International Journal of Global Science Research **ISSN: 2348-8344 (Online)**
Vol. 13, Issue. 1, April 2026, pp. 8969-9001 **DOI: https://doi.org/10.26540/ijgsr.v13.i1.2026.337**
***Available Online* at www.ijgsr.com**

Research Paper

Multiple sequence alingment and phylogenetic analysis of variants of MYD88 an adapter protein of immune mechanism

Madhulika Rout[1], Hitesh Prasad Sahoo[2] and Hongray Howrelia Patnaik[1*]
[1]Post Graduate Department of Zoology, Ravenshaw University, Cuttack, Odisha, India.
[2]School of Life Sciences, Jawaharlal Nehru University, New Delhi, India.
*****Corresponding author email: hhowrelia.patnaik@gmail.com

Received: 22/02/2026
Revised: 27/02/2026
Accepted: 08/03/2026

Abstract: Bio-informatics encompasses a multitude of domains, including comparative genomics, proteomics and systems biology. Comparative genomics entails the examination of genomic features across diverse organisms, with the objective of elucidating evolutionary relationships and functional conservation. This endeavour frequently necessitates methodologies for genome alignment and the identification of conserved sequences. This discipline employs mathematical modelling and simulations to forecast how alterations in one component can affect the entire system. With the advent of high-throughput technologies, researchers now possess access to a vast reservoir of genomic, transcriptomic, proteomic, and metabolomics data. Integrative methodologies, which synthesize these disparate data sources, facilitate the revelation of novel biological insights. The innate immune system serves as the primary line of defence for an organism against invading pathogens. Invertebrates, lacking adaptive immunity, rely exclusively on their innate immune mechanisms to combat pathogenic microorganisms. This system is activated by an array of pattern recognition receptors (PRRs), which are adapt at identifying pathogen-associated molecular patterns (PAMPs). The adaptor protein MyD88 sequences from *Homo sapiens, Mus musculus, Drosophila melanogaster, Paenibacillus amylolyticus, and Homo sapiens IL-1* collectively present a captivating array of downloaded (FASTAQ) sequences, amenable to phylogenetic tree analysis and Multiple Sequence Analysis (MSA). It provides profound insights into the structural and organizational nuances of the MyD88 (Myeloid Differentiation primary adapter protein of the innate immune mechanism) protein sequences across the five species, accentuating their similarities. Moreover, this analysis aspires to identify conserved regions among the quintet of species and reconstruct their phylogenetic relationships predicated on sequence data, thereby inferring the evolutionary trajectory of the protein.

International Journal of Global Science Research
Vol. 13, Issue. 1, April 2026, pp. 8969-9001
Available Online at www.ijgsr.com

ISSN: 2348-8344 (Online)
DOI: https://doi.org/10.26540/ijgsr.v13.i1.2026.337

Keywords: Bio-informatics, Innate Immunity, Adaptor Protein MYd88, Multiple sequence Alignment, Phylogenetic Relations.

Abbreviations: MSA: Multiple sequence analysis, **DNA**: Deoxy ribonucleic acid, **MYD88**: Myeloid differentiation factor 88, **RNA**: Ribonucleic acid, **FASTA**: FASTA-ALL, **NCBI**: National center for biotechnology information, **BLAST**: Basic local alignment search tool, **EMBL**: Europe Molecular Biology Laboratory, **EBI**: Europe bio-informatics institute, **Uniprot**: Universal protein sequences, **DmMyd88**: Drosophila melanogaster myeloid differentiation primary response 88, **TLR**: Toll like Receptor , **HGNC**: HUGO Gene Nomenclature Committee, **RPKM**: Reads Per Kilobase per Million mapped reads, **IL-1**: Interleukin-1**, IRAK**: Interleukin-1 receptor-associated kinase 1, **IRF**: Interferon Regulatory Factor, **TRAF**: Tumor Necrosis Factor Receptor-Associated Factor, **NF**: Nuclear factor, **IFN:** Interferons, **NOS**: Nitric Oxide Synthase, **INOS**: Inducible Nitric oxide synthase, **NLRP3**: NOD-like receptor family pyrin domain-containing 3, **REG3G**: Regenerating islets-derived protein 3-gamma, **GOR**: Garnier–Osguthorpe–Robson method, **MEGA**: Molecular Evolutionary Genetics Analysis, **APE**: Analysis of Phylogenetics and Evolution, **GNU**: General Public License, **CLUSTAL W**: Clustal-Weight

Introduction:

Molecular biology plays a pivotal role in understanding nucleotides and proteins, which are fundamental to the existence of life on Earth. This domain of biology remains a vast frontier with much yet to be explored. However, the research conducted thus far has significantly transformed our perspective as the world continues to evolve. Sequencing, which has now become the cornerstone of numerous scientific endeavors and investigations, has illuminated areas previously shrouded in obscurity. In recent years, the human genome has been meticulously sequenced and completed. Similarly, scientists have also delved into the genomic sequences of various other organisms. (Khan et al., 2023, Khan et al., 1988)

Moreover, the advancements in molecular biology have led to innovations such as CRISPR technology, which allows for precise genetic editing. This is not only holds the potential for correcting genetic defects but also opens doors to new therapeutic strategies against various diseases, including cancer and hereditary disorders. The application of bio-informatics has also greatly enhanced our understanding of gene interactions and functions. By employing computational tools, scientists can analyze vast amounts of genetic data, allowing them to identify patterns and make predictions about biological processes on a scale previously deemed impossible. As we continue to unveil the mysteries of life at the molecular level, interdisciplinary collaboration among biologists, chemists, and computer scientists will be crucial in addressing the complex challenges that lie ahead (Smith et al., 2024). This ongoing exploration is expected to revolutionize not only our understanding of biology but also the development of novel bio-technologies that could tackle global health and environmental issues.

The esteemed methodology of shotgun sequencing has significantly advanced the genomic analysis of numerous organisms, including *Haemophilus influenzae* in 1995, *Mycoplasma genitalium, Mycobacterium tuberculosis,* and more recently, *Yersinia pestis*. The vast troves of sequenced data amassed from diverse organisms necessitate meticulous analysis and systematic organization. This burgeoning

International Journal of Global Science Research
Vol. 13, Issue. 1, April 2026, pp. 8969-9001
Available Online at www.ijgsr.com

ISSN: 2348-8344 (Online)
DOI: https://doi.org/10.26540/ijgsr.v13.i1.2026.337

reservoir of information presents a formidable challenge in terms of storage and accessibility. Such requirements can be adeptly addressed through the field of bio-informatics. Paulien Hogeweg and Ben Hesper coined the term 'bio-informatics' in 1970, where 'bio' denotes biology and 'informatics' pertains to the systematic understanding and organization of information essentially, the elucidation of biological phenomena in a computational framework. This discipline serves as a computational tool that facilitates the efficient interpretation of extensive sequences within a condensed time frame. bio-informatics has significantly contributed to the storage and accessibility of vast datasets, allowing for ease of access from any location around the globe at any given moment.

Characterized as an interdisciplinary field, bio-informatics amalgamates elements of statistics, biology, mathematics, and physics (Luscombe and Greenbaum 2000). It can be regarded as an advanced iteration of biology. The discipline has adeptly integrated contemporary software, as the advent of digitalization has made it readily available; with adequate internet connectivity and computational resources, data can be analyzed seamlessly at any time and from anywhere (Bayat 2002).

Bio-informatics encompasses three primary objectives that are achieved through its application. The first objective primarily involves accessing information and subsequently augmenting the system with new data. However, mere submission of data is insufficient; comprehensive analysis is imperative (Khan et al., 2023). This necessity leads us to the second objective of bio-informatics, wherein various tools have been devised to facilitate data analysis. Bio-informatics not only stores vast volumes of information but also aids in its analysis through the utilization of these sophisticated tools. Finally, the third and ultimate objective is to employ the available tools to analyze the data and interpret it effectively. In this context, bio-informatics primarily evaluates the data on a global scale.

The data intended for processing cannot be effectively managed by a singular tool thus, there exist specialized instruments within bio-informatics tailored for distinct functions. This application necessitates the utilization of both the internet and computational resources. Numerous tools are available, each categorized according to its specific functionality (Bayat 2002).

As the volume of data expands, it becomes imperative to archive the older information while seamlessly incorporating new entries. The system should facilitate effortless modifications and ensure easy accessibility. The National Center for Biotechnology Information (NCBI) serves as a global repository for genetic data, allowing for the storage of information gathered from any corner of the globe. Analogous to NCBI, other applications such as Ensemble can also be utilized for similar purposes.

An application known as EMBL-EBI represents a globally recognized initiative by the European Bio-informatics Institute, offering a comprehensive array of freely accessible and up-to-date molecular data. Each data set is assigned a unique identifier in the form of an accession number, which serves as its address within the repository. Distinct applications are available for various bio molecular categories, including proteins and nucleotides. For proteomic data, resources such as UniProt and PRIDE are utilized, while ChEBI caters to chemical entities. For literature reviews and ontological studies, applications like Clexplore and Gene Ontology (GO) are employed. Additionally, there are tools designed to analyze protein families, motifs, and domains, such as InterPro, as well as

International Journal of Global Science Research
Vol. 13, Issue. 1, April 2026, pp. 8969-9001
Available Online at www.ijgsr.com

ISSN: 2348-8344 (Online)
DOI: https://doi.org/10.26540/ijgsr.v13.i1.2026.337

applications for visualizing protein structures, exemplified by PDBe. Reactome serves as a tool for elucidating protein pathways. SWISS-PROT is also a notable resource, comprising a vital protein database that includes sequence data from a multitude of organisms.

Bio-informatics has revolutionized the analysis of multiple organisms, offering a precise and time-efficient approach. It can be employed to formulate target-specific therapeutics for various diseases. Moreover, it aids in the identification of genetic defects through the utilization of extensive databases. There exists a promising future perspective wherein bio-informatics will further contribute to the development of treatments for ailments that currently lack cures. Additionally, it proves invaluable in elucidating the evolutionary relationships among organisms, facilitating this process by highlighting analogous and conserved regions across species (Sandhya et al., 2024). It delineates the localization of genes within specific chromosomes and elucidates the functionality of each region, thereby providing a comprehensive understanding of genetic architecture. Furthermore, it assists in tracing the origins of various related genes by establishing phylogenetic relationships among the extant genes of contemporary organisms.

The primary significance of bio-informatics lies in its ability to compile extensive datasets and analyze them, rendering this information accessible. It facilitates the expedited detection of diseases, thereby making it considerably more cost-effective than traditional methods. The conventional methodologies in biology have been supplanted by computational techniques. Wet lab techniques have been overshadowed by bio-informatics, which conserves time, resources, and energy. This field has notably advanced molecular docking processes, enabling predictions regarding reactivity. Furthermore, it assesses the toxicity of pharmaceuticals without the necessity for testing. In critical situations, it has proven invaluable by saving precious time.

Multiple sequence analysis is a fundamental technique wherein multiple genes undergo comparative examination based on their structural and functional attributes. In the present study, we have conducted a comprehensive analysis of the Multiple Sequence Alignment (MSA) of the MyD88 gene, an adapter protein integral to the Toll pathway of the innate immune response. The MSA was performed among five distinct organisms, ranging from lower to higher vertebrates, alongside an exploration of their phylogenetic relationships, employing advanced bio-informatics tools.

The innate immune system is the first line of defense of an organism against invading pathogens. To cope with pathogenic microorganisms, invertebrates rely solely on their innate immune system because they do not have adaptive immunity (Buchmann, 2014). The innate immune system is triggered by various pattern recognition receptors (PRRs) that can recognize pathogen-associated molecular patterns (PAMPs) (Akira and Hemmi; 2003). Toll-like receptors (TLRs), as one of the major types of PRRs, play a key role in innate immunity by recognizing PAMPs, such as lipopolysaccharides (LPS), peptidoglycans, polyinosinic cytidylic acid, β-glycan of fungi, and lipoproteins of various pathogens (Uematsu and Akira; 2008). Myeloid differentiation factor 88 (MyD88) is a key adapter protein of the TLR signal pathway, and it mediates all the signal pathways except TLR3. After the recognition of PAMPs by TLRs, MyD88 could recruit TLR through one of its conserved domains, toll/interleukin-1

International Journal of Global Science Research
Vol. 13, Issue. 1, April 2026, pp. 8969-9001
Available Online at www.ijgsr.com

ISSN: 2348-8344 (Online)
DOI: https://doi.org/10.26540/ijgsr.v13.i1.2026.337

receptor (TIR). And then, MyD88 uses its other conserved death domain (DD) to associate with the DD of interleukin-1 receptor-associated kinase, and triggers the activation of nuclear factor-kappa B (NF-κB) and mitogen-activated protein kinase pathways (Uematsu et al., 1994).

MyD88 genes are ubiquitous and conserved in the animal kingdom. MyD88-mediated signaling pathways play crucial roles in the immune response of invertebrates. Their connections with pathogen challenge (LPS or virus) were also investigated (Ishii et al 2006). Since the first discovery of MyD88 in 1990, MyD88 gens have been identified in many species, such as *Danio rerio* (Jault et al., 2019), *Crassostrea gigas* (Wang et al., 2013), *Chlamys farreri* (Qiu et al., 2013), *Ruditapes philippinarum* (Y. Lee., 2011), *Scylla paramamosain* **(**X.-C. Li et al., 2013), and *Rana dybowskii* (Niu et al., 2018). In contrast with insects and vertebrates that only have one copy of MyD88 (Bonnert et al., 1997), (A.M. van der Sar et al., 2015), many MyD88 genes have been identified in mollusks. A total of five MyD88 genes were found in the genome of Yesso scallop (*Patinopecten yessoensis*) [X. Ning et al., 2015], 10 were found in *C. gigas* (Zhang et al., 2015), and three were identified in Mediterranean mussel (*Mytilus galloprovincialis*) [M. Toubiana et al., 2013]. Two MyD88 duplications (HcMyD88-1 and HcMyD88-2) were found in the transcriptome of triangle-shell pearl mussel (*Hyriopsis cumingii*) (Ren et al., 2014). These reports suggest that many copies of MyD88 genes exist in the genomes of mollusks, but their evolution is still unclear. An increasing number of genomes of mollusks have been sequenced and released, thereby providing the basis for detailed analysis of MyD88 genes at the genomic level.

Sequence alignment is an active research area in the field of bio-informatics. It is also a crucial task as it guides many other tasks like phylogenetic analysis, function, and/or structure prediction of biological macromolecules like DNA, RNA, and Protein. Proteins are the building blocks of every living organism. Although protein alignment problem has been studied for several decades, unfortunately, every available method produces alignment results differently for a single alignment problem. Multiple sequence alignment is characterized as a very high computational complex problem. Many stochastic methods, therefore, are considered for improving the accuracy of alignment. Among them, many researchers frequently use Genetic Algorithm (Chowdhury and Garai, 2017).

Different types of the method applied in alignment and the recent trends in the multi objective genetic algorithm for solving multiple sequence alignment. MSA is more advantageous than PSA as it considers multiple members of a sequence family and thus provides more biological information. MSA is also (Morgenstern et al., 2006) a prerequisite to comparative genomic analyses for identification and quantification of conserved regions or functional motifs in a whole sequence family, estimation of evolutionary divergence between sequences and even for ancestral sequence profiling (Kumar and Filipski, 2007). Sequence alignment at amino acid level is more relevant than nucleotide level as protein is the key functional biological molecules and hence carries structural and/or functional information (Morgenstern et al., 2012).

