



OPEN ACCESS

EDITED BY

Pierre Boudinot,
Institut National de recherche pour
l'agriculture, l'alimentation et
l'environnement (INRAE), France

REVIEWED BY

Diego Robledo,
University of Edinburgh, United Kingdom
Pierre Pontarotti,
Centre National de la Recherche
Scientifique (CNRS), France

*CORRESPONDENCE

Katharine E. Magor
✉ kmagor@ualberta.ca

RECEIVED 10 May 2023

ACCEPTED 05 July 2023

PUBLISHED 09 August 2023

CITATION

Campbell LK, Peery RM and Magor KE
(2023) Evolution and expression of
the duck TRIM gene repertoire.
Front. Immunol. 14:1220081.
doi: 10.3389/fimmu.2023.1220081

COPYRIGHT

© 2023 Campbell, Peery and Magor. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Evolution and expression of the duck TRIM gene repertoire

Lee K. Campbell^{1,2}, Rhiannon M. Peery^{1,3}
and Katharine E. Magor^{1,2*}

¹Department of Biological Sciences, University of Alberta, Edmonton, AB, Canada, ²Li Ka Shing Institute of Virology, University of Alberta, Edmonton, AB, Canada, ³Department of Biology, Carleton University, Ottawa, ON, Canada

Tripartite motif (TRIM) proteins are involved in development, innate immunity, and viral restriction. TRIM gene repertoires vary between species, likely due to diversification caused by selective pressures from pathogens; however, this has not been explored in birds. We mined a *de novo* assembled transcriptome for the TRIM gene repertoire of the domestic mallard duck (*Anas platyrhynchos*), a reservoir host of influenza A viruses. We found 57 TRIM genes in the duck, which represent all 12 subfamilies based on their C-terminal domains. Members of the C-IV subfamily with C-terminal PRY-SPRY domains are known to augment immune responses in mammals. We compared C-IV TRIM proteins between reptiles, birds, and mammals and show that many C-IV subfamily members have arisen independently in these lineages. A comparison of the MHC-linked C-IV TRIM genes reveals expansions in birds and reptiles. The TRIM25 locus with related innate receptor modifiers is adjacent to the MHC in reptile and marsupial genomes, suggesting the ancestral organization. Within the avian lineage, both the MHC and TRIM25 loci have undergone significant TRIM gene reorganizations and divergence, both hallmarks of pathogen-driven selection. To assess the expression of TRIM genes, we aligned RNA-seq reads from duck tissues. C-IV TRIMs had high relative expression in immune relevant sites such as the lung, spleen, kidney, and intestine, and low expression in immune privileged sites such as in the brain or gonads. Gene loss and gain in the evolution of the TRIM repertoire in birds suggests candidate immune genes and potential targets of viral subversion.

KEYWORDS

TRIM protein, evolution, B30.2/PRY-SPRY domains, mallard duck, gene duplication

1 Introduction

Tripartite motif (TRIM) proteins comprise a large family with important roles in development (1, 2), cell cycle (3), immunity (4–7), autophagy (3, 7), and various other intracellular functions (7, 8). Comparisons of TRIM proteins across species demonstrate a rapidly expanding repertoire during eukaryote evolution. Humans have more than 80 TRIM genes, while mice have approximately 60, zebrafish have approximately 208, worms have more than 20, and flies have more than 10 (5, 9–12). No avian TRIM repertoire has been

systematically analyzed. An early estimate identified 37 in the chicken (9), but this has not been updated with improved genomic resources. The ambiguity of TRIM gene number in each species arises from their locations on many different chromosomes, each evolving independently, responding to pressures from pathogens and undergoing species-specific expansions (13–15).

TRIM proteins are characterized by the presence of three domains: the really interesting new gene (RING) domain, one or two B-box domains, and a coiled-coil (CC) domain. These domains together are termed the RBCC motif. TRIM proteins possess E3 ubiquitin ligase activity due to their RING domain, which helps conjugate polyubiquitin to target proteins (16, 17). B-Box domains can also perform E3 ligase activity (18–20), higher order multimerization, and binding substrate proteins (21–24). The CC domain is involved in homo- or heteromeric assemblies (25, 26). The C-terminal domains are often responsible for substrate recognition. Substrates can range from intracellular proteins, pathogen proteins, or nucleic acids (27–30).

TRIM proteins are commonly categorized into 11 subfamilies defined by variable C-terminal domains, termed C-I to C-XI (5, 12, 31, 32). Sardiello and colleagues suggested that the TRIM family broadly separates into two groups (9). Group I is more conserved through evolution and is composed of all 11 subfamilies. Group II contains only TRIM proteins belonging to the C-IV subfamily, which contain C-terminal B30.2/PRY-SPRY domains. The combined B30.2/PRY-SPRY domains arose later in evolution than the SPRY domain and is often associated with immune function (33). Members of group II appear to evolve faster than members in group I (9). Marin and colleagues demonstrated that the evolution of TRIM proteins is more complex and classified TRIM proteins into nine subfamilies, based on when the TRIM proteins arose in eukaryotes (8). The C-IV subfamily dramatically expanded in vertebrates. Members of the C-IV subfamily are regulated by immune responses (6, 9, 34), and many C-IV type TRIMs have direct roles in immunity and viral restriction (4, 5). Several C-IV type TRIM genes are present in the MHC locus of humans (35), chickens (36), ducks (37), and fish (38, 39). This suggests that TRIM genes were part of the ancestral MHC and have undergone duplication events in different species.

Many TRIM proteins are modulators of innate immunity and mediators of direct viral restriction. The repertoire of the mallard duck (*Anas platyrhynchos*) is of interest, as they are the reservoir host of influenza A viruses (40, 41). A comparison of the duck repertoire to chickens (*Gallus gallus*) is of value, since chickens are an important agricultural species and an established model for vertebrate development. Functional studies of individual TRIM proteins of birds demonstrate their importance in immunity. TRIM25 is functionally characterized in chicken (42–45), duck (46, 47), and goose (48). In mammals and ducks, TRIM25 catalyzes the addition of polyubiquitin for the activation of retinoic-acid-inducible gene I (RIG-I) in antiviral signaling (46, 49). Duck TRIM29 was identified as a negative regulator of the RIG-I signaling pathway (50). A gene described as TRIM39 in chicken is predominantly expressed in the spleen, but has not yet been

functionally characterized (51). Chicken TRIM62 has been characterized (52) and shown to have antiviral activity against reticuloendotheliosis virus (53) and avian leukosis virus subgroup J (54). TRIM32, known for its antiviral activity in mammals (55–57), can restrict influenza (58) and Tembusu virus in ducks (59).

Here, we characterized the duck TRIM gene repertoire by utilizing NCBI databases and *de novo* transcriptome assembly to identify candidates. We investigated TRIM protein domain architecture and phylogenetic relationships. We compared the duck repertoire to the chicken, to look at species-specific differences. We performed phylogenetic analyses of the C-IV subfamily members of reptiles, birds, and mammals, which allowed us to designate orthologous TRIM genes in ducks. We show that ducks have 57 TRIM genes, compared to 54 in chickens. Ducks and other birds have TRIM genes specific to their respective lineages. Finally, we investigated both abundance and relative expression of these TRIM gene sequences in duck tissues. While some TRIM genes were ubiquitously expressed in duck tissues, other TRIM genes had tissue-specific expression.

2 Materials and methods

2.1 Data mining

To generate a master transcriptome of duck sequence reads, the NCBI Short Read Archive (SRA) database (<https://www.ncbi.nlm.nih.gov/sra>) was mined for projects involving domestic mallard ducks (*Anas platyrhynchos*). Wild mallard and Muscovy ducks (*Cairina moschata*) were excluded. Project numbers and individual samples included in this study are listed in [Supplementary File 1](#). In addition to assembling a master duck transcriptome, a chicken (*Gallus gallus*) transcriptome was assembled using SRA project numbers listed in [Supplementary File 1](#). SRA libraries were uploaded to the Digital Research Alliance of Canada's research computing environment (formerly Compute Canada) (<https://alliancecan.ca/en>). SRA files with ambiguous descriptions, questionable content, or failed quality checks (i.e., more than 50% of reads were unpaired, or the files were corrupt) were excluded. To generate a reference dataset of all known avian TRIM sequences, we searched the NCBI protein databases for TRIM sequences. Redundant and misannotated sequences were removed. From the curated avian TRIM list, we made two additional databases, one composed of duck TRIM proteins and one composed of chicken TRIM proteins.

Members of the C-IV TRIM protein subfamily were also mined from NCBI to infer homology among TRIM proteins. We searched for each TRIM protein by name in mammals and generated a library of TRIM proteins from human, an additional placental mammal, and marsupial. For each avian TRIM protein, we ensured that there were at least three representative species. Reptile TRIM proteins were mined for a representative lizard species and a representative turtle species. C-IV protein accession numbers can be found in [Supplementary File 2](#).

2.2 Transcriptome assembly

SRA libraries were checked for quality and adaptor content using astQC Version 0.11.9 (60). Samples used passed fastQC analysis and were between 100 and 150 bp in length. Reads were trimmed using Trimmomatic version 0.36 (61) with the following parameters: a sliding base window size of 4 (bases removed if phred score is below 20) in a 15-base window and a minimum read length of 33 bp.

A total of 216 duck libraries and 107 chicken libraries were assembled using 7 Kmer values from 25 to 85 using TRANS-ABYSS 2.0.1 (62). Individual libraries were assembled and binned by tissue type, and duplicate contigs within each tissue type were collapsed into a consensus contig using CAP3 (63) with a cutoff value of 95% identity and CD-HIT-EST version 4.8.1 (64) with a cutoff value of 97% identity. Tissue-type assemblies were collapsed into one master assembly using CAP3 assuming 95% similarity. Singlet files were merged separately to reduce loss of genes due to excessive reduction in putative duplicate contigs. Merged singlet contigs were then compared back to the master transcriptome for a final master assembly. Sequences <200 bp were pruned from the master assembly using in-house scripts published at <https://github.com/rmpeery/dataProcessing>. Two quality control measures were employed to assess the final assembly. We checked the number and average size of contigs using the abyss-fac command in ABYSS v 2.0.1 (62). To ensure that transcriptome collapsing was not impacting gene content, BUSCO version 3.0.2 (65) was used to determine common orthologous gene content. To remove duplicate copies of genes (putative orthologs remaining due to assembly strategy), we used reciprocal blast hits (RBH) and applied a leave-one-out method to remove contigs with $\geq 97\%$ similarity (perl script published at <https://github.com/rmpeery/dataProcessing>).

2.3 Duck TRIM gene identification

The master duck and chicken assemblies were each translated into all six reading frames using EMBOSS 6.6.0 (66). Avian TRIM reference proteins were compiled into a BLAST+ database using the makeblastdb command in BLAST+ version 2.7.1 (67). We ran BLASTp (BLAST+) against the assembled duck transcriptome, and all hits were parsed from the master transcriptome using in-house scripts (published at <https://github.com/rmpeery/dataProcessing>). Our newly identified, putative avian TRIM proteins were aligned to the reference TRIM proteins, and neighbor-joining (NJ) phylogenetic trees were inferred using CLUSTAL OMEGA (68). We compared full-length transcriptome contigs to the assembled list of annotated duck TRIM proteins to both confirm identity and validate the master assembly. TRIM proteins detected by the BLAST search but were not present in ducks were aligned against the avian TRIM protein database to confirm identity. Other ambiguous hits were submitted to SMART (69) to verify protein domain composition. To identify TRIM genes missing in birds or ducks, TRIM sequences from various species were aligned using COBALT (<https://www.ncbi.nlm.nih.gov/tools/>

[cobalt/](#)) and submitted to HMMER (v3.3.2) (HMMER.org) to search for these missing TRIMs in our master assembly. Hits were further analyzed using HMMER (70), UniProt (71), and SMART (69). Domain composition of amino acid sequences were verified using SMART. Duck TRIM genes found in this analysis are summarized in [Supplementary Table S1](#), and nucleotide sequences used in downstream analyses can be found in [Supplementary File 3](#). Duck TRIM amino acid sequences used in these analyses were derived from translating the nucleotide sequences and can be found in [Supplementary File 4](#). All sequences used for the duck TRIM analysis resulted from interrogating the *de novo* transcriptome assembly, unless otherwise noted in [Supplementary Table S1](#). Some sequences obtained from NCBI were used because the contigs pulled from the *de novo* transcriptome assembly appeared to be chimeric reads. Genes and proteins currently annotated in the genome and their corresponding accession numbers are compiled into [Supplementary Table S1](#).

2.4 Comparison between duck and chicken TRIM protein repertoires

The newly identified duck TRIM nucleotide and protein sequences, generated from our *de novo* transcriptome assembly, were compared to chicken TRIM protein sequences from NCBI. Sequences present in the duck but presumed missing in the chicken were submitted to BLASTn and queried against the current chicken genome (version bGalGal1.mat.broilerGRCg7b, unpublished direct release) using the BLAST+ online portal (<https://blast.ncbi.nlm.nih.gov/Blast>). We used HMMER to further interrogate our chicken transcriptome for missing TRIM genes. TRIM genes found in the chicken transcriptome are available in [Supplementary File 5](#). Annotations and identification numbers of these genes were compiled into [Supplementary Table S2](#), and amino acid sequences were downloaded or added to ([Supplementary File 5](#)). Sequences were acquired from the annotations listed in [Supplementary Table S2](#) unless otherwise noted.

2.5 Mapping to chromosomes

To find chromosomal location of all duck TRIM genes, we queried nucleotide sequences against the current NCBI duck genome [assembly ZJU1.0, published (72)] using the online version of BLASTn. Chromosome lengths were taken from the reference duck genome, and locations were assigned from the start of the TRIM gene. Any genes not in the NCBI duck genome were submitted to blast against the Ensembl rapid release domestic *Anas platyrhynchos* genomes (GCA_017639305.1, GCA_015476345.1 and GCA_017639285.1). Genes present around any TRIM hits were compared back to the NCBI Pekin duck assembly to try to infer possible genomic locations. TRIM genes were assigned chromosomal location using karyoploteR (73) in the R studio environment (74). The resulting map was edited using Adobe Illustrator for clarity.

Approximate location and composition of MHC-linked TRIM genes were compared between duck, mouse (*Mus musculus*), human (*Homo sapiens*), Tasmanian devil (*Sarcophilus harrisii*), yellow pond turtle (*Mauremys mutica*), and Eastern fence lizard (*Sceloporus undulatus*) from data on NCBI. Genomic size and location of TRIM genes were extracted from reference genomes ZJU1.0 (Pekin duck), GRCm39 (mouse), GRCh38.p14 [human, published (75)], mSarHar1.11 [Tasmanian devil, published (76)], ASM2049712v1 (yellow pond turtle), and SceUnd_v1.1 [Eastern fence lizard, published (77)]. TRIM genes were assigned chromosomal location using karyoploteR (73) in the Rstudio environment (74). The resulting maps were edited graphically using Adobe Illustrator for clarity and readability.

Sizes and locations of TRIM genes located in the MHC-B locus of mallard ducks, tufted ducks (*Aythya fuligula*), chickens, kākāpō (*Stringops habroptilus*), barn swallow (*Hirundo rustica*), and European golden plover (*Pluvialis apricaria*) were approximated from genomic data on NCBI. Data were extracted from reference genomes ZJU1.0 (Pekin duck), bAytFul2.pri [tufted duck, published (78)], bGalGal1.mat.broiler.GRCg7b (chicken), bStrHab1.2.pri [kākāpō, published (79)], bHirRus1.pri.v2 (barn swallow, <https://vertebrategenomesproject.org/>), and pPluApr1.pri (European golden plover, <https://vertebrategenomesproject.org/>). The European golden plover genome is not yet annotated, so TRIM gene approximate locations were found using the NCBI blastn against the genome and coding genes predicted using GENSCAN (<http://hollywood.mit.edu/GENSCAN.html>). The resulting distances and approximations of gene sizes were edited using Adobe Illustrator.

Sizes and locations of TRIM genes located in the TRIM25 locus of duck, chicken, kākāpō, barn swallow, Adeline penguin (*Pygoscelis adeliae*), and Eastern fence lizard were approximated from genomic data on NCBI. Data were extracted from reference genomes ZJU1.0 (mallard duck), bGalGal1.mat.broiler.GRCg7b (chicken), bStrHab1.2.pri (kākāpō), bHirRus1.pri.v2 (barn swallow), ASM69910v1 (Adeline penguin), and SceUnd_v1.1 (Eastern fence lizard). The resulting distances and estimates of gene sizes were edited using Adobe Illustrator.

2.6 Phylogenetic trees and minimum spanning networks

We made two alignments of TRIM proteins, the first aligning all duck TRIM proteins and the second aligning both duck and chicken TRIM protein sequences. A third alignment was made using C-IV subfamily members from representative species. The species used were human, a representative placental mammal, marsupial, turtle, and lizard. Avian TRIM genes not found in chicken or duck were also added to our alignments. We mined two representative reptile species, the Eastern fence lizard and the yellow pond turtle, for MHC-linked TRIM proteins, by manually scanning the chromosomes containing MHC and collecting accession numbers for TRIM genes. Accession numbers for C-IV protein sequences used can be found in [Supplementary File 2](#). TRIM proteins were aligned using the online MAFFT alignment program (80).

Alignments used the L-INS-I refinement method and were further refined by eye in Unipro UGENE (81). The best protein model was determined for each tree using ModelFinder (82) in the IQTree environment. Maximum-likelihood (ML) phylogenetic trees of these alignments were inferred using IQTree (83) generating a tree for the duck TRIM protein; a tree for the duck and chicken TRIM proteins; a tree for all C-IV TRIM proteins in mammals, reptiles, and birds; and a tree for C-IV TRIM proteins in mammals, reptiles, and birds excluding the non-orthologous expansion of MHC-linked TRIM genes in reptiles. The Ultrafast bootstrap algorithm (84) was run with 10,000 bootstrap replications on the duck only and duck and chicken TRIM ML phylogenies. The Ultrafast bootstrap algorithm was run with 5,000 bootstrap replications on the C-IV TRIM protein subfamily trees. The consensus of these replicates was visualized in FigTree v1.4.3 (<http://tree.bio.ed.ac.uk/software/figtree/>). Phylogenetic trees were then further edited using Adobe Illustrator for clarity and to add additional information.

To make minimum spanning networks (MSN) with duck TRIM proteins, the duck TRIM protein alignment (as described above) was converted to a distance matrix using msa (85) and seqinr (86) in the Rstudio (v4.0) environment. A MSN was created using the prim algorithm provided by the RGLBL r package (87) and visualized using ggplot2 (88).

2.7 Assigning names to duck TRIM genes

To resolve ambiguous names and incorrect NCBI annotations, we assigned names to TRIM genes based on homology, determined through phylogenies. As many avian TRIM genes did not have orthologs in mammals, we assigned these TRIM genes numbers starting at 200. The C-IV subfamily of TRIM proteins is expanded in vertebrates, and many of the putative TRIM proteins have redundant descriptions and names on NCBI. Many names of the C-IV TRIM proteins needed to be resolved through phylogenetic analysis comparing these proteins to TRIM proteins from other species. The new names and previously published annotations can be found in [Supplementary Table S1](#) for ducks and [Supplementary Table S2](#) for chickens.

