Toll-Like Receptor Evolution in Birds: Gene Duplication, Pseudogenization, and Diversifying Selection

Hana Velová,*[,1] Maria W. Gutowska-Ding,[2] David W. Burt,[3] and Michal Vinkler[1]

[1]Department of Zoology, Faculty of Science, Charles University, Prague, Czech Republic

[2]Department of Genomics and Genetics, The Roslin Institute and Royal (Dick) School of Veterinary Studies, The Roslin Institute Building, University of Edinburgh, Midlothian, United Kingdom

[3]Office of DVC (Research), University of Queensland, St. Lucia, QLD, Australia

*Corresponding author: E-mail: bainova@natur.cuni.cz.

Associate editor: Meredith Yeager

Abstract

Toll-like receptors (TLRs) are key sensor molecules in vertebrates triggering initial phases of immune responses to pathogens. The avian TLR family typically consists of ten receptors, each adapted to distinct ligands. To understand the complex evolutionary history of each avian TLR, we analyzed all members of the TLR family in the whole genome assemblies and target sequence data of 63 bird species covering all major avian clades. Our results indicate that gene duplication events most probably occurred in TLR1 before synapsids diversified from sauropsids. Unlike mammals, ssRNA-recognizing TLR7 has duplicated independently in several avian taxa, while flagellin-sensing TLR5 has pseudogenized multiple times in bird phylogeny. Our analysis revealed stronger positive, diversifying selection acting in TLR5 and the three-domain TLRs (TLR10 [TLR1A], TLR1 [TLR1B], TLR2A, TLR2B, TLR4) that face the extracellular space and bind complex ligands than in single-domain TLR15 and endosomal TLRs (TLR3, TLR7, TLR21). In total, 84 out of 306 positively selected sites were predicted to harbor substitutions dramatically changing the amino acid physicochemical properties. Furthermore, 105 positively selected sites were located in the known functionally relevant TLR regions. We found evidence for convergent evolution acting between birds and mammals at 54 of these sites. Our comparative study provides a comprehensive insight into the evolution of avian TLR genetic variability. Besides describing the history of avian TLR gene gain and gene loss, we also identified candidate positions in the receptors that have been likely shaped by direct molecular host–pathogen coevolutionary interactions and most probably play key functional roles in birds.

Key words: **adaptive evolution, amino acid physicochemical properties, convergence, pattern recognition receptors, positive selection, pseudogene.**

Article

Introduction

The species richness and wide range of ecological adaptations are remarkable in birds (Jetz et al. 2012). Despite being similar to mammals in many aspects of their biology, birds evolved many similar traits (including homoiothermy and parental care) convergently through their analogous ecological life strategies (Farmer 2000; Emery and Clayton 2004; Olkowicz et al. 2016; Lovegrove 2017). Recent advance in avian genomic research, which started with the Avian Phylogenomics Project (http://avian.genomics.cn/en/; Zhang, Jarvis, et al. 2014) and continues with the Bird 10,000 Genomes Project (https://b10k.genomics.cn/) has created new possibilities to use publicly available whole genome sequencing data to resolve many questions regarding avian biology, evolution, and adaptations (Zhang, Li, et al. 2014). Among these, the evolution of avian immune function is of particular interest. Being challenged by similar classes of pathogens as found in mammals, highly diversified immune genes in birds are rich in mixed signatures of symplesiomorphy, convergence, and apomorphic adaptations to novel functions (Burri et al. 2010; Cheng et al. 2015).

The coevolution between host and pathogen likely involves mainly molecules that form direct interface between the host and the pathogen structures, for example, the pattern recognition receptors (PRRs) and their ligands, the pathogen-associated molecular patterns (PAMPs; Janeway and Medzhitov 2002). Forming the first line of the host immune defence, PRRs appear to constantly evolve toward specific and appropriate recognition of certain PAMPs (Wang, Zhang, Liu, et al. 2016). The Toll-like receptor (TLR) protein family belongs to one of the most essential and functionally most characterized PRRs (Palm and Medzhitov 2009; Coscia et al. 2011). After specific PAMP binding, TLRs trigger signaling pathways activating transcription factors, such as NF-κB to induce expression of target genes that are key to triggering an inflammatory immune response and subsequent activation of acquired immunity (Iwasaki and Medzhitov 2015).

TLRs are transmembrane proteins each with the characteristic horse-shoe-shaped ectodomain (ECD), where the direct contact between the receptor surface and specific microbe molecules occurs, the transmembrane domain and an intracellular Toll–interleukin 1 receptor (TIR) domain that

enables the downstream signal transmission (Kawai and Akira 2010). The capacity to bind a particular type of ligand in TLRs is dependent on the structure of the ECD. This has been used to divide the vertebrate TLRs into three different groups based on the pattern of hydrogen bonds formed among the asparagine residues in the neighboring leucine-rich repeats, which stabilize the shape of the ECD (Wang, Zhang, Liu, et al. 2016). The ten avian TLR family members belong either to the single-domain TLRs possessing a complete asparagine ladder (TLR3, TLR5, TLR7, TLR15, and TLR21) or to the three-domain TLRs with the ladder interrupted in the central part of the ECD (TLR1A, TLR1B, TLR2A, TLR2B, and TLR4; Wang, Zhang, Liu, et al. 2016). While some avian TLRs conservatively recognize the same ligands as mammalian TLRs (such as TLR4 detecting bacterial lipopolysaccharides, LPS; TLR5 flagellin; TLR3 dsRNA or TLR7 ssRNA; reviewed by Keestra et al. 2013), others were reported to form distinct paralogues with related ligand specificity (e.g., heterodimer-forming TLR1A/TLR1B together with TLR2A/TLR2B, which recognize di/triacylated lipopeptides; Keestra et al. 2007; Higuchi et al. 2008) or achieve recognition of similar ligands as their mammalian analogues through convergence (e.g., avian TLR21 that recognizes CpG DNA similarly as mammalian TLR9; Brownlie et al. 2009; Keestra et al. 2010). Finally, TLR15 that is unique to birds evolved to gain a novel function in recognition of extracellular proteases (de Zoete et al. 2011).

Gene duplication is an important mechanism allowing genes to evolve novel functions (Zhang 2003; Ellegren 2008). Gene duplications are common in TLRs, as with other immune genes. Duplicated TLR1 family members are known in mammals (TLR1, TLR6, TLR10, and TLR2) as well as in birds (TLR1A, TLR1B, TLR2A, and TLR2B; Coscia et al. 2011; Huang et al. 2011; Wang, Zhang, Liu, et al. 2016). Recently, duplication of *TLR7* has been described in several avian taxa (Cormican et al. 2009; Grueber et al. 2012; Raven et al. 2017), while other birds possibly lack the duplicated *TLR7*. This intriguing finding contrasts with that found in mammals, which contain two other TLR7-like family members, *TLR8* and *TLR9*, both missing in birds (Philbin et al. 2005; Temperley et al. 2008), suggesting that gene loss may also play a significant role in avian TLR evolution. *TLR5* pseudogenization occurred independently several times within the passerines (Bainova et al. 2014) and possibly in parrots (Alcaide and Edwards 2011). The general pattern of *TLR* pseudogenization and gene duplication is largely unknown and may be far more common across various *TLR* genes and avian taxa.

Although generally conserved in their structure, TLRs were reported to exhibit very high levels of inter- and intraspecific genetic variation in birds (Alcaide and Edwards 2011; Huang et al. 2011; Grueber et al. 2014). Several authors have documented the functional significance of *TLR* genetic variation (Leveque et al. 2003; Walsh et al. 2008), characterized associations between TLR variation and disease susceptibility (Netea et al. 2012; Medvedev 2013) and suggested the maintenance of intraspecific polymorphism through balancing selection mediated by pathogens (Ferrer-Admetlla et al. 2008). Despite several attempts to understand the evolutionary significance of *TLR* genetic variation in birds (Alcaide and Edwards 2011; Grueber et al. 2014; Vinkler et al. 2014), the question of its adaptive value for avian specific PAMP recognition remains unresolved. Here, we use various approaches to analyze natural selection in the most comprehensive attempt to investigate adaptive pathogen-driven evolution in avian TLRs.

Facilitated by the conservative structures of TLR molecules, they present a suitable model for studying pathogen-driven microevolutionary processes at the DNA level. The direct contact between TLRs and PAMPs is restricted to certain amino acid sites (Gay and Gangloff 2007) where even single amino acid substitutions may have profound effects on receptor binding properties (Keestra et al. 2008; Walsh et al. 2008; Resman et al. 2009; Meng et al. 2010). Thus, positively selected sites can be predicted to emerge from the generally negatively selected background and can be used for prediction of functionally relevant positions in nonmodel animals lacking precise protein crystallographic data. Several studies have investigated the nature of positively selected sites (PSS) at the interspecific level in various TLRs in selected bird taxa (Alcaide and Edwards 2011; Huang et al. 2011; Grueber et al. 2014; Vinkler et al. 2014; Wang, Zhang, Liu, et al. 2016). Yet, colocation study of PSS on the receptor surface with the predicted functional residues already known in the model species (human and mouse), has never been done across the avian phylogeny (but see Vinkler et al. 2014, for a similar approach in Galloanserae birds).

Given the current lack of comprehensive evidence on TLR family molecular evolution across the avian phylogeny, our comparative study reported here performed evolutionary analysis on whole-genome sequences for 48 species representing 34 avian orders (supplementary material S2: table S23, Supplementary Material online) mainly gained through the Avian Phylogenomics Project (Zhang, Jarvis, et al. 2014). The TLR sequences extracted from these genomes together with those added from other public resources (in total 63 species) allowed us to infer the history of avian *TLR* gene gain and gene loss in the context of TLR evolution in other vertebrates. By critical assessment of positive selection acting on all members of the TLR family in birds, we describe the adaptive microevolutionary changes in these immune receptors in their molecular context. Thus, this study is the first to comprehensively predict functionally relevant genetic variation in avian TLRs providing insights into the coadaptation between host and pathogen through evolution of ligand recognition.

Results and Discussion

Evolution of *TLR1* Family Gene Duplications

The *TLR1* gene family is widely duplicated in vertebrates including birds (*TLR1A*, *TLR1B*, *TLR2A*, and *TLR2B*; Temperley et al. 2008; Cormican et al. 2009; Huang et al. 2011; Wang, Zhang, Liu, et al. 2016). Presently, however, the timing of these gene duplication events is unclear: the paralogues within the *TLR1* and *TLR2* subfamily were suggested to have duplicated before the mammalian divergence from sauropsids (Huang et al. 2011) or after this divergence, independently in birds and mammals (Temperley et al. 2008; Cormican et al. 2009;

