## Database update

# PRRDB 2.0: a comprehensive database of pattern-recognition receptors and their ligands

**Dilraj Kaur[1], Sumeet Patiyal[1], Neelam Sharma[1], Salman Sadullah Usmani[1,2] and Gajendra P.S. Raghava[1,*]**

[1]Department of Computational Biology, Indraprastha Institute of Information Technology, Delhi, New Delhi 110020, India and [2]Bioinformatics Centre, CSIR-Institute of Microbial Technology, Chandigarh, Chandigarh 160036, India

*Corresponding author: Tel: +91 11 2690744; Email: raghava@iiitd.ac.in

## Abstract

PRRDB 2.0 is an updated version of PRRDB that maintains comprehensive information about pattern-recognition receptors (PRRs) and their ligands. The current version of the database has ~2700 entries, which are nearly five times of the previous version. It contains extensive information about 467 unique PRRs and 827 pathogens-associated molecular patterns (PAMPs), manually extracted from ~600 research articles. It possesses information about PRRs and PAMPs that has been extracted manually from research articles and public databases. Each entry provides comprehensive details about PRRs and PAMPs that includes their name, sequence, origin, source, type, etc. We have provided internal and external links to various databases/resources (like Swiss-Prot, PubChem) to obtain further information about PRRs and their ligands. This database also provides links to ~4500 experimentally determined structures in the protein data bank of various PRRs and their complexes. In addition, 110 PRRs with unknown structures have also been predicted, which are important in order to understand the structure–function relationship between receptors and their ligands. Numerous web-based tools have been integrated into PRRDB 2.0 to facilitate users to perform different tasks like (i) extensive searching of the database; (ii) browsing or categorization of data based on receptors, ligands, source, etc. and (iii) similarity search using BLAST and Smith–Waterman algorithm.

**Database URL:** https://webs.iiitd.edu.in/raghava/prrdb2/

## Introduction

Innate immunity is ubiquitous, naturally present in all classes of plants and animals, and acts as the first line of defense. Innate has originated from the Latin word 'Innatus', which stands for 'Inborn'. It comprises cells and mechanism providing the first line of defense against invading pathogens in a nonspecific manner. Besides macrophages and dendritic cells, nonprofessional cells such as fibroblasts and endothelial and epithelial cells play a significant role in pathogen recognition (1). Innate immunity also contributes to acute inflammation induced by tissue damage or microbial infection and also activates the acquired immunity. Innate immune cells utilize germline-encoded specialized receptors known as pattern recognition receptors (PRRs) to identify the pathogen- and damage-associated molecular patterns (PAMPs and DAMPs, respectively) present on invading microbes (2, 3).

In the past, different PRR families have been studied, which mainly include transmembrane proteins such as Toll-like receptors (TLRs) and C-type lectins receptors (CLRs) and cytoplasmic proteins such as nucleotide-binding oligomerization domain (NOD)-like receptors (NLRs) and retinoic acid-inducible gene-I-like receptors (RLRs). TLRs are type-1 transmembrane proteins and are the most extensively studied PRRs, which detect PAMPs associated with invading pathogens outside of the cell and in intracellular endosomes and lysosomes (4–8). CLRs are the signaling transmembrane CLRs, which play essential roles in antifungal immunity. CLRs are characterized by at least one C-type lectin-like domain and recognize both exogenous as well as endogenous ligands (6, 9). NLRs and RLRs are intracellular cytosolic sensors. NLRs, generally associated with bacterial recognition, are composed of a central nucleotide binding domain and C-terminal leucine-rich repeats, whereas RLRs are the helicases that sense PAMPs with viral RNA (10, 11). In addition to this, several other receptors such as scavenger receptors, mannose receptors and $\beta$-glucan receptors induce phagocytosis. Some secreted PRRs, such as complement receptors, collectins, ficolins, pentraxins for instances, serum amyloid and C-reactive protein, lipid transferases, peptidoglycan recognition proteins (PGRs), XA21D, etc., are not associated with their productive cells (12).

The primary function of all the PRRs is to recognize the essential microbial components, i.e. PAMPs or DAMPs. The association of PRRs with PAMPs leads to diverse phenomena such as maturation, migration and activation of immune cells, secretion of cytokines and chemokines (13). Most of the PRRs upregulate the transcription of genes regulating the proteins involved in inflammatory response such as type I interferons (IFNs), proinflammatory cytokines, chemokines, antimicrobial proteins, etc. They also upregulate the transcription and translation of proteins involved in the modulation of PRR signaling that may traverse to the adaptive immune response (2, 14, 15) (Figure 1). Adaptive immune system or 'specific immune system' is composed of specific cells and includes humoral and cell-mediated immunity, which aids in eliminating
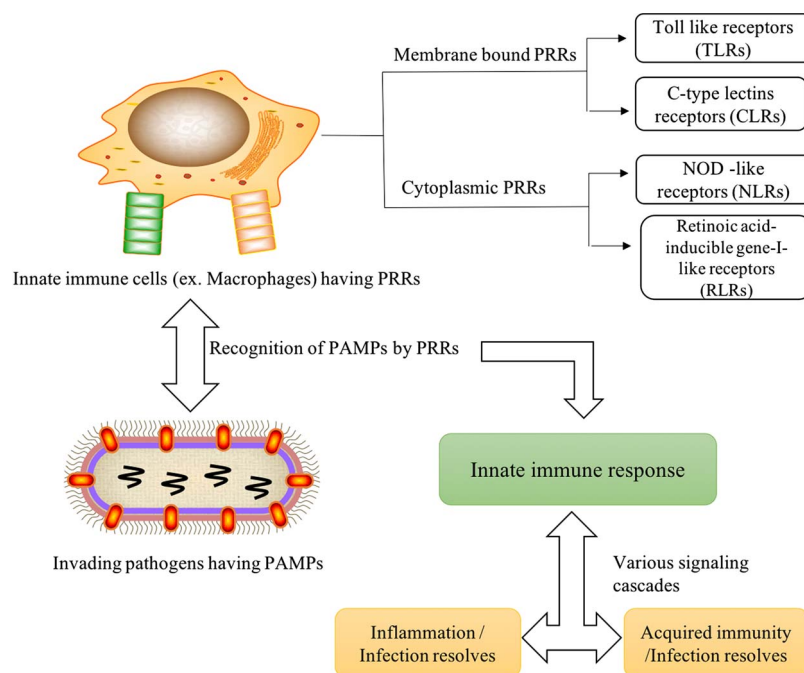


