

Using the SARS infection transcriptional signature to identify potential treatments for Covid-19

Richard Killick^a, Clive Ballard^b, Patrick Doherty^c and Gareth Williams^{c*}

a Maurice Wohl Clinical Neuroscience Institute, King's College London, UK

b University of Exeter Medical School, College of Medicine and Health, University of Exeter, UK

c Wolfson Centre for Age-related Diseases, King's College London, UK

Corresponding author: Gareth Williams (gareth.2.williams@kcl.ac)

Abstract

In the light of the SARS-CoV-2 pandemic there is a pressing need to trial as wide a range of drugs as possible as there is at present no effective treatments. There is limited time to develop drugs according to the standard pipeline. This makes repurposing, where existing therapeutics with known side-effect profiles and pharmacokinetic data are deployed for diseases other than those they were originally developed to ameliorate, an attractive endeavour. Here, publicly available gene expression data associated with SARS infection of human epithelial cell lines from multiple independent studies has been harnessed to delimit a SARS signature profile. The SARS signature exhibits a strong up-regulation of gene sets involved in viral defence mechanisms. The SARS signature has then been queried against gene expression data bases of cellular transcriptional response to approved drugs and candidate therapeutics hypothesised on the basis of their ability to bolster the host cell response to SARS infection. Candidate drugs thus identified, show a marked enrichment for reported anti-viral activity, with a number also reported to exhibit potent activity against SARS-CoV-2. It is hoped that this approach may lead to an effective and rapid repurposing strategy for treating the Covid-19 pandemic.

195 words

Introduction

The rapid evolution rate of certain coronavirus strains has, over the recent past, led to increasingly severe zoonotic human epidemics. In 2002 a severe acute respiratory syndrome (SARS) emerged in China resulting in 8,000 deaths (10% of infections) across 29 countries. Ten years later the more deadly (36% of infections) but less infectious middle east respiratory syndrome (MERS) developed in Saudi Arabia spreading to 27 countries with total infections currently at 2494 with 858 deaths. SARS-CoV-2, which emerged in China in 2019, appears to be less deadly than SARS-CoV or MERS-CoV, with current estimates of between 1 and 2%, but it has proved to be highly infectious with current infections standing at nearly four and half million and as many as 314,000 fatalities as of May 2020.

Efforts are underway to develop a SARS-CoV-2 vaccine (Amanat and Krammer, 2020), while inhibiting viral proteases (Jin et al., 2020) and the recent solution of the structure of the SARS-CoV-2 spike protein bound with its accessory protein ACE2 (Lan et al., 2020) offers the possibility of developing agents to block viral entry. However, the current pandemic trajectory calls for a more rapid intervention deployment. To this end the broad spectrum anti-viral Remdesivir targeting RNA polymerisation has been approved for emergency use in the USA and for severe cases in Japan (Ko et al., 2020). Recently, a more ambitious initiative has sought to repurpose existing therapeutics, not necessarily of an anti-viral class, for SARS-CoV-2 based on compiling evidence for 332 separate interactions between viral and host cell proteins (Gordon et al., 2020). This database was then used to identify existing drugs for potential repurposing based on their targeting the human proteins that interact with the viral proteins.

Drug repurposing is an attractive strategy because it circumvents the formidable hurdles that are inevitable with the exploration of novel chemical entities that result from high throughput screening of compound libraries against defined targets. With repurposing the target is rather a phenotype as there is little hope of finding a specific high affinity compound against a specified target from a limited set of a few 1000 drugs. Phenotype-based screening requires the development of high content quantitative descriptor of the underlying biological disease state and one such is transcriptional profiling. In addition to the 3.5 million samples curated by NCBI GEO (Barrett et al., 2007) there are large collections of publicly available compound associated expression databases developed as part

of the connectivity map (CMAP) project (Lamb et al., 2006). The CMAP2.0 database contains the transcriptional profiles of 1,309 drug-like compounds and therapeutics profiled across four human cancer cell lines (www.broadinstitute.org/connectivity-map-cmap). The CMAP database has recently been greatly expanded to include profiles for over 15,000 compounds defined across almost 100 different cell lines (Subramanian et al., 2017). The latter LINCS database is compiled from the expression levels for 1000 landmark genes from which the values for the entire genome are imputed with linear modelling. In this sense the two datasets are complimentary.