2.8 Read mapping and differential expression of TRIM genes

Our newly identified TRIM contigs were pruned to CDS regions and were concatenated into a multi-FASTA file and used as the reference for read mapping. We aligned RNA-seq reads to the reference TRIM genes using Bowtie2 v2.3.4.3 (89). We used featureCounts to summarize, count reads, and assign features (90) from the BowTie2 mapped outputs. Reads were normalized to individual library sizes using the EdgeR (91) trimmed mean of M values (TMM) method. Reads were compared using plotMDS.DGelist in the Rstudio environment. Normalized log2 counts per million log2(CPM) were visualized using the cim function in R (92). Raw data of normalized read counts for each

sample can be found in [Supplementary File 6](#). We analyzed relative patterns of tissue distribution by setting the matrix intercept as the overall mean and using the generalized linear model quasi-likelihood test (glmQL) to determine log₂ fold change (FC) values of individual tissues when compared to the mean in EdgeR. All log₂ (FC) values were visualized using a heatmap using the `gplots` heatmap.2 function (93) in the Rstudio environment. Dendrograms were added using heatmap.2 function in `gplots` in the Rstudio environment. All data for these experiments, including log₂(FC) and false discovery rate (FDR), can be found in [Supplementary File 7](#). To determine the highest and lowest relatively expressed TRIM genes in different duck tissues, we sorted the relative expression results for each tissue by false discovery rate FDR (<0.05) and filtered out any TRIM genes above this threshold. We then sorted these data by log₂(FC) and summarized the top 5 highest and top 5 lowest relatively expressed genes.

3 Results

3.1 Ducks have 57 genes in their TRIM repertoire

To determine how many TRIM genes are present in the domestic mallard duck (*Anas platyrhynchos*), we first mined the NCBI genome and protein databases for all annotated avian TRIM proteins. From this list, we identified TRIM proteins annotated in the duck and chicken. We made a *de novo* transcriptome assembly for ducks using 216 SRA libraries. We used both BLAST+ (67) and HMMER (hmmer.org) searches of our transcriptome to identify 57 TRIM genes in the duck. We used a neighbor-joining tree to cluster putative TRIM contigs with known avian TRIM proteins to verify identity and identify outliers. A flow chart of the workflow can be found in [Supplementary Figure 1](#). Genes that were found to be one to one orthologs of human TRIM genes were identified and named to reflect that ancestry. Duck TRIM protein annotations were assembled into a table ([Supplementary Table S1](#)). All but four of the TRIM sequences that we identified were previously annotated on NCBI. TRIM28 and fibronectin III and SPRY domain containing proteins 1 (FSD1) are annotated in the rapid release Ensembl duck genomes (TRIM28 in GCA_017639305.1 and FSD1 in both GCA_017639305.1 and GCA_017639285.1). TRIM39 and RING finger protein 39 (RNF39) are not yet annotated in any duck genome.

We included 11 genes classified as TRIM-like in this analysis, in that they do not have the traditional “tripartite motif” domain structure that the TRIM family is named for. Some of these, such as RNF207, FSD1 and FSD2, B-box and SPRY domain containing (BSPRY), and NHL repeat containing E3 ubiquitin protein ligase 1 (NHLRC1) are ancestral TRIM genes, which have lost domains during speciation events (8). RNF135, with the protein referred to as RIPILET, contains a RING, CC, and PRY-SPRY domains, and is lacking B-box domains. These TRIM-like genes will be included in our analyses of the duck TRIM repertoire. Sequences discovered to be butyrophilin (BTN) proteins were omitted, although we

acknowledge that many BTN proteins are highly similar in C-terminal domain composition to C-IV subfamily TRIM proteins, and thus, TRIM proteins and BTN proteins may share evolutionary ancestry.

Many TRIM genes identified did not appear to be direct orthologs to annotated mammalian TRIM genes and were renamed. Several non-orthologous TRIM genes are located in the MHC region of the duck. Additionally, NCBI annotations suggested ancestry to mammalian TRIM genes that our downstream analysis determined to be inaccurate. We renamed these genes TRIM200–213 to reflect their distinct sequences. For example, two genes (NCBI gene ID 101805457 and 101804875) were both described as TRIM39-Like on NCBI. Gene 101805457 was previously annotated as TRIM39.2 by Blaine and colleagues (37). Gene 101804875 was described as TRIM39-Like in ducks (NCBI) and TRIM39 in chickens (51). However, neither of these genes are an ortholog of TRIM39, so we amended the names to TRIM202 and TRIM212 (for gene 101805457 and 101804875, respectively). Our transcriptome interrogation did find a direct TRIM39 ortholog in the duck that is not present or annotated in Ensembl or NCBI genomes. Likewise, Gene ID 119718713 was annotated as RNF39-like in the NCBI duck genome. This gene does not appear to be orthologous to mammalian RNF39, so we renamed it TRIM200. Ducks do have an ortholog of RNF39; however, this gene is only present in the duck transcriptome and not in genomic resources. [Supplementary Table S1](#) indicates the duck TRIM genes with direct mammalian orthologs and the genes without clear mammalian orthologs that we renamed, and we will use these assigned names throughout.

3.2 The duck TRIM gene repertoire spans 21 chromosomes

To determine chromosomal locations of the 57 domestic mallard duck TRIM genes, we interrogated the Pekin duck genome (assembly ZJU1.0) using our newly generated TRIM genes. We assigned chromosomal locations to 54 of 57 TRIM genes, on 21 different duck chromosomes ([Figure 1](#)). Most duck TRIM genes are present in unique locations throughout the genome. However, two interesting clusters of genes are found on chromosomes 17 and 19. Ducks also have a duplication of promyelocytic leukemia protein (TRIM19/PML), resulting in two similar genes on chromosome 11 in opposite orientation.

A cluster of TRIM genes are found in the duck MHC region on chromosome 17 ([Figure 1](#), inset). We placed 10 TRIM or TRIM-like genes on chromosome 17, with nine of these found in the duck MHC region. All MHC-linked TRIM or TRIM-like proteins have C-terminal PRY-SPRY domains ([Supplementary Table S1](#)). We previously compared this region in ducks to the syntenic region in chickens and turkeys (37). The newest assembly of the Pekin duck genome allows TRIM207, another of the MHC-linked TRIMs, to be placed in this region. This gene was defined as a BTN in chicken (94); however, our analysis suggests that it is a TRIM-like gene. These data support the expansion of a group of related TRIM genes in the MHC region of birds.

A smaller cluster of TRIM genes is located on chromosome 19. The TRIM-like gene RNF135 is located on between TRIM25 and TRIM65. We will refer to this location as the TRIM25 locus. This region also includes a related gene, TRIM47. Another nucleotide sequence distantly related to TRIM25 is found on chromosome 14, which we have named TRIM211. However, the duck TRIM211 gene had premature stop codons in the sequence, suggesting that it may be a pseudogene. Indeed, its annotation on NCBI includes it in the 3' untranslated region of a separate gene, described as *PPARGC1B* (GeneID: 101790707). Whether this is due to mistakes in genome assembly/sequencing or gene fusion resulting in loss of function is unknown.

We identified TRIM28, TRIM39, RNF39R, and FSD1 within the duck transcriptome data but were unable to map them to chromosomes on the NCBI duck genome. TRIM28 is present in one and FSD1 is present in two of the Ensembl rapid release genomes (GCA_017639305.1 and GCA_017639285.1); however, neither of these genomes are chromosome level assemblies. In the Ensembl rapid release genomes, FSD1 is located in between the genes SH3GL1 and YJU1. Both these genes appear on chromosome 29 in the NCBI duck genome, suggesting that FSD1 is located on chromosome 29, and this chromosome is currently misassembled in the NCBI Pekin duck genome. We were unable to find TRIM39 or RNF39 in either the NCBI or Ensembl rapid release duck genomes.

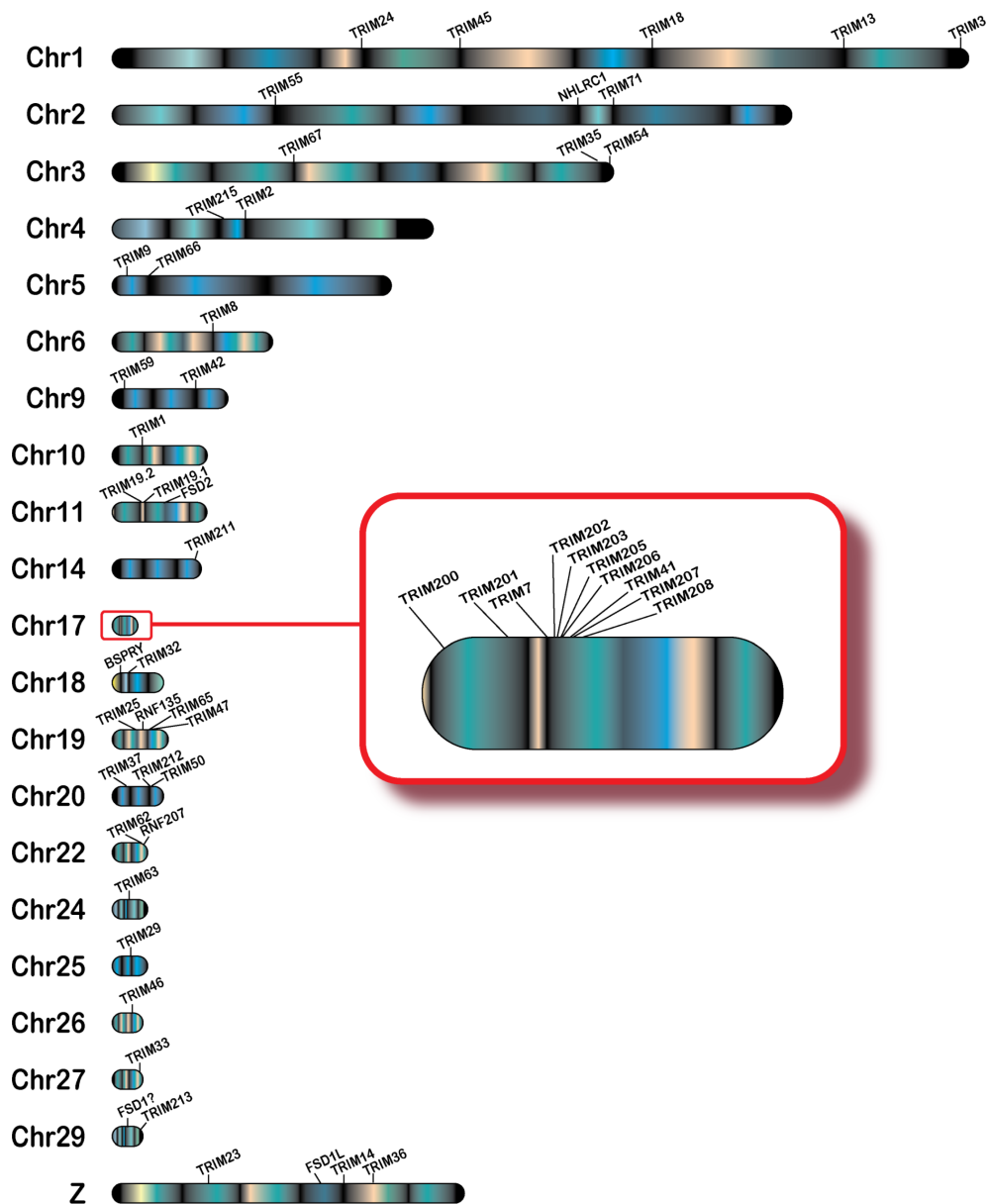


FIGURE 1
 Genomic locations of TRIM or TRIM-like genes in the duck. TRIM and TRIM-like genes were submitted to NCBI blast, and locations in the duck genome were mapped using karyoplotR in the R studio environment. Chromosome 17 is magnified to allow for visualization of the expansion of the TRIM genes in the duck MHC (B) locus. The gene FSD1 is followed by a "?" to reflect a presumed location due to data available in other genomic resources. The cluster of related genes on TRIM25 will be referred to as the "TRIM25 locus".

The current NCBI Pekin duck genome assembly has 33 chromosomes in addition to the sex chromosomes (72), and karyotyping shows 2N=40 for ducks (72, 95), suggesting that many chromosomes may still be unassembled.

3.3 Chicken and duck TRIM homologues split into two distinct clades

To compare the chicken and duck TRIM gene families, we searched for chicken orthologs for our 57 duck TRIM protein sequences using NCBI BLASTp. This search yielded 52 chicken TRIM or TRIM-like proteins (Supplementary Table S2). We created a *de novo* transcriptome assembly for chickens using 107 SRA RNA-seq libraries, and our interrogation of the transcriptome found RNF39R, TRIM39, and TRIM46, which were not present in the chicken genomes on NCBI or Ensembl rapid release. While many chicken TRIM proteins shared high percent identities to duck TRIM proteins (>90% identity), there were some exceptions. The TRIM19 paralogs were divergent at 75% and 55% identity, for TRIM19.1 and TRIM19.2, respectively. TRIM25 (71% identity), TRIM47 (81% identity), and TRIM65 (69% identity) were also divergent. Many of the MHC-linked TRIM proteins have also diverged between ducks and chickens, except TRIM201, which shares 97% amino acid identity. Both TRIM39 and RNF39R are significantly different between duck and chicken with 60.08% and 48.92% amino acid identity, respectively. However, the latter genes are only present in the duck or chicken transcriptomes and not yet confirmed in their genomes, so misassembly could contribute to sequence differences.

To compare the relationships of TRIM proteins within the duck and chicken repertoires, we aligned all duck and chicken amino acid sequences, then built an ML tree using 10,000 Ultrafast bootstrap replications. The sequences divide into two major clades (clades A and B) (Figure 2). In clade A, there are two subclades (clades C and D), clade D containing C-I and C-II TRIM proteins and clade C containing C-IV proteins. C-I subfamily members first arose in animals (8) and are defined as having both COS and FN3 domains and only a SPRY or both PRY and SPRY C-terminal domains. SPRY domains are present in animals, fungi, and plants, while PRY domains arose during vertebrate evolution (33). Some members of the C-I subfamily have both PRY-SPRY domains, while the more ancient members have only SPRY domains. TRIM9 is present in animals, including invertebrates, and has a C-terminal SPRY domain but not PRY domain in shrimp (96, 97), human (98), and ducks (Supplementary Table S1). The C-IV TRIM proteins, which expanded in vertebrates (8, 9) and have both PRY and SPRY C-terminal domains, split into two distinct subclades (clades E and F), clade E containing the MHC-linked TRIM proteins and clade F containing genes which cluster with members of the TRIM25 locus, to be referred to as the "TRIM25 cluster". While this suggests that the C-IV subfamily diverged from a common ancestor of the C-I subfamily, the bootstrap values are too low to support this hypothesis. In clade B, most duck proteins have direct orthologs with chicken on short branches. The duck and chicken TRIM19/

PML paralogs (TRIM19.1 and TRIM19.2) appear distant from their orthologous counterparts. This could be due to significant divergence in sequences or misassembly of these sequences. Many bootstrap values were too low to resolve distant ancestry of these TRIM proteins, for example the common ancestors of clades A and B, suggesting that further analysis using more taxa is needed.

We found some TRIM proteins only in chicken or only in duck. The TRIM-like gene RNF135 has previously been described as absent in chicken (99), quail (100), penguins, falcons, and petrels (101). We also did not find RNF135 in the chicken transcriptome. Two TRIM proteins found in chicken were without mammalian orthologs, so we named them TRIM210 and TRIM211. Phylogenetic analysis places TRIM210 and TRIM211 within clade F with the TRIM25 cluster, a group including RNF135, TRIM14, 25, and 65, known for RIG-I-like receptors (RLR) pathway modification (102–104). Ducks are missing TRIM210, and TRIM211 is likely a pseudogene, while the genes are intact and adjacent in the chicken genome. A closer inspection of the syntenic region on duck chromosome 14 revealed a deletion in that section. The deletion could be due to genomic misassembly; however, HMMER searches of our duck transcriptome did not find any contigs matching TRIM210. Duck and chicken TRIM29 also clusters within the clade containing the RLR pathway modifiers.

There were four TRIM genes that we did not find in chickens that we found in ducks: TRIM200, RNF135, TRIM213, and TRIM206. We did not find these genes in our chicken transcriptome. We previously noted TRIM206 to be absent in chickens and turkeys (formerly named TRIM27L) (37). We believe that TRIM213 is missing due to deletions in the chicken, from our analysis of the syntenic regions of the chromosome.

3.4 Subfamily classification of human, duck, and chicken TRIM proteins

To compare the duck TRIM repertoire to human and chicken, we grouped the subfamilies of TRIMs by the standard C-terminal domain nomenclature (5, 31, 32). Many TRIM proteins are shared between humans, chickens, and ducks, but there are some repertoire differences between species (Figure 3). All TRIM subfamilies are represented in the duck and chicken TRIM proteins, with most groups containing clear orthologs. Domain composition from duck TRIM proteins was analyzed using SMART, and missing domains are indicated by the color of the protein name. Additional information on duck TRIM domain composition can be found in Supplementary Table S1.

The C-IV subfamilies of birds and humans not only share presumed orthologs but also have many unique proteins. Humans have a cluster of MHC-linked TRIMs not present in birds including TRIM31, 40, 15, 26, and 38 (105). Human C-IV family TRIM genes, including TRIM5, 6, 22, and 34, appears absent from avian genomes. Humans and ducks have RNF135, which belongs in the C-IV subfamily, and it is missing in chickens. Chickens, however, have additional members of the TRIM25 lineage as they have TRIM210 and TRIM211.

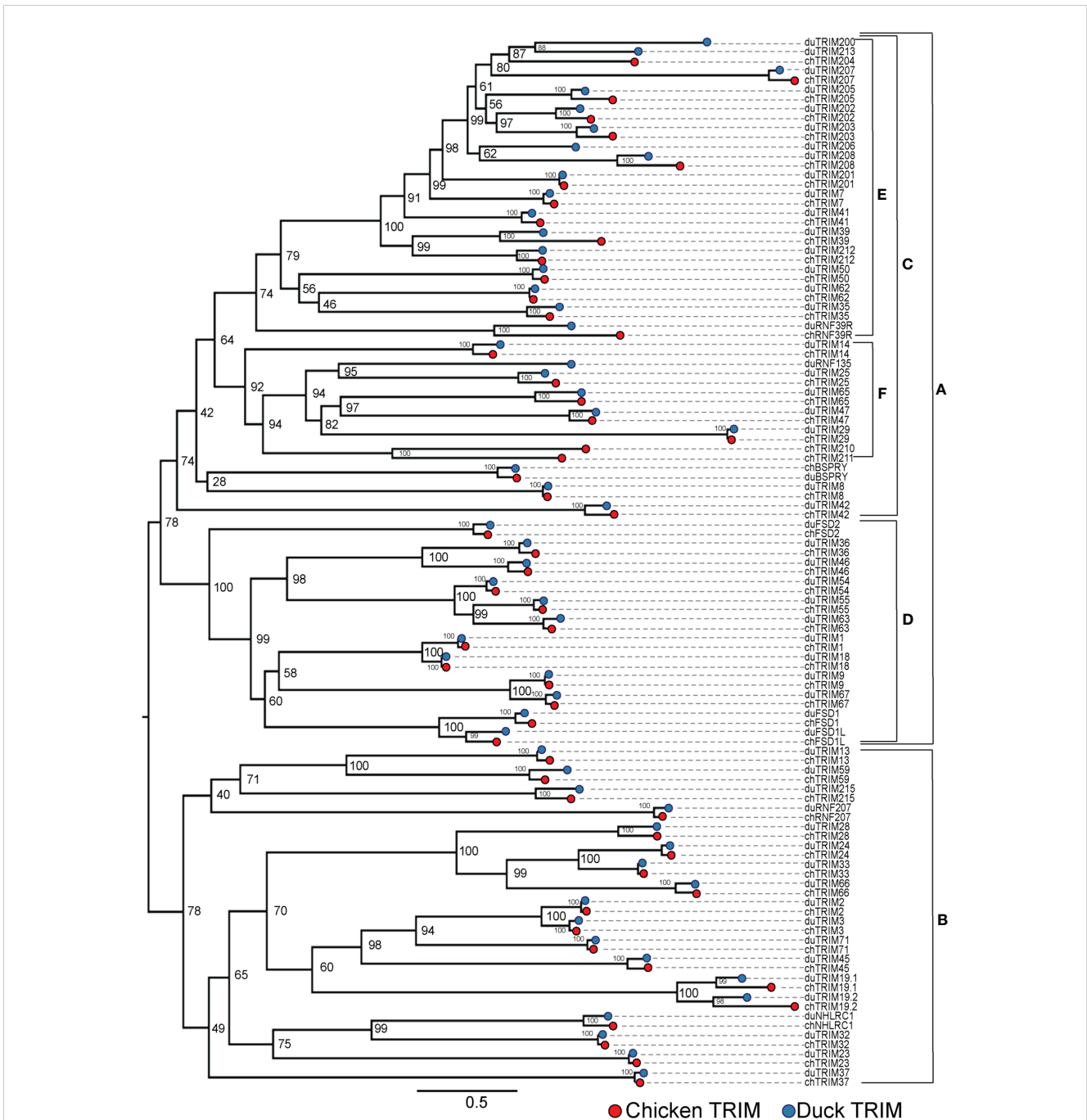


FIGURE 2
 Most duck TRIM proteins have chicken orthologs. Phylogenetic relationships between the duck and chicken TRIM protein sequences were investigated using maximum likelihood (ML) trees with 10,000 Ultrafast bootstrap replications. External nodes were color-coded to indicate species of origin for each TRIM protein. Clades A and B represent the two major subclades. Clade C represents the C-IV subfamily. Clade D represents the C-I and C-II subfamily. Clade E includes the MHC-linked TRIM proteins. Clade F represents genes present in the TRIM25 locus and related genes located elsewhere, referred to as the "TRIM25 cluster". Initial ML tree was referred using IQTree and visualized using FigTree. The resulting tree was edited for clarity using Adobe Illustrator.

