Adaptive evolution of Toll-like receptor 5 in domesticated mammals

Sarah A Smith1, Oliver C Jann2, David Haig1, George C Russell3, Dirk Werling4, Elizabeth J Glass2 and Richard D Emes1*

Abstract

Background: Previous studies have proposed that mammalian toll like receptors (TLRs) have evolved under diversifying selection due to their role in pathogen detection. To determine if this is the case, we examined the extent of adaptive evolution in the TLR5 gene in both individual species and defined clades of the mammalia.

Results: In support of previous studies, we find evidence of adaptive evolution of mammalian TLR5. However, we also show that TLR5 genes of domestic livestock have a concentration of single nucleotide polymorphisms suggesting a specific signature of adaptation. Using codon models of evolution we have identified a concentration of rapidly evolving codons within the TLR5 extracellular domain a site of interaction between host and the bacterial surface protein flagellin.

Conclusions: The results suggest that interactions between pathogen and host may be driving adaptive change in TLR5 by competition between species. In support of this, we have identified single nucleotide polymorphisms (SNP) in sheep and cattle TLR5 genes that are co-localised and co-incident with the predicted adaptive codons suggesting that adaptation in this region of the TLR5 gene is on-going in domestic species.

Keywords: Toll-like receptor, SNP, Adaptive evolution, Positive selection, Sheep, Cattle

Background

Toll-like receptors (TLRs) are type 1 transmembrane glycoproteins expressed on the cell surface and intracellular compartments of many cell types including epithelial cells and a variety of immune cells such as macrophages and dendritic cells. TLR ligands include pathogen-associated molecular pattern (PAMP) molecules, and TLRs are amongst the first receptors to respond to pathogen presence [1], hence their key role in innate immunity to infection. At least 10 TLRs have been identified in mammals and collectively these recognize a wide repertoire of microbial organisms and pathogens including bacteria, viruses, protozoa and fungi [2]. The TLR protein is comprised of three main regions: an extracellular pattern-recognition receptor domain (ECD), a transmembrane region and an intracellular TIR signalling domain [3]. The signalling domain is highly conserved across the TLRs. In contrast, the ECD involved in pathogen detection is often variable [2].

TLR5 is known to bind bacterial flagellin [4]. Both flagellin (e.g. of E.coli) [4]), and the ECD of TLR5 in primates [5] and other mammals [6] show evidence of adaptive positive selection. This suggests that interspecies competition between host and pathogen is likely to be driving the co-evolution of pathogen and host. In support of this, species-specific single nucleotide variations in the TLR5 gene exist and a single nucleotide polymorphism (SNP) in the ECD of mouse, chicken and human TLR5 is associated with a species-specific response to flagellin [7,8].

The domestication of livestock by selection of desirable traits gave rise to the concept of breeds over 200 years ago [9]. This formation of breeds by selective interbreeding offers a unique opportunity to examine an accelerated process of natural selection. To investigate the evolution of the TLR5 gene in domestic livestock compared to other mammals we used phylogenetic methods to identify species-specific and branch-specific evidence of positive selection. To investigate the potential role of recent
variation on evolution of the TLR5 gene we also identified known and novel SNPs in the coding region of TLR5 of sheep and cattle breeds.

Results

Evidence for adaptive evolution in mammalian TLR5

Positive diversifying selection acting on a gene can be inferred when the ratio of non-synonymous (dN) to synonymous (dS) substitution rates is greater than 1. This ratio dN/dS (also known as omega) provides a method to compare the evolutionary history of codons and lineages [10]. The parameters dN and dS can be estimated by a number of approaches. We applied the codon models of PAML [11] to infer estimates of parameters under a maximum likelihood framework. The results are summarized in Table 1 and Table 2. Complete results and parameter estimates for all PAML analyses are given as Additional file 1 and Additional file 2.

