Computers in Biology and Medicine 169 (2024) 107886

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

RV*victor*: Virus RNA-directed molecular interactions for RNA virus infection

Kuerbannisha Amahong [a,b,1], Wei Zhang [a,b,1], Yuhong Liu [c,1], Teng Li [a,1], Shijie Huang [a,b], Lianyi Han [d,**], Lin Tao [c,*], Feng Zhu [a,b,***]

[a] College of Pharmaceutical Sciences, The Second Affiliated Hospital, Zhejiang University School of Medicine, Zhejiang University, Hangzhou, 310058, China
[b] Innovation Institute for Artificial Intelligence in Medicine of Zhejiang University, Alibaba-Zhejiang University Joint Research Center of Future Digital Healthcare, Hangzhou, 330110, China
[c] Key Laboratory of Elemene Class Anti-cancer Chinese Medicines, School of Pharmacy, Hangzhou Normal University, Hangzhou, 311121, China
[d] Greater Bay Area Institute of Precision Medicine (Guangzhou), School of Life Sciences, Fudan University, Shanghai, 315211, China

ARTICLE INFO

Keywords:
RNA virus
Virus RNA
Host protein
RNA-Directed interactions
Enrichment analysis

ABSTRACT

RNA viruses are major human pathogens that cause seasonal epidemics and occasional pandemic outbreaks. Due to the nature of their RNA genomes, it is anticipated that virus's RNA interacts with host protein (INTPRO), messenger RNA (INTmRNA), and non-coding RNA (INTncRNA) to perform their particular functions during their transcription and replication. In other words, thus, it is urgently needed to have such valuable data on virus RNA-directed molecular interactions (especially INTPROs), which are highly anticipated to attract broad research interests in the fields of RNA virus translation and replication. In this study, a new database was constructed to describe the virus RNA-directed interaction (INTPRO, INTmRNA, INTncRNA) for RNA virus (RV*victor*). This database is unique in ***a***) unambiguously characterizing the interactions between viruses RNAs and host proteins, ***b***) providing, for the first time, the most systematic RNA-directed interaction data resources in providing clues to understand the molecular mechanisms of RNA viruses' translation, and replication, and ***c***) in RV*victor*, comprehensive enrichment analysis is conducted for each virus RNA based on its associated target genes/proteins, and the enrichment results were explicitly illustrated using various graphs. We found significant enrichment of a suite of pathways related to infection, translation, and replication, e.g., HIV infection, coronavirus disease, regulation of viral genome replication, and so on. Due to the devastating and persistent threat posed by the RNA virus, RV*victor* constructed, for the first time, a possible network of cross-talk in RNA-directed interaction, which may ultimately explain the pathogenicity of RNA virus infection. The knowledge base might help develop new anti-viral therapeutic targets in the future. It's now free and publicly accessible at: https://idrblab.org/rvvictor/.

1. Introduction

RNA viruses have led to global burden of diseases with substantial mortality over the past several years, the most devastating ones of which include: *human immunodeficiency virus* (HIV), *severe acute respiratory syndrome coronavirus 2* (SARS-CoV-2), *ebola virus* (EBOV), and so on [1–3]. The RNAs of these viruses are essential in both translation and replication during their infections by interacting with the protein, messenger RNA, and non-coding RNA (as illustrated in Fig. 1) of the infected host [4–6]. In contrast to the extensive research on their interactions with host messenger RNA (INTmRNA) and non-coding RNA (INTncRNA) [7–9], few studies have been performed to reveal interactions between virus RNA and host protein (INTPROs) over the past few decades [10]. As reported, INTPROs are critical for virus translation by recruiting diverse translation-essential host proteins, which are highly anticipated to offer new targets facilitating drug discovery and repurposing [11]. Moreover, all three types of virus RNA-directed interactions (INTPRO, INTmRNA, INTncRNA) should be collectively considered, since they are found to function as a complex network system in determining the molecular mechanism underlying the translation

* Corresponding author. School of Pharmacy, Hangzhou Normal University, Hangzhou, China.
** Corresponding author. School of Life Sciences, Fudan University, Shanghai, China.
*** Corresponding author. College of Pharmaceutical Sciences, Zhejiang University, Hangzhou, China.
E-mail addresses: hanlianyi@ipm-gba.org.cn (L. Han), taolin@hznu.edu.cn (L. Tao), zhufeng@zju.edu.cn (F. Zhu).
[1] These authors contributed equally to this work as co-first authors.

https://doi.org/10.1016/j.compbiomed.2023.107886
Received 6 August 2023; Received in revised form 14 December 2023; Accepted 18 December 2023
Available online 20 December 2023

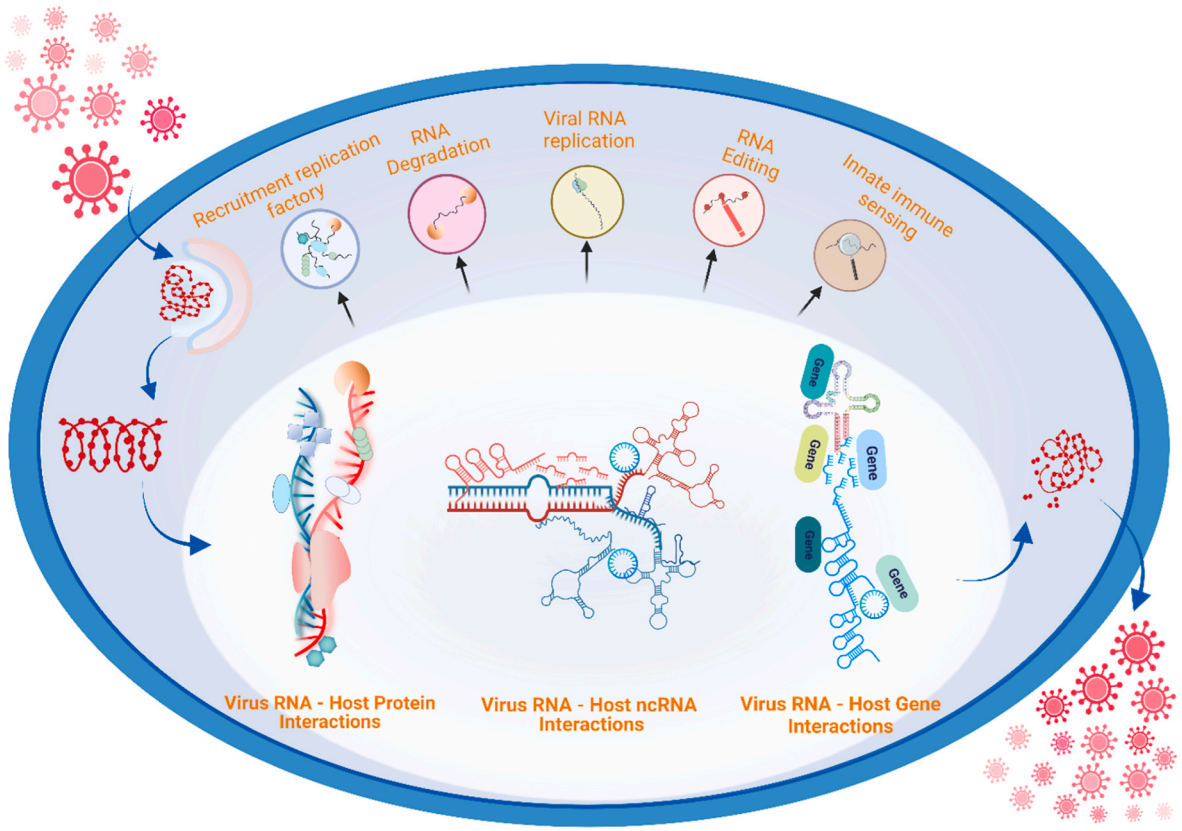


Fig. 1. Schematic representations of the RNA virus's life cycle, highlighting the RNA-directed interactions (INTPRO, INTmRNA, INTncRNA) are involved in key steps of a virus's life cycle such as translation and replication. Furthermore, RNA-directed interactions are important and emerging themes in biology of RNA viruses that have been shown to support or restrict virus infection. RNA-directed interactions can regulate virus RNA fate; however, emerging data suggest that virus RNA may, in some instances, regulate the function of host molecules, many of which have been linked to the cellular antiviral response, while others are associated with the virus replication cycles such as recruitment to the viral replication factory, degradation of RNA, splicing, and RNA editing as indicated by the colored circles and descriptions in orange above.

and replication of RNA virus [12–14]. Thus, it is urgently needed to have such valuable data on virus RNA-directed molecular interactions (especially INTPROs), which are highly expected to attract broad research interests from audiences who study the translation and replication of RNA viruses.

