

COMPARE analysis as an efficient bioinformatic approach to accelerate repurposing of existing drugs against Covid-19 and other emerging epidemics

Imad Naasani¹

¹Nanoco Group Plc

July 30, 2020

Abstract

A novel bioinformatic approach for drug repurposing against emerging viral epidemics like Covid-19 is described. It exploits the COMPARE algorithm, a public program from the NCI to sort drugs according to their patterns of growth inhibitory profiles from a diverse panel of human cancer cell lines. The data repository of the NCI includes the growth inhibitory patterns of more than 55000 molecules. When drugs with purported anti-SARS-CoV-2 activities were used as seeds (e.g., hydroxychloroquine and ritonavir) in COMPARE, the analysis uncovered several drugs with fingerprints similar to the seeded drugs. Interestingly, the uncovered drugs were all reportedly known to exert antiviral activities, confirming that COMPARE analysis is valuable for predicting antiviral drug repurposing. Unlike pure in-silico approaches, this approach is biologically more relevant and able to pharmacologically correlate compounds regard-less of their structures. This is an untapped resource, reliable and readily exploitable for drug repurposing against surprising viral outbreaks.

Abstract

A novel bioinformatic approach for drug repurposing against emerging viral epidemics like Covid-19 is described. It exploits the COMPARE algorithm, a public program from the NCI to sort drugs according to their patterns of growth inhibitory profiles from a diverse panel of human cancer cell lines. The data repository of the NCI includes the growth inhibitory patterns of more than 55000 molecules. When drugs with purported anti-SARS-CoV-2 activities were used as seeds (e.g., hydroxychloroquine and ritonavir) in COMPARE, the analysis uncovered several drugs with fingerprints similar to the seeded drugs. Interestingly, the uncovered drugs were all reportedly known to exert antiviral activities, confirming that COMPARE analysis is valuable for predicting antiviral drug repurposing. Unlike pure in-silico approaches, this approach is biologically more relevant and able to pharmacologically correlate compounds regardless of their structures. This is an untapped resource, reliable and readily exploitable for drug repurposing against surprising viral outbreaks.

Keywords: COVID-19; SARS-CoV-2; Antiviral; Drug Discovery, Algorithm; Bioinformatics; In-Silico; COMPARE Analysis; Drug Repurposing;

Introduction

The surprising outbreak of Covid-19 pandemic alarmed the world about the need for an agile approach to quickly tackle and mitigate unprecedented pathogenic diseases. Accelerated drug discovery by repurposing existing therapeutics is the most practical strategy because fast development of totally new drugs and vaccines can be hampered by the requirement of lengthy safety studies and regulatory processes. Most of all, newly developed vaccines and therapeutics can quickly lose their value if the pathogen is progressively mutating over the course of time, making fresh development of safe and effective vaccines and therapeutics a very difficult

goal to attain. This study elaborates on the effectiveness of COMPARE analysis, an algorithm used in anticancer drug screening, for uncovering and repurposing effective compounds against emerging viruses like SARS-CoV-2. COMPARE algorithm was originally introduced by the Developmental Therapeutic Program (DTP) of the National Cancer Institute (NCI) and is publicly accessible (1, 2). The DTP and its Japanese version, the Disease Oriented Screening (DOS) (3), were established nearly three decades ago in the hope of identifying the most effective compounds against specific cancers. The concept of these programs is based on relating in vitro growth inhibition of tested compounds (GI50) on a panel of diverse human cancer cell types with performance against corresponding clinical cancers (Figure 1).

Apart from the original objective of these programs, mining the collected data using bioinformatics can uncover new facts on pharmacological modes, biological traits, and entirely new prospects for anticancer drug-discovery (4). The author previously used the Japanese DOS program to establish molecules with strong inhibitory effects against the enzyme, telomerase (5, 6). In this study, the author demonstrates that COMPARE analysis using the current DTP data repository can be also useful for fast and low cost antiviral drug discovery. Indeed, by seeding compounds with presumed effects against SARS-CoV-2 virus in COMPARE, several correlated compounds (preys) with approved status for clinical use were easily identified. The value of this efficient resource seems to have been overlooked during the rush for discovering active antiviral drugs that are urgently needed to combat new epidemics like Covid-19.