When comparing rates of codon evolution in TLR5 since the divergence of the mammals (sites analysis, see methods) significant evidence (posterior probability Pb > 0.95) was found to suggest that positive selection has been acting on three codons (G104, H592, A659) Table 1. When using the branch-site test (see methods) to compare each lineage independently, significant evidence was obtained that the sloth, sheep, cattle and pig lineages were each evolving at an elevated rate compared to other branches of the phylogeny. Three positively-selected codons (I393, Q495, H720, Pb > 0.95) were detected in the cattle lineage. One codon (K326, Pb > 0.99) was identified in the sheep lineage and one codon (E630, Pb > 0.95) was identified in the pig lineage (Table 2). These sites are in addition to those identified as having evolved under positive selection across the mammalian lineage.

Using the multiple branch-sites analysis (see methods), positive selection of TLR5 was detected in the artiodactyls, which contain the domestic ungulate species (cow, sheep and pig), but not in the laurasiatheria (of which artiodactyls are a component clade) or in the separate euarchontoglires and primate clades (Figure 1). In the artiodactyls, eight codons of TLR5 were detected with significant evidence of positive selection in (L34, S268, I295, I307, I295, I307, I393, Q495, I621, E630, (Pb > 0.95). Five codons (L34, S268, I295, I307, I621) were only identified as having evolved under positive selection when the members of the artiodactyla were combined in the multiple branch-sites analysis. Three codons (I393, Q495, E630) were detected previously in ungulate species-specific lineages (Table 2).

Table 1 Detection of positive selection across mammalian TLR5

| Analysis          | LRT M1a versus M2a (significance) | LRT M7 versus M8 (significance) | dN/dS (model) | Sites under positive selection* |
|-------------------|----------------------------------|---------------------------------|---------------|----------------------------------|
| All Mammals       | 18.87 (p < 0.001)                | 28.61 (p < 0.001)              | 2.85 (M2), 1.36 (M8) | G104*                            |
|                   |                                  |                                 |               | H592*                            |
|                   |                                  |                                 |               | A659*                            |

Results of PAML site specific analysis of positive selection across mammalian TLR5 and branch-site analysis. Positively selected codons detected by both M2a and M8 models are shown * = Pb > 95%.

Table 2 Detection of positive selection of mammalian TLR5 using the branch-sites test

| Foreground branch(es) | LRT Foreground dN/dS = 1 versus foreground dN/dS > 1 | dN/dS (branch) | Sites under positive selection* |
|-----------------------|-----------------------------------------------------|---------------|----------------------------------|
| Artiodactyls          | 35.61 (p < 0.001)                                   | 4.76 (foreground) | L34*                            |
|                       |                                                    |               | S268*                            |
|                       |                                                    |               | I295*                            |
|                       |                                                    |               | I307*                            |
|                       |                                                    |               | I393*                            |
|                       |                                                    |               | Q495*                            |
|                       |                                                    |               | I621*                            |
|                       |                                                    |               | E630*                            |
| Cow                   | 21.84 (p < 0.001)                                   | 140.91 (foreground) | I393**                           |
|                       |                                                    |               | Q495*                            |
|                       |                                                    |               | H720*                            |
| Sheep                 | 18.10 (p < 0.01)                                   | 493.28 (foreground) | K326*                            |
| Pig                   | 11.89 (p < 0.05)                                   | 29.07 (foreground) | E630**                           |

Results of branch-sites analysis of positive selection within mammalian TLR5. * posterior probabilities (Pb) of codons existing in dN/dS > 1 site class * = Pb > 0.95 and ** Pb > 0.99.
When mapped onto the predicted tertiary structure of bovine TLR5 (PDB 2a0zA [12] and 1fyv [13]) the location of sites of positive selection detected by all approaches revealed a bias in their distribution. Eleven (L34, G104, S268, I295, I307, K326, I393, Q495, H592, I621, E630) of the thirteen positively-selected codons encode amino acids in the ECD (Figure 2a). When the total number of sites in the ECD compared to the rest of the protein are accounted for, this enrichment remains statistically-significant (Fisher Exact Test: \( P = 0.03 \)). Additionally five of these (S268, I295, I307, K326, I393) which exhibited evidence of positive selection in the artiodactyl clade are located within the putative flagellin-binding region of the ECD close to the conserved concave surface-associated with ligand binding [7] (Figure 2b).

