4 Results: Characterisation and Expression of IFITM3 in Chickens

4.1 Introduction

In humans, IFITM1, 2, and 3 are expressed in a wide range of tissues, whilst IFITM5 expression is limited to osteoblasts[105]. Mice have orthologues for *IFITM1, 2, 3*, and *5*, and additional IFITM genes, *Ifitm6* and *Ifitm7*[102,109]. Human IFITM1-3 have been shown to restrict a broad range of viruses, including IAV. Although the function of IFITM proteins has been well characterised in human and mouse, little compelling functional data exists for this ISG family in other species.

Avian IAVs represent a continuing threat to human populations both as a source for direct human infection and as a reservoir for IAV genetic variation. These reservoirs provide the conditions for the generation of reassorted IAVs with altered host ranges and pandemic potential[258]. Furthermore, poultry are an important source of both meat and eggs for a large proportion of the world population; current global production of chickens is over 30 billion per annum[259]. Endemic and emerging avian viral pathogens create major challenges for the poultry industry, through loss of productivity and mortality. Currently chicken vaccination programs against infectious bronchitis virus, infectious bursal disease, and Newcastle disease do exist[260]. However vaccination is very expensive on such a large scale and in the case of emerging viral pathogens, these vaccines are not always effective. Therefore, if chickens encode potent intrinsic antiviral factors, like IFITM3, variants with increased activity could be exploited in breeding programs to increase the innate protection of these birds.

Genome analysis of chickens has predicted the existence of two *IFITM* genes, orthologous to human *IFITM10* and *IFITM5*[132]. However, such *in silico* analysis is often confounded by inappropriate identification of pseudogenes and incorrect assignment of orthologues, due to an incomplete knowledge of *IFITM* gene duplication and evolutionary history of this locus during speciation. In such circumstances careful genome analysis of syntenic regions and functional characterisation of genes is required to unambiguously define orthologous genes. Although putative *IFITM* genes have been identified by database searching in many species[103,132] no formal genome analysis or functional assessment of avian *IFITM* genes has been undertaken.

The aims and objectives of this chapter are as follows:

i. Are the *IFITM* genes present in the Red Jungle Fowl chicken genome?
ii. Do chicken IFITM proteins have an antiviral effect?
iii. Do chicken IFITM proteins localise to the same sub-cellular regions as the human orthologues?
iv. Is C-terminal tagging an appropriate way to detect expression of IFITM proteins?
v. Are IFITM proteins transcribed in chicken cells? And does this vary across different tissues?
vi. Does suppression of these proteins *in vitro* affect potential antiviral activity?

4.2 Identifying the Chicken IFITM Locus

The chicken genome (ENSEMBL browser, version 68.2) contains two putative IFITM genes on chromosome 5, the so-called *IFITM5* (ENSGALG00000004239; chromosome 5:1600194-1601763) and *IFITM10* (ENSGALG00000020497; chromosome 5:15244061-15249351). The putative *IFITM5* gene is located next to an uncharacterised gene (ENSGALG00000004243) with which it shares 30 % amino acid identity. Immediately adjacent to this are three sequence gaps whose estimated sizes are 1 kb, 1 kb and 400 bp in the ENSEMBL chicken genome build (v68.2).

Importantly, the putative IFITM gene locus in chickens is flanked by the telomeric beta-1,4-N-acetyl-galactosaminyl transferase 4 (*B4GALNT4*) gene and the centromeric acid trehalase-like 1 (*ATHL1*) gene. The *B4GALNT4* and *ATHL1* genes flank the antiviral *IFITM1, 2, 3 and 5* gene block in mammalian genomes. Sequence similarity searches of the chicken genome (v4.0, NCBI) using TBLASTN analysis and the putative IFITM5 amino acid sequence, revealed several transcripts with high amino acid identity to IFITM5. Additionally, BLAST hits were also identified to putative genes *LOC770612 (*variant 1: XM_001233949.3; variant 2: XM_004941314.1) and *LOC422993* (XM_420925.4), within the locus flanked by *B4GALNT4* and *ATHL1* (Figure 41). A third BLAST hit matched an un-curated gene, "*gene-376074*", which is positioned between *LOC422993* and *IFITM5*. Further analysis of *gene-376074* showed it shared amino acid sequence identity with both *LOC422993* and *LOC770612* genes. Sequence similarity searches of the NCBI chicken EST database suggests *gene-376074* is expressed.

All of the chicken *IFITM* (*chIFITM*) paralogues, like mammalian *IFITMs*, are comprised of two exons and the location of the intron-exon boundary is conserved across all the *chIFITM* genes. Therefore the chicken genome contains an intact *IFITM* locus with four putative *IFITM* genes flanked by the genes *B4GALNT4* and *ATHL1*.

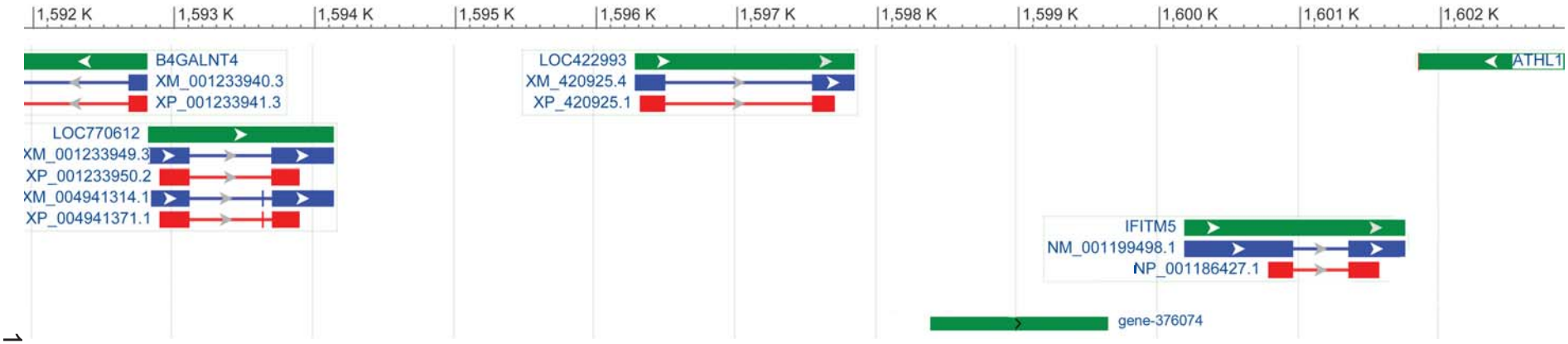

Figure 41: Genome locations of putative IFITM genes in the chicken genome build (NCBI v4)

Three unannotated genes were identified between *ATHL1* and *B4GALNT4* by TBLASTN searches using *IFITM5*; LOC770612 (Chr5:1592832-1594150), LOC422993 (Chr5:1596290-1597850), and un-curated gene-376074 (Chr5:1598390-1599634). Gene regions are shown in green, protein coding regions are shown in red, and mRNA coding regions are shown in blue.

4.3 Annotating the Chicken IFITM Genes

Using genome synteny we ascribed *chIFITM5* as orthologous to mammalian *IFITM5*, gene-376074 as orthologous to *IFITM2*, LOC422993 as orthologous to *IFITM1* and LOC770612 as orthologous to *IFITM3* (Figure 42). Multiple amino acid sequence alignments between the three predicted antiviral *chIFITM* genes and direct orthologues in primate species suggest this assignment is plausible. A number of conserved IFITM-family motifs are present in some of the chicken sequences (Figure 43) and although the chicken sequences differ significantly from the human and chimpanzee orthologues (42 % amino acid identity between chicken and human IFITM3), many amino acids in the CIL domain are conserved. Multiple sequence alignments also revealed important amino acids in the chicken IFITM proteins that help to categorise each sequence as either IFITM1 or IFITM2/3. Tyr20 is conserved in all primate IFITM2 or 3 sequences, and is also present in LOC770612, but none of the other IFITM1 orthologues. This, and the longer N-terminus, further supported our assessment of this gene as an *IFITM2* or *3,* and by synteny it is *IFITM3*. The alignment also revealed that other functionally significant amino acids are conserved in some of the chicken IFITM sequences, including the two cysteines (Cys75-76) in IM1 that are palmitoylation sites in other species[118] and are important for membrane positioning. Phe79, also in IM1, is conserved in LOC770612, which is believed to be important for mediating a physical association between IFITM proteins[261].

However, gene-376074 (*IFITM2*) has a shorter N-terminus than LOC422993 (*IFITM1)* so it could be argued that the labelling of these genes is inverted. Indeed the direction of transcription indicates that a simple inversion of *IFITM1* and *IFITM2,* relative to humans, would lead to this. This uncertainty is reflected in the labelling of Figure 42, the alternative nomenclature is shown in brackets.

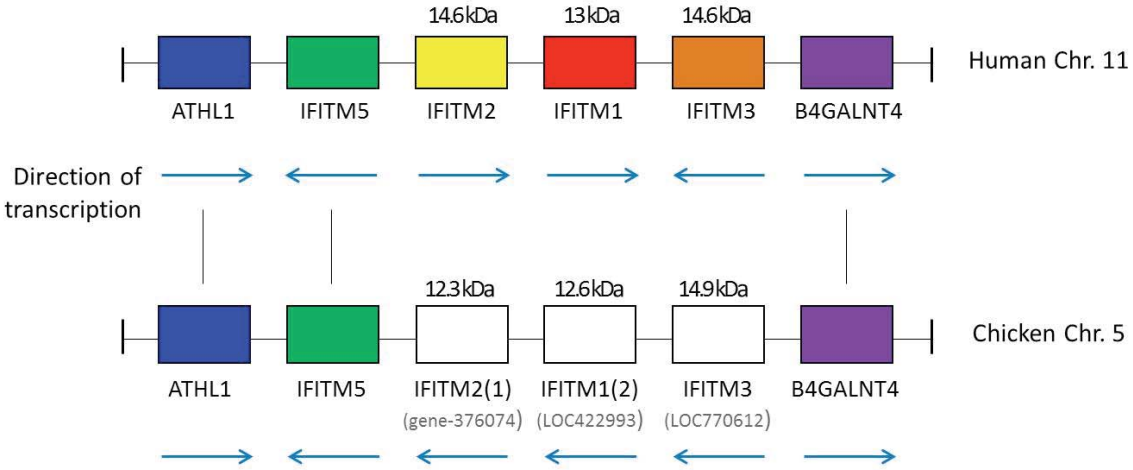

Figure 42: The chicken IFITM locus architecture

The IFITM gene cluster on *Gallus gallus* chromosome 5 is flanked by *ATHL1* and *B4GALNT4*. This region is syntenic with the *IFITM* gene cluster on Human chromosome 11. The orientation change of *chIFITM2* and *chIFITM1* make the assignment of orthology difficult. Therefore the chicken genes are named by gene order and conservation of specific functionally, defined amino acid residues, although the number in brackets reflects the uncertainty in differentiating between *chIFITM2* and *chIFITM1.* Predicted masses are shown above gene block.