Methods

The current version of DTP data repository comprises an aggregate of more than 88000 compounds tested on a panel of 59 different human cancer cell lines using sulforhodamine B assay (7). Within the repository, GI50 data from ~55000 compounds is publicly accessible. The SRB assay is by design a quantitative reporter on protein mass and is directly correlated with cell growth and division rate. The used cell panel is highly diverse and represents a variety of cancers representing leukemia, melanoma, lung, colon, brain, ovary, breast, prostate, and kidney malignancies. The full description of the cell panel and the screening method are described in details in the DTP link https://dtp.cancer.gov/discovery_development/nci-60/methodology.htm. Typically, all cell lines are grown on RPMI 1640 growth medium supplemented with 5% fetal bovine serum. Chemo-sensitivity is performed using a 48-h assay and then GI50 values are collected for each tested compound on the 59 different cell lines.

Data Collection was conducted by accessing the NCI's DTP link at <https://dtp.cancer.gov/databases.-tools/compare.htm>, then clicking on the tab [New Technology PUBLIC COMPARE]. This will open the COMPARE algorithm page <https://nci60.cancer.gov/publiccompare/>.

Running the COMPARE analysis was performed by identifying the NSC number (National Service Center number) for each drug molecule from the PubChem website (<https://pubchem.ncbi.nlm.nih.gov/>). For seeding a compound in COMPARE algorithm, the corresponding NSC code is entered in the "NSCs - any delimiter" tab, also the "GI50 data endpoint" box is checked. With these parameters, "SEARCH AS CONFIGURED BELOWS" can be clicked. After COMPARE algorithm finishes identifying the fingerprint of the seeded compound, it will ask for the option of Pearson minimum correlation to be selected in order to call and present the most correlated compounds. Once Pearson value is selected (e.g., >0.5), the algorithm will return a list of the correlated compounds ranked in a table with actual structures, CAS numbers (when available), SMILE codes and NSC codes. The results can be converted to excel. In order to easily convert the collected NSCs of the correlated compounds to useful information about their commercial availability, detailed identity and origin, the list is copied from excel sheet and then inserted into the search window of the site: Enhanced NCI Database Browser 2.2 (<https://cactus.nci.nih.gov/ncidb2.2/>). This site can convert a batch of NSC numbers to the corresponding molecules with useful details on identifiers and sources.

Results and Discussion

While the SRB assay is considered simple and straightforward, in effect, it accurately and categorically reports on the numerous events involved in the cellular growth machinery where hundreds of intra and extracellular events are involved. For example, it is reported that more than 850 genes are operating in

only cell cycle regulation activities (8). Any exogenous compound that changes the readout from the SRB assay is in effect a compound that interferes with one or more of the numerous events involved in cell growth. Meanwhile, the use of 59 different cell lines from various tissues warrants sufficient heterogeneity of the modes of interactions between each tested compound and the cell panel. It is well established in pharmacology that the molecular modes of interactions between drug molecules and biological targets are intrinsic properties produced by the various chemical attributes and descriptors of drug molecules like electron density, polarity, charge, hydrogen bonding, aromaticity, size and conformation (9). When a given compound is tested on the cell panel, its impact on the growth rate of each cell line is governed by a plethora of interactions and events like binding to membrane receptors and transporters, signaling, checkpoints, suppression or activation of enzymes, nucleic acids synthesis and the entire machineries involved in cell growth and division. Once the growth inhibitory values are plotted as mean graphs, a signature response (a pharmacological code or fingerprint) unique to each compound is generated. Figures 1 and 2 are descriptions on how the variable GI50 values from the cell panel encrypt a unique fingerprint for each tested compound. COMPARE algorithm was developed to enable the identification and stratification of compounds presenting similar mean graph fingerprints. What has been clearly observed is that the similarity of fingerprints is a reflection of similarities in the pharmacological mode(s) of action. This is because of the unique set of interactions between a given compound and a given type of cancer cells. In other words, for every biological readout, e.g., cell growth rate, there is a labyrinth of known and unknown multitude of events, all acts at variable levels in response to interaction(s) with the applied exogenous compound. The DTP's large pool of molecules (~55 000 thus far) and the heterogeneity of the cell panel both form a rich source for identifying drug similarities in pharmacological activities. Therefore, any compound with unknown pharmacological effect can be used as a seed molecule in COMPARE algorithm to enable the discovery of compounds that share similar fingerprints and hence have similar mode(s) of action. This approach has been used to shed light on the modes of actions of new anti-cancer drugs (4-6, 10). The current study shows that the same COMPARE-based approach can also be exploited for discerning previously unknown pharmacological activities of existing drugs as long as the screened drug molecules interfere with cell growth processes. Since viruses are known to affect host cells and hijack cellular machineries particularly those machineries involved in cell cycle and protein synthesis, the author demonstrates here that COMPARE analysis can also be useful for antiviral drug discovery and can be applied to uncover drugs suitable for repurposing against the SARS-CoV-2 virus.

