Mixed Signatures of Selection at Selected Immune and Heat Shock Protein Genes in Selected Poultry Species

Caroline Sigei[1], Daniel Kariuki[1], Emmanuel Ndiema[2], Eunice Wainaina[1], Simon Maina[3], Philip Oyier[3], Jacqueline Lichoti[4] and Sheila Ommeh[5]*

[1]Department of Biochemistry,
Jomo Kenyatta University of Agriculture and Technology,
P.O. Box 62000-00200, Nairobi, Kenya;
E-mails: carolinecsigey@gmail.com (SC); dkariuki@jkuat.ac.ke (DK);
wainainaeunice1@gmail.com (EW)

[2]Department of Earth Sciences,
National Museums of Kenya,
P.O. Box 40658-00100, Nairobi, Kenya;
E-mail: endiema@gmail.com (EN)

[3]Department of Information Technology,
Jomo Kenyatta University of Agriculture and Technology,
P.O. Box 62000-00200, Nairobi, Kenya;
Emails: oyier@itc.jkuat.ac.ke (PO); msimog@gmail.com (SM)

[4]State Department of Livestock,
Ministry of Agriculture, Livestock, and Fisheries,
P.O. Box 34188-00100, Nairobi, Kenya;
E-mail: kasiiti.orengo@gmail.com (JL)

[5] Institute for Biotechnology Research,
Jomo Kenyatta University of Agriculture and Technology,
P.O. Box 62000-00200, Nairobi, Kenya;
E-mail: sommeh@jkuat.ac.ke (SO)
***** Author to whom correspondence should be addressed; E-Mail: sommeh@jkuat.ac.ke

Abstract
Indigenous poultry represent a valuable micro-livestock resource that is particularly important to the rural-poor populace of Sub-Saharan Africa. However, they are threatened by virulent disease outbreaks and heat-stress conditions associated with the ongoing climate change. Based on the availability of genomic sequences in public databases, we selected genes and performed a reciprocal BLASTp to select inter-species homologs for comparative analysis. We used MUSCLE software to perform Multiple Sequence Alignments prior to phylogeny reconstruction using FastMe. To detect signatures of selection, we used nested codon substitution models of PAML package to compute the rate of non-synonymous (dN) to synonymous (dS) substitutions in different lineages and amino acid sites through likelihood Ratio Tests (LRTs) and Bayes Empirical Bayes (BEB)

posterior probabilities. Computational approach led us to detect signatures of adaptive and purifying selection at Protein KinaseR (PKR), 2'-5'-oligoadenylate synthetase (OAS), Toll-like Receptor7 (TLR7) and Toll-like Receptor3 (TLR3) immune genes. We also detected predominant purifying selection at Heat Shock Protein70 (HSP70), Heat Shock Protein90 (HSP90), and Small Heat Shock Protein (sHSP) genes. These results form an important foundation for further statistical testing and experimental validation through in vitro and in vivo studies and subsequent genetic development of better adapted poultry.

Keywords: *Adaptive evolution, candidate genes, computational molecular evolution, dN/dS, indigenous poultry, in silico*

Introduction

The Indigenous chickens of Africa are characterized by an extensive genetic and phenotypic diversity (Moraa et al., 2015; Mwacharo et al., 2013). This provides a base from which different alleles can be selected for genetic research and improvement of poultry. However, they are threatened by recurrent outbreaks of highly infectious diseases such as New Castle Disease and Avian Influenza that can cause up to 100 % mortality (Gardner, 2014; Jibril et al., 2014). In addition to inbreeding, local farmers and national breeding programs have often relied on cross-breeding with exotic breeds in an attempt to improve productivity. This has however resulted in compromised immunity and inability to adapt to local rural environmental conditions (Magothe, 2012). Use of expensive chemotherapeutic drugs and vaccinations to control poultry diseases is not a feasible approach for the rural-poor populace. Climatic extremes associated with climate change pose another challenge to poultry production in terms of heat-stress, infectious disease distribution, virulence and cross-species transmissions (Howard & Fletcher, 2012; Vandegrift et al., 2010).

Implementation of proper selection and molecular breeding schemes represent a viable approach. Advanced technologies like Genome-Wide Association Studies (GWAS), High-density Single Nucleotide Polymorphism (SNP) chips and QTL mapping have been successfully used in developed countries to provide insights into DNA variation and subsequently breed for desired traits among different livestock breeds (Mukhopadhyay, 2012; Dekkers, 2012; Kranis et al., 2013; Wolc et al., 2013). However, this approach is extremely costly, time-consuming and labor-intensive for most developing countries. We therefore proposed to utilize the freely available gene sequences and bioinformatics tools as a fast and cheap approach to identifying candidate genes and genomic regions that are targeted by natural selection. We used codon based maximum likelihood substitution models of PAML package which promoted the d_N/d_S ratio test to a parameter known as omega (ώ) (Yang 2007). This is popularly used to study function altering mutations along lineages and among amino acid sites in the coding region of a gene. Values of $\omega>1$, $=1$ and <1 indicate positive selection, neutral evolution, and purifying selection respectively. Using these models, Ommeh (2010) detected balanced selection in Mx gene of village chickens. Also, Lynn et al., (2004) identified signatures of positive selection on mammalian alpha defensin genes (Lynn et al., 2004) and also on the CD4 gene which encodes glycoproteins in the bovine genome (Lynn et al., 2005).

Methods

Identification of candidate genes and selection of homologs through reciprocal BLAST

Four chicken innate immune genes (TLR7, TLR3, OAS and PKR) and three heat-stress genes (HSP70, HSP90 and sHSP) were selected for analysis from bibliographic and biological databases. The complete chicken protein sequence for each gene was downloaded from NCBI's GenBank and saved for the next step of analysis. Pairwise sequence alignment was used to obtain homologs from other animal species that also express the select genes. In particular, BLASTp program was used to search the Non-redundant protein sequences database (nr) of NCBI's GenBank using the chicken protein query (http://blast.ncbi.nlm.nih.gov/Blast.cgi). We used $1e^{-10}$ as the expectation value cut off for selecting the homologs (Dataset 1 to 7). The scoring matrix used was BLOSUM62 and Gap costs were: Existence: 11 and Extensions: 1. A reciprocal BLAST was then performed to confirm homologs. From GenBank database, we downloaded the complete amino acid and corresponding coding sequences for each homolog. We then saved these sequences in FASTA format and renamed prior to the next step of analysis. Stop codons were also manually removed from the coding sequences thus avoiding interference in the downstream analysis.

Multiple Sequence Alignment (MSA) of homologous sequences

Multiple sequence alignment was performed on the selected gene sequences so as to assess and confirm homology. We used four different MSA programs, i.e, Clustal X version 2.0 (Larkin et al., 2007), MUSCLE version 3.8.31(Edgar, 2004), PRANK v.140603 (Löytynoja & Goldman, 2005) and MAFFT version 7 (Katoh & Standley, 2013). The alignment outputs were viewed and edited using Jalview version 2.8 (Waterhouse et al., 2009) and SeaView version 4.5.3 (Gouy et al., 2010). MUSCLE alignment outputs for each dataset were selected as inputs of the subsequent analyses.

Selection of substitution models

MEGA version 6 (Tamura et al., 2013) was used to select evolutionary protein substitution models as well as the Alpha Shape Parameter of gamma distribution. MUSCLE output files were first converted into MEG files before being tested against 48 different amino acid substitution models through Maximum Likelihood fits. The best fit model was selected as having the lowest Bayesian Information Criterion (BIC) scores and the lowest Akaike's Information Criterion (AIC) scores.

Reconstruction of Phylogeny

Phylogeny was built based on the maximum likelihood algorithm of MEGA version 6 (Tamura et al., 2013). We evaluated the reliability of the trees using 1000 bootstrap replicates. The trees produced were saved in Newick format.

Signatures of selection tests

To test the hypothesis that there are variable selective pressures acting on specific amino acid sites and on specific lineages of our target genes, we used the non-synonymous to synonymous substitution rate ratio (ω=dN/dS). To achieve this, we used nested codon-based substitution models of CODEML program of PAML4 v 4.2 package (Yang, 2007). The LRT statistic ($2\Delta\ell$) = 2 ($\ln LM_0$ - $\ln M_1$) was used and the results compared to a χ^2 distribution with NPm1-NPm2 degrees of freedom where NP is the number of parameters. To identify genetic signatures of

positive selection acting on different lineages along a phylogeny, two branch-based models were compared by LRTs; the one-ratio model (M0) and the free-ratio model (M1). The one-ratio model assumes that all branches have the same one ω-parameter, whereas the free-ratio model assigns a different ω -parameter to each branch in the tree for estimation. Where the free-ratio model is significantly better than the one-ratio model and lineages have ω values >1, this is evidence of adaptive evolution. Analyses of signatures at individual amino acid-sites were performed using site-based models which treat the ω ratio for any codon in the gene as a random variable from a statistical distribution, thus allows ω to vary among codons (Yang *et al*., 2000). Based on this, an LRT was constructed to compare the null model M7 which assumes a beta distribution of ω across sites with ω values between 0 and 1 to the alternative model M8, which adds an extra class of sites to M7 where ω can take values >1. Therefore, positive selection can be detected if a model allowing for positive selection is significantly more likely (as estimated by the LRT) than a null model without positive selection. When the LRT suggests positive selection, the BEB method is used to calculate the posterior probabilities that each codon is from the site class of positive selection under model M8. Codons are identified to be undergoing adaptive evolution where both tests are significant and the posterior probabilities under M8 model are ≥0.95. The CODEML settings for the null (neutral) model M0 were model = 0, NSsites = 0, and for the alternative (selection) model M1 were model = 1, NSsites = 0. The CODEML settings for the null model M7 were model = 0, NSsites = 7, and for the alternative (selection) model M8 were model = 0, NSsites = 8. After the analysis, the Log Likelihood Ratio Test (LRT) was used to test the confidence of the results obtained from all the models while the Bayes Empirical Bayes (BEB) posterior probabilities were used to test the confidence of the results obtained from amino acid sites analysis.