Transcription-based drug repurposing is founded upon the observation that transcriptional changes seen in disease states are largely recapitulated across independent samples and thereby can serve as effective disease descriptors or quantitative phenotypes (Golub et al., 1999; Lee and Young, 2013). This led the way for the hypothesis that drugs tending to reverse disease associated gene expression changes may thereby function as disease modulators and potential therapeutics (Hughes et al., 2000; Marton et al., 1998; Walf-Vorderwulbecke et al., 2018; Wei et al., 2006; Williams, 2012; Zhang et al., 2012). However, in some cases the expression changes seen in disease states, at least in the initial response phase, are beneficial and it could therefore be counterproductive for these changes to be reversed. In the particular case of viral infection, there is an established cellular response that may benefit from being enhanced by therapeutic intervention rather than inhibited.

The transcriptional consequences of viral infection have been assayed in multiple cellular contexts and animal models (Mitchell et al., 2013; Webster et al., 2009; Yoshikawa et al., 2010). The motivation for the present work is to harness this rich data source and; firstly to assess the extent to which gene expression changes are robust, being recapitulated across multiple studies; secondly, to inspect the class of genes being perturbed by viral infection and assess whether this response is a defence mechanism or a facilitating alteration of the cellular genome. Once the nature of the cellular response to the virus has been established compound libraries can be queried for candidate therapeutics on the basis of either reversing the virus induced changes or enhancing them.

The analysis presented below establishes a clear conservation of gene expression changes across independent SARS infection data sets for multiple human epithelial cell types. This leads on to the development of a SARS signature profile (SISP) of 192 up-regulated genes and 136 down-regulated genes. The biological nature of this response is investigated through an enrichment analysis against established pathway gene sets and gene ontology classes. It is shown that the SISP is dominated by a viral defence mechanism and it is therefore concluded that a drug repositioning endeavour should be based on recapitulating this response. To this end the SISP is queried against the CMAP and LINCS databases and a set of candidate therapeutics proposed. It is strikingly clear that drugs positively correlating with the viral response tend to have established anti-viral activities, with a number already shown to protect against SARS-CoV-2 infection in the low micromolar range in *in vitro* cell-based assays (Choy et al., 2020; Jeon et al., 2020).

Methods

Transcriptional data for SARS infection of human cells was obtained by querying NCBI GEO with relevant key words. In total 17 SARS-associated profiles have been generated sourced from two RNAseq series (GSE14507 (Webster et al., 2009) and GSE148729) and four microarray series (GSE17400 (Yoshikawa et al., 2010), GSE37827, GSE47963 (Mitchell et al., 2013), GSE48142). Profiles were generated based on grouping treatment and control samples and performing a student's t-test $p < 0.05$ filter. Profile correlation scores were generated by simple linear regression and measured with a Pearson correlation coefficient and significance given by the corresponding Z-score.

The relationship of the SISP to deposited transcriptional data was established through SPIED searches (www.spied.org.uk) of human and mouse samples. Pathway analysis was based on calculating the exact Fisher test enrichment of the up and down-regulated SISP genes for the Reactome dataset (Fabregat et al., 2017). Gene ontology (Carbon et al., 2009) enrichment analysis was performed in the same manner and based on gene sets downloaded from geneontology.org/.

CMAP profiles were generated as previously described (Williams et al., 2019). LINCS profiles were generated based on the deposited dataset series GSE92742 and GSE70138. Data is available in the form of Z scores for gene expression relative to plate average. The LINCS portal of SPIED hosts profiles in the form of categorical calls on the up/down status of genes generated based on combining expression data for drug/cell replicates. In the present work the LINCS profiles for different cell types are combined for each drug based on a given gene being assigned an up or down regulated status based on a majority vote across the cell types. Correlations of drug profiles against SISP were defined based on an exact Fisher test for the up/down category tables.

