Published in final edited form as:

Mol Biol Rep. 2012 March ; 39(3): 3185–3196. doi:10.1007/s11033-011-1085-7.

The mouse B cell-specific *mb-1* gene encodes an immunoreceptor tyrosine-based activation motif (ITAM) protein that may be evolutionarily conserved in diverse species by purifying selection

Richard Sims, Virginia Oberholzer Vandergon, and Cindy S. Malone

Department of Biology, California State University Northridge, 18111 Nordhoff St., Northridge, CA 91330-8303, USA

Cindy S. Malone: cmalone@csun.edu

Abstract

The B-lymphocyte accessory molecule Ig-alpha (Ig- α) is encoded by the mouse B cell-specific gene (*mb-1*), and along with the Ig-beta (Ig- β) molecule and a membrane bound immunoglobulin (mIg) makes up the B-cell receptor (BCR). Ig- α and Ig- β form a heterodimer structure that upon antigen binding and receptor clustering primarily initiates and controls BCR intracellular signaling via a phosphorylation cascade, ultimately triggering an effector response. The signaling capacity of Ig- α is contained within its immunoreceptor tyrosine-based activation motif (ITAM), which is also a key component for intracellular signaling initiation in other immune cell-specific receptors. Although numerous studies have been devoted to the *mb-1* gene product, Ig- α , and its signaling mechanism, an evolutionary analysis of the *mb-1* gene has been lacking until now. In this study, *mb-1* coding sequences from 19 species were compared using Bayesian inference. Analysis revealed a gene phylogeny consistent with an expected species divergence pattern, clustering species from the primate order separate from lower mammals and other species. In addition, an overall comparison of non-synonymous and synonymous nucleotide mutational changes suggests that the *mb-1* gene has undergone purifying selection throughout its evolution.

Keywords

Ig-alpha (Ig- α); B cell-specific; *mb-1*; B-cell receptor (BCR); Immunoreceptor tyrosine-based activation motif (ITAM); Evolutionary divergence; Bioinformatics

Introduction

The activation of the adaptive immune system is initiated via binding of a specific antigen to a lymphocyte's surface receptor. In the case of B lymphocytes (B cells), a membrane bound immunoglobulin (mIg) serves as the antigen binding molecule capable of activating the

Correspondence to: Cindy S. Malone, cmalone@csun.edu.

Electronic supplementary material The online version of this article (doi:10.1007/s11033-011-1085-7) contains supplementary material, which is available to authorized users.

cell's immunological response with the help of cytoplasmic signaling molecules. The mIg molecule is associated with two transmembrane polypeptides called Ig- α (mouse B cell-specific gene 1, *mb-1* gene product) and Ig- β (*B29* gene product) [1, 2], which form a heterodimer and undergo intracellular conformational changes in response to antigen binding of this cell surface mIg [3, 4]. Together the mIg, Ig- α and Ig- β form the B-cell receptor (BCR) [5–10] which is essential for mature B cell and terminally differentiated plasma cell function [11].

Ig- α and Ig- β are members of the Ig superfamily and both contain an extracellular Ig-like domain, a transmembrane alpha helical region and a cytoplasmic domain [2]. The cytoplasmic domains of these polypeptides contain a conserved region known as the immunoreceptor tyrosine-based activation motif (ITAM), commonly found on the cytoplasmic tails of various immune cell receptors including Fc γ RI, Fc γ RIIA—RIIA, Fc ϵ RI, and the T-cell co-receptor CD3 subunits ζ δ ϵ γ [7, 8, 12–17]. The ITAM region is the conserved structural motif D/E(X)₇-D/E(X)₂Y(X)₂L/I(X)_{6–8}Y(X)₂L/I, where D and E are aspartic acid and glutamic acid residues, Ys are tyrosine residues, X is any amino acid, and L and I are leucine and isoleucine residues [18]. The tyrosine amino acids within this motif are sites of phosphorylation and lie within what is termed the tyrosine-based activation motif (TAM) of the ITAM consensus structure [19, 20]. Phosphorylation of the tyrosines in TAM leads to the initiation of signal transduction via receptor crosslinking and phosphate modification making ITAM signaling the primary mode of activation of immune effector cells [6].

Upon multiple antigen binding and receptor cross linking, various protein tyrosine kinases (PTK), from the Src family of kinases, including Lyn and Hck [4, 21] are activated and proceed to phosphorylate the tyrosine residues within the cytoplasmic region of the Ig- α /Ig- β heterodimer. The tyrosine residues undergo both auto-phosphorylation and trans-phosphorylation initiated by the Src homology domain 2 (SH2) of the Src kinases, and this leads to further recruitment of downstream kinases including Syk and Btk [17, 22–25]. The phosphorylation cascade induced by elevated kinase activity leads to intracellular signaling involving a rise in intracellular calcium levels and other second messenger signaling that eventually culminates in the B-cell's effector response [22–24].

The B-cell's response to antigen activation via the BCR is greatly dependent upon the developmental stage of the cell itself [2]. The effector response may include apoptosis, receptor editing, clonal expansion involving cell cycle proliferation, cell differentiation, and antigen endocytosis [26–28]. The Ig- α /Ig- β heterodimer is absolutely required for the transport of mIg to the cell surface during early B-cell development [10]. In addition, the Ig- α /Ig- β heterodimer is necessary for the developmental transition from the pro-B cell stage to the pre-B cell stage. Null mutation in one or both co-receptors leads to a block at this transitional point and causes severe B cell immunodeficiency [5].

While Ig- β is encoded by the *B29* gene, Ig- α was first discovered in the mouse genome via subtractive cloning techniques, and was found to be encoded by the *mb-1* from a pre-B cell minus T-lymphocyte [2, 29]. The human *mb-1* gene itself contains 5 exons and is located on chromosome 19 at position 19q13.2. While the *mb-1* promoter lacks a TATA box (in human

and mice), it has been shown to initiate transcription via the initiator (INR) sequence immediately 5' of the first mRNA nucleotide [30–32]. Numerous transcription factor binding sites have been studied in the *mb-1* promoter [31, 33–36], including an octamer factor-binding motif in the *mb-1* core promoter region that demonstrates significant *mb-1* transcriptional regulation [37].

The *mb-1* gene and its product have been studied extensively. Research has focused on mechanisms of gene regulation, structural and conformational analyses, as well as biochemical signaling pathways [1, 9, 29, 31, 34–46]. As yet, little attention has been placed on evolutionary conservation of the immunologically essential *mb-1* gene and its product Ig- α . In fact, analysis has been limited to comparing the *mb-1* gene's intronic rearrangement with that of the *LMP7* gene to predict a recent common ancestor gene [47]. Phylogenetic analysis using maximum parsimony and distance-based methods on the constant and variable regions of β and α T cell receptors have suggested that ancestral immune cells may have contained $\gamma\delta$ -T type receptors and may have given rise to genes encoding both T and B cell antigen binding receptors [48]. Many analyses have been performed on the antigen binding portion of the BCR, all of which were focused on the mIg molecule and all of the various isotypes post antigen stimulation and B-cell activation [49–53].