**Figure 1.** Graphical representation of immune mechanism through PRRs and PAMPs association after the microbial invasion.
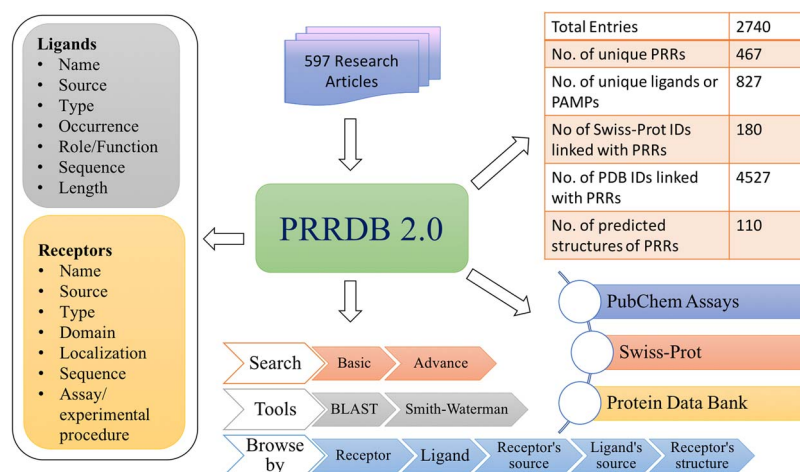
**Figure 2.** Architecture of PRRDB 2.0.

pathogens in the late phase of infection. Based on PRR, the innate immune system easily differentiates the type of pathogens and thus recruits the most effective form of adaptive immune response to destroy the pathogens and their toxic molecules (16–18). This field has been extensively studied in the past, and researchers have developed several computational resources like MHCBN, IEDB, Bcipep, etc. (19–21). BepiPred 2.0, Bcepred, Lbtope, IgPred, PEASE, etc. aid in predicting epitopes in humoral-mediated immunity (22–26). ProPred 1 and NetMHCstab-pan help in predicting MHC-I binder, whereas ProPred, MHC2Pred and EpiDOCK predict MHC-II binders (27–32). Furthermore, methods for predicting T-helper epitopes based on cytokines, like $IFN-\gamma$, IL-4 and IL-10 (33–35), have also been explored. Recently, we have published review articles, which provide detailed insights about the immunological resources and *in silico* tools (36, 37). We believe that computational resources in the field of PRRs–PAMPs need revival. How specific PRRs sense the invading pathogens, mechanisms involved in immune response against PAMPs, downstream signaling cascades involved in eliciting immune response, etc. are some questions/challenges that must be clearly inferred to retain better therapeutic strategies against several infectious diseases caused by invading pathogens.

The first version of PRRDB, a comprehensive database of PRRs and their ligands, was developed in 2008 (38). PRRDB aided in developing new resources such as Anti-genDB (39) and PolysacDB (40) as well as facilitated the prediction of pattern receptor recognition families (41). Since 2008, the understanding of innate immunity has been strengthened, and several other pathogen-associated molecules have been studied and discovered. Therefore, there is a necessity to enhance and upgrade the information

about PRRs and their ligands in as much detail as possible. PRRDB 2.0 is an updated comprehensive resource of PRRs and their ligands. The updated version has exhaustive information about receptors such as about their domain, localization, as well as elaborative functions with role, occurrence and sequence information of their ligands. In addition to this, PubChem assays and experimental procedures, while elucidating the PRRs and their ligands as well as their available structures, have been included in the updated version, which were previously lacking.

## Material and methods

### Data collection and compilation

PubMed was mined using keywords 'Pattern recognition receptors' and 'Pathogen-associated molecular patterns', specifically published from 2008 to 2018. The cumulative hits obtained were ∼30 000. We manually screened all the abstracts, and ∼3000 abstracts were selected for further data mining. Final data was obtained from 597 research articles.

### Database architecture and web interface

PRRDB 2.0 was built on an Apache HTTP server (version 2.4.7) installed on a machine with Ubuntu (version 14.04.5) as the operating system. MySQL (version 5.5.55) was used as the back end to manage the data, while HTML5, PHP5, CSS3 and JAVA scripts were used for developing responsive front ends, compatible for mobiles, tablets and desktops. PERL and PHP programming language were used to develop a common interface. The architecture of PRRDB 2.0 is depicted in Figure 2.