Results and discussion

Signatures of selection at innate immune genes

Understanding the selective pressures that have shaped the evolution of innate immunity genes can provide insights into resistance/susceptibility of organisms to infectious diseases. Although Mukherjee et al., (2009) have previously reported that innate immunity genes are under strong purifying selection, our study detected heterogeneous signatures across all the selected genes. Since these genes are PRRs that detect conserved PAMPs on the viral pathogens, we hypothesized two contrasting views. First is that the purifying signatures could have been driven by functional constraints aimed at removing disadvantageous mutations that can interfere with the host's ability to detect the conserved PAMPs of the invading pathogens. On the other hand, viral pathogens constantly evolve new strategies to counteract host defense and this results in genetic conflicts which can give rise to new alleles that can confer resistance to the rapidly mutating pathogens (Daugherty & Malik, 2012; Sawyer & Elde, 2012). The detected positive signatures could therefore be a result of co-evolution of host-restriction factors with the viral inhibitors through molecular "arms-race". This is supported by the fact the positively selected sites were found to occur in domains that directly interact with the viral pathogens. We also observed that the amino acid residues in these sites varied across the selected homologs for all the selected genes (Supplementary Figure 1 to 4). Since different amino acids have different physicochemical properties, such variabilities can have important functional consequences which can determine the receptor's binding capacity as well as species-specific ligand recognition and cross-species transmissions of pathogens.

Toll-like Receptor7

We obtained highly significant P-values for TLR7 which led us to reject the null hypothesis of selective neutrality (ω=1) (Table 1). From lineage analysis, we identified predominant purifying signatures and a few adaptive signatures along the lineages of bats, domestic poultry and wild chicken (fig. 1). Furthermore, we detected 3 positive signatures at the Leucine Rich Repeats (LRRs) domain (fig. 2). These findings concur with the findings of Alcaide & Edwards, (2011); who reported predominant purifying selection and significant positive signatures at amino acid sites in birds. Similar findings have also been reported in Galloanserae birds (Vinkler et al., 2014) and wild rodents (Fornůsková et al., 2013).

Table 1: Likelihood ratio tests (LRTs) for TLR7

Models	2Δℓ	χ2 Value	D. f	P-value	Model favored
Lineage Analysis (M0 v M1)	2(-23842.34 -23753.86)	176.96	55	P<0.001	M1
Codon Site Analysis (M7 v M8)	2(23118.85 -23102.63)	32.44	2	P<0.001	M8

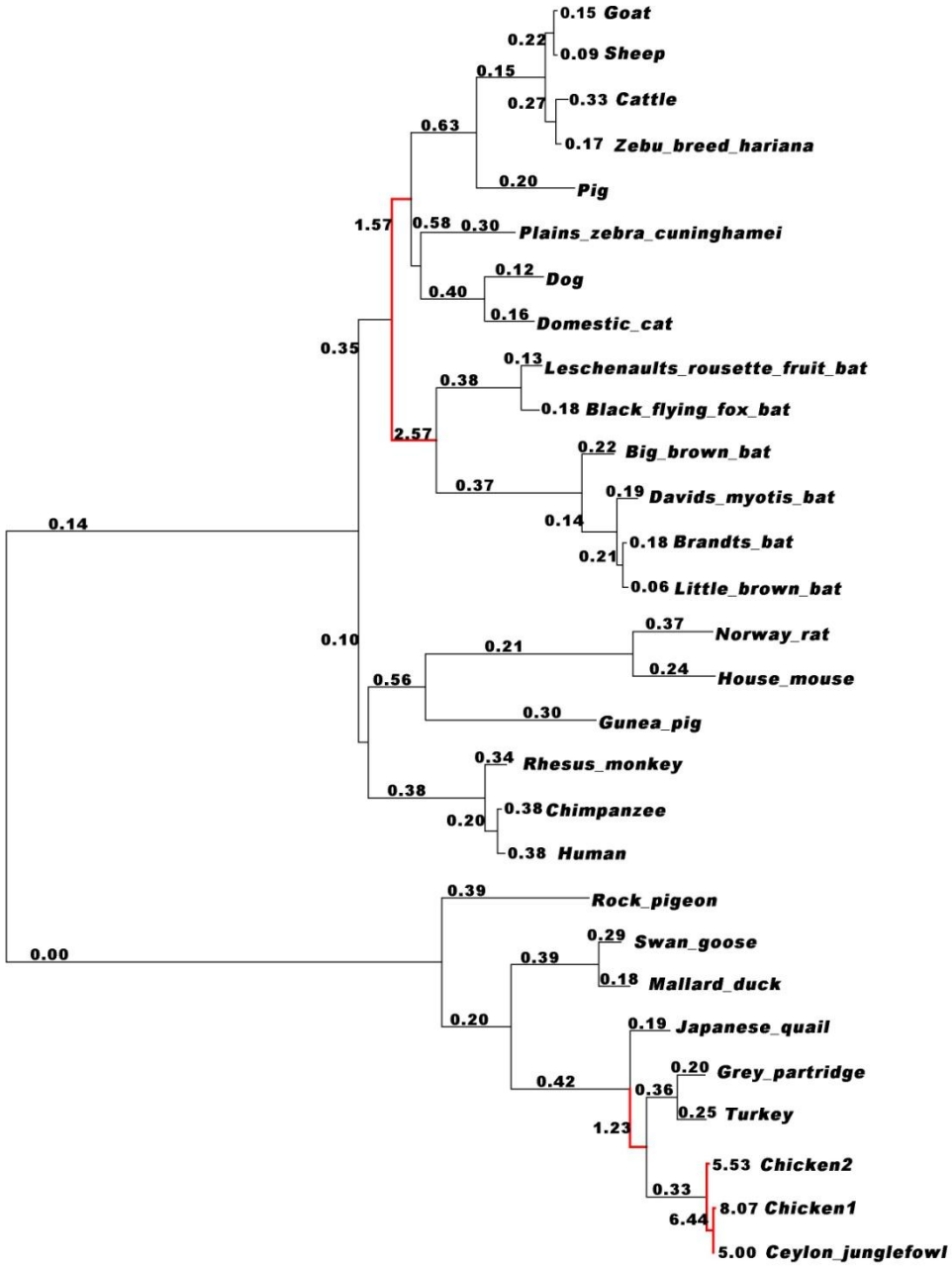


Figure 1: Phylogeny of TLR7 gene. Colored branches represent lineages undergoing adaptive evolution (ω>1).

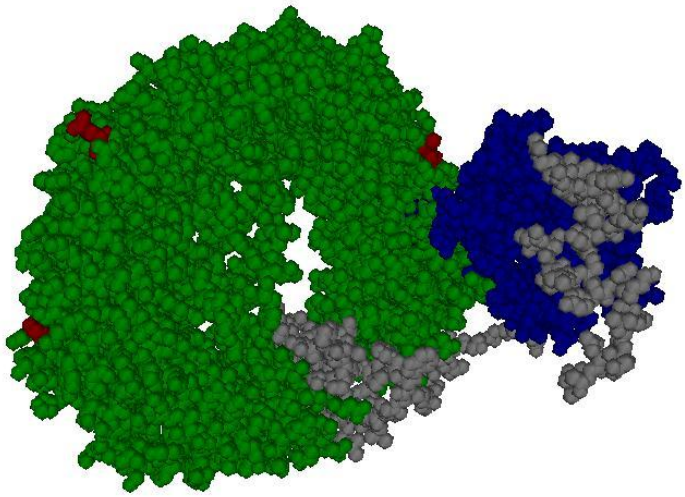


Fig. 2: The 3D structure of chicken TLR7. The blue residues represent the TIR domain while the green residues represent the LRR domain. The red residues represent sites under positive selection (Serine at position 275 with a BEB value of 0.981*, Glutamic acid at 380 with a BEB value of 0.969* and Leucine at 689 with a BEB value of 0.961*).