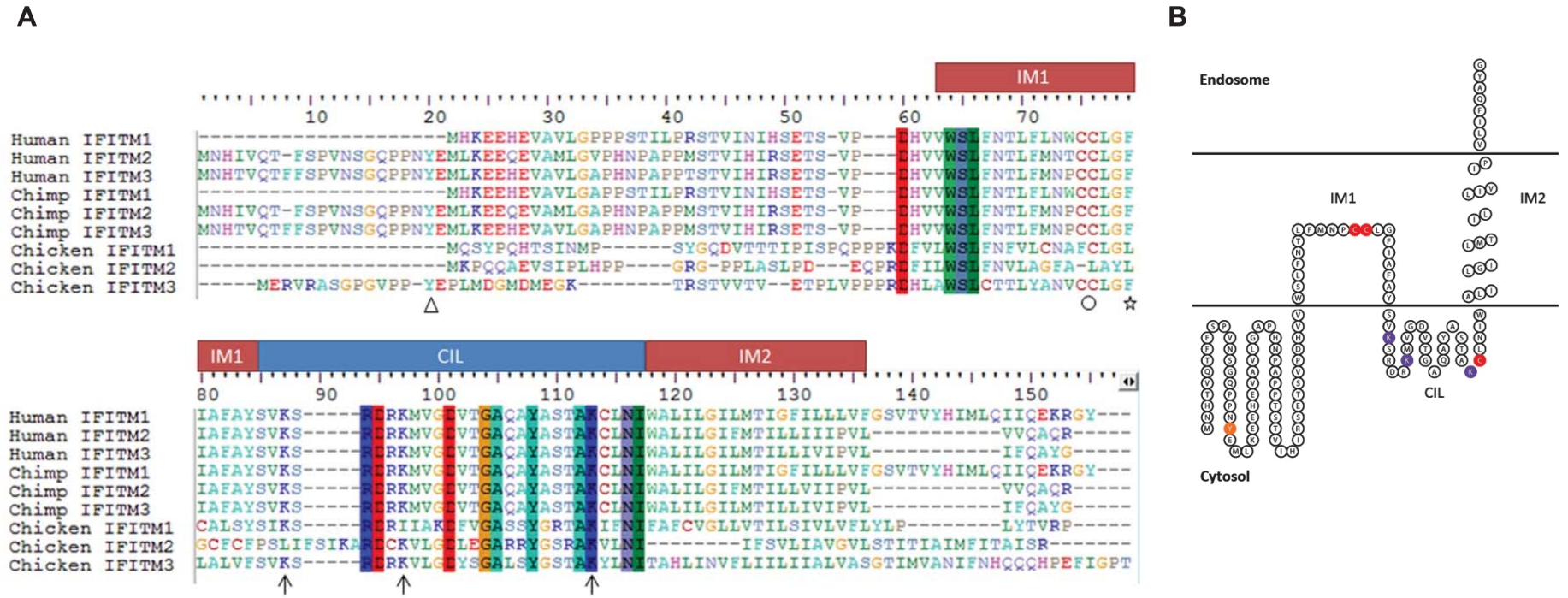

Figure 43: Protein sequence alignments of chicken IFITM sequences and their primate orthologues.

The coloured columns in the sequence alignment (A) show residues that are shared between all nine IFITM sequences from human, chimpanzee and chicken. Significant residues have been highlighted with a symbol below the sequence: Δ – tyrosine; O – double cysteine; ☆ - Phenylalanine important for multimerisation; ↑ - conserved ubiquitinated lysine. IM1 (Intramembrane 1), CIL (conserved intracellular loop), IM2 (Intramembrane 2). (B) shows the predicted structure of human IFITM3 (based on data from Bailey *et al.* [2013]) with the corresponding domains labelled, and important residues highlighted. Figure made using http://www.sacs.ucsf.edu/TOPO2.

4.4 Phylogenetic Analysis of Primate, Rodent, and Chicken IFITMs

A multiple sequence alignment of known primate, rodent, and chicken IFITMs was created and used to infer a phylogenetic tree in order to compare the given nomenclature to the relatedness of the sequences (Figure 44). The tree was created using an alignment of only the conserved intramembrane domains and the conserved intracellular loop (CIL). The N- and C-termini were excluded because their variability made it difficult to determine the homologous characters, which would reduce confidence in the inferred phylogeny. The tree shows that the primate sequences tend to cluster in clades of parologous genes, *i.e.* all the primate *IFITM1s* cluster together, such that human *IFITM1* is more similar to chimp *IFITM1* than to human *IFITM2*. This suggests that gene duplication happened prior to human/chimp speciation. The three chicken sequences cluster together, outside of the main part of the tree, but chicken IFITM2 is basal to the rest of the sequences, unlike the primate sequences where IFITM1 diverges separately, suggesting the nomenclature may be incorrect. However the tree is mid-point rooted and therefore is biased towards placing the sequence with the longest branch length as the out-group, but this could be due to a faster rate of evolution along the branch to chicken IFITM2 rather than an earlier divergence. This is further supported by the branch lengths for the primate IFITM2s being longer than IFITM1s and 3s. Therefore, due to the divergence between the chicken and mammalian orthologues, the sequence data alone is insufficient to confirm whether or not the nomenclature is correct for the chicken *IFITMs*.

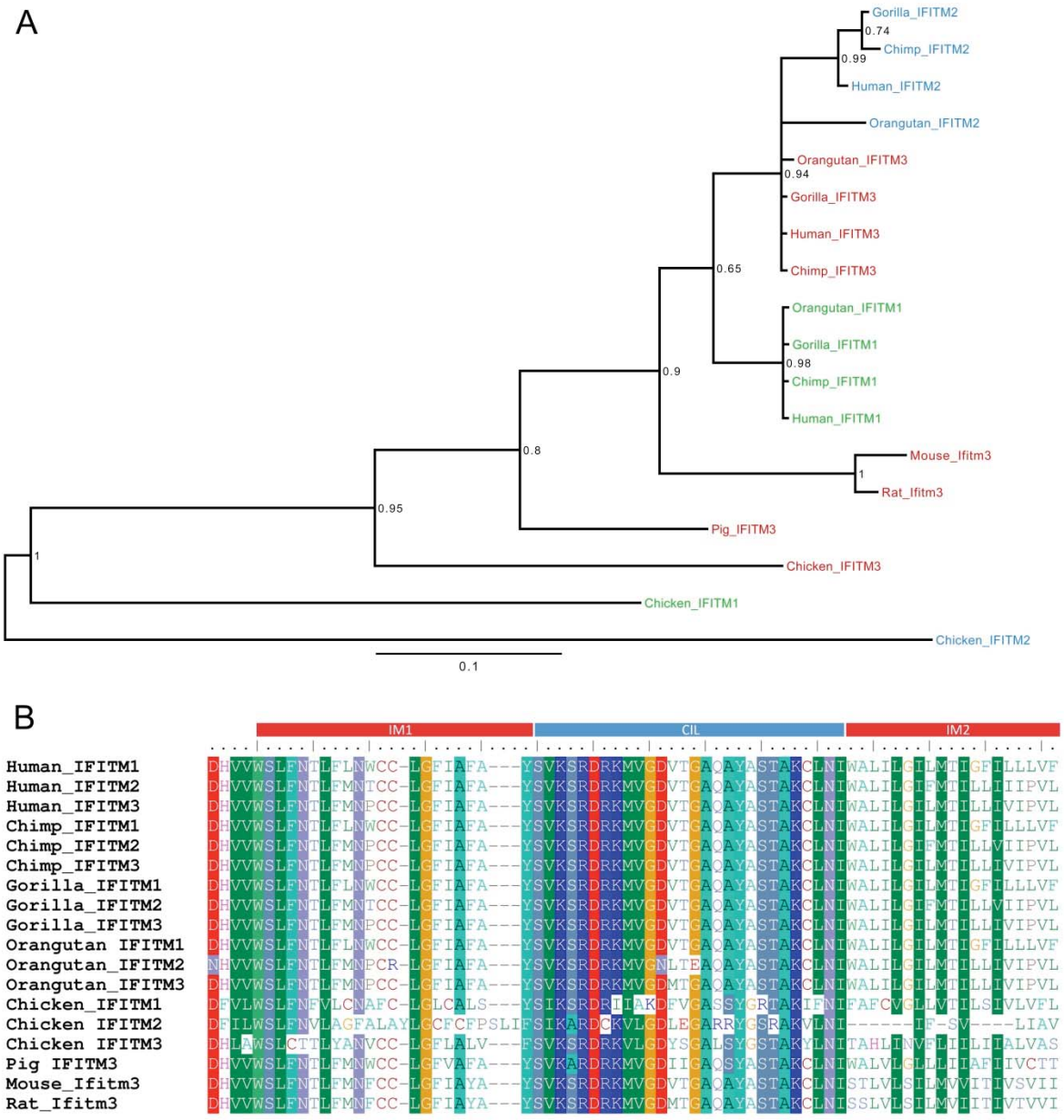

Figure 44: Phylogenetic tree showing relatedness of IFITM sequences

A mid-point rooted Baysian consensus tree (A) was created from an alignment of orthologous IFITM sequences trimmed to a region of high conservation (B). Vertical coloured bars denote conserved regions with a threshold of 85 %. Numbers at each node represent the posterior probability for that clade. The scale bar is in units of substitutions per site. Orthologous genes are grouped by colour.