The alignment, thus, has a strong connection to structural biology as well (Xiong, 2006). Hence, the SA, specifically MSA is the starting point of any biological macro molecular research field where MSA acts as an open window for viewing evolutionary, functional, and structural perspective of biological macro molecules

International Journal of Global Science Research
Vol. 13, Issue. 1, April 2026, pp. 8969-9001
Available Online at www.ijgsr.com

ISSN: 2348-8344 (Online)
DOI: https://doi.org/10.26540/ijgsr.v13.i1.2026.337

in a concise format. Different scoring methods are used in the sequence alignment to know the level of identity or similarity. Nucleotide scoring is a simple identification scheme where identical bases in both sequences are assigned positive scores. In contrast, for protein, a similarity score is also being counted (along with identity score) denoting the amino acids having similar physio-chemical properties. The substitution matrices mostly consulted for protein sequence alignment are Point Accepted Mutation (PAM), and Blocked Substitution Matrix (BLOSUM) (Henikoff and Henikoff, 1994). The method of SA can be of two types: global alignment and local alignment. Global alignment is done when the similarity is counted over the entire length of the sequences. Several MSA techniques accomplish global alignment but difficulties arise when sequences are only homologous over local regions where clear block of ungapped alignment common to all of the sequences or if there is presence of shuffled domains among the related sequences (Heringa and Taylor, 2006).

In such cases, local alignment is performed to know the local similar regions among the sequences. When there is a large difference in the lengths of the sequences to be compared, local alignment is generally performed (Higgins and Sharp, 1988). Three categories of approaches are frequently used in MSA namely, exact, progressive, and iterative alignment approaches. Exact algorithms usually deliver high quality alignment that very close to optimal. It tries to simultaneously align multiple sequences and thus need to depend on DP. Due to the drawback of the exact method in alignment as stated above, most MSA follows other two categories.

Progressive Alignment

Hogeweg and Hesper 1984 first formulated Progressive Alignment. Progressive is a heuristics approach where complex MSA problem is separated into sub problems. This solves direct MSA problem indirectly with PSA. This approach assembles all sequences progressively where best pairwise alignment is first taken into account. Progressive alignment uses guide tree (Huang, 1994) to solve MSA problem where each leaf represents a sequence to be aligned. Each visited internal node is associated with an MSA of the sequences in its corresponding sub-tree. Finally, MSA of all considered sequences is associated with the root node (Fig. 1). Progressive alignment technique is used in several alignment programs such as MULTAL (Taylor and Taylor, 1987, 1988), MAP (Corpet, 1988), PCMA (Pei et al., 2003, 2006), MULTALIGN (Higgins and Sharp, 1988), CLUSTAL (Higgins and Sharp, 1988, Thompson et al., 1994), T-Coffee (C.Notredame, 2003), KAlign Lassmann and Sonnhammer; 2005), MUMMALS/PROMALS (Pei and Grishin, 2006; 2007), and others. Among them, the most widely used method is ClustalW (Thompson et al., 1994). It first performs the global pairwise alignment (Sokal and Michener, 1958) of the sequences and develops a distance matrix. It then builds a guide tree based on the matrix values. Finally, it generates a consensus alignment by gradually adding sequences following the guide tree where the closest sequence pairs (smallest branch length in guide tree) are aligned first and thus, it gradually adds the next sequences. However, the greedy nature of these approaches cannot allow to modify the gaps and hence, the alignment cannot be altered in the later stage. There is a possibility that they can be trapped in local minima for this greediness. Fig. 1 shows that ClustalW failed to get the optimum MSA due to its greedy nature in modifying the positions of gaps thus trapped in a

International Journal of Global Science Research
Vol. 13, Issue. 1, April 2026, pp. 8969-9001
Available Online at www.ijgsr.com

ISSN: 2348-8344 (Online)
DOI: https://doi.org/10.26540/ijgsr.v13.i1.2026.337

local optimal alignment. Another drawback is that any progressive MSA is influenced by the initial alignment. As a result, any error made at that stage is propagated to the final MSA results. On the other hand, the iterative alignment methods iteratively modify the alignment by realigning the sequences or sequence groups and thus, overcome the drawback of the progressive method.

This is an example of how a progressive alignment performs MSA. The alignment structure and the guide tree are constructed by ClustalW. Each node in the guide tree is associated with an alignment. Based on the pairwise distance measures, s1 and s2 are aligned first because of smallest distance pairs (smallest branch length) and later s3 is added to them. Finally, the root node associates all the sequences considered in the MSA. The alignment shows that ClustalW failed to optimize the alignment of 'CAT 'region of s2 with other sequences' 'CAT' regions.

Guide tree

Guide tree guides the merging order of sequences based on the pairwise distances calculated for all the possible sequence pairs to be aligned in MSA (Fig. 1). For guide tree construction, UPGMA (Stoye, 1988) or neighbor-joining method is applied. However, the guide tree causes error in progressive alignment if an error is introduced while measuring the distances or at the time of tree construction. Ultimately, the error is reflected in the final alignment. The problem can be solved by iterative methods by repetitively modifying the guide tree and calculating the distance measurement. Hence, Iterative methods are biologically more sound approaches. Iterative approach

The iterative method generally performs post-processing by making changes in the alignment made by progressive methods. It modifies the construction of guide tree. Programs that use guide tree reestimation are MAFFT (Katoh, et al., 2002; 2005), MUSCLE (Stoye, 1988), PRIME (Yamada, et al., 2005), PRRP (Edgar, 2004) and MUMMALS/PROMALS. They compute new distance matrices using an MSA obtained by progressive alignment.

Hidden Markov model-based alignments

One of the popular statistical models like Hidden Markov model is also taken into account for sequence alignment. One of the popular methods is ProbCons (Yamada, et al., 2005), which uses a new scoring function based on probabilistic consistency. It is also a progressive approach that uses a combination of probabilistic modeling and consistency-based alignment techniques. Therefore, by incorporating multiple pairwise sequence conservation information along with probabilistic consistency model, probcons able to produce improved results compared to other consistent methods. MUMMALS extends the ProbCons approach to allow for more sophisticated HMM structures.

Pruning technique

To reduce the search space of MSA solution, some methods use pruning techniques. The principle of Divide and Conquer Algorithm (DCA) program is to divide and concatenate the MSA. This approach first identifies the "optimal cut" points using pairwise projected alignments for partitioning a large multiple alignment into smaller sub problems. Each of the small parts is aligned separately and then joins to produce the final MSA. Similarly, some MSA programs (Katoh, et al., 2002; 2005) uses the Carrillo Lipman bound technique to determine constraints of an optimal multiple alignment.

Alignment scoring technique

The sequence alignment techniques quantitatively measure the quality of an

International Journal of Global Science Research
Vol. 13, Issue. 1, April 2026, pp. 8969-9001
Available Online at www.ijgsr.com

ISSN: 2348-8344 (Online)
DOI: https://doi.org/10.26540/ijgsr.v13.i1.2026.337

alignment by considering a scoring model. The most commonly used scoring model chosen by several MSA methods is Sum of Pair (SOP)

K-mer based distance estimation

K-mer based distance estimation is an alternative to estimating evolutionary distance based on multiple alignments. At a high level, the distance between two sequences is defined by first collecting the set of k-mers (subsequences of length k) occurring in the two sequences. From these two sets, the evolutionary distance between the two organisms is now defined by measuring how different the two sets are. The more the two sets look alike, the smaller is the evolutionary distance. The main motivation for estimating evolutionary distance based on k-mers, is that it is computationally much faster than first constructing a multiple alignment. Experiments show that phylogenetic tree reconstruction using k-mer based distances can produce results comparable to the slower multiple alignment based methods (Blaisdell, 1989).

All of the k-mer based distance measures completely ignores the ordering of the k-mers inside the input sequences. Hence, if the selected k value (the length of the sequences) is too small, very distantly related organisms may be assigned a small evolutionary distance (in the extreme case where k is, two organisms will be treated as being identical if the frequency of each nucleotide/amino-acid is the same in the two corresponding sequences). In the other extreme, the k-mers should have a length (k) that is somewhat below the average distance between mismatches if the input sequences were aligned (in the extreme case of k the length of the sequences, two organisms have a maximum distance if they are not identical). Thus, the selected k value should not be too large and not too small. A general rule of thumb is to only use k-mer based distance estimation for organisms that are not too distantly related.

Phylogenetic Relationship:

The calculations for construction of phylogenetic trees can by distance matrix or from discrete character data In the first calculation, data based on evolutionary distances are set in a distance matrix. Most calculation methods do not weight each nucleotide mutation equally. The DNA structure plays an important role in the calculation procedures. It has been postulated that transversions are more easily recognized by the DNA repair system than are transitions because of the spherical DNA helix distortions (Kimura, 1980). These changes are therefore considered to be less frequent and result in a lower substitution rate, which can be taken into account when calculating distance values.

The MyD88 (Myeloid Differentiation Factor 88) protein sequence of *Homo sapiens,*
Mus musculus, Drosophila melanogaster, Paenibacillus amylolyticus, and *Homo sapiens* interleukin-1 (IL-1) exhibits remarkable conservation across species. These sequences encode instructions for synthesizing a protein integral to signaling within immune cells. The MyD88 protein functions as an adaptor, facilitating the connection between proteins that detect extracellular signals and those that transmit these signals intracellularly. Specifically, MyD88 is pivotal in relaying signals from particular proteins known as Toll-like receptors and interleukin-1 (IL-1) receptors, which play a crucial role in initiating an immune response against foreign invaders, such as bacteria. The MyD88 sequences from *Homo sapiens, Mus musculus, Drosophila melanogaster, Paenibacillus amylolyticus*, and *Homo sapiens* IL-1 collectively present a captivating array of downloaded (FASTAQ) sequences, suitable for

International Journal of Global Science Research
Vol. 13, Issue. 1, April 2026, pp. 8969-9001
Available Online at www.ijgsr.com

ISSN: 2348-8344 (Online)
DOI: https://doi.org/10.26540/ijgsr.v13.i1.2026.337

phylogenetic tree analysis and Multiple Sequence Analysis (MSA).

The present study elucidated the evolutionary relationships among the species examined. The Multiple Sequence Alignment (MSA) and phylogenetic analysis offered profound insights into the structure and organization of the Myd88 (Myeloid Differentiation primary adapter protein of the innate immune mechanism) protein sequences across the five species, highlighting any similarities. Furthermore, this analysis aimed to identify conserved regions among the five species and reconstruct their phylogenetic relationships based on sequence data, thereby inferring the evolutionary history of the protein.

Material and Methods:

The protein sequences were downloaded for the present study to find out the Multiple Sequence Alignment (MSA) and phylogenetic relationship among Myd88 an innate immune adapter gene of *Mus musculus, Drosophila melanogaste*r, *Paenibacillus amylolyticus*, *Homo sapiens* with the help of certain bio-informatics tool. These tools provide a great aid to collect data at ease. These are sometimes also called as Gene banks as the huge stored information of different types of gene like DNA, RNA and proteins. The tools used to collect data are mentioned below:

NCBI:

The National Center for Biotechnology Information (NCBI) serves as an invaluable bio-informatics resource utilized in experimental research. It stands as a national repository for molecular biology information. The NCBI website comprises several distinct sections, including a dedicated area for the submission of gene data, which is meticulously archived for future reference. Users can also access this stored data for subsequent applications. Furthermore, the platform is organized into segregated sections for Genome, Gene, Nucleotide, and Protein, where information is systematically cataloged.

Additionally, NCBI provides a suite of analytical tools, including software such as BLAST (Robert C. Edgar and Serafim Batzoglou, 2006), facilitating in-depth data analysis. A comprehensive literature section is available for educational purposes, enriching users' understanding of genetic information. Each gene is assigned a unique identifier, enabling direct searches that yield specific gene data. The information is formatted in FASTA, a widely used sequence format. Utilizing this tool, various datasets have been compiled from the protein section, further enhancing the depth of research capabilities.www.ncbi.nlm.nih.gov. After downloading the data, we proceeded to the subsequent tool.

EMBL:

Europe Molecular Biology Laboratory (EMBL) collaborates with the European bio-informatics Institute (EBI) to aggregate gene data along with the associated literature. In this study, EMBL, recognized as a sophisticated bio-informatics tool, has been employed to facilitate access to comprehensive gene information. Each gene is assigned a unique identifier, known as an accession number, which serves as a distinctive reference within this platform. This accession number enables streamlined access to the entirety of the gene collection. For research purposes, we procured protein data in FASTA format. The protein sequences acquired were also seamlessly integrated with UniProt software, enhancing the visualization of the collected protein sequences. This tool encompasses a global gene bank that permits both the upload and download of data, thereby fostering a collaborative

International Journal of Global Science Research
Vol. 13, Issue. 1, April 2026, pp. 8969-9001
Available Online at www.ijgsr.com

ISSN: 2348-8344 (Online)
DOI: https://doi.org/10.26540/ijgsr.v13.i1.2026.337

environment for data exchange. www.ebi.ac.uk

Uniprot:

The subsequent tool employed was UniProt, a comprehensive and freely accessible resource for protein data and structural information. Within this platform, protein sequences are presented in FASTA format. Additionally, users can access detailed 3D representations of protein structures, alongside pertinent information regarding their chromosomal locations and structural analyses. This tool is widely regarded as indispensable for visualizing the structures of collected protein sequences. For our materials, data encompassing protein structures and their associated characteristics were meticulously gathered. Here, the accession number serves as a unique identifier for protein sequences, having established a collaboration with EMBL. UniProt offers extensive information regarding protein structures, their interactions, and the evolutionary relationships among them. https://www.uniprot.org/

Once the data pertaining to various organisms has been meticulously gathered, it is subsequently organized into a taxonomic hierarchy, accompanied by detailed descriptions of the protein sequences. The organisms under consideration include *Drosophila melanogaster, Homo sapiens myd88, Homo sapiens IL-1 receptor, Mus musculus,* and *Paenibacillus amylolyticus.*

GOR method:

In this experiment this method was used to predict the secondary structures of protein by the Garnier–Osguthorpe–Robson method (GOR). The protein sequences where uploaded one by one to analyse. This tool in bio-informatics is best for protein structure analysis. It shows number of alpha helixes; beta sheets and turns in a structure. The information carried can be used to make a 3D structure of protein http://cib.cf.ocha.ac.jp/bitool/GOR/

MEGA:

MEGA or Molecular Evolutionary Genetics Analysis is a computer program that is used in this experiment to calculate the distance between genes and the phylogenetic relationship along with analyzing and making evolutionary trees for better analysis. Here after the analysis of the structure of protein these sequences were uploaded in the software to understand the phylogenetic relationship between them. It has tools for sequence alignment, estimating evolutionary distances, and testing evolutionary hypothesis. Its user friendly and its comprehensive features make it a valuable resource for researchers in genetic, evolutionary biology and computational biology. https://www.megasoftware.net/

MEGA, is a computer program that is used to calculate the distance between genes and the phylogenetic relationship along with analyzing and making evolutionary tree for better understanding of the given sequences .https://www.megasoftware.net/

R-PACKAGE is Analysis of Phylogenetics and Evolution (APE) is a package written in the R language for use in molecular evolution and phylogenetics. APE provides both utility functions for reading and writing data and manipulating phylogenetic trees, as well as several advanced methods for phylogenetic and evolutionary analysis (e.g. comparative and population genetic methods). APE takes advantage of the many R functions for statistics and graphics, and also provides a flexible framework for developing and implementing further statistical methods for the analysis of evolutionary processes (Paradis et al., 2004).

