The impact of pandemics, epidemics, and the proliferation of artificial intelligence on (computational) drug discovery

G. Narahari Sastry
CSIR – North East Institute of Science and Technology
Jorhat – 785006
Assam, India

18th Sept. 2023

XXIX Symposium on
Bioinformatics and Computer-Aided Drug Discovery

IBMC
Institute of Biomedical Chemistry

Pogodinskaya str., 119121 Moscow, Russia

Outline of the talk

- Challenges during a pandemic
- MPDS$^{COVID-19}$ A disease specific approach
- COVID directed CADD strategies
- Galaxy implementation of MPDS
- Summary

Epidemics and pandemics from time immemorial

Prehistoric epidemic: Circa 3000 B.C.: A 5,000-year-old house in China filled with skeletons is evidence of a deadly epidemic.

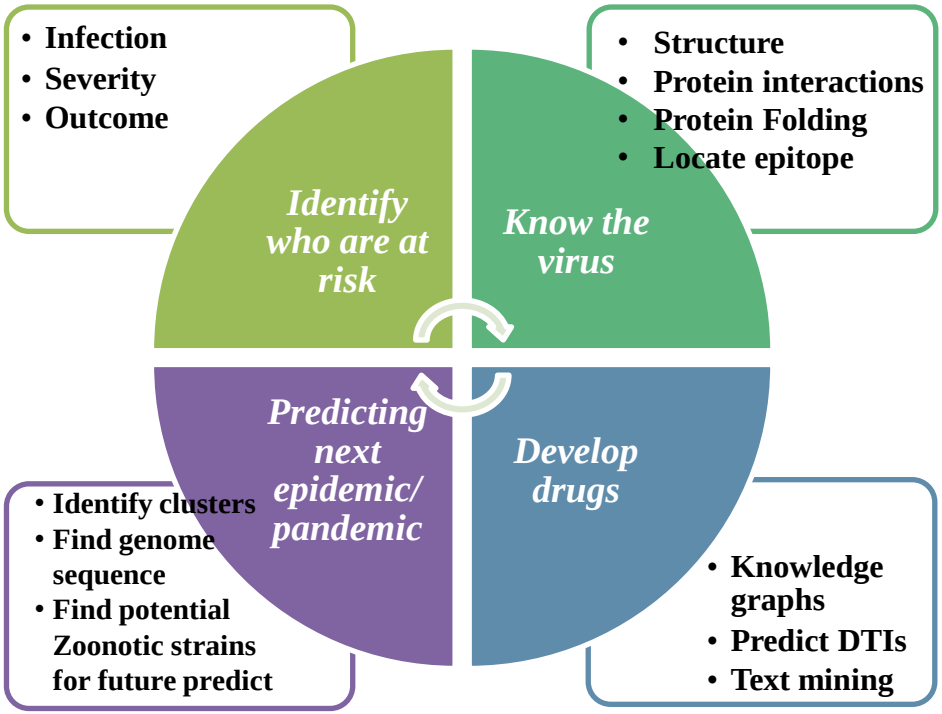
Plague of Athens: 430 B.C.: Remains of the Parthenon, one of the buildings on the acropolis of Athens. The city experienced a five year pandemic around 430 B.C.

Antonine Plague: A.D. 165-180 : When soldiers returned to the Roman Empire from campaigning, they brought back more than the spoils of victory. Roman soldiers likely brought smallpox home with them, giving rise to the Antonine Plague.

4. Plague of Cyprian: A.D. 250-271
5. Plague of Justinian: A.D. 541-542:
6. The Black Death: 1346-1353
7. Cocoliztli epidemic: 1545-1548
8. American Plagues: 16th century
9. Great Plague of London: 1665-1666
10. Great Plague of Marseille: 1720-1723
11. Russian plague: 1770-1772
12. Philadelphia yellow fever : 1793
13. Flu pandemic: 1889-1890
14. American polio : 1916
15. Spanish Flu: 1918-1920
16. Asian Flu: 1957-1958
17. AIDS pandemic and epidemic: 1981-?
18. H1N1 Swine Flu pandemic: 2009-2010
19. West African Ebola epidemic: 2014-2016
20. Zika Virus epidemic: 2015- ?