Several members of the C-V family with an uncharacterized C-terminal domain are not present in birds; however, both duck and chickens have two paralogs of TRIM19, co-orthologs to mammalian TRIM19. Ducks and chickens have an additional member of the C-XI family, without mammalian homologues, thus named TRIM215. The C-XI subfamily is classified by a C-terminal transmembrane

domain; however, both duck TRIM215 and TRIM13 appear to be missing this domain (Supplementary Table S1). TRIM215 appears distantly related to TRIM59 and TRIM13. The group C-II orthologs are missing the C-terminal domains for which they are classified in humans. Duck TRIM54, 55, and 63 appear to be missing the COS and acid domains that their human orthologous counterparts have.

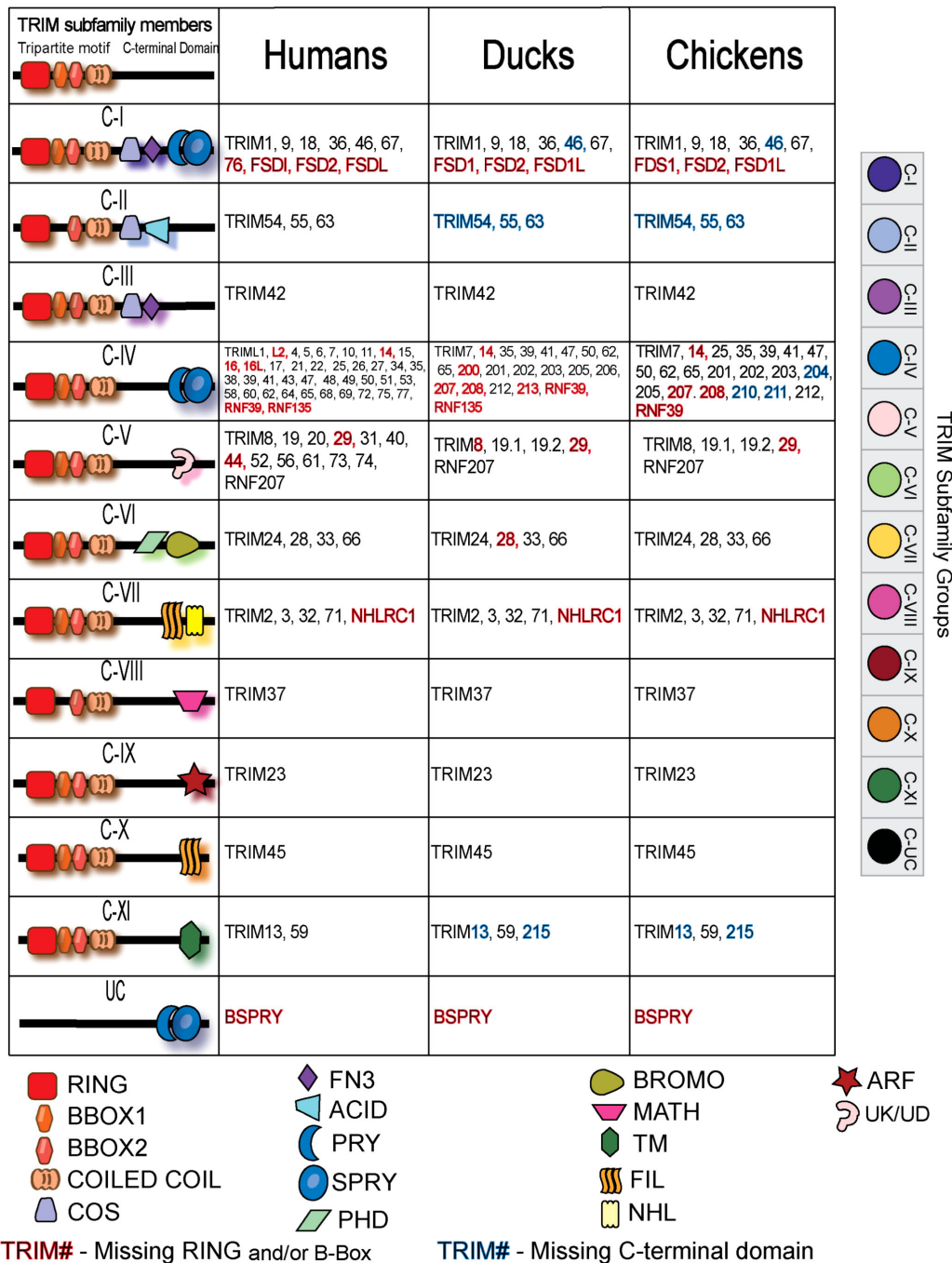


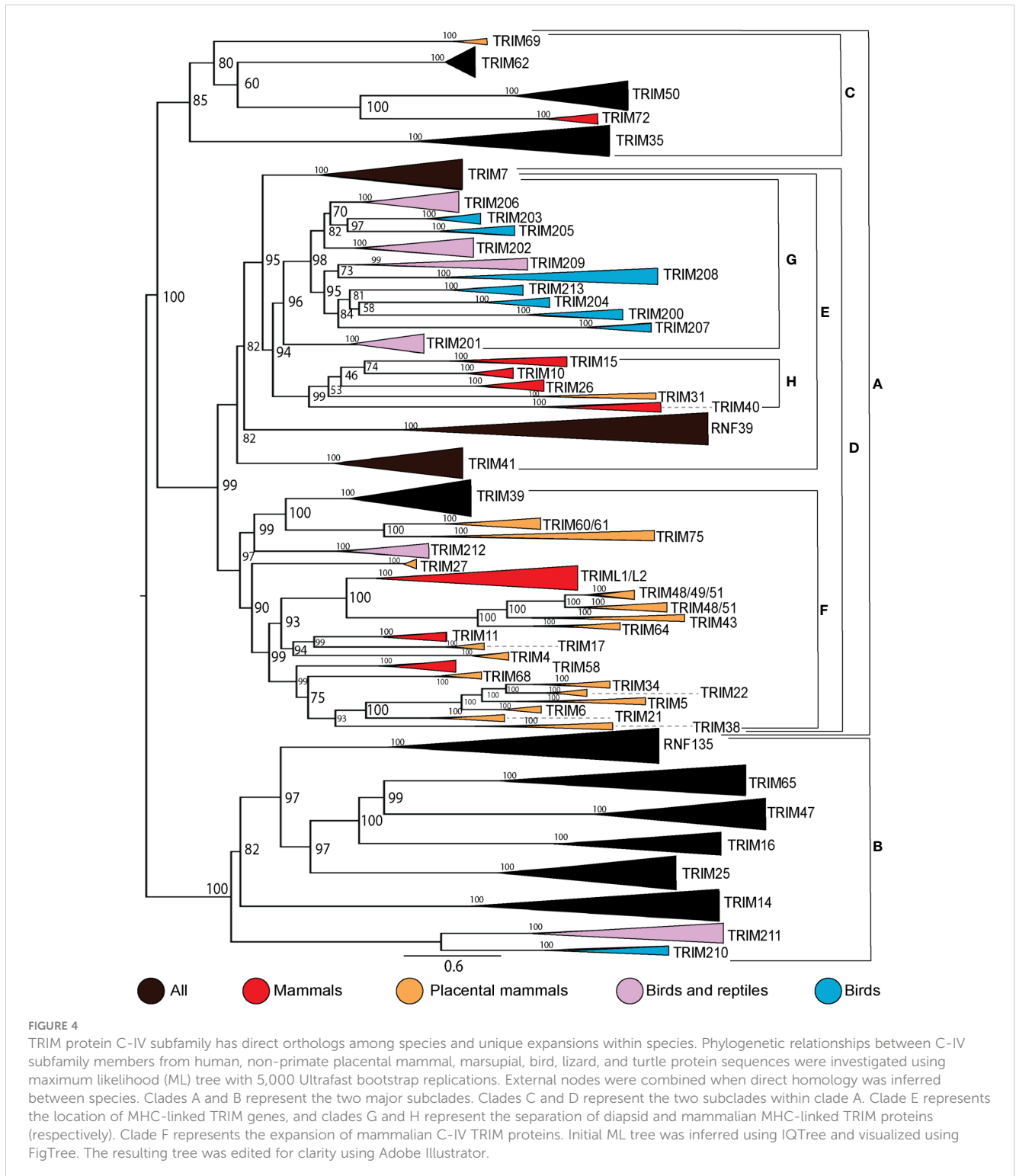
FIGURE 3

Classification of TRIM proteins by their C-terminal domains in human, duck, and chicken. RING, really interesting new gene; COS, C-terminal subgroup one signature; FN3, fibronectin, type III; SPRY, SPla, and Ryanodine receptor domain; PHD, Plant Homeo Domain; BROMO, Bromodomain; MATH, meprin and TRAF homology domain; TM, transmembrane domain; FIL, filamin domain; NHL, NCL-1, HT2A, and Lin-41 repeats; ARF, ADP ribosylation factor-like; UC, uncharacterized.

3.5 Birds, reptiles, and mammals have independent expansions of MHC-linked TRIM genes

To examine the expansion within the C-IV subfamily in higher-order vertebrates, we assembled protein sequences from human, mouse (or other placental mammal if mouse TRIM protein was

non-existent), marsupial, three birds, and two reptiles from the NCBI protein database. We aligned these sequences and inferred a ML phylogenetic tree with 5000 UltraFast bootstrap replications (Supplementary File 8). This tree shows both turtles and lizards appear to have independent expansions of C-IV TRIM genes in the MHC region, which were excluded from subsequent analysis. To assess shared and unique TRIM proteins between species, we



created a phylogenetic tree of C-IV subtype TRIM proteins from representative mammalian, reptile, and avian species (Figure 4). We combined external nodes when all members of the clade were clear orthologs and color-coded these branches by the group of higher-order vertebrates. Many TRIM C-IV subfamily genes are conserved in all vertebrate species, while some are present in only mammals and some unique to birds and/or reptiles. The C-IV subfamily broadly splits into two clades (clades A and B). Clade A splits into

two subclades (clades B and C), and clade C contains some highly conserved TRIM proteins, such as TRIM35, 50, and 63. Within clade A is the nested clade D, which contains the MHC-linked TRIM genes from vertebrates; however, these MHC-linked TRIM genes cluster separately between diapsids (clade G) and mammals (clade H). The human MHC-linked TRIM proteins, TRIM10, 15, 26, and 40, cluster closely with orthologs from both mouse and marsupial but not with any TRIM proteins from birds or reptiles.

Our results suggest lineage-specific expansions of TRIM genes in higher-order vertebrates and that the expansions of avian and mammalian MHC-linked genes diverged from an ancestor of TRIM7. The cluster of MHC-linked TRIM genes in mammals (clade H) appears to have expanded from a common ancestor shared with TRIM7. Many of the avian MHC-linked TRIM genes (clade G) expanded from a common ancestor shared with TRIM201, found in diapsids but not mammals. Our analysis also suggests that previous annotations of the avian MHC-linked TRIM genes were incorrect. Genes that we have assigned names TRIM203, 205, and 206 were previously designated as TRIM27.2, 27.1, and 27L, respectively (37). Our results demonstrate that these genes are not orthologous to human TRIM27. TRIM202 and TRIM204 were called TRIM39.2 and TRIM39.1, respectively. TRIM202 and TRIM204 exist in both birds and reptiles and appear to be missing in mammals. TRIM202 has been published under the name TRIM39 in chickens (51); however, we found a gene in the chicken and duck transcriptomes, which clusters with TRIM39 from other species. TRIM209 is in clade G with the avian MHC-linked TRIM genes and appears in both lizards and turtles. A single avian ortholog was found in the Kiwi (*Apteryx rowi*), but we were unable to locate an ortholog of this gene in other birds. TRIM213 is not located in the MHC region of ducks but clusters tightly with the other avian MHC-linked TRIM genes. Our results suggest that the diapsid TRIM genes in clade G are distant paralogs, expanding from a common ancestor shared with TRIM201, while the mammalian TRIM genes in clade H are distant paralogs originating from a common ancestor shared with TRIM7. These results suggest the diapsid and mammalian MHC-linked genes located in clades G and H are not co-orthologs but arose from distinct lineage-specific duplications. Reptiles appear to have lineage-specific duplications of TRIM genes in the MHC without orthologs in other vertebrates. We found 67 TRIM genes adjacent to the MHC class I genes in the eastern fence lizard and 228 in the yellow pond turtle. Many of these turtle and lizard TRIM genes appear not to have orthologous genes in birds or mammals and appear to be inparalogs within their lineages. It appears that independent duplications and diversification of the MHC-linked TRIM genes have happened often in reptilian, avian, and mammalian lineages.

Mammals have a large expansion of TRIM proteins in clade F. Many of these TRIM proteins are known to have expanded in eutherian mammals such as the expansion of the TRIM5/6/22/34 locus (15, 106, 107). Ducks have two TRIM proteins in this minor clade, TRIM39 and TRIM212. TRIM212 appears to be a result of lineage-specific duplication, as it appears present in diapsids only.

Clade B contains the TRIM25 cluster, C-IV TRIM proteins present in all animals surveyed, including TRIM14/16/25/47/65 and RNF135. This clade also includes TRIM210, which we could only find examples of in birds and TRIM211, which is present in birds and reptiles but not in mammals. This clade contains a TRIM protein found only in reptiles, which we have named TRIM214 (Supplementary File S8) and shares ancestry with TRIM47 and 65. Notably, RNF135 appears absent in species of bird, which also appear to be missing RIG-I (101) but is found in other birds, reptiles, and mammals. TRIM16 appears absent in chicken and

duck but is present in other birds, reptiles, and mammals. This TRIM25 cluster has remained conserved throughout vertebrate evolution, with additional TRIM genes in birds and reptiles.

3.6 The MHC-linked TRIM gene repertoire of birds, reptiles, and mammals has dramatically changed over time

To explore the evolution of TRIM genes in the MHC region, we generated maps of chromosomal locations of MHC-linked TRIM genes for representative vertebrate species (Figure 5), including duck, mouse, human, Tasmanian devil, yellow pond turtle, and Eastern fence lizard. The duck MHC-linked TRIM genes include orthologs of mammalian TRIM7 and TRIM41, and the diapsid lineage-specific TRIM200, 201, 202, 203, 205, 206, 207, and 208 (Figure 5A). MHC-linked TRIM genes show orthology between humans and mice. Mice have TRIM39, 26, 15, 10, 40, and 31, and RNF39 in the MHC region (Figure 5B), which humans share (Figure 5C). As previously reported, humans have TRIM38 and 27 telomeric to the TRIM cluster (35). To ascertain which MHC-linked TRIM genes predate eutherian mammal speciation, we examined the MHC region of a marsupial and two reptiles. Marsupials are an ancient mammal, diverging from eutherian mammals approximately 160 million years ago (108). Tasmanian devils have orthologs to MHC-linked TRIM genes present in either placental mammals or duck. Like humans and mice, Tasmanian devils have TRIM39, 26, 15, and 10, and RNF39 (Figure 5D). Like ducks, Tasmanian devil MHC-linked TRIM genes also include TRIM7 and 41, suggesting that these genes were lost from the mammalian MHC region during eutherian speciation. Different from birds and placental mammals, TRIM25, 47, and 65 are also located adjacent to the MHC-linked orthologous TRIM genes on Tasmanian devil chromosome 4. To see if this is an ancestral organization, we compared the Tasmanian devil to reptile MHC regions. Recent phylogenetic analysis suggests that turtles share a clade with birds and crocodiles, while lizards form a separate clade (109). We found a large expansion of 226 TRIM genes in the MHC region of the yellow pond turtle (Figure 5E) (Supplementary File S2). Many of these TRIM genes do not share obvious orthology to mammalian or avian TRIM genes. Like mammalian MHC-linked TRIM genes, the yellow pond turtle has RNF39 and TRIM39 in the MHC locus. Turtles appear to have an expansion of RNF39 co-orthologs, composed of six RNF39R genes. Some genes orthologous to avian MHC-linked TRIM genes also appear in turtles, including two TRIM206 genes, TRIM7, 41, 201, and 202. TRIM209, a TRIM gene that appears in reptiles and kiwi birds, is also in this region. Like the Tasmanian devil, the TRIM25 locus is adjacent to MHC-linked TRIM genes in turtles. We found 77 TRIM genes in the MHC locus and the TRIM25 locus in the Eastern fence lizard (Figure 5F). The Eastern fence lizard, more distantly related to birds than the yellow pond turtle, has two paralogs of TRIM206 and an ortholog of TRIM201, suggesting that TRIM201 and 206 are ancestral to the other MHC TRIM genes of birds. Like the yellow pond turtle and Tasmanian devil, the Eastern fence lizard MHC-