Until now, there has been a proliferation of databases concerning RNA viruses. Some of them provide general information on virus genome, species classification, sequence variation/annotation, and structural data of different RNA viruses, such as IMG/VR [15], BV-BRC [16], OrthoDB [17], pVOGs [18], Stanford HIV DB [19], and many others [20–28]. Some others contain molecular interaction data of RNA virus, which include (*a*) those describing protein-protein interaction between RNA virus and host, such as HIV Interaction [29], VirusMentha [30], VirHostNet [31], and VTcomplex [32], (*b*) those illustrating virus RNAs' interactions with host RNAs, such as miRSponge [33] focusing on INTmRNA data and ViRBase [34] specializing in INTncRNA data. However, none of them provide any interaction data between virus RNA & host protein (INTPRO), and none of them describe all three types of virus RNA-directed interaction data (INTPRO, INTmRNA, INTncRNA). Thus, it is highly demanded to construct a database that can cover all aspects of virus RNA-directed molecular interaction data to systematically describe the mechanisms underlying RNA virus' translation & replication.

In this study, a new database entitled "virus RNA-directed molecular interactions for RNA virus infection (RV*victor*)" was therefore constructed to describe virus RNA-directed molecular interaction for RNA virus infection from three different perspectives: INTPRO, INTmRNA, and INTncRNA. Compare with those existing databases, RV*victor* was unique in ***First***, providing, for the first time, the interaction data between virus RNA & host protein, a total of 16,223 INTPROs were manually collected and explicitly provided in RV*victor*, which came from 117 virus RNAs of 70 virus/strains and 3829 proteins of 11 host organisms. ***Second***, offering the most comprehensive RNA-directed interaction data of RNA virus' translation and replication among existing databases by collectively describing all three types of virus RNA-directed molecular interactions. A total of 101,345 pairs of interactions (94,875 pairs of INTmRNA and 6470 pairs of INTncRNA) were manually collected and described in our RV*victor*, which originated from 52 virus/strains and 17 host organisms. ***Third***, enabling the functional annotations of all virus RNAs in RV*victor* based on the enrichment analyses of their gene/protein targets. All enrichment results were explicitly illustrated using various tables & graphs of signaling pathways, gene ontologies, and disease classes. Moreover, both the virus species and host organisms collected to RV*victor* were diverse, which include a variety of well-known RNA viruses (such as HIV, SARS-CoV-2, HCV, Ebola, Dengue, Zika, West Nile, Influenza A, Epstein Barr, and Sindbis) and various kinds of popular organisms (such as human, monkey, mouse, rat, pig, chicken, and bovine).

Overall, the RV*victor* database was introduced to ***a***) explicitly describe the interaction data between virus RNAs and host proteins, ***b***) systematically provide the most comprehensive RNA-directed interaction data of RNA virus's translation and replication, and ***c***) enable the functional annotations of all virus RNAs in RV*victor* based on the enrichment analyses of their target gene/proteins. Given the devastating global burden of infectious disease caused by RNA viruses, the data shown in RV*victor* provide the most comprehensive network of RNA-

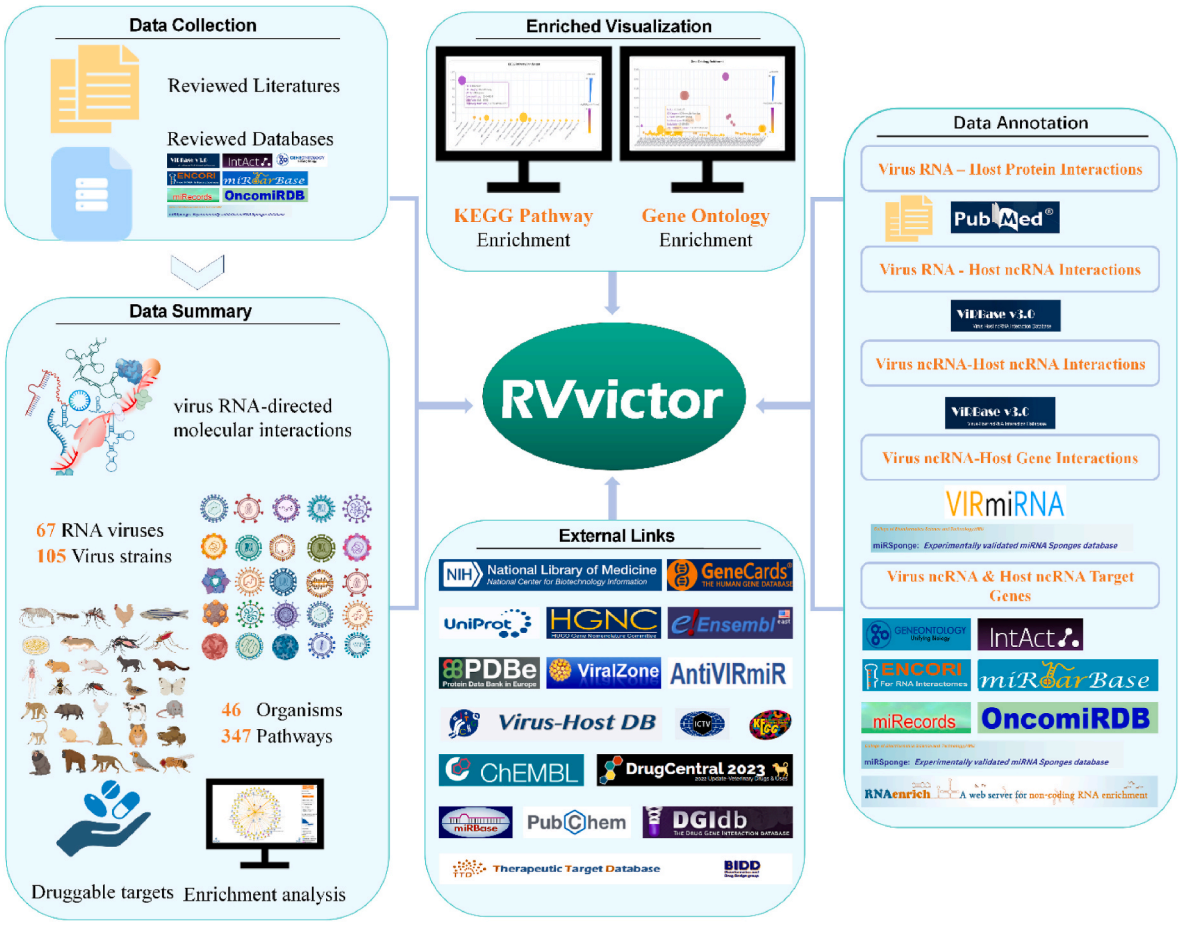


Fig. 2. The overall design, primary data components, and the corresponding statistics of RV*victor*. RV*victor* contains 117,568 pairs of RNA-directed interaction entries between 105 viruses/strains and 46 host organisms. It is classified into three categories: INTPRO data were manually curated, while INTmRNA, and INTncRNA data were extracted from ViRBase and miRSponge databases. Furthermore, target genes of these host ncRNAs were extracted from databases such as mirTarBase, miRSponge, starBase, miRecords, oncomiRDB, Gene Ontology, and InAct, etc. Afterward, enrichment analysis was performed for these host proteins, host mRNA, and target genes of host ncRNA through Gene Ontology and KEGG, and the results were visualized by Apache ECharts. In addition, RV*victor* provides basic general information about viruses and host molecules, as well as external linkage to other well-established databases.

directed interaction complex network between viral RNA and host molecules, thus facilitating the identification of novel therapeutic targets for drug discovery/repurposing.