Toll-like Receptor3

We obtained significant LRT results for TLR3 (Table 2). Like TLR7, we detected predominant purifying signatures and positive signatures which occurred along the zebu, bat, domestic and wild poultry lineages (fig. 3). A previous study in mammals by Areal at al., (2011) found similar results. In addition, we detected a positive signature at the Leucine Rich Repeats (LRRs) domain (fig. 4). However, our results contrast the research findings of Darfour-Oduro et al., (2015) who discovered only predominant purifying selection of TLR3 in the family Suidae. For both TLR7 and TLR3, we observed that the TIR domain is highly conserved unlike the LRR domain. This extends previous findings in humans, primates, avians, murines and some domesticates where the LRR domain was found to be more frequently targeted by positive selection than the TIR domain (Alcaide & Edwards, 2011; Barreiro et al., 2009; Fornůsková et al., 2013; Grueber et al., 2014; Matsushima et al., 2007; Quach et al., 2013; Vinkler et al., 2014; Werling et al., 2009). Other studies have also revealed that the ectodomains of all TLRs evolve more rapidly than the TIR domain across vertebrate species (Areal et al., 2011; Mikami et al., 2012; Wlasiuk & Nachman, 2010).

Table 2: Likelihood ratio tests (LRTs) for TLR3

Models	$2\Delta\ell$	$\chi 2$ Value	D. f	P-value	Model favored
Lineage Analysis (M0 v M1)	2(-20346.89 -20288.26)	117.26	59	P<0.001	M1
Codon Site Analysis (M7 v M8)	2(-19954.76 -19946.81)	15.9	2	P<0.001	M8

Fig. 3: Phylogeny of TLR3 gene. Colored branches represent lineages undergoing adaptive evolution (ω>1).

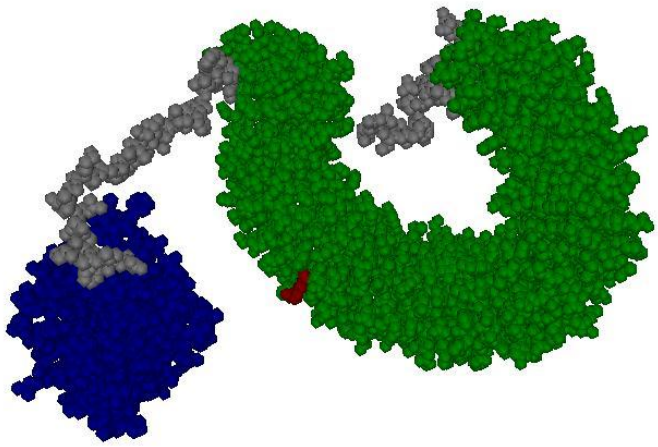


Fig. 4: The 3D structure of chicken TLR3. The blue residues represent the TIR domain while the green residues represent the LRR domain. The red residues represent sites under positive selection (Argine at position 440 with a BEB value of 0.956*).

2' 5' Oligoadenylate synthetase

Our LRT results for OAS were significant (Table 3). We detected positive signatures along the lineages of poultry, bats, domestic ferret, cat and dog (fig.5). In addition, codon site analysis detected 1 positive signature in the OAS1_C domain (fig. 6). The 2' 5' OAS gene family has been extensively studied in humans and mice and is characterized by extensive gene duplications and domain coupling that gave rise to several isoforms such as OAS1, OAS2, OAS3 and OASL (Kristiansen et al., 2011; Kumar, et al., 2000; Perelygin et al., 2006). In this study, the OASL isoform was selected since it's the only one that has been isolated in poultry. Although limited studies have been carried out for this isoform, studies in the OAS1 paralog have reported similar findings to our results. For instance, Hancks et al., (2015) and Mozzi et al., (2015) detected numerous positive signatures across the OAS1 gene of primates and bats which contrasted with OASL. Although many PRR genes have typical RNA binding domains, none has been identified for the 2'5' OAS genes and are therefore thought to interact with viral RNA in a sequence unspecific manner (Fierro-Monti & Mathews, 2000; Hartmann et al., 1998; J Justesen, 2000; Sarkar & Sen, 1998). The extensive duplications and domain couplings of this gene could also be another mechanism through which it escapes viral inhibitors as has previously been reported by Hancks et al., (2015).

Table 3: Likelihood ratio tests (LRTs) for OAS

Models	$2\Delta\ell$	$\chi 2$ Value	D. f	P-value	Model favored
Lineage Analysis (M0 v M1)	2(-13320.98 -13267.12)	107.72	53	P<0.001	M1
Codon Site Analysis (M7 v M8)	2(-13021.60 -13017.41)	8.38	2	P<0.002	M8

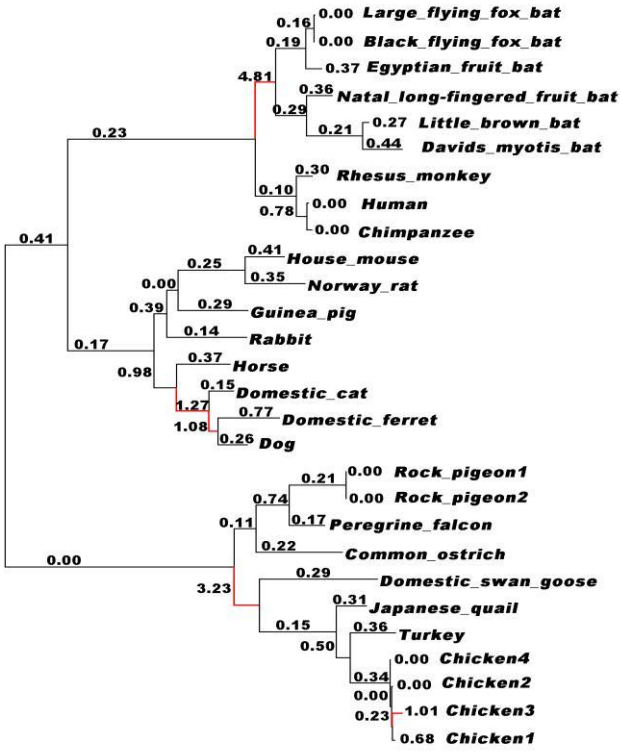


Fig. 5: Phylogeny of OAS gene. Colored branches represent lineages undergoing adaptive evolution (ω>1).

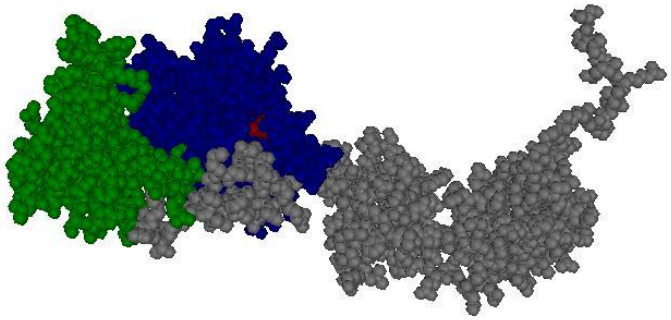


Fig. 6: The 3D structure of chicken OAS. The blue residues represent the OAS1_C domain while the green residues represent the NTase domain. The red residues represent sites under positive selection (Leucine at position 2272 with a BEB value of 0.992**).

Protein kinaseR

Like TLR7, TLR3 and OAS, we detected both purifying and adaptive signatures in PKR and the LRT results were highly significant (Table 4). We observed positive signatures in the poultry, bat, primate and livestock lineages (fig. 7). Site analysis revealed numerous adaptive signatures across the entire structure with a greater concentration in the dsRBM domain which binds viral dsRNA and PKC domain where substrate (eIF2α) phosphorylation occurs (fig. 8). Table 5 shows the amino acid substitutions, positions and BEB values > 0.95. This is consistent with the findings of Elde et al., (2009) who detected positive signatures in the three domains of PKR genes of primates with a greater concentration in the PKC domain. In yet another study among vertebrates which included chicken, Rothenburg et al., (2009) detected accelerated evolution in the Protein kinase C domain of PKR. This can be attributed to its multiple families of constantly evolving viral inhibitors/antagonists which exert strong selective pressures that subject it to strong adaptive evolution. In addition, viral mimicry of PKR substrate is another factor that could be driving the rapid evolution of the PKC domain. For instance, the rapidly evolving K3L that is encoded by Poxviruses has been shown to impose strong selective pressures at the PKC domain since it shares homology with the N-terminus of the PKR substrate, eIF2α, hence acts as a pseudo substrate (Elde et al., 2009). These numerous signatures could therefore be driven by the diverse viral antagonists and strong selective pressures aimed at evading viral mimicry.