4.5 Using A549s as a Cell Line for Over-Expression of IFITMs

To explore the function of IFITM proteins *in vitro* and make comparisons between proteins from different species, a reliable cell line low in IFITM expression, and permissible to lentiviral transduction, was required. A549 cells, a cancerous human lung adenocarcinoma cell line, are reported to be low in IFITM expression[1] and are a commonly used type II pulmonary epithelial cell model. Absence of human (hu) *IFITM1*, *2*, and *3* in A549s was assessed by RT-PCR (for primer design see Figure 27).

Total RNA was extracted from $1x10^6$ cells and quantified. 100 ng of RNA was used per RT-PCR reaction, allowing the copies per cell to be estimated by calculating the ng of RNA per cell. Five standards from 10^7 – 10^3 copies were made using plasmids encoding the non-optimised transcripts of human *IFITM1*, *2*, and *3*, to generate standard curves. The quantity of each transcript in A549s was determined relative to the standard curve. RT-PCR showed that without IFN stimulation, A549s transcribe between 1 and 2 copies of *IFITM1* and between 0 and 1 copy of *IFITM3* per cell, but up to 10 copies of *IFITM2* (Table 10). These numbers are in a similar range to IFITM expression in HEK293-Ts.

IFITM3 expression was also not detected in A549s by Western blot. An antibody specific for the NTD of IFITM3 (Abgent) was tested for efficacy against three controls; A549 cells over-expressing full-length wildtype IFITM3 with a C-terminal HA tag, cells over-expressing a human codon-optimised version of full-length IFITM3, and cells over-expressing IFITM3 with a 21 amino acid deletion at the N-terminus (ΔN-21) (Figure 45). The antibody against IFITM3 detected both the ΔN-21 truncated and full-length proteins; however two protein bands were detected by the N-terminal antibody for the full-length proteins (Figure 45). Since a faint band is still detected by the NTD antibody in the ΔN-21 cells, it suggests the antibody is specific for a larger region of the protein. When probed with an anti-HA antibody, only one band (17 kDa) was detected for all the cells tested. Therefore A549s were deemed a suitable cell line to test the function of IFITM proteins in.

Table 10: Quantification of *IFITM* transcripts in human A549s

Gene	Well Name	Biological replicate	Ct (dRn)	Quantity (copies)	Number of cells per input	Copies per cell	Average copies per cell	Standard deviation
IFITM1	A549s	1	27.3	4.50E+05	2.40E+05	1.87	1.57	0.29
		2	27.22	4.68E+05	2.81E+05	1.67		
		3	27.92	3.33E+05	2.82E+05	1.18		
	HEK293-T	1	24.07	2.16E+06	2.05E+05	10.52	7.92	2.30
		2	24.28	1.95E+06	2.35E+05	8.30		
		3	23.45	2.91E+06	5.90E+05	4.93		
	IFITM2 negative control	-	No Ct	No Ct	-	-	-	-
	IFITM3 negative control	-	37.99	2.53E+03	-	-	-	-
	NTC	-	39.31	1.33E+03	-	-	-	-
IFITM2	A549s	1	21.04	2.65E+06	2.40E+05	11.02	10.39	0.65
		2	20.76	2.99E+06	2.81E+05	10.64		
		3	21.02	2.68E+06	2.82E+05	9.50		
	HEK293-T	1	24.22	6.69E+05	2.05E+05	3.26	2.35	0.96
		2	24.29	6.50E+05	2.35E+05	2.77		
		3	24.44	6.08E+05	5.90E+05	1.03		
	IFITM1 negative control	-	37.87	1.82E+03	-	-	-	-
	IFITM3 negative control	-	35.61	4.82E+03	-	-	-	-
	NTC	-	38.64	1.30E+03	-	-	-	-
IFITM3	A549s	1	26.5	1.33E+05	2.40E+05	0.55	0.55	0.01
		2	26.16	1.52E+05	2.81E+05	0.54		
		3	26.11	1.55E+05	2.82E+05	0.55		
	HEK293-T	1	27.6	8.55E+04	2.05E+05	0.42	0.29	0.12
		2	27.95	7.45E+04	2.35E+05	0.32		
		3	27.89	7.62E+04	5.90E+05	0.13		
	IFITM1 negative control	-	36.13	2.85E+03	-	-	-	-
	IFITM2 negative control	-	35.07	4.35E+03	-	-	-	-
	NTC	-	36.89	2.11E+03	-	-	-	-

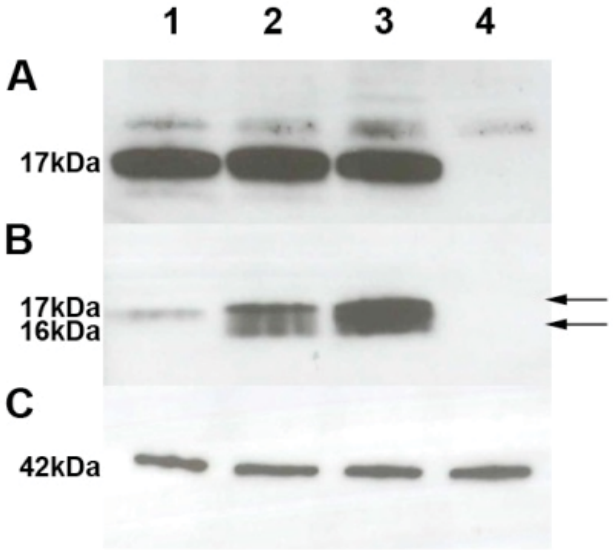

Figure 45: Testing IFITM3 and HA antibodies by Western blot.

The anti-HA antibody (A) and anti-IFITM3 antibody (B) were tested on A549 cells transduced with lentiviruses expressing either the truncated version of IFITM3 (ΔN-21, 1), full-length wildtype IFITM3 (2), or full-length human codon optimised IFITM3 (3). All constructs had a C-terminal HA tag. Black arrows show multiple bands observed when using the anti-IFITM3 antibody on full-length IFITM3. Samples were collected 24 h post transfection. Untransfected A549s were run as a control (4), as well as a β-actin loading control (C).

4.6 Testing the Stability of the C-terminal HA-tag on Human IFITM Proteins

Many studies that have explored the antiviral effects of IFITM proteins have been carried out in A549 cells[2-5,108,117,122], and in over-expression systems using HA tags. In collaboration with a group at University College London, we aimed to better characterise the location of human IFITM proteins during over-expression in A549 cells and determine if severing of the HA tag can occur in some instances.

As IFITM proteins are relatively short (less than 133 amino acids) co-staining for the NTD and CTD should give a near perfect co-localisation. A549 cell lines over-expressing human IFITM1 were incubated with antibodies against the NTD of IFITM1 (Sigma, HPA004810) and the C-terminal HA tag (Abcam, ab18181). Cell lines over-expressing human IFITM2 or 3 were incubated with antibodies against the NTD of IFITM3 (Abgent, AP1153a) and the C-terminal HA tag. Labelling with an NTD antibody shows that human IFITM1 expression occurs mainly on the cell surface and diffusely throughout the cytoplasm (Figure 46 i). Expression of human IFITM2 and 3 appears more punctate and clustered in the cytoplasm (Figure 46 ii and iii).

Human IFITM1 over-expressing cells showed a high degree of overlap for the two antibodies across multiple images, as demonstrated by the Mander's correlation coefficients M1 and M2 (0.97 and 0.99 respectively) (Table 11). This means that 97 % of the red pixels overlap with the green pixels and that 99 % of the green pixels overlap with red pixels. Furthermore, analysis of the areas of different pixel colours demonstrated that around 70 % of pixels were detectable as yellow. By contrast, in human IFITM2 and IFITM3 over-expressing cells, a lower level of co-localisation was observed (Figure 46 ii and iii). Importantly, clear red punctae, indicating the NTD, were visible. This suggests that in some of the organelles containing either IFITM2 or IFITM3, the IFITM proteins contain intact NTDs but lack the CTD-HA tag. This conclusion is supported by the quantification of multiple images that demonstrate a lower Mander's M1 and M2, compared to human IFITM1, and show an excess of red pixels for IFITM3 expressing cells. Data generated by Stuart Weston[262] (Marsh laboratory, University College London).