Availability: The program is free and available from the official R package

International Journal of Global Science Research
Vol. 13, Issue. 1, April 2026, pp. 8969-9001
Available Online at www.ijgsr.com

ISSN: 2348-8344 (Online)
DOI: https://doi.org/10.26540/ijgsr.v13.i1.2026.337

archive at http://cran.r project.org/src/contrib/ PACKAGES. html#ape. APE is licensed under the GNU General Public License.

CLUSTRAL W:

It is an application used for multiple sequence alignment. Here protein sequences of different organism were aligned with the help of this tool. It is a widely used bio-informatics tool. It aligns sequences by considering sequences similarities, gap between them and substituted sites to produce an accurate and meaningful alignment. Clustral W uses progressive alignment approach, staring with pairwise alignment to create a guide tree which is then further becomes useful to align sequence. The major benefits of Clustral W include high accuracy while aligning homologous sequences, it also can handle large datasets, it also helps to study the conserved region across sequences. https://www.genome.jp/tools-bin/clustalw

List of ClustalW Servers availability around the world:

- EBI Europe www.ebi.ac.uk/clustalw
- EMBnet Europe www.ch.embnet.org/software/ClustalW.html
- PIR USA pir.georgetown.edu/pirwww/search/multialn.shtml
- BCM USA searchlauncher.bcm.tmc.edu/multi-align/multi-align.html
- GenomeNet Japan align.genome.jp/
- DDBJ Japan www.ddbj.nig.ac.jp/search/clustalw-e.html
- Strasbourg Europe ftp://ftp-igbmc.u-strasbg.fr/pub/ClustalW/

Result:

Multiple Sequence Alignment (MSA) is acknowledged as an exceedingly computationally intricate endeavor. This vital instrument supports fundamental operations such as phylogenetic analysis, along side the assessment of function and structure for diverse biological macro-molecules, including DNA, RNA, and proteins. Various methodologies invariably produce differing alignment outcomes for a singular alignment scenario. MSA is esteemed for its capability to bolster the accuracy of these alignments.

The outcomes derived from this study utilizing Multiple Sequence Alignment signify a noteworthy advancement in the pursuit of alignment accuracy. The analytical process may encompass a spectrum of activities, ranging from the parallel measurement of samples to the minimization of manual effort, employing specialized software for quality assurance, or simply generating visual representations using widely-used programming languages such as Python or R. The Analysis of Phylogenetics and Evolution (APE) is a sophisticated package developed in the R programming language, tailored for applications in molecular evolution and phylogenetics. APE offers a suite of utility functions designed for the reading and writing of

data, as well as for the manipulation of phylogenetic trees. Furthermore, it encompasses several advanced methodologies for phylogenetic and evolutionary analysis, including comparative and population genetic approaches. APE leverages the extensive array of R functions dedicated to statistical analysis and graphical representation, while also providing a versatile framework for the development and implementation of additional statistical methodologies pertinent to the analysis of evolutionary processes.

The objective of this analysis (Figure 8) is to formulate hypotheses that will subsequently be validated, ultimately leading to novel insights regarding sequence similarities in relation to

International Journal of Global Science Research **ISSN: 2348-8344 (Online)**
Vol. 13, Issue. 1, April 2026, pp. 8969-9001 **DOI: https://doi.org/10.26540/ijgsr.v13.i1.2026.337**
***Available Online* at www.ijgsr.com**

biological activity, cellular responses, metabolic processes, and the growth and development of the five selected species: *Homo sapiens*, Myd88, *Mus musculus*, *Drosophila melanogaster, Paenibacillus amylolyticus,* and *Homo sapiens* of IL1.

The sequence data acquired from NCBI and EMBL-EBI underwent a meticulous series of methodological steps, including Multiple Sequence Alignment (MSA) and the construction of a phylogenetic tree employing the 'R' Package, CLUSTALW, and the BLOOSM substitution

matrix. This comprehensive analysis encompassed the species *A. trivirgatus, Oryctolagus cuniculus, Rattus norvegicus, Homo sapiens,* and *Mus musculus.* Phylogenetic tree plot was generated with APE 'R' Package Multiple Sequence Alignment view plot was generated with MSA 'R'Package. Downloaded nucleotide sequence of *Homo sapiens,* (myd88 and IL1) *Mus musculus Drosophila melanogaster, Paenibacillus amylolyticus.*

In multiple sequence alignment, different colors are used, which represents a specific characteristic. The red color denotes the widely conserved amino acid residues of the aligned sequence, whereas the blue color shows (acidic) the poorly conserved amino acid domains.

Dot (.) This symbol represents a position in the alignment that is not conserved but is similar. It indicates that the residues at this position in the aligned sequences are different but share some level of similarity, typically in their chemical properties.

Asterisk (*): An asterisk indicates that the residues at this position are fully conserved across all sequences in the alignment. This means that every sequence has the same amino acid or nucleotide at this position, suggesting a critical role in structure or function.

Colon (:): The colon signifies that there is a moderate level of conservation at this position. Specifically, it indicates that the residues are not identical but are similar enough to be grouped together, often due to having similar chemical properties (e.g., both are hydrophobic).

In this study of Multiple Sequence Alignment (MSA) of the myd88 protein sequences FASTAQ **(Figure 9)** of *Homo sapiens of* myd88 *Mus musculus Drosophila melanogaster, Paenibacillus amylolyticus, Homo sapiens of* IL1 the conserved region , non- conserved region and similar regions are identified. Understanding the color codes in Clustal alignments allows for efficient analysis of sequence similarities and differences, which is crucial for evolutionary studies, functional predictions, and more **(Figure 10)**.

Conserved regions in genes refer to segments that have remained relatively unchanged throughout evolution across different species. These regions tend to be highly similar or identical in sequence and are often critical for essential biological functions. Because they play such important roles, any changes or mutations in these regions are usually harmful .

These conserved regions in this study of Multiple Sequence Alignment (MSA) of myd88 *Mus musculus Drosophila melanogaster, Paenibacillus amylolyticus, Homo sapiens of* IL1. Among this analysis *Homo sapiens of myd88 and Mus musculus* are more than 50% and are essential for myd88 adapter (protein) gene expression and ensure essential functions such as protein binding, receptor activation, and gene regulation.

Myd88 protein sequence of *Drosophila melanogaster, and Homo sapiens of* IL1 retains conserved coding sequences necessary for plays essential roles in TLR-mediated inflammatory response of invertebrate and vertebrate animals.

International Journal of Global Science Research
Vol. 13, Issue. 1, April 2026, pp. 8969-9001
Available Online at www.ijgsr.com

ISSN: 2348-8344 (Online)
DOI: https://doi.org/10.26540/ijgsr.v13.i1.2026.337

MyD88 in other species of vertebrates share similar structural characteristics, genomic structures, and flanking genes, suggesting that MyD88 is structurally conserved in different phyla of vertebrates ranging from fish to mammals. In *Paenibacillus amylolyticus*, the myd88 receptor gene demonstrates strong conservation in its tyrosine kinase domain, essential for signal transduction. Also it offers protection against insect herbivores and phytopathogens, including bacteria, fungi, nematodes, and viruses through innate immune mechanism.

Non-conserved regions in genes are sequences that show significant variation between different species, or even among individuals within the same species. Unlike conserved regions, which are preserved through evolution due to their essential biological functions, non-conserved regions are more flexible and prone to change without causing major problems for the organism. mutations in non-conserved areas are more likely to be tolerated, as they usually don't disrupt critical functions. These regions contribute to variations in gene expression, protein structure, and receptor functionality. It allow for species-specific adaptations and regulatory mechanisms.

Similar regions of a gene refer to sequences that show a high degree of resemblance between different genes, either within the same organism or across different species. These similarities often indicate shared ancestry, structural function, or evolutionary conservation. Similar regions helps to regulating expression of insulin gene.

A phylogenetic tree is a diagram that represents evolutionary relationships among various species or genes based on similarities and differences in their genetic sequences.This study stated that, *Homo sapiens of* myd88 *Mus musculus Drosophila melanogaster, Paenibacillus amylolyticus, Homo sapiens of* IL1 the phylogenetic tree helps visualize how closely related these adapter protein related sequences are and can provide insight into the evolutionary divergence of these genes. *Homo sapiens of* myd88 and *Paenibacillus amylolyticus, Mus musculus and Homo sapiens of* IL1 genes are expected to cluster closely on the tree, as *Homo sapiens of* myd88 *Mus musculus Paenibacillus amylolyticus, Homo sapiens of* IL1 *myd88 adapter* genes shared *Drosophila melanogaster* lineage (Figure 11). In Close Relationship two species or genes have a close relationship if they share a recent common ancestor. In the tree, this is represented by Shorter branch lengths between them.

A distant relationship means the species or genes diverged longer ago and share a more ancient common ancestor. In this tree, this appears as longer branch lengths between them. In this tree, this appears as longer branch lengths between them. For example the phylogenetic tree of *Homo sapiens of* IL1and *Paenibacillus amylolyticus* are distantly related (Figure 12).

The most common secondary structures are alpha helices and beta sheets **(**Figure: 13.1, 2, 3, 4, 5). An alpha helix (or α-helix) is a sequence of amino acids in a protein that are twisted into a coil (a helix).The alpha helix is the most common structural arrangement in the secondary structure of proteins. It is also the most extreme type of local structure, and it is the local structure that is most easily predicted from a sequence of amino acids.The alpha helix has a right-handed helix conformation in which every backbone N−H group hydrogen bonds to the backbone C=O group of the amino acid that is four residues earlier in the protein sequence.

International Journal of Global Science Research
Vol. 13, Issue. 1, April 2026, pp. 8969-9001
Available Online at www.ijgsr.com

ISSN: 2348-8344 (Online)
DOI: https://doi.org/10.26540/ijgsr.v13.i1.2026.337

The beta sheet (β-sheet, also β-pleated sheet) is a common motif of the regular protein secondary structure. Beta sheets consist of beta strands (β-strands) connected laterally by at least two or three backbone hydrogen bonds, forming a generally twisted, pleated sheet. A β-strand is a stretch of polypeptide chain typically 3 to 10 amino acids long with backbone in an extended conformation. The supra molecular association of β-sheets has been implicated in the formation of the fibrils and protein aggregates.

The protein secondary structure of *Homo sapiens of* myd88, *Mus musculus, Drosophila melanogaster, Paenibacillus amylolyticus, Homo sapiens of* IL1consists of alpha helix and beta sheet. Myeloid differentiating factor 88 (MyD88) is a cytosolic adaptor protein (Figure: 14.1, 2, 3, 4, 5), that plays essential roles in both innate and acquired immune responses by mediating signal transduction pathways that are initiated by Toll-like receptors (TLRs) and IL-1 and IL-18 receptors (IL-1R and IL-18R). MyD88 consists of an N-terminal death domain (DD) (approximately 90 aa residues), a C-terminal Toll/Interleukin-1 receptor (TIR) domain (approximately 150 aa residues), and a short connecting linker. In innate immune responses, the TIR domain of MyD88 has pivotal functions in the formation of signal initiation complexes involving the cytosolic domain of TLRs.

Protein secondary structure is the local spatial conformation of the polypeptide backbone excluding the side chains. The two most common secondary structural elements are alpha helices and beta sheets, though beta turns and omega loops occur as well. Secondary structure elements typically spontaneously form as an intermediate before the protein folds into its three dimensional tertiary structure.Secondary structure is formally defined by the pattern of hydrogen bonds between the amino hydrogen and carboxyl oxygen atoms in the peptide backbone.

The GOR method is a promising and efficient alternative to other protein aggregation predicting tools.The GOR graph of protein structure analysis of MDY88 of *Homo sapiens of* myd88, *Mus musculus, Drosophila melanogaster, Paenibacillus amylolyticus, Homo sapiens of* IL1showed the alpha helix is as red slope and beta strand are as blue slope (Figure: 14.1, 2, 3, 4, 5).

Here, The GOR graph method of protein secondary structure prediction of *Homo sapiens of* myd88, *Mus musculus, Drosophila melanogaster, Paenibacillus amylolyticus, Homo sapiens of* IL1were described. The method was based on information theory, and an assumption that information function of a protein chain can be approximated by a sum of information from single residues and pairs of residues. The analysis of frequencies of occurrence of secondary structure for singlets and doublets of residues in a protein database enables prediction of secondary structure for new amino acid sequences.

Through the successive incorporation of observed frequencies of single, then pairs of residues on a local sequence of *Homo sapiens of* myd88, *Mus musculus, Drosophila melanogaster, Paenibacillus amylolyticus, Homo sapiens of* IL1 amino acid residues were identified. The accuracy of the GOR method has improved from about 55% up to 64.4%. The GOR method has the advantage over neural network-based methods or nearest-neighbor methods in that it clearly identifies what is taken into account for the prediction and what is neglected. The method provides estimates of probabilities for the three secondary structures at each residue position that can be useful for further application of the method.

International Journal of Global Science Research
Vol. 13, Issue. 1, April 2026, pp. 8969-9001
Available Online at www.ijgsr.com

ISSN: 2348-8344 (Online)
DOI: https://doi.org/10.26540/ijgsr.v13.i1.2026.337

Discussion:

Innate immunity constitutes the primary line of defense in multicellular organisms against infectious pathogens and is activated upon the recognition of microbially derived pathogen-associated molecular patterns (PAMPs) by host pattern recognition receptors (PRRs) (Akira et al., 2006). Toll-like receptors (TLRs), a pivotal family of PRRs that are extensively characterized in vertebrates, instigate a signaling cascade that culminates in the activation of Myeloid differentiation factor 88 (MyD88) and the transcription factor nuclear factor kappaB (NF-kB) (Kawai T and Akira S, 2007). MyD88 comprises a Toll/interleukin-1 receptor (TIR) domain and a death domain, serving as the universal signaling adaptor protein shared by all TLRs, with the exception of TLR3..

The TIR domain is pivotal for the interactions between TLRs and MyD88. The death domain, in turn, engages with the death domain of interleukin-1 receptor-associated kinase (IRAK), thereby initiating downstream signaling cascades that culminate in the activation of NF-kB (Akira and Takeda 2004). MyD88 has been discerned in both vertebrates and invertebrates; however, variants of MyD88 have been exclusively identified in humans, mice, and chickens. The MyD88 gene in humans and mice comprises five exons and four introns. Notably, alternative splicing of exon 2 results in a protein that is devoid of the inhibitor of DNA (ID) domain typically situated between the DD and TIR domains (Strausberg RL et al., 2002). An additional splice variant of human MyD88, designated EXC09, is characterized by the absence of 19 bp within nucleotides 659–677.