In this study, predicted cDNA sequences of the *mb-1* gene from 19 species found in the National Center for Biotechnology Information (NCBI) database and the European Molecular Biology Laboratory (EMBL) nucleotide database were aligned for a basic phylogenetic analysis. A cDNA multiple sequence alignment (MSA) was subjected to statistical analyses including model testing along with Bayesian inference and phylogenetic tree construction. In addition, the coding cDNA sequence (CDS) of the *mb-1* gene was translated and the amino acid MSA was assessed for protein domain conservation. These analyses suggest an ancestral relationship between orthologous *mb-1* coding sequences, they define selection pressures on the *mb-1* gene, and they define some rudimentary divergence profile between several species through which this gene has evolved.

Materials and methods

Multiple sequence alignment (MSA)

A general search for the *mb-1* cDNA sequences was carried out in the NCBI nucleotide database. A predicted human *mb-1* genomic sequence was obtained (accession number BC113731) along with its predicted CDS. This coding sequence was used in conjunction with a Basic Local Alignment Search Protocol (BLAST) [54–57] to obtain orthologous *mb-1* coding sequences from other species. A total of only 18 species were found in NCBI (including the human *mb-1* query sequence). One additional sequence was located in the EMBL Nucleotide Sequence Database for a total of 19 sequences. The species, accession numbers and sequence lengths obtained from the BLAST protocols are listed in Supplemental Table 1. Sequences were aligned using CLUSTAL W 3.2 and BioEdit 7.0.9.0 software [58]. CLUSTAL W is available through the Biology Workbench site: <http://workbench.sdsc.edu> [59]. BioEdit software is available from <http://www.mbio.ncsu.edu/BioEdit/bioedit.html>.

In addition, a translated cDNA MSA was performed to not only assure proper alignment but to also evaluate motif conservation among the species. This translated cDNA alignment, representing a human Ig- α (CD79A) polypeptide MSA, contains 238 positions from the N-terminus to the C-terminus. The Ig- α amino acid sequences from all 19 species were scanned using the protein MOTIF search server (<http://motif.genome.jp>) to determine conserved motifs within this protein.

Model testing

To determine the evolution model that best fit the *mb-1* dataset, model testing was performed on the cDNA alignment using the jModeltest [60–63]. Model testing included the hierarchical likelihood ratio test (hLRT), the Akaike information criterion (AIC) as well as the Bayesian information criterion (BIC) to assess a substitution model's goodness of fit to the dataset [60–63]. All three tests returned a 9 parameter general time-reversible model with a gamma distribution parameter (GTR + G) as the best-supported evolution model for the *mb-1* dataset. Supplemental Table 2 shows each parameter of the GTR model including the individual base frequencies, nucleotide substitution rates, transition/transversion bias (R) and the gamma distribution shape parameter (G), which was determined to be essentially one (0.9867). Invariable sites (I) were not detected (Supplemental Table 2).

Tree construction

Maximum likelihood (ML) along with parametric bootstrapping (5,000 replicates) [64–66] and Bayesian inference was performed using the GTR + G substitution model for tree construction. Bayesian inference was performed with Geneious 5.0 software platform (<http://www.geneious.com>) with a MrBayes 3.1.2 plugin component [62, 63, 67–71]. Three million generations were performed, generating a tree every 100 generations. The first 7,000 generations were discarded as burn-in. Sampling was performed using Markov chain Monte Carlo (MCMC) simulation algorithm for determining tree and clade posterior probabilities. In addition to ML, neighbor joining (NJ) and maximum parsimony (MP) were performed using the MEGA 4 package to re-evaluated some of the statistically supported groupings of ML. Bootstrapping [69] was also performed on NJ and MP tree topologies (5,000 and 10,000 replicates).

Protein conserved domain search

The *mb-1* coding sequences from the 19 species were translated in BioEdit while still aligned in the MSA. The Ig- α polypeptides from all species were scanned using the Prosit scan tool [72–76] to evaluate conserved domains within these amino acid sequences. Predicted conserved domains were obtained and the corresponding amino acid sequences were evaluated with visual inspection. All Prosit settings were left in default for each scan. The Prosit tool is accessible at expasy.chy/prosites.com.

dN/dS Analysis

To assess selection pressures on the *mb-1* coding sequences, non-synonymous and synonymous substitutions were determined using MEGA 4 [67, 77, 78] with pairwise comparisons between species, as well as being averaged over the entire alignment.

Substitution estimates were determined using the Nei–Gojobori methods with Jukes–Cantor. Statistical significance was determined using the Z test.

Results and discussion

mb-1 Phylogenetics

Given the manageable size of the *mb-1* dataset (19 sequences) and the results from our model testing (Supplemental Table 2), phylogeny was estimated with the statistically robust Bayesian analysis. Bayesian inference was chosen because it handles parameter-rich nucleotide substitution models and does not assume a simple model of evolution like parsimony. The tree produced by Bayesian analysis is shown in Fig. 1. The topology of the tree is what one would expect given the species involved, with the grouping of the primate order into one monophyletic clade with high Bayesian support values (Fig. 1). As expected, mouse and Norway rat clustered together with strong statistical support as did *mb-1* sequences from opossum and wallaby, and the monophyletic clade consisting of more domesticated species (horse, pig, cow and sheep). This gene tree is consistent with the divergence of classes in the animal kingdom one would expect. Interestingly, the ancestral *mb-1* sequence that gave rise to the amphibious Western Clawed frog *mb-1* gene appears to have diverged from a common ancestral sequence, and itself gave rise to the ancestral sequences of the species from the mammalian class. The lineage is displayed as in a multifurcating topology at the internal node diverging into the mammalian ancestral *mb-1* sequence, the ancestral sequence of the platypus *mb-1* gene, and the ancestral sequence of the chicken *mb-1* gene. This phylogeny, however, cannot describe the divergence of the *mb-1* gene in the chicken lineage with any significant statistical support. On the other hand, the platypus *mb-1* ancestral lineage appears to have evolved before the divergences of the other mammals in this dataset, suggesting monotremes are basal to other mammals. This could also be due to the high degree of substitution mutations and deletions in the Ig-like domain of the Ig- α in this species as demonstrated by a MSA of this region (Fig. 2). Resolution of the phylogeny in this region may be improved with sequences from other monotremes. The phylogeny did give good statistical support to the divergence of the lineages giving rise to the marsupials (wallaby and opossum) and the remainder of the placental mammals. The zebrafish was used as an outgroup in this analysis and clustered with channel catfish. Phylogeny of the *mb-1* coding sequence was also analyzed with ML and demonstrated an overall tree topology consistent with that of the Bayesian analysis, however, it was slightly less supported by bootstrap re-sampling (data not shown). Smaller bootstrap values does not under-cut the phylogeny predicted by the Bayesian inference, as these values are typically more conservative than posterior probabilities [79].