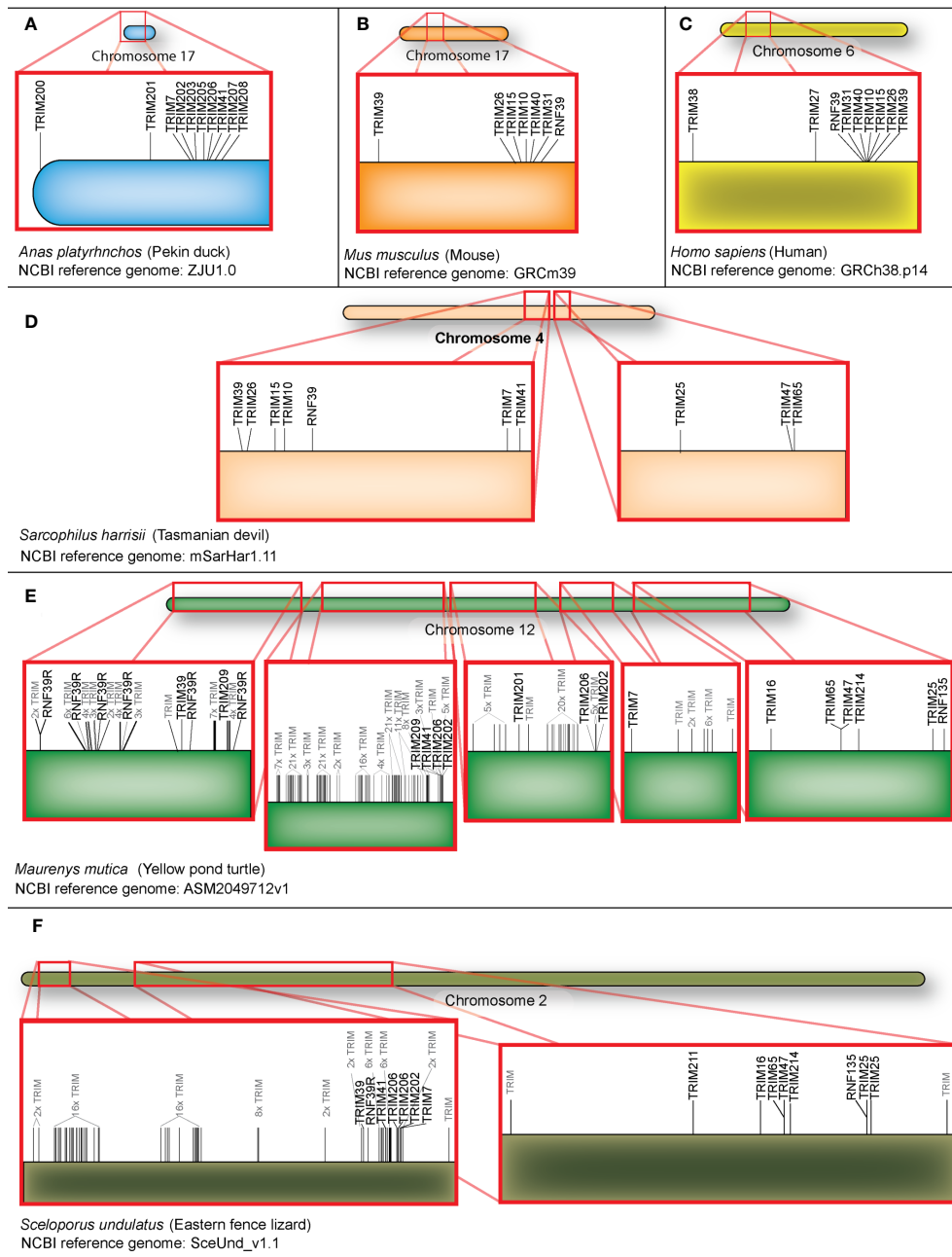


FIGURE 5

Orthologs of TRIM7, TRIM39, TRIM41 and RNF39 are present in all orders, but have been relocated from the MHC of humans and mice. MHC-linked TRIM genes from duck (A), mouse (B), human (C), Tasmanian devil (D), yellow pond turtle (E), and Eastern fence lizard (F) locations were mapped using karyoploteR in the R studio environment. TRIM genes without distinct homologs or names are shaded in gray.

linked TRIM repertoire also contains RNF39, TRIM7, 39, and 41 orthologs. Similar to the Tasmanian devil and yellow pond turtle, the Eastern fence lizard TRIM25 locus is adjacent to the MHC-linked TRIM genes. TRIM211 and TRIM214, both TRIM proteins that clustered with TRIM25, 47, and 65, are in this region in lizard. There were more TRIM genes downstream of the TRIM25 locus in the Eastern fence lizard; however, we decided to limit this comparison to the MHC and TRIM25 loci. Our results suggest the TRIM25 locus may have evolved alongside the MHC locus in lower vertebrates. As this is conserved between marsupials and

reptiles, it suggests that the TRIM25 locus moved away from the MHC independently in avian and placental mammals. Our results suggest that RNF39, TRIM7, 39, and 41 are ancestral MHC-linked TRIM genes with conserved orthologs in reptiles, birds, and mammals.

To investigate differences in the MHC-linked TRIM genes within the avian lineage, we compared the genomic arrangement of the MHC-linked TRIM genes of mallard duck, tufted duck, chicken, kākāpō, barn swallow, and European golden plover (Figure 6). Tufted duck and chicken both contain TRIM204, while this gene appears to be missing in

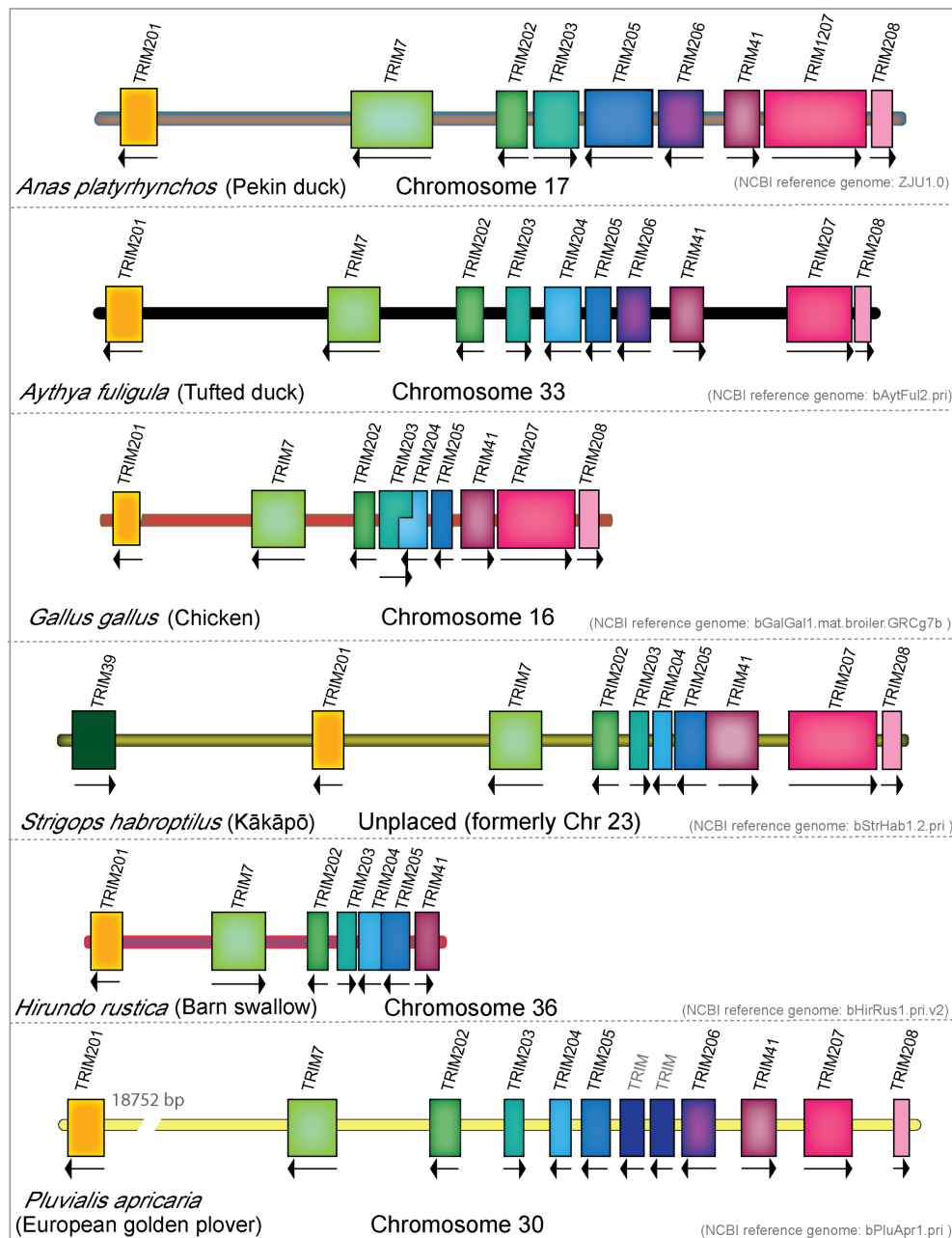


FIGURE 6

MHC-linked TRIM gene repertoire in birds has expanded and contracted during avian evolution. Genomic locations and direction of transcription of the MHC-linked TRIM genes from mallard duck, tufted duck, chicken, kākāpō, barn swallow, and European golden plover. All sizes of genes and chromosome length were normalized to the size of mallard duck gene TRIM208. Break shows gap of indicated size.

mallard duck. We can find the corresponding sequence of TRIM204 in the 5'-UTR of TRIM205 in the mallard duck, which suggests a misassembly. However, there are no detectable transcripts of this gene expressed in any mallard tissues, indicating that it may no longer be expressed in mallard ducks. Both species of mallard ducks have TRIM206, which is missing in chicken, kākāpō, and barn swallow. Barn swallows have lost TRIM206 and two TRIM genes downstream of TRIM41. We observe a species-specific expansion of the MHC-linked TRIM genes in the European golden plover between TRIM205 and 206, with two to three additional predicted TRIM-like genes. These TRIM

genes are closely related paralogs and do not appear to have direct orthologs in birds or reptiles. Phylogenetic analysis of these paralogs found the golden plover TRIM genes cluster in a clade with TRIM203 and TRIM205 and not with any of the reptile-specific MHC-linked TRIM genes (data not shown), suggesting that this may be a lineage-specific expansion. As the current golden plover genome is not annotated and has no accompanying transcriptome data, it is unknown how many of these genes are expressed. Our analysis of the MHC-linked TRIM genes in birds demonstrates these genes are continuously changing over time.

3.7 The TRIM25 locus has undergone gene loss and gain during diapsid evolution

The TRIM25 locus has gone through significant rearrangements during avian evolution (Figure 7). In the duck, genes are arranged TRIM25/RNF135/TRIM65/TRIM47, while in the chicken, they appear in the gene order of TRIM65/47/25. RNF135 was noted as missing in chickens (99, 110), and we could not find it in our chicken transcriptome interrogation. The

inactivation of chicken RNF135 appears to be independent of the chromosomal rearrangement, as only RNF135 is missing, while the surrounding genes ADAP2 and RHOT1 are still present. Like chickens, Adelie penguins appear to be missing RNF135; however, RHOT1 and ADAP2 are still present and near each other. Interestingly, Adelie penguins have an ortholog of TRIM16, which only appears in penguin, ostrich, kiwi, and hoatzin genomes (NCBI). A TRIM16 ortholog is also present in lizard and turtle genomes, suggesting that many birds, including

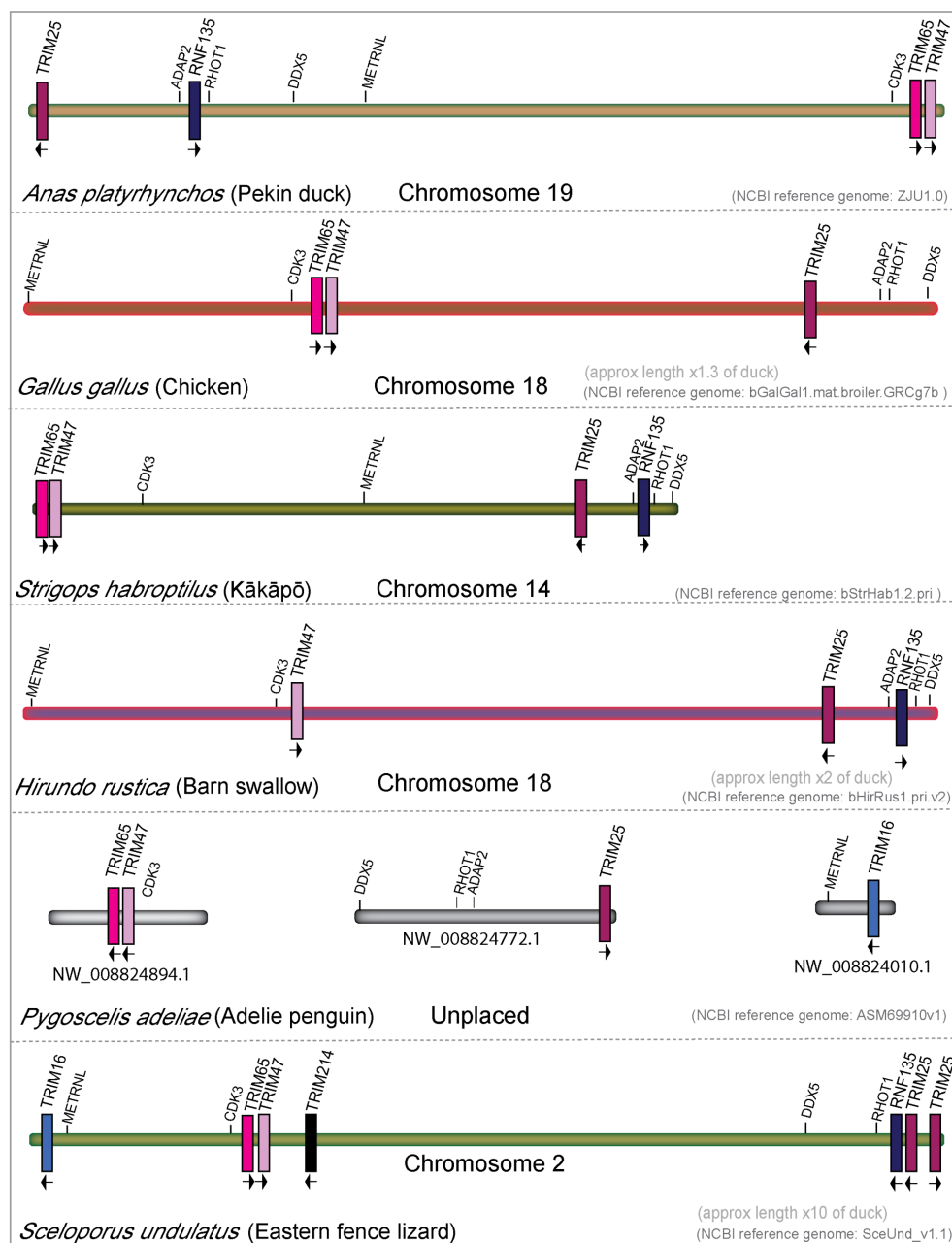


FIGURE 7 The TRIM25 locus has undergone rearrangement during diapsid evolution. Genomic locations and direction of transcription of TRIM genes in the TRIM25 locus from mallard duck, chicken, kākāpō, barn swallow, Adelie penguin, and Eastern fence lizard. All chromosome lengths were normalized to the size of the total TRIM25 locus length in mallard duck.

both chickens and ducks, may have lost TRIM16. Barn swallows appear to be missing TRIM65, which is present in all other birds that we surveyed in this study, and in reptiles. Reptiles have a novel TRIM gene in the TRIM25 locus, which we have named TRIM214, as it has no clear ortholog in mammals. In addition, Eastern fence lizards appear to have a duplication of TRIM25. The TRIM25 locus appears to have undergone significant rearrangement in both repertoire and gene placement throughout avian evolution.

3.8 Duck TRIM proteins cluster within TRIM subfamilies

To infer ancestry in the duck TRIM protein repertoire, we aligned all duck amino acid sequences and generated a ML tree. We color coded the tree based on the subfamily assigned by the C-terminal domain possessed by each TRIM protein or their orthologous TRIM protein in humans (Figure 8A). TRIM proteins segregate into two major clades, with one clade largely

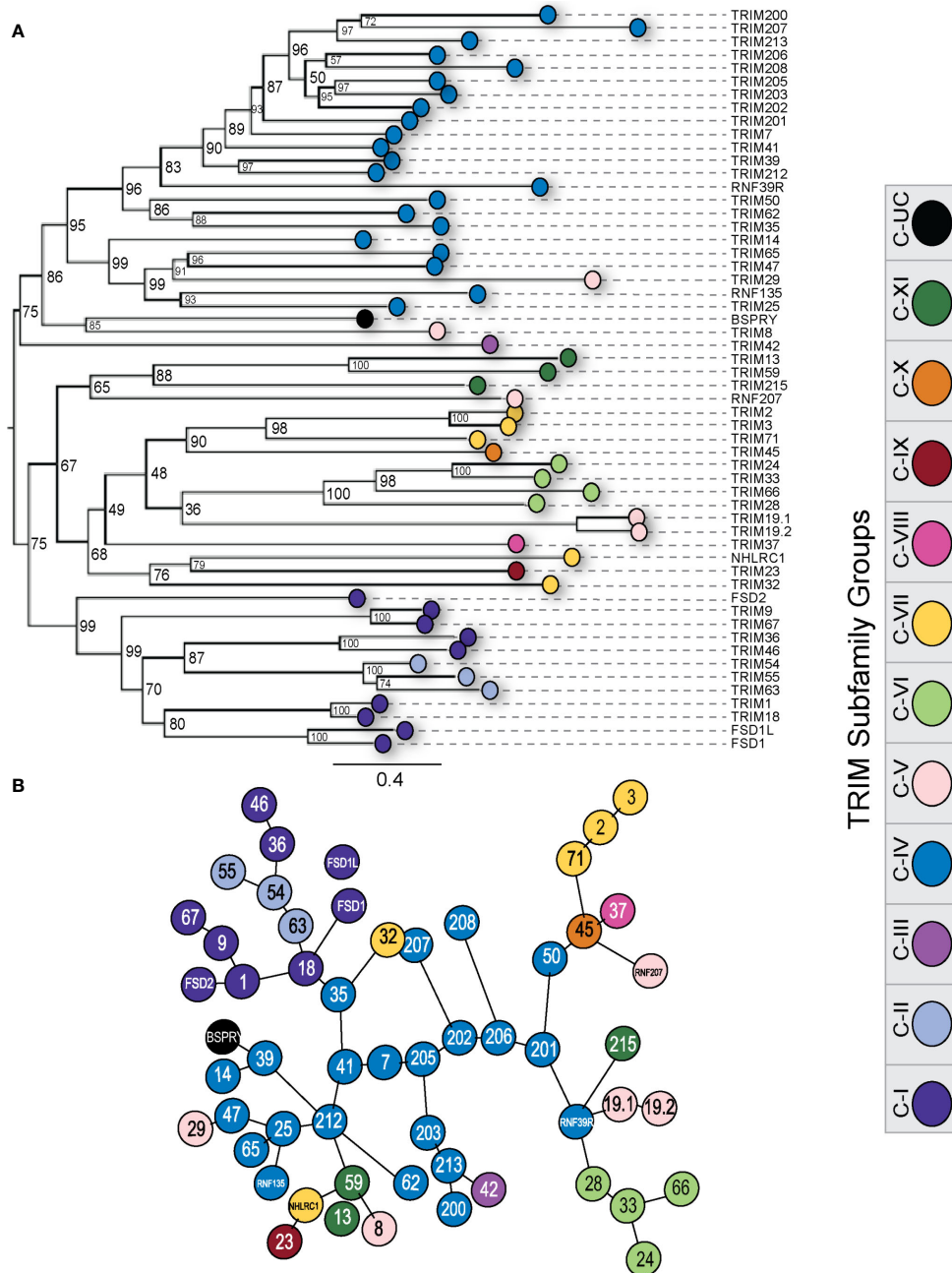


FIGURE 8

Duck TRIM and TRIM-like proteins cluster within subfamilies. Phylogenetic relationships between the duck TRIM protein sequences were investigated using maximum likelihood (ML) trees with 10,000 Ultrafast bootstrap replications (A). The distances between protein sequences were also investigated using a minimum spanning network to infer similarity and function (B). Each external node was color-coded according to TRIM subfamily designation. ML tree was made using IQTree, visualized using FigTree, and edited for clarity using Adobe Illustrator. MSN was made using msa and seqinR in RStudio and edited for clarity using Adobe Illustrator.

composed of the C-IV proteins. As previously mentioned, The C-IV subfamily is often associated with immune responses in mammals. TRIM8, 29, and 42 also cluster in this clade. TRIM8 and 29 belong to the C-V subfamily. The C-V subfamily has an unclassified C-terminal domain. TRIM42 is the sole member of the C-III subfamily and has a COS and FN3 C-terminal domain.

