biology driving targeted therapeutics. Cell Mol Life Sci. 2008;65(10):1565–84.

51. Miyazaki Y, Nakanishi Y, Hieda Y. Tissue interaction mediated by neuregulin-1 (Nrg1) regulates epithelial morphogenesis of mouse embryonic submandibular gland. Dev Dynam. 2004;230(4):591–601.

52. Nedevtski PI, Emmerson E, Finley JK, Ettinger A, Cruz-Pacheco N, Prochazka J, et al. Parasympathetic innervation regulates tubulogenesis in the developing salivary gland. Dev Cell. 2014;30(4):449–62.

53. Riese DJ, Stern DF. Specificity within the EGF family/ErbB receptor family signaling network. BioEssays. 1998;20(1):141–8.

54. Jaskoll T, Melnick M. Submandibular gland morphogenesis: stage-specific expression of TGFα, TGFβ, IGF, IGF-β, TNF, and IL-6 signal transduction in normal embryonic mice and the phenotypic effects of TGF-β2, TGF-β3, and EGF null mutations. Anat Rec. 1999;256(3):252–68.

55. Häärä O, Koivisto T, Miettinen PJ. EGF-receptor regulates salivary gland branching morphogenesis by supporting proliferation and maturation of epithelial cells and survival of mesenchymal cells. Differentiation. 2005;73(3):298–306.

56. Knox SM, Lombaert IM, Reed X, Vitale-Cross L, Gutkind JS, Hoffman MP. Parasympathetic innervation maintains epithelial progenitor cells during salivary organogenesis. Science. 2010;329(5999):1645–7.

57. Vliegenthart J, Kopf J, Hiepen C, Kraus P. Recent advances in BMP receptor signaling. Cytokine Growth Factor Rev. 2009;20(5):343–55.

58. Hoffman MP, Kidder BL, Steinberg ZL, Lakhani S, Ho S, Kleinman HK, et al. Gene expression profiles of mouse submandibular gland development: FGF-R1 regulates branching morphogenesis in vitro through BMP-and FGF-dependent mechanisms. Development. 2002;129(24):5767–78.

59. Zouvelou V, Luder HU, Mitsiadis TA, Graf D. Deletion of BMP7 affects the development of bones, teeth, and other ectodermal appendages of the crocrodilian complex. J Exp Zool B Mol Dev Evol. 2009;312(4):361–74.

60. Jaskoll T, Zhou YM, Chai Y, Makarenkova HP, Collinson JM, West JD, et al. Embryonic submandibular gland morphogenesis: stage-specific protein localization of FGFs, BMPs, Pax6 and Pax9 in normal mice and abnormal SMG phenotypes in Fgr2-iiic+/−, BMP7−/− and Pax6−/− mice. Cells Tissues Organs. 2001;7(2):3–83–98.

61. Tucker AS. Milletich I. Salivary glands: development, adaptations and disease. Switzerland: Karger Medical and Scientific Publishers, 2010.

62. Tung CH, Pan JQ, Clang HM, Chou SS. Authentic determination of bird’s nests by saccharides profile. J Food Drug Anal. 2008;16(4):86–91.

63. Huda MN, Zuki AB, Atmar K, Goh YM, Suhaimi H, Hazmi AA, et al. Proximate, elemental and fatty acid analysis of pre-processed edible birds’ nest (Aerodramus Fuciphagus): a comparison between regions and type of nest. J Food Technol. 2008;6(1):39–44.

64. Yang M, Cheung SH, Li SC, Cheung HY. Establishment of a holistic and scientific protocol for the authentication and quality assurance of edible bird’s nest. Food Chem. 2014;151:271–8.

65. Olivi S, Molinari M. Glycophore folding and the role of EDEM1, EDEM2 and EDEM3 in degradation of folding-defective glycoproteins. FEBS Lett. 2007;581(19):3658–64.

66. Ramji MF, Koon LC, Rahman MA. Roosting and nest-building behaviour of the white-nest swiftlet Aerodramus Fuciphagus (thunberg)aves: apodidae in farmed colonies. Raffles Bull Zool. 2013;30(29):225–35.

67. Soo koan S, Pirola CJ. Alarline and aspartate aminotransferase and glutamine-cycling pathway: their role in pathogenesis of metabolic syndrome. World J Gastroenterol. 2012;18(29):3775–81.