### SNP Detection

A total of 19 polymorphic sites were detected in cattle and 25 in sheep. No overlap between cattle and sheep was seen (Additional file 3 and Additional file 4). All but the Mongolian cattle breed and the Soay sheep breed showed variability within tested individuals (Figure 3).

Of the 19 polymorphisms detected in cattle, 8 are synonymous substitutions and 11 are non-synonymous substitutions. Twelve of the SNPs detected have been previously reported [14,15]. The remaining 7 SNPs were novel discoveries of this study, and three of these were non-synonymous (L34P, R59K and H262R). All except two polymorphisms are sub-species specific, with Bos indicus displaying the highest degree of genetic variability (details in supplementary information). Analysis of sheep breeds identified 25 novel SNPs, of which 13 were synonymous and 12 were non-synonymous substitutions (Figure 3). In addition, two SNPs were identified which are predicted to cause premature stop codons in cattle TLR5 (Figure 3 and Additional file 4). R125* was detected in the Jersey breed and has recently been reported [15]. The other putative stop codon (S431*) was detected in a selection of Bos taurus breeds and is a novel discovery of this study. Pseudogenes are predicted to evolve under neutral selection and as such are not subject to the same evolutionary constraints assumed for protein coding genes. Thus to avoid possible problems of including potential pseudogenes in evolutionary analyses, the stop codon variants were excluded from PAML analysis.

### Co-localisation of SNPs and Positively-Selected Sites

Two non-synonymous SNPs detected in the bovine species co-occurred at codons detected as evolving under positive selection. Codon L34 (\( d_N/d_S > 1 \) in artiodactyl) is positioned at the N-terminal region of the extracellular domain of TLR5. Amino acid A659 (\( d_N/d_S > 1 \) in all mammals) is located in the transmembrane region and is predicted to extend this domain (Figure 4 and Additional file 3 and Additional file 5). In sheep, two non-synonymous SNPs at position 2 and 3 of codon A659 were also detected (Figure 4 and Additional file 4). Due to the relatively low density of SNPs in close proximity within a single gene, these were not used to estimate population genetic measures such as linkage disequilibrium.
Figure 2 Localisation of TLR5 sites of positive selection. A) predicted tertiary protein structure of bovine TLR5 and sites of positive selection. B) predicted tertiary protein structure of ECD and sites of positive selection and conserved amino acids within putative flagellin binding region. Black backbone depicts putative flagellin binding region; Light blue amino acids = Positive selection acting on all mammals (G104, H592, A659); Purple amino acids = Positive selection acting on domestic ungulates (L34, S268, I295, I307, I393, Q495, I621, E630); Orange amino acid = positive selection in sheep lineage (K326), Green amino acids and black arrows = positive selection in cattle lineage (I393, Q495, H720); Dark blue amino acids = Conserved amino acids predicted to be involved in the detection of flagellin.

Figure 3 SNPs detected in TLR5. SNPs detected for each breed aligned to predicted secondary structure of TLR5. SP = Signal peptide; NT = LRR N-terminal; 1 – 22 in yellow = denotes each LRR; CT = LRR C-terminal; TM = transmembrane region; red bars = Non-synonymous polymorphisms; yellow bars = synonymous polymorphisms; White bars = putative stop codons; Black table = cattle breeds; Blue table = sheep breeds.
Discussion

Whilst many studies of the phylogeny and comparative genomics of the TLR gene family exist, this is the first study to characterize the lineage-specific adaptive evolution of the TLR5 gene across clades of the mammalian phylogeny. Three codons (G104, H592, A659) exhibited clear evidence for positive selection across the mammalian phylogeny. This extends previous discoveries that detected positive selection in TLR5 in primate species [5] or in selected mammalian species [6]. Areal et al. showed that positive selection is seen in a number of genes of the TLR family of proteins including multiple sites of positive selection in the TLR5 genes when using a subset of mammals [6]. Whilst our findings largely support those of their study, Areal et al. showed that positive selection was seen in a subset of mammalian species [6].