Within the second major clade, many of the subfamilies are represented, which supports the hypothesis that these subfamilies originated from a common ancestral TRIM gene. Typically, members of a subfamily group together in a clade. Subfamilies C-I and C-II appear to have descended from the same common ancestor, as they segregate into one large clade. Members of subfamily C-V do not share clades and are distant. This is perhaps not unexpected as the C-V subfamily is classified as RBCC domain containing TRIM proteins with unclassified C-terminal regions. The TRIM-like gene NHLRC1 and TRIM32 do not share a clade with the other C-VII subfamily members, suggesting they could have arisen through exon shuffling events independently from the other C-VII family members TRIM2, 3, and 71. We only included TRIM or TRIM-like genes that coded for full-length proteins in this analysis. TRIM211 had premature stop codons throughout the sequence and, as such, was excluded.

To assess the structural similarity of the TRIM proteins, we generated a minimum spanning network (MSN) (Figure 8B), which connects protein sequences (nodes) based on the distance between proteins without inferring ancestry and instead can infer shared function between these proteins (111). Most members of a subfamily cluster closely together, including members of the C-IV and C-VI subfamilies. Many C-V members are quite distant in the MSN, reflecting their disparate C-terminal domains and likely dissimilar functions. Most of the C-IV subfamily of TRIM proteins all cluster together in the center of the MSN. The MHC-linked C-IV TRIM mostly form the inner branches and cluster closely, apart from TRIM200, 207, and 208, which are quite far away from the other MHC-linked TRIM proteins on outer nodes. Proteins from the TRIM25 cluster group tightly together, expanding from the TRIM25 node, with the exception of TRIM14, which remains close but branches off of TRIM39 on a separate node. TRIM29, a C-V subfamily member, branches closely with TRIM25, 47, and 65, and RNF135. TRIM212, present in diapsids but appears missing in mammalian lineages, also clusters closely with the TRIM25 expansion. TRIM200 and TRIM213, both found in ducks but appear to be missing in chickens, share a distinct branch with TRIM203.

3.9 While most duck TRIM genes are ubiquitously expressed, some are tissue specific

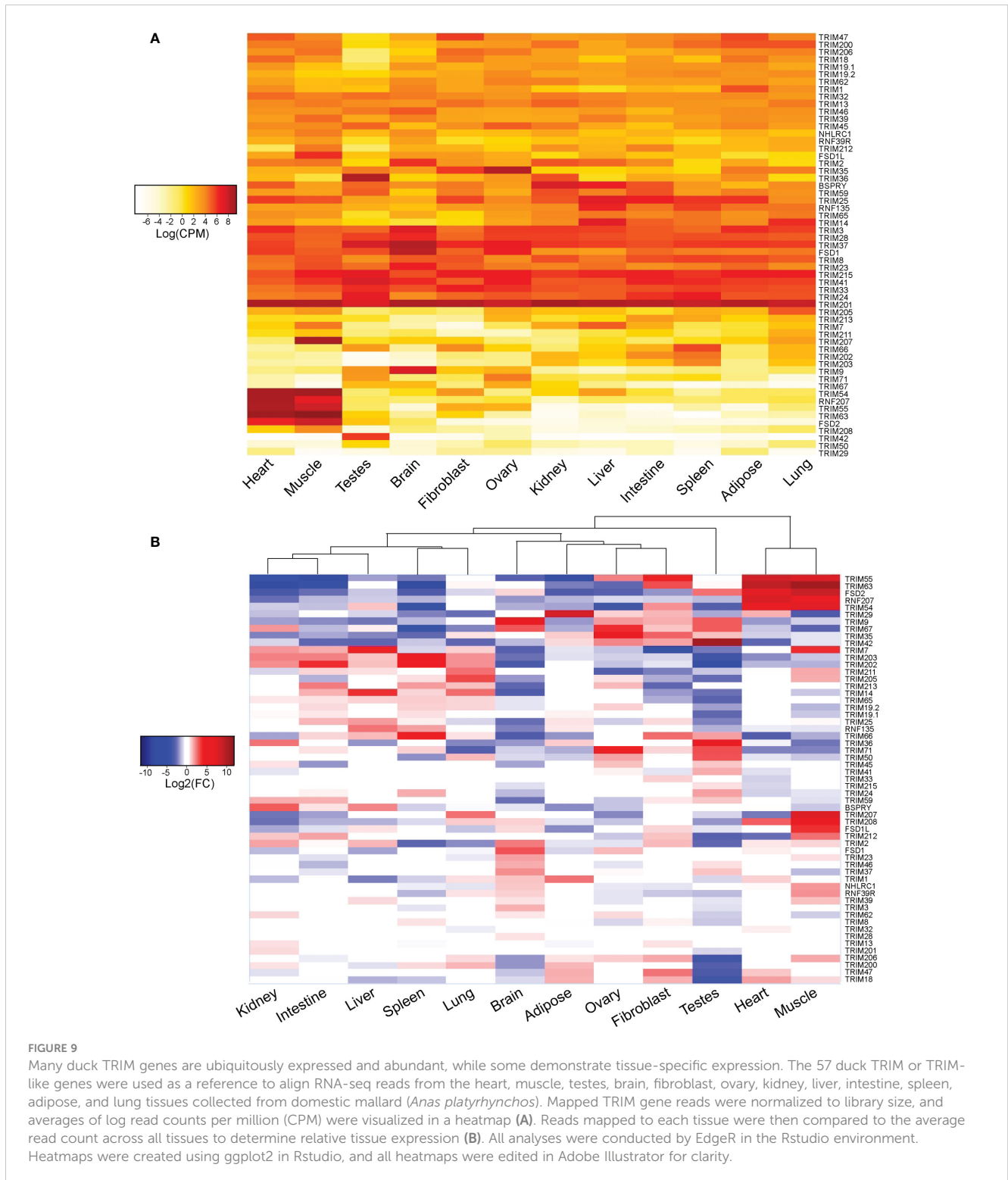
To visualize the expression of TRIM genes in duck tissues, we mined the NCBI SRA database for RNA-seq reads from various tissues and aligned these reads to our 57 TRIM or TRIM-like sequences. We analyzed pairwise differences in TRIM differential expression (DE) between tissues using an MDS plot and most tissues clustered together (Supplementary Figure S2). Muscle and

heart cluster very closely together when TRIM gene DE is compared between tissues. While all samples cluster together closely with like tissue samples, some tissues such as brain and testes tissues, for example, have more distinct clustering patterns. These results suggest that some TRIM genes are tissue specific, leading to distinct MDS plot clustering of tissues.

To visualize abundance and relative expression of TRIM genes in each duck tissue, we generated heatmaps from the analysis of our normalized read counts (Figure 9). While many of the 57 duck TRIM genes are ubiquitously expressed in the tissues sampled, with varying levels of read counts per gene, there are some which are more specific (Figure 9A). TRIM54, 55, and 63, FSD2, and RNF207 have more reads mapped to muscle and heart than to other tissues. TRIM36 and TRIM42 were abundantly expressed in testis. TRIM9 appears to be highly expressed in brain tissues. TRIM201 and 41, two of the MHC-linked TRIM genes, had high read counts in all tissues sampled. In contrast, TRIM29 had very few averaged reads counted in any tissues sampled. Although TRIM211 encoded a sequence with a premature stop codon, it was still expressed at a low level in all tissues sampled, suggesting that the promoter is still active.

To determine the relative expression of TRIM genes sampled in each tissue, we compared the expression of TRIM genes from each tissue to the average expression in all samples. White or near zero log₂ fold change (FC) values denote average expression of the gene in that tissue when compared to all other tissues, while negative (blue) or positive (red) log₂(FC) values suggest lower or higher relative expression (respectively) when compared to all other tissues. The resulting heatmap visualizes the relative expression of TRIM genes in each tissue and demonstrates some of the more subtle differences in TRIM gene expression between the tissues (Figure 9B). Hierarchical clustering separates the expression pattern of TRIM genes into two major clades. The first major clade contains testis, then branches into two smaller subclades. The first subclade contains the immune tissues: kidney, intestine, liver, spleen, and lung. The MHC-linked TRIM genes TRIM202 and 203 have higher relative expression in these tissues than in any of the other tissues sampled. The second subclade contains brain and adipose tissues, fibroblasts, and ovaries. Muscle and heart form the second major clade, with the pattern of TRIM expression between these two tissues being very similar. Notably, TRIM7, 207, 208, and 212, and FSD1L have high relative expression in the muscle, and only average or low relative expression in the heart.

We summarized the five highest and lowest statistically significant [false discovery rate (FDR) < 0.05] relatively expressed TRIM genes in each duck tissue analyzed (Figure 10). The C-IV subfamily of genes are often associated with inflammation and immune responses. Immune relevant tissues such as the lung, spleen, and intestine had a predominance of C-IV family members expressed at a higher level than in other tissues. Many MHC-linked genes also have higher relative expression in these immune relevant tissues. TRIM205, 207, and 202 have higher relative expression in lung. TRIM203 and 202 have higher relative expression in the spleen. TRIM202, 203, and TRIM7 have higher relative expression in the intestine. Immune privileged tissues such as brain and gonads, the C-IV subfamily genes, are among the least relatively expressed. Fibroblasts had low relative expression of many



C-IV family members. TRIM35 and 47 are C-IV subfamily members with higher relative expression in duck fibroblasts, while TRIM55, 63, and 66 have higher relative expression in fibroblasts but are not C-IV subfamily members. TRIM29 was flagged as having higher relative expression in the adipose tissue; however, on closer inspection, two duck adipose samples had an over-representation of reads mapped to TRIM29. As the adipose tissues sampled clustered together when analyzed for TRIM gene

expression (Supplementary Figure S2), it is likely that two of the birds sampled for this tissue were outliers.

4 Discussion

Here, we identify TRIM and TRIM-like genes in the duck by mining the SRA, gene, and protein NCBI databases and generating

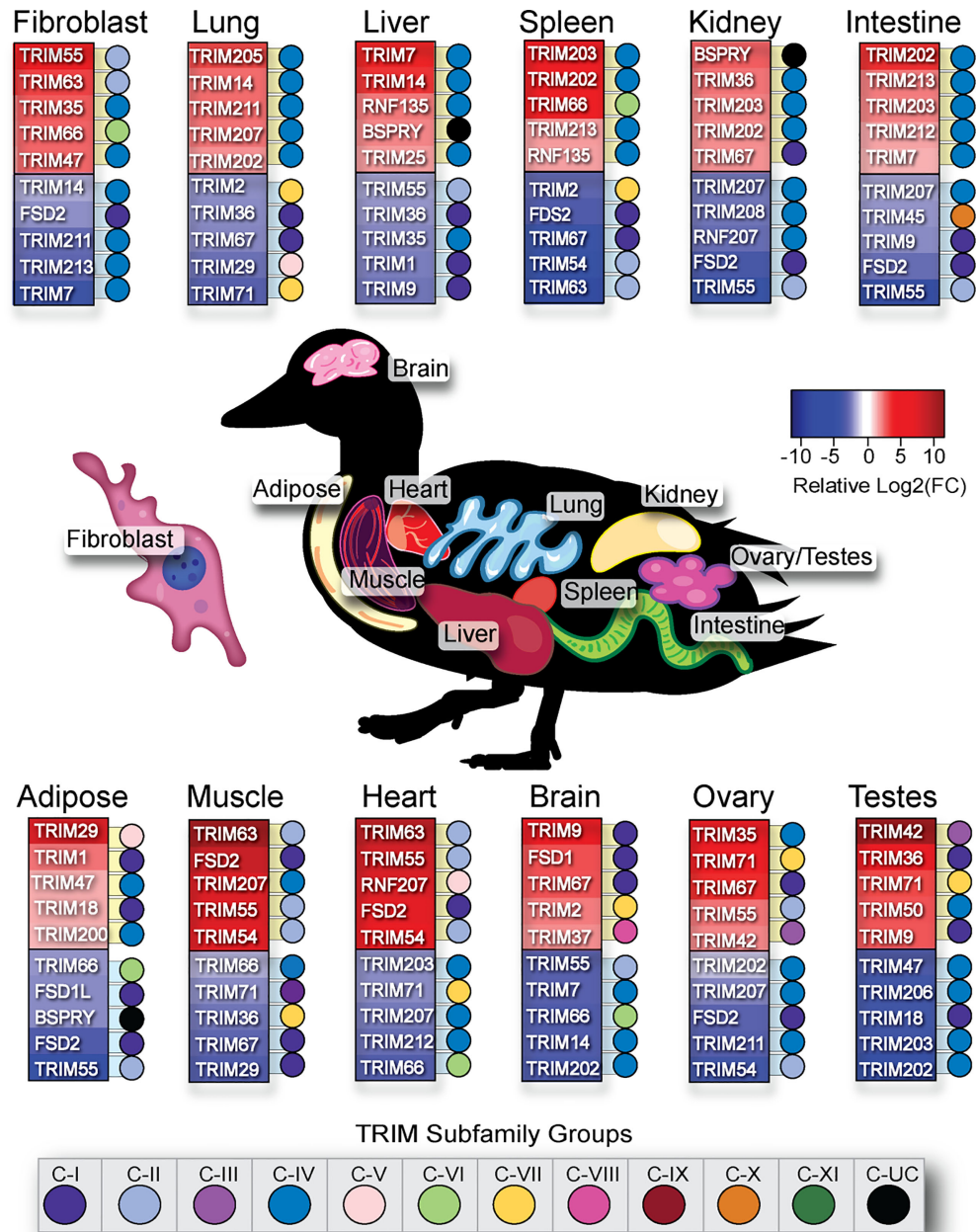


FIGURE 10
Immune relevant tissues have higher relative expression of C-IV TRIM family members. The relative expression of 57 duck TRIM or TRIM-like genes fibroblast, lung, liver, spleen, kidney, intestine, adipose, muscle, heart, brain, ovary, and testes were sorted by statistical significance (FDR<0.05) and then organized into top 5 highest or lowest relatively expressed TRIM genes per tissue. Dots next to TRIM gene are colour-coded according to TRIM subfamily designation.

and interrogating a *de novo* transcriptome assembly. We found 57 duck TRIM genes, classified by their C-terminal domains, and compared them to mammalian and chicken TRIM repertoires. Most duck TRIM proteins are similar to their chicken orthologs. The duck TRIM sequences were aligned, and phylogenetic relationships between the TRIM proteins were inferred. Phylogenetic analyses show expansion of the C-terminal PRY-SPRY containing C-IV TRIM subfamily. Many duck MHC-linked C-IV TRIM genes lack mammalian orthologs, and this repertoire

has expanded during diapsid evolution. The TRIM25 locus has also undergone rearrangement during vertebrate evolution, and our analyses suggest that this locus was adjacent to the MHC locus but separated onto different chromosomes in birds and eutherian mammals. Finally, we aligned RNA-seq reads from different duck tissues to the TRIM gene sequences to determine relative expression levels of these TRIMs in each tissue.

To our knowledge, only one study has previously classified the TRIM family genes in birds. Sardiello and colleagues listed 37

chicken TRIM genes during their analysis of the evolution of group I and group II TRIM genes in vertebrates (9). Other reports on avian TRIM genes characterized the expanded B30.2/PRY-SPRY TRIM genes in chickens (36, 94), turkeys (112), and ducks (37). Outside of birds, lineage-specific expansions of B30.2/PRY-SPRY TRIM genes have also been noted in fish (113, 114). Teleost fish had three separate expansions of B30.2/PRY-SPRY TRIM genes, with some specific to teleost fish as they are not found in amphibians, birds, or mammals. Zebrafish have 208 TRIM genes (*Danio rerio*), pufferfish have 67 (*Tetraodon nigroviridis*) (39), and grass carp (*Ctenopharyngodon idella*) have 42 TRIM genes (115). The differences in TRIM gene repertoire numbers in fish are partly due to the whole genome duplication events (116). Here, we document what appears to be a similar expansion of MHC-linked TRIM genes in reptiles, with many of these genes lacking direct orthologous counterparts in birds.

From our transcriptome interrogation, we found 57 TRIM genes in the duck. Of these TRIM genes, 54 could be found on their respective chromosomal locations in the NCBI Pekin duck genome (assembly ZJU1.0), and the location of FSD1 was inferred from synteny of surrounding genes in Ensembl rapid release genomes (GCA_017639305.1 and GCA_017639285.1). The duck TRIM gene repertoire spans 21 chromosomes, suggesting that many of these genes have evolved independently. Similar expansions can also be seen in human (9) and fish (39, 113, 115). As previously described, there is an expansion of PRY-SPRY containing TRIM genes within the MHC-B locus on chromosome 17 in the duck (37). The human MHC-linked TRIM genes are tightly linked and include TRIM10, 15, 26, 31, 39, and 40, and RNF39, while TRIM27 and 38 are telomeric to this region (35). In mice, TRIM27 and 38 are on chromosome 13 in the A3.1 region, while TRIM10, 15, 26, 31, 39, and 40, and RNF39 are located on chromosome 17 in the B1 region (105). In chickens, the MHC-B locus is located on chromosome 16 and contains many of the same TRIM genes found on chromosome 17 in ducks in syntenic organization, including TRIM 7, 201, 202, 203, 205, 207, 208, and 41 (36, 37, 94, 117). These genes were previously named to what was presumed to be orthologous human TRIM genes due to synteny and closest BLAST identity hits. Here, we updated the names of these genes to reflect that phylogenetic analysis suggests that they do not have clear orthologous genes in mammals. Previously, the chicken MHC-linked TRIM genes were named TRIM7.2, 7/7.1, 39/39.2, X/27.2, BR/39.1, 27/27.1, 41, B30.2-1/BTN-1, and B30.2-2/BTN-2 (36, 94). Ruby and colleagues presumed that synteny of the TRIM genes in the MHC-B locus of chicken was conserved with human and used exon-based identity to infer ancestry, although their phylogenetic analysis suggested that TRIM27 and TRIM39 in chicken did not share ancestry with their presumed human orthologs. With the inclusion of TRIM genes from genomes of other vertebrates in the phylogenies, we show that many avian MHC-linked TRIM genes arose independently from the human MHC-linked TRIM genes. While mammalian MHC-linked TRIM genes appear to have arisen from a duplication event involving the common ancestral gene shared with TRIM7, avian MHC-linked TRIM genes appear to have arisen from an ancient duplication event involving the common ancestor shared with TRIM201. Both lizards and turtles have TRIM7, 41, and 206,

while turtles, presumed to be a closer relation to birds than lizards (109), also have orthologs to duck TRIM201 and 202. While eutherian mammals do not share direct orthologs to duck TRIM genes in their MHC-locus, the Tasmanian devil has TRIM7 and 41 in their MHC locus, suggesting that TRIM7 and 41 are ancestral MHC-linked TRIM orthologs. TRIM206 appears missing in galliform lineages (37) but is present in ducks and many other birds. We document that TRIM204 appears to be missing in mallard duck and is present in chickens. This gene is annotated in tufted ducks and shares synteny with the chicken TRIM204 gene. Thus, it appears that the TRIM genes associated with MHC can change over time (117). Our results support a “birth and death” model of evolution of the MHC-linked TRIM genes (118), with lineage-specific duplication events increasing the TRIM diversity, and previously existing genes being removed by deletion, mutation, or inactivation (as in the case with TRIM204). We demonstrate that some MHC-linked genes have changed significantly between reptiles, birds, and mammals. It appears that TRIM7, 41, and RNF39 are ancestral to the MHC-linked TRIM genes in higher-order vertebrates.

Our phylogenetic trees show that ducks and chickens have direct orthologs for 53 TRIM genes, which segregate into distinct clades. Ducks have four proteins, which appear absent in chickens RNF135, TRIM200, 213, and 206. Chickens have TRIM204, 211, and 210, which are either incomplete or missing in ducks. TRIM211 orthologs are found in both birds and reptiles, while TRIM210 appears to be present only in birds. Duck RNF135, chicken TRIM211, and TRIM210 all fall within the clade containing TRIM14, 25, 47, and 65. No functions have yet been published for TRIM211 and TRIM210; however, due to their phylogenetic location, they may bind to helicase domains similar to the function of members of this clade in mammals (104). Some avian genes have been notoriously difficult to identify due to high GC content and lack of representation of these regions in genomic libraries (119). Indeed, this may be the reason that RNF39R is unannotated and unplaced in the chicken and the duck genomes. The newly identified RNF39R sequences have 72% and 74% GC content in the duck and chicken (respectively), making the contig harder to assemble and place in the genome.