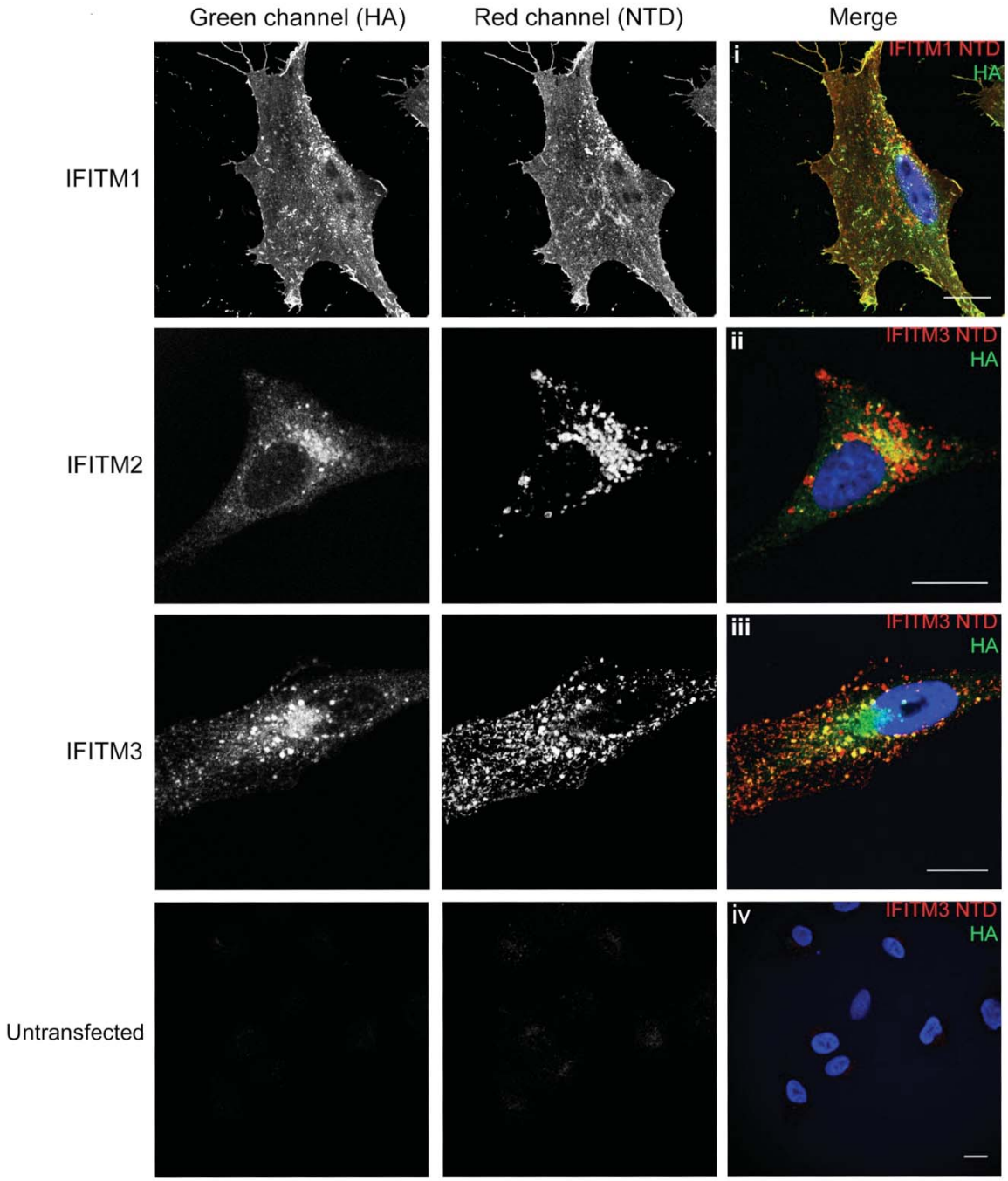

Figure 46: Co-staining with anti-NTD and anti-HA antibodies.

Permeabilised IFITM1 (i), 2 (ii) and 3 (iii) over-expressing A549 cells and untransduced A549s (iv) were stained with antibodies against the C-terminal HA-tag (green) and the NTD, using either the anti-IFITM1-NTD antibody for IFITM1 or the anti-IFITM3-NTD antibody for IFITM2 and 3 (red). Images represent a single optical slice (0.25µm thick) through the cell. Scale bars represent 15 µm.

Adapted from Weston *et al.*[262]

Table 11: Co-localisation analysis of anti-NTD and anti-HA staining of IFITM-expressing cells

Cell line	Number of cells imaged	Pearson's R value[+]	Mander's M1[§]	Mander's M2[§]
IFITM1	58	0.85 (±0.006)	0.97 (±0.12)	0.99 (±0.012)
IFITM2	57	0.73 (±0.13)	0.85 (±0.16)	0.86 (±0.14)
IFITM3	49	0.72 (±0.044)	0.75 (±0.21)	0.77 (±0.17)

Cell line	Number of cells imaged	Yellow relative area	Red relative area	Green relative area
IFITM1	14	0.70 (±0.18)	0.15 (±0.13)	0.15 (±0.13)
IFITM2	13	0.26 (±0.066)	0.49 (±0.11)	0.25 (±0.93)
IFITM3	15	0.27 (±0.077)	0.47 (±0.078)	0.26 (±0.081)

[+]Pearson's value represents the correlation in intensity between the red and green channels.

[§]Mander's correlation coefficients, M1 and M2, represent the overlap of red, in pixels that are green, and vice versa. Error given is of the standard deviation.

4.7 Subcellular Localisation of Human and Chicken IFITM Proteins

As human IFITM1, 2, and 3 have distinct subcellular localisations (Figure 46) we reasoned that assessing the localisation of putative chicken IFITM1, 2, and 3 would be a way to give further confidence to the orthologous predictions. Thus, the subcellular localisation of chIFITMs after over-expression in chicken cells was assessed and compared to the localisation of human IFITMs in A549 cells.

A549s were transiently transfected with human IFITM1, 2, or 3 and DF-1 cells (chicken fibroblasts) were transfected with non-codon-optimised chIFITM1, 2 or 3. Using confocal microscopy and two antibodies against HA and LAMP1 (a late endosomal marker), it is clear that the human proteins localise distinctly in the cell IFITM1 is expressed predominantly on the cell surface, whereas IFITM2 and 3 localise intracellularly. Previous studies have suggested that these proteins are trafficked to late endosomes, however we only see moderate co-localisation with Lamp1 (Figure 47). ChIFITM1 is diffusely expressed throughout the cytoplasm, whereas chIFITM2 is present in the cytoplasm and the cell membrane, which looks similar to the expression of human IFITM1 (Figure 48A). However, the localisation of human IFITM1 is somewhat inconsistent between Figure 46 and Figure 47. ChIFITM3 localises peri-nuclearly, which is consistent with expression of huIFITM3 (Figure 48C). However, some peri-nuclear staining may be an artefact of proteins being produced in the secretory pathway, but not enclosed in endosomes. ChIFITM3 therefore shares synteny, amino acid similarity, and subcellular localisation with huIFITM3. In the case of the other two chIFITMs, their localisation is less clearly paired with the human IFITMs, thus our nomenclature is founded on the gene order.

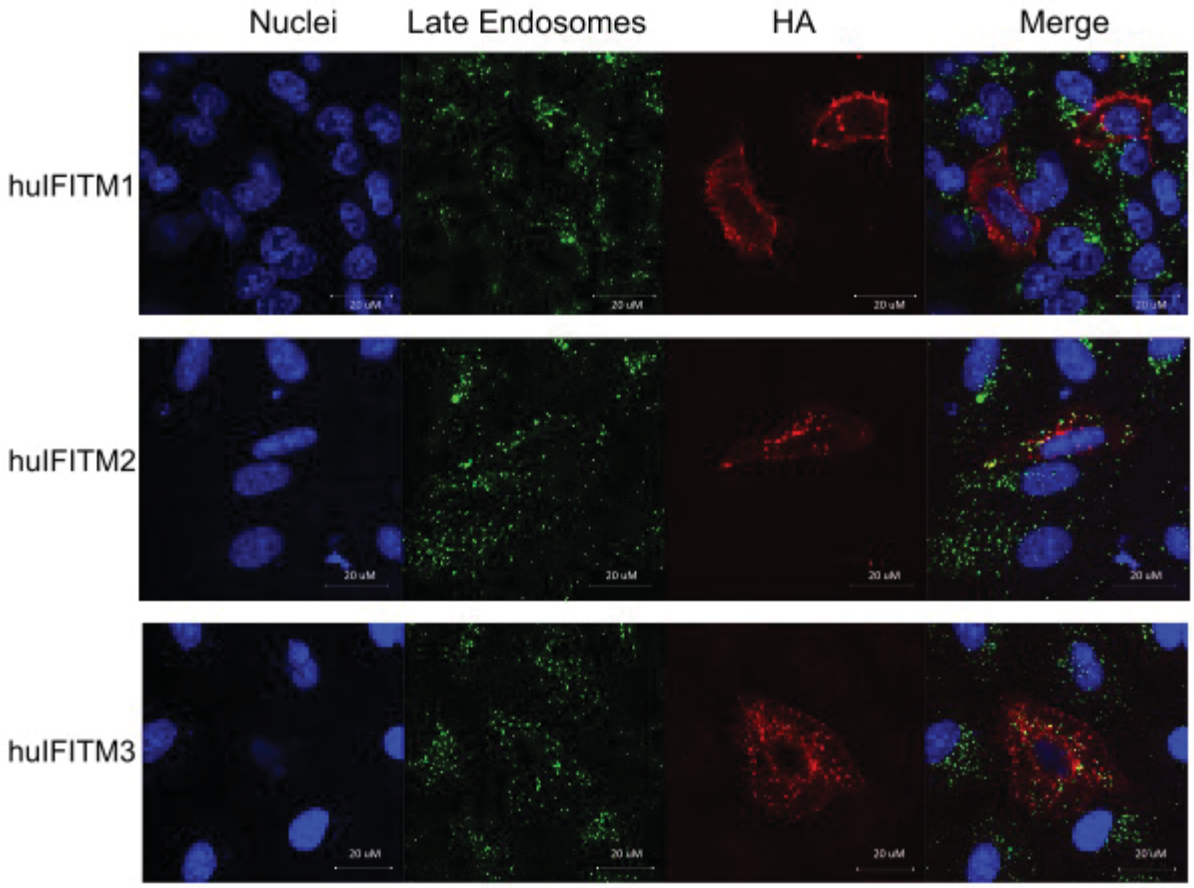

Figure 47: Cellular localisation of over-expressed human IFITM proteins in A549s

Confocal microscopy of A549s transduced with human IFITM proteins 1-3 (pBNHA_huIFITMX) in the absence of infection. Panels show nuclei stained with DAPI (blue), late endosomes marked with an antibody against lamp1 (green), IFITM proteins marked by an antibody against the HA tag (red), and a merged image. The scale bar represents 20 μm in each instance.