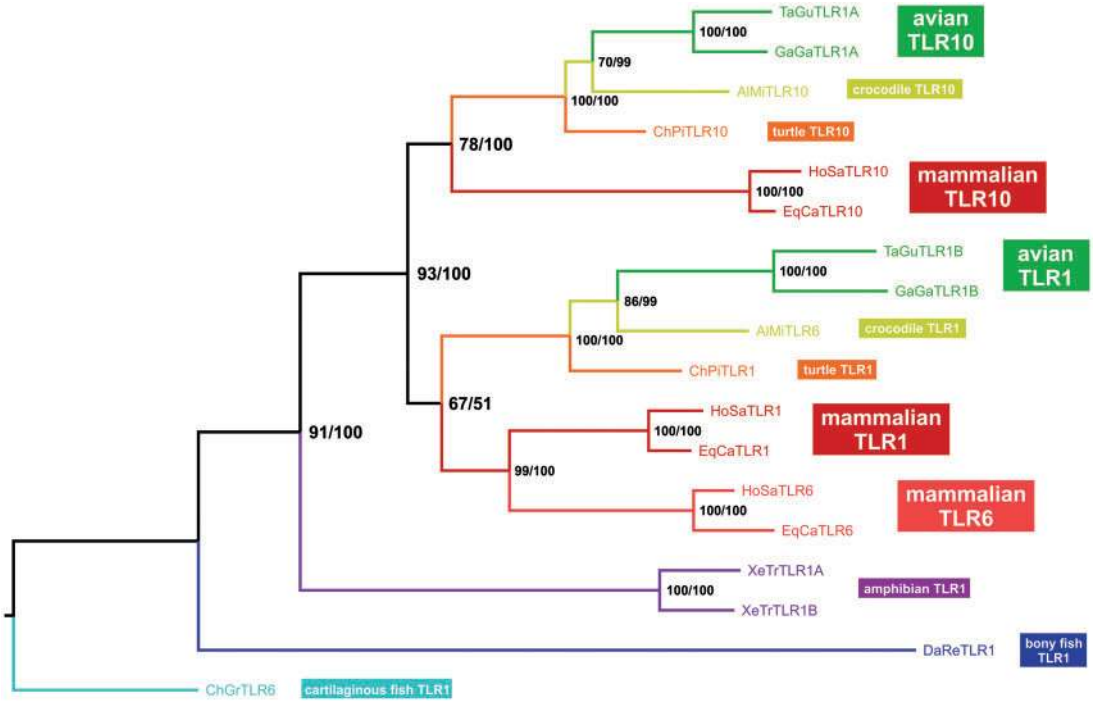

Fig. 1. Phylogenetic tree based on nonconverted regions of *TLR1* subfamily members. The bootstrap values of maximum likelihood analysis obtained using PhyML and the posterior probability of Bayesian analysis obtained using MrBayes (in percentage per each node) are provided. Birds are represented by zebra finch (TaGu, *Taeniopygia guttata*) and chicken (GaGa, *Gallus gallus*), crocodiles by alligator (AlMi, *Alligator mississippiensis*), turtles by painted turtle (ChPi, *Chrysemys picta*), mammals by human (HoSa, *Homo sapiens*) and horse (EqCa, *Equus caballus*), amphibians by clawed frog (XeTr, *Xenopus tropicalis*), bony fish by zebrafish (DaRe, *Danio rerio*) and cartilaginous fish by shark (ChGr, *Chiloscyllium griseum*). The analysis was performed using a single amino acid sequence per TLR and species. Based on the results, we suggest renaming TLR1A to TLR10 and TLR1B to TLR1 in birds.

Mikami et al. 2012; Wang, Zhang, Liu, et al. 2016). This disagreement results from the lack of accounting for gene conversion as a mechanism of paralogue sequence homogenization in some studies (Temperley et al. 2008; Cormican et al. 2009; Wang, Zhang, Liu, et al. 2016). We found significant support for gene conversion between the two paralogues both in the *TLR1* subfamily (*TLR1A* and *TLR1B*) and *TLR2* subfamily (*TLR2A* and *TLR2B*; supplementary material S1: table S1, Supplementary Material online; for detailed info see supplementary material S3, Supplementary Material online) in birds. Therefore, we constructed both *TLR1* and *TLR2* phylogenetic trees based on the protein sequences of the nonconverted regions only (fig. 1; supplementary material S1: figs. S1 and S2, Supplementary Material online).

Our results based on the analysis of 179 sequences of 83 vertebrate species (for the list see supplementary material S2: table S24, Supplementary Material online) show that in the *TLR1* family the sequences of avian species cluster together based on paralogue identity (supplementary material S1: fig. S1, Supplementary Material online). Avian *TLR1A* and reptilian *TLR10* cluster together with mammalian *TLR10*, while avian *TLR1B* and reptilian *TLR1*/*TLR6* (inconsistent nomenclature in reptilian *TLR1*) cluster together with mammalian *TLR1* and *TLR6* paralogues (shown in fig. 1; detailed in supplementary material S1: fig. S1, Supplementary Material online). This is consistent with the results previously published by Huang et al. (2011) confirming that the first duplication event within the *TLR1* subfamily occurred before mammal–reptile divergence. On the basis of these findings, we suggest renaming avian *TLR1A* to *TLR10* (from now on marked as *TLR10* [*TLR1A*] in the text), while modifying the name of the avian *TLR1B* gene to *TLR1* (marked as *TLR1* [*TLR1B*]). Our analysis also indicates that the duplication of *TLR1* in amphibians (namely in *Xenopus*) apparently arose independently of the *TLR1* duplication in amniotes (supplementary material S1: fig. S1, Supplementary Material online).

In addition, in the *TLR2* subfamily there are two copies present in birds (*TLR2A* and *TLR2B*). In mammalian genomes, on the other hand, only one functional *TLR2* is maintained and for some species *TLR2*-like pseudogenes (a second copy) were described (Roach et al. 2005). The duplication of *TLR2* in sauropsids could have arisen independently of mammals, or alternatively, a duplication event predating divergence between birds and mammals could have occurred (similar to the *TLR1* subfamily), followed with pseudogenization leaving only a single *TLR2* copy in mammals (Huang et al. 2011). Our analysis of 137 homologous sequences in 78 vertebrate species (for the list see supplementary material S2: table S24, Supplementary Material online) contradicts the second scenario (supplementary material S1: fig. S2, Supplementary Material online) and supports the independent duplication of *TLR2* in sauropsids. This is also supported by the phylogenetic analysis of a ca. 80 amino-acid-long alignment of the conserved sequence in human *TLR2*/*TLR2*-like pseudogene

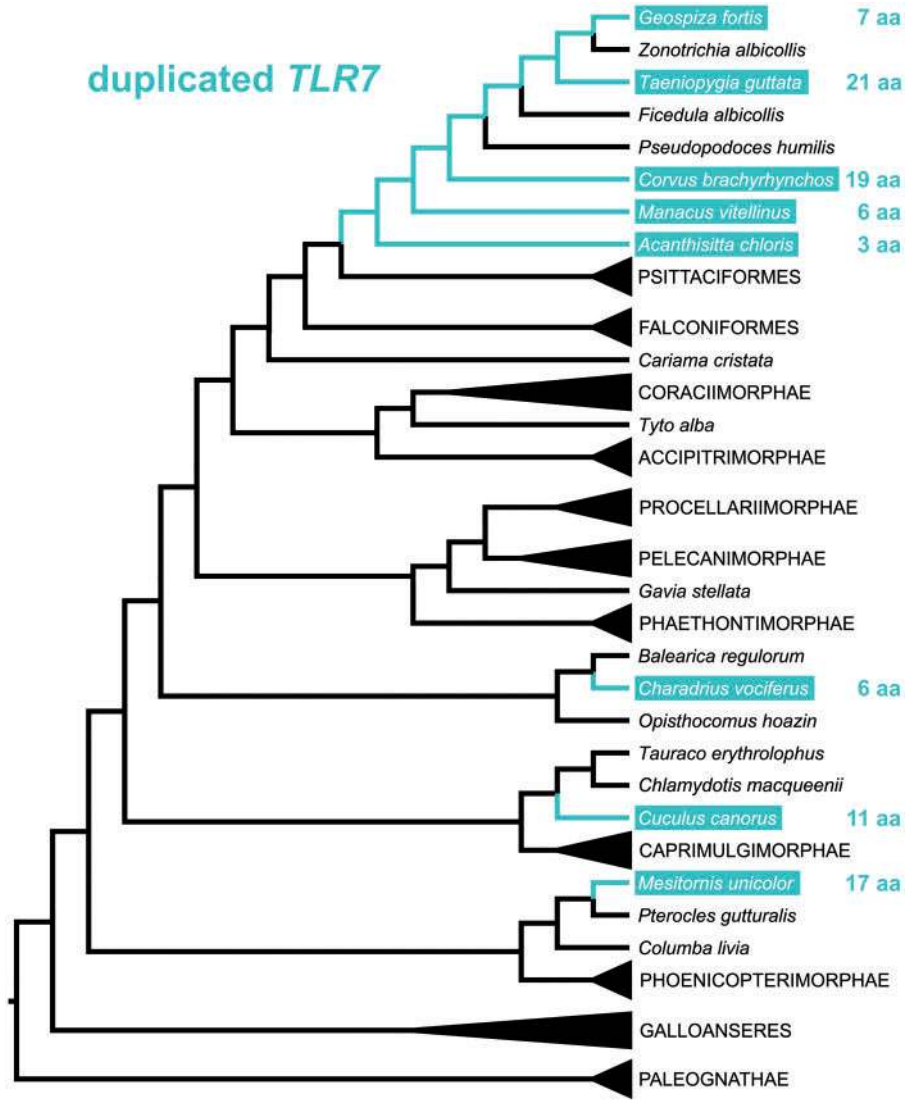

Fig. 2. Avian *TLR7* duplication. In the schematic avian phylogenetic tree the species with duplicated *TLR7* are highlighted in teal rectangles. The numbers of amino acid substitutions distinguishing the two copies of the duplicated *TLR7* are shown behind the species name. The analysis was performed using a single sequence per gene and species.

and chicken TLR2A/TLR2B (supplementary material S1: fig. S3, Supplementary Material online); the resulting phylogenetic tree suggests that *TLR2* duplication in birds was independent of a parallel mammalian *TLR2* duplication that was followed by a pseudogenization of the second *TLR2* copy (supplementary material S1: fig. S4, Supplementary Material online). However, it is difficult to draw any final conclusions on the validity of the two scenarios proposed for *TLR2* evolution in amniotes since the support for the branching of the avian *TLR2* paralogues is limited (bootstrap support for the mammal-sauropsid split = 47, supplementary material S1: fig. S2, Supplementary Material online).

Avian *TLR7* Gene Duplications

Birds lack two of the three vertebrate *TLR7* family members (i.e., *TLR8* and *TLR9*; Philbin et al. 2005; Temperley et al. 2008). Instead, the avian *TLR7* locus appears to have recently duplicated in some passerines (Cormican et al. 2009; Grueber et al. 2012) and waders (Raven et al. 2017). Within the genome sequences investigated in this study, we also found two *TLR7* copies in other avian taxa: Cuculiformes and Mesiornithiformes (fig. 2). The number of amino acid substitutions distinguishing the two *TLR7* loci ranges between species from 21 in zebra finch (*Taeniopygia guttata*) to only 3 in rifleman (*Acanthisitta chloris*; see fig. 2; supplementary material S1: table S2 and supplementary material S4, Supplementary Material online). Furthermore, the two *TLR7* copies likely differ slightly in their functions, because 11 of the sites identified as distinct between the two copies in any avian species match the known ligand-binding positions (Wei et al. 2009; Gupta et al. 2016; Zhang et al. 2016), five are situated in the Z-loop region responsible for ligand binding and dimerization and the other 10 variable sites are identical with PSS detected in birds (supplementary material S1: table S3, Supplementary Material online). To bring independent support to the evidence of *TLR7* gene duplication, we