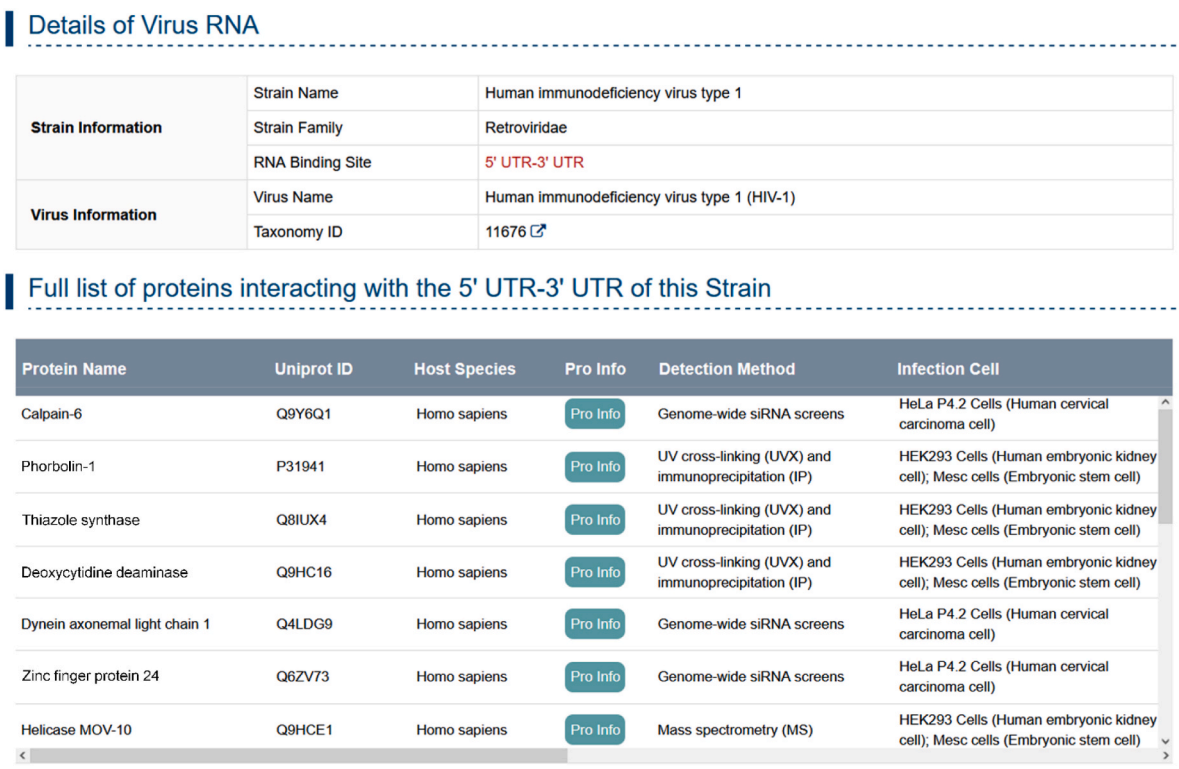
2. Materials and methods

2.1. Data collection, curation, and processing

The data of virus RNA-directed interactions (INTPRO, INTmRNA, INTncRNA) in RV*victor* were collected from experimental-supported literature and databases, including 16,223 INTPROs, 94,875 INTmRNA, and 6470 INTncRNA entries involving 107 virus/strains and 46 host organisms. Respectively, to obtain INTPRO, a comprehensive literature review on the interactions between virus RNA and host protein was conducted by PubMed searching using the keywords of "RNA virus interaction", "Virus RNA binding", "HIV RNA host protein", "Ebola RNA host RBP", "Dengue virus-host interactions", "Zika host infection", "Influenza A host factors" and so on. For the INTmRNA and INTncRNA data, corresponding experimentally validated data were gathered from a thorough review of comprehensive published databases. Furthermore, the general information about virus, virus RNA region, mRNA/protein, and ncRNA were recorded, and a variety of experimental data were collected and described, which contained virus infection time, infection cells, cell-originated tissue, detection methods, interaction type, binding type, and so on. On the whole, RV*victor* as a valuable source of RNA-directed interaction data which will then aid in the discovery of new antiviral drug targets (as shown in Fig. 2).

2.2. Detailed description of each virus and host moleculars

Moreover, the additional reference data for virus RNAs and host proteins and their corresponding interactions were systematically collected and provided in the RV*victor* database. For a *virus RNA*, its referencing data included virus name, strains name, strains family, RNA binding site, and taxonomy ID [35]. In the case of a *host protein*, a variety of reference data was given, which included protein name, protein family, gene name, EC number (if available), subcellular location, sequence, UniProt ID [36], gene ID [37], Ensembl ID [38], HGNC ID [39]. In addition, the downloadable 2D and 3D structures of these proteins from PDB [40] and AlphaFold [41] were also collected. For a *host mRNA*, RV*victor* provides mRNA name, host species, gene ID [37], Ensembl ID [38], and gene card ID [42]. For a *virus ncRNA*, INTncRNA name, ncRNA category, miRbase ID [43], and AntiVIRmiR [44]. Moreover, for a *host ncRNA*, ncRNA name, ncRNA category, host species, lncRNA ID, miRbase ID [43], AntiVIRmiR [44], and Ensembl ID [38] are provided. Detailed information about the interacting viral RNA or host proteins, host mRNAs, and host ncRNAs can be found by clicking the corresponding blue button, "mRNA Info", "HncRNA Info" or "VncRNA Info". All of this data were fully downloadable and could be viewed on the RV*victor*.



