Network-based identification of drugs and drug combinations targeting host dependency factors across multiple pathogens

Marcus Oswald[1], André Dietz[1], Mohasehe Naseri[1], Alicia Hiemisch[1], Itunu Isewon[1], Mathieu Mateo[2], Seyoum Ayehunie[3], Holger Erfle[4], Tuna Toptan Grabmair[5], Caroline Goujon[6], and Rainer König[1]

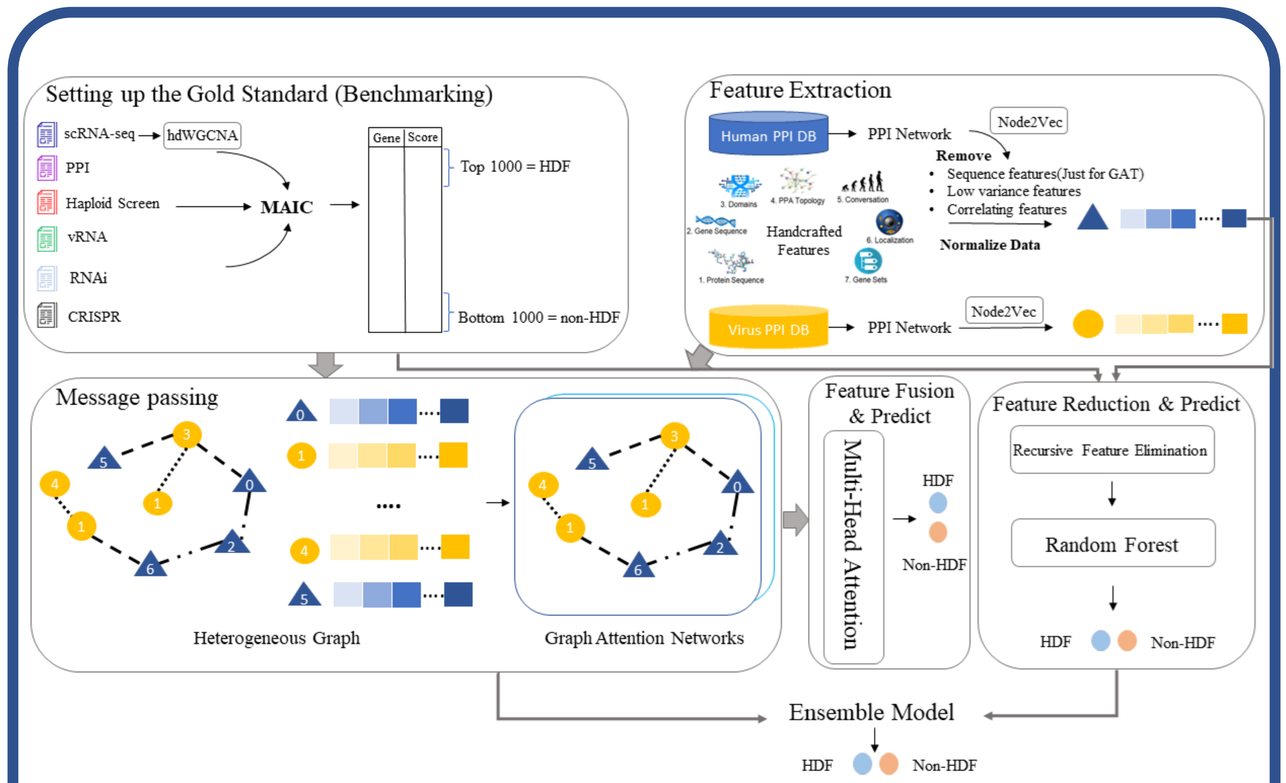
1. Systemsbiology Research Group, Institute for Infectious Diseases and Infection Control, Jena University Hospital, Jena
2. Institut Pasteur, Lyon, France
3. Mattek Europe, Bratislava, Slovakia
4. Ruprecht Karls University, Heidelberg, Germany
5. Technical University of Munich, Munich, Germany, München
6. Institut de Recherche en Infectiologie de Montpellier (IRIM), CNRS, France, Montpellier

Introduction

The emergence and re-emergence of viral pathogens have the potential to cause life-threatening epidemics, particularly in our increasingly globalized and ecologically challenging world. A critical aspect of pandemic preparedness is the development of effective antiviral therapies, particularly broad-spectrum antivirals that target host dependency factors (HDFs), thereby reducing the risk of resistance development. The EU-funded project APPEAL (Antivirus Pandemic Preparedness EuropeAn pLatform) integrates computational and experimental approaches to identify broad-spectrum antiviral HDFs along with corresponding drugs. The aim of APPEAL is to implement a drug selection pipeline. This work is one part of the pipeline that helps selecting promising drug candidates for our experimental collaboration partners.