Species-specific differences in TRIM repertoires are seen in other non-avian vertebrates, especially when comparing C-IV TRIM genes. There is a large expansion of C-IV TRIMs on chromosome 11 in primates (120), including TRIM5, 6, 22, and 34. TRIM5 and 22 restrict retroviral replication in primates (13, 14, 121, 122). Humans have a single copy of TRIM5, cattle have five co-orthologs of TRIM5, while TRIM5 is deleted in dogs (15). Primates have additional TRIM duplications on chromosome 11 close to the TRIM5/6/22/34 cluster, including TRIM49 and 64 (120). Bats host many types of viruses (123–125) and have duplications of TRIM5 and 22 (107, 126). There are other species-specific differences of C-IV members in bats, including duplications of TRIM25, 41, 60, and 75 (126). Duplication and deletion events of these genes may be a response to selective pressures from viruses. Large duplication events have occurred in fish C-IV TRIM proteins, and the number of these TRIM gene vary greatly, even within orders of fish. For example, *Perciformes* (perch-like fish) have from 107 to 672 C-IV TRIM genes in their

genomes, depending on species (114). It is unknown if the fish C-IV TRIM expansions are duplicating from a similar common ancestor or expanding from different common ancestors in different types of fish. TRIM repertoires in vertebrates are likely rapidly evolving due to selective pressures from pathogens.

Our identification of two TRIM19 co-orthologs in ducks and chickens prompted us to search available avian genomes, which indicate two paralogs of TRIM19. In mammals TRIM19/PML is a key component of PML nuclear bodies (127). PML nuclear bodies regulate many important processes in mammals, such as the DNA damage response, apoptosis, and gene expression (128–130). PML is also involved in immune responses to viruses, by regulating signaling pathways during infection (131, 132), and are targeted by viruses to inhibit this signaling (133, 134). PML has not been annotated in any published fish (39, 113) or amphibian lineages (NCBI). The PML co-orthologs, however, appear in both birds and reptiles, suggesting this duplication happened in the common ancestor of diapsids. It is unknown if either of the avian TRIM19 paralogs form PML nuclear bodies or has antiviral activity.

Our phylogenetic tree of duck TRIM proteins demonstrated that most TRIM subfamilies form separate clades. The C-IV subfamily, the largest and most diverse of the subfamilies, has two major clades, the first containing the MHC-linked TRIM proteins. The avian-specific TRIM213 also clusters in the MHC-clade. This TRIM potentially translocated to or from the MHC region earlier in vertebrate evolution. TRIM39 and RNF39R are unplaced in the duck and chicken genomes, however, are located in the MHC-region in reptiles, mammals, and, in the case of TRIM39, also in the kākāpō. This suggests that RNF39 and TRIM39 may also be located in the MHC region of the duck and chicken. The MHC region has rapidly duplicated and expanded throughout vertebrate evolution, and many genes in this region are involved in adaptive or innate immune responses (135, 136). In chickens, the MHC region was dubbed as the “minimal essential MHC” due to it being much more compact and simpler than MHC regions found in mammals (137, 138), and our comparisons of this region between duck and chicken demonstrate that the chicken MHC region is more condensed than in ducks. Reptiles have large expansions of uncharacterized TRIM genes in the MHC region without any obvious orthologs in birds or mammals.

In humans, the MHC-linked TRIM genes attenuate innate immune signaling pathways (105), thus regulate responses to infection. As the avian MHC-linked TRIM genes are closely linked, they likely have co-evolved and may also be involved in innate immunity. Duck TRIM206 and TRIM205 both modulate signaling downstream of constitutively active RIG-I CARD domains when overexpressed in chicken cells, with TRIM206 increasing IFN β promoter activity and TRIM205 decreasing it (37). It is not yet clear which components in the innate signaling pathway are targeted by these TRIM proteins. In birds, an ortholog of human TRIM41 was found in the MHC-B locus of chicken, with orthologs later found in the turkey (112) and duck (37) MHC-B locus. We also found TRIM41 in the MHC region of reptiles and marsupials. Human TRIM41 is not found in the MHC and is instead located on chromosome 5. TRIM41 in humans is known to restrict viral replication by selective targeting and ubiquitination of

viral proteins (139, 140) and by augmenting antiviral signaling pathways (141). It is unknown if the avian ortholog of TRIM41 can restrict virus; however, TRIM41 is well conserved between birds, reptiles, and mammals.

A second group of duck TRIM genes within C-IV subfamily having PRY-SPRY domains includes the closely related TRIM25, 47, and 65, and RNF135. Our phylogenies support TRIM14, 25, 47, and 65, and RNF135 all originating from the same common ancestral TRIM gene in birds, reptiles, and mammals, which is consistent with what other groups have found when analyzing these TRIM proteins in mammals (104). Our analysis also suggests that TRIM29, a member of the C-V subfamily, belongs in this clade. In mammals, RNF135, TRIM14, 25, and 65 bind helicases involved in immune signaling using their PRY-SPRY domains, while the function of TRIM47 is unknown (104). Duck TRIM29 inhibits signaling pathways downstream of these helicases by catalyzing the addition of K29-linked ubiquitin to the signaling adaptor MAVS, resulting in a decrease in IFN signaling downstream of MAVS (50). RNF135 has previously been excluded from TRIM repertoires because it does not have the classical RBCC motif, as it is missing the B-box domain. Previous studies performed in our lab could not locate a RING domain in RNF135, and we had suggested that this protein would be largely inactive (110). However, the recent duck genome assembly and HMMER searches of the sequence obtained from our *de novo* transcriptome indicate that duck RNF135 does have a RING domain. The minimum spanning network places RNF135, TRIM25, 29, 47, and 65 on the same branch, while TRIM14 is branched separately, but still in proximity. While there is no documented function of TRIM47 as a modifier RLR pathways, the phylogenetic relationships and the placement in the MSN suggest that this protein may be involved in immune function in lower vertebrates. We demonstrated that the TRIM25 locus is adjacent to the MHC locus in reptiles and remains adjacent to the MHC locus in Tasmanian devils, thus likely was the ancestral organization. The TRIM25 locus is no longer adjacent to the MHC of birds or eutherian mammals because of genomic rearrangement during vertebrate evolution. Penguins and reptiles have TRIM16, which appears to be missing in many other birds, including chicken and duck. This suggests that the TRIM25 locus has undergone significant rearrangement throughout evolution. Many viruses target TRIM25 (142–145) and RNF135/RIPLET (45, 146) in order to evade antiviral responses in mammals. It is possible that the changes in this locus are due to selective pressures from pathogens targeting TRIM proteins.

Modifiers of the RIG-I signaling pathway appear missing in avian lineages. RNF135 was reported missing in chicken (110), Japanese quail (*Coturnix japonica*) (100), Procellariiformes (petrel), Sphenisciformes (penguin), and Falconiformes (falcon) genomes (101). RIG-I, the cytoplasmic detector of single-stranded RNA viruses, also appears missing in chickens (147), other Galliformes, petrels, penguins, and falcons (101). RIG-I is stabilized by ubiquitinylation by RNF135/RIPLET in mammals to increase type I interferon signaling during infection (104, 148–150). The loss of RNF135 corresponds with the loss of RIG-I in birds (101). Recently, Krchlíková and colleagues identified remnants of the RNF135 gene with partial exons with frameshifts in chickens.

Pseudogenization of RNF135 has happened in many galliform birds, which have lost RIG-I, while the intact gene is present in two basal galliform birds that also have RIG-I. TRIM206 augments the RIG-I signaling pathway when cotransfected with constitutively active RIG-I in chicken cells; however, the mechanism is unknown (37). Kākāpō and barn swallow both also appear to be missing TRIM206; however, both species have RIG-I (NCBI). Penguins appear to be missing RNF135 but have TRIM206, and most species of penguin also appear to be missing RIG-I (101) with the exception of the Adelie penguin, which has RIG-I (NCBI). Falcons also appear to be missing RIG-I (101); however, we found an ortholog of TRIM206 in the Saker Falcon. While the loss of RIG-I and RNF135 appears to be linked in most bird species, the loss of TRIM206 and RIG-I does not seem to be correlated; however, a thorough search of avian genomes should be done to confirm this. It is important to note that while many genes appear missing in avian genomes, our data suggests that even with new long read genomic sequencing technology, some avian genes remain “hidden” in dark DNA (119) and can only be detected in transcriptome data. This should also be taken into consideration when analyzing avian genomic resources.

Many duck TRIM genes are ubiquitously expressed in all tissues sampled; however, some do show tissue-specific expression. Muscle and heart tissues in the duck had the most similar expression of TRIM genes of any of the tissues sampled. TRIM54, 55, and 63 are also known as muscle-specific ring finger (MURF) genes and are primarily expressed in muscle fibers in mammals (151, 152). Duck muscle and heart tissues highly express TRIM54, 55, and 63. TRIM9, 46, and 67 are all associated with neuronal development and brain tissues in humans (1, 152–154), and this pattern of higher relative expression in brain tissue is consistent with what we see in the duck. TRIM42 has the highest relative expression in the testis of the duck, which is consistent with human (152). We reported that duck TRIM71 has higher relative expression in both testis and ovary, compared to other tissues. In human, TRIM71 is highly expressed in testis but not ovary (152), suggesting that even with highly conserved TRIM genes such as TRIM71, tissue specificity can change during evolution.

Immune-relevant tissues had higher relative expression of the C-IV TRIM subfamily members. TRIM14, 202, 205, and 207 had the highest relative expression in lung tissues, while TRIM203, 202, and 213 had the highest relative expression in both the spleen and intestine. TRIM202, 203, and 213 are not yet functionally characterized, making them potential candidates for future immunological studies. TRIM genes which encode immune modulating proteins, had much less relative expression in immune privileged sites such as the brain and gonads. TRIM25, for example, which is known to help increase RIG-I signaling during viral infection in mammals (49) and ducks (46, 49), had much less relative expression in brain. TRIM206, known to increase IFN- β signaling when cotransfected with RIG-I in chicken cells (37), had lower relative expression in the testes. Immune privileged sites also had most C-IV subfamily TRIM genes expressed at a lower relative level than other tissues. It is possible that these transcripts have lower expression in immune-privileged sites to prevent inflammatory responses. Ducks have high relative expression of

TRIM14 in both the lung and liver. TRIM14 has not been functionally characterized in ducks but is an antiviral protein in mammals, which can target multiple viruses for degradation such as hepatitis B virus (155) and influenza A virus (30). Ducks act as the natural host and reservoir to influenza A virus (40, 41, 156) and usually have reduced symptoms when infected with low pathogenic strains (40, 157). Highly pathogenic strains of IAV, however, replicate in the lungs of infected ducks and can cause disease (158–160). Young ducks are also susceptible to duck hepatitis virus (DHV), which replicates in the liver and results in liver damage and mortality (161–163). Potentially, the higher expression of duck TRIM14 in the lung and liver is a conserved protective mechanism that allows a quick response to viral infections from RNA viruses such as IAV and DHV.

Some TRIM genes are notably absent in birds, as they were not found in our genome searches or transcriptome. These include the cluster of C-IV TRIM genes located on chromosome 11 in humans, including TRIM6, 5, 22, and 34 (15). Many of these genes have direct antiviral activities, first noted in TRIM5alpha, shown to restrict HIV in non-human primates (23, 121). TRIM22 is known to restrict influenza virus in mammals (164). TRIM6 augments antiviral signaling pathways (165). These genes have undergone expansion and contraction in the mammalian genome, presumably in response to pathogen pressures (15). Our phylogenies suggest that the TRIM5/6/22/34 expansion in mammals happened within the expansion of other mammalian specific TRIM genes including TRIM21, 27, 38, 58, and 68. Birds also appear to be missing TRIM20/PYRIN, a C-V TRIM with a PYRIN domain in the N-terminus. TRIM20 has a proinflammatory role in mammals due to its interactions with the inflammasome component apoptosis-associated speck-like protein containing caspase recruitment domain (ASC) (166, 167). We found no TRIM20 ortholog in our duck transcriptome or in genomes of other birds. We also did not find an ASC protein in our transcriptome or on NCBI, suggesting that the “pyrin-inflammasome” may be absent in ducks.

More TRIM genes may be present in the duck but have greatly diverged and were not detected by our transcriptome interrogation. The incompleteness of the duck genome paired with an incomplete set of tissues to use for *de novo* assembly, leaves the possibility that we have missed some TRIM genes. Our study lacks tissues from the eye, stomach, pancreas, and bursa. If these tissues have tissue-specific expression of TRIM genes, we likely would not be able to find them in our transcriptome. TRIM genes are highly regulated during development, and screening embryonic tissues at various stages of development might help to classify TRIM genes predominantly expressed during development that would otherwise be rare. Additionally, recent duplications or highly similar genes cannot be resolved using *de novo* transcriptome assembly. If highly similar genes are not annotated in the available genomes, we would have missed them.

Throughout our analysis of TRIM genes, the issue of mis- or improper annotations arose. Many of the diapsid TRIM genes are assigned locus numbers and have computer-generated descriptions. Many gene names are used redundantly to describe multiple independent TRIM genes (i.e., TRIM39L, RFPL, etc.). We suggest

that pipeline-assigned gene names should not be taken as proper annotations, especially in lower vertebrates, as these gene names are assigned on most similar human hit, which often is a human TRIM gene of low similarity. Naming of TRIM genes in non-human vertebrates has relied on BLAST, identity analysis, and presumed synteny to humans. Our results demonstrate that TRIM gene evolution is too complex to rely on these methods alone, especially when naming the C-IV TRIM subfamily members. To determine orthology, phylogenetic analysis should include characterized TRIM members from multiple taxa. We have implemented new rules in naming TRIM genes in non-human vertebrates and offer some suggestions when naming newly found TRIM genes in non-human species. We assigned diapsid TRIM genes without orthology to human TRIM genes as TRIM2##. The start at TRIM200 was to ensure there was no overlap in names between current TRIM annotations. TRIM genes with close paralogs should be given the same name, with numbers identifying them. For example, we named the TRIM19 paralogs TRIM19.1 and TRIM19.2. This will help inform the ancestry of these genes and preserve numbers for new genes. With the increasing number of non-model organism genomes available on public databases, it has become much easier to compare complex gene families between species.

We found 57 TRIM genes in the duck, with evidence that one of these genes is a TRIM-like pseudogene. We found key differences between the duck and chicken TRIM gene repertoires that highlight the complex and understudied mechanics of TRIM gene evolution. We show evidence that TRIM proteins in the C-IV family are rapidly changing in avian species, with important differences in the organization of MHC-linked genes and TRIM25 locus between species. Remarkably, these two regions are adjacent in a marsupial and reptiles, suggesting that they were linked in the ancestral vertebrate MHC region. This is the first major study in TRIM gene classification in birds, where species of birds were compared. As more complete genomes in other birds and vertebrates are sequenced and become available, we can better trace the expansions and deletions of TRIM genes in the vertebrate lineages.

Data availability statement

The original contributions presented in the study are included in the article/**Supplementary Materials**. Further inquiries can be directed to the corresponding author.

Author contributions

LC and KM conceived of the study. LC performed the computational analysis and originally drafted and edited the manuscript. RP completed the RBH analysis and provided scripts

for analysis, provided feedback and editing. KM provided funding for the project, feedback on analyses and edited the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This research was funded by the Natural Sciences and Engineering Research Council Discovery grant NSERC RGPIN-2020-04668 (KM).

Acknowledgments

We performed bioinformatic analyses using computational resources of the Digital Research Alliance of Canada's research computing environment (formerly Compute Canada). LC was supported in part by the QEII Scholarship, Alberta Innovates Technology Futures Scholarship, and the Alberta Graduate Excellence Scholarship.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2023.1220081/full#supplementary-material>

SUPPLEMENTARY FIGURE 1

Flow chart documenting workflow used to generate *de novo* transcriptome and find *Anas platyrhynchos* TRIM genes.

SUPPLEMENTARY FIGURE 2

Multidimensional scaling plot (MDS) of duck TRIM and TRIM-like gene expression in tissues sampled. All RNA-seq libraries were accessed from the NCBI short sequence read archive (SRA), normalized by library size, compared using the TMM method in EdgeR in the RStudio environment and edited for clarity in Adobe Illustrator. Individual libraries were compared using

dimensions 1, 2, 3 and 4. Individual samples were colour coded by tissue that library was sequenced from.

SUPPLEMENTARY TABLE 1

TRIM and TRIM-like genes present in the domestic mallard duck (*Anas platyrhynchos*).

SUPPLEMENTARY TABLE 2

TRIM and TRIM-like genes present in the chicken (*Gallus gallus*).

SUPPLEMENTARY FILE 1

Project numbers, tissue and data included from the NCBI SRA database.

SUPPLEMENTARY FILE 2

Accession numbers of C-IV TRIM proteins from representative mammals, reptiles and birds.

SUPPLEMENTARY FILE 3

Nucleotide sequences of duck TRIM or TRIM-like genes used in this study in fasta format.

SUPPLEMENTARY FILE 4

Amino acid sequences of duck TRIM or TRIM-like genes used in this study in fasta format.

SUPPLEMENTARY FILE 5

Nucleotide and amino acid sequences of chicken TRIM or TRIM-like genes used in this study in fasta format. Nucleotide sequences only for genes pulled from *de novo* chicken transcriptome.

SUPPLEMENTARY FILE 6

Raw read counts for all samples used after library normalization in EdgeR.

SUPPLEMENTARY FILE 7

Relative expression of TRIM or TRIM like genes in ducks.

SUPPLEMENTARY FILE 8

Maximum-likelihood tree of C-IV TRIM proteins from birds, reptiles and mammals, including MHC-linked TRIM genes from reptiles with no known orthologs in birds or mammals.