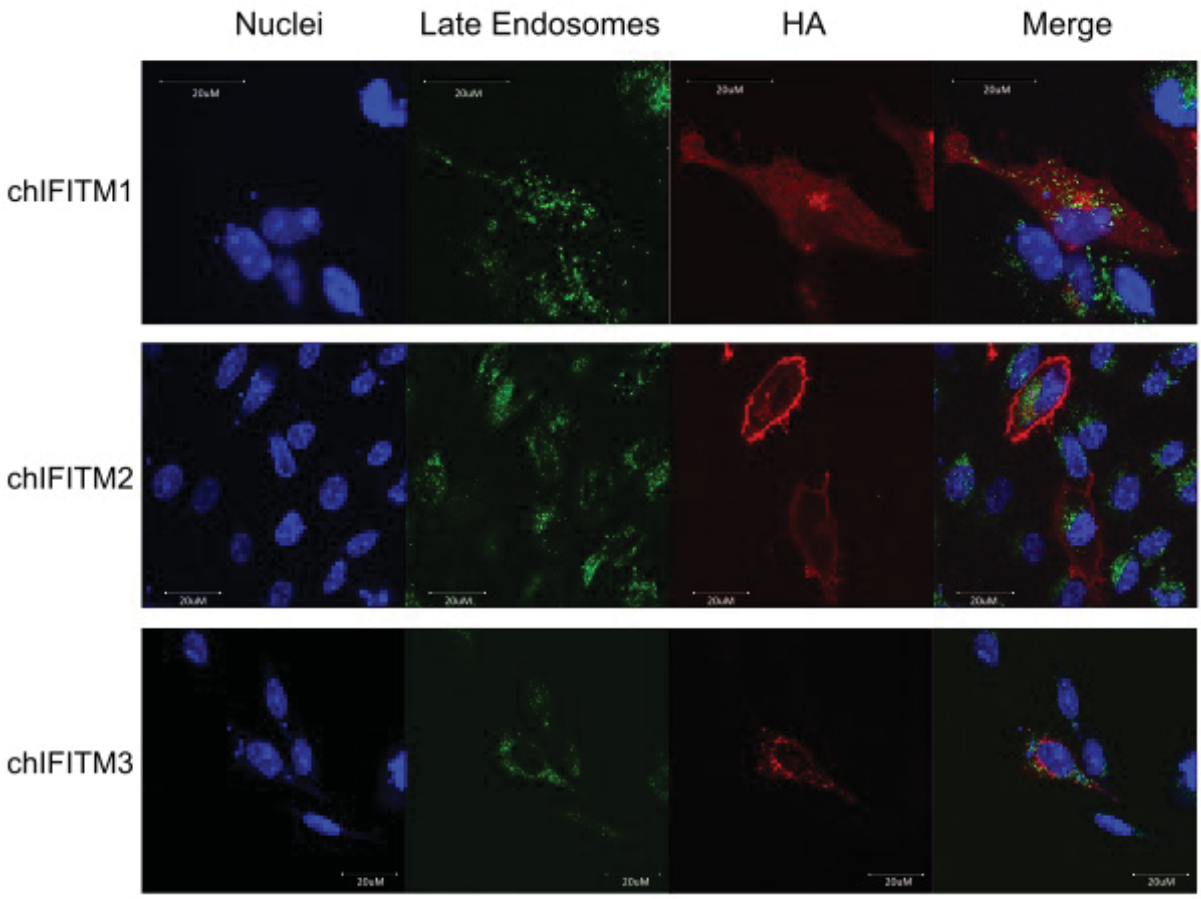

Figure 48: Cellular localisation of over-expressed chicken IFITM proteins in DF1 cells

Confocal microscopy of DF-1 cells transiently transfected with chIFITM proteins 1-3 (pBNHA_chIFITMX) in the absence of infection. Panels show nuclei stained with DAPI (blue), late endosomes marked with an antibody against chicken lamp1 (green), IFITM proteins marked by an antibody against the HA tag (red), and a merged image. The scale bar represents 20 µm in each instance.

4.8 Chicken IFITM Proteins Restrict Diverse Virus Infection

We investigated if, despite considerable amino acid sequence divergence, chicken IFITMs could function as restriction factors. Human codon-optimised chicken IFITM1, 2, and 3 were cloned into lentivirus vectors and these were used to transduce A549 cells. Single cell clones were isolated and developed from the bulk transformations, and expression of the clones tested by flow cytometry against the HA tag (Figure 49). Pure clones were obtained for both chIFITM2 and 3, but after several attempts, a clonal cell line expressing equivalent protein levels could not be made for chIFITM1 (Figure 49D). This could be due to C-terminal HA tag degradation preventing detection. Therefore as accurate comparisons could not be made, data for chIFITM1 is not included in further experiments.

Over-expression of huIFITM3 in A549s resulted in 98.3 % and 98.8 % reduction in infection by pseudoviruses expressing the lyssavirus envelopes from Rabies virus (RABV) and Lagos bat virus (LBV), and over-expression of chIFITM3 resulted in 79.4 % and 85 % reduction, respectively. This is similar to the level of restriction by huIFITM3 to the same viruses (Figure 50A) even though chickens are rarely infected by lyssaviruses[263]. ChIFITM2 also restricts lyssavirus LBV and RABV infection to a comparable level as chIFITM3. These experiments are the first to show restriction of lyssaviruses by any IFITM protein. Detection of chIFITM3 by western blot (Figure 50C) identifies a protein that runs at a higher molecular weight than predicted compared to human IFITM3 (predicted 14.9kDa and 14.6kDa respectively) and two bands are present, the reasons for which are unclear, but perhaps post-translational phosphorylation or myristoylation are responsible.

A similar pattern of restriction is seen for lentiviruses pseudotyped with IAV H1, H5, H7 and H10 (Figure 50B). HuIFITM3 restricted viral infection of all influenza HAs, reducing infection by greater than 90 %, and chIFITM3 restricted H1 and H10 pseudotypes as effectively, but restricted H5 and H7 less well. ChIFITM2 restricts more moderately, like huIFITM2, as shown by others[1]. Consistent with previous studies on huIFITM3[1,2], chIFITM3 failed to restrict MLV-A (Figure 50D). Overall, although chIFITM3 and huIFITM3 only share 42 % amino acid identity, the level of viral restriction of chIFITM3 is similar to huIFITM3.

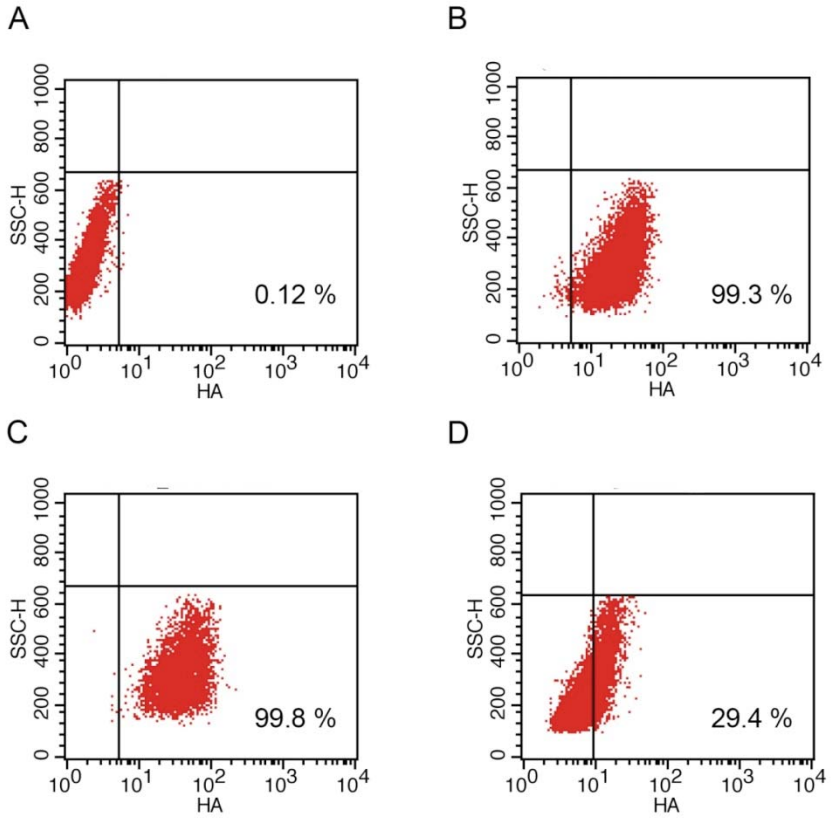

Figure 49: Flow cytometry of A549 single cell clones expressing chicken IFITM proteins

Clonal cell populations were assessed by flow cytometry using antibodies against the HA tag of the IFITM protein. Quandrants were defined by assessing the fluorescence of untransduced A549s (A) and 10,000 cells per gate were measured. The percentage of transduced cells is represented in the lower right quandrant of each graph for chicken IFITM2 (B), 3 (C), and 1 (D). N.B. a different negative control gate was used for chIFITM1 (D) as shown by the shifted quadrant.