In chickens, MyD88 variants were not produced through alternative splicing of MyD88 pre-mRNA (Wheaton et al., 2007). Chicken MyD88-2 exhibited a single nucleotide deletion at position 859, resulting in a protein that is 77 amino acids shorter than MyD88-1, the wild-type MyD88 comprising 376 amino acids. Additionally, chicken MyD88-3 displayed an 8 amino acid deletion at the N-terminus. The expression of MyD88 was modulated following the activation of TLR signaling in both vertebrates and invertebrates. Rock bream MyD88 has been demonstrated to be significantly up-regulated in blood, spleen, and head kidney in response to experimental challenges involving LPS and *Edwardsiella tarda* (Whang et al., 2011).

As we seen in the results it is evident that MYD88 in *Homo sapiens* and *Mus musculus* are more related having weakly conserved regions that can be seen in the multiple sequence analysis. Protein sequences between *Homo sapiens* IL-1 and *Paenibacillus amylolyticus* MYD88 are more linked than others protein sequences that are taken. As for *Drosophila melanogaster* it forms a common part before which some divergence took place in the ancestor of all the protein sequences to form what they are today.

For instance, amphibians serve as an exemplary model for investigating immunological evolution, as they are the first organisms to inhabit the interface between terrestrial and aquatic environments. As amphibians transition from aquatic habitats to terrestrial ones, there is a concomitant alteration in the pathogens with which they interact, influencing their innate immunity. Notably, positive selection has been observed in TLR2 and MYD88 (Zhang et al., 2024). This phenomenon is posited to arise from the diverse pathogens encountered as these species embarked on their journey from an aquatic to a terrestrial existence.

Detailed analysis revealed that the five intron positions were conserved among mollusks, four of which were also found

International Journal of Global Science Research
Vol. 13, Issue. 1, April 2026, pp. 8969-9001
Available Online at www.ijgsr.com

ISSN: 2348-8344 (Online)
DOI: https://doi.org/10.26540/ijgsr.v13.i1.2026.337

and conserved in human and zebrafish. (Toubiana, 2013 and Yipeng Ren et al., 2017) During evolution, some species obtained some species-specific intron For example, the MyD88 genes from *P. f. martensii* (Pfm-10008089) and C. gigas (Cgi-10026092) have the same introns at similar
positions, indicating that this intron occurred before the separation of *P. f. martensii* and *C. gigas* (Yu Jiaoa, et al., 2020).

A comparative study of immunology conducted on metazoans, particularly within Mollusca and Insecta, has demonstrated that commercially significant shellfish, such as *Scapharca subcrenata*, possess two distinct domains in MYD88: a death domain and a TLR domain (Yipeng Ren et al., 2017). Phylogenetic analysis indicates that during the course of evolution, positive selection has been observed in the TLR region, while the death domain remains unchanged. This suggests that as metazoans evolved, there was a notable alteration in the TLR region, (Kawai and Akira 2007) which is integral to innate immunity, in contrast to the death domain, which is primarily associated with cellular apoptosis.

From the aforementioned evolutionary relationship observed in Homo sapiens and various organisms, one salient conclusion emerges: the site of expression varies significantly across different species. For instance, in *Mus musculus*, MYD88 exhibits heightened expression in the lungs, whereas in *Homo sapiens*, MYD88 is predominantly expressed in the intestine. (Du, Yet al., 2013) This disparity underscores how pathogenic interactions have intricately influenced the evolutionary trajectories of diverse organisms.

The investigation of phylogenetic relationships within MYD88 (Ning, et al., 2015) is of paramount importance, as it plays a pivotal role in advancing our understanding of the evolution of innate immunity in response to various pathogenic interactions. A more profound exploration in this area may elucidate the mechanisms by which specific pathogens, such as COVID-19, operate and how they might influence contemporary organisms in the years to come.

The comprehensive genome-wide identification of MyD88 genes across eight mollusk genomes, specifically *P. f. martensii, C. gigas, M. yessoensis, M. philippinarum, B. platifrons, A. californica, O. bimaculoides, and L. gigantean,* alongside two vertebrate genomes, *D. rerio* and *H. sapiens*, revealed a notable expansion of MyD88 genes within bivalves. Phylogenetic tree analysis elucidated (Emmanuel Paradis et al., 2003) that this proliferation of MyD88 genes originated from an ancestral gene shared among mollusks, subsequently undergoing expansion within the bivalve lineage. The genomic architecture and domain composition suggested that MyD88 genes (Mayameei, 2007 and Strausberg et al., 2002) exhibit remarkable conservation in both invertebrates and vertebrates. We successfully acquired the cDNA sequence of PfmMyD88-2 from P. f. martensii and elucidated its role in the immune response, demonstrating its involvement in the NF-κB signaling pathway and its regulation by PfmmiR-4047 (Yu Jiaoa, et al., 2020, Lee, et al., 2011).

Future directions:

Multiple alignment of protein and nucleotide sequences will continue to be a pivotal application in the foreseeable future. The proliferation of newly available protein sequences significantly outstrips the number of elucidated three-dimensional protein structures, thereby rendering sequence homology the predominant methodology for inferring

ISSN: 2348-8344 (Online)
DOI: https://doi.org/10.26540/ijgsr.v13.i1.2026.337

protein structure, function, active sites, and evolutionary lineage. In recent years, tools for protein multiple sequence alignment (MSA) have advanced markedly in both scalability and precision. Prospective enhancements are likely to emerge from the integration of sequence alignment with supplementary data, such as the known structures of certain proteins under alignment or homology to an expansive repository of proteins.

References:

Khan Rahat, Amit Hasan Anik, Shabiha Hossain, Khamphe Phoungthong, Abu Reza Md. Towfiqul Islam, Narottam Saha, Abubakr M. Idris, Md. Harunor Rashid Khan, Saad Aldawood, Mahbub Alam. (1988) Receptor model-based source tracing and risk assessment of elements in sediment of a transboundary Himalayan River,Chemosphere,Volume 339,2023,139733,ISSN 0045-6535, *https://doi.org/10.1016/j.chemosphere.* 2023.139733. Res. 16 10881–10890.Khan Rayyan, Xinghua Ma, Quaid Hussain, Keling Chen, Saqib Farooq, Muhammad Asim, Xiaochun Ren, Shahen Shah, Yi Shi. (2023)Transcriptome and anatomical studies reveal alterations in leaf thickness under long-term drought stress in tobacco, Journal of Plant Physiology,Volume 281,153920,ISSN 0176-1617,https://doi.org/10.1016/j.jplph.2023.153920.

Smith, J.H. and Singh, M. (2024). Unlocking Secrets: bio-informatics' Impact on Forensic Bio-Examinations. *Int. J. Netw. Secur. Its Appl*, *16*, pp.1-15.

Luscombe, NM. Greenbaum, D. Gerstein, M. (2001) . What is bioinformatics? A proposed definition and overview of the field Methods of information in medicine 40 (4), 346-358.

Bayat, A. Science, medicine, and the future: bio-informatics. (2002) BMJ. Apr 27;324(7344):1018-22. doi: 10.1136/bmj.324.7344.1018. PMID: 11976246; PMCID: PMC1122955.

Sandhya Sharma, Kumari Arpita, Machindra Nirgude, Harsha Srivastava, Kuldeep Kumar, Rohini Sreevathsa, Ramcharan Bhattacharya, Kishor Gaikwad. (2024).Genomic insights into cytokinin oxidase/dehydrogenase (CKX) gene family, identification, phylogeny and synteny analysis for its possible role in regulating seed number in Pigeonpea (Cajanus cajan (L.) Millsp.),International Journal of Biological Macromolecules,Volume 277, Part 3,2024,134194,ISSN 0141-8130,https://doi.org/10.1016/j.ijbiomac.134194.

Buchmann, K. (2014). Evolution of innate immunity: clues from invertebrates via fish to mammals, Front. Immunol. 5 459, https://doi.org/10.3389/fimmu.2014.00459.

Akira, S. Hemmi, H. (2003) Recognition of pathogen-associated molecular patterns by TLR family, Immunol. Lett. 85 85–95, https://doi.org/10.1016/S0165-2478(02)00228-6.

Uematsu, S. Akira, S. (2008). Toll-Like Receptors (TLRs) and Their Ligands. Toll-like Receptors (TLRs) and Innate Immunity, Springer, pp. 1–20, https://doi.org/ 10.1007/978-3-540-72167-3_1.

Uematsu, S. Akira, S. Akira, S. Henzel,W.J. Shillinglaw, W. Li, S. Cao, Z. (1997). MyD88: an adapter that recruits IRAK to the IL-1 receptor complex, Immunity 7 837–847, https://doi.org/10.1016/S1074-7613(00)80402-1.

Ishii KJ, Uematsu S, Akira S. (2006) 'Toll'gates for future immunotherapy. Current Pharmaceutical Design. 2006 Nov 1;12(32):4135-42.

Jault, C. Pichon, L.Chluba, J. (2003)Toll-like receptor gene family and TIR-domain adapters in Danio rerio, Mol. Immunol. 40

(2004) 759–771, https://doi.org/10. 1016/j.molimm.10.001

Whang I, Lee Y, Kim H, Jung SJ, Oh MJ, et al. (2011). Characterization and expression analysis of the myeloid differentiation factor 88 (MyD88) in rock bream Oplegnathus fasciatus. Mol Biol Rep 38: 3911–3920.

Qiu, L. Song, L. Yu,Y. Xu, W. Ni, D. Zhang, Q. (2007). Identification and characterization of a myeloid differentiation factor 88 (MyD88) cDNA from Zhikong scallop Chlamys farreri, Fish Shellfish Immunol. 23 614–623, https://doi.org/10.1016/j.fsi. 2007.01.012.

Li, X.-C. Zhu,L. Li, L.-G. Ren,Q. Huang, Y.-Q.Lu, J.-X. et al., (2013). A novel myeloid differentiation factor 88 homolog, SpMyD88, exhibiting SpToll-binding activity in the mud crab Scylla paramamosain, Dev. Comp. Immunol. 39 313–322, https://doi.org/10.1016/j.dci.2012.11.011.

Liu, X.C., Li, S.J., Zhang, X.Y., et al. (2013) Effect of Tongxinluo Capsule on Vascular Endothelial Function, Brachial-Ankle and Pulse Wave Velocity in Stable Angina. Chinese Journal of Integrative Medicine on Cardio/Cerebrovascular Disease, 11, 408.

Niu, S. Shi, X. Zhang, J. Chai, L. Xiao, X. (2016) Cloning, characterization, and expression analysis of MyD88 in Rana dybowskii, Appl. Biochem. Biotechnol. 179 294–306, https://doi.org/10.1007/s12010-016-1994-y

Bonnert, T.P. Garka, K.E. Parnet, P. Sonoda,G. Testa,J.R. Sims, J.E. (1997). The cloning and characterization of human MyD88: a member of an IL-1 receptor related family 1The nucleotide sequences reported in this paper have been submitted to the Genbank/EMBL Data Bank with accession numbers U84408 and U84409.1, FEBS (Fed. Eur. Biochem. Soc.) Lett. 402 81–84, https://doi.org/10.1016/S0014-5793(96)01506-2.

Van der Sar, A.M. Stockhammer, O.W. van der Laan,C. Spaink, H.P. Bitter, W. Meijer, A.H. (2006) MyD88 innate immune function in a zebrafish embryo infection model, Infect. Immun. 74 2436–2441, https://doi.org/10.1128/iai.74.4.2436-2441.2006.

Zhang, L.. Li, L. Guo,X. Litman, G.W. Dishaw, L.J. Zhang, G. (2015). Massive expansion and functional divergence of innate immune genes in a protostome, Sci. Rep. 5 8693, https://doi.org/10.1038/srep08693 https://www.nature.com/articles/ srep08693#supplementary-information.

Ren, Q. Chen, Y.-H.Ding, Z.-F. Huang, Y. Shi, Y.-R. (2014). Identification and function of two myeloid differentiation factor 88 variants in triangle-shell pearl mussel (Hyriopsis cumingii), Dev. Comp. Immunol. 42 286–293, https://doi.org/ 10.1016/j.dci.2013.09.012

Chowdhury, B. and Gautam Garai (2017). A review on m ultiple sequence alignment from the perspective of genetic algorithm. Genomics 419-429

Kumar, S. Filipski, A. (2007). Multiple sequence alignment: in pursuit of homologous DNA positions, Genome Res. 17 127–135.

Morgenstern, B. Prohaska,S.J. Pöhler, D. Stadler, P.F. (2006) Multiple sequence alignment with user-defined anchor points, Algorithms Mol. Biol. 1 6.

Morgenstern I, Powlowski J, Ishmael N, Darmond C, Marqueteau S, Moisan MC, Quenneville G, Tsang A. (2012) A molecular phylogeny of thermophilic fungi. Fungal Biology. Apr 1;116(4):489-502.

Xiong, J. (2006). Essential bio-informatics, Cambridge University Press, NY,

Henikoff, S. Henikoff, J.G.(1994) Protein family classification based on searching a database of blocks. Genomics. Jan 1;19(1):97-107.

Heringa, J. . Taylor, W.R (1997). Three-dimensional domain duplication, swapping and stealing, Curr. Opin. Struct. Biol. 7 416–421.

Higgins, D.G.. Sharp, P.M. (1988). CLUSTAL: a package for performing multiple sequence alignment on a microcomputer, Gene 73 237–244.

Hogeweg, P. Hesper, B. (1984) .The alignment of sets of sequences and the construction of phylogenetic trees: an integrated method, J. Mol. Evol. 20 175–186.

Huang,X. (1994). On global sequence alignment, Comput. Appl. Biosci. 10 227–235.

Taylor, W.R. (1987). Multiple sequence alignment by a pairwise algorithm, Comput. Appl. Biosci. 3 81–87.

Taylor, W.R. (1988). A flexible method to align large numbers of biological sequences, J. Mol. Evol. 28 (161–169.

Corpet F. (1988) Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res.Nov 25;16(22):10881-90. doi: 10.1093/nar/16.22.10881. PMID: 2849754; PMCID: PMC338945.

Pei, J. Grishin, N.V. (2007) PROMALS: towards accurate multiple sequence alignments of distantly related proteins, bio-informatics 23 802–808.

Thompson, J.D. Higgins, D.G.Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice, Nucleic Acids Res. 22 4673–4680.

Lassmann, T. Sonnhammer, E.L. Kalign. (2005). An accurate and fast multiple sequence alignment algorithm, BMC Bioinf. 6 298.

Notredame, C. Higgins, D.G. Heringa, J. (2000) T-Coffee: a novel method for fast and accurate multiple sequence alignment, J. Mol. Biol. 302 205–217.

Jimin Pei, Ruslan Sadreyev, Nick, V. Grishin, PCMA: (2003) Fast and accurate multiple sequence alignment based on profile consistency, *bio-informatics*, Volume 19, Issue 3, february Pages 427–428, https://doi.org/10.1093/bio-informatics/btg008

Pei, J. and Grishin, N.V. (2006) MAMMALS: multiple sequence alignment improved by using hidden Markov models with local structural information, Nucleic Acids Res. 34 4364–4374.

Sokal, R.R. Michener, C.D. (1958). A statistical method for evaluating systematic relationships, Univ. Kans. Sci. Bull. 28 1409–1438.

Stoye, J. Moulton, V. Dress, A.W. (1998) DCA: An efficient implementation of the divideand conquer approach to simultaneous multiple sequence alignment, Comput. Appl. Biosci. 13 625–626.