## Data content

Primary data contains the sole information extracted from the corresponding research articles, which has been linked under head 'PMID'. Information has been provided in the tabular form for both ligands as well as PRRs. The main fields illustrating about ligand (PAMPs/DAMPs) are (i) Ligand name: represents the name of ligands used in the literature; (ii) Ligand source: describes the actual source or origin of that ligand; (iii) Ligand type: represents the category of ligands such as lipid, peptidoglycan, lipopolysaccharide, protein etc.; (iv) Occurrence of ligand: represents either natural or synthetic occurrence of ligands; (v) Role of ligands: provides extensive information about corresponding ligands' role in activating the immune system. Besides this information, PRRs have been curated under heads such as (vi) Receptor's name: represents the name of PRRs used in the literature; (vii) Receptor source: describes the actual source or origin of that receptor; (viii) Receptor type: represents the major classification of PRRs such as TLRs, CLRs, RLRs, NLRs, etc.; (ix) Localization of receptor: conveys its localization or cell type from which it has been found; (x) Domain: represents the specific domain present within particular PRRs such as Leucine-rich domain in TLR9, lectin domain in CLRs, etc. and (xi) Function: describes the role or function of PRRs in eliciting the innate immune system by signal cascades, when associated with their specific ligands. In addition to this, the experimental procedure or specific assays used in the corresponding literature is also curated under the head 'Assay used'. As a supplement to above information, all the PubChem assays known till date, of that particular PRR, have also been provided by a hyperlink (42). PubChem and Swiss-Prot were mined to curate several important information that was not mentioned in that particular research article, such as sequences of PRRs, as well as their ligands (43). Protein Data Bank (PDB) was also queried to extract the experimentally known structures of PRRs (44). In cases where the structures were not experimentally characterized, we tried to provide the predicted structures of PRRs by virtue of structure prediction methods mainly by PHYRE2 (45).

## Implementation of tools

*Data search*  Basic and advanced search modules have been implemented in PRRDB 2.0 to provide effortless searching. The user can query against the name, source and type of ligand. Similarly, queries can also be put against name, type, source and domain of receptors. The basic search module allows the output customized according to the search query, while the user can give multiple queries simultaneously with Boolean expressions (e.g. AND, OR and NOT) in the advanced search module.

*Data browse*  Data can be browsed by (i) receptors, (ii) ligands, (iii) receptor's source and (iv) ligand's source. Apart from this, the receptor's structure is an additional service to provide extensive structural information of all PRRs known till date. A user can browse all the PRRs by major type of PRRs, notably: (i) CLRs, (ii) Gram-negative binding proteins, (iii) 1,3-$\beta$-D-Glucan pattern recognition protein (GRP), (iv) Lectins, (v) Mannose receptors, (vi) NLRs, (vii) Peptidoglycan recognition proteins, (viii) Rig-like receptors, (ix) Scavenger receptors, (x) TLRs and few others. Browsing by receptor's structure leads to a list of the particular type of PRRs and their all experimentally validated structures with PDBID, description and the source organism.

*Data similarity or alignment*  BLAST (46) and Smith–Waterman algorithm (47) have been implemented to facilitate data similarity-based search. In BLAST module, the user has to submit FASTA format of protein ligands or PAMPs/DAMPs with default or chosen parameters, and the server automatically performs the BLAST search against stored data. Similarly, Smith–Waterman algorithm also performs similarity-based search against ligands.

## Results

### Data analysis

The updated version of PRR database, PRRDB 2.0, holds a total of 2740 entries extracted from 597 research articles. We have performed a major update on PRR and their ligands. We have incorporated 2374 new receptors in addition to 353 receptors from the first version of PRRDB, making the total tally of 2727 total PRRs. Similarly, PRRDB 2.0 contains 2197 total ligands studied in the last 10 years as well as 353 ligands from the previous database, making the total count of ligands to 2550. In summary, PRRDB 2.0 comprises 2740 entries providing information about 2727 total, 467 unique PRRs and 2550 total and 827 unique ligands.

Figure 3A shows the major type of receptors available in PRRDB 2.0. TLRs are the most prominent and widely studied among PRRs, so 62% of entries provide information regarding them. Other entries provide information about NLRs—241, CLRs—135, Scavenger—88, Syk-coupled CLRs—63, RLRs—40, Mannose receptor—33, PGRPs—25 and RAGE—22.