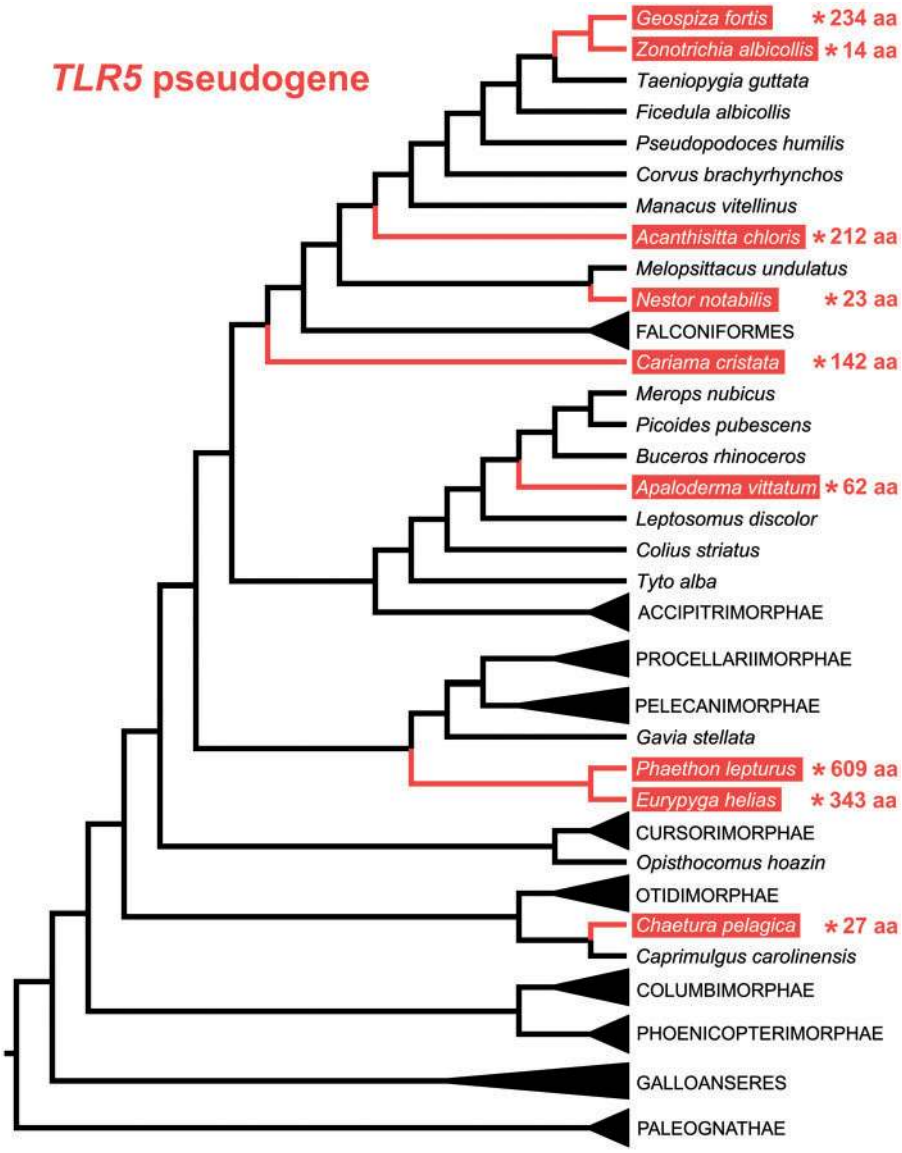

Fig. 3. *TLR5* pseudogenization in birds. Species possessing only *TLR5* pseudogene are highlighted in red rectangles within the schematic representation of avian phylogenetic tree. The position of the first stop codon is indicated by the number provided behind the species name (position numbering according to the chicken reference). The analysis was performed using a single sequence per gene and species.

performed a PCR-based quantitative copy number variation (CNV) analysis comparing *TLR7* with two single-copy *TLRs* (*TLR3* and *TLR4*) in selected species (collared flycatcher, *Ficedula albicollis*; ground tit, *Pseudopodoces humilis*; white-throated sparrow, *Zonotrichia albicollis* and zebra finch). Our results confirm that the zebra finch (a species with two *TLR7* loci based on the genome sequence analysis) possesses two *TLR7* copies (supplementary material S1: tables S4–S6 and fig. S5, Supplementary Material online), while the other investigated species (with a predicted single *TLR7* locus) possess only one *TLR7* gene copy. In contrast to the *TLR1* family, we did not detect any gene conversion between the two copies of the duplicated *TLR7* genes (see supplementary material S3, Supplementary Material online). Despite this, the constructed tree of avian TLR7 does not show any evidence of two separate TLR7 phylogenetic clusters, but surprisingly both copies of *TLR7* always clustered separately for each species (supplementary material S1: fig. S6 and table S7, Supplementary Material online). Therefore, we infer that the duplication in *TLR7* occurred several times independently in recent avian history and possibly even multiple times in some cases (as, e.g., in ruddy turnstone, *Arenaria interpres*, where *TLR7* was reported to be triplicated; Raven et al. 2017).

TLR5 Pseudogenization in Birds

Recently, it has been shown that there is no functional gene for TLR5 in some passerine species (Bainova et al. 2014). In this study, we have also found several avian taxa possessing only nonfunctional *TLR5* pseudogenes with stop-codons in their sequences. Apart from passerines, we revealed *TLR5* pseudogenes in other clades: that is, Psittaciformes, Cariamiformes, Trogoniformes, Phaethontiformes, Eurypygiformes, and Apodiformes (fig. 3). The sequence data show that the stop-codon positions differ among the

investigated species with differences also in the pseudogenization mechanism (single nucleotide substitution or frameshifting indels, see supplementary material S1: table S8, Supplementary Material online). The stop-codons in *TLR5*, thus, arose independently in the evolutionary history of the different avian taxa. This, however, does not appear to be the result of relaxed selection acting in *TLR5*. Firstly, the number of single nucleotide variants (SNVs) in the alignment of the *TLR5* functional sequences is comparable to other *TLRs* (indicating comparable mutation rate; supplementary material S1: table S9, Supplementary Material online). Secondly, the total number of PSS is higher than in other TLRs and ω (dN/dS ratio) is again similar to other TLRs (see below and supplementary material S1: table S10, Supplementary Material online) and higher than the genome average (Zhang, Li, et al. 2014). However, since most of the PSS contain only conservative substitutions without dramatic effects on the amino acid site physicochemical properties, it appears that negative selection may prevent the loss of function in those species where the *TLR5* functional gene is maintained. Altogether this suggests that there may be a selection for TLR5 loss of function in certain avian evolutionary lineages. The loss of a *TLR5* functional gene is not limited only to birds. The *TLR5* pseudogene has also been described in one human allele, where its presence is associated with an increased risk of pneumonia infection caused by flagellated bacteria *Legionella pneumophila* (Hawn et al. 2003; Zhang et al. 2013), while at the same time possessing a *TLR5* nonfunctional allele might be advantageous for decreasing the probability of autoimmune disease development (Hawn et al. 2005). *TLR5* pseudogenization may be allowed by the high redundancy of pathogen detection, where apart from TLR5, flagellin is also recognized by other PRRs, for example, NLRC4 inflammasome (Zhao et al. 2011; Yang et al. 2014). In birds other flagellin-recognizing PRRs have not yet sufficiently been studied to support this hypothesis.

Diversifying Selection in Avian TLRs

We tested pervasive, positive, and diversifying selection acting in all avian TLRs. Numbers of PSS differ between *TLRs* (from 0.3% to 5.8% PSS per TLR; table 1; supplementary material S1: fig. S7 and supplementary material S5: table S26, Supplementary Material online). The results suggest that positive selection is acting more on TLRs exposed toward the cell surface (mainly TLR1 [TLR1B], TLR2A, TLR5, TLR4, and TLR2B), than in the TLRs situated in endosomes (TLR21, TLR3, and TLR7). This may be because the endosomal TLRs are specialized for detection of less complex ligands, which show low structural variation (such as ssRNA in TLR7, dsRNA in TLR3 or CpG DNA regions in TLR21; Brownlie and Allan 2011) interacting with the genetic variability in TLRs. This supports the previous findings by Mikami et al. (2012) in vertebrates. The only exception is TLR15, which is unique to birds and reptiles (Boyd et al. 2012). Although situated on the cell surface, TLR15 is activated (unlike other TLRs) by the ECD proteolytic cleavage with pathogen-derived proteases (de Zoete et al. 2011). Since *TLR15* harbors only low numbers of PSS (table 1), we hypothesize that this pathogen-recognition mechanism may be linked to reduced positive selection acting on the ECD. Nonetheless, the cleavage site is variable across bird taxa and contains two PSS (supplementary material S1: table S11, Supplementary Material online; see also supplementary material S1: text S1 and fig. S8, Supplementary Material online).

Table 1. The Number of Positively Selected Sites in Avian TLRs.

	Species[a]	aa length[b]	PSS[c]	PSS/TLR[d] (%)
TLR10 [TLR1A]	45	794	31	3.9
TLR1 [TLR1B]	44	638	37	5.8
TLR2A	40	793	43	5.4
TLR2B	42	781	35	4.5
TLR3	51	895	19	2.1
TLR4	54	843	38	4.5
TLR5	46	861	45	5.2
TLR7	51	1041	31	3.0
TLR15	53	868	24	2.8
TLR21	14	907	3	0.3

[a]Number of species (one sequence per species).
[b]The protein sequence length in reference chicken TLRs (NCBI accession numbers are listed in supplementary material S1: table S14, Supplementary Material online).
[c]Number of positively selected sites detected in investigated species per each gene.
[d]The percentage of PSS per whole receptor amino acid sequence.

Our results are also mostly consistent with the findings of Wang, Zhang, Liu, et al. (2016), who suggested grouping of TLRs based on their ECD architecture and showed that in mammals the single-domain TLRs (TLR3, TLR5, TLR7, TLR15, and TLR21) are under stronger purifying selection than the three-domain TLRs (TLR10 [TLR1A], TLR1 [TLR1B], TLR2A, TLR2B, and TLR4). Here, we found that also in birds the positive selection is acting more on the three-domain TLRs, with the exception of TLR5, which is under strong positive selection in birds (table 1 and supplementary material S1: fig. S7, Supplementary Material online). Since some avian species have lost a functional TLR5 through pseudogenization (Bainova et al. 2014; and this study, fig. 3), we hypothesize that TLR5 has a specific role in avian immunity when compared, for example, to mammals. The variation in selection acting at TLR5 in different taxa might, for example, reflect the differences in selection against overactivation of gut immunity with flagellated symbiotic microbiota (Iqbal et al. 2005).

Being responsible for direct and specific recognition of structurally heterogeneous PAMPs (Reddick and Alto 2014), the pathogen-mediated selective pressures are particularly diversified in the TLR ECDs. As previously shown in general for vertebrates (Mikami et al. 2012), we also find the majority of TLR PSS in birds to be situated in ligand-binding ECDs (supplementary material S1: fig. S7, Supplementary Material online; the exact positions of PSS are visualized in supplementary material S1: fig. S9, Supplementary Material online and listed in supplementary material S6: table S27, Supplementary Material online). We may predict stronger functional effects of PSS with nonconservative amino acid substitutions that change the physicochemical properties of the particular residues. The highest numbers of these nonconservative PSS are in TLR4 and TLR2A (fig. 4; see also supplementary material S1: table S12 and supplementary materials S7 and S8, Supplementary Material online). Especially in TLR4, most