Seed compounds with presumed effect against Covid-19 were chosen from ongoing clinical experience and reports from professional sites like Medscape and Cure ID. Among these compounds, chloroquine, hydroxychloroquine and ritonavir (11) were found enlisted in the publically accessible ~55000 chemicals screened by the DTP. Each of the three candidates was used as a seed "a bait" and a Pearson correlation factor of >0.55 was set for each screening attempt. Using the above criteria, the algorithm typically returned 100-250 compounds per seed. Only compounds that are clinically relevant or considered as an active principle in commonly used medicinal plants are shown in this study. Because hydroxychloroquine as a seed yielded only two clinically relevant compounds (lapachol, an active principle in antimalarial and antiviral medicinal plants (12), and ethacizine, an antiarrhythmic and psychotropic agent (13)), a second cycle of COMPARE analysis was conducted on each of these two compounds. As shown in table 1, several known natural and synthetic compounds were extracted based on mean graph similarities. They all belong to one of the following pharmacological classes: anti-viral, anti-parasitic, anti-arrhythmic, and psychotropic agents. Interestingly, the pharmacological classes of the extracted compounds are the same classes that were described in recent screening studies seeking potentially effective therapeutics for viral infections like SARS-CoV-2 (14-17). In particular, it is striking that the extracted compounds in this study are very similar to the compounds identified by the recent extensive QSAR based anti Covid-19 screening study conducted by a consortium of 49 research centers (17), confirming that the COMPARE-based repurposing strategy presented here is capable of achieving comparable results but in a way that is easier, faster and at a far lower cost as it uses an already established database. The results also validate the concept that the cell type-dependent variability of the GI50 values in the DTP repository is indeed correlated pharmacologically at the molecular level and reflects a unique pharmacological pattern for each compound. It is noteworthy that although the resulting hits share common pharmacological effects, they do not always share common pharmacophore structures,

indicating that the conventional view of using dedicated pharmacophores in QSAR *in-silico* drug discovery studies may not always be relevant and the results from COMPARE screening have more biological relevance. Observations on the extracted compounds are summarized in table 1. Of particular interest are the currently evaluated compounds for the treatment of Covid-19, like retinol (16, 18, 19), methotrexate (20, 21), and didanosine (21). Interesting to see also that the psychotropic compounds ethacizine, a phenothiazine with antiarrhythmic effects, and thiothixene, a thioxanthene are in the list. This type of compounds was found to be effective in inhibiting virus entry and virus-cell fusion (22) This finding further validates the vision that COMPARE is sophisticated enough to enable the identification of compounds with clinical relevance even when they belong to different pharmacological classes. Also interesting to notice that flucloxacillin is among the extracted molecules being a synthetic penicillin. This suggests that the currently trialed penicillin antibiotics on Covid-19 patients (e.g., piperacillin and amoxicillin, CURE ID updates)) might be assisting the patients via additional mechanisms beyond their anti-bacterial effects. Finally, although this study is not providing direct experimental evidence on the anti SARS-CoV-2 effect of the compounds in Table 1, the fact that all the identified compounds are reportedly involved in antiviral effects proves that this approach is valid and accurate in guiding drug repurposing efforts.

Conclusion

Despite the fact that the purpose of the DTP repository and COMPARE analysis is anticancer drug discovery, it is demonstrated here that these unique assets of bioinformatics are valuable also in guiding drug repurposing to combat new viral diseases like SARS-CoV-2. Unlike QSAR-based *in-silico* screening methods, the COMPARE-based approach is biologically more relevant, easier, faster, and more economical. Moreover, this approach is flexible and can continuously evolve to respond to mutational changes or potential surge of infections. A variety of compounds were singled out using this approach. Some are existing drugs and can be quickly evaluated on volunteer patients. Others are components of existing medicinal plants already known for antiviral activities. The findings also warrant the need for extending the DTP repository to include a wider range of clinically approved therapeutics so that more possibilities are allowed. Finally, the study urges health response authorities to consider COMPARE analysis and DTP as an additional tool in the fight against current and future viral outbreaks.

Abbreviations

DTP, Developmental Therapeutic Program; NCI, National Cancer Institute; SRB, Sulforhodamine B; DOS; Disease Oriented Screening, QSAR, Quantitative Structure Activity Relationship; GI50, 50-percent growth-inhibitory concentration.

Acknowledgement

The author is grateful to Dr Mark W. Kunkel of the NCI for the helpful information on the latest versions of COMPARE algorithm.