Ig- α conserved polypeptide motifs

The *mb-1* coding sequences were translated into amino acids without alteration of the original alignment to aid in accurate positioning of the base pairs in MSA and to briefly evaluate the conserved regions in the Ig- α polypeptide. The predicted polypeptides were individually scanned using a sequence motif search tool at <http://motif.genome.jp/>. This search engine integrates several protein and domain databases including Prosite, Pfam and ProDom. There were four major conserved regions of the Ig- α co-receptor. The first (Fig. 2)

is an Ig-like domain that is positioned extra-cellularly and is stabilized by disulfide bonding [24, 80]. The second (Fig. 3) is the ITAM, which aids in the B cell activation via a kinase initiated cascade of intracellular phosphorylation [6, 17]. The third (Fig. 4) is the Ig- α transmembrane domain and the fourth (Fig. 5) is the Ig- α non-ITAM cytoplasmic tail serine threonine motifs thought to negatively regulate signal transduction [81].

Ig- α conserved Ig-like domain

The Ig-like domain is defined by a region that folds by cysteine disulfide bonding. These folds are usually found in proteins that form receptors, and are named for the folds seen in immunological receptor systems [82]. Cysteine residues predicted for disulfide bonding are shown at amino acid positions 54 and 109 (Fig. 2). The Ig domain was found in all Ig- α sequences except for the platypus (Fig. 2). Interestingly, the signaling ITAM portion of the platypus is present. The platypus *mb-1* gene missing the Ig-like domain could be due to the extreme variation in the platypus genome in this particular region encoding the Ig-like portion of the protein. Multiple mutations in this region may be preventing the alignment algorithms from matching the Ig- α from the platypus with commonly known Ig-like motifs (Fig. 2). The frog sequence also had a large number of missing data in this region, however, an Ig-like region was still detected (Fig. 2). The frog Ig- α amino acid sequence only contains the C-terminal region on the Ig-like domain, which has the predicted conserved cysteine residue within the sequence LEIKNAQKNDSALY/TRCRV. These 18 amino acids may form a signature confirmation in this region that contributes to the overall cell surface structure of the Ig- α molecule. Stabilization of this structure likely involves the strictly conserved leucine residues at position 99, the tyrosine residue at position 112 in 18 of the 19 species, as well as the C-terminal cysteine residue at position 114. In the rat, tyrosine is replaced with threonine at position 112. Ig- α is believed not to contain disulfide bonding in channel catfish [83], however, the first cysteine residue is conserved at position 54 in both fish species. The Ig domain also contains a J-like motif at the C-terminal end of the Ig domain with the consensus of TXLX(V/L/I) or FGXG, where X is any amino acid residue [83, 84]. All of the mammalian species (except for platypus) contain the sequence TYLRV. The two fish species also contain the J-like motif in the form of FGXG, however, the first glycine in the catfish is replaced with threonine (FTPG), and serine in the zebrafish (FSHG). The remainder of the Ig- α Ig-like regions from the other species are fairly well conserved in the N-terminal region between the non-fish species. The conserved cysteine residues, which allow for the intra-molecular disulfide bonding of this molecule, are at positions 46 and 114. It should be noted that these positions are relative to the Ig-like region and not the entire Ig- α molecule. Disulfide bonding cysteine residues are conserved in the Ig- α and Ig- β co-receptors across different species, however, the number of cysteine residues have been shown to differ across organisms. Some of these cysteine residues form intramolecular disulfide bonds to stabilize the Ig fold of the co-receptor, and others form disulfide bridges across both the Ig- α and Ig- β to stabilize the Ig- α/β heterodimer [85]. The conserved cysteine residues depicted in Fig. 2 are predicted and likely involved in intra-molecular disulfide bonding given the equal spacing of the N-terminal and C terminal regions, which would allow for extracellular folding of Ig-like domain.

Ig- α conserved ITAM

The ITAM is a signaling molecule found in the receptors of T and B lymphocytes, as well as various other Fc receptor harboring cells in the vertebrate immune system [7, 8, 86]. The overall ITAM consensus is D/E(X)₇D/E(X)₂Y(X)₂L/I(X)₆₋₈Y(X)₂L/I where X is any amino acids, Y is tyrosine, L and I are leucine and isoleucine residues and D and E are aspartic acid and glutamic acid residues [6, 12, 18]. The kinase activating tyrosines reside within what has come to be known as the TAM region of this consensus [87]. The TAM tyrosines are precisely spaced forming a motif where the tyrosine residues are separated from a leucine or iso-leucine residue by two amino acids (YXXL/I). These tyrosine structures flank a 6–8 amino acid bridge giving the TAM a consensus of Y(X)₂L/I(X)₆₋₈Y(X)₂L/I. The TAMs are usually found in multiples and have been suggested to cause an amplifying effect when initiating intracellular signaling when in series [17]. Ig- α contains only one TAM, and along with the one possessed by its heterodimer partner Ig- β , stimulates activation of antigen-bound B cells through a series of phosphorylation of the TAM tyrosines initiated by the Src family of kinases. These kinases are located tethered to the membrane via a Src homology domain 3 (SH3) [88]. Attached to this, is a SH2 and a kinase enzyme, which is held inactive by phosphorylation of a C-terminal tyrosine within the protein [25]. Upon dephosphorylation of this terminal tyrosine, the Src kinase becomes primed to phosphorylate the tyrosine of the TAM upon a conformational shift of the co-receptors caused by antigen binding. The active Src kinase also activates other kinase families, which leads to activation and clustering of more receptors, and amplification of the intracellular signal [22].

The Ig- α TAM domain is strongly conserved in all species (Fig. 3). Again, the Ig- α protein MSA was a direct translation of the *mb-1* cDNA MSA and was unchanged after conversion. Figure 3 demonstrates a strictly conserved Ig- α TAM among the mammalian Ig- α sequences. Thirteen of the 15 amino acid positions are identical in the Ig- α TAM region in the mammal species (Fig. 3). Upon omission of the infraclass Marsupialia, the remaining mammals are 100% identical in their Ig- α TAM amino acid sequence (Fig. 3). Within the non-placental mammals, 14 of 15 amino acid positions are identical (Fig. 3) with an alanine to serine substitution at position 191 in the wallaby and opossum and a conserved aspartate to glutamate substitution at position 188 in the platypus Ig- α TAM region. The substitution in the platypus's Ig- α TAM sequence is likely of little consequence in terms of kinase docking and phosphorylation. Both aspartate and glutamate share similar polar properties, which would presumably lead to similar protein folding patterns within the Ig- α protein. Despite the apparent absence of the Ig-like domain in the platypus Ig- α protein, the platypus Ig- α TAM appears to be functional. Immunological studies in the platypus have demonstrated mitogen induced B-cell proliferation [89], as well as B-cell intracellular signaling activation [90, 91], both of which require Ig- α /Ig- β heterodimers with functional ITAMs [92, 93]. The remainder of the amino acids in the Ig- α TAM region are exactly the same across all species tested in the mammalian class (YEGLNLDDCSMYEDI) [6]. The frog Ig- α TAM amino acid sequences diverged from mammals at three locations. In addition to the position 188 conserved substitution, a glutamate for the mammalian aspartate like the platypus, the frog substitutes glutamate for aspartate at position 189 as well (Fig. 3). Additionally, the frog substitutes threonine for the mammalian methionine at position 192 (Fig. 3). The chicken Ig- α TAM diverges from the other species, substituting an aspartate