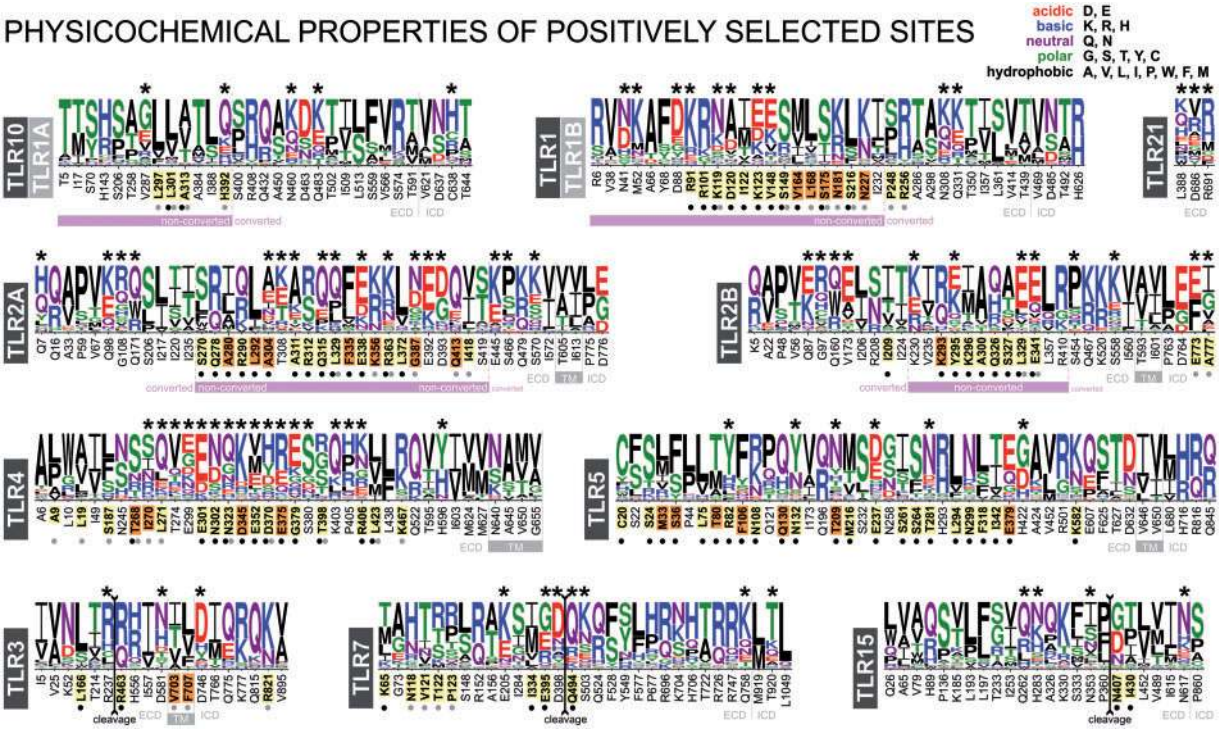

Fig. 4. Physicochemical properties of the positively selected sites (PSS). All PSS are shown in all avian TLRs—amino acid substitutions are colored according to their physicochemical properties: acidic in red, basic in blue, neutral in purple, polar in green, and hydrophobic in black. The size of an letter corresponds to the procentual proportion of that particular amino acid within the sequence alignment. The numbering is adopted from reference chicken TLRs (for NCBI IDs see supplementary material S1: table S14, Supplementary Material online). PSS which correspond to functionaly important residues (black dot—ligand binding; gray dot—dimerization) are highlighted in bold and orange (identical site) or yellow (topological proximity closer than 5 Å from a functionaly important residue). Ectodomain (ECD), intracellular domain (ICD), and transmembrane (TM) region are visualized; nonconverted region in TLR1/2 is highlighted in pink; cleavage site in TLR3, TLR7, and TLR15 is indicated by a black line tipped with arrows (supplementary material S1: table S11 and text S1, Supplementary Material online). Nonconservative PSS are marked by stars (supplementary material S1: table S12, Supplementary Material online).

PSS with dramatic changes in charge and hydrophobicity fall into the known functionally important sites of the receptor, with high probability of changing the TLR4 ligand-binding properties. On the other hand, most substitutions in PSS in TLR15, TLR3, and TLR10 [TLR1A] do not significantly change the amino acid physicochemical properties. Despite avian TLR1 [TLR1B] being significantly shorter than TLR10 [TLR1A], we found more PSS in TLR1 [TLR1B] and these were less conservative in their charge than in TLR10 [TLR1A], especially in the predicted ligand-binding region that is spanning the region avoiding gene conversion (supplementary material S1: text S2 and supplementary material S7, Supplementary Material online). This, together with high TLR2A PSS variation, can be possibly explained by distinct binding capacities of the heterodimers formed by the duplicated TLR1/TLR2 subfamily members, where the combination TLR1 [TLR1B]/TLR2A (unlike any other combination) is able to recognize peptidoglycans and efficiently recognizes diacylated bacterial lipoproteins (BLP; Higuchi et al. 2008). Surprisingly, despite its high number of PSS, TLR5 is the most PSS-conservative TLR across all avian taxa (supplementary material S1: table S12, Supplementary Material online). Given that flagellin is the only known TLR5 ligand (Hayashi et al. 2001), this may indicate functional constraints in those species with preserved functional TLR5, limiting any adaptations to larger numbers of relatively minor changes. The PSS identified in this study are consistent with PSS identified by other studies and other taxa especially in TLR15, TLR4, and TLR5 (PSS listed in supplementary material S6: table S27, Supplementary Material online and those located in ECD visualized in fig. 5; visualization of ECDs along with intracellular domains, ICDs, and transmembrane, TM, regions is given in supplementary material S9, Supplementary Material online). Interestingly, from the five amino acid substitutions related to *Salmonella enterica* resistance variation in chickens (Leveque et al. 2003) only position E301 is also positively selected within birds in general, where negatively charged E/D are substituted with positively charged K/R (fig. 4). The relevance of this position was also supported in other studies of birds (Grueber et al. 2014) and mammals (Vinkler et al. 2009; Wlasiuk and Nachman 2010) and partially with the results of numerous human studies (Arbour et al. 2000; Zareparsi et al. 2005; Shen et al. 2010; Cario 2013; Rupasree et al. 2015; but see also Ohto, Yamakawa, et al. 2012) showing potential importance of variation at a neighboring position D299G (D303 in chicken sequence).

We found the highest agreement between the PSS and known functional site distribution in TLR5, TLR4, TLR2A, and TLR1 [TLR1B] (23, 18, 18, and 16 positions, respectively; for details see supplementary material S1: table S13 and fig. S10,

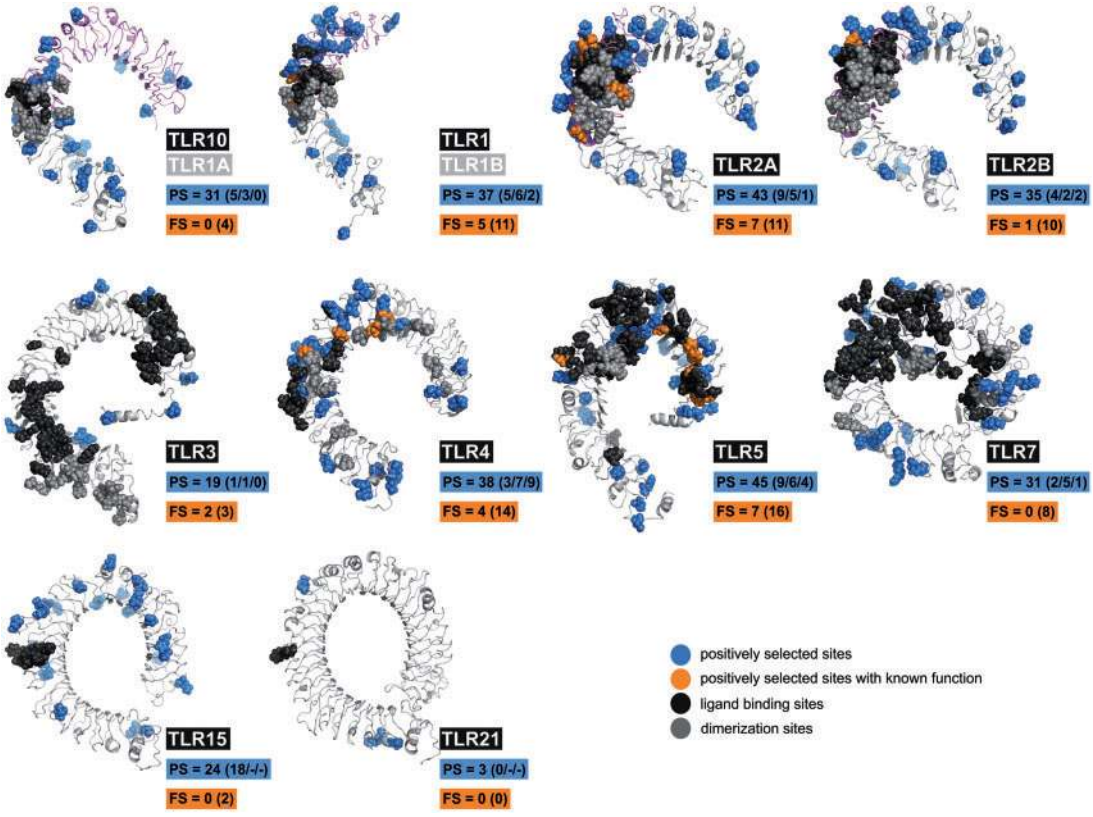

Fig. 5. Positively selected sites and functionaly important sites visualized on 3D extracellular domain structures of avian TLRs. PSS detected in this study are shown in blue. By orange coloration are highlighted the PSS identified at sites with previously described function. Other previously reported functionaly important residues are highlighted in black (ligand binding residues) or in gray (dimerization residues). The total numbers of PSS for each TLR are shown in blue rectangles, where in parentheses are numbers of PSS detected also in other avian/mammalian/both studies (for references see supplementary material S6: table S27, Supplementary Material online). The numbers of PSS with previously described function are provided in orange rectangles, where the number of PSS in topological proximity <5 Å is shown in parentheses. For more detailed information (including ICD and TM), see supplementary material S9, Supplementary Material online.

Supplementary Material online). We show that the PSS frequently evolve toward an amino acid of similar physicochemical properties being gained in distantly related avian taxa, while distinct properties can be found in closely related taxa (supplementary material S1: text S2 and supplementary material S8, Supplementary Material online). This convergent evolution may result from analogous selective pressures induced by partially similar microbial communities in different avian taxa (Waite and Taylor 2014). For example, there are several PSS in avian TLR4 (D345, E375, and G379) that are good candidates for functionally important sites evolving under convergence either toward negative or toward positive charge (supplementary material S8, Supplementary Material online). Being either lipid IVa-binding sites (Meng et al. 2010; Ohto, Fukase, et al. 2012; Scior et al. 2013) or LPS-binding sites (Park et al. 2009; Garate and Oostenbrink 2013; Paramo et al. 2013; Scior et al. 2013), these positions have been identified as positively selected also in several other studies in birds (Grueber et al. 2014) and mammals (Nakajima et al. 2008; Vinkler et al. 2009; Wlasiuk and Nachman 2010; Areal et al. 2011; Shen et al. 2012; Fornůsková et al. 2013).

The lack of agreement between the identified PSS and previously described TLR functional sites in other TLRs may be caused either by the lack of functional studies in some TLRs (e.g., TLR15 and TLR21) or by interspecific variation in TLR-ligand binding (e.g., TLR1, TLR3, and TLR7, where the structural and experimental studies were performed in mammals and not in birds; see supplementary material S1: text S2, Supplementary Material online). In both cases the analysis of PSS may indicate the target sites that are candidates for a functional role and may be the focus for further research. Despite the low number of PSS among the previously reported functional sites of TLR7, which may result from the constraints of conservative ligand binding, two nonconservative PSS located on the TLR7 Z-loop at the proximity of the TLR7 cleavage site (supplementary material S1: text S2, Supplementary Material online; Ewald et al. 2008; Kanno et al. 2013) might affect formation of secondary ligand-binding sites (Zhang et al. 2016).

Although most of the PSS functional sites are located in the ECD and involved in direct ligand binding, in TLR3 we found two PSS that are lying in the transmembrane region and are located in close proximity to the known receptor dimerization residues (fig. 4). Since TLR3 dimerization depends on the dsRNA length (Pirher et al. 2008) and the number of TLR3 molecules involved in the interaction (TLR3 dimer or trimer; Mineev et al. 2014), selection on these residues may play a crucial role in ligand recognition in TLR3. In

general, however, most PSS in avian TLR3 are very conservative (fig. 4; for details see also supplementary material S1: table S12 and supplementary material S7, Supplementary Material online), which is consistent with the results reported by Wang, Zhang, Liu, et al. (2016) who identified the TLR3 family as the most conservative TLR family within vertebrates. The avian PSS which can be suggested as candidates for relevant functional effect based on the nonconservativeness of the changes in their physicochemical properties are described in supplementary material S1: text S2, Supplementary Material online.