Prevention :: Cure :: Crisis Management

Drug Discovery :: Supply Chain :: Hospital Management :: Policy Decisions

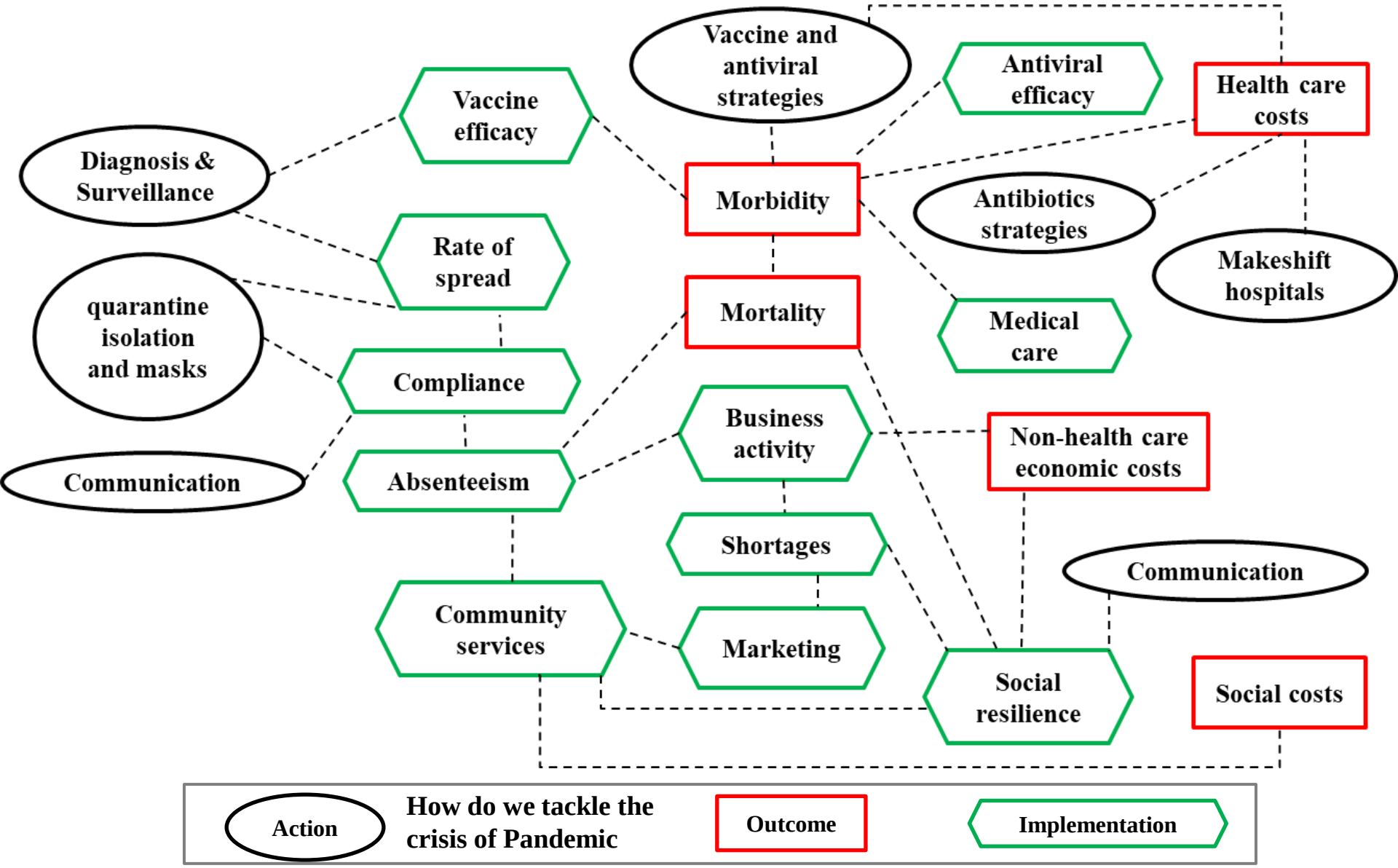


Preventive management

Adopting measures to reduce the impact of pandemics is challenging, thorough planning is vital in mitigating impacts on socio-economic and environmental systems,

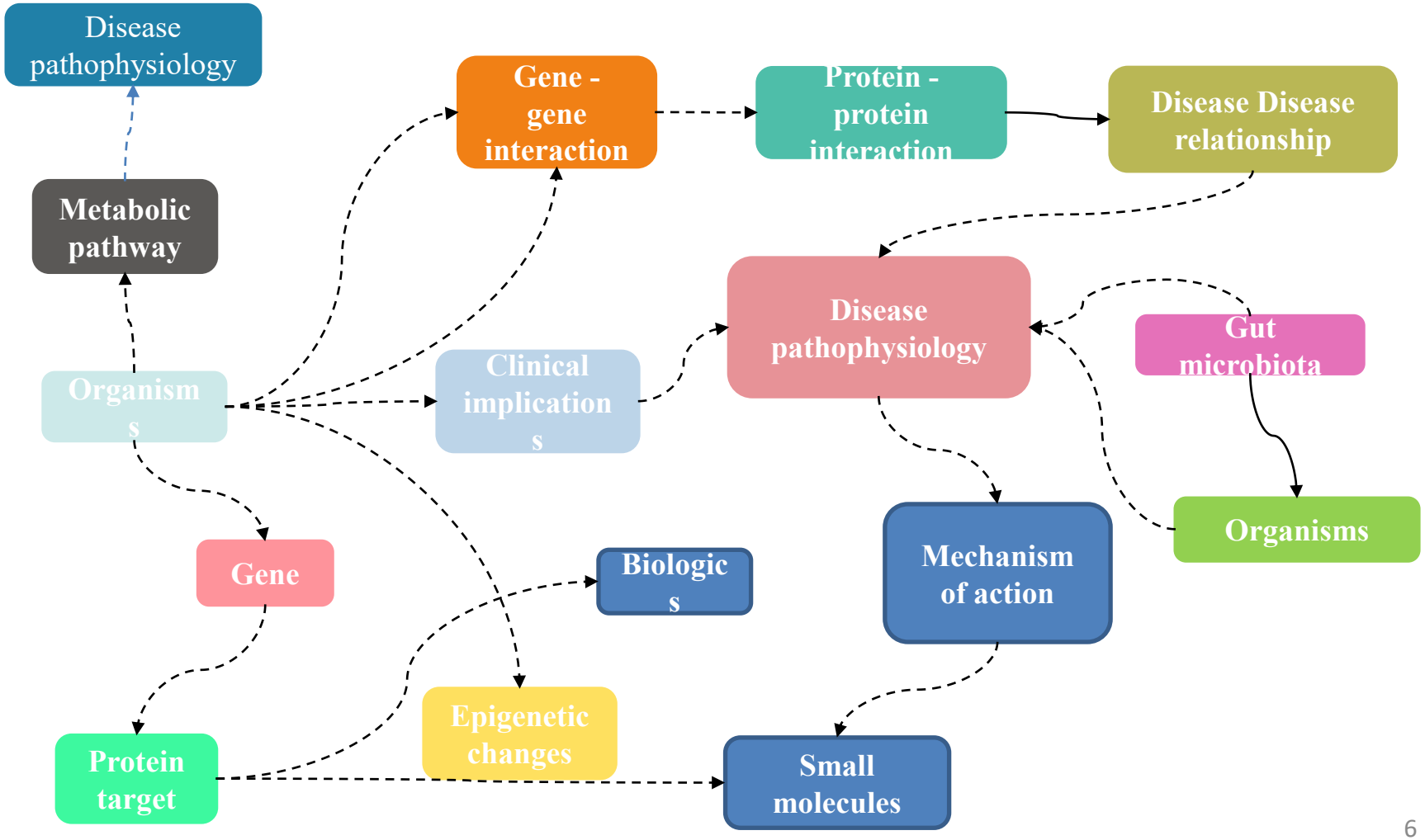
Role of Artificial intelligence (AI) : As COVID-19 spreads, raising fears of a worldwide lockdown, international organizations and scientists have been using artificial intelligence (AI) to track the epidemic in real-time, so as to be able to predict where the virus might appear next and develop an effective response.

Action, Outcome and Implementation Modes



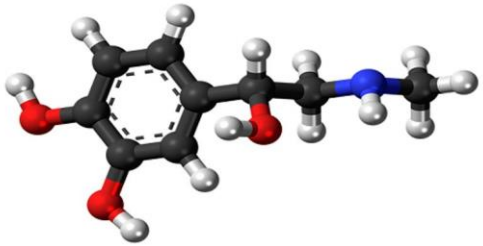
Disease-specific Drug Development

A disease cure may have a unique way, and generic strategy of drug development for all diseases has its limitations



Genesis of Molecular Property Diagnostic Suite: MPDS

What molecule to synthesize?