for an asparagine at position 186. Both positions 188 and 189 diverge as well with position 188 substituting an alanine in place of an aspartate residue and position 189 substituting a glutamine for the mammalian aspartates in the Ig- α TAM (Fig. 3). The most divergent species in our lineup of the Ig- α TAM are the catfish and zebrafish having 4 residue changes (Fig. 3). The entire Ig- α ITAM region in catfish and zebrafish is one amino acid longer than that of the mammals with an asparagine residue residing between positions 190 and 191 of the Ig- α TAM. Both catfish and zebrafish also replaced methionine found in mammals with threonine and alanine at position 192. Positions 194 and 195 in the fish species replace glutamate and aspartate with histidine and glutamine (Fig. 3). The 11 amino acids making up the negative charged amino acid domain, consensus D/E(X)₇D/E(X)₂, directly proximal to the TAM portion of the ITAM consist of mainly negatively charged amino acids aspartate and or glutamate at positions 171, 177, 178, and 179. Positions 171, 177, 178, 179, 180 and 181 are highly conserved among most species. Position 171 contains the acidic amino acid aspartate or glutamate in all species except chicken and frog, while positions 177, 178, and 179 contains aspartate or glutamate in all species except the frog 177 isoleucine and 178 glycine. Asparagine is conserved in all species at position 180. At position 181, leucine is conserved in all mammalian species but is substituted with isoleucine in the frog and fish species. All the mammals contain an aspartate at position 174. The D174 site is substituted with a threonine residue in the frog and zebrafish, a methionine residue in catfish, and a serine residue in chicken (Fig. 3). Position 176 contains a conserved tyrosine residue found in all mammalian species. This Y176 in the negatively charged amino acid domain of the ITAM in the Ig- α polypeptide has been shown to be involved with recruitment of a B cell linker (BLNK) protein during the kinase driven intracellular signaling in B cells [90]. This tyrosine is not seen in the chicken, frog or fish species, but instead is replaced by a glycine residue in the chicken, a glutamine residue in the frog, and arginine residues in both fish species (Fig. 3).

Ig- α conserved transmembrane domain

Ig- α is known to contain a transmembrane domain and to exist as an amphipathic protein in B cells [29]. The transmembrane region of the Ig- α molecule shows 79% identity and 92% conservation across the placental mammalian amino acid sequences with only five positions (8%) having substitutions that are not conserved. Two of these positions are towards the N-terminal portion (positions 10 and 11) and the other three are at the C-terminal portion (positions 55, 57, 62) of the Ig- α transmembrane domain (Fig. 4). Adding the wallaby and the opossum to the alignment only drops the identity calculation to 76% but does not affect the 92% conservation. The platypus Ig- α transmembrane domain sequences are highly divergent from the rest of the mammal species studied here and were not considered in any of the calculations of identity or conservation (Fig. 4). Inclusion of the chicken sequences to the non-monotreme mammals results in a 52% identity of residues and a 76% conservation of residues among species studied here. The frog and the fish species further diverge from the mammals, but there are clearly areas within the Ig- α transmembrane domain that show identity and conservation of amino acids, particularly in the hydrophobic area between the N-terminal and C-terminal regions of the domain (Fig. 4).

Ig- α conserved cytoplasmic tail

Figure 5 shows a portion of the cytoplasmic tail of Ig- α distal and proximal to the TAM signaling region. This region contains two conserved tyrosine residues (Y176 and Y204, Fig. 5) flanking the two tyrosine residues of the TAM region (Y182 and Y193, Fig. 5) that are crucial for the downstream B cell intracellular signaling initiated by Src kinases [94]. The Src activated tyrosine in the TAM (Y182 and Y193, Fig. 5) recruits an active Syk tyrosine kinase via the B cell linker adaptor protein BLNK [95, 96]. BLNK is recruited and binds to a phosphorylated tyrosine (Y204, Fig. 5) located on the C-terminal side of the Ig- α TAM and a non-phosphorylated tyrosine (Y176, Fig. 5) located on the N-terminal side of the TAM [94]. Mutation of Y176 and Y204 disrupts the downstream BLNK dependent signaling pathways of Syk kinase [94]. It is interesting that Y176 is identical in the mammalian species including platypus, but divergent in the frog, chicken and fish species (Fig. 5). This may suggest that the divergent species do not use the BLNK dependent pathway of B cell activation, but may employ some other means of relaying downstream intracellular signaling [94]. Y176 also lies within the negatively charged amino acid consensus of the ITAM but is proximal to its kinase activating TAM portion. This suggests the negatively charged amino acid consensus plays some role in the TAM driven intracellular signaling in mammals, perhaps through stabilization of Y176 [94].

Serine and threonine phosphorylation within and just downstream of the Ig- α TAM, along with the well known importance of tyrosine phosphorylation, have been shown to regulate Ig- α TAM signal transduction. Mutation of the two serines (S191 and S197, Fig. 5) and single threonine (T203, Fig. 5) increases signal transduction from the BCR, suggesting that these residues negatively regulate the Ig- α ITAM under normal circumstances [81]. In our analysis, S191 is 89% (17/19) identical, S197 is 84% (16/19) identical, and T203 is 89% (17/19) identical among all species with the placental mammals showing 100% identity at all three sites (Fig. 5). Interestingly, the divergences among these three residues do not occur in all of the same species. Specifically, the divergence in S191 occurs in the wallaby and opossum, but not the platypus, and results in a substitution of alanine for serine. At S197, the conserved substitution occurs in the frog as a threonine, and in the fish species as a glutamine substitution. The divergence in T203 occurs in the fish species where proline replaces threonine (Fig. 5). It seems clear that these changes must affect the negative regulation of signaling in the fish species because they contain only one of the three sites (S191, Fig. 5). Even at S197 where glutamine replaces serine (Fig. 5), it has not been shown that glutamine is a site of phosphorylation and even if it were, it might not respond in the same way. P203 is substituted for T203 in the fish, and has also not been shown to be a site of phosphorylation. In both the wallaby and the opossum, only S191 is replaced with alanine, but both S197 and T203 remain identical (Fig. 5). It is not known whether these two sites alone are competent for negative signaling [81].