Table 4: Likelihood ratio tests (LRTs) for PKR

Models	2Δℓ	χ2 Value	D. f	P-value	Model favored
Lineage Analysis (M0 v M1)	2(-15072.85 -15009.09)	127.52	53	P<0.001	M1
Codon Site Analysis (M7 v M8)	2(-14530.76	136.38	2	P<0.001	M8

	-14462.57)				

Table 5: BEB results for PKR (*: P>0.95%; **: P>0.99%)

Position	Amino Acid	BEB	Position	Amino Acid	BEB	Position	Amino Acid	BEB
59	Lysin	0.975*	134	Glutamine	0.980*	265	Proline	0.987*
71	Asparagine	0.956*	136	Glutamine	0.997**	266	Asparagine	0.989*
76	Proline	0.995**	145	Alanine	0.994**	365	Histidine	0.965*
116	Glutamine	1.000**	167	Glutamic acid	1.000**	369	Asparagine	0.999**
120	Serine	0.991**	170	Argine	0.995**	379	Aspartic acid	0.998**
124	Valine	0.996**	171	Glutamine	0.987*	396	Glutamic acid	0.994**
125	Histidine	0.966*	233	Threonine	0.996**			
133	Glycine	0.965*	260	Asparagine	0.974*			

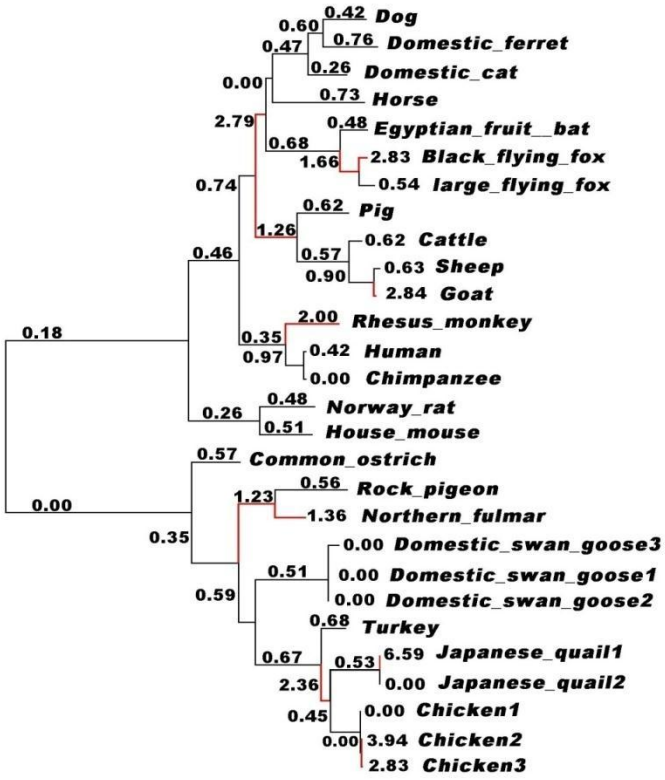


Fig. 7: Phylogeny of PKR gene. Colored branches represent lineages undergoing adaptive evolution (ω>1).

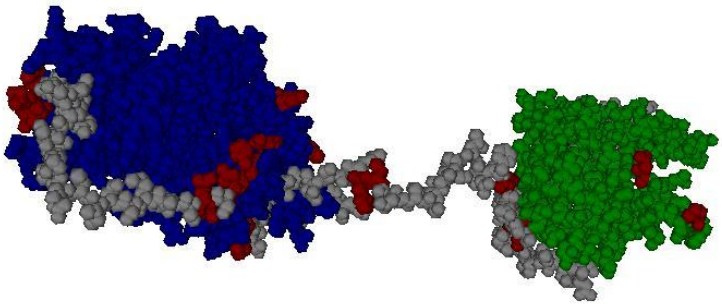


Fig. 8: The 3D structure of chicken PKR. The blue residues represent the Protein Kinase domain while the green residues represent the dsRBM (Double-stranded RNA binding motif) domain. The red residues represent sites under positive selection. The residues, positions and BEB values are shown in Table 5.

Signatures of selection at heat shock protein genes

Heat shock proteins have been thought to play an evolutionary and ecologically important role in thermal adaptation of organisms to extreme temperatures (Feder and Hofmann 1999; parsel and Lindquist 1993). However, no molecular evolution studies have previously been reported in exotic and indigenous poultry species. Across all our select genes, we detected predominant purifying selection. This is suggestive of evolutionary conservation that could be driven by functional constraints aimed at maintaining the structural and functional integrity of the genes and gene products.

Heat shock protein70

Among all families of HSPs, HSP70 has been studied widely and found to be highly conserved (Wang et al., 2015). Likewise, though we detected a positive signature along the chicken and other avian lineages, results of this study revealed strong purifying selection in all the select poultry species and homologs (fig. 9). Although M8 was favored over M7, codon site analysis detected no positive signatures (Table 6). This is in line with previous findings of Gade et al., (2010) who concluded that HSP70 gene is highly conserved in domestic animals. The detected positive signature may have occurred as a result of the various mechanisms of gene evolution such as gene conversion and gene duplication events. Similarly, these events have previously been reported in other vertebrate species such as humans, pigs, mice and rats (Günther & Walter, 1994). Alternatively, episodic positive selection and domestication processes can result in diversification and functional adaptations that are later maintained through purifying selection.

Table 6: Likelihood ratio tests (LRTs) for HSP70

Models	2Δℓ	χ2 Value	D. f	P-value	Model favored
Lineage Analysis (M0 v M1)	2(-8528.33 -8477.54)	101.58	45	P<0.001	M1
Codon Site Analysis (M7 v M8)	2(-8481.10 -8476.51)	9.18	2	P<0.002	M8

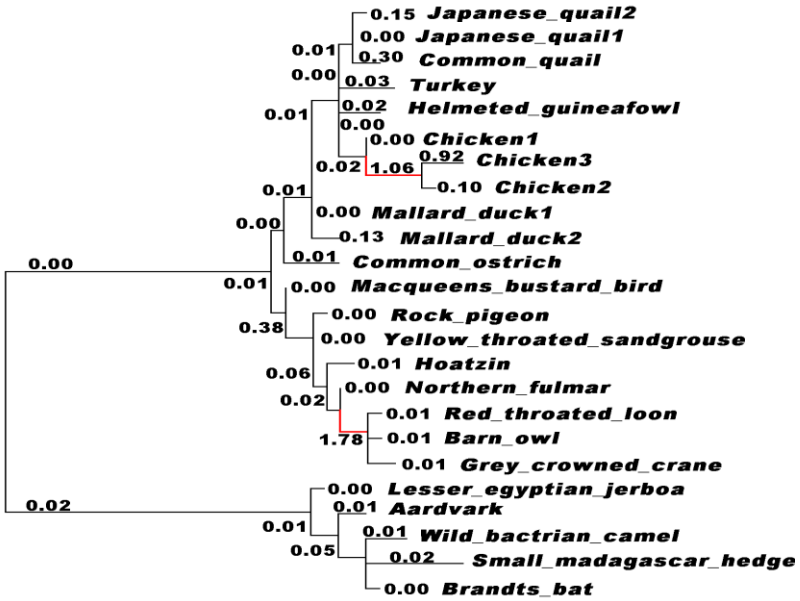


Fig. 9: Phylogeny of HSP70 gene. All branches had values purifying signatures (ω>1).

Heat shock protein90

Except for the positive signature along brandts bat lineage, we detected strong purifying selection in HSP90 (fig. 10). In addition, we did not detect any positive signatures from codon site analysis and M7 was favored over M8 (Table 7). Similar to HSP70, limited evolutionary studies have been carried out for HSP90. However, similar to the findings of this study, it has been observed that the HSP90 gene is highly conserved across all organisms (Csermely et al., 1998; Pantzartzi et al., 2013). Nevertheless, contrary to our findings, a study in 54 species of the main eukaryotic lineages (vertebrates, seed plant and yeast) revealed signatures of positive selection which were associated with gene duplications and subsequent functional diversifications (Carretero-Paulet et al., 2013).

Table 7: Likelihood ratio tests (LRTs) for HSP90

Models	2Δℓ	χ2 Value	D. f	P-value	Model favored
Lineage Analysis (M0 v M1)	2(-10533.68 -10494.20)	78.96	45	P<0.002	M1

Codon Site Analysis (M7 v M8)	2(-10502.96 -10502.27)	1.38	2	P>0.20	M7

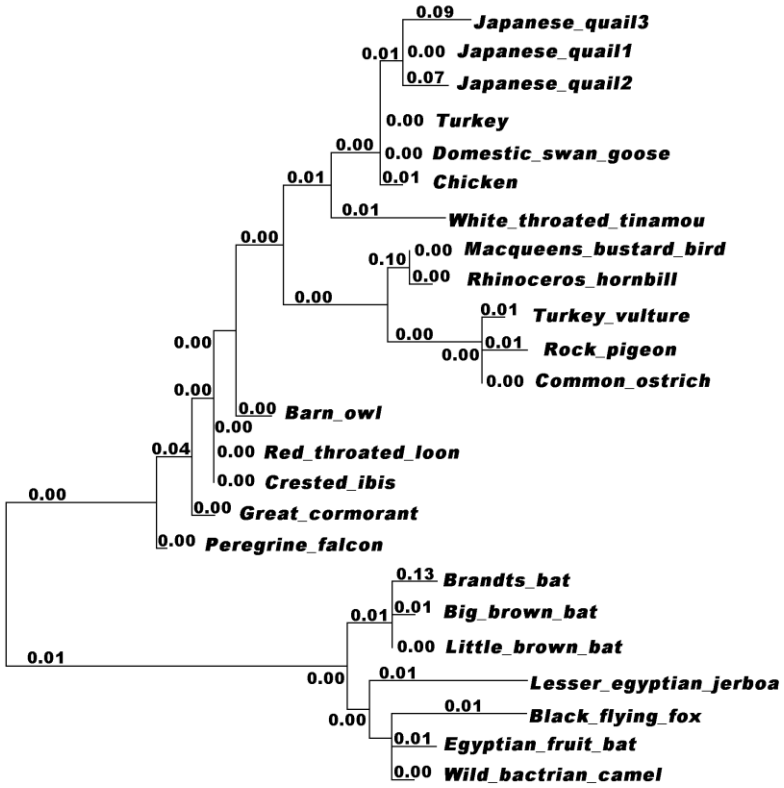


Fig. 10: Phylogeny of HSP90 gene. All branches had purifying signatures (ω>1).