Details of Virus RNA

Strain Information	Strain Name	Human immunodeficiency virus type 1
	Strain Family	Retroviridae
	RNA Binding Site	5' UTR-3' UTR
Virus Information	Virus Name	Human immunodeficiency virus type 1 (HIV-1)
	Taxonomy ID	11676

Full list of proteins interacting with the 5' UTR-3' UTR of this Strain

Protein Name	Uniprot ID	Host Species	Pro Info	Detection Method	Infection Cell
Calpain-6	Q9Y6Q1	Homo sapiens	Pro Info	Genome-wide siRNA screens	HeLa P4.2 Cells (Human cervical carcinoma cell)
Phorbolin-1	P31941	Homo sapiens	Pro Info	UV cross-linking (UVX) and immunoprecipitation (IP)	HEK293 Cells (Human embryonic kidney cell); Mesc cells (Embryonic stem cell)
Thiazole synthase	Q8IUX4	Homo sapiens	Pro Info	UV cross-linking (UVX) and immunoprecipitation (IP)	HEK293 Cells (Human embryonic kidney cell); Mesc cells (Embryonic stem cell)
Deoxycytidine deaminase	Q9HC16	Homo sapiens	Pro Info	UV cross-linking (UVX) and immunoprecipitation (IP)	HEK293 Cells (Human embryonic kidney cell); Mesc cells (Embryonic stem cell)
Dynein axonemal light chain 1	Q4LDG9	Homo sapiens	Pro Info	Genome-wide siRNA screens	HeLa P4.2 Cells (Human cervical carcinoma cell)
Zinc finger protein 24	Q6ZV73	Homo sapiens	Pro Info	Genome-wide siRNA screens	HeLa P4.2 Cells (Human cervical carcinoma cell)
Helicase MOV-10	Q9HCE1	Homo sapiens	Pro Info	Mass spectrometry (MS)	HEK293 Cells (Human embryonic kidney cell); Mesc cells (Embryonic stem cell)

Fig. 3. A typical RV*victor* page for virus RNA describes a comprehensive list of host proteins that interacted with this RNA (scrolling up and down to get more information). Detailed experimental information was provided and explicitly discussed, which included the virus infection time, infection cell, cell-originated tissue, detection methods, interaction types, interaction binding type, and so on (scrolling right and left to get this information). All the interactions were validated using diverse living systems including 31 cell lines from 11 tissues and various model organisms. Detailed information on the interacting proteins can be found by clicking the blue button.

2.3. Functional enrichment analysis and visualization of virus RNA-directed interactions

To understand the purpose of virus RNA-directed interactions, we perform enrichment analysis for mRNA/protein targets interacting with each virus RNA as follows. The first step was to match the mRNA/protein targets of these host ncRNAs with the query databases such as mirTarBase [45], miRSponge [33], starBase [46], miRecords [47], Gene Ontology [48], InAct [49], etc., through comprehensive database review and organizing the data. Once the host proteins, host mRNA, and query ncRNA target mRNAs were obtained, the second step was to perform an enrichment analysis through enrichment databases, such as Gene Ontology [48] and KEGG [50]. Finally, rich results will be described in tabular and bar charts illustrating by Apache ECharts. The tables and plots drawn above can be viewed online from the website.

2.4. Online platform implementation

The RV*victor* is programmed using PHP and deployed on the Apache HTTP Server and the Ubuntu operating system. All data in RV*victor* is stored and managed with MySQL v15.1 for easy custom database searches. The web user interface is developed with JavaScript, HTML5, and CSS. RV*victor* has been tested on different browsers, such as Google Chrome, Mozilla Firefox, and Safari. It's now publicly accessible at: https://idrblab.org/rvvictor/.

3. Results and discussion

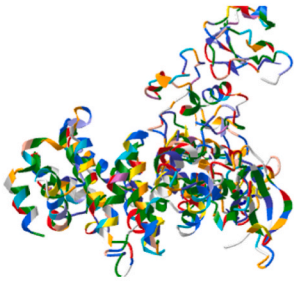
3.1. Statistics of interaction data between virus RNAs and host proteins (INTPROs)

INTPROs are an integral part not only of the regulation of viral translation and replication but also of the defense against virus infection [51–53]. In addition, deciphering INTPRO is beneficial for understanding the pathogenesis of RNA viruses and informing the development of antiviral therapy targets [54–56]. However, insight into INTPRO is limited, and these studies have been of limited impact regarding revealing how the viral RNA is regulated during infection. Thus, in this study, the INTPROs data of 107 types of virus/strains were collected, which included: *human immunodeficiency virus* (HIV), *ebola virus* (EBOV), *Dengue virus* (DENV), and a total of 16,223 INTPROs between 118 virus RNAs and 3829 host proteins were identified. All these interactions were validated using diverse living systems, which included 31 cell lines from 11 tissues and various model organisms. As can be seen in Fig. 3 (a typical *RVvictor* page for virus RNA), a complete list of host proteins interacting with this virus RNA, as well as detailed experimental information, which is explicitly outlined.

3.2. Explicit description of the interacting protein from multiple perspectives

Recent studies identifying or analyzing novel INTPROs provide insights into the development of novel antiviral strategies [57–59]. Furthermore, detailed descriptions of interacting proteins from multiple perspectives are also required. For this reason, we provide additional data for these host proteins, such as the pathways in which the proteins were involved and the available molecular regulators (especially drugs) of the proteins. In particular, to achieve an in-depth insight into the host response pathway activated by host protein during RNA virus infection, a total of 348 infection-associated pathways that 2235 host proteins were involved in were identified by literature review. Some of the typical infectious pathways included such as HIV infection, Coronavirus disease, single-stranded Viral RNA Replication, *etc.* All infectious pathway maps are readily viewable online and can be freely

Host Protein General Information

Protein Name	Cyclin-dependent kinase 1	**Gene Name**	CDK1
Host Species	Homo sapiens	**Uniprot Entry Name**	CDK1_HUMAN
Protein Families	Protein kinase superfamily, CMGC Ser/Thr protein kinase family, CDC2/CDKX subfamily		
EC Number	2.7.11.22; 2.7.11.23		
Subcellular Location	Nucleus Cytoplasm		
External Link	NCBI Gene ID	983	
	Uniprot ID	P06493	
	Ensembl ID	ENSG00000033327	
	HGNC ID	1722	
Related KEGG Pathway	Human immunodeficiency virus 1 infection	hsa05170	Pathway Map
	Cell cycle	hsa04110	Pathway Map
	Viral carcinogenesis	hsa05203	Pathway Map
3D Structure		PDB ID: 4Y72	
		FASTA Download	
		2D PNG Download	
		PDB File Download	

Potential Drug(s) that Targets This Protein

Drug Name	DrunkBank ID	Pubchem ID	TTD ID	REF
abemaciclib	DB12001	46220502	D05SBO	DrugCentral
AG-24322	DB13035	135413565	D0IX4B	DGIdb
ALSTERPAULLONE	DB04014	5005498	D0H3EV	DGIdb
AT-7519	DB08142	11338033	D07DTV	DGIdb
CINNARIZINE	DB00568	1547484	D0Q3YO	DGIdb
CLOFIBRATE	DB00636	2796	D0J5DC	DGIdb

Fig. 4. A typical RV*victor* webpage provides detailed descriptions of general information of each host interacting protein. A total of 3828 host proteins and their general information, such as protein name, protein family, gene name, EC number (if available), subcellular location, structures (downloadable in both 2D and 3D formats), and so forth, external linkage to other molecular biological databases (such as UniProt ID, gene ID, Ensembl ID, HGNC ID) are provided. In addition, we identified a total of 348 infection–associated pathways in which 2235 host proteins have been implicated. Available molecular regulators (especially drugs) for host proteins were also collected which yielded a total of 2000 drugs targeting 110 host proteins. All data are freely available for download from the RV*victor* website.

downloadable directly from the RV*victor* website. Moreover, available molecular regulators (especially drugs) of host-interacting proteins were also gathered through related database reviews. This resulted in the collection of a total of 7031 molecular regulators that targeted 703 host proteins and also provided the linkages for these drugs with several popular databases such as DrugBank [60], PubChem [61], and TTD [62]. In the RV*victor* webpage describing both pathways and regulators of host proteins (shown in Fig. 4), a variety of data from the corresponding protein were systematically described, and all data are freely available for download from the website.