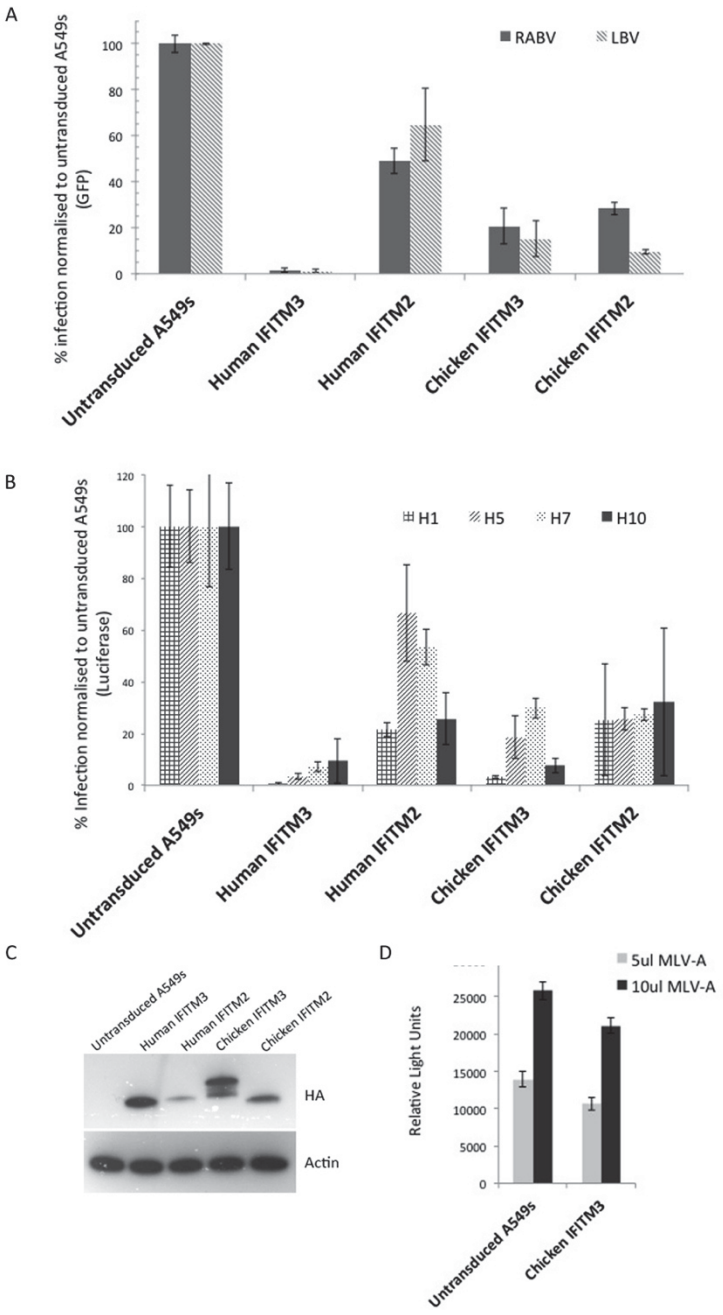

Figure 50: Human and chicken IFITM proteins restrict cell infection

Stable cell lines expressing hu and chIFITM2 and 3 were infected by pseudotyped viruses with either lyssavirus glycoprotein envelopes RABV (CVS-11); LBV (LBV.NIG56-RV1) (A) or IAV haemagglutinin envelopes (H1 [human], H5 [human], H7 [bird], H10 [bird]) (B). The relative level of infection compared to untransduced A549s was measured by GFP expression or luciferase activity for the lyssavirus and IAV envelope pseudotypes respectively. Error bars represent standard deviation across two biological replicates each performed in triplicate. Expression levels of each cell line are shown by Western blot (C) relative to endogenous β-actin. Stable cell line expressing chIFITM3 was infected with a pseudotyped virus expressing a luciferase reporter gene and the murine leukaemia virus (MLV-A) envelope as a control (D).

We hypothesised that cells expressing more IFITM proteins would restrict virus replication more effectively than clones expressing a small amount of protein. To test this, seven clones over-expressing chIFITM3 to varying levels were infected by a lentivirus vector pseudotyped with the lyssavirus LBV envelope (Figure 51). We show that there is a strong expression-level dependent correlation between chIFITM3 expression and the percentage of cells infected.

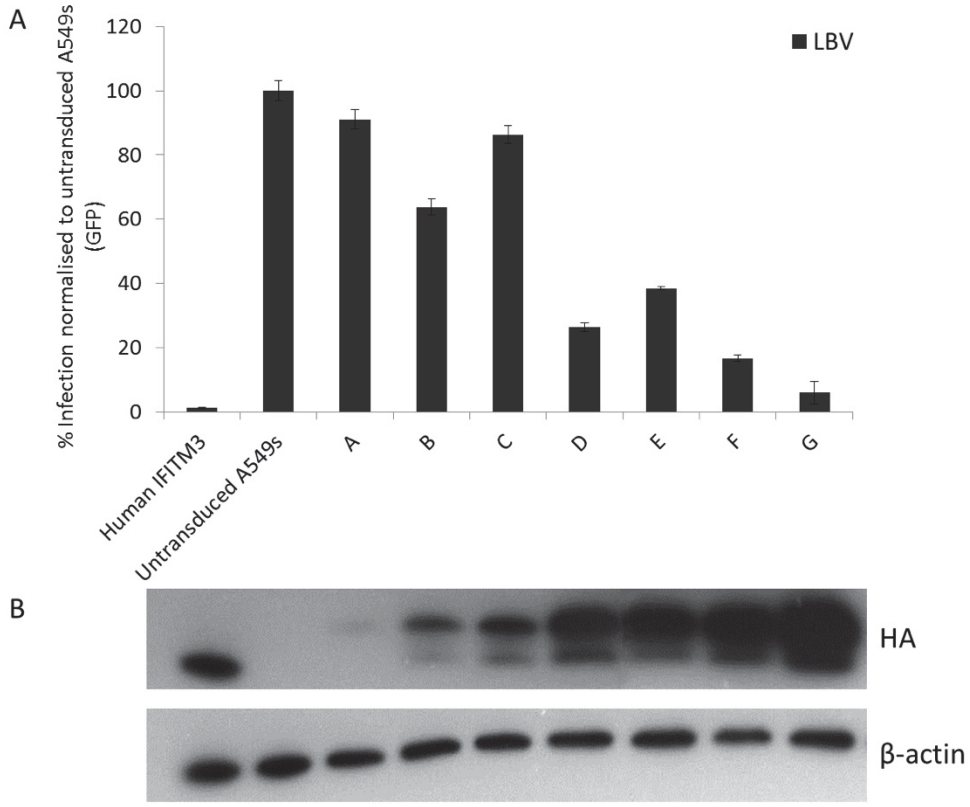

Figure 51: An increase in the expression of chicken IFITM3 is associated with a decrease in viral infection

A range of clonal A549 cell populations expressing increasing levels of chIFITM3 protein (bars A to G) were assessed by Western blotting of the HA tag (B). These cell lines were infected by a lentivirus pseudotyped with the Lagos bat virus (LBV) glycoprotein, and the replication was measured by GFP expression relative to that in untransduced A549s (A). Error bars show standard deviations of the means (n=3).

4.9 Ablation of IFITM Expression in Chicken DF-1 Cells Increases Infection

Although chIFITM proteins could be successfully over-expressed in human epithelial cells, it was still unclear whether or not these proteins were endogenously transcribed and translated in chicken cells.

We assessed the constitutive level of expression of *chIFITM3* in DF-1 cells (chick embryo fibroblast cell line), by quantitative RT-PCR with probes and primers specific for *chIFITM3* (Life Technologies). The results showed that DF-1 cells expressed high levels of *chIFITM3* compared to the *GAPDH* control (*IFITM3* Ct 20, *GAPDH* Ct 22). Despite being IFN inducible, addition of IFN-γ resulted in only a moderate induction, whereas addition of IFN-α (a type-I IFN) caused a 2.67 $\log_2$ (6.4 fold) increase in *chIFITM3* expression (Figure 52A). We assessed our ability to knockdown *chIFITM3* expression in DF-1 cells using an siRNA designed to the *chIFITM3* transcript. Treatment with this siRNA on unstimulated DF-1 cells resulted in a 1.23 $\log_2$ (2.4 fold) reduction in the transcript level, with no change in *chIFITM3* transcript abundance with a non-specific siRNA. Knockdown of endogenous *chIFITM3* resulted in a greater than two fold increase in infection of DF-1 cells by replication competent influenza A (A/WSN/1933) (Figure 52B), assayed by flow cytometric analysis of nucleoprotein expression.

Furthermore, DF-1 cells were transfected with chIFITM3_HA and subsequently infected with influenza A (A/WSN/1933). Cells over-expressing chIFITM3_HA and NP were detected by flow cytometry (Figure 53A). Over-expression of chIFITM3 in DF-1 cells reduced viral replication by an average of 55 % (Figure 53B) and plaque assays show that viral load was reduced from 1.3×10^6 plaque forming units (pfu) ml^{-1} to 3.1×10^5 pfu ml^{-1} when chIFITM3 was transiently overexpressed (Figure 53C). Together, these results show chIFITM3 is able to restrict IAV entry into DF-1 cells.