Seventeen species were studied compared to the 37 species of mammal used in this study. The current study is proposed to have increased power to resolve true signature of positive selection. Importantly, the Areal study included the chicken TLR5 gene sequence (NM_001024586.1) in the analysis of TLR5 (see Table S5 of [6]). This would potentially introduce the problem of saturation (see Methods section). Additionally, our study included the novel analysis of groups of lineages to investigate changes in evolutionary constraint in different lineages within the mammals. When $d_{SN}/d_S$ between lineages or clades were compared (single branch-sites and multiple branch-sites analysis) the artiodactyl lineage and the individual porcine, ovine and bovine lineages comprising this clade exhibited significant evidence of adaptive evolution ($d_{SN}/d_S > 1$). Within the artiodactyl clade, eight positively-selected sites were detected. Positive selective pressure on genes is symptomatic of functional adaptations acquired during the evolution of species and can promote species functional diversification [16]. This suggests that positive selection observed within the artiodactyl clade is different when compared to that seen in other mammals. We postulate that adaptive evolution observed in TLR5 of domestic livestock is a result of the breeding process. In support of this, it has been previously proposed that ruminant species are undergoing differential selective pressure in the related TLR2 genes [17].

This phenomenon may be directly caused by selective breeding resulting in a rapidly restricted population. The effective population size of all cattle breeds is known to have decreased in recent history and this may reflect initial domestication, breed formation or selection for breed specific production traits (beef or milk) [18]. However, the genetic diversity of cattle as opposed to other species such as dogs is not as low as the effective population size would suggest and high levels of divergent selection associated with immune genes amongst others are detectable.
S. typhimurium response to recombinant FliC of and bovine TLR5 has been shown to have a reduced recognition compared to human or murine TLR5 [8], chicken TLR5 has been shown to recognise different flagellins from different species to flagellin. Indeed, TLR5-flagellin interaction has recently been mapped at the single amino acid level [7,19,20] suggesting that changes of amino acids in either, TLR5 or flagellin can alter the species-specific ligand recognition [7]. Evidence of increased positive selection within the extracellular domain supports the hypothesis that competition with flagellated bacterial pathogens is driving adaptation in specific host TLR5 ECD and more precisely in the flagellin-binding region. This may be counterintuitive as many livestock will be expected to share similar microbiota and pathogens. However, TLR5-flagellin interaction is known to have been mapped at the single amino acid level [7,19,20] suggesting that changes of amino-acids within either, TLR5 or flagellin can alter the species-specific TLR5 response. This in turn may influence the host range and susceptibility of infection. Such amino-acid changes could explain some of the biological differences seen in the response of different species to flagellin. Indeed, chicken TLR5 has been shown to recognise different flagellin-forms compared to human or murine TLR5 [8], and bovine TLR5 has been shown to have a reduced response to recombinant FliC of S. typhimurium [21]. These differences are partially based on flagellin-amino-acid differences, and alteration of these amino acids in different flagellins alters their interaction with TLR5 from different species [8,22]. It has been proposed that PAMP ligands engage with the concave surface of their cognate TLR ECDs [23]. Five codons (S268, I295, I307, K326, I393) under positive selection in the artiodactyl clade are in close proximity to this conserved cavity. The non-synonymous SNPs detected within the putative flagellin-binding region are good candidates for genetic variants most likely to impact upon the immune response mediated by TLR5. The function of some of these variants is currently being pursued. However, sites more distal to the flagellin-binding region may also be of importance in TLR5 ligand recognition and function, for example by altering the shape of the molecule or interfering with signal transduction. For example, in TLR3 a SNP outside of the extracellular region was found to impair receptor signalling [24]. Also a polymorphism in the transmembrane region of human TLR1 is found to regulate the innate immune response indicating that the transmembrane region plays a role in function [25].