Conclusions

To our knowledge, this is the first study investigating adaptive evolution in all members of the avian TLR protein family in high number of species and most avian orders. Although we were unable to extract data to all *TLR* genes from all avian whole-genome assemblies investigated, our study covers 87% of the genes in these species, providing the most comprehensive analysis published so far. Our results provide important novel insights into the history of TLR family evolution including the processes of gene duplication, pseudogenization and diversifying selection. Most importantly, we show that the main duplication event for *TLR1* arose before avian and mammalian lineages split into separate clades (giving rise to *TLR1* [avian *TLR1B*] and *TLR10* [avian *TLR1A*]), while avian gene duplication of *TLR2* to *TLR2A* and *TLR2B* may have occurred in parallel to *TLR2* duplication in some mammals that resulted in pseudogenization of the second copy of *TLR2*. The latter conclusion however remains uncertain due to the power limitation imposed by the short sequences (not affected by gene conversion) aligned in our analysis. Furthermore, we confirmed the unique avian gene duplication in *TLR7* based on qPCR copy-number variation analysis. Surprisingly, this recent *TLR7* gene duplication in birds emerged independently in many species and representing several clades (Passeriformes, Charadriiformes, Cuculiformes, and Mesiornithiformes). Similarly, the *TLR5* pseudogenization, previously reported in several passerines and parrots, is seemingly more widespread in birds. Independent loss of functional TLR5 occurred also in seriemas (Cariamiformes), trogons (Trogoniformes), tropicbirds (Phaethontiformes), sunbitterns (Eurypygiformes), and swifts (Apodiformes).

The results of our analysis of positive selection in avian TLRs allowed us to predict functionally important, interspecifically variable positions. Consistent with some previous findings, these sites were located mainly at the ligand binding extracellular domain, mostly in TLRs where the extracellular domain is exposed to the cell surface and binds structurally diverse ligands (mainly in TLR1 [TLR1B], TLR2A, TLR2B, TLR4, and TLR5). The level of positive selection was low in endosomal TLRs which bind structurally more conservative oligonucleotide ligands (TLR3, TLR7, and TLR21). Positive selection also appears to act more on three-domain than on single-domain TLRs (except for TLR5 that is under very strong positive selection in birds). We also compared the PSS detected in this study with those previously described on an interspecific level in birds and mammals and with all sites previously identified as functionally relevant in TLRs to show strongest evidence for functional effect of the predicted PSS in TLR1 [TLR1B], TLR2A, TLR4, and TLR5. In agreement with the hypothesis that the variation in the TLR ligands is a driving force for selection in TLRs, nonconservative substitutions with the potential to dramatically change surface physicochemical properties were found mainly in two TLRs (TLR2A and TLR4) with highly variable ligands known in mammals. Altogether, our results indicate strong positive selection driving TLR evolution in birds. To better understand the significance of this adaptive evolution in avian *TLR* genes, further research would benefit from both in silico structural modeling allowing closer prediction of the positively selected physicochemical changes and functional testing of the variant effects in model in vitro systems.

Materials and Methods

Data Set

To create our data set, we used sequence data publicly available for up to 63 avian species per gene representing all orders of Neoaves (supplementary material S2: table S23, Supplementary Material online). Most of the *TLR* sequences were extracted from the whole-genome data generated by the Avian Phylogenomics Consortium (http://avian.genomics.cn/en/; Zhang, Jarvis, et al. 2014), genomes included in the B10K Project (Avianbase; http://b10k.genomics.cn/; Eöry et al. 2015). The nucleotide sequences of all avian *TLRs* were obtained by blasting (BLAST v 2.2.25+, NCBI, Zhang et al. 2000; blastn settings: E = 0.1, hits with the highest score and lowest E were accepted) the reference *TLR* sequences of chicken and zebra finch (for sequence ID number see supplementary material S1: table S14, Supplementary Material online) against the CDS database of all avian whole-genome sequences of the species included in the first B10K study (Zhang, Jarvis, et al. 2014). In cases where the blast search in the CDS databases was not successful, we ran blat search against the genomic scaffolds (e.g., the case of *TLR5* pseudogenes; see supplementary material S1: text S3 and table S15, Supplementary Material online; settings: tile size = 11, step size = 5, allowing one mismatch in tile, the number of repetitions of a tile = 1,000,000). In several avian species, where *TLRs* were previously sequenced, the sequences were found using on-line Web BLAST search (blastn searching cds database with default settings; https://blast.ncbi.nlm.nih.gov/Blast.cgi; accessed in March 2013) and downloaded directly from the GenBank nucleotide database (Benson et al. 2012). The list of NCBI accession numbers of the sequences used is given in supplementary material S2: table S23, Supplementary Material online. Basic description of the *TLR* CDS analyzed is provided in supplementary material S1: text S4, tables S16 and S17 and more detailed in supplementary material S2: table S25, Supplementary Material online.

Multiple Sequence Alignment

First, all CDSs were translated to extract the protein sequences of all representatives of each TLR family member. Then,

the multiple amino acid sequence alignment was done for each TLR separately using the ClustalW alignment tool implemented in Geneious v. 9.0.4 (Biomatters Limited; Kearse et al. 2012), or using MAFFT v. 6.850 (Katoh and Standley 2013) in the case of the duplicated genes from the vertebrate *TLR1* family. For the construction of nucleotide codon alignments corresponding to the protein sequence alignments, we used the PAL2NAL webtool v. 14 (http://www.bork.embl.de/pal2nal/; accessed in July 2014; Suyama et al. 2006). All created multiple alignments are shown in supplementary material S10, Supplementary Material online.

Gene Conversion Analysis

The occurrence of gene conversion was statistically tested on nucleotide sequence alignment of all duplicated genes (*TLR1*, *TLR2*, and *TLR7*) using GENECONV v. 1.81 (S. A. Sawyer, Washington University in St. Louis, https://www.math.wustl.edu/∼sawyer/geneconv/). Full program setting is provided for all investigated genes in supplementary material S3, Supplementary Material online.

Phylogenetic Tree Analyses

Phylogenetic analysis was done only in the duplicated avian *TLRs* (i.e., in *TLR1*, *TLR2*, and *TLR7*), also including other vertebrate taxa, that is, mammals, reptiles, amphibians, and fishes (involved species and their sequences IDs are listed in supplementary material S2: table S24, Supplementary Material online). For the *TLR1* family, the phylogenetic analysis was based on the sequence alignment of nonconverted amino acid regions only. Amino acid sequences were used instead of nucleotide sequences to avoid the biasing effect of rapid evolution on the third codon positions on broad evolutionary scales (i.e., higher vertebrate taxa). For the phylogenetic analysis, we used two approaches: maximum likelihood calculated in PhyML v. 3.0 software package (Bootstrap: 1,000; Tree type: SPR&NNI; Substitution model: LG; Guindon et al. 2010) and Bayesian estimation of phylogeny calculated using MrBayes v. 3.2.1 (number of generations: 1,000; burn-in fraction was set to default of 25%; the critical value for topological convergence diagnostic was 0.01; the best suited aa model for TLR1: Fixed(Jones) and for TLR2: Fixed(Wag); Ronquist et al. 2012). Schematic phylogenetic trees for visualization of *TLR7* duplication (fig. 2) and *TLR5* pseudogenization (fig. 3) were constructed using previously published avian phylogeny (Jarvis et al. 2014). All trees were graphically adjusted in FigTree v1.3.1 (A. Rambaut, University of Edinburgh, United Kingdom; http://tree.bio.ed.ac.uk/software/figtree/).

Copy Number Variation Analysis

To verify the increased number of *TLR7* copies in selected passerine species, we performed a qPCR copy number variation (CNV) analysis. Tissue samples of four model species (*Ficedula albicollis*, *Pseudopodoces humilis*, *Taeniopygia guttata*, and *Zonotrichia albicollis*) used for this analysis were obtained from various genetic banks (for details see supplementary material S1: table S18, Supplementary Material online). DNA was extracted using the DNeasy Blood and Tissue Kit (QIAGEN) and stored at −20°C. As reference single copy genes, we used *TLR3* and *TLR4*. Primers were designed to amplify a conserved region of a similar length for all studied genes (for details see supplementary material S1: table S19, Supplementary Material online). The specificity of the primers was previously verified by Sanger sequencing (Applied Biosystems 3130xl Genetic Analyzer) of a broader surrounding region of each gene in all investigated species (for PCR conditions see supplementary material S1: table S20, Supplementary Material online; NCBI accession numbers are listed in supplementary material S1: table S21, Supplementary Material online). Each sample was run in triplicate in LightCycler 480 Instrument II (Roche) using EvaGreen Dye (Biotum; Mao et al. 2007). The qPCR efficiency was calculated based on a dilution series (5-times dilution) for each gene and sample in LightCycler 480 software v1.5.1 using both 2nd Derivate function and automatic Fit Point method (the values measured are shown in supplementary material S1: table S4, Supplementary Material online). The *TLR7* copy numbers were then calculated based on a modified version of the formula proposed by Pfaffl (2001; eq. 1, where Eff stands for the PCR efficiency and Cp stands for the crossing point).

$$\mathbf{R} = 2* \frac{\mathbf{Eff}_{\mathrm{TLR7}}{}^{-\mathbf{Cp}_{\mathrm{TLR7}}}}{\mathbf{Eff}_{\mathrm{TLR3}}{}^{-\mathbf{Cp}_{\mathrm{TLR3}}} + \mathbf{Eff}_{\mathrm{TLR4}}{}^{-\mathbf{Cp}_{\mathrm{TLR4}}}} \quad (1)$$

Positive Selection Analysis and Estimates of Conservativeness of Amino Acid Substitutions

For the detection of selective pressures acting on each avian *TLR*, the codon alignment generated by PAL2NAL tool (http://www.bork.embl.de/pal2nal/; last accessed July 2014; Suyama et al. 2006) was used to ensure correct alignment of codons. All regions involving gaps were removed before the analysis and tested separately. The problem of missing sequence data did not impact our estimates since the number of sequences sampled was not too small in any of the TLRs and thus the dN/dS ratios could have been estimated. The position numbering of PSS followed the chicken reference sequence (for NCBI IDs see supplementary material S1: table S14, Supplementary Material online). To test for positive selection acting on individual residues at the interspecific level in avian TLRs, we used two methods based on the hierarchical Bayes (Bayes Empirical Bayes, BEB) approach implementing the Markov chain Monte Carlo routine—PAML (Phylogenetic Analysis by Maximum Likelihood; Yang 2007) and FUBAR (A Fast, Unconstrained Bayesian AppRoximation for Inferring Selection; Murrell et al. 2013). Being based on unrelated preconditions, these two tests provide independent estimates of positive selection. In PAML (Version 4.7), the codon-based substitution models (codeml) using comparison of neutral M8a (beta&$\omega = 1$) with alternative M8 (beta&ω) model were adopted. The likelihood ratio test (LRT) for comparison of two nested models was calculated using the chi-square approximation: $\chi^2 = 2\times(\mathrm{lnLM8}-\mathrm{lnLM8a})$, where LM8 and LM8a are the likelihood values. The degrees of freedom (df) were established as the difference in the numbers of parameters in used models (for details see supplementary material S1: table

S10, Supplementary Material online). If the LRT is significant ($\leq$0.05), positive selection is considered to be detected. The BEB approach (Yang 2007) was then used to determine site specific posterior probabilities indicating positive selection ($\geq$0.9) at specific codons. The phylogenetic tree needed for PAML analysis was constructed based on the previously published avian phylogeny (Jarvis et al. 2014). FUBAR (Fast Unconstrained Bayesian AppRoximation) analysis was performed at the Datamonkey server (http://www.datamonkey.org/; accessed in July 2014; Pond and Frost 2005) with the significance level of posterior probability established by default to 0.9. The FUBAR algorithm was used because it is more robust and much faster than other available selection tests, which are based on the random effect likelihood (REL; Murrell et al. 2013) methods. The degree of dissimilarity in biochemical properties of amino acid substitutions was tested using the PRIME tool (PRoperty Informed Model of Evolution; accessed in July 2014) available at the Datamonkey server (Pond and Frost 2005). This tool builds on the same conceptual frameworks as MEME (Murrell et al. 2012), but allows the nonsynonymous substitution rate β to depend not only on the site in question but also on what type of residues are being exchanged. Both predefined sets of five amino-acid composite properties were used for PRIME analysis: 1) Polarity index, Secondary structure factor, Volume, Refractivity/Heat Capacity, and Charge/Iso-electric point (Atchley et al. 2005) and 2) Chemical Composition, Polarity, Volume, Iso-electric point, and Hydropathy (Conant et al. 2007) on the significant level of posterior probabilities $\geq$0.9. Amino acid physicochemical properties (chemistry, charge, and hydrophobicity) at all PSS were graphically visualized using a web-based application Weblogo v. 3.5 (http://weblogo.threeplusone.com/create.cgi; last accessed October 2017; Crooks et al. 2004).