Molecule

→ Properties → Synthesize molecule with desired properties

Idea was conceived in 2004

MPDS Team: Started working in 2011

First publication: J. Chem. Sci. 2017 may issue (cover page)

Core developers

1. ANSHU BHARDWAJ
2. PRASAD V BHARATAM
3. M RAM VIVEK
4. NEHA TRIPATHI
5. RAKESH KUMAR
6. ARUN SHARMA
7. ANAMIKA SINGH GAUR
8. SRIDHARA JANARDHAN
9. LIJO JOHN

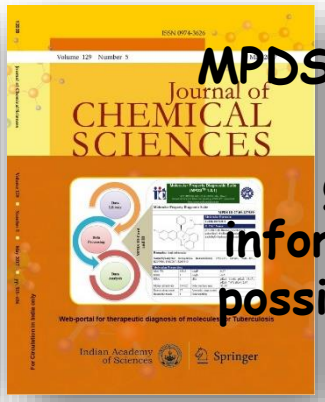
Co-developers

1. CHINMAYEE CHOUDHURY
2. ANIRBAN BANERJI**
3. D. ARUN KUMAR
4. M. PRASANTHI
5. P. SRI SARANYA
6. S. HEMASRI
7. KARUNAKAR TANNEERU
8. NANDAN KUMAR
9. KUMARDEEP CHAUDHARY
10. SANDEEP SINGH
11. ASHEESH KUMAR
12. RUCHI MISHRA
13. SURESH KUMAR
14. CHARUVAKA MUVVA
15. ER AZHAGIYA SINGAM
16. CHINMAI MADHURI
17. DEEPAK PANDIT
18. VIJAY KHEDKAR
19. YOGESH JOSHI
20. ANMOL HEMROM
21. DEEPAK BHARTI
22. PANKAJ NARANG**
23. ABHAYSINH MORI
24. HARISH JANGRA
25. R VENKAT KRISHNAN
26. APARNA SINGH
27. REETU SHARMA
28. HARI SAILAJA
29. KAAMINI RAITHATHA
30. PRASUN DUTTA
31. NEERAJ K RAJPUT
32. UCA JALEEL
33. ANURAG PASSI
34. ANURAG PASSI

PI: Dr. G. NARAHARI SASTRY

Co-Principal Investigators

Dr. G.P.S RAGHAV
Dr. ANSHU BHARDWAJ
Dr. PRASAD V BHARATAM
Dr. V SUBRAMANIAN
Dr. DEVESH KUMAR
Dr. M KARTHIKEYAN
Dr. UC ABDUL JALEEL
Dr. ANDREW LYNN



MPDS disease-specific web portals are aimed to gather, find & discover all possible information on a particular disease from all possible resources along with CADD tools at one place.

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

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

Molecular property diagnostic suite for diabetes mellitus (MPDS^{DM}): An integrated web portal for drug discovery and drug repurposing

Anamika Singh Gaur[a], Selvaraman Nagamani[a], Karunakar Tanneeru[a], Dmitry Druzhilovskiy[b], Anastassia Rudik[b], Vladimir Poroikov[b], G. Narahari Sastry[a,*]

[a] Centre for Molecular Modeling, CSIR-Indian Institute of Chemical Technology, Tarnaka, Hyderabad 500 007, India
[b] Institute of Biomedical Chemistry, Pogodinskaya Str., 10 Bldg. 8, Moscow, 119121, Russia

SAR AND QSAR IN ENVIRONMENTAL RESEARCH, 2017
VOL. 28, NO. 11, 913–926
https://doi.org/10.1080/1062936X.2017.1402819

Taylor & Francis
Taylor & Francis Group

Molecular property diagnostic suite (MPDS): Development of disease-specific open source web portals for drug discovery$

S. Nagamani[a], A. S. Gaur[a], K. Tanneeru[a], G. Muneeswaran[a], S. S. Madugula[a], MPDS Consortium, D. Druzhilovskiy[b], V. V. Poroikov[b] and G. N. Sastry[a]

[a]Centre for Molecular Modeling, CSIR-Indian Institute of Chemical Technology, Hyderabad, India; [b]Institute of Biomedical Chemistry, Moscow, Russia

Towards systematic exploration of chemical space: building the fragment library module in molecular property diagnostic suite
A.S. Gaur,..GNS. *Mol. Divers.*, 2022, pp.1-10

Molecular Property Diagnostic Suite Compound Library (MPDS-CL): A Structure based Classification of the Chemical Space
L. John, .. GNS, *Mol. Divers.*, 2023

Evolution of MPDS…

Literature

- Mtb Drug Info
- Mtb Gene Info
- Polypharmacology with Mtb
- DM Type 1 Drug Info
- DM Type 1 Biomarker Info
- DM Type 2 Drug Info
- DM Type 2 Biomarker Info
- Pathways

Data Processing

- File Format Converter
  - Compound conversion
  - Converter with OpenBabel
  - Generate 3D coordinates
- Descriptor Calculator
  - Padel Descriptor Tool
  - CDK Descriptor Tool

Data Library

- Protein Info
- Drug & Drug Repurposing
- Pathways
- Clinical Interventions
- Mutations
- Polypharmacology
- Protein-Protein interactions
- Drug-Protein Interactions
- Epidemiology
- Case Studies

Data Analysis

- Active site analysis
  - Search similar protein pockets
  - Find potential binding sites
- Pharmacophore
  - Optimization
  - Generation
- Scaffold Analysis
- Virtual Screening
- Sequence Alignment
  - TCoffee
  - ClustalW
- Phylogenetic Analysis
  - Build phylogenetic trees using PhyML, FASTTREE, GraPhiAn, TreeBeST

Target Library

- Mtb Target Search
- DM Target Search

$MPDS^{TB}$ (2017)

$MPDS^{DM}$ (2018)

$MPDS^{COVID-19}$ (2023)

Compound Library

- MPDS ID Search
- Exact Search
- Substructure Search
- Property based search
- Fingerprint based search

Gene Library

- Gene Library Search
- MPDS Gene ID Search

Data Analysis

- QSAR
- Docking
- Screening
  - BCS classification
  - Toxicity filter
  - NP likeliness calculator
- Visualization
  - 3D visualization with Jmol
  - Protein-ligand interaction
- Drug repurposing Tool
  - PASS online

Advanced Module (Machine Learning)

- Ensemble methods
- Linear classification model
- Linear regression model
- Support vector machine
- Nearest neighbor classification
- Numerical clustering
- Calculate classification metrics
- Calculate regression metrics
- Split datasets

Fragment Library

- Fragmenter
- Fragment based search

Predictive Models

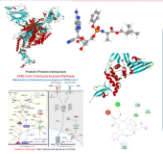
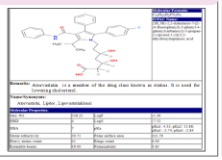
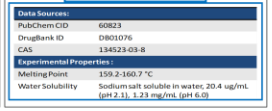
- Antiviral Prediction
- BBB Permeability Prediction