Site-specific co-incidence of adaptive codons \(d_\text{A659} > 1\) and detection of SNPs suggests that adaptive evolution within these regions is on-going. A good example is codon A659, which is identified as evolving under positive selection across the mammalian phylogeny and also three non-synonymous SNPs (2 in sheep, 1 in Bos indicus) that were found in this codon. SNPs detected at these positions in both sheep and cattle argue against A659 variation being stochastic but is likely to convey an advantage. A previous study of bovine TLR5 SNPs surmised that this site may be of functional significance (SNPdb ID: rs55617251) [26]. This finding was supported using the SIFT program which predicts the functional relevance of SNPs by comparison of conservation at that site [27]. We used the same approach to confirm this result but found that the SNP is now predicted to be a tolerated replacement. This conflict in results suggests caution when solely relying on algorithmic methods to predict putative sites of functional relevance. However, secondary structure analysis on bovine TLR5 revealed this SNP is predicted to alter the alpha helix structure (Additional file 5). This suggests that this amino acid site may indeed be of functional significance.

As expected, we found the internal signaling TLR domain to have evolved predominantly under purifying selection. In cattle a single codon (H720) was predicted as evolving subject to adaptive evolution. This positively-selected site is close to the BB-loop which is predicted to be involved in TLR dimerization and adaptor protein recognition [28].

Of the 25 breeds investigated in our study two breeds were found to have no detectable SNPs in TLR5. The Soay sheep breed was found to be homozygous in the 10 individuals analysed. The Soay breed is a primitive domestic breed which was introduced to Soay island of the St. Kilda archipelago in the Outer Hebrides of northern Scotland. This breed has experienced expansive population growth followed by seasonal periods of population crashes thought to be associated with parasitic helminth disease [29]. Our results indicate that this breed may have lost heterozygosity within TLR5. This raises the possibility that opportunistic bacterial infections may occur due to a poor TLR5 repertoire in this breed. However, sampling of a larger sample set for this breed should be carried out to verify this result.

Four cattle breeds (Kankrej, Nelore, Turkmen Zebu, Yemini Zebu) included in this study were Bos indicus subspecies and were found to contain a higher proportion of the total number of non-synonymous SNPs along with all, except two, of the total number of synonymous SNPs (Additional file 6) suggesting that the SNPs are evolutionarily recent events following the divergence of Bos indicus and Bos taurus (estimated to have diverged as early as 200,000 years ago) [30].

**Conclusions**

We have shown that in agreement with other studies, positive selection is acting on mammalian TLR5.
However, our analyses have revealed that the domesticated species in the artiodactyl clade have undergone detectable diversifying selection compared with the rest of the mammals. With the ubiquity of bacterial infection why should the ungulate clade be different to the rest of the mammals? We suggest that artificial selection may have accelerated the evolutionary process resulting in positive selective pressure driving adaptation of the TLR5 gene. The nature of this selection is not known but may be due to selection for cattle resistant to bacteria during domestication or due to an increased bacterial load due to maintaining animals in closer groups and promoting bacterial infection. The concentration of positively-selected sites in close proximity to the conserved cavity of the protein supports the hypothesis that on-going competition is a driving force shaping both the bacterial flagellin and host TLR5 genes.