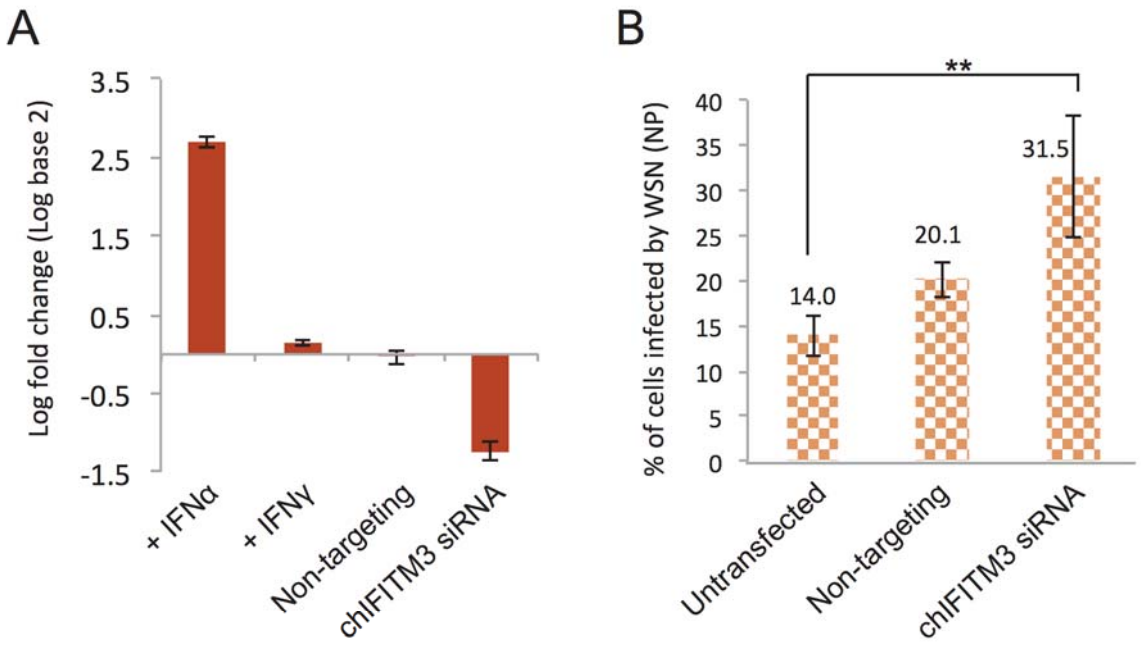

Figure 52: Endogenous chicken IFITM3 has antiviral activity against IAV in DF-1 cells

The expression level and log fold change of chIFITM3 was measured using quantitative RT-PCR after stimulation with IFNα and IFNγ or after pre-incubation with a non-targeting siRNA or one specific to chIFITM3 (A). The effect of knocking down endogenous chIFITM3 expression in DF-1 cells infected with influenza A virus (A/WSN/1933 [WSN/33]), was measured by flow cytometry using an antibody against nucleoprotein (B) p=0.01, Student's *t*-test. Error bars represent standard deviation across each condition performed in triplicate.

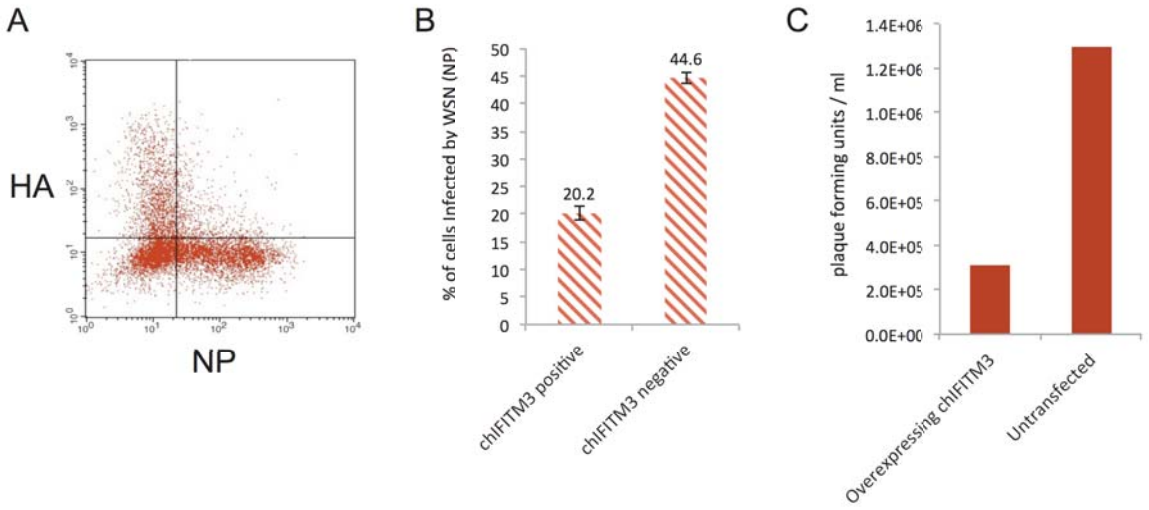

Figure 53: Over-expression of chicken IFITM3 in DF-1 cells reduces infection by influenza A

DF-1 cells transfected with pBNHA_chIFITM3 were infected by WSN/33. Expression of the HA tag and influenza NP was detected by flow cytometry (A and B), and viral titres were measured by calculating the number of pfu ml^{-1} of cell culture supernatant (C). Error bars represent standard deviations across each condition performed in triplicate.

4.10 Differential Expression of IFITMs in Chicken Tissues

We assessed the tissue specific gene expression pattern in chickens using a panel of RNA extracted from tissues of three week old Rhode Island red (RIR) chickens. This tissue panel included: thymus, spleen, bursa of Fabricius, caecal tonsil, trachea, gastro-intestinal tract, bone marrow, brain, muscle, heart, liver, kidney, lung, and skin. Three primer-pairs were designed to specifically amplify to *chIFITM1, 2* or *3* (Figure 54) and primer specificity was tested on plasmid controls encoding each chicken gene (Figure 55). The maximum percent sequence identity of each primer to the other chIFITMs was calculated and is shown in Table 12.

Table 12: Primer binding affinity for chicken IFITM sequences

	Primer	**IFITM (% IDENTITY)**		
		1	**2**	**3**
FORWARD	CHIFITM1_F'	100.00	61.90	66.67
	CHIFITM2_F'	65.00	100.00	70.00
	CHIFITM3_F'	57.89	57.89	100.00
REVERSE	CHIFITM1_R'	100.00	61.90	57.14
	CHIFITM2_R'	55.00	100.00	55.00
	CHIFITM3_R'	52.63	52.63	100.00

Expression of *IFITM2* and *3* was detected in all tissues, although with lower expression levels in the muscle and brain and higher levels in the caecal tonsils (Figure 55). In contrast, expression of *IFITM1* was more restricted and confined to the bursa of Fabricius, the gastro-intestinal tract, and the caecal tonsil.

Amplifying Chicken *IFITM1*

ATGCAGAGCTACCCTCAGCACACCAGCATCAACATGCCTTCCTACGGGCAGGATGTGACCACCACTAT
TCCCATCTCTCCGCAGCCGCCCCCCAAGGATTTTGTACTCTGGTCCCTCTTCAACTTTGTGCTGTGCA
ACGCCTTCTGCCTGGGCTTATGTGCTCTCTCATACTCCATCAAG*GTA…CAG*TCCAGGGATAGGATCAT
CGCCAAGGACTTCGTAGGCGCCAGCAGCTATGGGAGGACAGCGAAGATCTTTAACATCTTTGCATTCT
GTGTGGGACTTCTTGTGACCATCCTCTCCATCGTCCTGGTGTTTCTCTACCTCCCGTTGTACACTGTG

Predicted size: 198 bp

Amplifying Chicken *IFITM2*

ATGAAGCCGCAACAGGCGGAGGTGAGCATCCCGCTGCACCCACCCGGGCGGGGGCCGCCCCTCGCCAG
CCTCCCCGACGAGCAGCCCCGCGACTTCATCCTCTGGTCCCTCTTCAACGTCCTGGCGGGCTTCGCTC
TCGCCTACCTCGGCTGCTTCTGCTTCCCCTCGCTCATCTTCTCCATCAAG*GTG…TAG*GCCCGCGACTG
CAAAGTGCTGGGCGACCTGGAAGGTGCTCGGCGGTATGGAAGCCGGGCCAAGGTGCTGAACATCATCT
TCTCTGTGCTGATAGCCGTCGGTGTGTTGTCCACCATCACCATTGCCATCATGTTCATCACCGCGATC

Predicted size: 213 bp

Amplifying Chicken *IFITM3*

ATGGAGCGGGTACGCGCTTCGGGTCCGGGAGTCCCACCGTATGAACCCCTGATGGACGGGATGGACAT
GGAGGGGAAGACCCGCAGCACGGTGGTGACGGTGGAGACGCCCCTGGTGCCTCCTCCCCGCGACCACC
TGGCCTGGTCGCTGTGCACCACGCTGTACGCCAACGTCTGCTGCCTCGGCTTCCTGGCGCTCGTCTTC
TCCGTGAAG*GTT…CAG*TCCAGGGATCGCAAAGTCCTGGGTGACTACAGCGGGGCGCTCAGCTATGGCT
CCACTGCGAAGTACCTGAACATCACGGCCCATCTGATCAACGTCTTCCTCATCATCCTCATCATCGCC

Predicted size: 83 bp

Figure 54: Location of primers to uniquely amplify chicken IFITM1, 2, and 3

Forward and reverse primers were designed to distinguish between *chIFITM1*, *2*, and *3*. Sequence in orange indicates where *IFITM1* primer pairs bind, red indicates *IFITM2* primer pair binding, and blue indicates *IFITM3* primer pair binding. Grey italicised letters indicate intronic sequence. Predicted sizes are for mRNA.

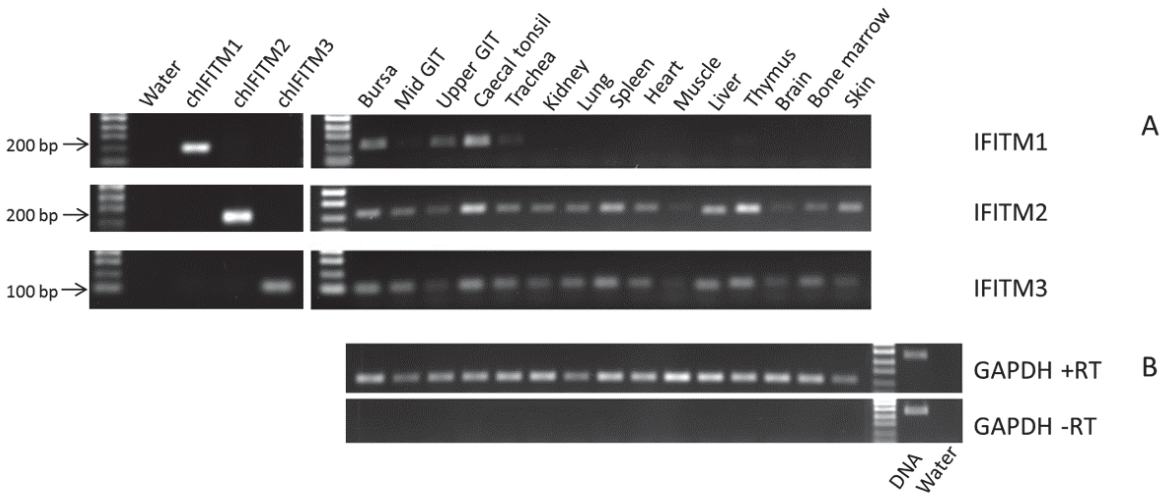

Figure 55: Differential expression of *chicken IFITM* transcripts in chicken tissues

Expression levels of *IFITM1*, *2*, and *3* were determined by RT-PCR across a range of chicken tissues (A) and compared to the expression level of *GAPDH* (B). *GAPDH* PCR was also performed without reverse transcriptase (−RT) to control for genomic DNA contamination.