Selective pressures on Ig- α sequences

In order to determine what kind of selective pressures, if any, have been acting on Ig- α over evolutionary time, we analyzed our 19 sequences for mutational changes at the DNA level. In this analysis, mutations in nucleotides that result in silent mutations, or no change in amino acid, are calculated against nucleotide mutations that result in missense mutations, or

mutations that do change the resulting amino acid. The number of missense mutations, or non-synonymous mutations per non-synonymous sites (dN), are divided by the number of silent, or synonymous mutations per synonymous sites (dS), giving the dN/dS ratio [97–99]. In essence, this analysis determines whether positive evolutionary selection has occurred to allow change in the protein over time or whether purifying (negative) evolutionary selection has occurred to have the protein remain the same over time. Neutral selection suggests that neither positive nor negative selection occurred and mutations happened randomly throughout the gene in about equal number at synonymous versus nonsynonymous sites [97–99]. We evaluated the selective pressure on the Ig- α co-receptor by estimating dN and dS over all sequence pairs (Supplemental Table 3). A dN/dS ratio (ω) of 0.336 was calculated using the Nei–Gojobori method with Jukes–Cantor for the entire *mb-1* gene. The estimated average number of synonymous sites were 63, while the average number of non-synonymous sites were 186. Bootstrapping and the Z-test were used to determine statistical significance ($P < 0.03$). The 0.336 ω value indicates that a purifying selection mechanism may have acted on the Ig- α coding sequence during its evolution. In addition to the tests run on the entire alignment of *mb-1* sequences, both the Ig-like domain and the ITAM region were tested individually using the same techniques. The ITAM region defined as the D/E(X)₇D/E(X)₂Y(X)₂L/I(X)_{6–8}Y(X)₂L/I consensus (Fig. 3), and the varied Ig-like domain (Fig. 2) both demonstrated purifying selection with an estimated ω of 0.156 ($P < 0.03$) for the 26 site ITAM sequence and an ω of 0.798 for the Ig-like domain. Given the variability of the Ig-like domain, it is possible that this was subjected to more mutations over time and purifying selective mechanisms may be largely responsible for maintaining this variable region without introducing mutations that would hinder the function of Ig- α . However, an ω value of the Ig-like domain may indicate that both positive and negative selection may have begun to equilibrate as a dN/dS ratio of essentially 0.8 nears neutrality. In fact, analysis did not demonstrate statistical support to reject neutral selection on this domain (Supplemental Table 3). In general, DNA mutations are usually more deleterious than beneficial over time, and negative (purifying) selection is important for eliminating these deleterious changes and maintaining the stability of the molecule [100]. Amino acids in the Ig-like region are far less conserved than in the ITAM region, suggesting this region was subjected to a larger amount of mutation during the evolution of Ig- α . These changes probably paralleled those seen in immunoglobulin variable gene cassettes. However, unlike the immunoglobulin genes, *mb-1* does not encode for an antigen binding molecule. Nevertheless, nature has likely fixed the Ig- α molecule into a stable and usable structure through evolution, and it appears that selective pressures have acted to stabilize this protein's structure in the presence of multiple mutational events directed to enhance variability in the vertebrate immune system. In the case of the ITAM coding region of the *mb-1* gene, strict conservation of the TAM region has been maintained, as this region contains the functional region of the Ig- α polypeptide. The negatively charged amino acid region of the ITAM is strongly conserved among mammalian species, and to a lesser extent in fish, frog, and chicken. Besides the involvement of BLNK recruitment by Y176 in kinase-driven intracellular B cell signaling, how this region contributes to the overall signaling of the ITAM is less understood and remains to be elucidated. Given the large number of negatively charged amino acids in the region, particularly aspartate and glutamate, this region of the ITAM may be simply involved in stabilization of the TAM domain [94]. In any event, the ITAM structure as defined in Fig. 3

demonstrates purifying selection as does the whole Ig- α coding sequence. This suggests that despite subjection to multiple mutations evolution maintained the entire Ig- α sequence, likely due to the importance in maintaining immune system function.

Summary

This study demonstrates a baseline phylogeny of the *mb-1* coding sequence containing 19 orthologs. Despite the limited sample size, a reasonable portion of vertebrate classes were represented including, fish, aves, amphibians, and mammals. In addition, multiple of species were included within the mammalian class ranging from lower monotreme mammals to primates. Phylogenetic inference using Bayesian analysis provided an evolutionary relationship for the *mb-1* gene using 19 species that had good statistical support for the major branching monophyletic clades including the divergence of the frog from the mammals, divergence of the marsupials from the placental mammals, and divergence of the primates from the remainder of the mammals. This gene tree is consistent with the accepted mode of evolution of the species, and the ancestral states of the *mb-1* gene itself may have evolved parallel to the expected divergence of these classes. In addition, comparing synonymous to non-synonymous mutational changes across all 19 species suggested that *mb-1* evolved under overall purifying selection pressures, most likely attributed to its ITAM region. This study is the first to reveal the possible divergence pathway(s) the *mb-1* orthologs analyzed have taken through evolution. Of course, as the coding sequences from more species become available, this phylogeny will be better resolved. Additional species from the fish, amphibian and reptilian classes will be especially useful in better defining the divergence of this gene.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Torres RM, Hafen K. A negative regulatory role for Ig- α during B cell development. *Immunity*. 1999; 11(5):527–536. [PubMed: 10591178]
2. Burrows PD, et al. The transient expression of pre-B cell receptors governs B cell development. *Semin Immunol*. 2002; 14(5):343–349. [PubMed: 12220935]
3. Flaswinkel H, et al. Signaling and internalisation function of the B cell antigen receptor complex. *Adv Exp Med Biol*. 1994; 365:1–8. [PubMed: 7887295]
4. Zhang Y, et al. Induction of the antigen receptor expression on B lymphocytes results in rapid competence for signaling of SLP-65 and Syk. *EMBO J*. 1998; 17(24):7304–7310. [PubMed: 9857187]
5. Gong JH, et al. RANTES and MCP-3 antagonists bind multiple chemokine receptors. *J Biol Chem*. 1996; 271(18):10521–10527. [PubMed: 8631850]
6. Cambier JC. Antigen and Fc receptor signaling. The awesome power of the immunoreceptor tyrosine-based activation motif (ITAM). *J Immunol*. 1995; 155(7):3281–3285. [PubMed: 7561018]
7. Reth M. Antigen receptors on B lymphocytes. *Annu Rev Immunol*. 1992; 10:97–121. [PubMed: 1591006]
8. Reth M. B cell antigen receptors. *Curr Opin Immunol*. 1994; 6(1):3–8. [PubMed: 8172677]
9. Roth PE, DeFranco AL. Receptor tails unlock developmental checkpoints for B lymphocytes. *Science*. 1996; 272(5269):1752–1754. [PubMed: 8650572]