References

- Berti C, Messali S, Ballabio A, Reymond A, Meroni G. TRIM9 is specifically expressed in the embryonic and adult nervous system. *Mech Dev* (2002) 113(2):159–62. doi: 10.1016/S0925-4773(02)00013-8
- Reymond A, Meroni G, Fantozzi A, Merla G, Cairo S, Luzi L, et al. The tripartite motif family identifies cell compartments. *EMBO J* (2001) 20(9):2140–51. doi: 10.1093/emboj/20.9.2140
- Venuto S, Merla G. E3 ubiquitin ligase TRIM proteins, cell cycle and mitosis. *Cells* (2019) 8(5). doi: 10.3390/cells8050510
- Nisole S, Stoye JP, Saib A. TRIM family proteins: retroviral restriction and antiviral defence. *Nat Rev Microbiol* (2005) 3(10):799–808. doi: 10.1038/nrmicro1248
- Ozato K, Shin DM, Chang TH, Morse HC3rd. TRIM family proteins and their emerging roles in innate immunity. *Nat Rev Immunol* (2008) 8(11):849–60. doi: 10.1038/nri2413
- Rajsbaum R, Garcia-Sastre A, Versteeg GA. TRIMmunity: the roles of the TRIM E3-ubiquitin ligase family in innate antiviral immunity. *J Mol Biol* (2014) 426(6):1265–84. doi: 10.1016/j.jmb.2013.12.005
- Hatakeyama S. TRIM family proteins: roles in autophagy, immunity, and carcinogenesis. *Trends Biochem Sci* (2017) 42(4):297–311. doi: 10.1016/j.tibs.2017.01.002
- Marin I. Origin and diversification of TRIM ubiquitin ligases. *PLoS One* (2012) 7(11):e50030. doi: 10.1371/journal.pone.0050030
- Sardiello M, Cairo S, Fontanella B, Ballabio A, Meroni G. Genomic analysis of the TRIM family reveals two groups of genes with distinct evolutionary properties. *BMC Evol Biol* (2008) 8(1):225. doi: 10.1186/1471-2148-8-225
- Carthagena L, Bergamaschi A, Luna JM, David A, Uchil PD, Margottin-Goguet F, et al. Human TRIM gene expression in response to interferons. *PLoS One* (2009) 4(3):e4894. doi: 10.1371/journal.pone.0004894
- Hatakeyama S. TRIM proteins and cancer. *Nat Rev Cancer* (2011) 11(11):792–804. doi: 10.1038/nrc3139
- Meroni G. Genomics and evolution of the TRIM gene family. *Adv Exp Med Biol* (2012) 770:1–9. doi: 10.1007/978-1-4614-5398-7_1
- Sawyer SL, Wu LI, Emerman M, Malik HS. Positive selection of primate TRIM5alpha identifies a critical species-specific retroviral restriction domain. *Proc Natl Acad Sci USA* (2005) 102(8):2832–7. doi: 10.1073/pnas.0409853102
- Song B, Javanbakht H, Perron M, Park DH, Stremlau M, Sodroski J. Retrovirus restriction by TRIM5alpha variants from old world and new world primates. *J Virol* (2005) 79(7):3930–7. doi: 10.1128/JVI.79.7.3930-3937.2005
- Sawyer SL, Emerman M, Malik HS. Discordant evolution of the adjacent antiretroviral genes TRIM22 and TRIM5 in mammals. *PLoS Pathog* (2007) 3(12):e197. doi: 10.1371/journal.ppat.0030197
- Metzger MB, Pruneda JN, Klevit RE, Weissman AM. Ring-type E3 ligases: master manipulators of E2 ubiquitin-conjugating enzymes and ubiquitination. *Biochim Biophys Acta* (2014) 1843(1):47–60. doi: 10.1016/j.bbamcr.2013.05.026
- Joazeiro CA, Weissman AM. Ring finger proteins: mediators of ubiquitin ligase activity. *Cell* (2000) 102(5):549–52. doi: 10.1016/S0092-8674(00)00077-5
- Bell JL, Malyukova A, Holien JK, Koach J, Parker MW, Kavallaris M, et al. TRIM16 acts as an E3 ubiquitin ligase and can heterodimerize with other TRIM family members. *PLoS One* (2012) 7(5):e37470. doi: 10.1371/journal.pone.0037470
- Zheng Q, Hou J, Zhou Y, Yang Y, Xie B, Cao X. Siglec1 suppresses antiviral innate immune response by inducing tbk1 degradation via the ubiquitin ligase TRIM27. *Cell Res* (2015) 25(10):1121–36. doi: 10.1038/cr.2015.108
- Han X, Du H, Massiah MA. Detection and characterization of the in vitro E3 ligase activity of the human mid1 protein. *J Mol Biol* (2011) 407(4):505–20. doi: 10.1016/j.jmb.2011.01.048
- Massiah MA, Simmons BN, Short KM, Cox TC. Solution structure of the rbcc/TRIM B-box1 domain of human mid1: B-box with a ring. *J Mol Biol* (2006) 358(2):532–45. doi: 10.1016/j.jmb.2006.02.009
- Massiah MA. *Zinc-binding B-Box Domains with Ring Folds Serve Critical Roles in the Protein Ubiquitination Pathways in Plants and Animals*. Summers M, editor. London: IntechOpen (2019).
- Wagner JM, Roganowicz MD, Skorupka K, Alam SL, Christensen D, Doss G, et al. Mechanism of B-Box 2 domain-Mediated higher-Order assembly of the retroviral restriction factor TRIM5alpha. *Elife* (2016) 5. doi: 10.7554/eLife.16309
- Li Y, Ma X, Chen Z, Wu H, Wang P, Wu W, et al. B1 oligomerization regulates PML nuclear body biogenesis and leukemogenesis. *Nat Commun* (2019) 10(1):3789. doi: 10.1038/s41467-019-11746-0
- Sanchez JG, Okreglicka K, Chandrasekaran V, Welker JM, Sundquist WI, Pornillos O. The tripartite motif coiled-coil is an elongated antiparallel hairpin dimer. *Proc Natl Acad Sci USA* (2014) 111(7):2494–9. doi: 10.1073/pnas.1318962111
- Weinert C, Morger D, Djekic A, Grütter MG, Mittl PRE. Crystal structure of TRIM20 C-terminal coiled-coil/B30.2 fragment: implications for the recognition of higher order oligomers. *Sci Rep* (2015) 5(1):10819. doi: 10.1038/srep10819
- James LC, Keeble AH, Khan Z, Rhodes DA, Trowsdale J. Structural basis for prySPRY-mediated tripartite motif (TRIM) protein function. *Proc Natl Acad Sci* (2007) 104(15):6200. doi: 10.1073/pnas.0609174104
- D'Cruz AA, Babon JJ, Norton RS, Nicola NA, Nicholson SE. Structure and function of the SPRY/B30.2 domain proteins involved in innate immunity. *Protein Sci* (2013) 22(1):1–10. doi: 10.1002/pro.2185
- Choudhury NR, Heikel G, Trubitsyna M, Kubik P, Nowak JS, Webb S, et al. Rna-binding activity of TRIM25 is mediated by its PRY/SPRY domain and is required for ubiquitination. *BMC Biol* (2017) 15(1):105. doi: 10.1186/s12915-017-0444-9
- Wu X, Wang J, Wang S, Wu F, Chen Z, Li C, et al. Inhibition of influenza A virus replication by TRIM14 via its multifaceted protein-protein interaction with np. *Front Microbiol* (2019) 10:344(FEB). doi: 10.3389/fmicb.2019.00344
- Watanabe M, Hatakeyama S. TRIM proteins and diseases. *J Biochem* (2017) 161(2):135–44. doi: 10.1093/jb/mvw087
- Short KM, Cox TC. Subclassification of the rbcc/TRIM superfamily reveals a novel motif necessary for microtubule binding*. *J Biol Chem* (2006) 281(13):8970–80. doi: 10.1074/jbc.M512755200
- Rhodes DA, de Bono B, Trowsdale J. Relationship between SPRY and B30.2 protein domains. Evolution of a component of immune defence? *Immunology* (2005) 116(4):411–7. doi: 10.1111/j.1365-2567.2005.02248.x
- Versteeg GA, Rajsbaum R, Sánchez-Aparicio MT, Maestre AM, Valdiviezo J, Shi M, et al. The E3-ligase TRIM family of proteins regulates signaling pathways triggered by innate immune pattern-recognition receptors. *Immunity* (2013) 38(2):384–98. doi: 10.1016/j.immuni.2012.11.013

35. Meyer M, Gaudieri S, Rhodes DA, Trowsdale J. Cluster of TRIM genes in the human MHC class I region sharing the B30.2 domain. *Tissue Antigens* (2003) 61(1):63–71. doi: 10.1034/j.1399-0039.2003.610105.x
36. Ruby T, Bed'Hom B, Wittzell H, Morin V, Oudin A, Zoorob R. Characterisation of a cluster of TRIM-B30.2 genes in the chicken MHC B locus. *Immunogenetics* (2005) 57(1-2):116–28. doi: 10.1007/s00251-005-0770-x
37. Blaine AH, Miranzo-Navarro D, Campbell LK, Aldridge JR Jr., Webster RG, Magor KE. Duck TRIM27-L enhances MAVS signaling and is absent in chickens and turkeys. *Mol Immunol* (2015) 67(2 Pt B):607–15. doi: 10.1016/j.molimm.2015.07.011
38. Du Pasquier L. Fish 'N' TRIMs. *J Biol* (2009) 8(5):50. doi: 10.1186/jbiol150
39. Boudinot P, van der Aa LM, Jouneau L, Du Pasquier L, Pontarotti P, Briolat V, et al. Origin and evolution of TRIM proteins: new insights from the complete TRIM repertoire of zebrafish and pufferfish. *PLoS One* (2011) 6(7):e22022. doi: 10.1371/journal.pone.0022022
40. Webster RG, Yakhno M, Hinshaw VS, Bean WJ, Murti KG. Intestinal influenza: replication and characterization of influenza viruses in ducks. *Virology* (1978) 84(2):268–78. doi: 10.1016/0042-6822(78)90247-7
41. Webster RG, Bean WJ, Gorman OT, Chambers TM, Kawaoka Y. Evolution and ecology of influenza A viruses. *Microbiol Rev* (1992) 56(1):152–79. doi: 10.1128/mr.56.1.152-179.1992
42. Feng ZQ, Cheng Y, Yang HL, Zhu Q, Yu D, Liu YP. Molecular characterization, tissue distribution and expression analysis of TRIM25 in gallus gallus domesticus. *Gene* (2015) 561(1):138–47. doi: 10.1016/j.gene.2015.02.025
43. Wang S, Yu M, Liu A, Bao Y, Qi X, Gao L, et al. TRIM25 inhibits infectious bursal disease virus replication by targeting vp3 for ubiquitination and degradation. *PLoS Pathog* (2021) 17(9):e1009900. doi: 10.1371/journal.ppat.1009900
44. Zhou J-R, Liu J-H, Li H-M, Zhao Y, Cheng Z, Hou Y-M, et al. Regulatory effects of chicken TRIM25 on the replication of ALV-A and the MDA5-mediated type I interferon response. *Veterinary Res* (2020) 51(1):145. doi: 10.1186/s13567-020-00870-1
45. Rajsbaum R, Albrecht RA, Wang MK, Maharaj NP, Versteeg GA, Nistal-Villán E, et al. Species-specific inhibition of RIG-I ubiquitination and IFN induction by the influenza A virus NS1 protein. *PLoS Pathog* (2012) 8(11):e1003059. doi: 10.1371/journal.ppat.1003059
46. Miranzo-Navarro D, Magor KE. Activation of duck RIG-I by TRIM25 is independent of anchored ubiquitin. *PLoS One* (2014) 9(1):e86968. doi: 10.1371/journal.pone.0086968
47. Kaikai H, Zhao D, Liu Y, Liu Q, Huang X, Yang J, et al. The E3 ubiquitin ligase TRIM25 inhibits tembusu virus replication in vitro. *Front Vet Sci* (2021) 8:722113. doi: 10.3389/fvets.2021.722113
48. Wei Y, Zhou H, Wang A, Sun L, Wang M, Jia R, et al. TRIM25 identification in the Chinese goose: gene structure, tissue expression profiles, and antiviral immune responses in vivo and in vitro. *BioMed Res Int* (2016) 2016:1403984. doi: 10.1155/2016/1403984
49. Gack MU, Shin YC, Joo CH, Urano T, Liang C, Sun L, et al. TRIM25 ring-finger E3 ubiquitin ligase is essential for RIG-I-mediated antiviral activity. *Nature* (2007) 446(7138):916–20. doi: 10.1038/nature05732
50. Li W, Song Y, Du Y, Huang Z, Zhang M, Chen Z, et al. Duck TRIM29 negatively regulates type I IFN production by targeting MAVS. *Front Immunol* (2022) 13:1016214. doi: 10.3389/fimmu.2022.1016214
51. Pan C, Zhao H, Shen L, Sheng J. Molecular characterization and expression pattern of tripartite motif protein 39 in gallus gallus with a complete PRY/SPRY domain. *Int J Mol Sci* (2011) 12(6):3797–809. doi: 10.3390/ijms12063797
52. Wang X, Zhao Y, Li L, Cheng Z, Wang G. Preparation and application of a monoclonal antibody against chicken TRIM62. *Monoclon Antib Immunodiagn Immunother* (2018) 37(3):134–8. doi: 10.1089/mab.2017.0062
53. Li L, Niu D, Yang J, Bi J, Zhang L, Cheng Z, et al. TRIM62 from chicken as a negative regulator of reticuloendotheliosis virus replication. *Front Veterinary Sci* (2020) 7:152. doi: 10.3389/fvets.2020.00152
54. Li L, Feng W, Cheng Z, Yang J, Bi J, Wang X, et al. TRIM62-mediated restriction of avian leukosis virus subgroup J replication is dependent on the SPRY domain. *Poult Sci* (2019) 98(11):6019–25. doi: 10.3382/ps/pez408
55. Zhang J, Hu MM, Wang YY, Shu HB. TRIM32 protein modulates type I interferon induction and cellular antiviral response by targeting mita/sting protein for K63-linked ubiquitination. *J Biol Chem* (2012) 287(34):28646–55. doi: 10.1074/jbc.M112.362608
56. Fu B, Wang L, Ding H, Schwamborn JC, Li S, Dorf ME. TRIM32 senses and restricts influenza A virus by ubiquitination of PB1 polymerase. *PLoS Pathog* (2015) 11(6):e1004960. doi: 10.1371/journal.ppat.1004960
57. Cui H, Liu Y, Huang Y. Roles of TRIM32 in corneal epithelial cells after infection with herpes simplex virus. *Cell Physiol Biochem* (2017) 43(2):801–11. doi: 10.1159/000481563
58. Wu S, Zhang J, Xue Q, Liu J, Huang B, He Z, et al. Duck TRIM32 functions in IFN- β signaling against the infection of H5N6 highly pathogenic avian influenza virus. *Front Immunol* (2020) 11:377. doi: 10.3389/fimmu.2020.00377
59. Li T, Hu X, Zhang T, Song X, Zhang H, Dai N, et al. Cloning, analysis, and anti-duck tembusu virus innate immune response of cherry valley duck tripartite motif-containing 32. *Poult Sci* (2021) 100(5):101048. doi: 10.1016/j.psj.2021.101048
60. Andrews S. FastQC. A quality control tool for high throughput sequence data. (2010). Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastq>
61. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for illumina sequence data. *Bioinformatics* (2014) 30(15):2114–20. doi: 10.1093/bioinformatics/btu170
62. Birol I, Jackman SD, Nielsen CB, Qian JQ, Varhol R, Stazyk G, et al. De novo transcriptome assembly with abyss. *Bioinformatics* (2009) 25(21):2872–7. doi: 10.1093/bioinformatics/btp367
63. Huang X, Madan A. Cap3: A DNA sequence assembly program. *Genome Res* (1999) 9(9):868–77. doi: 10.1101/gr.9.9.868
64. Li W, Godzik A. Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* (2006) 22(13):1658–9. doi: 10.1093/bioinformatics/btl158
65. Simao FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* (2015) 31(19):3210–2. doi: 10.1093/bioinformatics/btv351
66. Rice P, Longden I, Bleasby A. EMBOSS: the European molecular biology open software suite. *Trends Genet* (2000) 16(6):276–7. doi: 10.1016/s0168-9525(00)02024-2
67. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. *BMC Bioinf* (2009) 10:421. doi: 10.1186/1471-2105-10-421
68. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* (2011) 7:539. doi: 10.1038/msb.2011.75
69. Letunic I, Doerks T, Bork P. SMART: recent updates, new developments and status in 2015. *Nucleic Acids Res* (2015) 43(Database issue):D257–60. doi: 10.1093/nar/gku949
70. Finn RD, Clements J, Eddy SR. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res* (2011) 39(Web Server issue):W29–37. doi: 10.1093/nar/gkr367
71. UniProt Consortium T. UNIPROT: the universal protein knowledgebase. *Nucleic Acids Res* (2018) 46(5):2699–. doi: 10.1093/nar/gky092
72. Li J, Zhang J, Liu J, Zhou Y, Cai C, Xu L, et al. A new duck genome reveals conserved and convergently evolved chromosome architectures of birds and mammals. *GigaScience* (2021) 10(1):giaa142. doi: 10.1093/gigascience/giaa142
73. Gel B, Serra E. Karyoploter: an R/bioconductor package to plot customizable genomes displaying arbitrary data. *Bioinformatics* (2017) 33(19):3088–90. doi: 10.1093/bioinformatics/btx346
74. RStudio Team. *Rstudio: Integrated Development Environment for R*. Boston, MA: RStudio, Inc (2020).
75. Nurk S, Koren S, Rhie A, Rautiainen M, Bizakdzé AV, Mikheenko A, et al. The complete sequence of a human genome. *Science* (2022) 376(6588):44–53. doi: 10.1126/science.abc6987
76. Brandies PA, Tang S, Johnson RSP, Hogg CJ, Belov K. The first Antechinus reference genome provides a resource for investigating the genetic basis of semelparity and age-related neuropathologies. *GigaByte* (2020) 2020:gigabyte7. doi: 10.1101/2020.09.21.305722
77. Westfall AK, Telemeco RS, Grizante MB, Waits DS, Clark AD, Simpson DY, et al. A chromosome-level genome assembly for the eastern fence lizard (*Sceloporus undulatus*), a reptile model for physiological and evolutionary ecology. *Gigascience* (2021) 10(10):. doi: 10.1093/gigascience/giab066
78. Mueller RC, Ellström P, Howe K, Uliano-Silva M, Kuo RI, Miedzinska K, et al. A high-quality genome and comparison of short versus long read transcriptome of the palaeartic duck *Aythya fuligula* (tufted duck). *Gigascience* (2021) 10(12):giab081. doi: 10.1093/gigascience/giab081
79. Foster Y, Dutoit L, Grosser S, Dussex N, Foster BJ, Dodds KG, et al. Genomic signatures of inbreeding in a critically endangered parrot, the kakāpō. *G3 (Bethesda)* (2021) 11(11). doi: 10.1093/g3journal/jkab307
80. Katoh K, Rozewicki J, Yamada KD. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Briefings Bioinf* (2019) 20(4):1160–6. doi: 10.1093/bib/bbx108
81. Rose R, Golosova O, Sukhomlinov D, Tiunov A, Prosperi M. Flexible design of multiple metagenomics classification pipelines with UGENE. *Bioinformatics* (2019) 35(11):1963–5. doi: 10.1093/bioinformatics/bty901
82. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermini LS. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods* (2017) 14(6):587–9. doi: 10.1038/nmeth.4285
83. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* (2014) 32(1):268–74. doi: 10.1093/molbev/msu300
84. Minh BQ, Nguyen MAT, von Haeseler A. Ultrafast approximation for phylogenetic bootstrap. *Mol Biol Evol* (2013) 30(5):1188–95. doi: 10.1093/molbev/mst024
85. Bodenhofer U, Bonatesta E, Horejš-Kainrath C, Hochreiter S. MSA: an R package for multiple sequence alignment. *Bioinformatics* (2015) 31(24):3997–9. doi: 10.1093/bioinformatics/btv494