It is a well-known fact that TLRs and CLRs are membrane-bound pathogen receptors and around 72% of the PRRs curated in PRRDB 2.0 fall in this membrane-bound category as compared to 10% of the cytoplasmic PRRs. Similarly, Figure 3B represents the graphical distribution of entries of different ligand types that include 496 entries about nucleic acids, 353 for protein-type ligands,
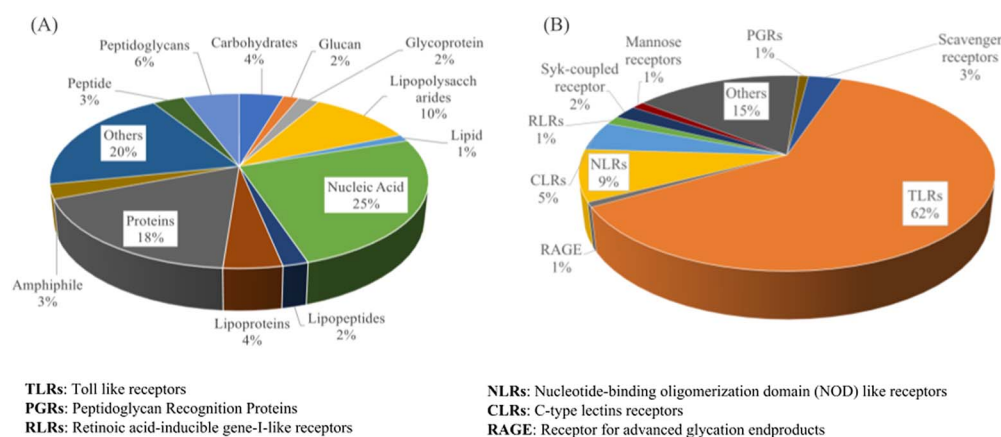
**Figure 3.** Distribution of various categories of PRRs (**A**) and ligands (**B**) stored in PRRDB 2.0.
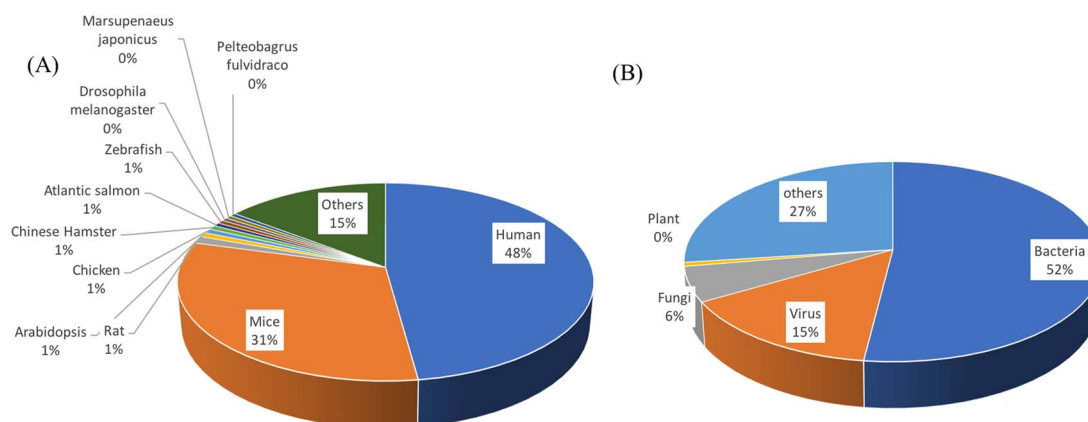


**Figure 4.** Distribution of source of PRRs (**A**) and ligands (**B**).

as well as lipopolysaccharides—207, peptidoglycan—111, carbohydrates—88, lipoproteins—85, glycoprotein—41, lipopeptide—37, glucan—31, lipid—25, polysaccharide—16, amphiphile—53 and few others. Most of the ligands (79%) stored in PRRDB 2.0 have natural origins. Data analysis also revealed that sources of most of the PRRs are human (48%) and mice (31%). These PRRs bind to ligands originated mostly from bacteria (52%), viruses (15%) and fungi (6%) (Figure 4).

## Comparison with the previous version

PRRDB was developed in 2008 and comprises two different tables; one of PRRs containing 491 entries provides information such as receptor's name, source organism, sequence and their length, family and type. Another table is of ligands containing 266 entries, providing information like name, source, class of the ligand, origin and its receptor (38). In the updated version, we have incorporated more information about each PRR as well as ligand, so the final data consist of 2740 entries. We have tried to provide the role of ligand in activating the immune system in addition to its name, source, type and origin. Similarly, the updated version

contains extensive information about each PRR such as its name, source, type, sequence and its length, localization and domain as well as its function. A comparative statistics is shown in Table 1. It shows only the major update of the database. The comprehensive statistics have been given in Supplementary Table S1. The experimental procedure or assay is also curated in the updated version. To add more to this, data has been linked with Swiss-Prot, PubChem and PDB for maximum information.

## Utility of database

PRRDB 2.0 can be used to get exhaustive information about any PRR on a single platform. For example, if a user is interested in TLR 4, which mostly recognizes lipopolysaccharides associated with gram-negative bacteria, one has to type TLR4 in the search box given at basic search page and check the name of receptor, as shown in Figure 5A. By a simple click on the search button, one will be directed towards a list of 475 entries stored in PRRDB 2.0, which are differentiated with a unique ID, as shown in Figure 5B. A click on each ID will direct to a detailed display page, providing all the information as well as hyperlinks to

**Table 1.** Comparison of data in two versions of PRRDB database, total entries/information in PRRDB and PRRDB 2.0 is shown

| Overall information | | | Entries for different types of receptors | | |
|---|---|---|---|---|---|
| Field/Information | PRRDB | PRRDB 2.0 | Receptors | PRRDB | PRRDB 2.0 |
| Total receptors | 353 | 2727 | TLR* | 185 | 1737 |
| Total ligands | 354 | 2550 | CLR* | 27 | 135 |
| Sequence of Receptors | 221 | 1784 | NLR* | 15 | 241 |
| Sequence of ligands | 241 | 1583 | Mannose | 26 | 33 |
| | | | Scavenger | 53 | 88 |

| Entries for major types of ligands | | | Entries for major sources of receptors | | |
|---|---|---|---|---|---|
| Field/Information | PRRDB | PRRDB 2.0 | Receptors | PRRDB | PRRDB 2.0 |
| Peptide | 15 | 62 | Human | 146 | 1092 |
| Nucleic acid | 68 | 496 | Mice | 102 | 717 |
| PAMP* | 54 | 376 | Chicken | 0 | 17 |
| DAMP* | 0 | 247 | Hamster | 15 | 16 |
| Protein | 60 | 353 | Rat | 3 | 27 |
| LPS* | 16 | 207 | Zebrafish | 0 | 13 |
| Peptidoglycan | 8 | 111 | *Arabidopsis* | 1 | 17 |
| Carbohydrates | 37 | 88 | | | |