3.3. Systematic collection of interaction data of INTmRNA and INTncRNA

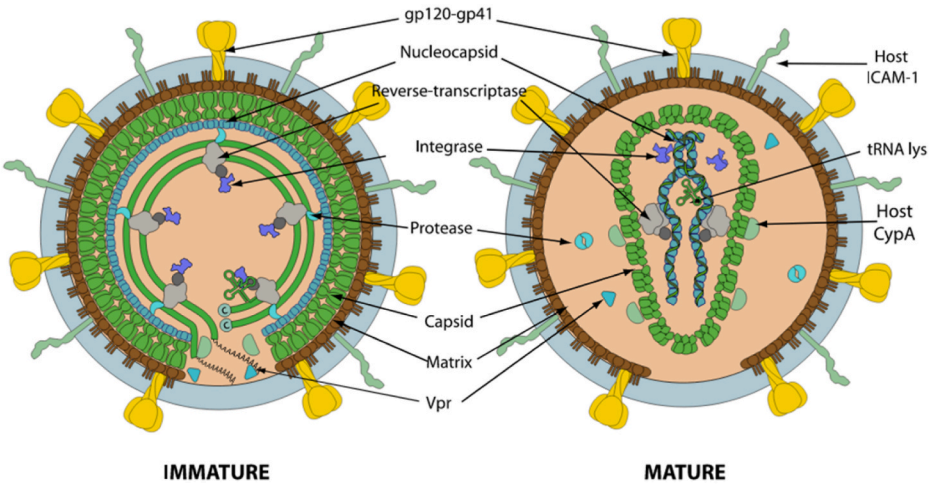
In the RNA virus's life cycle, replication and transcription are mediated by a replication transcription complex that involves RNA-directed interactions (INTPRO, INTmRNA, INTncRNA) that should be collectively considered to function in determining the molecular mechanism that underlies are critical for the infection of RNA viruses [63,64]. However, an integrated map of these interactions from all previous studies is lacking. Here, RV*victor* integrated the experimentally validated INTmRNA and INTncRNA data from ViRBase [51], and the miRSponge [33] database. As a result, RV*victor* contains a total of 94, 875 INTmRNA entries and 6470 INTncRNA entries involving 70 virus/strains and 11 host Organisms. Furthermore, we collected a variety of general information about each virus, including virus name, lineage, virus molecule type, genome type, related host, related disease information and Taxon ID [35], ICTV ID [65], KEGG BRITE ID [50], KEGG GENOME ID [66] virion, and genome illustrated picture from the ViralZone [27] on RV*victor* page for RNA virus (as provided in Fig. 5). In addition, a separate table of each RNA-directed interaction associated with this virus is located below the virus's general information on RV*victor* page for RNA virus (as shown Fig. 6). Herein, a comprehensive atlas of RNA-directed interactions (INTPRO, INTmRNA, INTncRNA)

Virus General Information

Virus Name	Human immunodeficiency virus type 1 (HIV-1)
Lineage	Viruses; Riboviria; Pararnavirae; Artverviricota; Revtraviricetes; Ortervirales; Retroviridae; Orthoretrovirinae; Lentivirus
Virus Molecule Type	ssRNA-RT
Sequence ID	U43141.1
Genome Type	Non-segmented
Related Host	Homo sapiens
Related Disease	Acquired immunodeficiency syndrome (AIDS)

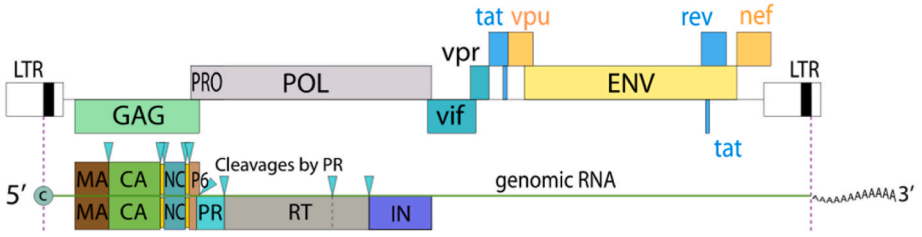
Virus Virion and Genome Information

Virion Information



Enveloped, spherical to pleomorphic in shape, 80-100 nm in diameter.

GENOME Information



Monopartite, linear, dimeric, ssRNA(+) genome of about 7-11 kb,, with a 5'-cap and a 3'poly-A tail. There are two long terminal repeats (LTRs) at the 5' and 3' ends. The LTRs contain the U3, R, and U5 regions. There are also a primer binding site (PBS) at the 5'end and a polypurine tract (PPT) at the 3'end.

Fig. 5. A typical RVvictor page for RNA viruses showing the detailed description of each RNA virus. General information is provided in the upper section, which includes virus name, lineage, virus molecule type, genome type, related host, related disease information, and external linkage to other molecular biological databases (such as Taxon ID, ICTV ID, KEGG BRITE ID, KEGG GENOME ID, and Virus-Host DB link). In addition, illustrated virion and genome images from the ViralZone database were also collected and provided.

were provided, which would a valuable resource in providing clues to understand the molecular mechanisms of viral infection as well as for developing antiviral strategies and drug repurposing. Furthermore, additional reference information on host mRNA, host ncRNA, and viral ncRNA was also collected and provided to RV*victor*. Overall, the RNA-directed interactions between virus RNA and host molecules were concentrated, providing an extensive database that aids users in studying RNA-directed interactions, particularly those associated with dynamic drug target discovery [67–69]and machine learning techniques [70–72].

3.4. Enrichment analysis of virus RNA-directed interactions

All these RNAs of these viruses generally interact with specific mRNAs or proteins and manipulate their activity for their survival, eventually affecting mRNAs/proteins that may regulate immune signaling pathways and signaling pathways induced by viral infection processes [73]. Therefore, RV*victor* uses experiment-supported ncRNA -