Stoye,J. (1988) Multiple sequence alignment with the divide-and-conquer method, Gene 211 GC45–GC56.

Katoh, K. Kuma, K. Toh, H. Miyata, T. (2005). MAFFT version 5: improvement in accuracy of multiple sequence alignment, Nucleic Acids Res. 33 511–518.

Katoh, K.Misawa, K. Kuma,K. Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform, Nucleic Acids Res. 30 3059–3066.

Yamada, S. Gotoh, O. Yamana, H. (2006). Improvement in accuracy of multiple sequence alignment using novel group-to-group sequence alignment algorithm with piecewise linear gap cost, BMC Bioinf. 7 524.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy

and high throughput, Nucleic Acids Res. 32 1792–1797.

Blaisdell,J.O. Hatahet, Z. Wallace, S.S.(1999) A novel role for Escherichia coli endonuclease VIII in prevention of spontaneous G→ T transversions. Journal of bacteriology. 1999 Oct 15;181(20):6396-402.

Kimura, M. (1981) Estimation of evolutionary distances between homologous nucleotide sequences. Proceedings of the National Academy of Sciences. 1981 Jan;78(1):454-8.

Paradis, E. Claude, J. Strimmer, K. (2004) APE: analyses of phylogenetics and evolution in R language. Bioinformatics. 2004 Jan 22;20(2):289-90.

Akira S, Uematsu S, Takeuchi O (2006) Pathogen recognition and innate immunity. Cell 124: 783–801.

Akira, S. and Takeda, K. (2004) Toll-like receptor signalling. Nature Reviews Immunology 4: 499–511.

Wheaton, S., Lambourne M.D., Sarson A.J., Brisbin J.T., Mayameei. A. et al. (2007). Molecular cloning and expression analysis of chicken MyD88 and TRIF genes. DNA Seq 18: 480–486.

Toubiana, M. Gerdol,M. Rosani, U. Pallavicini, A.Venier, P. Roch, P. (2013). Toll-like receptors and MyD88 adaptors in Mytilus: complete cds and gene expression levels, Dev. Comp. Immunol. 40 158–166, https://doi.org/10.1016/j.dci.2013.02. 006.

Yipeng Ren , Junli Xue , Huanhuan Yang , Baoping Pan , Wenjun Bu (2017) Compar ative and evolutionary analysis of an adapter molecule MyD88 in invertebrate metazoans.

Jie Zhang, Ruinan Zhao, Hongyan Bi, Jiaoying He, Yang Guo, Dian Liu, Ganggang Yang, Xiaohong Chen (2024). Positive Selection of TLR2 and MyD88 Genes ProvidesInsights Into the Molecular Basis of Immunological adaptation in Amphibians.

Yu Jiaoa, b. Zefeng Gua , Shaojie Luoa , Yuewen Deng. (2020). Evolutionary and functional analysis of MyD88 genes in pearl oyster *Pinctada fucata martensii*. Fish and Shellfish Immunology 99 322–330.

Emmanuel Paradis, Julien Claude and Korbinian Strimmer (2004). Analyses of Phylogenetics and Evolution in R language. bio-informatics Applications Note. Vol. 20 no. 2 pages 289–290. DOI: 10.1093/bio-informatics/btg412

Ning, X. Wang,R. Li, X. Wang,S. Zhang, M. Xing,Q. et al., (2015) Genome-wide identification and characterization of five MyD88 duplication genes in Yesso scallop (Patinopecten yessoensis) and expression changes in response to bacterial challenge, Fish Shellfish Immunol. 46 181–191, https://doi.org/10.1016/j.fsi. 2015.06.028.

Kawai, Tand Akira, S. (2007). Signaling to NF-(kappa) B by Toll-like receptors. Trends in molecular medicine 13: 460–469.

Mayameei, A. (2007) Molecular cloning and expression analysis of chicken MyD88 and TRIF genes, J. DNA Sequencing Mapp. https://doi.org/10.1080/ 10425170701295856.

Strausberg, R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D, et al. (2002) Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. Proc Natl Acad Sci U S A 99: 16899.

Du, Y. Zhang, L. Huang, B. Guan, X. Li, L. Zhang,G. (2013). Molecular cloning, characterization, and expression of two myeloid differentiation Factor 88 (Myd88) in Pacific Oyster, Crassostrea gigas, J. World Aquacult. Soc. 44 759–774, https://doi.org/10.1111/jwas.12077

Lee, Y. Whang, I. Umasuthan, N. De Zoysa, M. Oh, C. Kang, D.-H. et al., (2011). Characterization of a novel molluscan MyD88 family protein from

ISSN: 2348-8344 (Online)
DOI: https://doi.org/10.26540/ijgsr.v13.i1.2026.337

manila clam, Ruditapes philippinarum, Fish Shellfish Immunol. 31 887–893, https://doi. org/10.1016/j.fsi.2011.08.003.

International Journal of Global Science Research
Vol. 13, Issue. 1, April 2026, pp. 8969-9001
Available Online at www.ijgsr.com

ISSN: 2348-8344 (Online)
DOI: https://doi.org/10.26540/ijgsr.v13.i1.2026.337

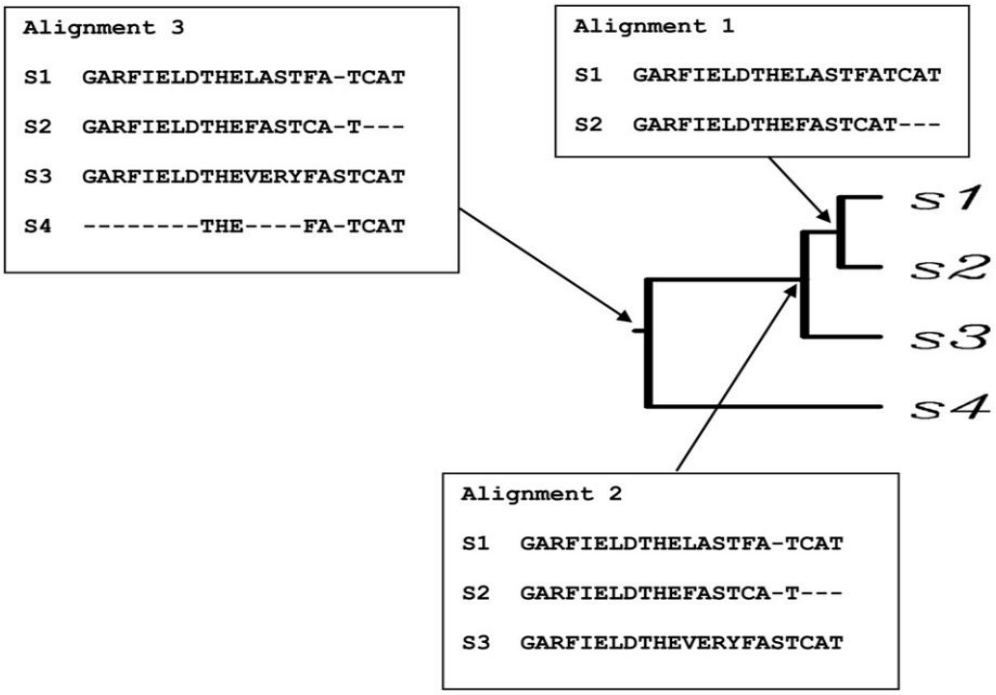

Fig 1: Guide tree guides the merging order of sequences based on the pairwise distances calculated for all the possible sequence pairs to be aligned in MSA

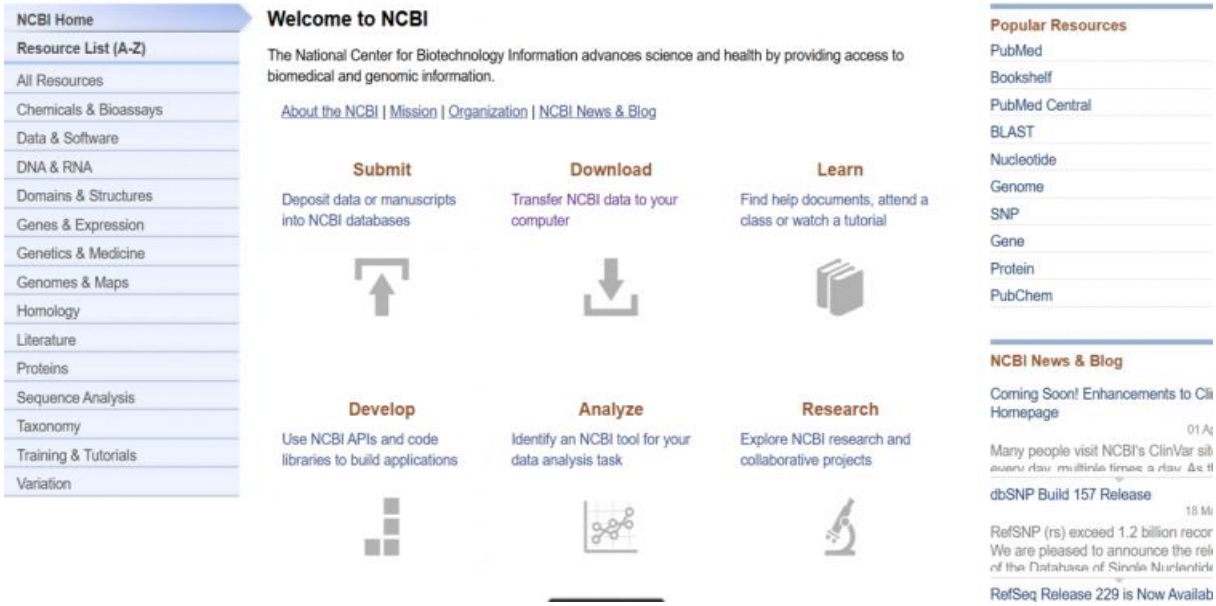

Fig 2: Website page of NCBI- The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information.

International Journal of Global Science Research **ISSN: 2348-8344 (Online)**
Vol. 13, Issue. 1, April 2026, pp. 8969-9001 **DOI: https://doi.org/10.26540/ijgsr.v13.i1.2026.337**
***Available Online* at www.ijgsr.com**

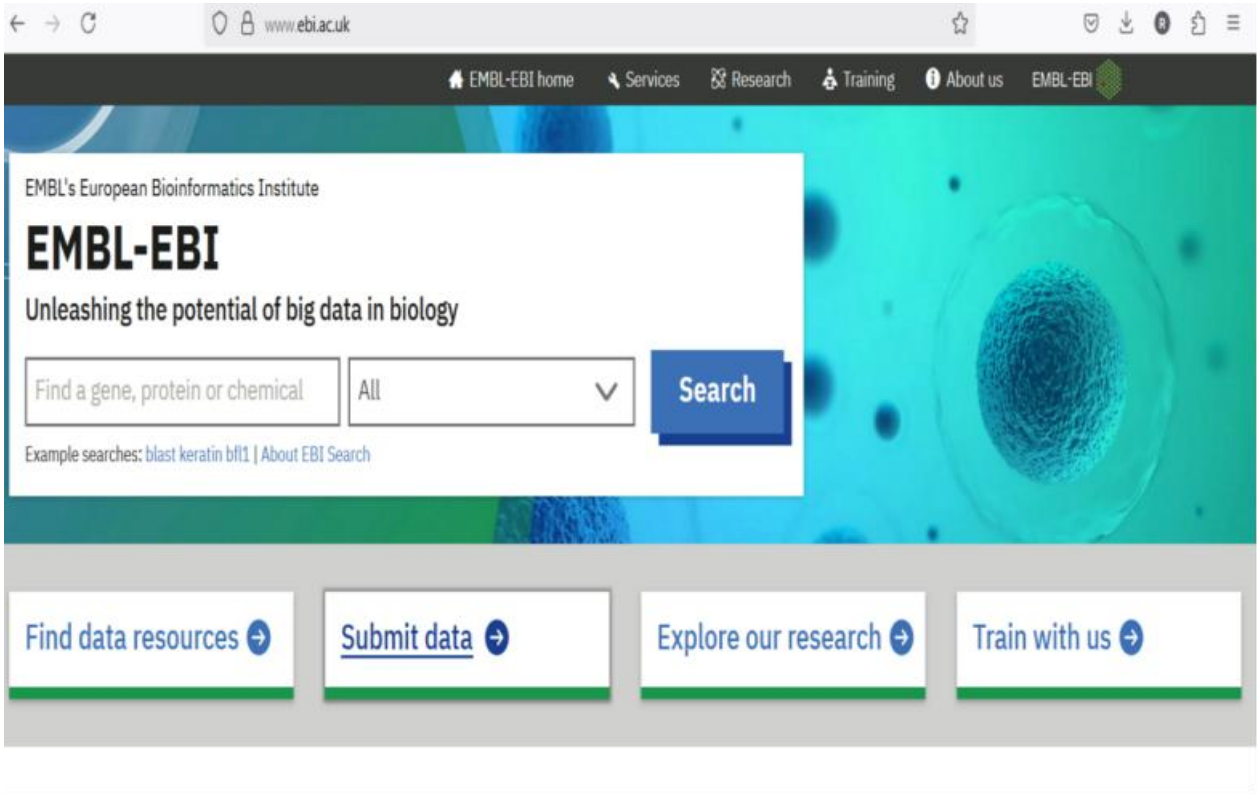

Fig 3: Website page of EMBL

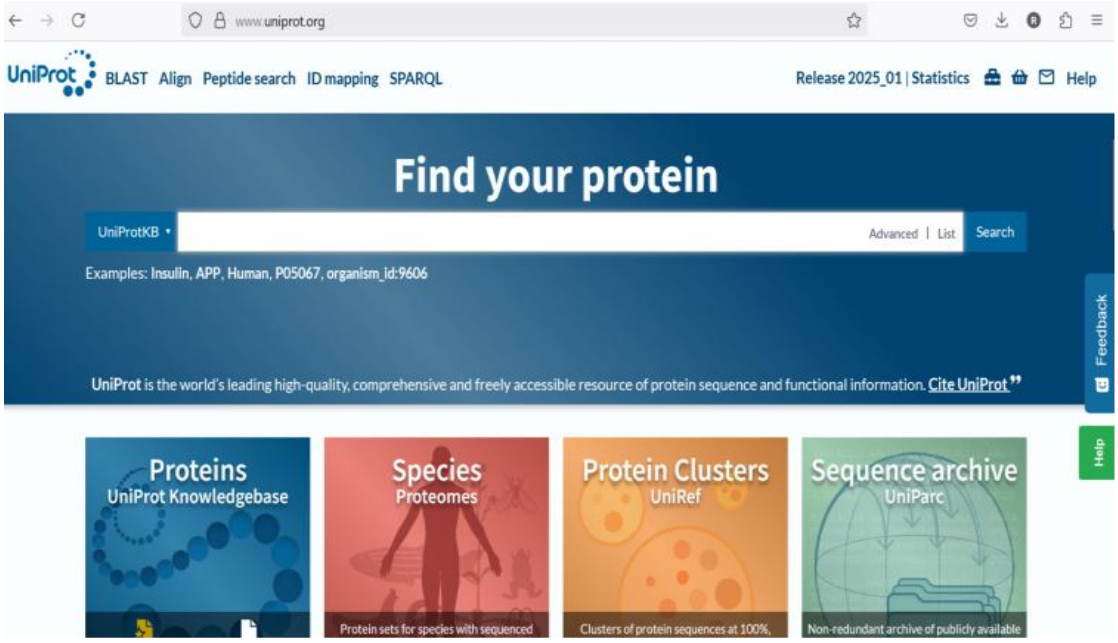

Fig 4: Website page of Uniprot

International Journal of Global Science Research
Vol. 13, Issue. 1, April 2026, pp. 8969-9001
Available Online at www.ijgsr.com