References

1. Paull, K. D., Shoemaker, R. H., Hodes, L., Monks, A., Scudiero, D. A., Rubinstein, L., Plowman, J., & Boyd, M. R. (1989). Display and analysis of patterns of differential activity of drugs against human tumor cell lines: development of mean graph and COMPARE algorithm. *Journal of the National Cancer Institute*, 81(14), 1088–1092.
2. Monks, A., Scudiero, D., Skehan, P., Shoemaker, R., Paull, K., Vistica, D., Hose, C., Langley, J., Cronise, P., & Vaigro-Wolff, A. (1991). Feasibility of a high-flux anticancer drug screen using a diverse panel of cultured human tumor cell lines. *Journal of the National Cancer Institute*, 83(11), 757–766.
3. Yamori T. *Gan To Kagaku Ryoho*. 1997;24(2):129-135.
4. Weinstein, J. N., Myers, T. G., O'Connor, P. M., Friend, S. H., Fornace, A. J., Jr, Kohn, K. W., Fojo, T., Bates, S. E., Rubinstein, L. V., Anderson, N. L., Buolamwini, J. K., van Osdol, W. W., Monks, A. P., Scudiero, D. A., Sausville, E. A., Zaharevitz, D. W., Bunow, B., Viswanadhan, V. N., Johnson, G. S., Wittes, R. E., ... Paull, K. D. (1997). An information-intensive approach to the molecular

- pharmacology of cancer. *Science (New York, N.Y.)*, 275(5298), 343–349.
- Naasani I, Seimiya H, Yamori T, Tsuruo T. FJ5002: a potent telomerase inhibitor identified by exploiting the disease-oriented screening program with COMPARE analysis. *Cancer Res.* 1999;59(16):4004-4011.
 - Naasani I, Yamori T, Tsuruo T. Screening with COMPARE analysis for telomerase inhibitors. *Methods Mol Biol.* 2002;191:197-207.
 - Vichai V, Kirtikara K. Sulforhodamine B colorimetric assay for cytotoxicity screening. *Nat Protoc.* 2006;1(3):1112-1116.
 - Kitajima, Y., Ishii, T., Kohda, T., Ishizuka, M., Yamazaki, K., Nishimura, Y., Tanaka, T., Dan, S., & Nakajima, M. (2019). Mechanistic study of PpIX accumulation using the JFCR39 cell panel revealed a role for dynamin 2-mediated exocytosis. *Scientific reports*, 9(1), 8666.
 - Cherkasov, A., Muratov, E. N., Fourches, D., Varnek, A., Baskin, I. I., Cronin, M., Dearden, J., Gramatica, P., Martin, Y. C., Todeschini, R., Consonni, V., Kuz'min, V. E., Cramer, R., Benigni, R., Yang, C., Rathman, J., Terfloth, L., Gasteiger, J., Richard, A., & Tropsha, A. (2014). QSAR modeling: where have you been? Where are you going to?. *Journal of medicinal chemistry*, 57(12), 4977–5010.
 - Luzina EL, Popov AV. Synthesis, evaluation of anticancer activity and COMPARE analysis of N-bis(trifluoromethyl)alkyl-N'-substituted ureas with pharmacophoric moieties. *Eur J Med Chem.* 2012;53:364-373.
 - Li H, Zhou Y, Zhang M, Wang H, Zhao Q, Liu J. Updated Approaches against SARS-CoV-2. *Antimicrob Agents Chemother.* 2020;64(6):e00483-20. Published 2020 May 21.
 - Sacau, E. P., Estévez-Braun, A., Ravelo, A. G., Ferro, E. A., Tokuda, H., Mukainaka, T., & Nishino, H. (2003). Inhibitory effects of lapachol derivatives on epstein-barr virus activation. *Bioorganic & medicinal chemistry*, 11(4), 483–488.
 - Kaverina NV, Sokolov SF. Pharmacology and clinical use of a new group of antiarrhythmic drugs: derivatives of tricyclic nitrogen-containing systems. *Pharmacol Res.* 1992;25(3):217-225. doi:10.1016/s1043-6618(05)80070-2
 - Bleasel MD, Peterson GM. Emetine, Ipecac, Ipecac Alkaloids and Analogues as Potential Antiviral Agents for Coronaviruses. *Pharmaceuticals (Basel).* 2020;13(3):51. Published 2020 Mar 21.
 - Choy KT, Wong AY, Kaewpreedee P, et al. Remdesivir, lopinavir, emetine, and homoharringtonine inhibit SARS-CoV-2 replication in vitro. *Antiviral Res.* 2020;178:104786
 - Riva, L., Yuan, S., Yin, X., Martin-Sancho, L., Matsunaga, N., Burgstaller-Muehlbacher, S., Pache, L., De Jesus, P. P., Hull, M. V., Chang, M., Chan, J. F., Cao, J., Poon, V. K., Herbert, K., Nguyen, T. T., Pu, Y., Nguyen, C., Rubanov, A., Martinez-Sobrido, L., Liu, W. C., ... Chanda, S. K. (2020). A Large-scale Drug Repositioning Survey for SARS-CoV-2 Antivirals. *bioRxiv : the preprint server for biology*, 2020.04.16.044016.
 - Gordon, D. E., Jang, G. M., Bouhaddou, M., Xu, J., Obernier, K., White, K. M., O'Meara, M. J., Rezelj, V. V., Guo, J. Z., Swaney, D. L., Tummino, T. A., Huettenhain, R., Kaake, R. M., Richards, A. L., Tutuncuoglu, B., Foussard, H., Batra, J., Haas, K., Modak, M., Kim, M., ... Krogan, N. J. (2020). A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature*, 10.1038/s41586-020-2286-9. Advance online publication.
 - Trasino S. E. (2020). A role for retinoids in the treatment of COVID-19?. *Clinical and experimental pharmacology & physiology*, 10.1111/1440-1681.13354. Advance online publication.
 - Yuan, S., Chan, J., Chik, K., Chan, C., Tsang, J., Liang, R., Cao, J., Tang, K., Chen, L. L., Wen, K., Cai, J. P., Ye, Z. W., Lu, G., Chu, H., Jin, D. Y., & Yuen, K. Y. (2020). Discovery of the FDA-approved drugs bexarotene, cetilistat, diiodohydroxyquinoline, and abiraterone as potential COVID-19 treatments with a robust two-tier screening system. *Pharmacological research*, 159, 104960. Advance online publication.
 - Beck, S., Zhu, Z., Oliveira, M. F., Smith, D. M., Rich, J. N., Bernatchez, J. A., & Siqueira-Neto, J. L. (2019). Mechanism of Action of Methotrexate Against Zika Virus. *Viruses*, 11(4), 338.
 - Perricone, C., Triggianese, P., Bartoloni, E., Cafaro, G., Bonifacio, A. F., Bursi, R., Perricone, R.,