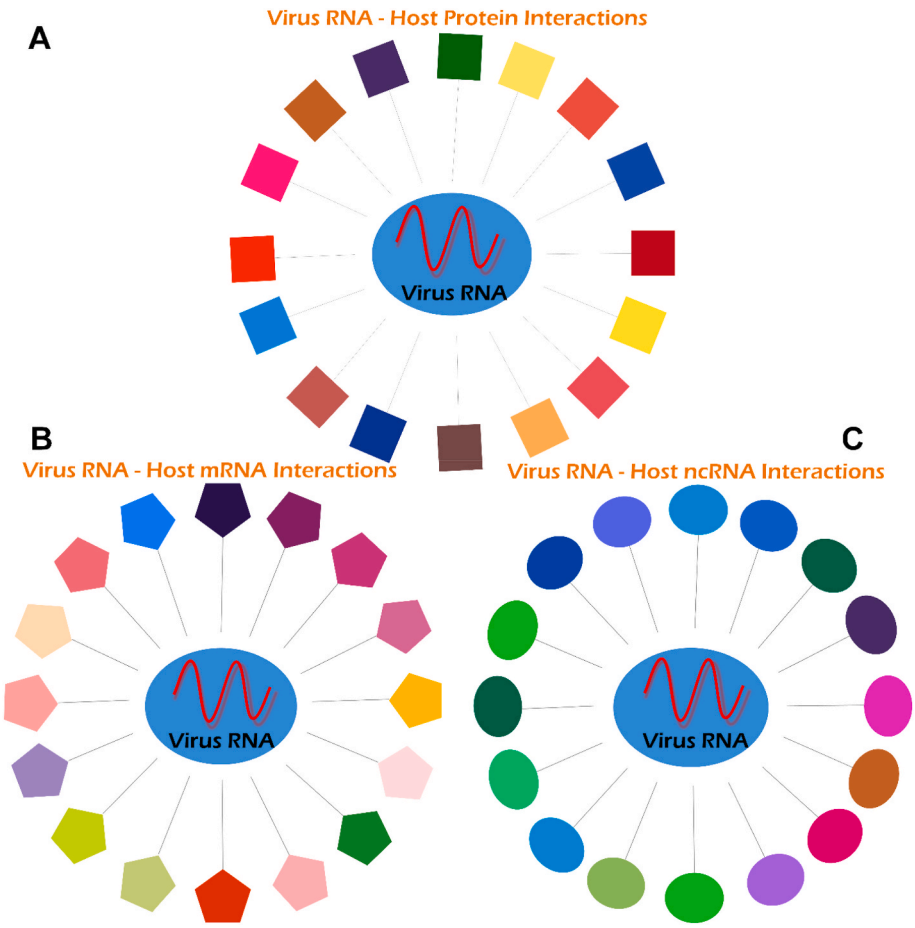


Fig. 6. Schematic picture of RNA-directed interaction (INTPRO, INTmRNA, INTncRNA) for each virus RNA, including **A:** interactions between virus RNA and host protein (INTPRO), **B:** interactions between virus RNA and host messenger RNA (INTmRNA), and **C:** interactions between virus RNA and host non-coding RNA (INTncRNA). A separate table of each RNA-directed interaction associated with this virus (each interaction is listed separately in the table) is below the virus's general information on RV*victor* page for RNA virus. Detailed information about the interacting viral RNA or host proteins, host mRNAs, and host ncRNAs can be found by clicking the corresponding blue button, such as "RNA info", "Pro-info", "mRNA info", "HncRNA Info" or "VncRNA Info".

target interactions as built-in data to map their targets of query ncRNA list and capture the function of these targets by Gene Ontology [74] and KEGG databases [50], which will generate a more comprehensive enrichment analysis based on the associated targets. We found significant enrichment of a suite of pathways related to infection, translation, and replication, e.g., HIV infection, coronavirus disease, regulation of viral genome replication, and so on (shown in Fig. 7).

4. Conclusion

Herein, RV*victor* was developed to unambiguously characterize the interaction data between virus RNAs and host molecules, offering, for the first time, the most comprehensive RNA-directed interaction (INTPRO, INTmRNA, INTncRNA) data of RNA virus' translation and replication among the existing databases by collectively describing the three types of RNA directed molecular interactions of viruses, and allowing functional annotations of all RV*victor* viral RNAs based on enrichment analyses of their target gene/proteins. RV*victor* has been smoothly running for months and tested by different research labs, and its data can now be fully accessed without any login requirement by all users at: http://rvvictor.idrblab.net/.

Funding

Funded by National Natural Science Foundation of China (82373790, 22220102001, 81872798 and U1909208); Natural Science Foundation of Zhejiang Province (LR21H300001); National Key R&D Program of China (2022YFC3400501); Leading Talent of the 'Ten Thousand Plan' - National High-Level Talents Special Support Plan of China; Fundamental Research Fund for Central Universities (2018QNA7023); "Double Top-Class" University Project (181201*194232101); Key R&D Program of Zhejiang Province (2020C03010). This work was supported by Westlake Laboratory (Westlake Laboratory of Life Sciences and Biomedicine); Alibaba-Zhejiang University Joint Research Center of Future Digital Healthcare; Alibaba Cloud; Information Technology Center of Zhejiang University.

CRediT authorship contribution statement

Kuerbannisha Amahong: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. **Wei Zhang:** Software, Visualization. **Yuhong Liu:** Data curation, Investigation. **Teng Li:** Data curation, Formal analysis, Investigation. **Shijie Huang:** Conceptualization, Data curation, Investigation. **Lianyi Han:** Conceptualization, Methodology, Project administration, Supervision, Validation. **Lin Tao:** Funding acquisition, Investigation, Methodology, Project administration, Supervision. **Feng Zhu:** Project administration, Supervision, Validation, Writing – original draft, Writing – review & editing.

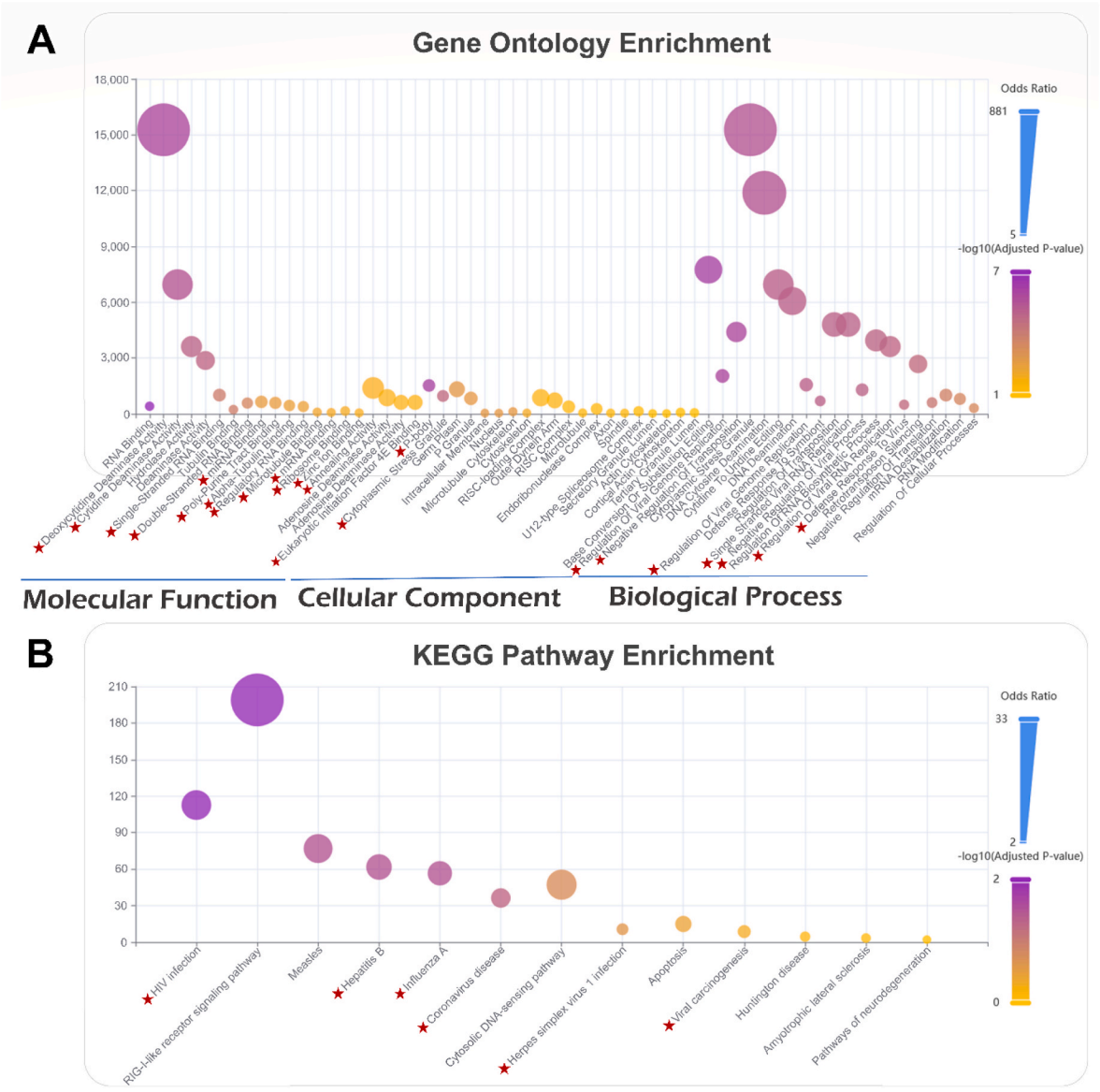


Fig. 7. A typical RVvictor page for virus RNA describes pathways enrichment analysis and visualization of enrichment results (top twenty) of A: Go ontology (Molecular Function, Cellular Component, and Biological Process) and B: KEGG (down) for RNA virus interacting host proteins. Virus infection-related results are marked by asterisks. All the host proteins, host mRNA, and target genes of host ncRNA that interact with virus RNA have performed enrichment analysis through Gene Ontology and KEGG, and the results were visualized by Apache ECharts, which can be readily viewed online from the RVvictor website.