*TLR: Toll-like receptors, CLR: C-type lectin receptor, NLR: nucleotide-binding oligomerization domain (NOD)-like receptor, PAMP: pathogens-associated molecular pattern, DAMP: damage-associated molecular pattern, LPS: lipopolysaccharide.

PubChem, PubMed and Swiss-Prot. In addition, sequence of TLR4 is also available in FASTA format as depicted in Figure 5C. Furthermore, all 27 experimentally proven structures about TLR4 and its complexes stored in PDB can also be viewed by user in PRRDB 2.0.

## Limitation and update

We have tried to provide the complete information of PRRs and their ligands, but sequences of some PRRs and ligands are neither available in the literature nor in any other chemical database. The data is manually curated and thoroughly checked to minimize the error, but claiming perfect accuracy would be unfair due to human errors that may have crept in. The first version was released 10 years ago, but we will try to update the database frequently, preferably after 3 years.

## Discussion

Innate immune system acts as the first line of defense for host and is responsible for pathogen recognition and elicitation of proinflammatory responses against invading pathogens in the initial phases of infection, whereas, in the late phase, adaptive immune system eliminates the pathogen and generates immunological memory. As elaborated in the introduction, the repertoire of PRRs is very vast and recognizes various classes of pathogens. A surprising aspect is the similar mechanism of host PRRs in recognizing a wide range of microbes in different life cycles and with different biochemical compositions. Another astonishing fact is that all classes of pathogens are sensed by more than one type of PRRs through various ligands and lead to a rapid proinflammatory response through various intracellular signal cascades (14).

Inflammation plays an important role in various infectious diseases and autoimmunity; therefore, targeting signal cascades and process of activation of the innate immune system holds a tremendous medical promise (48, 49). The concept of pattern recognition by a large family of receptors redefined the specificity and complexity of innate immune system, which was largely ignored in the early immunological studies, but emerged as a hotspot in medical interventions. Immunotherapy and vaccine have always proved to be a rescue for mankind; as a result of which, several deadly diseases have been eradicated in the past (50). Various computational tools have aided in immunotherapy and vaccine designing; a few studies have also focused on innate immune system to design therapeutics (33, 51–55). In the last decade, several databases and *in silico* tools related to therapeutic entities have been developed (56–62) and successfully used in designing newer therapeutic entities (63, 64).

Despite immense progress in innate immune-related research in the last few decades, uncertainty still exists. TLRs are the most studied PRRs, but cytoplasmic PRRs also play a major role in accumulating various immune responses, which needs more investigation. Similarly, other PRRs such as mannose receptors, scavenger receptors, as well as few secreted PRRs need to be explored further. Understanding of cross-talk between various PRRs needs
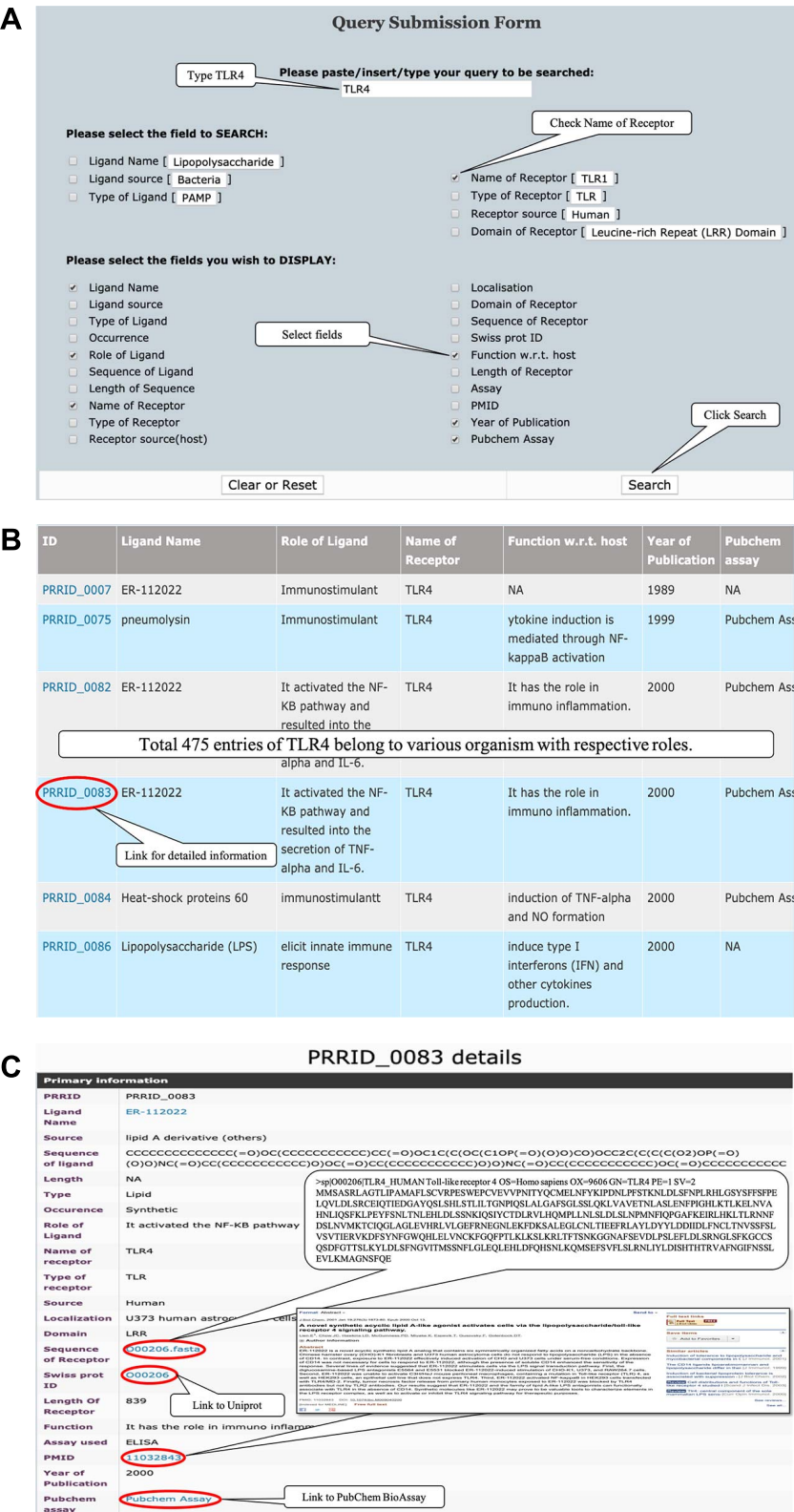
**Figure 5.** (**A**) Representative screenshot of PRRDB 2.0 demonstrating the submission of a query in simple search page. (**B**) Screenshot of PRRDB 2.0 demonstrating the result page after submission of a query in simple search page. (**C**) Screenshot of PRRDB 2.0 demonstrating the detailed information page after the result page.