Integrating protein interaction data to improve HDF prediction performance

We developed a machine-learning pipeline that integrates datasets of different entities using the MAIC algorithm [1] enabling to assemble a robust gold standard of 1,000 host-dependency factors (HDF) and 1,000 non-HDF. For each gene, >60,000 cellular features were computed describing sequence motifs, gene sets, cellular compartment, evolutionary conservation, network features, transcriptional profiles. Features were filtered by variance, correlation and recursive feature elimination. Random Forest and Graph Attention Network models were trained through a stacking ensemble, achieving cross-validated ROC-AUC >0.90 for different pathogens. The analysis identified key hub factors like HMGB1, SRC, CSNK2A1, and HNRNPK for future experimental validation.



The PPI network base on data from BioGRID [3] and STRING [4]

The drug database was compiled from Drugbank [5], ChEMBL [6], TTD [7], PharmGKB [8] and BindingDB [9] Drugbank [5], ChEMBL [6], TTD [7], PharmGKB [8] and BindingDB [9]

Identifying drugs and drug combinations using protein interaction

Given a list of HDF, the PPI-network and the drug database, we use a modified version of the method of Cheng et al. [2] to assemble a list of drugs that fit best to the given HDF as targets. The method minimizes the distances of the targets of a drug to the HDF, in comparison with distances calculated from randomly selected proteins with similar degree. With this method. we obtain drugs that are locally close to our HDF within the network. In addition, we obtain a rather different order of the drugs if we put their targets into three distance groups and perform tests for enrichment of HDF. This latter approach rewards drugs that are globally reflecting our HDF-list. The final ordering is obtained by a combination of the local and the global scoring scheme.

Defined distances within the PPI

0	1	2
HDF	Neighbours of HDF	All other proteins

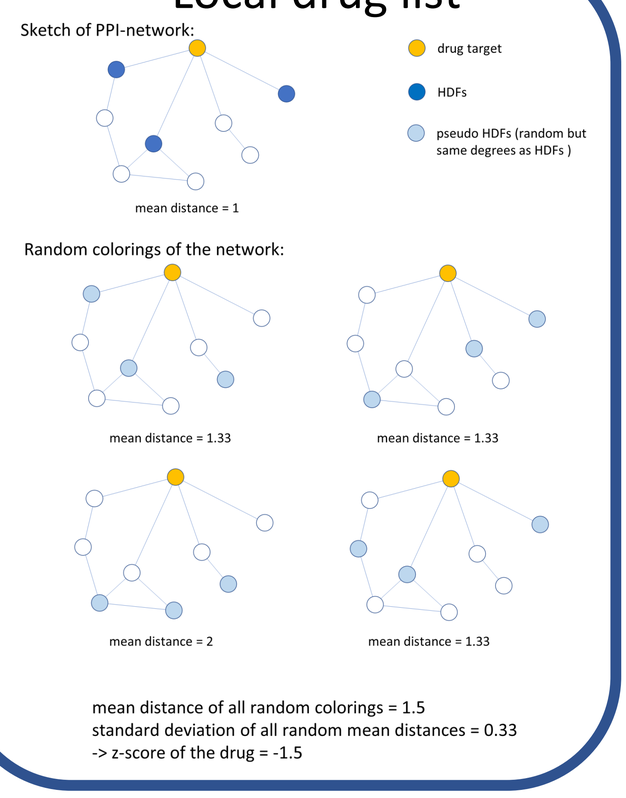
Global drug list

HDF enrichment tests for the drug targets

Predicted drugs targeting HDF of invasive bacteria

1. Guanosine-5'-Diphosphate
2. Fedratinib
3. Fostamatinib
4. Botulinum toxin type B
5. Ruboxistaurin
6. Hexachlorophene
7. Sunitinib
8. Imatinib
9. Masitinib
10. Calcium citrate

Local drug list



Project progress to date

The resulting lists of ordered drugs were compared with the literature evidencing several top ranking hits as inhibitors of known HDFs of pathogens of our concern, but also so far not known HDFs. Lists of drugs and drug-combinations are passed to our experimental collaboration partners.

References

[1] Li, Bo, et al. "Genome-wide CRISPR screen identifies host dependency factors for influenza A virus infection." Nature communications 11.1 (2020): 164., [2] Cheng F, Kovács IA, Barabási AL. "Network-based prediction of drug combinations." Nat Commun. 2019 Mar 13;10(1):1197. doi: 10.1038/s41467-019-09186-x. Erratum in: Nat Commun. 2019 Apr 15;10(1):1806. doi: 10.1038/s41467-019-09692-y. PMID: 30867426; PMCID: PMC6416394.[3] https://thebiogrid.org, [4] https://string-db.org, [5] https://go.drugbank.com,,[6] https://www.ebi.ac.uk/chembl,,[7] https://idrblab.net/ttd, [8] https://www.pharmgkb.org, [9] https://www.bindingdb.org

Acknowledgements

This work was supported by the Federal Ministry of Education and Research (BMBF) Germany, 13 N15711 (LPI), 13 N15745 (SARS-CoV2-DX) and the European Union (iRecords, 01KU2209).

Gefördert durch: Bundesministerium für Forschung, Technologie und Raumfahrt

Funded by the European Union

Antivirus Pandemic Preparedness EuropeAn pLatform

Horizon Europe research and innovation programme (Grant Agreement No. 101137311) Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the granting authority can be held responsible for them. **www.project-appeal.eu**