Methods
Identification of positive selection
Genes encoding mammalian TLR5 were obtained from Ensembl and GenBank (Additional file 7). Translated protein sequences were aligned using Muscle [31] and back translated to obtain a codon alignment. Sequences with an in frame stop codon were removed, and 37 sequences remained (Additional file 8). Phylogenetic analysis was based on an accepted mammalian phylogeny [32] from which an unrooted tree of aligned species was created (Additional file 9). Using this tree topology branch lengths were calculated by codeML using the M0 codon model in PAML package version 4 [11]. This tree was used as the fixed tree topology for subsequent analysis. The F3 × 4 codon frequency model calculated using the nucleotide frequencies at the three codon positions was used throughout the analysis. To detect positive selection at individual codons within a gene (sites analysis) two pairs of models were compared using codeML: M1a (neutral model) was compared with M2a (adaptive model) [33,34] and M7 (beta) was compared with M8 (beta plus omega) [35]. Statistically-significant evidence of positive selection was inferred by a likelihood ratio test (LRT) comparing 2 × the log likelihood of the data (P < 0.05). Codons subject to adaptive evolution were then inferred using the Bayes empirical Bayes algorithm [42]. Codons with a posterior probability of > 95% (P0 > 0.95) of belonging to the class dNS/dSN = 1 are reported as having been subject to positive selection. All positions reported correspond to bovine TLR5 (accession: NP_001035591.1). There has been recent discussion of the appropriateness of the branch-sites test to detect positive selection particularly by Nozawa et al. [43] who suggested that the branch-sites test is particularly susceptible to false positives when the number of nucleotide substitutions is small. Nozawa et al. proposed that a minimum of 9 substitutions are required to accurately infer positive selection along a foreground branch [43]. Whilst the validity of the approach used to reach these conclusions has also been challenged [44,45], the data analyzed here does fulfill Nozawa's suggestions. Caution should also be exercised when the number of substitutions along a lineage is too great and can lead to errors in predictions. This is due to the problem of saturation of substitutions where the true number of substitutions is masked by multiple nucleotide changes at a single site. The branches of the species of interest are not considered to be a concern in this
regard. For example, the combined branch lengths of the artiodactyl clade (cow, sheep and pig) are less than one substitution per codon. Whilst the results of any inference should be subject to future validation we are confident of the appropriateness of the tests used.

Protein Structure Prediction

The domain architecture for TLR5 was determined by LRR finder [46], SMART6 [47] and TMHMM [48]. This largely follows published predictions [3]. Secondary Structure predictions were completed using PSI PRED [49]. Tertiary structure predictions were generated using Swiss-Model [50] using known crystallised TLR structures of highest similarity: extracellular domain template 2a0zA and TIR domain template 1fyvA. Positions of sites of positive selection and SNPs affecting particular amino acid positions were visualised using MolSoft ICM Browser [51].

Sequence Assembly and SNP Detection

Bovine genomic DNA (n = 110) representing 15 breeds (Additional file 10), and ovine genomic DNA (n = 87) representing 10 breeds (Additional file 11), were extracted from blood samples using standard procedures (Qiagen, UK). Samples were not obtained for the sole purpose of this project, but were either obtained (see acknowledgements) or from existing DNA banks at the Roslin Institute. The single exon encoding TLR5 was amplified using primers designed using the bovine sequence (Btau 4.0) (Additional file 12 Table S6 and Additional file 13). Amplification was performed using Taq polymerase (cattle) and KOD proof reading enzyme (sheep). Sequence reads were generated using 6 sequen- polymerase (cattle) and KOD proof reading enzyme (sheep). Sequence reads were generated using 6 sequen-

Additional files

Additional file 1: Results and parameter estimates of all single branch-sites tests.
Additional file 2: Results and parameter estimates of all multiple branch-site analysis.

Competing interests

The authors declare that they have no competing interests.

Authors’ contributions

SAS designed and conducted the molecular genetic studies, performed the statistical analysis and drafted the manuscript. OCI designed amplification primers for TLR5 and facilitated amplification and sequencing of cattle samples. OCI, DH, GCR, DW and EJG participated in the design of the study. DH and EJG conceptualised and coordinated the study. RDE participated in conceptualisation and study design, performed the statistical analysis and helped to draft the manuscript. All authors have read and approved the final manuscript.