ISSN: 2348-8344 (Online)
DOI: https://doi.org/10.26540/ijgsr.v13.i1.2026.337

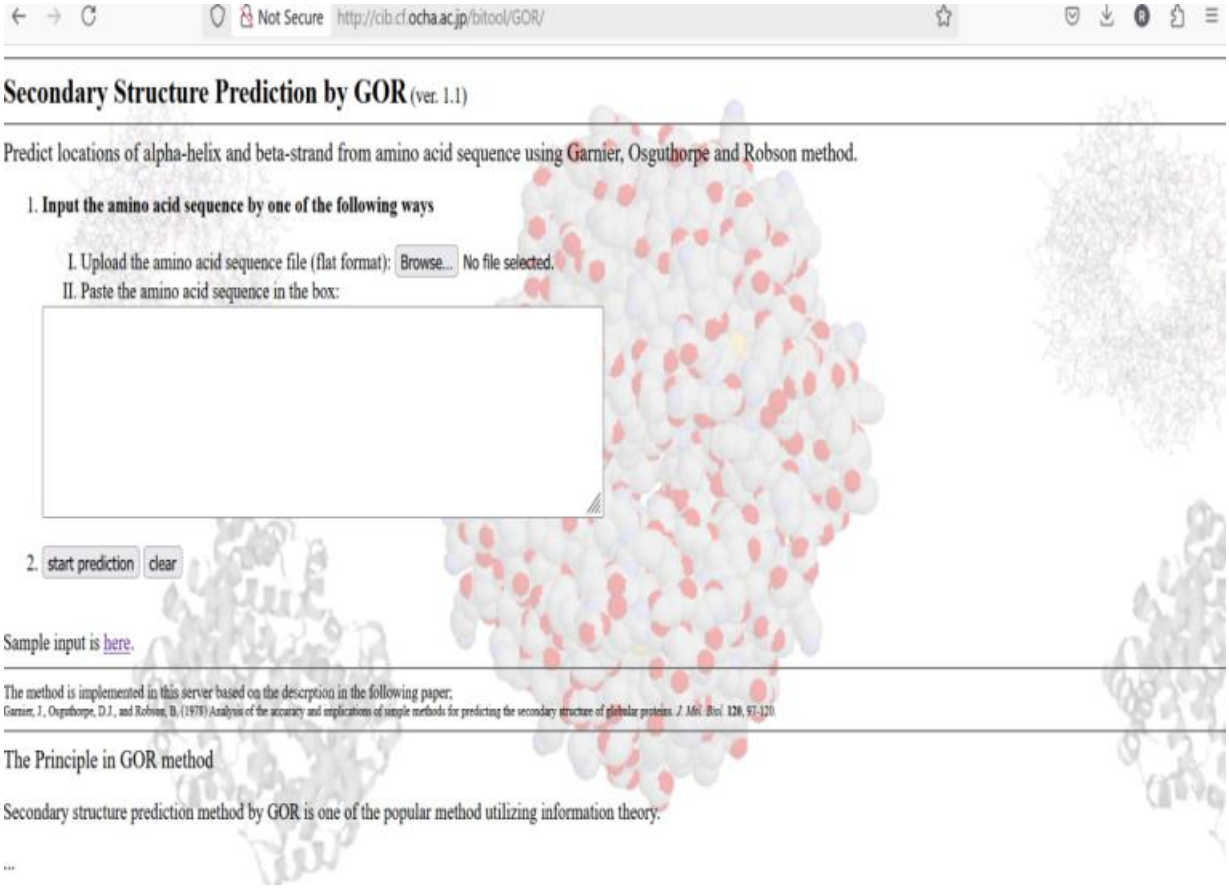

Fig 5: Website of GOR

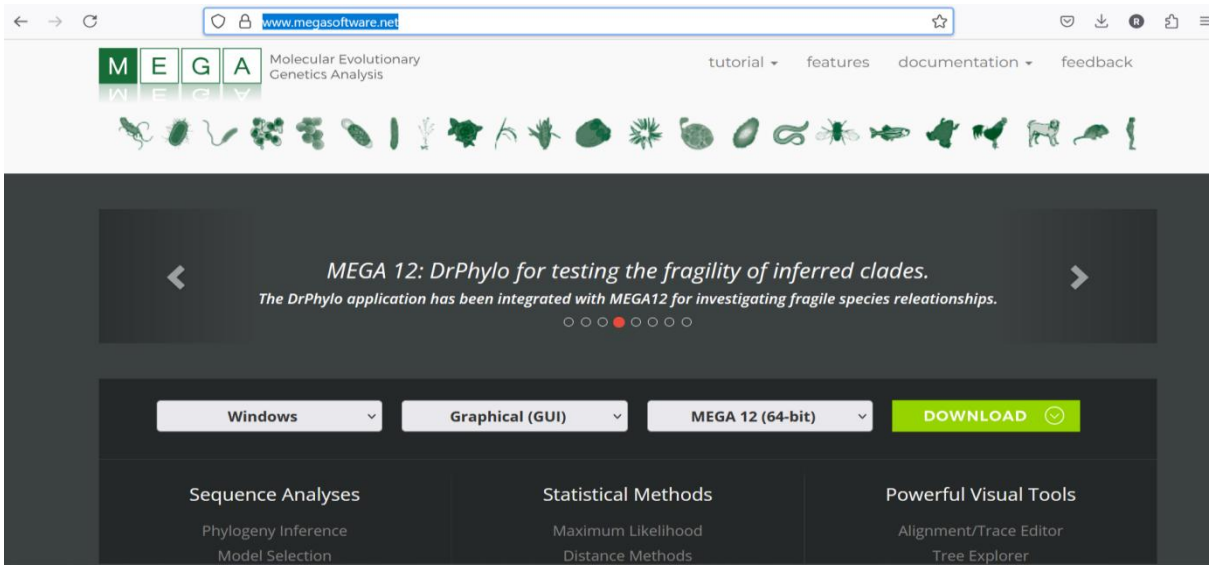

Fig 6: Website of MEGA

International Journal of Global Science Research
Vol. 13, Issue. 1, April 2026, pp. 8969-9001
Available Online at www.ijgsr.com

ISSN: 2348-8344 (Online)
DOI: https://doi.org/10.26540/ijgsr.v13.i1.2026.337

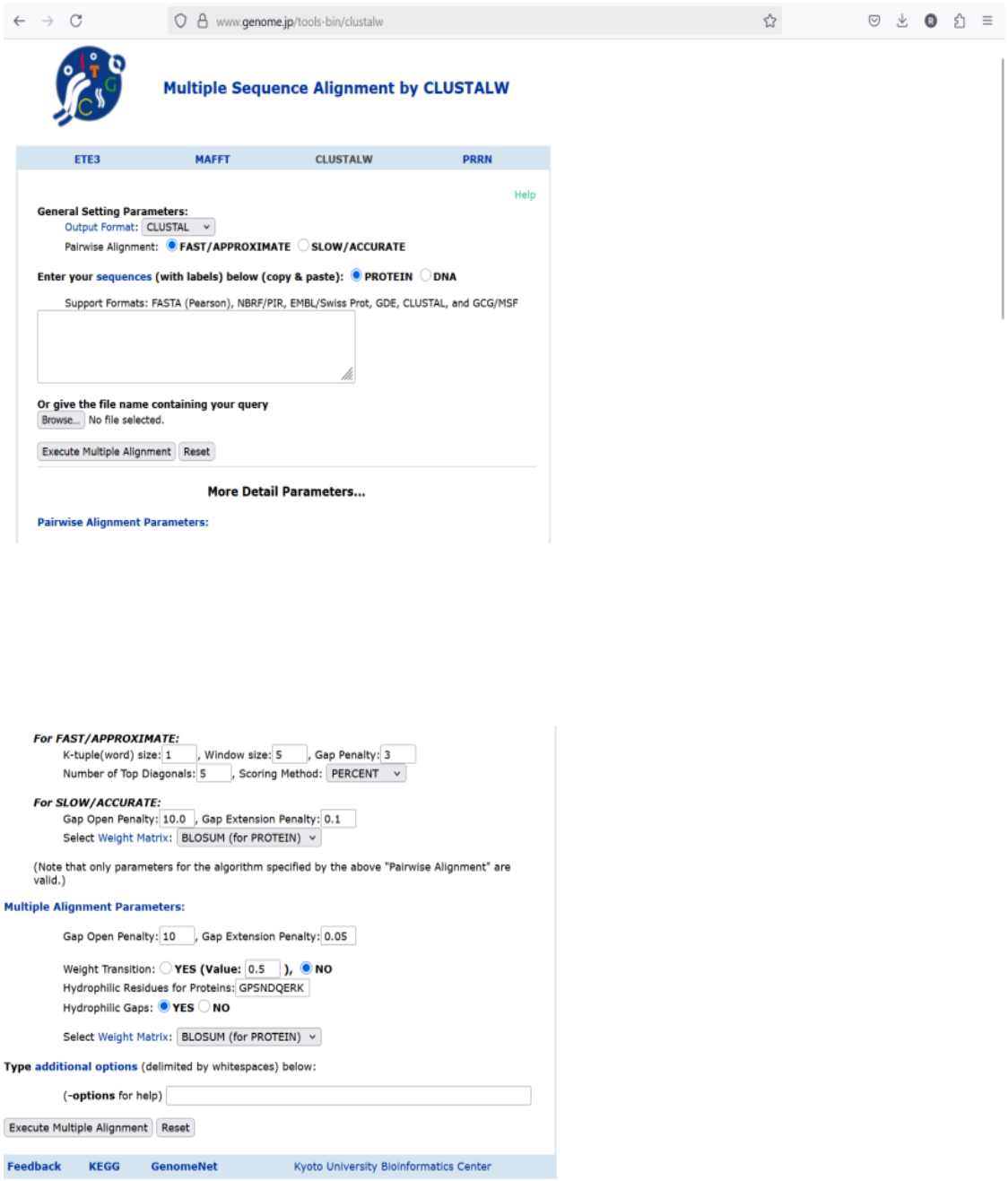

Fig 7: Website of Clustral W

International Journal of Global Science Research
Vol. 13, Issue. 1, April 2026, pp. 8969-9001
Available Online at www.ijgsr.com

ISSN: 2348-8344 (Online)
DOI: https://doi.org/10.26540/ijgsr.v13.i1.2026.337

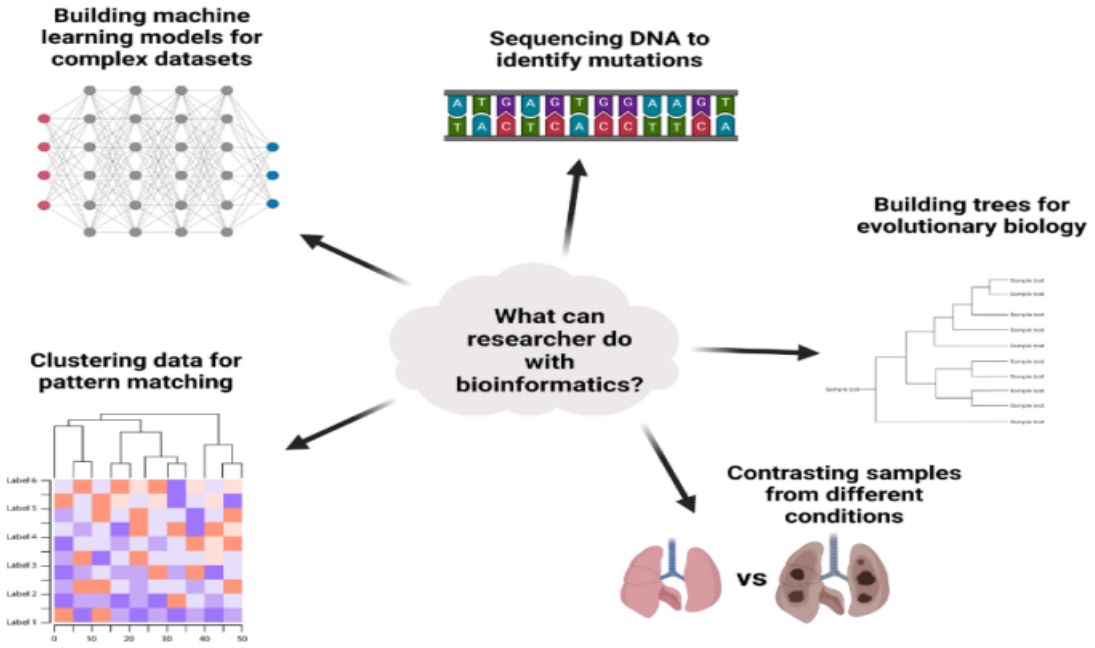

Figure 8: The bio-informatics Pipeline to Biology encompasses the development of machine learning models, the sequencing of DNA samples, the construction of evolutionary trees, and the analysis and visualization of expression data.

```
>UNIPROT:MYD88_HUMAN Q99836 Myeloid differentiation primary response protein MyD88 {ECO:0000305}
MAAGGPGAGSAAPVSSTSSLPLAALNMRVRRRLSLFLNVRTQVAADWTALAEEMDFEYLE
IRQLETQADPTGRLLDAWQGRPGASVGRLLELLTKLGRDDVLLELGPSIEEDCQKYILKQ
QQEEAEKPLQVAAVDSSVPRTAELAGITTLDDPLGHMPERFDAFICYCPSDIQFVQEMIR
QLEQTNYRLKLCVSDRDVLPGTCVWSIASELIEKRCRRMVVVVSDDYLQSKECDFQTKFA
LSLSPGAHQKRLIPIKYKAMKKEFPSILRFITVCDYTNPCTKSWFWTRLAKALSLP

>UNIPROT:MYD88_MOUSE P22366 Myeloid differentiation primary response protein MyD88
MSAGDPRVGSGSLDSFMFSIPLVALNVGVRRRLSLFLNPRTPVAADWTLLAEEMGFEYLE
IRELETRPDPTRSLLDAWQGRSGASVGRLLELLALLDREDILKELKSRIEEDCQKYLGKQ
QNQESEKPLQVARVESSVPQTKELGGITTLDDPLGQTPELFDAFICYCPNDIEFVQEMIR
QLEQTDYRLKLCVSDRDVLPGTCVWSIASELIEKRCRRMVVVVSDDYLQSKECDFQTKFA
LSLSPGVQQKRLIPIKYKAMKKDFPSILRFITICDYTNPCTKSWFWTRLAKALSLP

>tr|Q7K105|Q7K105_DROME LD20892p OS=Drosophila melanogaster OX=7227 GN=Myd88 PE=1 SV=1
MRPRFVCHQQHSVAHSHYQPHSHFHHHTHRHPNPPHHHHIYGATDVSYRRYRTAGMVVAE
GVMDSGSGSGTGTGLGHFNETPLSALGIETRTQLSRMLNRKKVLRSEEGYQRDWRGISEL
AKQKGFVDENANNPMDLVLISWSQRSPQTAKVGHLEHFLGIIDRWDVCDDIQENLAKDTQ
RFIMKQEQRQTALVEACPPPPSDCFETNNNYSSSNNITVGQSVQILSDEDQRCVQMGQPL
PRYNACVLYAEADIDHATEIMNNLESERYNLRLFLRHRDMLMGVPFEHVQLSHFMATRCN
HLIVVLTEEFLRSPENTYLVNFTQKIQIENHTRKIIPILYKTDMHIPQTLGIYTHIKYAG
DSKLFNFWDKLARSLHDLDAFSIYSTRQVQTPSPVEESAPRVTTPSIRIQINDKDVTDMP
NYNSCKVPEAETTIVSVSGDTGSPLPEHKPKKKDRFLRRITHSFGKTARSDGASGKTLRH
AHSVSTINVTERERTLSASSSNISTTSESKKSFIKWQPNILKKALFSRSSSKLQTPG

>tr|A0A5M9WX28|A0A5M9WX28_PAEAM Antifreeze protein type I OS=Paenibacillus amylolyticus OX=1451 GN=EC604_19725 PE=4 SV=1
MAIIDVIKYDGSPDVFAWKHPETELGTWTQLIVNQSQQAILFKDGRALDLFGPGRHTLST
ANIPILNRLINLPFGGKSPFAAEVWYVNQVSAMDVKWGTANPIQVQDPKYNIIVPVRAFG
QMGMKISDSRKFLVKLVGTLPEFNQLNMVNYFRGLITMNINSMLSSYLIHRKVSVLEINA
YIAEISRHFADNVASTFDEFGIELINLYIHNVNLPEEDPSVIRLREALARKAEMDIIGYT
YQQERSFDTLEGAAKNEGSMQSDIMGAGLGMGMGVGLGGSFSSGMSQMSKVMSTSTETSA
VRLCGHCQHPNQEHSSFCSKCGKSLAEKSATTDCNHCGHTMEKGTKFCPNCGDKYYACPS
CGADNAENASECVKCHEPMPRPCPNCNHMNGPHTKFCGNCGTGLTLKCSQCQHEVQPGQK
FCLDCGNNLQEGGQA
```