10. DeFranco AL. The two-headed antigen. B-cell co-receptors. *Curr Biol.* 1996; 6(5):548–550. [PubMed: 8805272]
11. Castro I, et al. B cell receptor-mediated sustained c-Rel activation facilitates late transitional B cell survival through control of B cell activating factor receptor and NF-kappaB2. *J Immunol.* 2009; 182(12):7729–7737. [PubMed: 19494297]
12. Siraganian RP. Mast cell signal transduction from the high-affinity IgE receptor. *Curr Opin Immunol.* 2003; 15(6):639–646. [PubMed: 14630197]
13. Stachowski J, et al. Signalling via the TCR/CD3 antigen receptor complex in uremia is limited by the receptors number. *Nephron.* 1993; 64(3):369–375. [PubMed: 8101978]
14. Ravetch JV, Bolland S. IgG Fc receptors. *Annu Rev Immunol.* 2001; 19:275–290. [PubMed: 11244038]
15. Call ME, Wucherpennig KW. Molecular mechanisms for the assembly of the T cell receptor-CD3 complex. *Mol Immunol.* 2004; 40(18):1295–1305. [PubMed: 15072848]
16. Alarcon B, et al. The CD3-gamma and CD3-delta subunits of the T cell antigen receptor can be expressed within distinct functional TCR/CD3 complexes. *EMBO J.* 1991; 10(4):903–912. [PubMed: 1826255]
17. Underhill DM, Goodridge HS. The many faces of ITAMs. *Trends Immunol.* 2007; 28(2):66–73. [PubMed: 17197236]
18. Reth M. Antigen receptor tail clue. *Nature.* 1989; 338(6214):383–384. [PubMed: 2927501]
19. Daeron M, et al. Tyrosine-containing activation motif-dependent phagocytosis in mast cells. *J Immunol.* 1994; 152(2):783–792. [PubMed: 8283051]
20. Flaswinkel H, Reth M. Dual role of the tyrosine activation motif of the Ig-alpha protein during signal transduction via the B cell antigen receptor. *EMBO J.* 1994; 13(1):83–89. [PubMed: 8306975]
21. Chan VW, et al. Characterization of the B lymphocyte populations in Lyn-deficient mice and the role of Lyn in signal initiation and down-regulation. *Immunity.* 1997; 7(1):69–81. [PubMed: 9252121]
22. Clevers HC, et al. Synergistic action of A23187 and phorbol ester on human B cell activation. *J Immunol.* 1985; 135(6):3827–3830. [PubMed: 3934268]
23. Choquet D, et al. Cross-linking of IgG receptors inhibits membrane immunoglobulin-stimulated calcium influx in B lymphocytes. *J Cell Biol.* 1993; 121(2):355–363. [PubMed: 8468351]
24. Kirkham PM, Schroeder HW Jr. Antibody structure and the evolution of immunoglobulin V gene segments. *Semin Immunol.* 1994; 6(6):347–360. [PubMed: 7654992]
25. Liu X, et al. Regulation of c-Src tyrosine kinase activity by the Src SH2 domain. *Oncogene.* 1993; 8(5):1119–1126. [PubMed: 7683128]
26. Rolink A, et al. Two pathways of B-lymphocyte development in mouse bone marrow and the roles of surrogate L chain in this development. *Immunol Rev.* 1994; 137:185–201. [PubMed: 8034335]
27. Sater RA, Sandel PC, Monroe JG. B cell receptor-induced apoptosis in primary transitional murine B cells: signaling requirements and modulation by T cell help. *Int Immunol.* 1998; 10(11):1673–1682. [PubMed: 9846696]
28. Neuberger MS, Scott J. Immunology. RNA editing AIDs antibody diversification? *Science.* 2000; 289(5485):1705–1706. [PubMed: 11001738]
29. Sakaguchi N, et al. B lymphocyte lineage-restricted expression of mb-1, a gene with CD3-like structural properties. *EMBO J.* 1988; 7(11):3457–3464. [PubMed: 2463161]
30. Quon DV, et al. Similarity between a ubiquitous promoter element in an ancient eukaryote and mammalian initiator elements. *Proc Natl Acad Sci USA.* 1994; 91(10):4579–4583. [PubMed: 8183951]
31. Travis A, Hagman J, Grosschedl R. Heterogeneously initiated transcription from the pre-B- and B-cell-specific mb-1 promoter: analysis of the requirement for upstream factor-binding sites and initiation site sequences. *Mol Cell Biol.* 1991; 11(11):5756–5766. [PubMed: 1922076]
32. Javahery R, et al. DNA sequence requirements for transcriptional initiator activity in mammalian cells. *Mol Cell Biol.* 1994; 14(1):116–127. [PubMed: 8264580]

33. Fitzsimmons D, et al. Pax-5 (BSAP) recruits Ets proto-oncogene family proteins to form functional ternary complexes on a B-cell-specific promoter. *Genes Dev.* 1996; 10(17):2198–2211. [PubMed: 8804314]
34. Feldhaus AL, et al. BLyF, a novel cell-type- and stage-specific regulator of the B-lymphocyte gene mb-1. *Mol Cell Biol.* 1992; 12(3):1126–1133. [PubMed: 1545794]
35. Hagman J, Travis A, Grosschedl R. A novel lineage-specific nuclear factor regulates mb-1 gene transcription at the early stages of B cell differentiation. *EMBO J.* 1991; 10(11):3409–3417. [PubMed: 1915300]
36. Nutt SL, et al. Identification of BSAP (Pax-5) target genes in early B-cell development by loss- and gain-of-function experiments. *EMBO J.* 1998; 17(8):2319–2333. [PubMed: 9545244]
37. Malone CS, Patrone L, Wall R. An essential octamer motif in the mb-1 (Igalpha) promoter. *Mol Immunol.* 2000; 37(6):321–328. [PubMed: 11000405]
38. Antony P, et al. B-cell antigen receptor activates transcription factors NFAT (nuclear factor of activated T-cells) and NF-kappaB (nuclear factor kappaB) via a mechanism that involves diacylglycerol. *Biochem Soc Trans.* 2004; 32(Pt 1):113–115. [PubMed: 14748726]
39. Campbell KS, et al. IgM antigen receptor complex contains phosphoprotein products of B29 and mb-1 genes. *Proc Natl Acad Sci USA.* 1991; 88(9):3982–3986. [PubMed: 2023945]
40. Maier H, et al. Requirements for selective recruitment of Ets proteins and activation of mb-1/Ig-alpha gene transcription by Pax-5 (BSAP). *Nucleic Acids Res.* 2003; 31(19):5483–5489. [PubMed: 14500810]
41. Maier H, et al. Activation of the early B-cell-specific mb-1 (Ig-alpha) gene by Pax-5 is dependent on an unmethylated Ets binding site. *Mol Cell Biol.* 2003; 23(6):1946–1960. [PubMed: 12612069]
42. Malone CS, Wall R. Bob1 (OCA-B/OBF-1) differential transactivation of the B cell-specific B29 (Ig beta) and mb-1 (Ig alpha) promoters. *J Immunol.* 2002; 168(7):3369–3375. [PubMed: 11907094]
43. Minegishi Y, et al. Mutations in Igalpha (CD79a) result in a complete block in B-cell development. *J Clin Invest.* 1999; 104(8):1115–1121. [PubMed: 10525050]
44. Nomura J, et al. Signal transmission through the B cell-specific MB-1 molecule at the pre-B cell stage. *Int Immunol.* 1991; 3(2):117–126. [PubMed: 1709046]
45. van Noesel CJ, et al. The membrane IgM-associated heterodimer on human B cells is a newly defined B cell antigen that contains the protein product of the mb-1 gene. *J Immunol.* 1991; 146(11):3881–3888. [PubMed: 2033258]
46. Zhang M, Srivastava G, Lu L. The pre-B cell receptor and its function during B cell development. *Cell Mol Immunol.* 2004; 1(2):89–94. [PubMed: 16212894]
47. Abdulla S, et al. Divergent intron arrangement in the MB1/LMP7 proteasome gene pair. *Immunogenetics.* 1996; 44(4):254–258. [PubMed: 8753855]
48. Nelson KE, et al. Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of *Thermotoga maritima*. *Nature.* 1999; 399(6734):323–329. [PubMed: 10360571]
49. Dishaw LJ, et al. The basis for haplotype complexity in VCBPs, an immune-type receptor in amphioxus. *Immunogenetics.* 2010; 62(9):623–631. [PubMed: 20652563]
50. Cannon JP, et al. Variable domains and a VpreB-like molecule are present in a jawless vertebrate. *Immunogenetics.* 2005; 56(12):924–929. [PubMed: 15650874]
51. Litman GW, Cannon JP, Rast JP. New insights into alternative mechanisms of immune receptor diversification. *Adv Immunol.* 2005; 87:209–236. [PubMed: 16102575]
52. Litman GW, Cannon JP, Dishaw LJ. Reconstructing immune phylogeny: new perspectives. *Nat Rev Immunol.* 2005; 5(11):866–879. [PubMed: 16261174]
53. Zhang SM, et al. Diversification of Ig superfamily genes in an invertebrate. *Science.* 2004; 305(5681):251–254. [PubMed: 15247481]
54. Altschul SF, et al. Basic local alignment search tool. *J Mol Biol.* 1990; 215(3):403–410. [PubMed: 2231712]
55. Altschul SF, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 1997; 25(17):3389–3402. [PubMed: 9254694]