Declaration of competing interest

None declared.

References

[1] I.C. Clark, P. Mudvari, S. Thaploo, et al., HIV silencing and cell survival signatures in infected T cell reservoirs, Nature 614 (2023) 318–325.

[2] C. Momont, H.V. Dang, F. Zatta, et al., A pan-influenza antibody inhibiting neuraminidase via receptor mimicry, Nature 618 (2023) 590–597.

[3] O. Goethals, S.J.F. Kaptein, B. Kesteleyn, et al., Blocking NS3-NS4B interaction inhibits dengue virus in non-human primates, Nature 615 (2023) 678–686.

[4] M. Behnia, S.B. Bradfute, The host Non-coding RNA response to alphavirus infection, Viruses 15 (2023) 562.

[5] B. Kim, S. Arcos, K. Rothamel, et al., Viral crosslinking and solid-phase purification enables discovery of ribonucleoprotein complexes on incoming RNA virus genomes, Nat. Protoc. 16 (2021) 516–531.

[6] S. Zhang, K. Amahong, X. Sun, et al., The miRNA: a small but powerful RNA for COVID-19, Briefings Bioinf. 22 (2021) 1137–1149.

[7] A. Slonchak, L.E. Hugo, M.E. Freney, et al., Zika virus noncoding RNA suppresses apoptosis and is required for virus transmission by mosquitoes, Nat. Commun. 11 (2020) 2205.

[8] D.W. Trobaugh, C.L. Gardner, C. Sun, et al., RNA viruses can hijack vertebrate microRNAs to suppress innate immunity, Nature 506 (2014) 245–248.

[9] S. Zhang, K. Amahong, C. Zhang, et al., RNA-RNA interactions between SARS-CoV-2 and host benefit viral development and evolution during COVID-19 infection, Briefings Bioinf. 23 (2021) bbab397.

[10] S. Zhang, W. Huang, L. Ren, et al., Comparison of viral RNA-host protein interactomes across pathogenic RNA viruses informs rapid antiviral drug discovery for SARS-CoV-2, Cell Res. 32 (2022) 9–23.

[11] B. Kim, S. Arcos, K. Rothamel, et al., Discovery of widespread host protein interactions with the Pre-replicated genome of CHIKV using VIR-CLASP, Mol. Cell. 78 (2020) 624–640.

[12] E.V. Koonin, Viruses and mobile elements as drivers of evolutionary transitions, Philos. Trans. R. Soc. Lond. B Biol. Sci. 371 (2016) 110.

[13] R. Mizuuchi, T. Furubayashi, N. Ichihashi, Evolutionary transition from a single RNA replicator to a multiple replicator network, Nat. Commun. 13 (2022) 1460.

[14] S.L. Bonilla, M.E. Sherlock, A. MacFadden, et al., A viral RNA hijacks host machinery using dynamic conformational changes of a tRNA-like structure, Science 374 (2021) 955–960.

[15] S. Roux, D. Paez-Espino, I.A. Chen, et al., IMG/VR v3: an integrated ecological and evolutionary framework for interrogating genomes of uncultivated viruses, Nucleic Acids Res. 49 (2021) D764–D775.

[16] R.D. Olson, R. Assaf, T. Brettin, et al., Introducing the bacterial and viral bioinformatics resource center (BV-BRC): a resource combining PATRIC, IRD and ViPR, Nucleic Acids Res. 51 (2023) D678–D689.