better insight. We believe that PRRDB 2.0 will aid in getting all the information and queries discovered in the past.

## Conclusion

PRRDB 2.0, comprising more than 2700 entries, provides better coverage of all the PRRs and their ligands studied till now. We have improved the data coverage by incorporating additional fields and also highlighting the role and specificity of PRRs and ligands in eliciting the immune response. The hyper linkage with Swiss-Prot, PDB and PubChem will provide maximum information at a single place. We believe that the updated version will be very helpful to the scientific community.

### Availability

PRRDB 2.0 is freely available at https://webs.iiitd.edu.in/raghava/prrdb2/ as a user-friendly, display compatible interface. The previous version can be accessed at http://crdd.osdd.net/raghava/prrdb/.

## Statement of Ethics

The authors have no ethical conflicts to disclose.

## Author contributions

D.K., S.P. and N.S. manually collected and curated all the data, as well as experimentally verified structures. D.K. predicted the structures. D.K., S.P., N.S. and S.S.U. developed the web interface. S.S.U. and G.P.S.R. prepared the manuscript. G.P.S.R. conceived the idea, planned and coordinated the entire project.

## Supplementary data

Supplementary data are available at *Database* Online.

## Acknowledgement

## Funding

*Conflict of interest*. None declared.

## References

1. Riera Romo,M., Pérez-Martínez,D. and Castillo Ferrer,C. (2016) Innate immunity in vertebrates: an overview. *Immunology*, **148**, 125–139.

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

3. Takeuchi,O. and Akira,S. (2010) Pattern recognition receptors and inflammation. *Cell*, **140**, 805–820.

4. Tartey,S. and Takeuchi,O. (2017) Pathogen recognition and Toll-like receptor targeted therapeutics in innate immune cells. *Int. Rev. Immunol.*, **36**, 57–73.

5. Kawai,T. and Akira,S. (2010) The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. *Nat. Immunol.*, **11**, 373–384.

6. Hoving,J.C., Wilson,G.J. and Brown,G.D. (2014) Signalling C-type lectin receptors, microbial recognition and immunity. *Cell. Microbiol.*, **16**, 185–194.

7. Franchi,L., Warner,N., Viani,K. *et al.* (2009) Function of Nod-like receptors in microbial recognition and host defense. *Immunol. Rev.*, **227**, 106–128.

8. Loo,Y.-M. and Gale,M. (2011) Immune signaling by RIG-I-like receptors. *Immunity*, **34**, 680–692.

9. Tang,J., Lin,G., Langdon,W.Y. *et al.* (2018) Regulation of C-type lectin receptor-mediated antifungal immunity. *Front. Immunol.*, **9**, 123.

10. Kim,Y.K., Shin,J.S. and Nahm,M.H. (2016) NOD-like receptors in infection, immunity, and diseases. *Yonsei Med. J.*, **57**, 5–14.

11. Onoguchi,K., Yoneyama,M. and Fujita,T. (2011) Retinoic acid-inducible gene-I-like receptors. *J. Interferon Cytokine Res.*, **31**, 27–31.