56. Morgulis A, et al. Database indexing for production MegaBLAST searches. *Bioinformatics*. 2008; 24(16):1757–1764. [PubMed: 18567917]
57. Zhang J, Madden TL. PowerBLAST: a new network BLAST application for interactive or automated sequence analysis and annotation. *Genome Res*. 1997; 7(6):649–656. [PubMed: 9199938]
58. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*. 1994; 22(22):4673–4680. [PubMed: 7984417]
59. Subramaniam S. The Biology Workbench—a seamless database and analysis environment for the biologist. *Proteins*. 1998; 32(1):1–2. [PubMed: 9672036]
60. Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol*. 2003; 52(5):696–704. [PubMed: 14530136]
61. Posada D, Crandall KA. MODELTEST: testing the model of DNA substitution. *Bioinformatics*. 1998; 14(9):817–818. [PubMed: 9918953]
62. Posada D, Buckley TR. Model selection and model averaging in phylogenetics: advantages of akaike information criterion and bayesian approaches over likelihood ratio tests. *Syst Biol*. 2004; 53(5):793–808. [PubMed: 15545256]
63. Posada D. jModelTest: phylogenetic model averaging. *Mol Biol Evol*. 2008; 25(7):1253–1256. [PubMed: 18397919]
64. Steel M, Penny D. Parsimony, likelihood, and the role of models in molecular phylogenetics. *Mol Biol Evol*. 2000; 17(6):839–850. [PubMed: 10833190]
65. Gadagkar SR, Kumar S. Maximum likelihood outperforms maximum parsimony even when evolutionary rates are heterotactous. *Mol Biol Evol*. 2005; 22(11):2139–2141. [PubMed: 16014870]
66. Felsenstein J. Evolutionary trees from DNA sequences: a maximum likelihood approach. *J Mol Evol*. 1981; 17(6):368–376. [PubMed: 7288891]
67. Tamura A, et al. Re-evaluation and functional classification of non-synonymous single nucleotide polymorphisms of the human ATP-binding cassette transporter ABCG2. *Cancer Sci*. 2007; 98(2): 231–239. [PubMed: 17297656]
68. Huelsenbeck JP, Ronquist F. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*. 2001; 17(8):754–755. [PubMed: 11524383]
69. Alfaro ME, Zoller S, Lutzoni F. Bayes or bootstrap? A simulation study comparing the performance of Bayesian Markov chain Monte Carlo sampling and bootstrapping in assessing phylogenetic confidence. *Mol Biol Evol*. 2003; 20(2):255–266. [PubMed: 12598693]
70. Bollback JP. Bayesian model adequacy and choice in phylogenetics. *Mol Biol Evol*. 2002; 19(7): 1171–1180. [PubMed: 12082136]
71. Huelsenbeck JP, et al. Potential applications and pitfalls of Bayesian inference of phylogeny. *Syst Biol*. 2002; 51(5):673–688. [PubMed: 12396583]
72. de Castro E, et al. ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Res*. 2006; 34:W362–W365. Web Server issue. [PubMed: 16845026]
73. Hulo N, et al. Recent improvements to the PROSITE database. *Nucleic Acids Res*. 2004; 32:D134–D137. Database issue. [PubMed: 14681377]
74. Hulo N, et al. The PROSITE database. *Nucleic Acids Res*. 2006; 34:D227–D230. Database issue. [PubMed: 16381852]
75. Hulo N, et al. The 20 years of PROSITE. *Nucleic Acids Res*. 2008; 36:D245–D249. Database issue. [PubMed: 18003654]
76. Sigrist CJ, et al. PROSITE, a protein domain database for functional characterization and annotation. *Nucleic Acids Res*. 2010; 38:161–166. Database issue.
77. Nei M, Gojobori T. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol Biol Evol*. 1986; 3(5):418–426. [PubMed: 3444411]
78. Gojobori T, Nei M. Relative contributions of germline gene variation and somatic mutation to immunoglobulin diversity in the mouse. *Mol Biol Evol*. 1986; 3(2):156–167. [PubMed: 3444398]