Small heat shock protein

We obtained only predominant purifying signatures in sHSP gene (fig. 11). Also, we did not detect positive signatures from site analysis and M7 was favored over M8 (Table 8). No studies have previously been reported on the evolution of sHSps in specific organisms, However, reports have been published that indicate that sHSPs are highly conserved across species (Bakthisaran et al., 2015; Jong et al, 1998; Franck et al., 2004; Haslbeck & Vierling, 2015; Jakob et al., 1993b). This is consistent with the findings of this study.

Table 8: Likelihood ratio tests (LRTs) for sHSP

Models	**2Δℓ**	**χ2 Value**	**D. f**	**P-value**	**Model favored**
Lineage Analysis (M0 v M1)	2(-3066.10 -3033.98)	64.24	49	P<0.05	M0

Codon Site Analysis (M7 v M8)	2(-3062.84 -3062.83)	0	2	P>0.975	M7

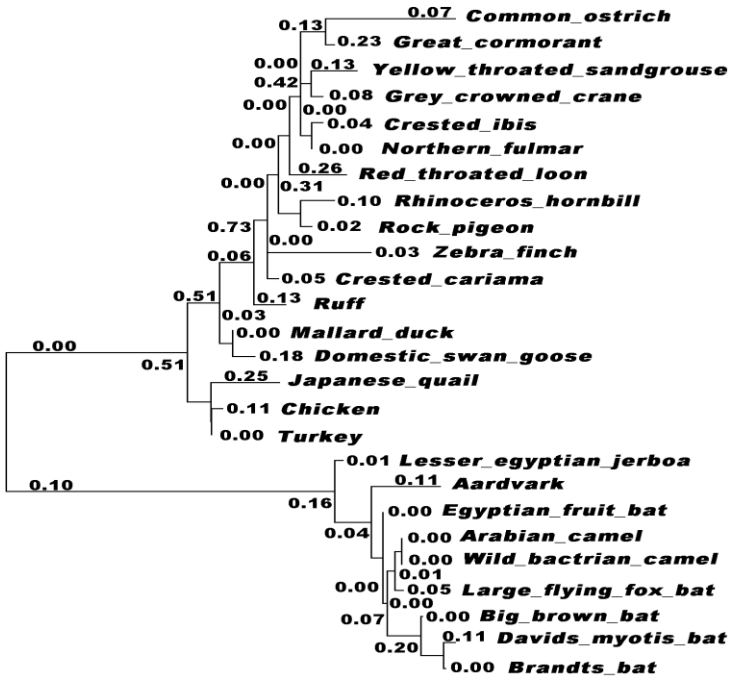


Fig. 11: Phylogeny of sHSP gene. All branches had purifying signatures (ω>1).

Conclusions

We concluded that heat-stress genes evolve under strong purifying selection which could be driven by functional constraints. On the other hand, we found evidence of adaptive evolution in all our select innate immune genes. The location and distribution of the positively selected codons strongly suggest the role of pathogens in exerting selective pressures and shaping the diversity and variability of these genes. This makes them a promising target for further experimental validation through *in vitro* and *in vivo* studies. The results obtained can be used in genetic improvement and conservation of poultry, a species that is threatened by existing and emerging infectious viral diseases.

Data availability

Dataset 1: Homologs, Protein accession numbers and Expectation values for TLR7

Animal Species	**Protein Accession Nos.**	**E-values**
Gallus gallus1 (chicken)	ACR26208.1	0.0
Gallus gallus2 (chicken)	XP_015129128.1	0.0
Gallus lafayetii (ceylon junglefowl	ACR26206.1	0.0

Perdix perdix (grey partridge)	AGO86775.1	0.0
Coturnix japonica (japanese quail)	XP_015742697.1	0.0
Meleagris gallopavo (turkey)	XP_003203134.1	0.0
Anser cygnoides domesticus (swan goose)	XP_013046996.1	0.0
Anas platyrhynchos (mallard)	XP_005029236.1	0.0
Columba livia (rock pigeon)	AIK67344.1	0.0
Sus scrofa (pig)	ABG47422.1	0.0
Equus burchellii cuninghamei (plains zebra)	AGK25872.1	0.0
Canis lupus familiaris (dog)	ABC69204.1	0.0
Rousettus leschenaultii (leschenault's rousette fruitbat)	BAH02556.1	0.0
Felis catus (domestic Cat)	NP_001073602.1	0.0
Capra hircus (goat)	XP_005701170.1	0.0
Bos taurus (cattle)	ABN71673.1	0.0
Bos indicus (bos taurus indicus-zebu breed harianna)	ACY25086.1	0.0
Macaca mulatta (rhesus monkey)	NP_001123898.1	0.0
Pan troglodytes (chimpanzee)	NP_001123605.1	0.0
Eptesicus fuscus (big brown bat)	XP_008154799.1	0.0
Myotis brandtii (brandt's bat)	XP_005881008.1	0.0
Myotis lucifugus (little brown bat)	XP_006088669.1	0.0
Myotis davidii (davids myotis bat)	XP_006763859.1	0.0
Mus musculus (house mouse)	NP_001277687.1	0.0
Homo sapiens (human)	AAF78035.1	0.0
Ovis aries (sheep)	NP_001128531.1	0.0
Pteropus alecto (black flying fox)	NP_001277093.1	0.0
Rattus norvegicus (Norway rat)	NP_001091051.1	0.0
Cavia porcellus (domestic guinea pig)	XP_003462941.2	0.0

Dataset 2: Homologs, Protein accession numbers and Expectation values for TLR3

Animal Species	**Protein Accession Nos.**	**E-values**
Gallus gallus1 (chicken)	ABL74502.1	0.0
Gallus gallus2 (chicken)	XP_015140918.1	0.0
Meleagris gallopavo (turkey)	XP_003205822.1	0.0
Anas platyrhynchos (mallard duck)	XP_005009038.1	0.0
Cairina moschata (Muscovy duck)	AFK29094.1	0.0
Coturnix japonica (japanese quail)	XP_015717606.1	0.0
Columba livia (rock pigeon)	XP_005500267.1	0.0
Anser cygnoides domesticus (domestic Swan goose)	XP_013035222.1	0.0
Gallus varius (green junglefowl)	ACR26347.1	0.0
Gallus lafayetii (sri lankan junglefowl)	ACR26327.1	0.0
Gallus sonneratii (grey junglefowl)	ACR26351.1	0.0
Struthio camelus australis (common ostrich)	XP_009674995.1	0.0
Anser anser (greylag goose)	AGJ98456.1	0.0
Felis catus (domestic cat)	XP_006930623.1	0.0
Myotis lucifugus (little brown bat)	XP_006092716.1	0.0

Eptesicus fuscus (big brown bat)	XP_008150129.1	0.0
Myotis davidii (davids myotis bat)	XP_006772770.1	0.0
Macaca mulatta (rhesus monkey)	ABY64988.1	0.0
Canis lupus familiaris (dog)	XP_005630024.1	0.0
Rousettus leschenaultii (leschenault's rousette fruitbat	BAH02555.1	0.0
Rattus norvegicus (norway rat)	XP_008769488.1	0.0
Capra hircus (goat)	AHJ90636.1	0.0
Sus scrofa (pig)	ADQ00195.1	0.0
Pan troglodytes verus (west african chimpanzee)	ADH84437.1	0.0
Bos indicus (bos taurus indicus)	ACU16426.1	0.0
Bos taurus (cattle)	ABN71661.1	0.0
Homo sapiens (human)	ABC86908.1	0.0
Ovis aries (sheep)	NP_001129400.1	0.0
Mus musculus (house mouse)	AAH99937.1	0.0
Pteropus alecto (black flying fox)	NP_001277098.1	0.0
Oryctolagus cuniculus algirus (rabbit)	AGU70373.1	0.0

Dataset 3: Homologs, Protein accession numbers and Expectation values for OAS1

Animal Species	**Protein Accession Nos.**	**E-values**
Gallus gallus1 (chicken)	BAB19016.1	0.0
Gallus gallus2 (chicken)	XP_015148492.1	0.0
*Gallus gallus*3 (chicken) 3	BAB19015.1	0.0
Gallus gallus4 (chicken)	NP_990372.1	0.0
Anser cygnoides domesticus (domestic swan goose)	XP_013047372.1	0.0
Struthio camelus australis (common ostrich)	XP_009671383.1	0.0
Columba livia1 (rock pigeon)	XP_005508920.1	0.0
Columba livia2 (rock pigeon)	EMC83969.1	0.0
Meleagris gallopavo (turkey)	XP_010716729.1	0.0
Coturnix japonica (japanese quail)	XP_015730133.1	0.0
Falco peregrinus (peregrine falcon)	XP_005229859.1	0.0
Felis catus (domestic cat)	XP_003994784.1	7e-127
Canis lupus familiaris (dog)	NP_001041558.1	9e-126
Equus caballus (horse)	XP_001488427.3	1e-125
Mustela putorius furo (domestic ferret)	XP_004753444.1	2e-125
Mus musculus (house mouse)	NP_035984.2	1e-108
Rattus norvegicus (norway rat)	NP_001009682.1	3e-105
Macaca mulatta (rhesus monkey)	XP_001091486.1	7e-105
Pan troglodytes (chimpanzee)	NP_001267398.1	3e-101
Oryctolagus cuniculus (rabbit)	XP_002722162.1	9e-119
Cavia porcellus (domestic guinea pig)	XP_003477780.1	7e-117
Homo sapiens (human)	NP_003724.1	2e-99
Myotis lucifugus (little brown bat)	XP_006096634.1	7e-82
Pteropus vampyrus (large flying fox)	XP_011354201.1	5e-100
Myotis davidii (davids myotis)	XP_006770414.1	3e-71