We compared the PSS identified in this study with the results of other studies that focused on avian and mammalian TLR evolution identifying positive selection at the interspecific level (Nakajima et al. 2008; Vinkler et al. 2009; Wlasiuk et al. 2009; Wlasiuk and Nachman 2010; Alcaide and Edwards 2011; Areal et al. 2011; Huang et al. 2011; Shen et al. 2012; Fornůsková et al. 2013; Grueber et al. 2014; Vinkler et al. 2014; Wang, Zhang, Chang, et al. 2016) and studies that described the functionally relevant residues for TLR1 (Jin et al. 2007; Omueti et al. 2007), TLR2 (Underhill et al. 1999; Lorenz et al. 2000; Xu et al. 2000; Kang and Chae 2001; Tao et al. 2002; Gautam et al. 2006; Kang et al. 2009), TLR3 (Sarkar et al. 2003; Choe et al. 2005; de Bouteiller et al. 2005; Bell et al. 2006; Sun et al. 2006; Ranjith-Kumar et al. 2007; Liu et al. 2008; Pirher et al. 2008; Luo et al. 2012; Mineev et al. 2014), TLR4 (Poltorak et al. 1998; Ronni et al. 2003; Nishitani et al. 2006; Kim et al. 2007; Walsh et al. 2008; Park et al. 2009; Resman et al. 2009; Meng et al. 2010; Ohto, Fukase, et al. 2012; Ohto, Yamakawa, et al. 2012; Garate and Oostenbrink 2013; Paramo et al. 2013; Scior et al. 2013; Wang, Su, et al. 2016), TLR5 (Jacchieri et al. 2003; Andersen-Nissen et al. 2007; Keestra et al. 2008; Yoon et al. 2012; Ivicak-Kocjan et al. 2013; Song et al. 2017), TLR7 (Wei et al. 2009; Yu et al. 2013; Tseng et al. 2014; Gentile et al. 2015; Gupta et al. 2016; Zhang et al. 2016), TLR10 (Hasan et al. 2005; Nyman et al. 2008; Guan et al. 2010; Jang and Park 2014), TLR15 (Wang, Zhang, Chang, et al. 2016), and TLR21 (Keestra et al. 2010).

Protein Structure Modeling

To predict 3D structures of all avian TLRs, the I-TASSER v. 5.0 server (https://zhanglab.ccmb.med.umich.edu/I-TASSER/; last accessed November 2016; Roy et al. 2010) was used. Since all TLRs are transmembrane proteins, the ECD and ICD (and in the case of TLR3 also TM region) were modeled separately, always based on the chicken reference sequence (for NCBI IDs see the supplementary material S1: table S14, Supplementary Material online). The domains were identified by SMART v. 7.0 web tool (http://smart.embl-heidelberg.de/; last accessed November 2016; Letunic and Bork 2018), the amino acid ranges of ECD and ICD for each TLR are provided in supplementary material S1: table S22, Supplementary Material online. The I-TASSER model with the highest C value reflecting the confidence score for estimating the quality of predicted models was always downloaded and used for further analysis. The graphical visualization of important residues was then done using the PyMOL Molecular Graphics System (Version 1.7.6, Schrödinger, LLC).

Assessing Function of Positively Selected Sites

The list of the previously reported functionally important positions as well as PSS detected in other interspecific studies for all TLRs were obtained by a detailed review of the published literature (for the complete list of references see supplementary material S6: table S27, Supplementary Material online). The distance of any PSS detected in our study from these previously described functionally important residues was measured on the 3D structural models of TLRs (obtained in previous step) in the PyMOL Molecular Graphics System (Version 1.7.6, Schrödinger, LLC) using python command iterate and plugin distancetoatom, where PSS lying in distance closer to 5 Å were considered as closely connected to the functionally important residues, that is, having potential influence on the receptor function.

Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

Acknowledgments

We would like to thank especially Zuzana Świderská and Anna Bryjová for their helpful advice during laboratory experiments and Tatiana Aghová and Petri Kemmpainen for their advices in phylogenetic analyses. Tissue samples for this analysis were kindly provided from The DNA Bank of the Natural History Museum of Oslo (NHMO), The Museum of Zoology, University of Michigan (UMMZ), and The Genetic bank of the Department of Zoology, Charles University (ZCU). The data sets generated and/or analyzed during the current study are available in the NCBI GenBank repository (see supplementary material S1: tables S14 and S21 and supplementary material S2: tables S23 and S24, Supplementary Material online) or are included in this published article and

its supplementary material files online. This work was supported by the Czech Science Foundation (P505/10/1871 and P506/15-11782S); by institutional research support (SVV-260434/2018); by the Charles University (GAUK 540214, UNCE 204069 and PRIMUS/17/SCI/12); by the travel grant provided by the Charles University Mobility Fund to H.V. and by BBSRC CASE award to M.W.G.D. The Avian Phylogenetic and B10K consortium kindly provided the access to avian genome sequences, with support from Institute Strategic Grant funding from the BBSRC. Computing and data storage facilities were provided by the National Grid Infrastructure MetaCentrum (project CESNET LM2015042). The authors' contribution to this article was as follows: H.V. (70%)—study design, performed all given analyses, prepared all output files, wrote the first draft of the article; M.W.G.D. (5%)—involvement in BLAST of genomic databases and PAML analysis; D.W.B. (10%)—supervision of M.W.G.D. and study design and M.V. (15%)—study design, revised the article. All authors contributed by their comments to the article preparation, read, and approved the final article.

References

Alcaide M, Edwards SV. 2011. Molecular evolution of the Toll-like receptor multigene family in birds. *Mol Biol Evol*. 28(5):1703–1715.

Andersen-Nissen E, Smith KD, Bonneau R, Strong RK, Aderem A. 2007. A conserved surface on Toll-like receptor 5 recognizes bacterial flagellin. *J Exp Med*. 204(2):393–403.

Arbour N, Lorenz E, Schutte B, Zabner J, Kline J, Jones M, Frees K, Watt J, Schwartz D. 2000. TLR4 mutations are associated with endotoxin hyporesponsiveness in humans. *Nat Genet*. 25(2):187–191.

Areal H, Abrantes J, Esteves PJ. 2011. Signatures of positive selection in Toll-like receptor (TLR) genes in mammals. *BMC Evol Biol*. 11:368.

Atchley WR, Zhao J, Fernandes AD, Drüke T. 2005. Solving the protein sequence metric problem. *Proc Natl Acad Sci U S A*. 102(18):6395–6400.

Bainova H, Kralova T, Bryjova A, Albrecht T, Bryja J, Vinkler M. 2014. First evidence of independent pseudogenization of Toll-like receptor 5 in passerine birds. *Dev Comp Immunol*. 45(1):151–155.

Bell JK, Askins J, Hall PR, Davies DR, Segal DM. 2006. The dsRNA binding site of human Toll-like receptor 3. *Proc Natl Acad Sci U S A*. 103(23):8792–8797.

Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. 2012. GenBank. *Nucleic Acids Res*. 41(D1):D36–D42.

Boyd AC, Peroval MY, Hammond JA, Prickett MD, Young JR, Smith AL. 2012. TLR15 is unique to avian and reptilian lineages and recognizes a yeast-derived agonist. *J Immunol*. 189(10):4930–4938.

Brownlie R, Allan B. 2011. Avian Toll-like receptors. *Cell Tissue Res*. 343(1):121–130.

Brownlie R, Zhu J, Allan B, Mutwiri GK, Babiuk LA, Potter A, Griebel P. 2009. Chicken TLR21 acts as a functional homologue to mammalian TLR9 in the recognition of CpG oligodeoxynucleotides. *Mol Immunol*. 46(15):3163–3170.

Burri R, Salamin N, Studer RA, Roulin A, Fumagalli L. 2010. Adaptive divergence of ancient gene duplicates in the avian MHC class II β. *Mol Biol Evol*. 27(10):2360–2374.

Cario E. 2013. The human TLR4 variant D299G mediates inflammation-associated cancer progression in the intestinal epithelium. *Oncoimmunology* 2(7):e24890.

Cheng Y, Prickett MD, Gutowska W, Kuo R, Belov K, Burt DW. 2015. Evolution of the avian β-defensin and cathelicidin genes. *BMC Evol Biol*. 15:188.

Choe J, Kelker MS, Wilson IA. 2005. Crystal structure of human Toll-like receptor 3 (TLR3) ectodomain. *Science* 309(5734):581–585.

Conant GC, Wagner GP, Stadler PF. 2007. Modeling amino acid substitution patterns in orthologous and paralogous genes. *Mol Phylogenet Evol*. 42(2):298–307.

Cormican P, Lloyd AT, Downing T, Connell SJ, Bradley D, O'Farrelly C. 2009. The avian Toll-Like receptor pathway-Subtle differences amidst general conformity. *Dev Comp Immunol*. 33(9):967–973.

Coscia MR, Giacomelli S, Oreste U. 2011. Toll-like receptors: an overview from invertebrates to vertebrates. *Invertebr Surviv J*. 2011:210–226.

Crooks GE, Hon G, Chandonia JM, Brenner SE. 2004. WebLogo: a sequence logo generator. *Genome Res*. 14(6):1188–1190.

de Bouteiller O, Merck E, Hasan U, Hubac S, Benguigui B, Trinchieri G, Bates E, Caux C. 2005. Recognition of double-stranded RNA by human Toll-like receptor 3 and downstream receptor signaling requires multimerization and an acidic pH. *J Biol Chem*. 280(46):38133–38145.

de Zoete MR, Bouwman LI, Keestra AM, van Putten JPM. 2011. Cleavage and activation of a Toll-like receptor by microbial proteases. *Proc Natl Acad Sci U S A*. 108(12):4968–4973.

Ellegren H. 2008. Comparative genomics and the study of evolution by natural selection. *Mol Ecol*. 17(21):4586–4596.

Emery NJ, Clayton NS. 2004. The mentality of crows: convergent evolution of intelligence in corvids and apes. *Science* 306(5703):1903–1907.

Eöry L, Gilbert MTP, Li C, Li B, Archibald A, Aken BL, Zhang G, Jarvis E, Flicek P, Burt DW. 2015. Avianbase: a community resource for bird genomics. *Genome Biol*. 16:21.

Ewald SE, Lee BL, Lau L, Wickliffe KE, Shi G-P, Chapman HA, Barton GM. 2008. The ectodomain of Toll-like receptor 9 is cleaved to generate a functional receptor. *Nature* 456(7222):658–662.

Farmer CG. 2000. Parental care: the key to understanding endothermy and other convergent features in birds and mammals. *Am Nat*. 155(3):326–334.

Ferrer-Admetlla A, Bosch E, Sikora M, Marques-Bonet T, Ramirez-Soriano A, Muntasell A, Navarro A, Lazarus R, Calafell F, Bertranpetit J, Casals F. 2008. Balancing selection is the main force shaping the evolution of innate immunity genes. *J Immunol*. 181(2):1315–1322.

Fornůsková A, Vinkler M, Pagès M, Galan M, Jousselin E, Cerqueira F, Morand S, Charbonnel N, Bryja J, Cosson J-F. 2013. Contrasted evolutionary histories of two Toll-like receptors (Tlr4 and Tlr7) in wild rodents (MURINAE). *BMC Evol Biol*. 13:194.

Garate JA, Oostenbrink C. 2013. Lipid a from lipopolysaccharide recognition: structure, dynamics and cooperativity by molecular dynamics simulations. *Proteins* 81(4):658–674.