79. Yang Z, Rannala B. Bayesian phylogenetic inference using DNA sequences: a Markov Chain Monte Carlo Method. *Mol Biol Evol.* 1997; 14(7):717–724. [PubMed: 9214744]
80. Thies MJ, et al. Folding and oxidation of the antibody domain C(H)3. *J Mol Biol.* 2002; 319(5): 1267–1277. [PubMed: 12079363]
81. Muller R, Wienands J, Reth M. The serine and threonine residues in the Ig-alpha cytoplasmic tail negatively regulate immunoreceptor tyrosine-based activation motif-mediated signal transduction. *Proc Natl Acad Sci USA.* 2000; 97(15):8451–8454. [PubMed: 10900006]
82. Bork P, Holm L, Sander C. The immunoglobulin fold. Structural classification, sequence patterns and common core. *J Mol Biol.* 1994; 242(4):309–320. [PubMed: 7932691]
83. Sahoo M, et al. B cell receptor accessory molecules in the channel catfish, *Ictalurus punctatus*. *Dev Comp Immunol.* 2008; 32(11):1385–1397. [PubMed: 18572245]
84. Desai S, et al. The medaka novel immune-type receptor NITR gene clusters reveal an extraordinary degree of divergence in variable domains. *BMC Evol Biol.* 2008; 8:177. [PubMed: 18565225]
85. Siegers GM, et al. Identification of disulfide bonds in the Ig-alpha/Ig-beta component of the B cell antigen receptor using the *Drosophila* S2 cell reconstitution system. *Int Immunol.* 2006; 18(9): 1385–1396. [PubMed: 16877534]
86. Muller B, Cooper L, Terhorst C. Cloning and sequencing of the cDNA encoding the human homologue of the murine immunoglobulin-associated protein B29. *Eur J Immunol.* 1992; 22(6): 1621–1625. [PubMed: 1534761]
87. Samelson LE, Klausner RD. Tyrosine kinases and tyrosine-based activation motifs. Current research on activation via the T cell antigen receptor. *J Biol Chem.* 1992; 267(35):24913–24916. [PubMed: 1459994]
88. Mayer BJ, Baltimore D. Mutagenic analysis of the roles of SH2 and SH3 domains in regulation of the Abl tyrosine kinase. *Mol Cell Biol.* 1994; 14(5):2883–2894. [PubMed: 8164650]
89. Stewart NJ, et al. Mitogen-induced responses in lymphocytes from platypus, the Tasmanian devil and the eastern barred bandicoot. *Aust Vet J.* 2008; 86(10):408–413. [PubMed: 18826514]
90. Wilkinson GW, Akrigg A. Constitutive and enhanced expression from the CMV major IE promoter in a defective adenovirus vector. *Nucleic Acids Res.* 1992; 20(9):2233–2239. [PubMed: 1317548]
91. Parra G, et al. Tandem chimerism as a means to increase protein complexity in the human genome. *Genome Res.* 2006; 16(1):37–44. [PubMed: 16344564]
92. Isakov N. Immunoreceptor tyrosine-based activation motif (ITAM), a unique module linking antigen and Fc receptors to their signaling cascades. *J Leukoc Biol.* 1997; 61(1):6–16. [PubMed: 9000531]
93. Isakov N. ITIMs and ITAMs. The Yin and Yang of antigen and Fc receptor-linked signaling machinery. *Immunol Res.* 1997; 16(1):85–100. [PubMed: 9048210]
94. Kabak S, et al. The direct recruitment of BLNK to immunoglobulin alpha couples the B-cell antigen receptor to distal signaling pathways. *Mol Cell Biol.* 2002; 22(8):2524–2535. [PubMed: 11909947]
95. Cheng AM, et al. Syk tyrosine kinase required for mouse viability and B-cell development. *Nature.* 1995; 378(6554):303–306. [PubMed: 7477353]
96. Fu JM, Castelhana AL. Design and synthesis of a pyri-done-based phosphotyrosine mimetic. *Bioorg Med Chem Lett.* 1998; 8(19):2813–2816. [PubMed: 9873628]
97. Kryazhimskiy S, Plotkin JB. The population genetics of dN/dS. *PLoS Genet.* 2008; 4(12):e1000304. [PubMed: 19081788]
98. Kryazhimskiy S, Bazykin GA, Dushoff J. Natural selection for nucleotide usage at synonymous and nonsynonymous sites in influenza A virus genes. *J Virol.* 2008; 82(10):4938–4945. [PubMed: 18321967]
99. Hurst LD. The Ka/Ks ratio: diagnosing the form of sequence evolution. *Trends Genet.* 2002; 18(9): 486. [PubMed: 12175810]
100. Hoffmann R, et al. Neutrality, compensation, and negative selection during evolution of B-cell development transcriptomes. *Mol Biol Evol.* 2007; 24(12):2610–2618. [PubMed: 17890238]

Fig. 1.

Bayesian inference phylogenetic tree of the *mb-1* gene. Nineteen species were analyzed by Bayesian inference and a phylogenetic tree was produced that demonstrates posterior probabilities shown between 0 and 1, with equaling 100% agreement. The analysis was performed using the Geneious 5.0 software platform and the MrBayes 3.1.2 plugin. Three million generations were performed with a tree produced every 100 generations. The first 7,000 generations were discarded as burn-in and LnL of the one run was -8013.13. Tree branch lengths are in proportion to estimated evolutionary distances. *Scale bar* represent units of genetic distance (0.2 mutations per site)

Fig. 2.

MSA of the amino acids in the Ig-like domains of the Ig-α co-receptor. The Ig-α Ig-like N-terminal domain is present in all species except the frog and platypus, while the C-terminal Ig-like domain is only missing in the platypus. The N-terminal domain spans residues 17–51 and the C-terminal domain spans residues 99–116 and are *boxed*. The disulfide bonding identical cysteine residues are located at positions 46 and 114 and are shown highlighted yellow and *boxed*. Ig-α polypeptide sequences were scanned with MOTIF search server (<http://motif.genome.jpz>) and aligned with CLUSTAL W3.2 and BioEdit version 7.0.9.0

Fig. 3.

MSA of the amino acids in the ITAM region of the Ig- α co-receptor. The Ig- α ITAM domain consensus sequence, D/E(X)₇D/E(X)₂Y(X)₂L/I(X)₆₋₈Y(X)₂L/I, is highly conserved across the 19 species analyzed. The ITAM domain consists of the TAM and a negatively charged amino acid domain as indicated by *boxes*. Missense mutations substituted with conserved amino acids are highlighted in *yellow* while nonconserved missense mutations are highlighted in *pink*. Ig- α polypeptide sequences were scanned with MOTIF search server (<http://motif.genome.jpz>) and aligned with CLUSTAL W3.2 and BioEdit version 7.0.9.0

Fig. 4.

MSA of the amino acids in the transmembrane region of the Ig-α co-receptor. The Ig-α transmembrane domain shows high conservation across the placental mammals shown *boxed in red*. The marsupial animals are shown *boxed in orange* and the addition of these species retains the high 92% conservation seen in the placental mammals. Only five positions contain unconserved amino acid substitutions and are *boxed in pink* (10, 11 55, 57, and 62). Ig-α polypeptide sequences were scanned with MOTIF search server (<http://motif.genome.jp>) and aligned with CLUSTAL W3.2 and BioEdit version 7.0.9.0

Fig. 5.

MSA of the amino acids in the cytoplasmic tail of the Ig- α co-receptor. The Ig- α cytoplasmic tail shows high conservation across tyrosine, serine, and threonine residues, and contains the ITAM domain. The TAM region, YXX(L/I)X6-12YXX(L/I), of the Ig- α ITAM is highly conserved across the 19 species analyzed and is *boxed in black*. This TAM region is involved in the primary kinase signaling through phosphorylation of Y182 and Y193, *boxed in yellow*. Two conserved tyrosine residues flanking the TAM region, Y176 within the ITAM and Y204 downstream of the ITAM are *boxed in red*, and are crucial for the downstream B cell intracellular signaling initiated by the ITAM and Src kinases. Serine and threonine phosphorylation within and just downstream of the TAM region of the ITAM, S191, S197, and T203, have been shown to regulate Ig- α ITAM signal transduction and are *boxed in pink*. Ig- α polypeptide sequences were scanned with MOTIF search server (<http://motif.genome.jp>) and aligned with CLUSTAL W3.2 and BioEdit version 7.0.9.0