[17] E.V. Kriventseva, D. Kuznetsov, F. Tegenfeldt, et al., OrthoDB v10: sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs, Nucleic Acids Res. 47 (2019) D807–D811.
[18] A.L. Grazziotin, E.V. Koonin, D.M. Kristensen, Prokaryotic virus orthologous groups (pVOGs): a resource for comparative genomics and protein family annotation, Nucleic Acids Res. 45 (2017) D491–D498.
[19] S.Y. Rhee, M.J. Gonzales, R. Kantor, et al., Human immunodeficiency virus reverse transcriptase and protease sequence database, Nucleic Acids Res. 31 (2003) 298–303.
[20] A. Canakoglu, P. Pinoli, A. Bernasconi, et al., ViruSurf: an integrated database to investigate viral sequences, Nucleic Acids Res. 49 (2021) D817–D824.
[21] P. Li, X. Zhou, K. Xu, et al., RASP: an atlas of transcriptome-wide RNA secondary structure probing data, Nucleic Acids Res. 49 (2021) D183–D191.
[22] D. Montiel-Garcia, N. Santoyo-Rivera, P. Ho, et al., VIPERdb v3.0: a structure-based data analytics platform for viral capsids, Nucleic Acids Res. 49 (2021) D809–D816.
[23] D. Chen, C. Tan, P. Ding, et al., VThunter: a database for single-cell screening of virus target cells in the animal kingdom, Nucleic Acids Res. 50 (2022) D934–D942.
[24] S. Zhou, B. Liu, Y. Han, et al., ZOVER: the database of zoonotic and vector-borne viruses, Nucleic Acids Res. 50 (2022) D943–D949.
[25] E.L. Hatcher, S.A. Zhdanov, Y. Bao, et al., Virus variation resource - improved response to emergent viral outbreaks, Nucleic Acids Res. 45 (2017) D482–D490.
[26] J. Hayer, F. Jadeau, G. Deléage, et al., HBVdb: a knowledge database for hepatitis B virus, Nucleic Acids Res. 41 (2013) D566–D570.
[27] C. Hulo, E. de Castro, P. Masson, et al., ViralZone: a knowledge resource to understand virus diversity, Nucleic Acids Res. 39 (2011) D576–D582.
[28] Y. Xiang, Q. Zou, L. Zhao, VPTMdb: a viral posttranslational modification database, Briefings Bioinf. 22 (2021) bbaa251.
[29] D. Ako-Adjei, W. Fu, C. Wallin, et al., HIV-1, human interaction database: current status and new features, Nucleic Acids Res. 43 (2015) D566–D570.
[30] A. Calderone, L. Licata, G. Cesareni, VirusMentha: a new resource for virus-host protein interactions, Nucleic Acids Res. 43 (2015) D588–D592.
[31] T. Guirimand, S. Delmotte, V. Navratil, VirHostNet 2.0: surfing on the web of virus/host molecular interactions data, Nucleic Acids Res. 43 (2015) D583–D587.
[32] S. Yang, C. Fu, X. Lian, et al., Understanding human-virus protein-protein interactions using a human protein complex-based analysis framework, mSystems 4 (2019), e00303.
[33] P. Wang, H. Zhi, Y. Zhang, et al., miRSponge: a manually curated database for experimentally supported miRNA sponges and ceRNAs, Database 2015 (2015) bav098.
[34] J. Cheng, Y. Lin, L. Xu, et al., ViRBase v3.0: a virus and host ncRNA-associated interaction repository with increased coverage and annotation, Nucleic Acids Res. 50 (2022) D928–D933.
[35] S. Federhen, The NCBI taxonomy database, Nucleic Acids Res. 40 (2012) D136–D143.
[36] C. UniProt, UniProt: a worldwide hub of protein knowledge, Nucleic Acids Res. 47 (2019) D506–D515.
[37] G.R. Brown, V. Hem, K.S. Katz, et al., Gene: a gene-centered information resource at NCBI, Nucleic Acids Res. 43 (2015) D36–D42.
[38] K.L. Howe, P. Achuthan, J. Allen, et al., Ensembl 2021, Nucleic Acids Res. 49 (2021) D884–D891.
[39] E.A. Bruford, B. Braschi, P. Denny, et al., Guidelines for human gene nomenclature, Nat. Genet. 52 (2020) 754–758.
[40] S.K. Burley, H.M. Berman, C. Bhikadiya, et al., RCSB protein data bank: biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy, Nucleic Acids Res. 47 (2019) D464–D474.
[41] M. Varadi, S. Anyango, M. Deshpande, et al., AlphaFold protein structure database: massively expanding the structural coverage of protein-sequence space with high-accuracy models, Nucleic Acids Res. 50 (2022) D439–D444.
[42] M. Safran, I. Dalah, J. Alexander, et al., Genecards Version 3: the Human Gene Integrator, 2010, Database (Oxford), 2010, p. baq020.
[43] A. Kozomara, M. Birgaoanu, S. Griffiths-Jones, miRBase: from microRNA sequences to function, Nucleic Acids Res. 47 (2019) D155–D162.
[44] A. Thakur, M. Kumar, AntiVIRmiR: a repository of host antiviral miRNAs and their expression along with experimentally validated viral miRNAs and their targets, Front. Genet. 13 (2022), 971852.
[45] H.Y. Huang, Y.C. Lin, J. Li, et al., miRTarBase 2020: updates to the experimentally validated microRNA-target interaction database, Nucleic Acids Res. 48 (2020) D148–D154.
[46] J.H. Li, S. Liu, H. Zhou, et al., starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data, Nucleic Acids Res. 42 (2014) D92–D97.
[47] F. Xiao, Z. Zuo, G. Cai, et al., miRecords: an integrated resource for microRNA-target interactions, Nucleic Acids Res. 37 (2009) D105–D110.
[48] Gene ontology consortium: going forward, Nucleic Acids Res. 43 (2015) D1049–D1056.
[49] N. Del Toro, A. Shrivastava, E. Ragueneau, et al., The IntAct database: efficient access to fine-grained molecular interaction data, Nucleic Acids Res. 50 (2022) D648–D653.
[50] M. Kanehisa, M. Furumichi, Y. Sato, et al., KEGG: integrating viruses and cellular organisms, Nucleic Acids Res. 49 (2021) D545–D551.
[51] J. Wang, D. Sun, M. Wang, et al., Multiple functions of heterogeneous nuclear ribonucleoproteins in the positive single-stranded RNA virus life cycle, Front. Immunol. 13 (2022), 989298.
[52] C. Haddad, J. Davila-Calderon, B.S. Tolbert, Integrated approaches to reveal mechanisms by which RNA viruses reprogram the cellular environment, Methods 183 (2020) 50–56.
[53] N. Chen, B. Zhang, L. Deng, et al., Virus-host interaction networks as new antiviral drug targets for IAV and SARS-CoV-2, Emerg. Microb. Infect. 11 (2022) 1371–1389.
[54] L. Iselin, N. Palmalux, W. Kamel, et al., Uncovering viral RNA-host cell interactions on a proteome-wide scale, Trends Biochem. Sci. 47 (2022) 23–38.
[55] E. Girardi, S. Pfeffer, T.F. Baumert, et al., Roadblocks and fast tracks: how RNA binding proteins affect the viral RNA journey in the cell, Semin. Cell Dev. Biol. 111 (2021) 86–100.
[56] A. Singh, K.K. Pandey, S.K. Agrawal, et al., The SARS-CoV-2 UTR's intrudes host RBP's and modulates cellular splicing, Adv Virol 2023 (2023), 2995443.
[57] Y. Yao, H. Sun, Y. Chen, et al., RBM24 inhibits the translation of SARS-CoV-2 polyproteins by targeting the 5'-untranslated region, Antivir. Res. 209 (2023), 105478.
[58] R. Xiong, L. Zhang, S. Li, et al., Novel and potent inhibitors targeting DHODH are broad-spectrum antivirals against RNA viruses including newly-emerged coronavirus SARS-CoV-2, Protein Cell 11 (2020) 723–739.
[59] K. Amahong, W. Zhang, Y. Zhou, et al., CovInter: interaction data between coronavirus RNAs and host proteins, Nucleic Acids Res. 51 (2023) D546–D556.
[60] D.S. Wishart, Y.D. Feunang, A.C. Guo, et al., DrugBank 5.0: a major update to the drugbank database for 2018, Nucleic Acids Res. 46 (2018) D1074–D1082.
[61] S. Kim, J. Chen, T. Cheng, et al., PubChem in 2021: new data content and improved web interfaces, Nucleic Acids Res. 49 (2021) D1388–D1395.
[62] F. Zhu, Z. Shi, C. Qin, et al., Therapeutic target database update 2012: a resource for facilitating target-oriented drug discovery, Nucleic Acids Res. 40 (2012) D1128–D1136.
[63] S. Boersma, H.H. Rabouw, L.J.M. Bruurs, et al., Translation and replication dynamics of single RNA viruses, Cell 183 (2020) 1930–1945.
[64] C.R. Reid, A.M. Airo, T.C. Hobman, The virus-host interplay: biogenesis of +RNA replication complexes, Viruses 7 (2015) 4385–4413.
[65] E.J. Lefkowitz, D.M. Dempsey, R.C. Hendrickson, et al., Virus taxonomy: the database of the International committee on taxonomy of viruses, Nucleic Acids Res. 46 (2018) D708–D717.
[66] M. Kanehisa, S. Goto, S. Kawashima, et al., The KEGG databases at genomenet, Nucleic Acids Res. 30 (2002) D42–D46.
[67] F. Li, J. Yin, M. Lu, et al., DrugMAP: molecular atlas and pharma-information of all drugs, Nucleic Acids Res. 51 (2023) D1288–D1299.
[68] Y.H. Li, X.X. Li, J.J. Hong, et al., Clinical trials, progression-speed differentiating features and swiftness rule of the innovative targets of first-in-class drugs, Briefings Bioinf. 21 (2020) 649–662.
[69] X. Wang, F. Li, W. Qiu, et al., SYNBIP: synthetic binding proteins for research, diagnosis and therapy, Nucleic Acids Res. 50 (2022) D560–D570.
[70] M. Mou, Z. Pan, M. Lu, et al., Application of machine learning in spatial proteomics, J. Chem. Inf. Model. 62 (2022) 5875–5895.
[71] Y. Wang, Z. Chen, Z. Pan, et al., RNAincoder: a deep learning-based encoder for RNA and RNA-associated interaction, Nucleic Acids Res. (2023), https://doi.org/10.1093/nar/gkad1404.
[72] H. Zhang, Y. Wang, Z. Pan, et al., ncRNAInter: a novel strategy based on graph neural network to discover interactions between lncRNA and miRNA, Briefings Bioinf. 23 (2022) bbac411.
[73] Z. Guo, Y. Li, S.W. Ding, Small RNA-based antimicrobial immunity, Nat. Rev. Immunol. 19 (2019) 31–44.
[74] Consortium, Gene ontology consortium: going forward, Nucleic Acids Res. 43 (2015) D1049–D1056.