Acknowledgements

Our gratitude is extended to the providers of samples used in this study: Albano Beja-Pereira CIBIO Portugal for the donation of cattle samples; Olivier Hanneote, University of Nottingham, and Milka Tapio, ILR, for sheep samples (Red Maasai, Dorper, Djallonké); and to both Esmeralda Minguinjo and Ramon Juste, NIKER Tecnalia, for donating Laba sheep samples. Our thanks also to ARK Genomics (The Roslin Institute, University of Edinburgh) for sequencing of cattle TLR5 genes.

Sarah Smith was funded by a University of Nottingham studentship grant awarded to David Haig. This work was supported by the Biotechnology and Biological Sciences Research Council (BBSRC) (Institute Strategic Programme Grant funding) (EJG) a Genesis Faraday (now KTN Biosciences) facilitated project grant BB/DS24040/1 jointly funded by BBSRC, The Scottish Government Rural and Environment Research and Analysis Directorate (RERAD) and Pfizer Animal Health (OCI, EJG, DW, DH).
Author details
1School of Veterinary Medicine and Science, University of Nottingham, Sutton Bonington LE12 9SD, UK. 2The Roslin Institute and Royal Dick School of Veterinary Studies, The University of Edinburgh, Easter Bush, Midlothian EH25 9RG, UK. 3Royal Veterinary College, University of London, Hawkshead Lane, Hatfield, Herts AL9 7TA, UK. 4Medicines Research Institute, Pirriants Science Park, Bush Loan, Penicuik, Midlothian EH26 0PZ, UK.