12. Paveley,R.A., Aynsley,S.A., Turner,J.D. *et al.* (2011) The Mannose Receptor (CD206) is an important pattern recognition receptor (PRR) in the detection of the infective stage of the helminth Schistosoma mansoni and modulates IFNγ production. *Int. J. Parasitol.*, **41**, 1335–1345.

13. Taghavi,M., Khosravi,A., Mortaz,E. *et al.* (2017) Role of pathogen-associated molecular patterns (PAMPS) in immune responses to fungal infections. *Eur. J. Pharmacol.*, **808**, 8–13.

14. Mogensen,T.H. (2009) Pathogen recognition and inflammatory signaling in innate immune defenses. *Clin. Microbiol. Rev*, **22**, 240–273, Table of Contents.

15. Fearon,D.T. and Locksley,R.M. (1996) The instructive role of innate immunity in the acquired immune response. *Science*, **272**, 50–53.

16. Iwasaki,A. and Medzhitov,R. (2015) Control of adaptive immunity by the innate immune system. *Nat. Immunol.*, **16**, 343–353.

17. Jain,A. and Pasare,C. (2017) Innate control of adaptive immunity: beyond the three-signal paradigm. *J. Immunol.*, **198**, 3791–3800.

18. Palm,N.W. and Medzhitov,R. (2009) Pattern recognition receptors and control of adaptive immunity. *Immunol. Rev.*, **227**, 221–233.

19. Bhasin,M., Singh,H. and Raghava,G.P.S. (2003) MHCBN: a comprehensive database of MHC binding and non-binding peptides. *Bioinformatics*, **19**, 665–666.

20. Vita,R., Mahajan,S., Overton,J.A. *et al.* (2019) The immune epitope database (IEDB): 2018 update. *Nucleic Acids Res.*, **47**, D339–D343.

21. Saha,S., Bhasin,M. and Raghava,G.P. (2005) Bcipep: a database of B-cell epitopes. *BMC Genomics*, **6**, 79.

22. Jespersen,M.C., Peters,B., Nielsen,M. *et al.* (2017) BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Res.*, **45**, W24–W29.

23. Saha,S. and Raghava,G.P.S. (2004) BcePred: prediction of continuous B-cell epitopes in antigenic sequences using physicochemical properties. In *International Conference on Artificial Immune Systems*. Springer, Berlin, Heidelberg, pp. 197–204.

24. Lian,Y., Huang,Z.C., Ge,M. *et al.* (2015) An improved method for predicting linear B-cell epitope using deep maxout networks. *Biomed. Environ. Sci.*, **28**, 460–463.

25. Gupta,S., Ansari,H.R., Gautam,A. *et al.* (2013) Identification of B-cell epitopes in an antigen for inducing specific class of antibodies. *Biol. Direct*, **8**, 27.

26. Sela-Culang,I., Ofran,Y. and Peters,B. (2015) Antibody specific epitope prediction-emergence of a new paradigm. *Curr. Opin. Virol.*, **11**, 98–102.

27. Singh,H. and Raghava,G.P.S. (2003) ProPred1: prediction of promiscuous MHC Class-I binding sites. *Bioinformatics*, **19**, 1009–1014.

28. Rasmussen,M., Fenoy,E., Harndahl,M. *et al.* (2016) Pan-specific prediction of peptide–MHC class I complex stability, a correlate of T cell immunogenicity. *J. Immunol.*, **197**, 1517–1524.

29. Bhasin,M. and Raghava,G.P.S. (2007) A hybrid approach for predicting promiscuous MHC class I restricted T cell epitopes. *J. Biosci.*, **32**, 31–42.

30. Singh,H. and Raghava,G.P. (2001) ProPred: prediction of HLA-DR binding sites. *Bioinformatics*, **17**, 1236–1237.

31. Lata,S., Bhasin,M. and Raghava,G.P.S. (2007) Application of machine learning techniques in predicting MHC binders. *Immunoinformatics*, **409**, 201–215.

32. Atanasova,M., Patronov,A., Dimitrov,I. *et al.* (2013) EpiDOCK: a molecular docking-based tool for MHC class II binding prediction. *Protein Eng. Des. Sel.*, **26**, 631–634.

33. Nagpal,G., Usmani,S.S., Dhanda,S.K. *et al.* (2017) Computer-aided designing of immunosuppressive peptides based on IL-10 inducing potential. *Sci. Rep.*, **7**, 42851.

34. Dhanda,S.K., Gupta,S., Vir,P. *et al.* (2013) Prediction of IL4 inducing peptides. *Clin. Dev. Immunol.*, **2013**, 263952.

35. Dhanda,S.K., Vir,P. and Raghava,G.P. (2013) Designing of interferon-gamma inducing MHC class-II binders. *Biol. Direct*, **8**, 30.

36. Dhanda,S.K., Usmani,S.S., Agrawal,P. *et al.* (2017) Novel in silico tools for designing peptide-based subunit vaccines and immunotherapeutics. *Brief. Bioinform.*, **18**, 467–478.

37. Usmani,S.S., Kumar,R., Bhalla,S. *et al.* (2018) In silico tools and databases for designing peptide-based vaccine and drugs. *Adv. Protein Chem. Struct. Biol.*, **112**, 221–263.

38. Lata,S. and Raghava,G.P.S. (2008) PRRDB: a comprehensive database of pattern-recognition receptors and their ligands. *BMC Genomics*, **9**, 180.