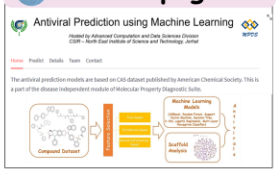
MPDS$^{COVID-19}$ A drug discovery portal

http://mpds.neist.res.in:8085

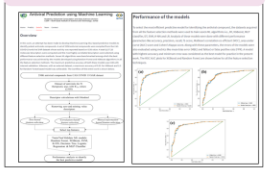
A ONE STOP SOLUTION

MPDS$^{COVID-19}$

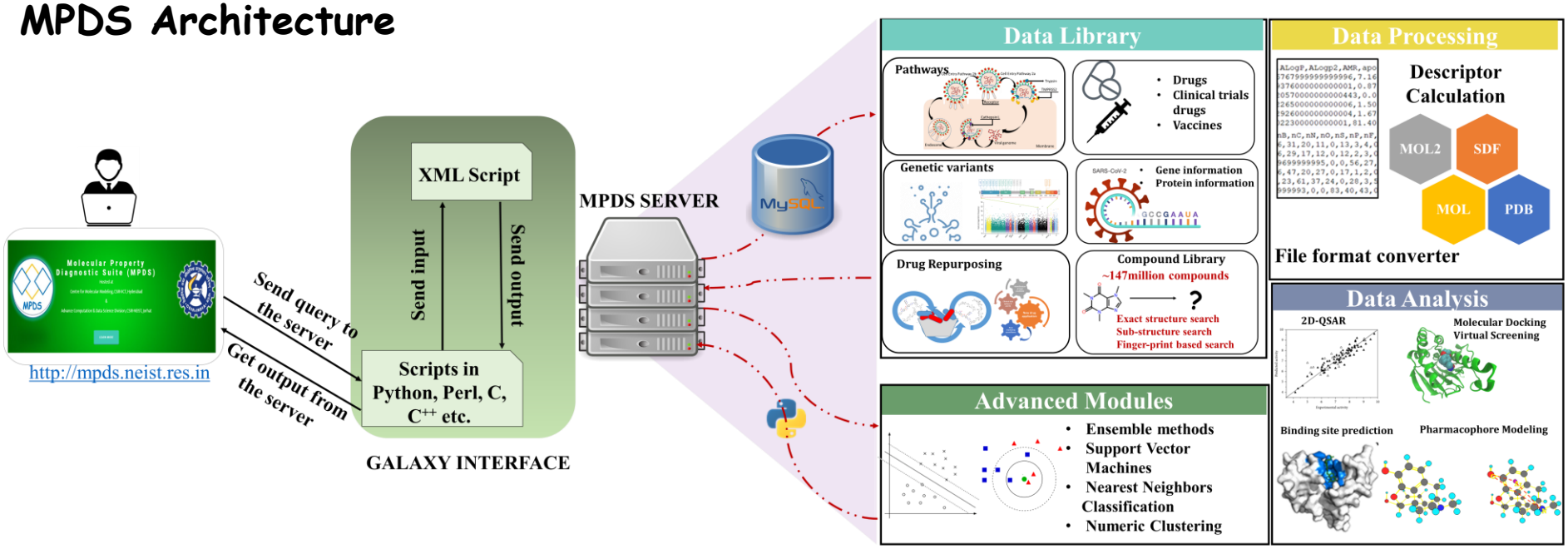
- **Home** → 
- **Search tools** → Quickly search for the desired tool
- **Browse** → Browse the data to get the analysis result
- **Developed databases and models**
  - **Biological Space** → **Target library**
    
    PDB IDs SARS-CoV2 | Structure information
    - **SARS CoV2 Disease library**
      
      - Polypharmacology
      - Drug repurposing
      - Mutation
      - Drug-Drug interaction
      - PPI
      - Host-protein Info.
      - Epidemiology
      - Variation
      - Inhibitors database
      - Literatures
  - **Compound library search** → **Search molecule**
    - Database ID search
    - Exact structure search
    - Sub-structure search
    - Property based search
    - Fingerprint based search
    - **Molecule ID card**
      1st page
      
    - **(Aadhar card)**
      2nd page
      

  - **Fragment library search** → **Fragmenter**
    Splits molecules in fragments
    
    - **Fragments based search**
      Searches using MPDS platform
      Nature of compound: Cyclic
      Type of fragments: 1. Monocyclic 2. Bicyclic 3. Tricyclic 4.Tetracyclic 5. Polycyclic
      
      Tricyclic
  - **Antiviral prediction machine learning model** → **Home page**
    

    - **Predict**
      

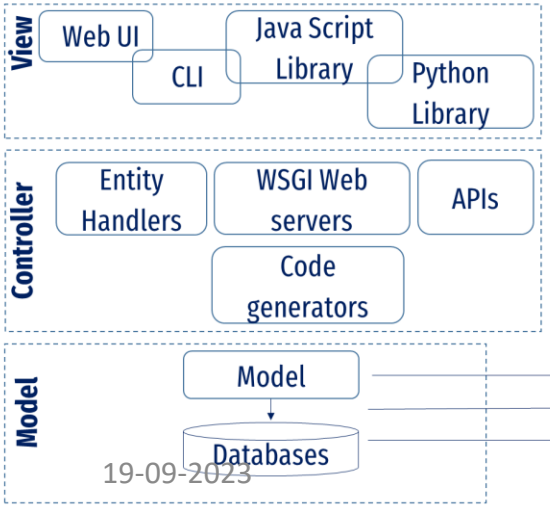
    - **Details**
      
- **Web service** → Enables researchers to upload their data/local files and make use of the available tools for various approaches
- **Submission** → The researchers are encouraged to submit their jobs
- **Download** → All the data can be downloaded freely
- **Manual** → Detailed tutorial on the usage of MPDS$^{COVID-19}$