Received: 21 March 2012 Accepted: 6 July 2012

Published: 24 July 2012

References
1. Gay NJ, Gangloff M, Weber AN: Toll-like receptors as molecular switches. Nat Rev Immunol 2006, 6:693–698.
2. Welbing D, Jann OC, Oforo V, Glass C, Coffey TJ: Variation matters: TLR structure and species-specific pathogen recognition. Trends Immunol 2009, 30:124–130.
3. Masutihara T, Nakanaka T, Erkibabay P, Mikami T, Taga M, Yamada K, Kuroki Y: Comparative sequence analysis of leucine-rich repeats (LRRs) within vertebrate toll-like receptors. BMC Genomics 2007, 8:124.
4. Reid SD, Selander RK, Whitlock RH, Schukken YH, Schnabel RD, Taylor JF, Hamernik DL, Kappes SM, Lien S, et al: A maximum likelihood method for detecting adaptive evolution of mammalian ejaculate proteins. J Mol Evol 2009, 69:59–69.
5. Areal H, Abrantes J, Esteves P: Signatures of positive selection in Toll-like receptor (TLR) genes in mammas. BMC Evol Biol 2011, 11:368.
6. Andersen-Nissen E, Smith KD, Bonneau R, Strong RK, Aderem A: Functional characterization of chicken TLRs reveals species-specific recognition of flagellin. Mol Immunol 2008, 45:1298–1307.
7. Keestra AM, de Zoete MR, van Aubel RA, van Putten JP: Novel sequence diversity of flagellin (fliC) alleles in pathogenic Escherichia coli. Mol Biol Evol 2009, 26:993–999.
8. Kasiski G, Khan S, Switzer WM, Nachman MW: A history of recurrent positive selection at the toll-like receptor 5 in primates. Mol Biol Evol 2009, 26:995–1002.
9. Bancroft DR, Pemberton JM, King P: Extensive protein and microsatellite variability in an isolated, cyclic ungulate population. Heredity (Edinb) 1995, 74(Pt 3):326–336.
10. Ng PC, Henikoff S: Prediction of amino acid sites under positive selection. Nucleic Acids Res 2004, 32:6127–6132.
11. Murphy WJ, Eizikovitz E, Johnson WE, Zhang YP, Ryder OA, O’Brien SJ: Mammalian phylogenetics and the origins of placental mammals. Nature 2001, 409:614–618.
12. Nielsen R, Yang Z: Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. Genetics 1998, 148:929–936.
13. Bengtsson-Palme J, Persson K, Nyman T, Stenmark P, Flodin S, Johansson I, Hammarstrom M, Nordlund P: Structural basis of TLR5-flagellin recognition and signaling. J Biol Chem 2008, 283:11861–11865.
14. Brodbin DR, Pemberton JM, King P: Extensive protein and microsatellite variability in an isolated, cyclic ungulate population. Heredity (Edinb) 1995, 74(Pt 3):326–336.
15. Murphy WJ, Eizikovitz E, Johnson WE, Zhang YP, Ryder OA, O’Brien SJ: Mammalian phylogenetics and the origins of placental mammals. Nature 2001, 409:614–618.
16. Nielsen R, Yang Z: Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. Genetics 1998, 148:929–936.
17. Bengtsson-Palme J, Persson K, Nyman T, Stenmark P, Flodin S, Johansson I, Hammarstrom M, Nordlund P: Structural basis of TLR5-flagellin recognition and signaling. J Biol Chem 2008, 283:11861–11865.
18. Bancroft DR, Pemberton JM, King P: Extensive protein and microsatellite variability in an isolated, cyclic ungulate population. Heredity (Edinb) 1995, 74(Pt 3):326–336.
19. Murphy WJ, Eizikovitz E, Johnson WE, Zhang YP, Ryder OA, O’Brien SJ: Mammalian phylogenetics and the origins of placental mammals. Nature 2001, 409:614–618.
20. Nielsen R, Yang Z: Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. Genetics 1998, 148:929–936.
21. Bengtsson-Palme J, Persson K, Nyman T, Stenmark P, Flodin S, Johansson I, Hammarstrom M, Nordlund P: Structural basis of TLR5-flagellin recognition and signaling. J Biol Chem 2008, 283:11861–11865.
22. Bancroft DR, Pemberton JM, King P: Extensive protein and microsatellite variability in an isolated, cyclic ungulate population. Heredity (Edinb) 1995, 74(Pt 3):326–336.
45. Yang Z, Nielsen R, Goldman N: In defense of statistical methods for detecting positive selection. *Proc Natl Acad Sci U S A* 2009, 106:E95. author reply E96.

46. Offord V, Coffey TJ, Werling D: LRRfinder: a web application for the identification of leucine-rich repeats and an integrative Toll-like receptor database. *Dev Comp Immunol* 2010, 34:1035–1041.

47. Letunic I, Doerks T, Bork P: SMART 6: recent updates and new developments. *Nucleic Acids Res* 2009, 37:D229–D232.

48. Krogh A, Larsson B, von Heijne G, Sonnhammer EL: Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 2001, 305:567–580.

49. McGuffin LJ, Bryson K, Jones DT: The PSIPRED protein structure prediction server. *Bioinformatics* 2000, 16:404–405.

50. Arnold K, Bordoli L, Kopp J, Schwede T: The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* 2006, 22:195–201.

51. Abagyan R, Lee WH, Raush E, Budagyan L, Totrov M, Sundstrom M, Marsden BD: Disseminating structural genomics data to the public: from a data dump to an animated story. *Trends Biochem Sci* 2006, 31:76–78.

52. Staden R: The Staden sequence analysis package. *Mol Biotechnol* 1996, 5:233–241.

53. Tamura K, Dudley J, Nei M, Kumar S: MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 2007, 24:1596–1599.

doi:10.1186/1471-2148-12-122

Cite this article as: Smith et al: Adaptive evolution of Toll-like receptor 5 in domesticated mammals. *BMC Evolutionary Biology* 2012 12:122.