Pteropus alecto (black flying fox)	XP_006908689.1	5e-100
Miniopterus natalensis (natal long-fingered bat)	XP_016057043.1	2e-71
Rousettus aegyptiacus (egyptian fruitbat)	XP_015982708.1	4e-95

Dataset 4: Homologs, Protein accession numbers and Expectation values for PKR

Animal Species	**Protein Accession Nos.**	**E-values**
Gallus gallus1 (chicken)	NP_989818.1	0.0
Gallus gallus2 (chicken)	XP_015139096.1	0.0
Gallus gallus3 (chicken)	XP_015139098.1	0.0
Meleagris gallopavo (turkey)	XP_003204013.1	0.0
Coturnix japonica1 (japanese quail)	XP_015713050.1	0.0
Coturnix japonica2 (japanese quail)	XP_015713051.1	0.0
Anser cygnoides domesticus1 (domestic swan goose)	XP_013026149.1	0.0
Anser cygnoides domesticus2 (domestic swan goose)	XP_013026158.1	0.0
Anser cygnoides domesticus3 (domestic swan goose)	XP_013026172.1	0.0
Fulmarus glacialis (northern fulmar)	XP_009579791.1	0.0
Struthio camelus australis (common ostrich)	XP_009684950.1	0.0
Columba livia (rock pigeon)	XP_005503897.1	0.0
Mustela putorius furo (domestic ferret)	XP_004812890.1	9e-87
Felis catus (domestic cat)	XP_003984379.1	5e-86
Homo sapiens (human)	AAF13156.1	1e-85
Sus scrofa (pig)	NP_999484.1	3e-85
Canis lupus familiaris (dog)	NP_001041600.1	9e-85
Pan troglodytes (chimpanzee)	NP_001138509.1	3e-83
Rattus norvegicus (norway rat)	NP_062208.1	3e-83
Bos taurus (cattle)	XP_005212627.1	2e-82
Capra hircus (goat)	XP_005686545.1	2e-82
Equus caballus (horse)	NP_001137272.1	3e-79
Ovis aries (sheep)	XP_004007349.1	1e-72
Macaca mulatta (rhesus monkey)	NP_001077417.1	2e-70
Mus musculus (house mouse)	NP_035293.1	1e-61
Pteropus vampyrus (large flying fox)	XP_011359712.1	7e-78
Rousettus aegyptiacus (egyptian fruitbat)	XP_016017983.1	5e-76
Pteropus alecto (black flying fox)	XP_006910452.1	1e-79

Dataset 5: Homologs, Protein accession numbers and Expectation values for HSP70

Animal Species	**Protein Accession Nos.**	**E-values**
Gallus gallus1 (chicken1)	AAP37959.1	0.0
Gallus gallus2 (chicken)	ACI31545.1	0.0
Gallus gallus3 (chicken)	ACE79190.1	0.0
Anas platyrhynchos1 (mallard duck)	XP_005022715.1	0.0
Coturnix japonica1 (Japanese quail)	BAF37039.1	0.0
Coturnix japonica2 (Japanese quail)	NP_001310127.1	0.0

Anas platyrhynchos2 (Mallard duck)	NP_001297704.1	0.0
Meleagris gallopavo (turkey)	XP_003206814.1	0.0
Fulmarus glacialis (northern fulmar)	XP_009574733.1	0.0
Columba livia (rock pigeon)	XP_005506432.1	0.0
Coturnix coturnix (common quail)	ACC85671.1	0.0
Numida meleagris (helmeted guineafowl)	BAC24791.1	0.0
Gavia stellata (red-throated loon)	KFV42297.1	0.0
Struthio camelus australis (common ostrich)	XP_009673875.1	0.0
Jaculus jaculus (lesser egyptian jerboa)	XP_004649317.1	0.0
Orycteropus afer afer (aardvark)	XP_007940478.1	0.0
Echinops telfairi (small Madagascar hedgehog)	XP_004698706.1	0.0
Camelus ferus (wild Bactrian camel)	XP_006177814.1	0.0
Myotis brandtii (brandt's bat)	XP_005874088.1	0.0
Chlamydotis macqueenii (macqueens bustard bird)	XP_010125825.1	0.0
Tyto alba (barn owl)	XP_009973348.1	0.0
Opisthocomus hoazin (hoatzin)	XP_009941163.1	0.0
Pterocles gutturalis (yellow throated sandgrouse)	XP_010077649.1	0.0
Balearica regulorum gibbericeps (grey crowned crane)	XP_010296761.1	0.0

Dataset 6: Homologs, Protein accession numbers and Expectation values for HSP90

Animal Species	**Protein Accession Nos.**	**E-values**
Gallus gallus (chicken)	NP_001103255.1	0.0
Coturnix japonica1 (japanese quail)	BAI23206.1	0.0
Meleagris gallopavo (turkey)	XP_010710229.1	0.0
Anser cygnoides domesticus (domestic swan goose)	XP_013053698.1	0.0
Coturnix japonica2 (japanese quail)	NP_001310124.1	0.0
Tyto alba (barn owl)	XP_009972463.1	0.0
Coturnix japonica3 (japanese quail)	BAI23210.1	0.0
Gavia stellata (red-throated loon)	XP_009817830.1	0.0
Buceros rhinoceros silvestris (rhinoceros hornbill)	XP_010132756.1	0.0
Tinamus guttatus (white-throated tinamou)	XP_010213538.1	0.0
Nipponia nippon (crested ibis)	XP_009465435.1	0.0
Phalacrocorax carbo (great cormorant)	XP_009508855.1	0.0
Falco peregrinus (peregrine falcon)	XP_005243017.1	0.0
Chlamydotis macqueenii (macqueens bustard bird)	KFP45671.1	0.0
Struthio camelus australis (common ostrich)	XP_009673798.1	0.0
Cathartes aura (turkey vulture)	KFP47363.1	0.0
Columba livia (rock pigeon)	XP_005506148.1	0.0
Myotis lucifugus (little brown bat)	XP_006100595.1	0.0
Eptesicus fuscus (big brown bat)	XP_008152985.1	0.0
Camelus ferus (wild Bactrian camel)	XP_006175722.1	0.0
Myotis brandtii (brandt's bat1)	XP_005877923.1	0.0
Jaculus jaculus (lesser egyptian jerboa)	XP_004665641.1	0.0

Rousettus aegyptiacus (egyptian fruitbat)	XP_016021504.1	0.0
Pteropus alecto (black flying fox)	XP_006925646.1	0.0

Dataset 7: Homologs, Protein accession numbers and Expectation values for sHSP

Animal Species	**Protein Accession Nos.**	**E-values**
Gallus gallus (chicken)	NP_990507.1	7e-121
Meleagris gallopavo (turkey)	XP_003212808.1	2e-120
Anas platyrhynchos (mallard duck)	NP_001297295.1	4e-119
Anser cygnoides domesticus (domestic swan goose)	XP_013042703.1	1e-118
Coturnix japonica (japanese quail)	XP_015739320.1	3e-119
Fulmarus glacialis (northern fulmar)	XP_009578108.1	1e-115
Pterocles gutturalis (yellow-throated sandgrouse)	XP_010080930.1	2e-115
Balearica regulorum gibbericeps (crowned crane)	KFO09036.1	4e-115
Phalacrocorax carbo (great cormorant)	XP_009506073.1	5e-115
Buceros rhinoceros silvestris (rhinoceros hornbill)	XP_010140908.1	9e-114
Taeniopygia guttata (zebra finch)	XP_002192920.1	9e-113
Gavia stellata (red-throated loon)	XP_009808997.1	4e-112
Struthio camelus australis (common ostrich)	XP_009668516.1	2e-111
Cariama cristata (crested cariama)	XP_009694886.1	9e-116
Nipponia nippon (crested ibis)	XP_009470546.1	1e-116
Calidris pugnax (ruff)	XP_014805009.1	5e-117
Columba livia (rock pigeon)	XP_005500801.1	3e-112
Jaculus jaculus (lesser egyptian jerboa)	XP_004666540.1	1e-92
Camelus ferus (wild bactrian camel)	XP_006190435.1	2e-90
Eptesicus fuscus (big brown bat)	XP_008148383.1	3e-90
Orycteropus afer afer (aardvark)	XP_007934744.1	5e-90
Myotis brandtii (brandt's bat)	XP_005856898.1	2e-89
Pteropus vampyrus (large flying fox)	XP_011363711.1	7e-92
Rousettus aegyptiacus (egyptian fruitbat)	XP_016004541.1	8e-92
Camelus dromedaries (arabian camel)	XP_010984284.1	2e-91
Myotis davidii (davids myotis bat)	XP_006761589.1	2e-90

AUTHOR CONTRIBUTIONS

SO conceived the concept. SC, EW, SO, PO, and SM performed the analysis. DK, EN, SM, MM, PO, JL, and SO guided research concept. All authors participated in writing.