MPDS Architecture

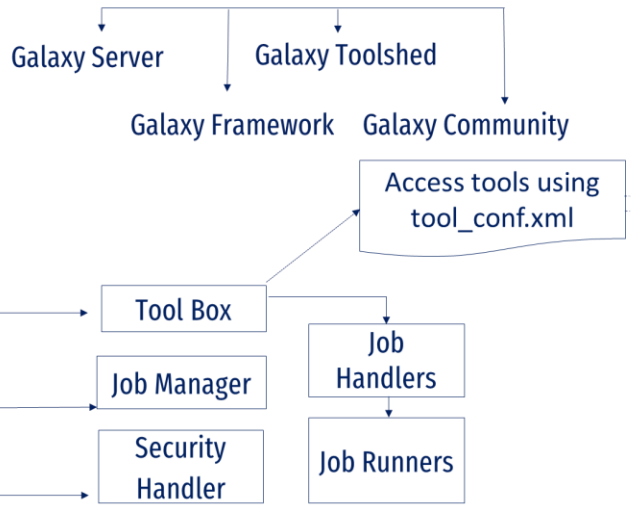


- The Data library information are stored in MPDS server and the information are being retrieved using MySQL database query
- The Data processing and Data analysis are drug discovery modules that are connected through Perl and python scripts with Galaxy
- The advanced modules are incorporated from "Galaxy ToolShed"