86. Charif D, Lobry JR. Seqnr 1.0-2: A contributed package to the R project for statistical computing devoted to biological sequences retrieval and analysis. In: Bastolla U, Porto M, Roman HE, Vendruscolo M, editors. *Structural Approaches to Sequence Evolution: Molecules, Networks, Populations*. Berlin, Heidelberg: Springer Berlin Heidelberg (2007). p. 207–32.
87. Caery V. *LLaRG: Rbg: An Interface to the Boost Graph Library* (2021). Available at: <http://www.bioconductor.org>.
88. Wickham H. *ggplot2: Elegant Graphics for Data Analysis*. New York: Singer-Verlag (2016).
89. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* (2012) 9(4):357–9. doi: 10.1038/nmeth.1923
90. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* (2014) 30(7):923–30. doi: 10.1093/bioinformatics/btt656
91. Robinson MD, McCarthy DJ, Smyth GK. EdgeR: A bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* (2010) 26(1):139–40. doi: 10.1093/bioinformatics/btp616
92. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci* (1998) 95(25):14863. doi: 10.1073/pnas.95.25.14863
93. Gregory R, Warnes BB, Bonebakker L, Gentleman R, Huber W, Liaw A, et al. *Gplots: Various R Programming Tools for Plotting Data*. (2022). [online] Available at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>
94. Shiina T, Briles WE, Goto RM, Hosomichi K, Yanagiya K, Shimizu S, et al. Extended gene map reveals tripartite motif, C-type lectin, and Ig superfamily type genes within a subregion of the chicken MHC-B affecting infectious disease. *J Immunol* (2007) 178(11):7162–72. doi: 10.4049/jimmunol.178.11.7162
95. Wójcik E, Smalec E. Description of the mallard duck (*Anas platyrhynchos*) karyotype. *Folia Biol (Krakow)* (2007) 55(3-4):115–20. doi: 10.3409/173491607781492588
96. Sun M, Li S, Yu K, Xiang J, Li F. An E3 ubiquitin ligase TRIM9 is involved in wssv infection via interaction with β -trcp. *Dev Comp Immunol* (2019) 97:57–63. doi: 10.1016/j.dci.2019.03.014
97. Sun M, Li S, Jin S, Li X, Xiang J, Li F. A novel TRIM9 protein promotes NF- κ B activation through interacting with Ivmid in shrimp during WSSV infection. *Front Immunol* (2022) 13. doi: 10.3389/fimmu.2022.819881
98. Tanji K, Kamitani T, Mori F, Kakita A, Takahashi H, Wakabayashi K. TRIM9, a novel brain-specific E3 ubiquitin ligase, is repressed in the brain of parkinson's disease and dementia with lewy bodies. *Neurobiol Dis* (2010) 38(2):210–8. doi: 10.1016/j.nbd.2010.01.007
99. Smith J, Smith N, Yu L, Paton IR, Gutowska MW, Forrest HL, et al. A comparative analysis of host responses to avian influenza infection in ducks and chickens highlights a role for the interferon-induced transmembrane proteins in viral resistance. *BMC Genomics* (2015) 16(1):574. doi: 10.1186/s12864-015-1778-8
100. Morris KM, Hindle MM, Boitard S, Burt DW, Danner AF, Eory L, et al. The quail genome: insights into social behaviour, seasonal biology and infectious disease response. *BMC Biol* (2020) 18(1):14. doi: 10.1186/s12915-020-0743-4
101. Krchliková V, Hron T, Těšický M, Li T, Ungrová L, Hejnar J, et al. Dynamic evolution of avian RNA virus sensors: repeated loss of RIG-I and RIPLET. *Viruses* (2022) 15(1). doi: 10.3390/v15010003
102. Zhou Z, Jia X, Xue Q, Dou Z, Ma Y, Zhao Z, et al. TRIM14 is a mitochondrial adaptor that facilitates retinoic acid-inducible gene-1-like receptor-mediated innate immune response. *Proc Natl Acad Sci USA* (2014) 111(2):E245–54. doi: 10.1073/pnas.1316941111
103. Tan P, He L, Cui J, Qian C, Cao X, Lin M, et al. Assembly of the WHIP-TRIM14-PPP6C mitochondrial complex promotes RIG-I-mediated antiviral signaling. *Mol Cell* (2017) 68(2):293–307. doi: 10.1016/j.molcel.2017.09.035
104. Kato K, Ahmad S, Zhu Z, Young JM, Mu X, Park S, et al. Structural analysis of RIG-I-like receptors reveals ancient rules of engagement between diverse RNA helicases and TRIM ubiquitin ligases. *Mol Cell* (2021) 81(3):599–613. doi: 10.1016/j.molcel.2020.11.047
105. Jia X, Zhao C, Zhao W. Emerging roles of MHC class I region-encoded E3 ubiquitin ligases in innate immunity. *Front Immunol* (2021) 12:687102. doi: 10.3389/fimmu.2021.687102
106. Hattlmann CJ, Kelly JN, Barr SD. TRIM22: A diverse and dynamic antiviral protein. *Mol Biol Int* (2012) 2012:153415. doi: 10.1155/2012/153415
107. Fernandes AP, Águeda-Pinto A, Pinheiro A, Rebelo H, Esteves PJ. Evolution of TRIM5 and TRIM22 in bats reveals a complex duplication process. *Viruses* (2022) 14(2). doi: 10.3390/v14020345
108. Luo Z-X, Yuan C-X, Meng Q-J, Ji Q. A jurassic eutherian mammal and divergence of marsupials and placentals. *Nature* (2011) 476(7361):442–5. doi: 10.1038/nature10291
109. Lind AL, Lai YYY, Mostovoy Y, Holloway AK, Iannucci A, Mak ACY, et al. Genome of the komodo dragon reveals adaptations in the cardiovascular and chemosensory systems of monitor lizards. *Nat Ecol Evol* (2019) 3(8):1241–52. doi: 10.1038/s41559-019-0945-8
110. Magor KE, Miranzo Navarro D, Barber MR, Petkau K, Fleming-Canepa X, Blyth GA, et al. Defense genes missing from the flight division. *Dev Comp Immunol* (2013) 41(3):377–88. doi: 10.1016/j.dci.2013.04.010
111. Keiser MJ, Roth BL, Armbruster BN, Ernsberger P, Irwin JJ, Shoichet BK. Relating protein pharmacology by ligand chemistry. *Nat Biotechnol* (2007) 25(2):197–206. doi: 10.1038/nbt1284
112. Chaves LD, Krueh SB, Reed KM. Defining the turkey MHC: sequence and genes of the B locus. *J Immunol* (2009) 183(10):6530–7. doi: 10.4049/jimmunol.0901310
113. van der Aa LM, Levraud J-P, Yahmi M, Lauret E, Briolat V, Herbomel P, et al. A large new subset of TRIM genes highly diversified by duplication and positive selection in teleost fish. *BMC Biol* (2009) 7(1):7. doi: 10.1186/1741-7007-7-7
114. Suurväli J, Garraway CJ, Boudinot P. Recurrent expansions of B30.2-associated immune receptor families in fish. *Immunogenetics* (2022) 74(1):129–47. doi: 10.1007/s00251-021-01235-4
115. Qin B, Xiao T, Ding C, Deng Y, Lv Z, Su J. Genome-wide identification and expression analysis of potential antiviral tripartite motif proteins (TRIMs) in grass carp (*Ctenopharyngodon idella*). *Biol (Basel)* (2021) 10(12). doi: 10.3390/biology10121252
116. Glasauer SM, Neuhauss SC. Whole-genome duplication in teleost fishes and its evolutionary consequences. *Mol Genet Genomics* (2014) 289(6):1045–60. doi: 10.1007/s00438-014-0889-2
117. Kaufman J. Innate immune genes of the chicken MHC and related regions. *Immunogenetics* (2022) 74(1):167–77. doi: 10.1007/s00251-021-01229-2
118. Nei M, Gu X, Sitnikova T. Evolution by the birth-and-death process in multigene families of the vertebrate immune system. *Proc Natl Acad Sci USA* (1997) 94(15):7799–806. doi: 10.1073/pnas.94.15.7799
119. Hron T, Pajer P, Paces J, Bartunek P, Elleder D. Hidden genes in birds. *Genome Biol* (2015) 16(1):164. doi: 10.1186/s13059-015-0724-z
120. Qiu S, Liu H, Jian Z, Fan Z, Liu S, Xing J, et al. Characterization of the primate TRIM gene family reveals the recent evolution in primates. *Mol Genet Genomics* (2020) 295(5):1281–94. doi: 10.1007/s00438-020-01698-2
121. Stremelau M, Owens CM, Perron MJ, Kiessling M, Autissier P, Sodroski J. The cytoplasmic body component TRIM5alpha restricts HIV-1 infection in old world monkeys. *Nature* (2004) 427(6977):848–53. doi: 10.1038/nature02343
122. Barr SD, Smiley JR, Bushman FD. The interferon response inhibits HIV particle production by induction of TRIM22. *PLoS Pathog* (2008) 4(2):e1000007. doi: 10.1371/journal.ppat.1000007
123. Koch LK, Cunze S, Kochmann J, Klimpel S. Bats as putative Zaire ebolavirus reservoir hosts and their habitat suitability in Africa. *Sci Rep* (2020) 10(1):14268. doi: 10.1038/s41598-020-71226-0
124. Latinne A, Hu B, Olival KJ, Zhu G, Zhang L, Li H, et al. Origin and cross-species transmission of bat coronaviruses in China. *Nat Commun* (2020) 11(1):4235. doi: 10.1038/s41467-020-17687-3
125. Wiederkehr MA, Qi W, Schoenbaechler K, Fraefel C, Kubacki J. Virus diversity, abundance, and evolution in three different bat colonies in Switzerland. *Viruses* (2022) 14(9). doi: 10.3390/v14091911
126. Xie J, Tan B, Zhang Y. Positive selection and duplication of bat TRIM family proteins. *Viruses* (2023) 15(4). doi: 10.3390/v15040875
127. Lallemand-Breitenbach V, de Thé H. PML nuclear bodies: from architecture to function. *Curr Opin Cell Biol* (2018) 52:154–61. doi: 10.1016/j.cob.2018.03.011
128. Hoischen C, Monajembashi S, Weisshart K, Hemmerich P. Multimodal light microscopy approaches to reveal structural and functional properties of promyelocytic leukemia nuclear bodies. *Front Oncol* (2018) 8. doi: 10.3389/fonc.2018.00125
129. Hsu KS, Guan BJ, Cheng X, Guan D, Lam M, Hatzoglou M, et al. Translational control of PML contributes to TNF α -induced apoptosis of MCF7 breast cancer cells and decreased angiogenesis in HUVECs. *Cell Death Differ* (2016) 23(3):469–83. doi: 10.1038/cdd.2015.114
130. Kurihara M, Kato K, Sanbo C, Shigenobu S, Ohkawa Y, Fuchigami T, et al. Genomic profiling by ALA-Seq reveals transcriptional regulation by PML bodies through DNMT3A exclusion. *Mol Cell* (2020) 78(3):493–505. doi: 10.1016/j.molcel.2020.04.004
131. Yan H-Y, Wang H-Q, Zhong M, Wu S, Yang L, Li K, et al. PML Suppresses Influenza Virus Replication By Promoting FBXW7 Expression. *Virologica Sin* (2021) 36(5):1154–64. doi: 10.1007/s12250-021-00399-3
132. El Bougrini J, Dianoux L, Chelbi-Alix MK. PML positively regulates interferon gamma signaling. *Biochimie* (2011) 93(3):389–98. doi: 10.1016/j.biochi.2010.11.005
133. El McHichi B, Regad T, Maroui M-A, Rodriguez MS, Aminev A, Gerbaud S, et al. Sumoylation promotes PML degradation during encephalomyocarditis virus infection. *J Virol* (2010) 84(22):11634–45. doi: 10.1128/JVI.01321-10
134. Scherer M, Schilling EM, Stamminger T. The human CMV IE1 protein: an offender of PML nuclear bodies. *Adv Anat Embryol Cell Biol* (2017) 223:77–94. doi: 10.1007/978-3-319-53168-7_4
135. Abi Rached L, McDermott MF, Pontarotti P. The MHC big bang. *Immunol Rev* (1999) 167:33–44. doi: 10.1111/j.1600-065x.1999.tb01380.x
136. Flajnik MF, Kasahara M. Comparative genomics of the MHC: glimpses into the evolution of the adaptive immune system. *Immunity* (2001) 15(3):351–62. doi: 10.1016/s1074-7613(01)00198-4
137. Kaufman J, Völk H, Wallny HJ. A "Minimal essential MHC" and an "Unrecognized MHC": two extremes in selection for polymorphism. *Immunol Rev* (1995) 143:63–88. doi: 10.1111/j.1600-065x.1995.tb00670.x

138. Kaufman J, Milne S, Göbel TWF, Walker BA, Jacob JP, Auffray C, et al. The chicken B locus is a minimal essential major histocompatibility complex. *Nature* (1999) 401(6756):923–5. doi: 10.1038/44856
139. Patil G, Zhao M, Song K, Hao W, Bouchereau D, Wang L, et al. TRIM41-mediated ubiquitination of nucleoprotein limits influenza A virus infection. *J Virol* (2018) 92(16). doi: 10.1128/JVI.00905-18
140. Patil G, Xu L, Wu Y, Song K, Hao W, Hua F, et al. TRIM41-mediated ubiquitination of nucleoprotein limits vesicular stomatitis virus infection. *Viruses* (2020) 12(2). doi: 10.3390/v12020131
141. Yu Z, Li X, Yang M, Huang J, Fang Q, Jia J, et al. TRIM41 is required to innate antiviral response by polyubiquitinating BCL10 and recruiting NEMO. *Signal Transduction Targeted Ther* (2021) 6(1):90. doi: 10.1038/s41392-021-00477-8
142. Gack MU, Albrecht RA, Urano T, Inn KS, Huang IC, Camero E, et al. Influenza A virus ns1 targets the ubiquitin ligase TRIM25 to evade recognition by the host viral RNA sensor RIG-I. *Cell Host Microbe* (2009) 5(5):439–49. doi: 10.1016/j.chom.2009.04.006
143. Hu Y, Li W, Gao T, Cui Y, Jin Y, Li P, et al. The severe acute respiratory syndrome coronavirus nucleocapsid inhibits type I interferon production by interfering with TRIM25-mediated RIG-I ubiquitination. *J Virol* (2017) 91(8). doi: 10.1128/jvi.02143-16
144. Manokaran G, Finol E, Wang C, Gunaratne J, Bahl J, Ong EZ, et al. Dengue subgenomic RNA binds TRIM25 to inhibit interferon expression for epidemiological fitness. *Science* (2015) 350(6257):217–21. doi: 10.1126/science.aab3369
145. Pauli EK, Chan YK, Davis ME, Gableske S, Wang MK, Feister KF, et al. The ubiquitin-specific protease USP15 promotes RIG-I-mediated antiviral signaling by deubiquitylating TRIM25. *Sci Signal* (2014) 7(307):ra3. doi: 10.1126/scisignal.2004577
146. Oshiumi H, Miyashita M, Matsumoto M, Seya T. A distinct role of RIPLET-mediated K63-linked polyubiquitination of the RIG-I repressor domain in human antiviral innate immune responses. *PLoS Pathog* (2013) 9(8):e1003533. doi: 10.1371/journal.ppat.1003533
147. Barber MR, Aldridge JR Jr., Webster RG, Magor KE. Association of RIG-I with innate immunity of ducks to influenza. *Proc Natl Acad Sci USA* (2010) 107(13):5913–8. doi: 10.1073/pnas.1001755107
148. Oshiumi H. Recent advances and contradictions in the study of the individual roles of ubiquitin ligases that regulate RIG-I-like receptor-mediated antiviral innate immune responses. *Front Immunol* (2020) 11:1296. doi: 10.3389/fimmu.2020.01296
149. Cadena C, Ahmad S, Xavier A, Willemsen J, Park S, Park JW, et al. Ubiquitin-dependent and -independent roles of E3 ligase riplet in innate immunity. *Cell* (2019) 177(5):1187–200.e16. doi: 10.1016/j.cell.2019.03.017
150. Hayman TJ, Hsu AC, Kolesnik TB, Dagley LF, Willemsen J, Tate MD, et al. RIPLET, and not TRIM25, is required for endogenous RIG-I-dependent antiviral responses. *Immunol Cell Biol* (2019) 97(9):840–52. doi: 10.1111/imcb.12284
151. Perera S, Mankoo B, Gautel M. Developmental regulation of MURF E3 ubiquitin ligases in skeletal muscle. *J Muscle Res Cell Motil* (2012) 33(2):107–22. doi: 10.1007/s10974-012-9288-7
152. Uhlen M, Fagerberg L, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Proteomics. Tissue-based map of the human proteome. *Science* (2015) 347(6220):1260419. doi: 10.1126/science.1260419
153. van Beuningen SFB, Will L, Harterink M, Chazeau A, van Battum EY, Frias CP, et al. TRIM46 controls neuronal polarity and axon specification by driving the formation of parallel microtubule arrays. *Neuron* (2015) 88(6):1208–26. doi: 10.1016/j.neuron.2015.11.012
154. Boyer NP, Monkiewicz C, Menon S, Moy SS, Gupton SL. Mammalian TRIM67 functions in brain development and behavior. *eNeuro* (2018) 5(3). doi: 10.1523/eneuro.0186-18.2018
155. Tan G, Xu F, Song H, Yuan Y, Xiao Q, Ma F, et al. Identification of TRIM14 as a type I IFN-stimulated gene controlling hepatitis B virus replication by targeting HBX. *Front Immunol* (2018) 9:1872. doi: 10.3389/fimmu.2018.01872
156. Taubenberger JK, Kash JC. Influenza virus evolution, host adaptation, and pandemic formation. *Cell Host Microbe* (2010) 7(6):440–51. doi: 10.1016/j.chom.2010.05.009
157. Daoust PY, Kibenge FS, Fouchier RA, van de Bildt MW, van Riel D, Kuiken T. Replication of low pathogenic avian influenza virus in naturally infected mallard ducks (*Anas platyrhynchos*) causes no morphologic lesions. *J Wildl Dis* (2011) 47(2):401–9. doi: 10.7589/0090-3558-47.2.401
158. Sturm-Ramirez KM, Ellis T, Bousfield B, Bissett L, Dyrting K, Rehg JE, et al. Reemerging H5N1 influenza viruses in Hong Kong in 2002 are highly pathogenic to ducks. *J Virol* (2004) 78(9):4892–901. doi: 10.1128/jvi.78.9.4892-4901.2004
159. Haider N, Sturm-Ramirez K, Khan SU, Rahman MZ, Sarkar S, Poh MK, et al. Unusually high mortality in waterfowl caused by highly pathogenic avian influenza A (H5N1) in Bangladesh. *Transbound Emerg Dis* (2017) 64(1):144–56. doi: 10.1111/tbed.12354
160. Bingham J, Green DJ, Lowther S, Klippel J, Burggraaf S, Anderson DE, et al. Infection studies with two highly pathogenic avian influenza strains (Vietnamese and Indonesian) in Pekin Ducks (*Anas platyrhynchos*), with particular reference to clinical disease, tissue tropism and viral shedding. *Avian Pathol* (2009) 38(4):267–78. doi: 10.1080/03079450903055371
161. Niu Y, Ma H, Ding Y, Li Z, Sun Y, Li M, et al. The pathogenicity of duck Hepatitis A virus types 1 and 3 on ducklings. *Poultry Sci* (2019) 98(12):6333–9. doi: 10.3382/ps/pez455
162. Hisham I, Ellakany HF, Selim AA, Abdalla MAM, Zain El-Abideen MA, Kilany WH, et al. Comparative pathogenicity of duck Hepatitis A virus-1 isolates in experimentally infected Pekin and Muscovy ducklings. *Front Veterinary Sci* (2020) 7. doi: 10.3382/ps/pez455
163. Liu R, Shi S, Huang Y, Chen Z, Chen C, Cheng L, et al. Comparative pathogenicity of different subtypes of duck hepatitis A virus in Pekin ducklings. *Vet Microbiol* (2019) 228:181–7. doi: 10.1016/j.vetmic.2018.11.030
164. Di Pietro A, Kajaste-Rudnitski A, Oteiza A, Nicora L, Towers GJ, Mechti N, et al. TRIM22 inhibits influenza A virus infection by targeting the viral nucleoprotein for degradation. *J Virol* (2013) 87(8):4523–33. doi: 10.1128/JVI.02548-12
165. Rajsbaum R, Versteeg Gijs A, Schmid S, Maestre Ana M, Belicha-Villanueva A, Martínez-Romero C, et al. Unanchored K48-linked polyubiquitin synthesized by the E3-ubiquitin ligase TRIM6 stimulates the Interferon-IKKε Kinase-mediated antiviral response. *Immunity* (2014) 40(6):880–95. doi: 10.1016/j.immuni.2014.04.018
166. Samukawa S, Yoshimi R, Kirino Y, Nakajima H. The PRY/SPRY domain of ppyrin/TRIM20 interacts with β(2)-microglobulin to promote inflammasome formation. *Sci Rep* (2021) 11(1):23613. doi: 10.1038/s41598-021-03073-6
167. Yu JW, Wu J, Zhang Z, Datta P, Ibrahim I, Taniguchi S, et al. Cryopyrin and PYRIN activate CASPASE-1, but not NF-κβ, via Asc Oligomerization. *Cell Death Differentiation* (2006) 13(2):236–49. doi: 10.1038/sj.cdd.4401734