39. Ansari,H.R., Flower,D.R. and Raghava,G.P.S. (2010) AntigenDB: an immunoinformatics database of pathogen antigens. *Nucleic Acids Res.*, **38**, D847–D853.

40. Aithal,A., Sharma,A., Joshi,S. *et al.* (2012) PolysacDB: a database of microbial polysaccharide antigens and their antibodies. *PLoS One*, **7**, e34613.

41. Gao,Q.-B., Zhao,H., Ye,X. *et al.* (2012) Prediction of pattern recognition receptor family using pseudo-amino acid composition. *Biochem. Biophys. Res. Commun.*, **417**, 73–77.

42. Kim,S., Chen,J., Cheng,T. *et al.* (2019) PubChem 2019 update: improved access to chemical data. *Nucleic Acids Res.*, **47**, D1102–D1109.

43. Bateman,A., Martin,M.J., O'Donovan,C. *et al.* (2017) UniProt: the universal protein knowledgebase. *Nucleic Acids Res.*, **45**, D158–D169.

44. Burley,S.K., Berman,H.M., Bhikadiya,C. *et al.* (2019) RCSB Protein Data Bank: biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy. *Nucleic Acids Res.*, **47**, D464–D474.

45. Kelley,L.A., Mezulis,S., Yates,C.M. *et al.* (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.*, **10**, 845–858.

46. Altschul,S.F., Gish,W., Miller,W. *et al.* (1990) Basic local alignment search tool. *J. Mol. Biol.*, **215**, 403–410.

47. Smith,T.F. and Waterman,M.S. (1981) Identification of common molecular subsequences. *J. Mol. Biol.*, **147**, 195–197.

48. Mahajan,S., Saini,A., Chandra,V. *et al.* (2015) Nuclear receptor Nr4a2 promotes alternative polarization of macrophages and confers protection in sepsis. *J. Biol. Chem.*, **290**, 18304–18314.

49. Bhagyaraj,E., Nanduri,R., Saini,A. *et al.* (2016) Human xenobiotic nuclear receptor PXR augments mycobacterium tuberculosis survival. *J. Immunol.*, **197**, 244–255.

50. Greenwood,B. (2014) The contribution of vaccination to global health: past, present and future. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, **369**, 20130433.

51. Chaudhary,K., Nagpal,G., Dhanda,S.K. *et al.* (2016) Prediction of Immunomodulatory potential of an RNA sequence for designing non-toxic siRNAs and RNA-based vaccine adjuvants. *Sci. Rep.*, **6**, 20678.

52. Nagpal,G., Chaudhary,K., Agrawal,P. *et al.* (2018) Computer-aided prediction of antigen presenting cell modulators for designing peptide-based vaccine adjuvants. *J. Transl. Med.*, **16**, 181.

53. Nagpal,G., Usmani,S.S. and Raghava,G.P.S. (2018) A web resource for designing subunit vaccine against major pathogenic species of bacteria. *Front. Immunol.*, **9**, 2280.

54. Dhanda,S.K., Vir,P., Singla,D. *et al.* (2016) A web-based platform for designing vaccines against existing and emerging strains of mycobacterium tuberculosis. *PLoS One*, **11**, e0153771.

55. Nagpal,G., Gupta,S., Chaudhary,K. *et al.* (2015) VaccineDA: prediction, design and genome-wide screening of oligodeoxynucleotide-based vaccine adjuvants. *Sci. Rep.*, **5**, 12478.

56. Kumar,V., Agrawal,P., Kumar,R. *et al.* (2018) Prediction of cell-penetrating potential of modified peptides containing natural and chemically modified residues. *Front. Microbiol.*, **9**, 725.

57. Usmani,S.S., Bedi,G., Samuel,J.S. *et al.* (2017) THPdb: database of FDA-approved peptide and protein therapeutics. *PLoS One*, **12**, e0181748.

58. Singh,S., Chaudhary,K., Dhanda,S.K. *et al.* (2016) SATPdb: a database of structurally annotated therapeutic peptides. *Nucleic Acids Res.*, **44**, D1119–D1126.

59. Bhalla,S., Verma,R., Kaur,H. *et al.* (2017) CancerPDF: a repository of cancer-associated peptidome found in human biofluids. *Sci. Rep.*, **7**, 1511.

60. Usmani,S.S., Kumar,R., Kumar,V. *et al.* (2018) AntiTbPdb: a knowledgebase of anti-tubercular peptides. *Database (Oxford).*, **2018**, bay025, 1–8.

61. Agrawal,P., Bhalla,S., Usmani,S.S. *et al.* (2016) CPPsite 2.0: a repository of experimentally validated cell-penetrating peptides. *Nucleic Acids Res.*, **44**, D1098–D1103.

62. Usmani,S.S., Bhalla,S. and Raghava,G.P.S. (2018) Prediction of antitubercular peptides from sequence information using ensemble classifier and hybrid features. *Front. Pharmacol.*, **9**, 954.

63. Gautam,A., Sharma,M., Vir,P. *et al.* (2015) Identification and characterization of novel protein-derived arginine-rich cell-penetrating peptides. *Eur. J. Pharm. Biopharm.*, **89**, 93–106.

64. Gautam,A., Chaudhary,K., Kumar,R. *et al.* (2015) Computer-aided virtual screening and designing of cell-penetrating peptides. *Methods Mol. Biol.*, **1324**, 59–69.