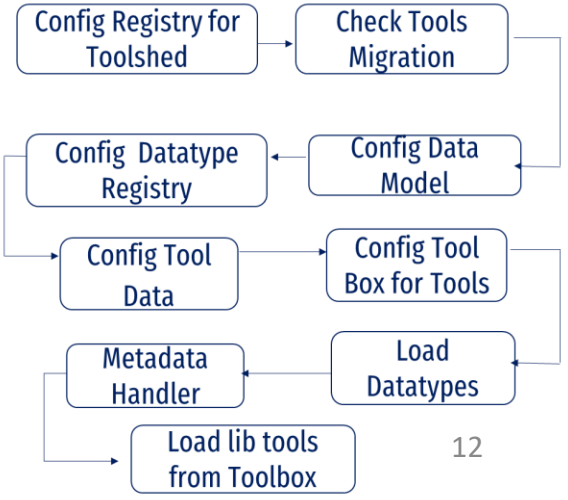
The MVC Architecture



Galaxy Complementary Components

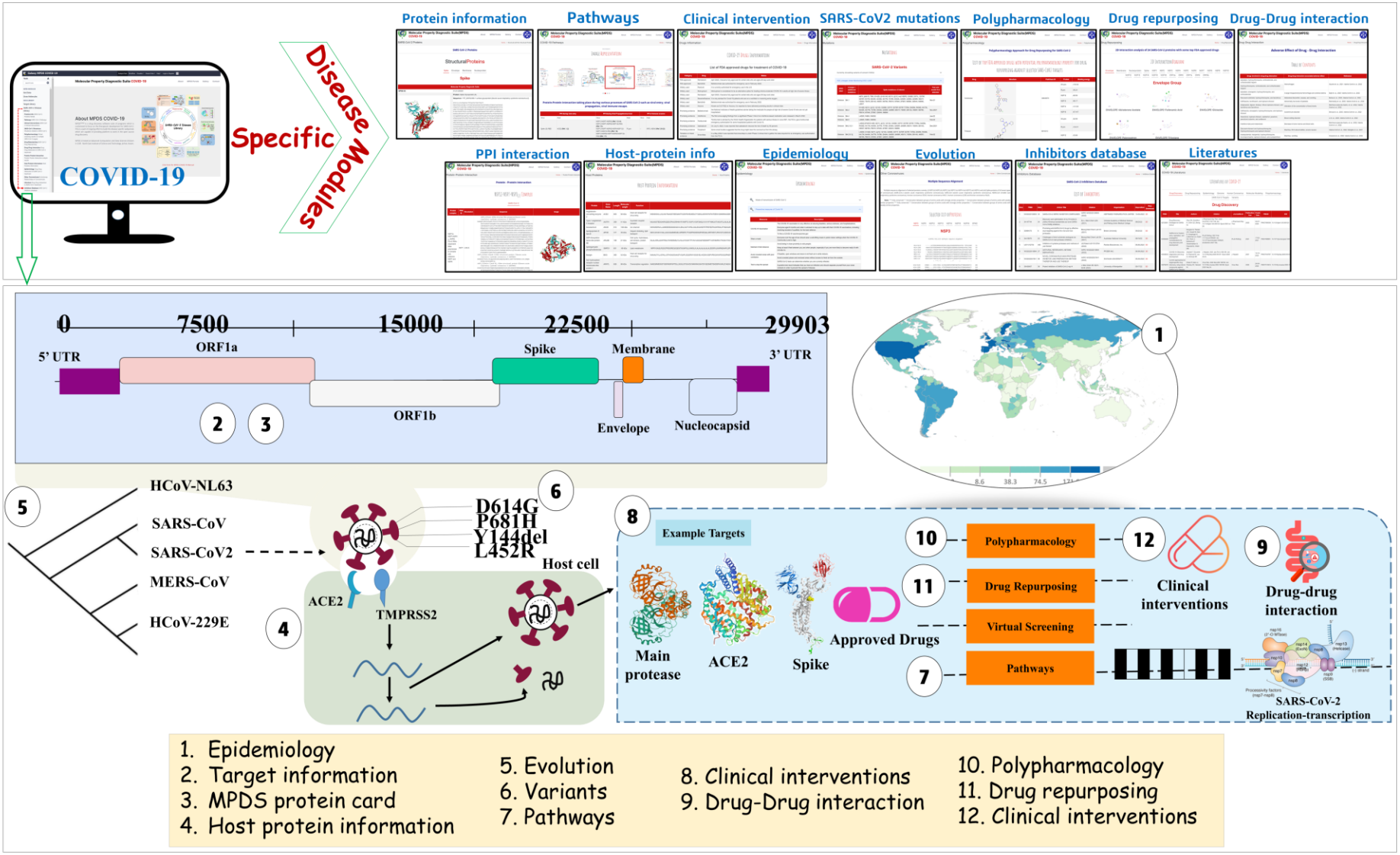


Configuring Tools from Galaxy Toolshed

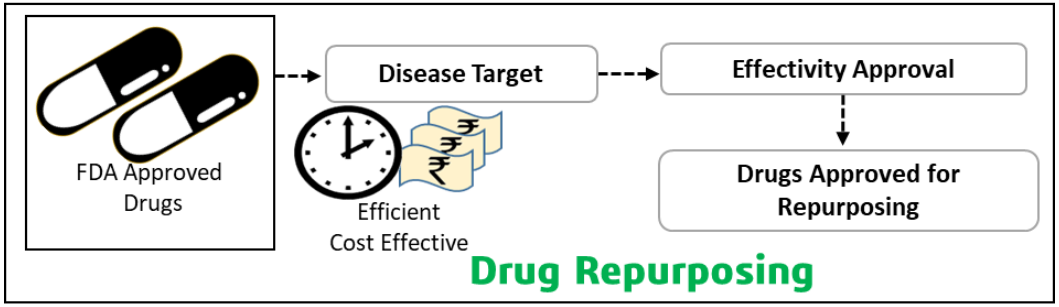


Covid-19 specific modules employed in MPDS$^{COVID-19}$

http://mpds.neist.res.in:8085

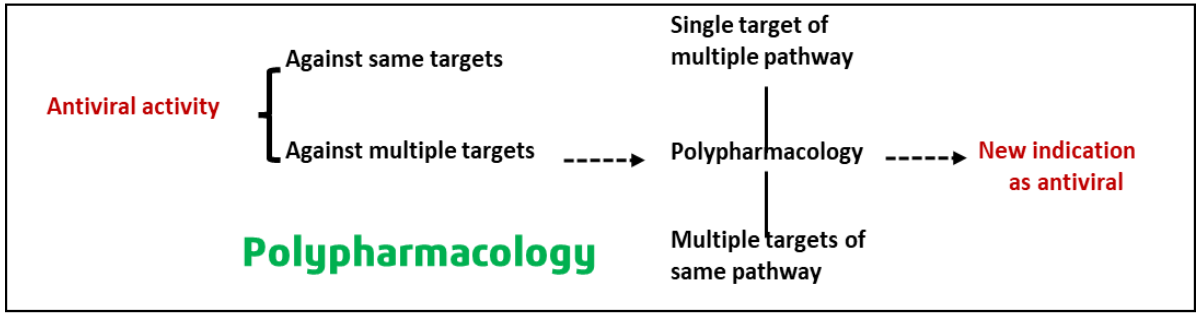


Drug Repurposing and Polypharmacology Strategy



Drug repurposing have become indispensable in recent times as compared to conventional drug discovery which is an expensive and time-consuming process with high risk of failure

Identifying drugs with property for targeting multiple targets that can be an effective method for combating the viral disease undergoing rapid mutational changes.



A unique combination of polypharmacology with drug repurposing serve as a novel approach for drug development against SARS-CoV2

E Jamir.. GNS J. Chem. Sci. 2022, 134, 57 , E Jamir.. GNS Plos one, 2023, 18 , e0289890
MH Bohari and GNS J. Mol. Model., 2012, 18, 4263-4274, SS Madugula.. GNS Mol. Divers., 2022, 26, 1675-1695

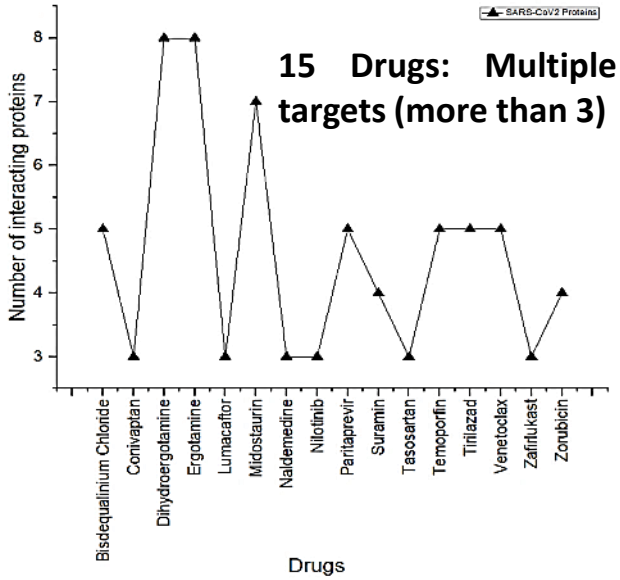
Drug Repurposing and Polypharmacology for COVID-19

4139 FDA Approved Drugs + **24 SARS-CoV2 Targets NSP (15), Structural proteins (4), ORFs (5)**

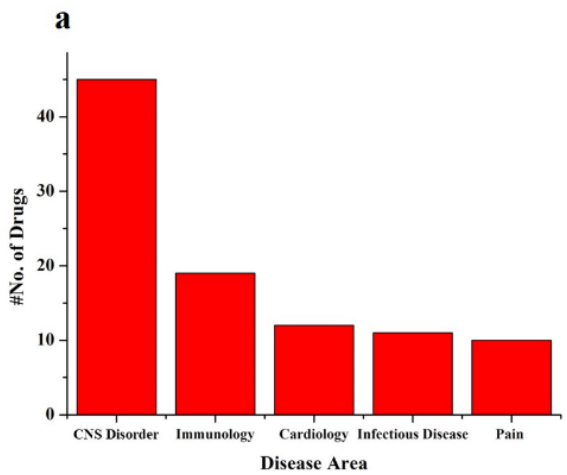
- Virtual Screening
- Distribution of therapeutic Indications
- Protein-Ligand Interaction
- Pathway Analysis