Figure:9. Protein sequences of (FASTAQ) Myd88 (*Homo sapiens of* myd88 *Mus musculus Drosophila melanogaster, Paenibacillus amylolyticus, Homo sapiens of* IL1).

International Journal of Global Science Research **ISSN: 2348-8344 (Online)**
Vol. 13, Issue. 1, April 2026, pp. 8969-9001 **DOI: https://doi.org/10.26540/ijgsr.v13.i1.2026.337**
***Available Online* at www.ijgsr.com**

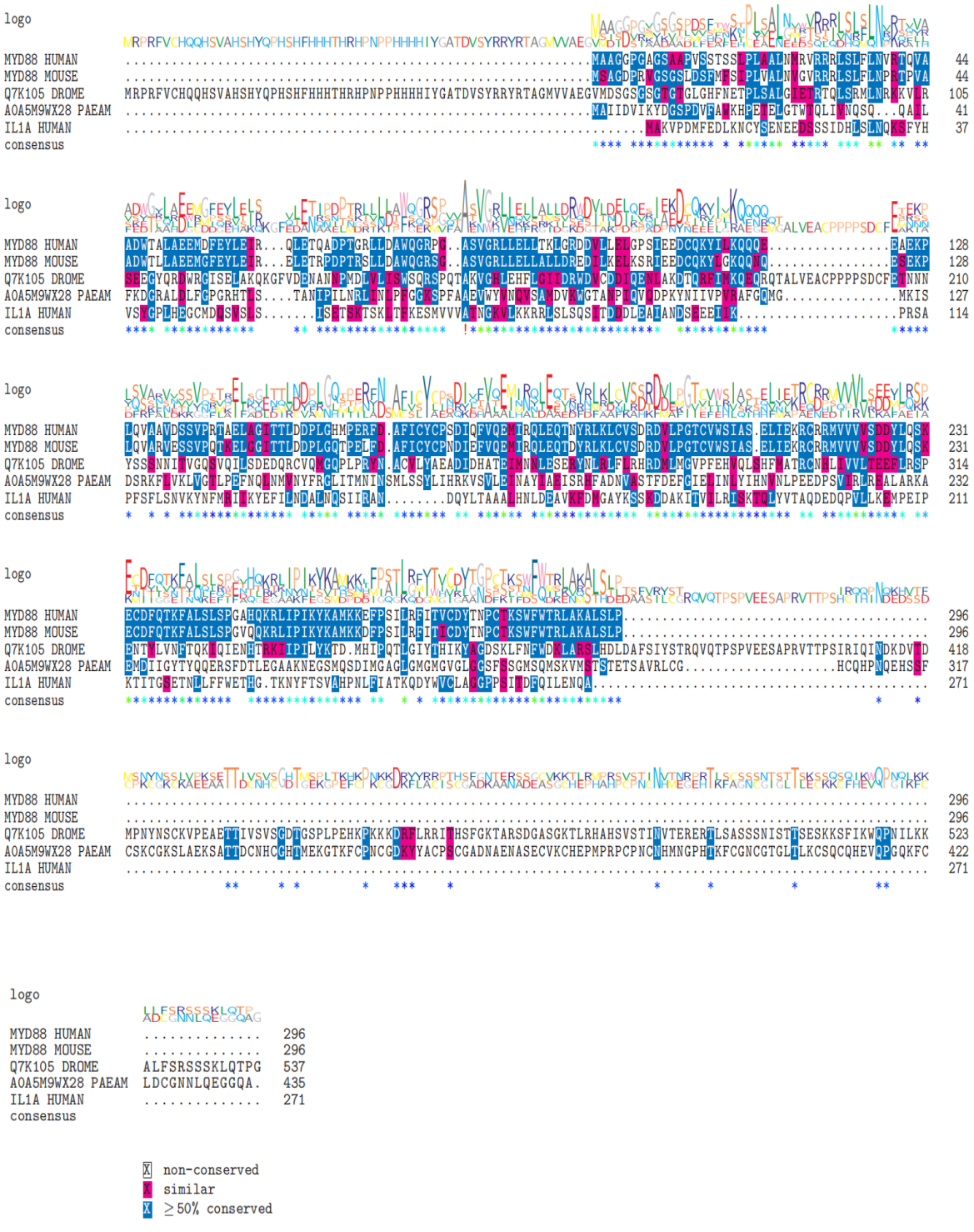

Figure:10. Multiple Sequence Alignment of Myd88 (*Homo sapiens of* myd88 *Mus musculus Drosophila melanogaster, Paenibacillus amylolyticus, Homo sapiens of* IL1).

International Journal of Global Science Research
Vol. 13, Issue. 1, April 2026, pp. 8969-9001
Available Online at www.ijgsr.com

ISSN: 2348-8344 (Online)
DOI: https://doi.org/10.26540/ijgsr.v13.i1.2026.337

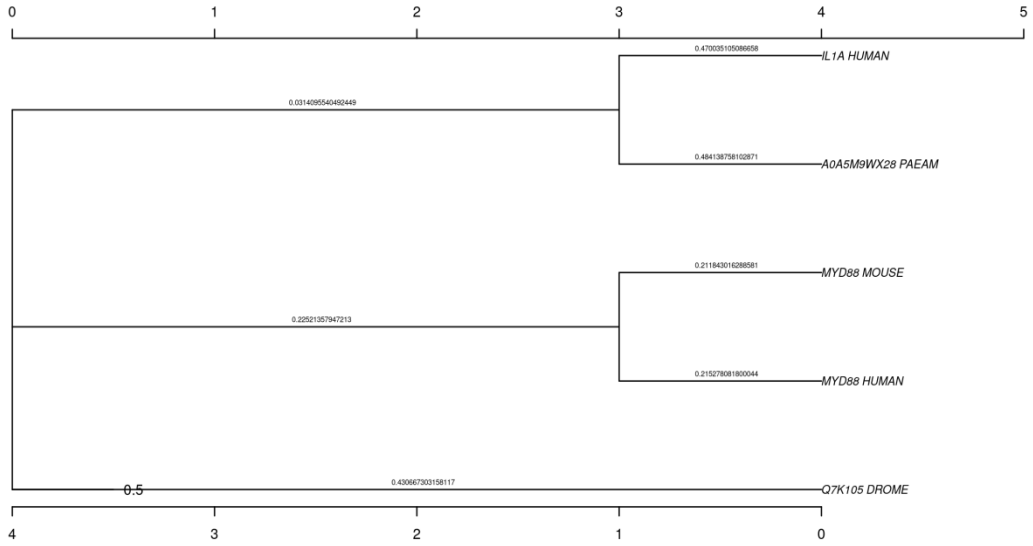

Figure:11. Phylogenetic tree showing similarities among *Homo sapiens of* myd88 *Mus musculus Drosophila melanogaster, Paenibacillus amylolyticus, Homo sapiens of* IL1

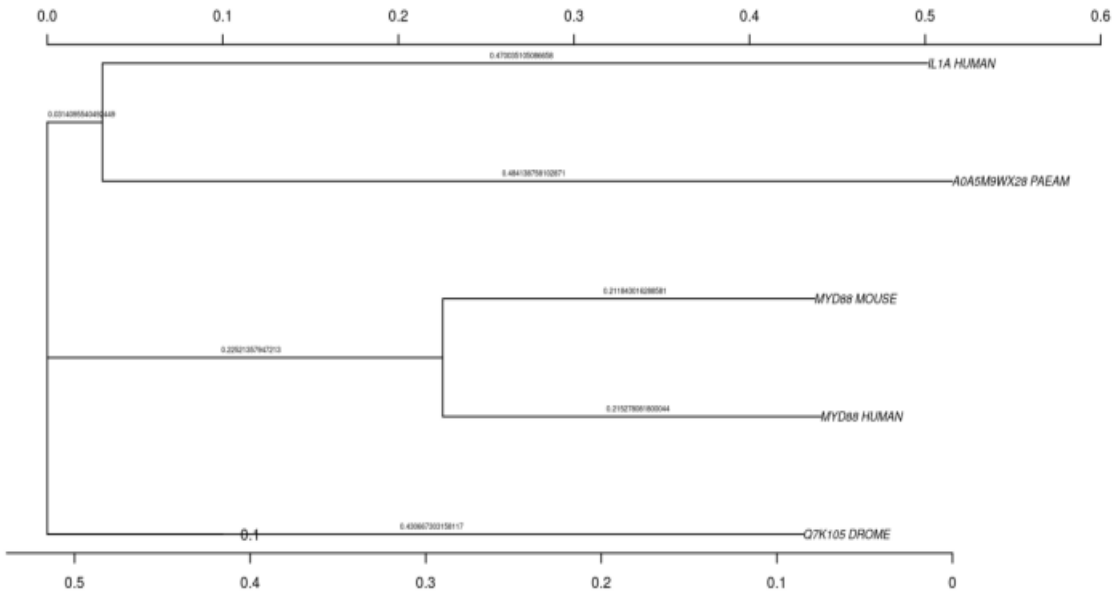

Figure:12. Phylogenetic tree showing distance among *Homo sapiens of* myd88 *Mus musculus Drosophila melanogaster, Paenibacillus amylolyticus, Homo sapiens of* IL1

International Journal of Global Science Research
Vol. 13, Issue. 1, April 2026, pp. 8969-9001
Available Online at www.ijgsr.com

ISSN: 2348-8344 (Online)
DOI: https://doi.org/10.26540/ijgsr.v13.i1.2026.337

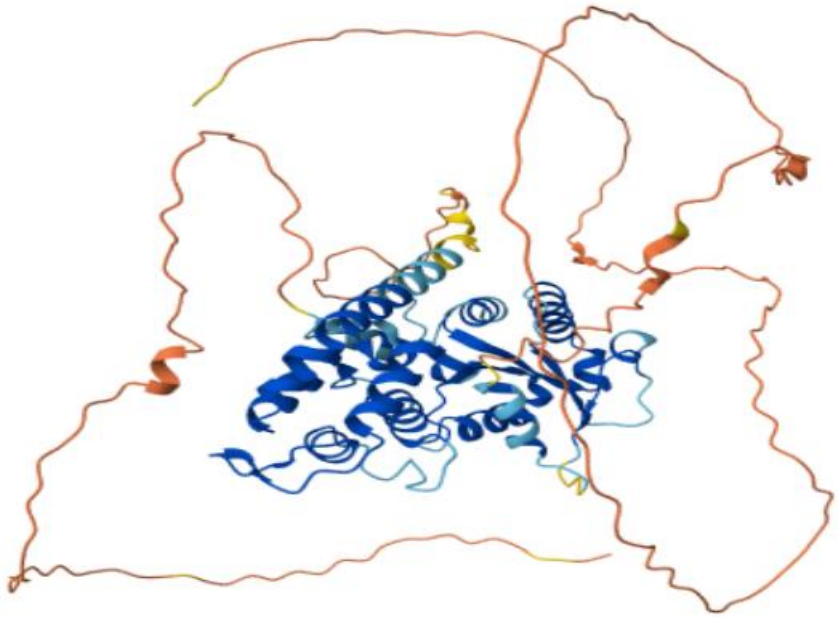

Figure:13.1. Protein structure- (a) MYD88 of *Homo sapiens*

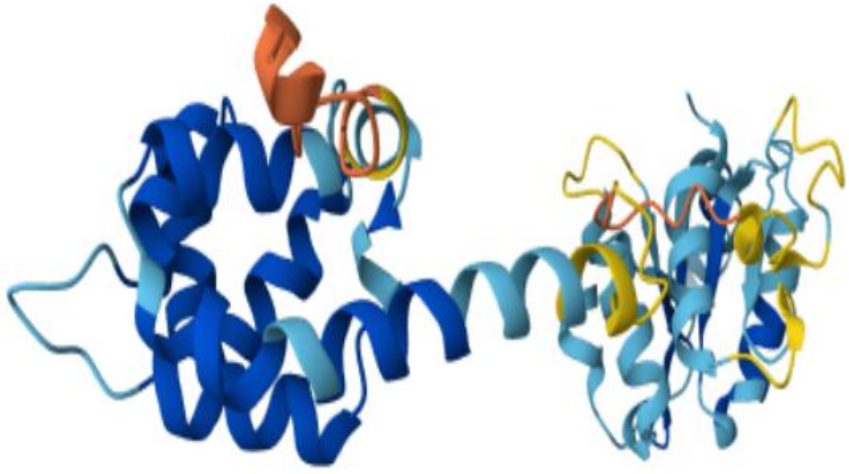

Figure:13.2. Protein structure (b) MYD88 of *Mus musculus*

International Journal of Global Science Research
Vol. 13, Issue. 1, April 2026, pp. 8969-9001
Available Online at www.ijgsr.com