COMPETING INTERESTS

No competing interests were disclosed.

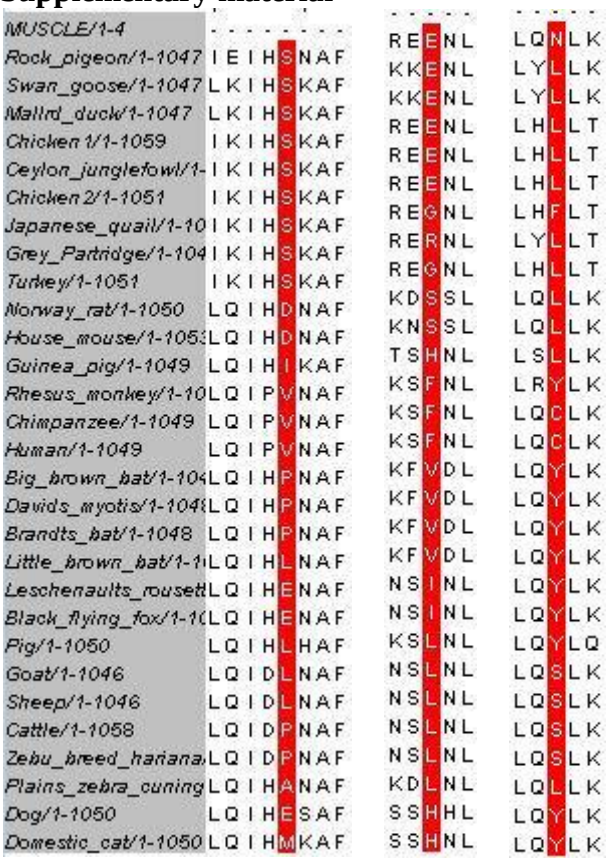
GRANT INFORMATION

The authors wish to acknowledge financial support awarded to Dr. Sheila Ommeh by the Jomo Kenyatta University of Agriculture and Technology research grant no. JKU/2/4/RP/181.

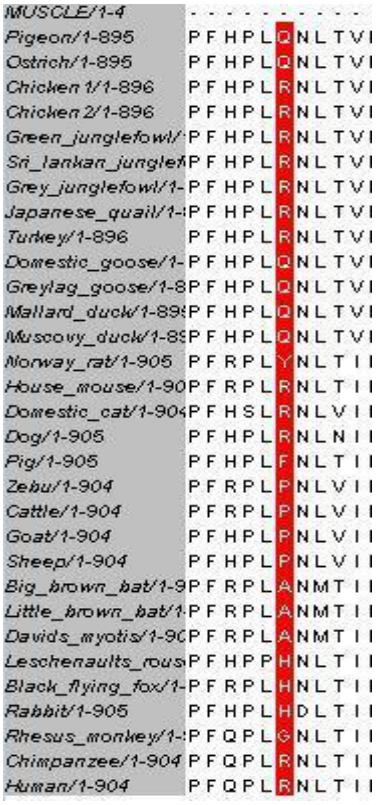
ACKNOWLEDGEMENTS

Sincere gratitude goes to Jomo Kenyatta University of Agriculture and Technology, Ministry of Agriculture Livestock, and Fisheries, the State Department of Veterinary services, and the National Museums of Kenya for timely provision of all the necessary facilities that made this study a success.

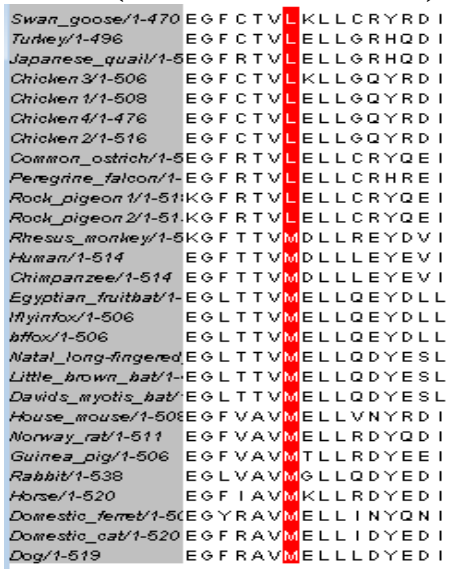
Supplementary material



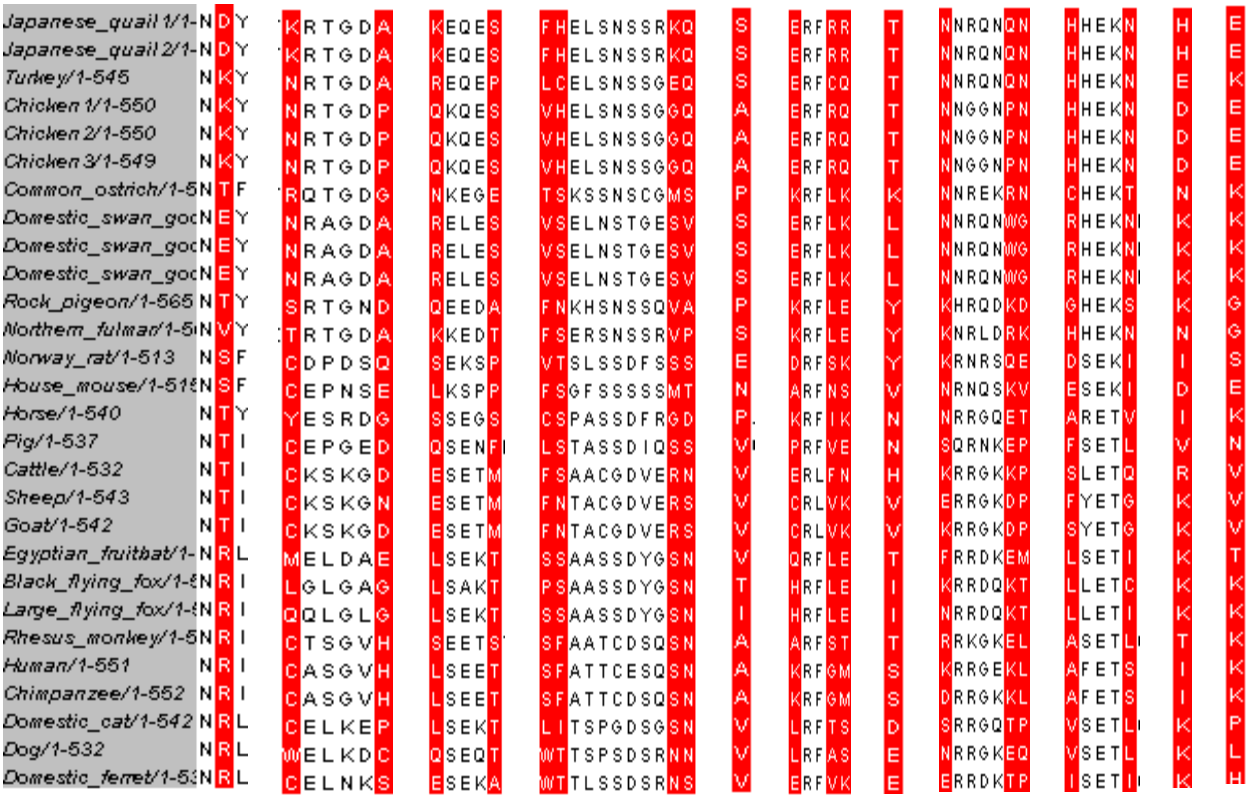
Supplementary figure 1: Codon sites under positive selection at TLR7 gene as visualized in JalView (Waterhouse et al., 2009).



Supplementary figure 2: Codon sites under positive selection at TLR3 gene as visualized in JalView (Waterhouse et al., 2009)



Supplementary figure 3: Codon sites under positive selection at OAS gene as visualized in JalView (Waterhouse et al., 2009).



Supplementary figure 4: Codon sites under positive selection at PKR gene as visualized in JalView (Waterhouse et al., 2009).

References

Alcaide, M., & Edwards, S. V. (2011). Molecular evolution of the toll-like receptor multigene family in birds. *Molecular Biology and Evolution*, *28*(5), 1703–1715.

Areal, H., Abrantes, J., & Esteves, P. J. (2011). Signatures of positive selection in Toll-like receptor (TLR) genes in mammals. *BMC Evolutionary Biology*, *11*(1), 368.

Bakthisaran, R., Tangirala, R., & Rao, C. M. (2015). Small heat shock proteins: Role in cellular functions and pathology. *Biochimica Et Biophysica Acta*, *1854*(4), 291–319.

Brennan, G., Kitzman, J. O., Rothenburg, S., Shendure, J., & Geballe, A. P. (2014). Adaptive Gene Amplification As an Intermediate Step in the Expansion of Virus Host Range. *PLoS Pathog*, *10*(3), e1004002.

Child, S. J., Brennan, G., Braggin, J. E., & Geballe, A. P. (2012). Species Specificity of Protein Kinase R Antagonism by Cytomegalovirus TRS1 Genes. *Journal of Virology*, *86*(7), 3880–3889.