4.11 Discussion of Results

To date, the antiviral activity of IFITM2 and IFITM3 proteins have only been demonstrated in mammals, with a single report characterising the function of chicken IFITM1 and IFITM5[2]. Computational analysis of vertebrate genomes suggests the *IFITM* gene family is present throughout vertebrates. However this analysis, and any phylogenetic reconstruction of gene history, is complicated by the paralogous nature of the *IFITM* gene family, the presence of copy number variations and the presence of numerous processed pseudogenes[103]. Indeed, the identification of avian IFITM proteins as part of the Dispanin protein family failed to identify chicken IFITMs in the antiviral IFITM1-3 subfamily defined as DSP2a-c[106]. Similarly, a more thorough analysis of vertebrate *IFITM* genes identified distantly related *IFITMs* in reptiles and birds, but primarily focused on eutherian sequences for a detailed phylogenetic analysis[132]. Hickford *et al.*[104] have undertaken a comprehensive analysis of IFITM genes across a broad range of chordates. The authors showed that all of the species analysed, including 'lower' vertebrates such as lampreys, possess at least one IFITM-like gene. Phylogenetic analysis of all the *IFITM* paralogues they identified revealed that *IFITM5* emerged first in bony fish whilst *IFITM10* appears restricted to tetrapods.

This study resolved the antiviral *IFITM* locus on chromosome 5 of the chicken genome, expanding the number of *IFITM* genes to four in this locus, and confirmed that the locus is flanked by the genes *ATHL1* and *B4GALNT4*[132]. Crucially, we have shown that anti-viral activity is conserved in chicken IFITM proteins. The low-level sequence identity and orientation change of *chIFITM2* and *chIFITM1* make the phylogenetic assignment of orthology problematic. The revised nomenclature of the chicken IFITM locus presented here is based on the syntenic gene order and functional data where possible. However, given chIFITM2 is localised to the plasma membrane, and the lack of an N-terminal extension (characteristic of huIFITM2/3) it is possible that it is analogous to huIFITM1. The direction of transcription of chicken *IFITM1*, *2*, *3* and *5* are all on the reverse strand, whereas in the human genome *IFITM1* and *2* are on the forward strand and *IFITM3* and *5* are on the reverse strand. A simple inversion of the gene block containing chicken *IFITM1* and *2* would lead to the gene arrangement seen on chicken chromosome 5. In addition chIFITM1, unlike chIFITM2, has a tyrosine residue in the N-terminus (Y4), which could also lead to

some endosomal localisation, suggesting the nomenclature of these two proteins has been inverted. However, the chicken appears to express one longer protein and two truncated proteins, unlike the human orthologues so regardless, there are some differences in these protein families. Furthermore although chIFITM2 was an outlier on the phylogenetic tree, suggesting it was more dissimilar to chIFITM1 and 3, this could be due to mid-point rotting. If the tree was rooted on chIFITM1, the phylogeny is inverted. Also, the expression in tissues suggests that chIFITM2 and 3 are more similar. It is likely that similar extensive genetic and functional analyses will be essential to characterise the *IFITM* loci in other vertebrate species and define unambiguously *IFITM1*, *2* and *3* orthologues.

A stable clonal cell line expressing chIFITM1 could not be made; this lack of stability at high expression levels is supported by Hach *et al.*[264] who show that over-expression of unpalmitoylated murine IFITM1 is difficult to achieve. It is possible therefore, that enforced expression of chicken and human IFITM1 results in cellular toxicity.

Control of animal pathogens, especially those with zoonotic potential is a key component of ensuring human health and food security. RABV is responsible for approximately 70,000 human deaths each year[265] while other lyssaviruses have only been conclusively shown to cause a handful of fatalities[266], although this could be due to poor surveillance. Our results are the first to show diverse members of this genus of virus are sensitive to the inhibitory action of human IFITM proteins. Furthermore, although most warm-blooded animals are susceptible to RABV, domestic birds are rarely infected by lyssaviruses[263]. Despite this, chIFITM2 and 3 were able to significantly reduce cell lyssavirus infection.

Avian IAV infections however, pose significant threats to human health, to the international poultry industry, and to small scale poultry farmers[267]. Our identification and functional characterisation of the avian *IFITM* locus, together with knowledge that this gene family exists with copy number and allelic variants in other species[3,132,247], should provide a focus for identifying *IFITM* variants with enhanced antiviral activity for use in farm-animal breeding strategies to improve animal infectious disease resistance. Specifically, we hypothesise that certain wild or outbred chicken *IFITM* allelic variants will confer enhanced levels of protection to pathogenic avian viruses that enter through acidic endosomes, and that

breeding for enhanced activity in IFITM variants will improve disease resistance in chickens. Similarly, should chicken IFITM proteins restrict IAV infection in chick embryos the ablation of IFITM protein expression could improve vaccine production and boost yield.

We showed, by both Western blot and RT-PCR, that A549 cells are low in IFITM3 expression and, as lung epithelial cells, are a suitable cell line to perform over-expression experiments in. Co-localisation experiments showed that cleavage of the HA tag does occur, and thus this should always be considered when drawing conclusions from the data; experiments using HA tags will underestimate IFITM protein expression. However, specific chIFITM antibodies were not available, and specificity in designing an antibody is difficult to achieve because of the high sequence similarity between the homologues. Therefore IFITM-specific PCR primers were designed to assess gene expression and primer specificity, confirmed on plasmids encoding each IFITM transcript.

We have shown that chIFITM proteins, expressed in human A549 cells are capable of restricting diverse viruses that enter cells through the acidic endosome pathway. Further, we show that DF-1 chicken cells constitutively express chIFITM3 and this protein is able to restrict influenza infection *in vitro*. Despite sharing less than 50 % amino acid identity, both chIFITM3 and huIFITM3 effectively restrict the entry of all lyssavirus and IAV envelope pseudotypes tested. Nevertheless, certain key amino acids in the N-terminus, IM1, and the CIL domain are conserved in chicken and human IFITM3, suggesting a functional importance.

The immunofluorescence studies showed that the location of IFITM1, 2, and 3 varies for both humans and chickens. Human IFITM1 does not have the YxxΦ motif that enables IFITM2 and 3 to be trafficked to endosomal compartments, which is likely to be why expression of IFITM1 is mainly on the cell surface. Previous studies have shown the IFITM1, 2, and 3 in humans preferentially restrict viruses to differing degrees; IFITM2 and 3 restrict Semliki Forest Virus much more effectively than IFITM1[115], but IFITM1 restricts Jaagsiekte sheep retrovirus, whilst neither IFITM2 or 3 do[115]. Moreover, IFITM2 and 3 can restrict a range of Bunyaviruses, but not Crimean-Congo haemorrhagic fever virus[117]. It is possible that the differences in restrictive capabilities are reliant on the location and trafficking of the protein within the cell.

Although human IFITM3 restricted the different HA influenza subtypes to a similar degree, there were substantial differences in restriction by chIFITM3. This could be due to differences in fusion pH for the HAs[268]. The structure of human IFITM3 may be more rigid than chicken IFITM3, so that the fusion peptide of some HAs are able to penetrate the restriction of the chicken protein, but not the human protein.

Expression of chicken IFITM1 was restricted to the respiratory and gastrointestinal tract, unlike chicken IFITM2 and 3, which were expressed systemically. Klymiuk *et al*. carried out a comprehensive study on mouse IFITM1 expression during development and in adult tissue. The study looked for expression in dissected adult brain, intestine, kidney, liver, lung, ovary, pancreas, spleen, tongue, and thymus of *Ifitm1*$^{tm1IEG/wt}$ mice, but only found reproducible expression in the lung and thymus[269]. However, it is expressed in many tissue types during mouse embryo development. The authors also detected IFITM1 in human bronchial epithelium and increased expression in human lung carcinomas by immunohistochemistry, but did not do a thorough analysis of multiple tissue types. Everitt *et al.* and Bailey *et al.* also showed that *Ifitm3* is expressed ubiquitously in the mouse intestine, liver, spleen, lung, bronchioles, trachea, leuckocytes and lymph nodes[107,109], which is in accordance with IFITM3 expression in chickens.

Many key questions remain; it is unclear how genes such as *IFITM3* in humans and chickens, separated by 310 million years of evolution[270], sharing less that 50 % amino acid identity, maintain a conserved cellular location and a strong antiviral activity against a diverse range of viruses. It is of equal importance to determine, given the level of antiviral activity and the proposed indirect mechanism of IFITM protein restriction[115,122], how viruses overcome the restriction either within or between species. Investigating appropriately defined IFITM loci from different host species where cross species transfer of virus infection occurs may help explain barriers and vulnerabilities to infection by diverse viruses.

4.11.1 Conclusions

Work here has identified the chicken IFITM locus on chromosome 5, and investigated the antiviral function of chIFITM1-3. These proteins localise to similar sub-cellular regions as their human orthologues, and can restrict infection by influenza and lyssavirus infections in human cell culture systems. Chicken IFITM3

was also identified as type I IFN-inducible and induction can reduce influenza A infection in DF-1 cells. Expression of chIFITM1, 2, and 3 was also confirmed in several chicken tissue types.