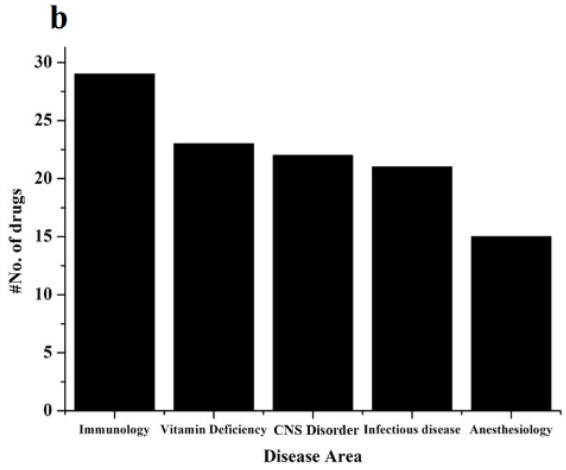
→ **Potential Leads for Repurposing**

Multi-Targeting Drugs



Classification of Drugs Based on Their Therapeutic Areas

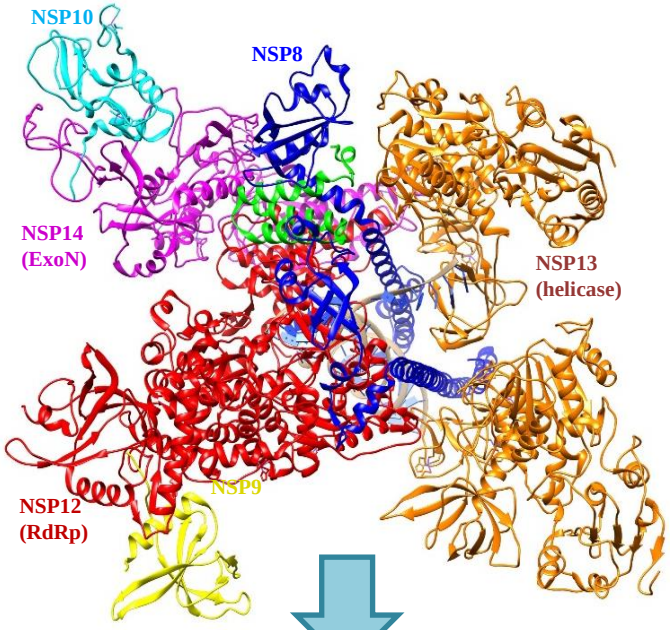


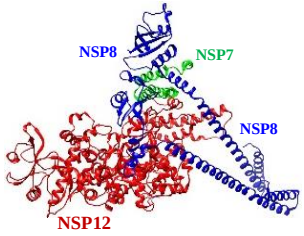


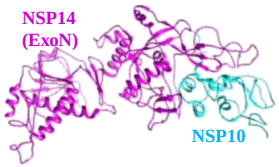
- The repurposeable drugs with original indications against neurological disorders (tirilazad), pain (dihydroergotamine, ergotamine), cancer (midostaurin, venetoclax, temoporfin) and bacterial infections (bisdequalinium chloride) may emerge as effective anti SARS-CoV2 drugs, owing to their multi-targeting nature.

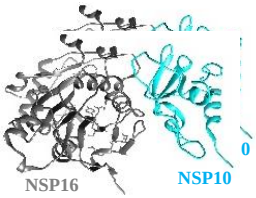
E Jamir.. GNS J. Chem. Sci. 2022, 134, 57 , E Jamir.. GNS Plos one, 2023, 18 , e0289890
MH Bohari and GNS J. Mol. Model., 2012, 18, 4263-4274, SS Madugula.. GNS Mol. Divers., 2022, 26, 1675-1695

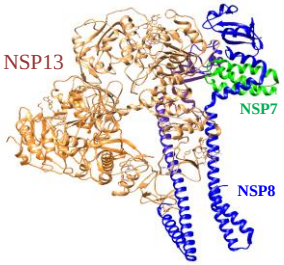
Pinning down PPIs in SARS-CoV-2 Infection and Propagation











Host-Virus PPI
ACE2-Spike

SARS-CoV-2 Entry

Viral entry

ACE2

ACE2

Lungs

Gut

Tissue Damage

Hyper inflammation

Cytokine Storm

Increasing cytokines and interleukins

Viral-Viral PPIs

Viral Propagation Replication-transcription

IMMUNE SYSTEM

Antigen Processing and Presentation

Cellular Immunity

Humoral Immunity

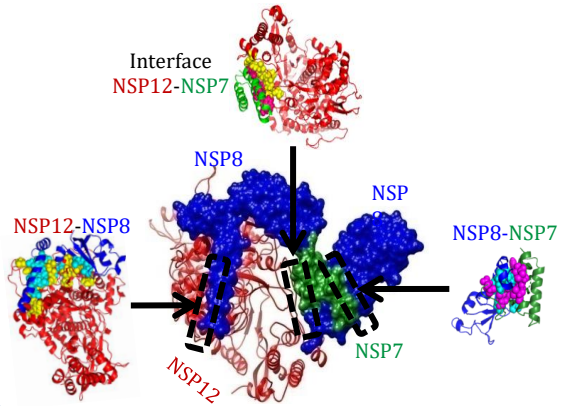
Immune evasion

Host-Virus PPI

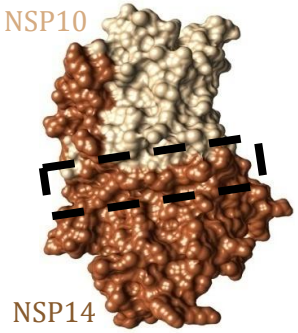
Ribosome-NSP1, Plpro-IRF/MDA5, 3CLpro-STING, NSP6-TKB1, NSP13-TKB1

Protein-Protein interactions in deciphering viral infection mechanism: replication-transcription and RNA proofreading, etc.

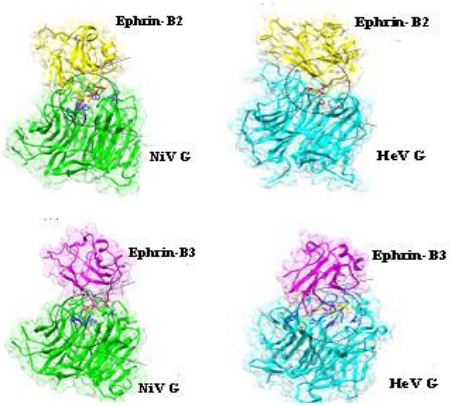
NSP12-NSP7-NSP8(2) (propagation)
of SARS-CoV2 and CoV
Function: Replication-Transcription