- & Gerli, R. (2020). The anti-viral facet of anti-rheumatic drugs: Lessons from COVID-19. *Journal of autoimmunity*, 111, 102468.
22. Dyllal, J., Coleman, C. M., Hart, B. J., Venkataraman, T., Holbrook, M. R., Kindrachuk, J., Johnson, R. F., Olinger, G. G., Jr, Jahrling, P. B., Laidlaw, M., Johansen, L. M., Lear-Rooney, C. M., Glass, P. J., Hensley, L. E., & Frieman, M. B. (2014). Repurposing of clinically developed drugs for treatment of Middle East respiratory syndrome coronavirus infection. *Antimicrobial agents and chemotherapy*, 58(8), 4885–4893.
 23. Chander V, Aswal JS, Dobhal R, Uniyal DP. A review on Pharmacological potential of Berberine; an active component of Himalayan *Berberis aristata*. *The Journal of Phytopharmacology* 2017; 6(1): 53-58.
 24. Enkhtaivan G, Maria John KM, Ayyanar M, Sekar T, Jin KJ, Kim DH. Anti-influenza (H1N1) potential of leaf and stem bark extracts of selected medicinal plants of South India. *Saudi J Biol Sci*. 2015;22(5):532-538.

Hosted file

image1.wmf available at <https://authorea.com/users/347394/articles/473106-compare-analysis-as-an-efficient-bioinformatic-approach-to-accelerate-repurposing-of-existing-drugs-against-covid-19-and-other-emerging-epidemics>

Hosted file

image2.wmf available at <https://authorea.com/users/347394/articles/473106-compare-analysis-as-an-efficient-bioinformatic-approach-to-accelerate-repurposing-of-existing-drugs-against-covid-19-and-other-emerging-epidemics>

Hosted file

image3.emf available at <https://authorea.com/users/347394/articles/473106-compare-analysis-as-an-efficient-bioinformatic-approach-to-accelerate-repurposing-of-existing-drugs-against-covid-19-and-other-emerging-epidemics>

Hosted file

image4.wmf available at <https://authorea.com/users/347394/articles/473106-compare-analysis-as-an-efficient-bioinformatic-approach-to-accelerate-repurposing-of-existing-drugs-against-covid-19-and-other-emerging-epidemics>