ISSN: 2348-8344 (Online)
DOI: https://doi.org/10.26540/ijgsr.v13.i1.2026.337

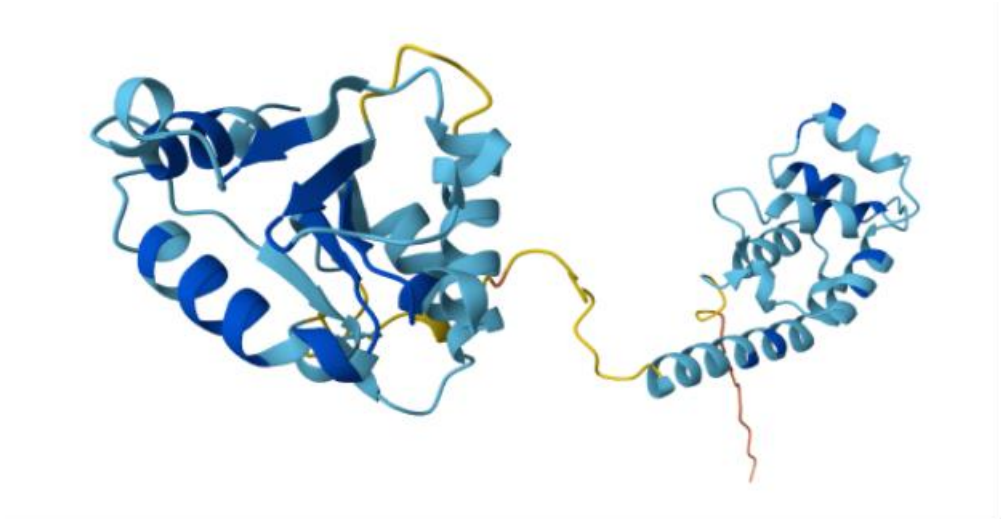

Figure:13.3. Protein structure (c) MYD88 of *Drosophila melanogaster*

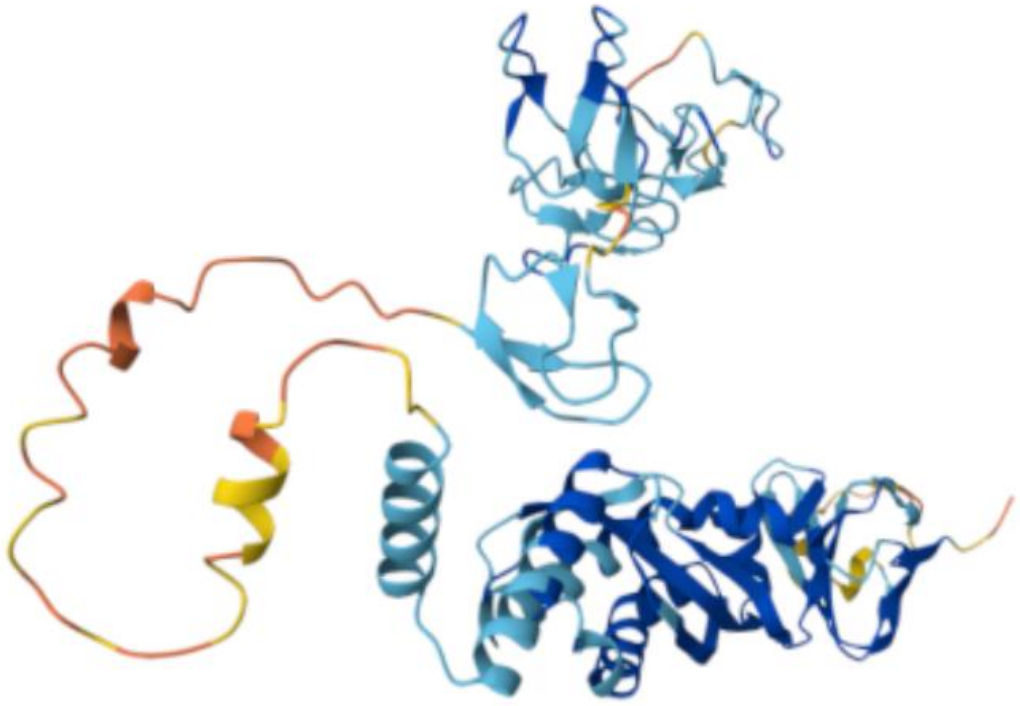

Figure:13.4. Protein structure (d) MYD88 of *Paenibacillus amylolyticus*

International Journal of Global Science Research
Vol. 13, Issue. 1, April 2026, pp. 8969-9001
Available Online at www.ijgsr.com

ISSN: 2348-8344 (Online)
DOI: https://doi.org/10.26540/ijgsr.v13.i1.2026.337

Figure:13.5. Protein structure (e)Interleukin alpha-1 of *Homo sapiens*

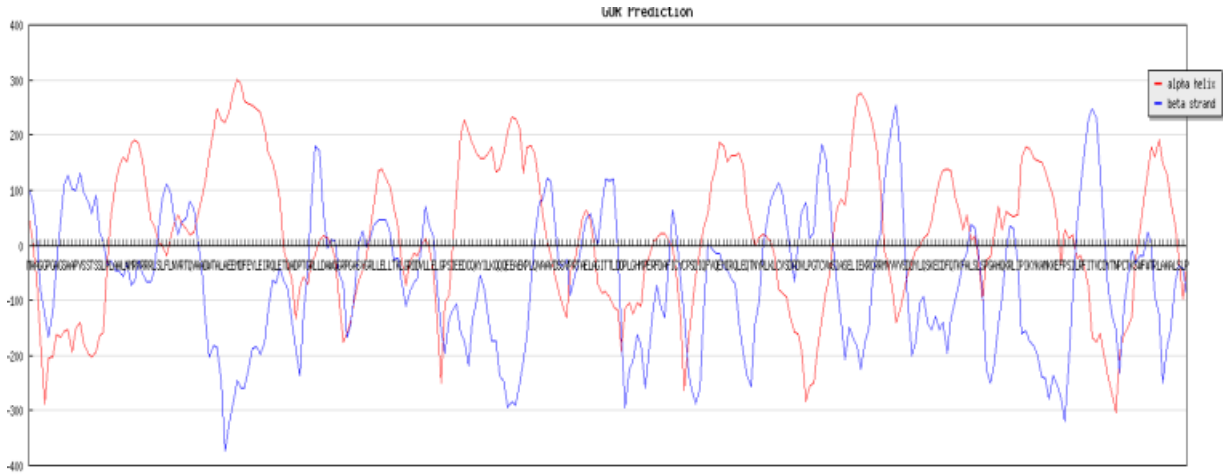

Figure:14.1. GOR graph of secondary structure (a)MYD88 of *Homo sapiens*

International Journal of Global Science Research
Vol. 13, Issue. 1, April 2026, pp. 8969-9001
Available Online at www.ijgsr.com

ISSN: 2348-8344 (Online)
DOI: https://doi.org/10.26540/ijgsr.v13.i1.2026.337

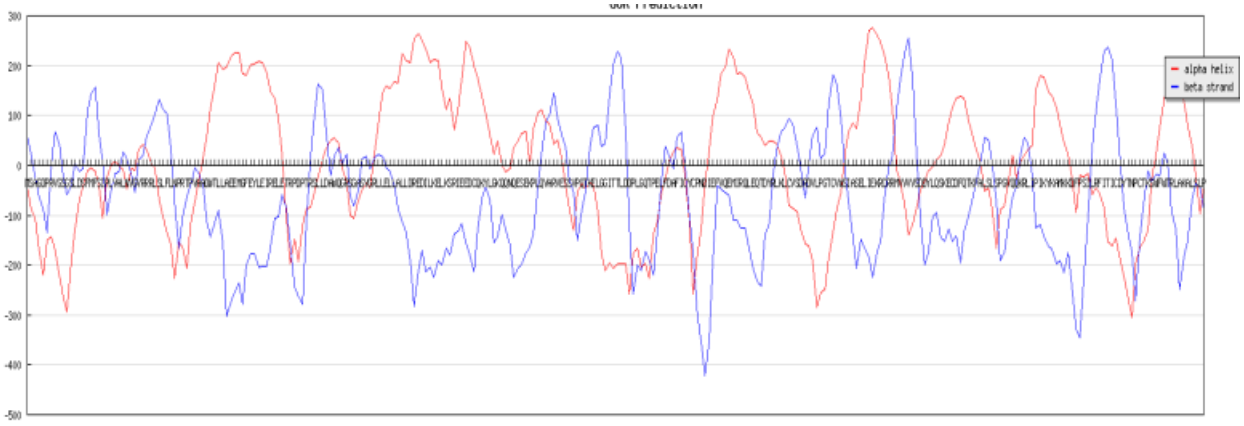

Figure:14.2.GOR graph of secondary structure (b)MYD88 of *Mus musculus*

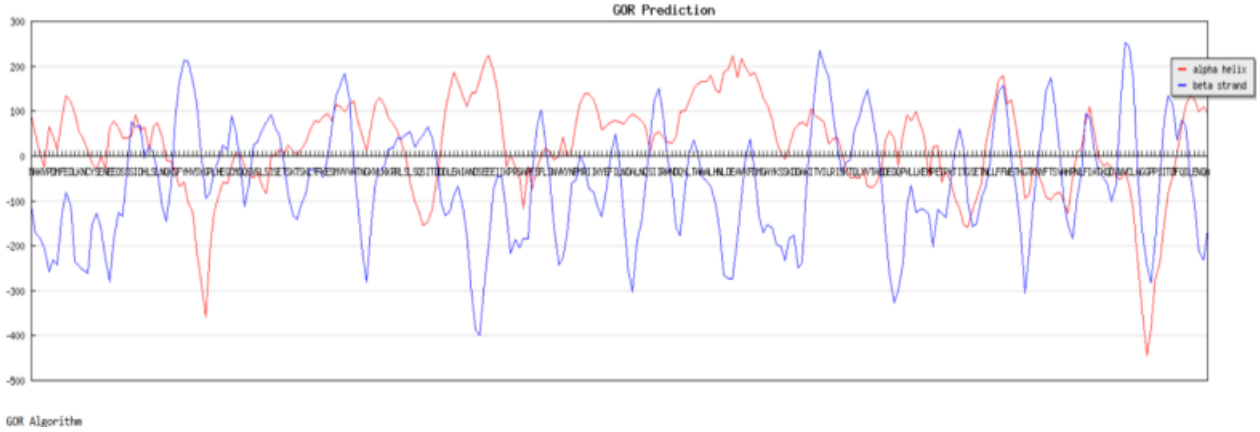

Figure:14.3.GOR graph of secondary structure MYD88 of *Drosophila melanogaster.*

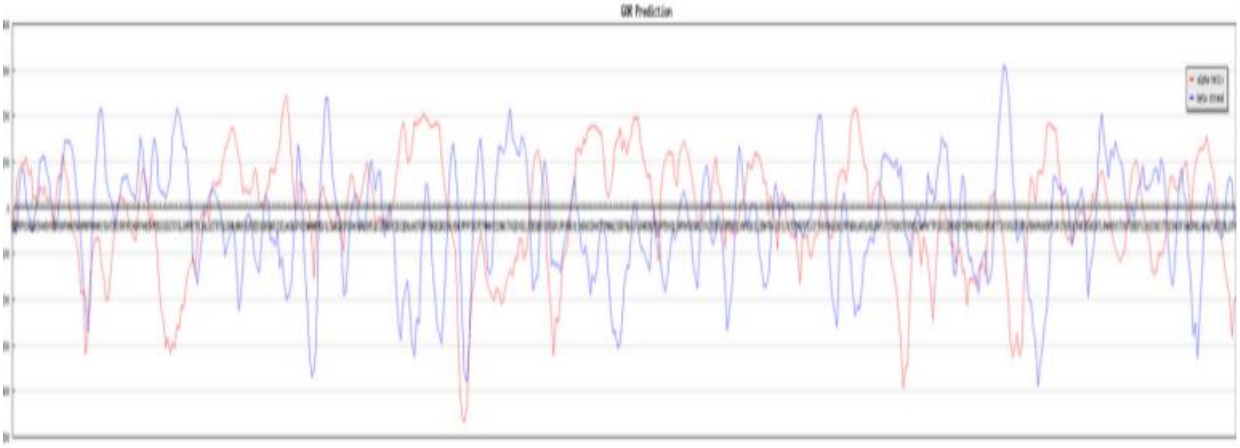

Figure:14.4.GOR graph of secondary structure (b)MYD88 of *Paenibacillus amylolyticus.*

International Journal of Global Science Research
Vol. 13, Issue. 1, April 2026, pp. 8969-9001
Available Online at www.ijgsr.com

ISSN: 2348-8344 (Online)
DOI: https://doi.org/10.26540/ijgsr.v13.i1.2026.337

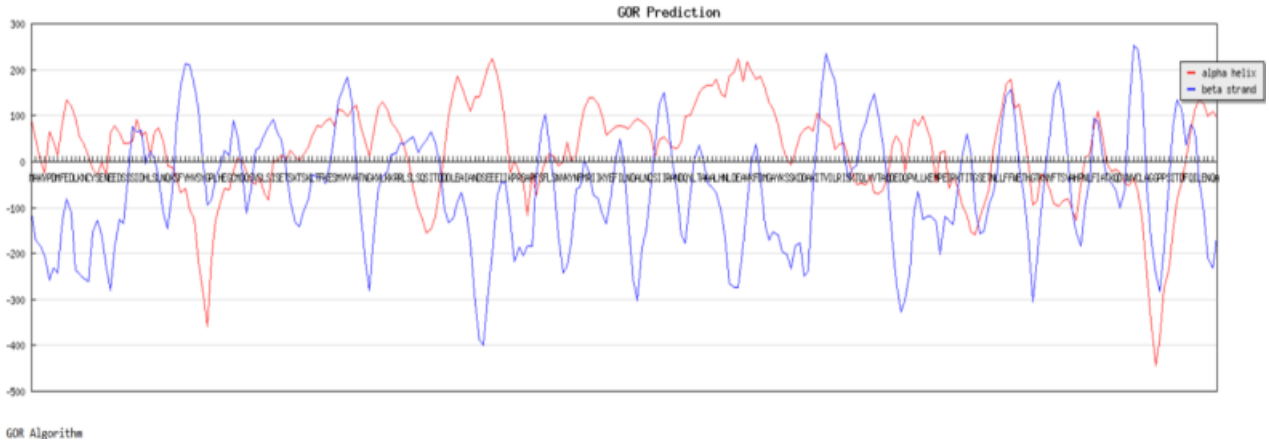

Figure: 14.5.GOR graph of secondary structure (b)MYD88 of *Homo sapiens of* IL1

Table 1: Myd88 gene information of *Mus musculus, Drosophila melanogaster, Paenibacillus amylolyticus, Homo sapiens*

Variants of Myd88	**Organism Name**	**Uniport ID**	**Biological function**
Myeloid differentiation primary response protein MyD88	Homo sapiens	Q99836	Adapter protein involved in the Toll-like receptor and IL-1 receptor signalling pathway in the innate immune response.
Myeloid differentiation primary response protein MyD88	*Mus musculus*	P22366	Adapter protein involved in the Toll-like receptor and IL-1 receptor signalling pathway in the innate immune response.
LD20892p	*Drosophila melanogaster*	Q7K105	Cyclase activity, molecular adapter and other molecular function
Antifreeze protein type I	*Paenibacillus amylolyticus*	A0A5M9WX28	Multi drug resistance
Interleukin-1 alpha	*Homo sapiens*	P01583	Plays an important role in inflammation and bridges the innate and adaptive immune systems.