C. S. Mukhopadhyay, D. K. (2012). SNP chip development and genome wide association studies in livestock (Theory & Practical Session).

Darfour-Oduro, K. A., Megens, H.-J., Roca, A. L., Groenen, M. A. M., & Schook, L. B. (2015). Adaptive Evolution of Toll-Like Receptors (TLRs) in the Family Suidae. *PLoS ONE*, *10*(4).

Dekkers, J. C. M. (2012). Application of Genomics Tools to Animal Breeding. *Current Genomics*, *13*(3), 207–212.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792–1797.

Elde, N. C., Child, S. J., Geballe, A. P., & Malik, H. S. (2009). Protein kinase R reveals an evolutionary model for defeating viral mimicry. *Nature*, *457*(7228), 485–489.

FAO. (2014). *Decision Tools for Family Poultry Development* (1 edition). Rome: FAO.

Fay, J. C., & Wu, C. I. (2000). Hitchhiking under positive Darwinian selection. *Genetics*, *155*(3), 1405–1413.

Ferguson, W., Dvora, S., Fikes, R. W., Stone, A. C., & Boissinot, S. (2012). Long-Term Balancing Selection at the Antiviral Gene OAS1 in Central African Chimpanzees. *Molecular Biology and Evolution*, *29*(4), 1093–1103.

Fornůsková, A., Vinkler, M., Pagès, M., Galan, M., Jousselin, E., Cerqueira, F., Cosson, J.-F. (2013). Contrasted evolutionary histories of two Toll-like receptors (Tlr4 and Tlr7) in wild rodents (MURINAE). *BMC Evolutionary Biology*, *13*, 194.

Gade, N., Mahapatra, R. K., Sonawane, A., Singh, V. K., Doreswamy, R., & Saini, M. (2010). Molecular Characterization of Heat Shock Protein 70-1 Gene of Goat (Capra hircus). *Molecular Biology International*, *2010*.

Gardner, E. G. (2014). Livestock Risks and Opportunities: Newcastle Disease and Avian Influenza in Africa. *Planet@Risk*, *2*(4).

Gouy, M., Guindon, S., & Gascuel, O. (2010). SeaView Version 4: A Multiplatform Graphical User Interface for Sequence Alignment and Phylogenetic Tree Building. *Molecular Biology and Evolution*, *27*(2), 221–224.

Hancks, D. C., Hartley, M. K., Hagan, C., Clark, N. L., & Elde, N. C. (2015). Overlapping Patterns of Rapid Evolution in the Nucleic Acid Sensors cGAS and OAS1 Suggest a Common Mechanism of Pathogen Antagonism and Escape. *PLoS Genetics*, *11*(5).

Howard, C. R., & Fletcher, N. F. (2012). Emerging virus diseases: can we ever expect the unexpected? *Emerging Microbes & Infections*, *1*(12), e46.

http://www.ncbi.nlm.nih.gov/. (n.d.). National Center for Biotechnology Information. Retrieved July 26, 2015, from http://www.ncbi.nlm.nih.gov/

Jibril, A. H., Umoh, J. U., Kabir, J., Saidu, L., Magaji, A. A., Bello, M. B., & Raji, A. A. (2014). Newcastle Disease in Local Chickens of Live Bird Markets and Households in Zamfara State, Nigeria. *International Scholarly Research Notices*, *2014*, e513961.

Justus, O., Owuor, G., & Bebe, B. O. (2013). Management practices and challenges in smallholder indigenous chicken production in Western Kenya. *Journal of Agriculture and Rural Development in the Tropics and Subtropics (JARTS)*, *114*(1), 51–58.

Katoh, K., & Standley, D. M. (2013). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution*, *30*(4), 772–780.

Kim, Y., & Nielsen, R. (2004). Linkage Disequilibrium as a Signature of Selective Sweeps. *Genetics*, *167*(3), 1513–1524.

Kingori, A. M., Wachira, A. M., & Tuitoek, J. K. (2010). Indigenous Chicken Production in Kenya: A Review. *International Journal of Poultry Science*, *9*(4), 309–316.

Kranis, A., Gheyas, A. A., Boschiero, C., Turner, F., Yu, L., Smith, S., Burt, D. W. (2013). Development of a high density 600K SNP genotyping array for chicken. *BMC Genomics*, *14*(1), 59.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics (Oxford, England)*, *23*(21), 2947–2948.

Löytynoja, A., & Goldman, N. (2005). An algorithm for progressive multiple alignment of sequences with insertions. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(30), 10557–10562.

Magothe, T. M. (2012). Indigenous chicken production in Kenya: I. Current status. *World's Poultry Science Journal*, *68*, 119–132.

Marcos-Carcavilla, A., Mutikainen, M., González, C., Calvo, J. H., Kantanen, J., Sanz, A., Serrano, M. (2010). A SNP in the HSP90AA1 gene 5′ flanking region is associated with the adaptation to differential thermal conditions in the ovine species. *Cell Stress & Chaperones*, *15*(1), 67–81.

McDonald, J. H., & Kreitman, M. (1991). Adaptive protein evolution at the Adh locus in Drosophila. *Nature*, *351*(6328), 652–654.

Moraa, G. K., Oyier, P. A., Maina, S. G., Makanda, M., Ndiema, E. K., Alakonya, A. E., Ommeh, S. C. (2015). Assessment of phenotypic traits relevant for adaptation to hot environments in indigenous chickens from four agro-climatic zones of Kenya.

Mozzi, A., Pontremoli, C., Forni, D., Clerici, M., Pozzoli, U., Bresolin, N., Sironi, M. (2015). OASes and STING: Adaptive Evolution in Concert. *Genome Biology and Evolution*, *7*(4), 1016–1032.

Mwacharo, J. M., Nomura, K., Hanada, H., Han, J. L., Amano, T., & Hanotte, O. (2013). Reconstructing the origin and dispersal patterns of village chickens across East Africa: insights from autosomal markers. *Molecular Ecology*, *22*(10), 2683–2697.

Nekrutenko, A., Makova, K. D., & Li, W.-H. (2002). The KA/KS Ratio Test for Assessing the Protein-Coding Potential of Genomic Regions: An Empirical and Simulation Study. *Genome Research*, *12*(1), 198–202.

Nielsen, R. (2005). Molecular signatures of natural selection. *Annual Review of Genetics*, *39*, 197–218.

Ommeh S, J. L. N. (2010). Geographic and Breed Distribution Patterns of an A/G Polymorphism resent in the Mx Gene Suggests Balanced Selection in Village Chickens. *International Journal of Poultry Science*, *9*(1).

Ramadan, S. (2015). A study on genetic diversity of Egyptian native livestock.

Rothenburg, S., Seo, E. J., Gibbs, J. S., Dever, T. E., & Dittmar, K. (2009). Rapid evolution of protein kinase PKR alters sensitivity to viral inhibitors. *Nature Structural & Molecular Biology*, *16*(1), 63–70.

Sadler, A. J., & Williams, B. R. G. (2008). Interferon-inducible antiviral effectors. *Nature Reviews Immunology*, *8*(7), 559–568.

Salces-Ortiz, J., González, C., Martínez, M., Mayoral, T., Calvo, J. H., & Serrano, M. M. (2015). Looking for adaptive footprints in the HSP90AA1 ovine gene. *BMC Evolutionary Biology*, *15*(1), 7.

Tag-El-Din-Hassan, H. T., Sasaki, N., Moritoh, K., Torigoe, D., Maeda, A., & Agui, T. (2012). The chicken 2'-5' oligoadenylate synthetase A inhibits the replication of West Nile virus. *The Japanese Journal of Veterinary Research*, *60*(2-3), 95–103.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, *123*(3), 585–595.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution*, *30*(12), 2725–2729.

Vandegrift, K. J., Sokolow, S. H., Daszak, P., & Kilpatrick, A. M. (2010). Ecology of avian influenza viruses in a changing world. *Annals of the New York Academy of Sciences*, *1195*, 113–128.

Vinkler, M., Bainová, H., & Bryja, J. (2014). Protein evolution of Toll-like receptors 4, 5 and 7 within Galloanserae birds. *Genetics Selection Evolution*, *46*(1), 72.

Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., & Barton, G. J. (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics (Oxford, England)*, *25*(9), 1189–1191.

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-Statistics for the Analysis of Population Structure. *Evolution*, *38*(6), 1358–1370.

Werling, D., Jann, O. C., Offord, V., Glass, E. J., & Coffey, T. J. (2009). Variation matters: TLR structure and species-specific pathogen recognition. *Trends in Immunology*, *30*(3), 124–130.

Wolc, A., Arango, J., Jankowski, T., Settar, P., Fulton, J. E., O'Sullivan, N. P., Dekkers, J. C. M. (2013). Genome-wide association study for Marek's disease mortality in layer chickens. *Avian Diseases*, *57*(2 Suppl), 395–400.

Yang, Z., & Nielsen, R. (2008). Mutation-selection models of codon substitution and their use to estimate selective strengths on codon usage. *Molecular Biology and Evolution*, *25*(3), 568–579.

Yang, Z., Nielsen, R., Goldman, N., & Pedersen, A. M. (2000). Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics*, *155*(1), 431–449.