Gautam JK, Ashish CLD, Krueger JK, Smith MF. Jr. 2006. Structural and functional evidence for the role of the TLR2 DD loop in TLR1/TLR2 heterodimerization and signaling. *J Biol Chem*. 281:30132–30142.

Gay NJ, Gangloff M. 2007. Structure and function of Toll receptors and their Ligands. *Annu Rev Biochem*. 76:141–165.

Gentile F, Deriu MA, Licandro G, Prunotto A, Danani A, Tuszynski JA. 2015. Structure based modeling of small molecules binding to the TLR7 by atomistic level simulations. *Molecules* 20(5):8316–8340.

Grueber CE, Wallis GP, Jamieson IG. 2014. Episodic positive selection in the evolution of avian Toll-like receptor innate immunity genes. *PLoS One* 9(3):e89632.

Grueber CE, Wallis GP, King TM, Jamieson IG. 2012. Variation at innate immunity Toll-like receptor genes in a bottlenecked population of a New Zealand robin. *PLoS One* 7(9):e45011.

Guan Y, Ranoa DRE, Jiang S, Mutha SK, Li X, Baudry J, Tapping RI. 2010. Human TLRs 10 and 1 share common mechanisms of innate immune sensing but not signaling. *J Immunol*. 184(9):5094–5103.

Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol*. 59(3):307–321.

Gupta CL, Akhtar S, Sayyed U, Pathak N, Bajpai P. 2016. In silico analysis of human Toll-like receptor 7 ligand binding domain. *Biotechnol Appl Biochem*. 63(3):441–450.

Hasan U, Chaffois C, Gaillard C, Saulnier V, Merck E, Tancredi S, Guiet C, Briere F, Vlach J, Lebecque S, et al. 2005. Human TLR10 is a functional

receptor, expressed by B cells and plasmacytoid dendritic cells, which activates gene transcription through MyD88. *J Immunol*. 174(5):2942–2950.
Hawn TR, Verbon A, Lettinga KD, Zhao LP, Li SS, Laws RJ, Skerrett SJ, Beutler B, Schroeder L, Nachman A, et al. 2003. A common dominant TLR5 stop codon polymorphism abolishes flagellin signaling and is associated with susceptibility to Legionnaires' disease. *J Exp Med*. 198(10):1563–1572.
Hawn TR, Wu H, Grossman JM, Hahn BH, Tsao BP, Aderem A. 2005. A stop codon polymorphism of Toll-like receptor 5 is associated with resistance to systemic lupus erythematosus. *Proc Natl Acad Sci U S A*. 102(30):10593–10597.
Hayashi F, Smith K, Ozinsky A, Hawn T, Yi E, Goodlett D, Eng J, Akira S, Underhill D, Aderem A. 2001. The innate immune response to bacterial flagellin is mediated by Toll-like receptor 5. *Nature* 410(6832):1099–1103.
Higuchi M, Matsuo A, Shingai M, Shida K, Ishii A, Funami K, Suzuki Y, Oshiumi H, Matsumoto M, Seya T. 2008. Combinational recognition of bacterial lipoproteins and peptidoglycan by chicken Toll-like receptor 2 subfamily. *Dev Comp Immunol*. 32(2):147–155.
Huang Y, Temperley ND, Ren L, Smith J, Li N, Burt DW. 2011. Molecular evolution of the vertebrate TLR1 gene family – a complex history of gene duplication, gene conversion, positive selection and co-evolution. *BMC Evol Biol*. 11(149):1–17.
Iqbal M, Philbin VJ, Withanage GSK, Wigley P, Beal RK, Goodchild MJ, Barrow P, McConnell I, Maskell DJ, Young J, et al. 2005. Identification and functional characterization of chicken Toll-like receptor 5 reveals a fundamental role in the biology of infection with *Salmonella enterica* serovar typhimurium. *Infect Immun*. 73(4):2344–2350.
Ivicak-Kocjan K, Panter G, Bencina M, Jerala R. 2013. Determination of the physiological 2: 2 TLR5: flagellin activation stoichiometry revealed by the activity of a fusion receptor. *Biochem Biophys Res Commun*. 435(1):40–45.
Iwasaki A, Medzhitov R. 2015. Control of adaptive immunity by the innate immune system. *Nat Immunol*. 16(4):343–353.
Jacchieri S, Torquato R, Brentani R. 2003. Structural study of binding of flagellin by Toll-like receptor 5. *J Bacteriol*. 185(14):4243–4247.
Janeway CA, Medzhitov R. 2002. Innate immune recognition. *Annu Rev Immunol*. 20:197–216.
Jang T, Park HH. 2014. Crystal structure of TIR domain of TLR6 reveals novel dimeric interface of TIR-TIR interaction for Toll-like receptor signaling pathway. *J Mol Biol*. 426(19):3305–3313.
Jarvis ED, Mirarab S, Aberer AJ, Li B, Houde P, Li C, Ho SYW, Faircloth BC, Nabholz B, Howard JT, et al. 2014. Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science* 346(6215):1320–1331.
Jetz W, Thomas GH, Joy JB, Hartmann K, Mooers AO. 2012. The global diversity of birds in space and time. *Nature* 491(7424):444–448.
Jin MS, Kim SE, Heo JY, Lee ME, Kim HM, Paik S-G, Lee H, Lee J-O. 2007. Crystal structure of the TLR1-TLR2 heterodimer induced by binding of a tri-acylated lipopeptide. *Cell* 130(6):1071–1082.
Kang JY, Nan X, Jin MS, Youn S-J, Ryu YH, Mah S, Han SH, Lee H, Paik S-G, Lee J-O. 2009. Recognition of lipopeptide patterns by Toll-like receptor 2-Toll-like receptor 6 heterodimer. *Immunity* 31(6):873–884.
Kang T, Chae G. 2001. Detection of Toll-like receptor 2 (TLR2) mutation in the lepromatous leprosy patients. *FEMS Immunol Med Microbiol*. 31(1):53–58.
Kanno A, Yamamoto C, Onji M, Fukui R, Saitoh S, Motoi Y, Shibata T, Matsumoto F, Muta T, Miyake K. 2013. Essential role for Toll-like receptor 7 (TLR7)-unique cysteines in an intramolecular disulfide bond, proteolytic cleavage and RNA sensing. *Int Immunol*. 25(7):413–422.
Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 30(4):772–780.
Kawai T, Akira S. 2010. The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. *Nat Immunol*. 11(5):373–384.
Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28(12):1647–1649.
Keestra AM, de Zoete MR, Bouwman LI, Vaezirad MM, van Putten JPM. 2013. Unique features of chicken Toll-like receptors. *Dev Comp Immunol*. 41(3):316–323.
Keestra AM, de Zoete MR, Bouwman LI, van Putten JPM. 2010. Chicken TLR21 is an innate CpG DNA receptor distinct from mammalian TLR9. *J Immunol*. 185(1):460–467.
Keestra AM, de Zoete MR, van Aubel RAMH, van Putten JPM. 2007. The central Leucine-rich repeat region of chicken TLR16 dictates unique ligand specificity and species-specific interaction with TLR2. *J Immunol*. 178(11):7110.
Keestra AM, de Zoete MR, van Aubel RAMH, van Putten JPM. 2008. Functional characterization of chicken TLR5 reveals species-specific recognition of flagellin. *Mol Immunol*. 45(5):1298–1307.
Kim HM, Park BS, Kim J-I, Kim SE, Lee J, Oh SC, Enkhbayar P, Matsushima N, Lee H, Yoo OJ, Lee J-O. 2007. Crystal structure of the TLR4-MD-2 complex with bound endotoxin antagonist Eritoran. *Cell* 130(5):906–917.
Letunic I, Bork P. 2018. 20 years of the SMART protein domain annotation resource. *Nucleic Acids Res*. 46(D1):D493–D496.
Leveque G, Forgetta V, Morroll S, Smith A, Bumstead N, Barrow P, Loredo-Osti J, Morgan K, Malo D. 2003. Allelic variation in TLR4 is linked to susceptibility to Salmonella enterica serovar typhimurium infection in chickens. *Infect Immun*. 71(3):1116–1124.
Liu L, Botos I, Wang Y, Leonard JN, Shiloach J, Segal DM, Davies DR. 2008. Structural basis of Toll-like receptor 3 signaling with double-stranded RNA. *Science* 320(5874):379–381.
Lorenz E, Mira J, Cornish K, Arbour N, Schwartz D. 2000. A novel polymorphism in the Toll-like receptor 2 gene and its potential association with staphylococcal infection. *Infect Immun*. 68(11):6398–6401.
Lovegrove BG. 2017. A phenology of the evolution of endothermy in birds and mammals. *Biol Rev*. 92(2):1213–1240.
Luo J, Obmolova G, Malia TJ, Wu S-J, Duffy KE, Marion JD, Bell JK, Ge P, Zhou ZH, Teplyakov A, et al. 2012. Lateral clustering of TLR3: dsRNA signaling units revealed by TLR3ecd: 3Fabs quaternary structure. *J Mol Biol*. 421(1):112–124.
Mao F, Leung W-Y, Xin X. 2007. Characterization of EvaGreen and the implication of its physicochemical properties for qPCR applications. *BMC Biotechnol*. 7:76.
Medvedev AE. 2013. Toll-like receptor polymorphisms, inflammatory and infectious diseases, allergies, and cancer. *J Interferon Cytokine Res off J Int Soc Interferon Cytokine Res*. 33(9):467–484.
Meng J, Lien E, Golenbock DT. 2010. MD-2-mediated ionic interactions between lipid A and TLR4 are essential for receptor activation. *J Biol Chem*. 285(12):8695–8702.
Mikami T, Miyashita H, Takatsuka S, Kuroki Y, Matsushima N. 2012. Molecular evolution of vertebrate Toll-like receptors: evolutionary rate difference between their leucine-rich repeats and their TIR domains. *Gene* 503(2):235–243.
Mineev KS, Goncharuk SA, Arseniev AS. 2014. Toll-like receptor 3 transmembrane domain is able to perform various homotypic interactions: an NMR structural study. *FEBS Lett*. 588(21):3802–3807.
Murrell B, Moola S, Mabona A, Weighill T, Sheward D, Kosakovsky Pond SL, Scheffler K. 2013. FUBAR: a fast, unconstrained Bayesian AppRoximation for inferring selection. *Mol Biol Evol*. 30(5):1196–1205.
Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K, Pond SLK. 2012. Detecting individual sites subject to episodic diversifying selection. *PLoS Genet*. 8(7):e1002764.
Nakajima T, Ohtani H, Satta Y, Uno Y, Akari H, Ishida T, Kimura A. 2008. Natural selection in the TLR-related genes in the course of primate evolution. *Immunogenetics* 60(12):727–735.
Netea MG, Wijmenga C, O'Neill LAJ. 2012. Genetic variation in Toll-like receptors and disease susceptibility. *Nat Immunol*. 13(6):535–542.

Nishitani C, Mitsuzawa H, Sano H, Shimizu T, Matsushima N, Kuroki Y. 2006. Toll-like receptor 4 region Glu24-Lys47 is a site for MD-2 binding: importance of CYS29 and CYS40. *J Biol Chem*. 281(50):38322–38329.

Nyman T, Stenmark P, Flodin S, Johansson I, Hammarström M, Nordlund P. 2008. The crystal structure of the human Toll-like receptor 10 cytoplasmic domain reveals a putative signaling dimer. *J Biol Chem*. 283(18):11861–11865.

Ohto U, Fukase K, Miyake K, Shimizu T. 2012. Structural basis of species-specific endotoxin sensing by innate immune receptor TLR4/MD-2. *Proc Natl Acad Sci U S A*. 109:7421–7426.

Ohto U, Yamakawa N, Akashi-Takamura S, Miyake K, Shimizu T. 2012. Structural analyses of human Toll-like receptor 4 polymorphisms D299G and T399I. *J Biol Chem*. 287(48):40611–40617.