NSP14-NSP10 (propagation)
SARS-CoV2, CoV & MERS
Function: RNA proofreading

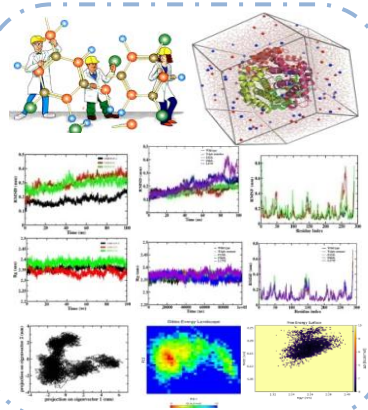


G-protein-EFNB2/B3 (Entry)
Hendra & Nipah virus
Function: Viral entry

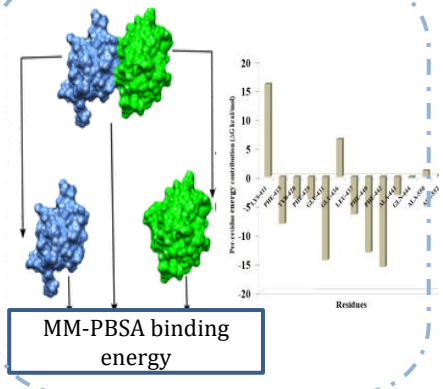


Computational Methods Employed

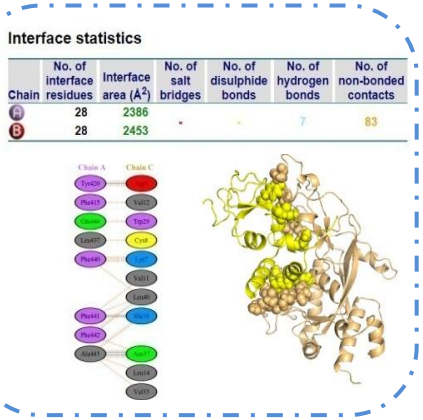
MD Simulation
& Mutational analysis



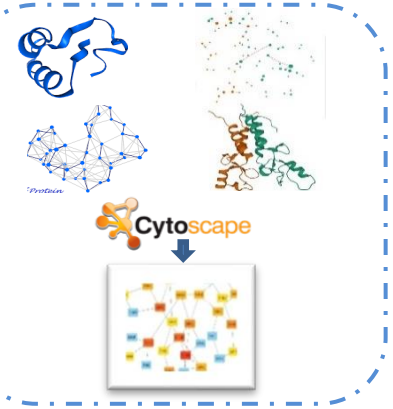
Total binding free
energy/residue-wise



PPI profiling &
Hotspot detection



Res-res networking:
Hub residue identification

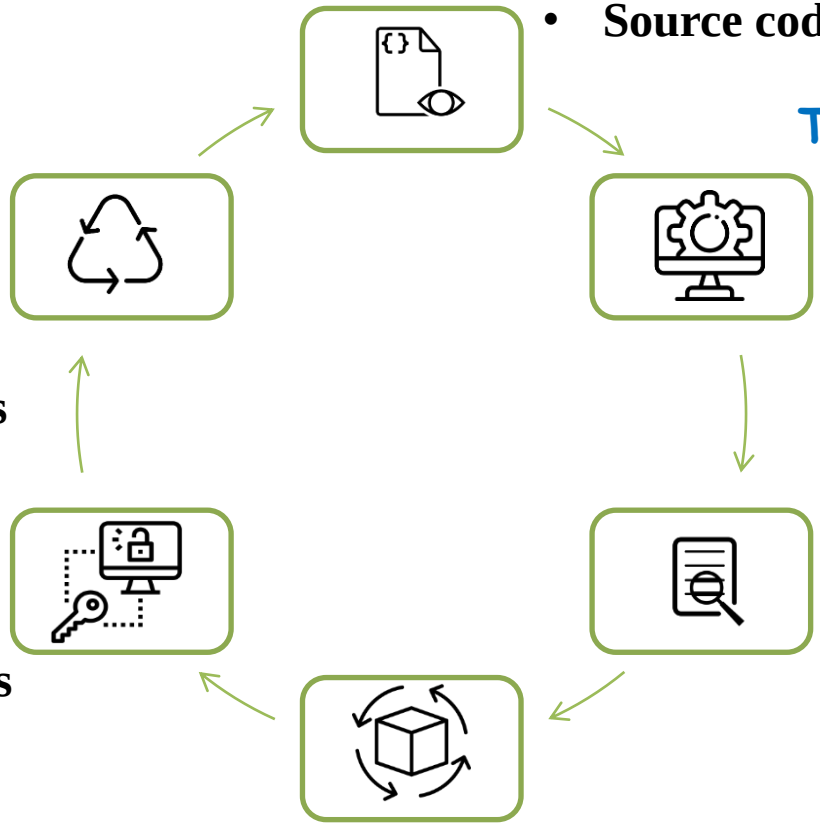


H Sarma, GNS; ACS Omega, 7:30041 (2022); H Sarma, E. Jamir, GNS, J. Mol. Struct. 1257:132602 (2022); L Priyadarsinee, H Sarma, GNS, J. Chem. Sci, 134:114 (2022)

Galaxy : A community led development

Open source platform that enables integrating and customizing sites

- A web service integrating a wealth of tools, computer resources, terabytes of data and permanent storage. Across the Globe; Australian, Indian, European, Czech, France, Japan, etc Galaxy communities with users and developers were spread.



Open-source

- **Source code is freely accessible**

Tools

- **Various domains (that can be plugged into workflows) through its graphical web interface.**

Transparent

- **Users can share histories, workflows and visualization**

Interactive environments

- **Run codes in Rstudio, Jupyter, etc.**

Accessible

- **Run complex tools and workflows**
- **Visualization of results**

Reproducible

- **Repeat and understand computational analysis**
- **Tool parameters**

Custom tools integration to MPDS Galaxy

Script or a program written for any given problem in CADD and related area can be added into MPDS galaxy and made available as a tool.

```perl
#!/usr/bin/perl -w
open (IN, "<$ARGV[0]");
open (OUT, ">$ARGV[1]");
while (<IN>) {
    chop;
    if (m/^>/) {
        s/^>//;
        if ($. > 1) {
            print OUT sprintf("%.3f", $gc/$length) . "\n";
        }
        $gc = 0;
        $length = 0;
    } else {
        $gc += tr/[gc]//;
        $length += length $_;
    }
}
print OUT sprintf("%.3f", $gc/$length) . "\n";
close( IN );
close( OUT );
```

toolExample.pl

```xml
<tool id="fa_gc_content_1" name="Compute GC content" version="0.1.0">
  <description>for each sequence in a file</description>
  <command interpreter="perl">toolExample.pl $input $output</command>
  <inputs>
    <param format="fasta" name="input" type="data" label="Source file"
  </inputs>
  <outputs>
    <data format="tabular" name="output" />
  </outputs>

  <tests>
    <test>
      <param name="input" value="fa_gc_content_input.fa"/>
      <output name="out_file1" file="fa_gc_content_output.txt"/>
    </test>
  </tests>

  <help>
This tool computes GC content from a FASTA file.
  </help>

</tool>
```

toolExample.xml

Outlook

- An approach based on disease specific approach appear to be important and it is very relevant during a pandemic.
- Pandemics also have a great potential to push disruptive research. AI/ML have a great potential for the future preparedness.
- MPDS$^{COVID\text{-}19}$, is an attempt to bring multifarious computational modeling, data driven approaches for the drug discovery research. It is an open access initiative.

Acknowledgements

Lipsa Priyadarsinee, Esther Jamir, Selvaraman Nagamani, Hridoy Jyoti Mahanta, Nandan Kumar, Lijo John, Himakshi Sarma, Asheesh Kumar, Anamika Singh Gaur, Rosaleen Sahoo, S. Vaikundamani, Kripa Dristi Dihingia, Jyotirmoy Dev, D. Gogoi, A. Pandey, D. Saikia, Bikram, H. Chutia

Vladimir Poroikov, N. Arul Murugan, U. Deva Priyakumar, G.P.S. Raghava, Prasad V. Bharatam, Ramakrishnan Parthasarathi, V. Subramanian, G. Madhavi Sastry, Pankaj Bharali, Lakshi Saikia, Swapnali Hazarika, Santanu Baruah, Mohan Lal

We thank DBT centre of excellence, Centre of excellence of petroleum research, DBT-BIRAC-BIONEST, CSIR, CSIR-NEIST and other funding agencies for supporting research

Thank you

Revitalizing NEIST for Strengthening North East