Olkowicz S, Kocourek M, Lučan RK, Porteš M, Fitch WT, Herculano-Houzel S, Němec P. 2016. Birds have primate-like numbers of neurons in the forebrain. *Proc Natl Acad Sci U S A*. 113(26):7255–7260.

Omueti KO, Mazur DJ, Thompson KS, Lyle EA, Tapping RI. 2007. The polymorphism P315L of human Toll-like receptor 1 impairs innate immune sensing of microbial cell wall components. *J Immunol*. 178(10):6387–6394.

Palm NW, Medzhitov R. 2009. Pattern recognition receptors and control of adaptive immunity. *Immunol Rev*. 227(1):221–233.

Paramo T, Piggot TJ, Bryant CE, Bond PJ. 2013. The structural basis for endotoxin-induced allosteric regulation of the Toll-like receptor 4 (TLR4) innate immune receptor. *J Biol Chem*. 288(51):36215–36225.

Park BS, Song DH, Kim HM, Choi B-S, Lee H, Lee J-O. 2009. The structural basis of lipopolysaccharide recognition by the TLR4-MD-2 complex. *Nature* 458(7242):1191–1195.

Pfaffl MW. 2001. A new mathematical model for relative quantification in real-time RT–PCR. *Nucleic Acids Res*. 29(9):e45.

Philbin VJ, Iqbal M, Boyd Y, Goodchild MJ, Beal RK, Bumstead N, Young J, Smith AL. 2005. Identification and characterization of a functional, alternatively spliced Toll-like receptor 7 (TLR7) and genomic disruption of TLR8 in chickens. *Immunology* 114(4):507–521.

Pirher N, Ivicak K, Pohar J, Bencina M, Jerala R. 2008. A second binding site for double-stranded RNA in TLR3 and consequences for interferon activation. *Nat Struct Mol Biol*. 15(7):761–763.

Poltorak A, He X, Smirnova I, Liu M-Y, Huffel CV, Du X, Birdwell D, Alejos E, Silva M, Galanos C. 1998. Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene. *Science* 282(5396):2085–2088.

Pond SLK, Frost SDW. 2005. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics* 21(10):2531–2533.

Ranjith-Kumar CT, Miller W, Xiong J, Russell WK, Lamb R, Santos J, Duffy KE, Cleveland L, Park M, Bhardwaj K, et al. 2007. Biochemical and functional analyses of the human Toll-like receptor 3 ectodomain. *J Biol Chem*. 282(10):7668–7678.

Raven N, Lisovski S, Klaassen M, Lo N, Madsen T, Ho SYW, Ujvari B. 2017. Purifying selection and concerted evolution of RNA-sensing Toll-like receptors in migratory waders. *Infect Genet Evol*. 53:135–145.

Reddick LE, Alto NM. 2014. Bacteria fighting back: how pathogens target and subvert the host innate immune system. *Mol Cell* 54(2):321–328.

Resman N, Vasl J, Oblak A, Pristovsek P, Gioannini TL, Weiss JP, Jerala R. 2009. Essential roles of hydrophobic residues in both MD-2 and Toll-like receptor 4 in activation by endotoxin. *J Biol Chem*. 284(22):15052–15060.

Roach JC, Glusman G, Rowen L, Kaur A, Purcell MK, Smith KD, Hood LE, Aderem A. 2005. The evolution of vertebrate Toll-like receptors. *Proc Natl Acad Sci U S A*. 102(27):9577–9582.

Ronni T, Agarwal V, Haykinson M, Haberland M, Cheng G, Smale S. 2003. Common interaction surfaces of the Toll-like receptor 4 cytoplasmic domain stimulate multiple nuclear targets. *Mol Cell Biol*. 23(7):2543–2555.

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol*. 61(3):539–542.

Roy A, Kucukural A, Zhang Y. 2010. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc*. 5(4):725–738.

Rupasree Y, Naushad SM, Rajasekhar L, Uma A, Kutala VK. 2015. Association of TLR4 (D299G, T399I), TLR9-1486T < C, TIRAP S180L and TNF-alpha promoter (-1031,-863,-857) polymorphisms with risk for systemic lupus erythematosus among South Indians. *Lupus* 24:50–57.

Sarkar S, Smith H, Rowe T, Sen G. 2003. Double-stranded RNA signaling by Toll-like receptor 3 requires specific tyrosine residues in its cytoplasmic domain. *J Biol Chem*. 278(7):4393–4396.

Scior T, Lozano-Aponte J, Figueroa-Vazquez V, Yunes-Rojas JA, Zähringer U, Alexander C. 2013. Three-dimensional mapping of differential amino acids of human, murine, canine and equine TLR4/MD-2 receptor complexes conferring endotoxic activation by lipid A, antagonism by Eritoran and species-dependent activities of Lipid IVA in the mammalian LPS sensor system. *Comput Struct Biotechnol J*. 7:e201305003.

Shen T, Xu S, Wang X, Yu W, Zhou K, Yang G. 2012. Adaptive evolution and functional constraint at TLR4 during the secondary aquatic adaptation and diversification of cetaceans. *BMC Evol Biol*. 12:39.

Shen X, Shi R, Zhang H, Li K, Zhao Y, Zhang R. 2010. The Toll-like receptor 4 D299G and T399I polymorphisms are associated with Crohn's disease and ulcerative colitis: a meta-analysis. *Digestion* 81(2):69–77.

Song WS, Jeon YJ, Namgung B, Hong M, Yoon S. 2017. A conserved TLR5 binding and activation hot spot on flagellin. *Sci Rep*. 7:40878.

Sun J, Duffy KE, Ranjith-Kumar CT, Xiong J, Lamb RJ, Santos J, Masarapu H, Cunningham M, Holzenburg A, Sarisky RT, et al. 2006. Structural and functional analyses of the human Toll-like receptor 3 – role of glycosylation. *J Biol Chem*. 281(16):11144–11151.

Suyama M, Torrents D, Bork P. 2006. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res*. 34(Web Server):W609–W612.

Tao X, Xu Y, Zheng Y, Beg AA, Tong L. 2002. An extensively associated dimer in the structure of the C713S mutant of the TIR domain of human TLR2. *Biochem Biophys Res Commun*. 299(2):216–221.

Temperley ND, Berlin S, Paton IR, Griffin DK, Burt DW. 2008. Evolution of the chicken Toll-like receptor gene family: a story of gene gain and gene loss. *BMC Genomics* 9:62.

Tseng C-Y, Gajewski M, Danani A, Tuszynski JA. 2014. Homology and molecular dynamics models of Toll-like receptor 7 protein and its dimerization. *Chem Biol Drug Des*. 83(6):656–665.

Underhill D, Ozinsky A, Hajjar A, Stevens A, Wilson C, Bassetti M, Aderem A. 1999. The Toll-like receptor 2 is recruited to macrophage phagosomes and discriminates between pathogens. *Nature* 401(6755):811–815.

Vinkler M, Bainova H, Bryja J. 2014. Protein evolution of Toll-like receptors 4, 5 and 7 within Galloanserae birds. *Genet Sel Evol*. 46:72.

Vinkler M, Bryjova A, Albrecht T, Bryja J. 2009. Identification of the first Toll-like receptor gene in passerine birds: tLR4 orthologue in zebra finch (Taeniopygia guttata). *Tissue Antigens* 74(1):32–41.

Waite DW, Taylor MW. 2014. Characterizing the avian gut microbiota: membership, driving influences, and potential function. *Front Microbiol*. 5:223.

Walsh C, Gangloff M, Monie T, Smyth T, Wei B, McKinley TJ, Maskell D, Gay N, Bryant C. 2008. Elucidation of the MD-2/TLR4 interface required for signaling by lipid IVa. *J Immunol*. 181(2):1245–1254.

Wang J, Zhang Z, Chang F, Yin D. 2016. Bioinformatics analysis of the structural and evolutionary characteristics for Toll-like receptor 15. *PeerJ* 4:e2079.

Wang J, Zhang Z, Liu J, Zhao J, Yin D. 2016. Ectodomain architecture affects sequence and functional evolution of vertebrate Toll-like receptors. *Sci Rep*. 6:26705.

Wang Y, Su L, Morin MD, Jones BT, Whitby LR, Surakattula MMRP, Huang H, Shi H, Choi JH, Wang K-w. 2016. TLR4/MD-2 activation by a synthetic agonist with no similarity to LPS. *Proc Natl Acad Sci U S A*. 113(7):E884–E893.

Wei T, Gong J, Jamitzky F, Heckl WM, Stark RW, Rössle SC. 2009. Homology modeling of human Toll-like receptors TLR7, 8, and 9 ligand-binding domains. *Protein Sci*. 18(8):1684–1691.

Wlasiuk G, Khan S, Switzer WM, Nachman MW. 2009. A history of recurrent positive selection at the Toll-like receptor 5 in primates. *Mol Biol Evol*. 26(4):937–949.

Wlasiuk G, Nachman MW. 2010. Adaptation and constraint at Toll-like receptors in primates. *Mol Biol Evol*. 27(9):2172–2186.

Xu Y, Tao X, Shen B, Horng T, Medzhitov R, Manley JL, Tong L. 2000. Structural basis for signal transduction by the Toll/interleukin-1 receptor domains. *Nature* 408(6808):111–115.

Yang J, Zhang E, Liu F, Zhang Y, Zhong M, Li Y, Zhou D, Chen Y, Cao Y, Xiao Y, et al. 2014. Flagellins of Salmonella typhi and nonpathogenic Escherichia coli are differentially recognized through the NLRC4 pathway in macrophages. *J Innate Immun*. 6(1):47–57.

Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol*. 24(8):1586–1591.

Yoon S, Kurnasov O, Natarajan V, Hong M, Gudkov AV, Osterman AL, Wilson IA. 2012. Structural basis of TLR5-flagellin recognition and signaling. *Science* 335(6070):859–864.

Yu H, Jin H, Sun L, Zhang L, Sun G, Wang Z, Yu Y. 2013. Toll-like receptor 7 agonists: chemical feature based pharmacophore identification and molecular docking studies. *PLoS One* 8(3):e56514.

Zareparsi S, Buraczynska M, Branham KEH, Shah S, Eng D, Li M, Pawar H, Yashar BM, Moroi SE, Lichter PR, et al. 2005. Toll-like receptor 4 variant D299G is associated with susceptibility to age-related macular degeneration. *Hum Mol Genet*. 14(11):1449–1455.

Zhang F, Gao X-D, Wu W-W, Gao Y, Zhang Y-W, Wang S-P. 2013. Polymorphisms in Toll-like receptors 2, 4 and 5 are associated with Legionella pneumophila infection. *Infection* 41(5):941–948.

Zhang G, Jarvis ED, Gilbert MTP. 2014. A flock of genomes. *Science* 346(6215):1308–1309.

Zhang G, Li C, Li Q, Li B, Larkin DM, Lee C, Storz JF, Antunes A, Greenwold MJ, Meredith RW, et al. 2014. Comparative genomics reveals insights into avian genome evolution and adaptation. *Science* 346:1311–1320.

Zhang J. 2003. Evolution by gene duplication: an update. *Trends Ecol Evol*. 18(6):292–298.

Zhang Z, Ohto U, Shibata T, Krayukhina E, Taoka M, Yamauchi Y, Tanji H, Isobe T, Uchiyama S, Miyake K, et al. 2016. Structural analysis reveals that Toll-like receptor 7 is a dual receptor for guanosine and single-stranded RNA. *Immunity* 45(4):737–748.

Zhang Z, Schwartz S, Wagner L, Miller W. 2000. A greedy algorithm for aligning DNA sequences. *J Comput Biol*. 7(1–2):203–214.

Zhao Y, Yang J, Shi J, Gong Y-N, Lu Q, Xu H, Liu L, Shao F. 2011. The NLRC4 inflammasome receptors for bacterial flagellin and type III secretion apparatus. *Nature* 477(7366):596–600.