Results

The SARS infection signature profile (SISP)

There is a high degree of correlation between the 17 profiles demonstrated by the significant positive Pearson correlations and the associated Z scores, see Table 1. The high degree of mutual similarity facilitates the definition of a composite signature for SARS infection based on at least 2/3 of genes being modulated in the same direction. This consists of 192 up regulated and 136 down regulated genes. The spectrum of expression changes for the genes across the profiles is shown in Table 2. The pathways and gene ontology classes associated with the up and down regulated genes are shown in Tables 3 and 4. As expected, the up regulated set is dominated by an immune response. In particular, pathways associated with interferon signalling, cytokine response and anti-viral mechanisms dominate the significantly enriched pathways. This is largely recapitulated with gene ontology class set enrichment analysis. In contrast, the down regulated gene set harbours fewer significantly enriched pathways and ontology classes, with the cell cycle set enriched as a pathway and ontology class.

SISP in a wider context

The SISP profile can be examined in the wider context of deposited transcriptional data through search engines like SPIED (Williams, 2013). The results are shown in Table 5 and highlight a similarity to other virus induced gene expression changes. It is important to note here that this diversity in virus type with similar transcriptional effects upon diverse cell types is a reflection of the broad-brush nature of the signature. When it comes to delimiting candidate therapeutics based on the SISP it is to be expected that drugs with general, as well as specific, anti-viral properties will be returned. The repurposing methodology does not have the aim to discover drugs with specificity against novel targets, which is rather the aim of the traditional and protracted target-based novel chemical entity approach. In contrast, repurposing aims to harness the global phenotypic effects of drugs with the hope that these will tend to shift the underlying biology, in the present case by enhancing the cells natural defence response.

Drug repurposing with the SISP

The top correlating drugs from CMAP and LINCS searches are listed in Table 6. There is considerable overlap in the CMAP and LINCS outputs, see column on the right of Table 6. As expected, there is a conspicuous enrichment for drugs with established anti-viral activities, these are highlighted in yellow.

In particular, there are 20 compounds with published antiviral activity in the top 45 CMAP hits and 6 of these compounds are also in the top 45 LINCS hits, which has an additional 16 compounds with established antiviral activity, see Table 6 for references. Furthermore, four compounds that are hits for both CMAP and LINCS, Emetine, Ouabain, Digoxin and Niclosamide, have each been found to display potent anti-SARS-CoV-2 activities. In the report from Choy et al., 2020, Emetine effectively inhibited SARS-CoV-2 replication in Vero cells at 0.5 μM , while in that same study the top hit in the LINCS set, homoharringtonine, inhibited the virus with an EC_{50} of 2.10 μM (Choy et al., 2020). In a second report, (Jeon et al., 2020), Ouabain, Digoxin and Niclosamide each potently inhibited SARS-CoV-2 replication, again in Vero cells, with IC_{50} 's of 0.097 μM , 0.19 μM and 0.28 μM , respectively (Jeon et al., 2020). Also tested by Jeon et al., were two further drugs present at 4th and 9th position in the CMAP data set, Terfenadine and Mefloquine, which exhibited IC_{50} 's of 3.0 μM and 4.33 μM , respectively.

Conclusion

The SARS-CoV-2 pandemic is the most serious medical emergency since the Spanish flu of over 100 years ago. At present there are ongoing efforts to develop a vaccine and harness the increasing amount known about the mechanism of SARS-CoV-2 infection in terms of structure and protein targets to embark on a drug discovery pipeline. In the interim researchers have started to look at repurposing existing therapeutics, whether these be broad spectrum anti-viral drugs or more general therapeutics with indirect beneficial effects. The present paper seeks to highlight the effectiveness of transcriptional profiling to firstly define the activated pathways upon viral infection and secondly to use the infection associated expression changes as a quantitative phenotype to query drug databases for possible therapeutics. This approach has identified a broad panel of therapeutic entities ranging from plant extracts; the expectorant (and at higher doses, emetic) emantine from the rhizome of *Carapichea ipecacuanha* and the foxglove phytosteroid, digoxin, used for atrial fibrillation; through antibiotics such as puromycin, an aminonucleoside derived from the bacterium *Streptomyces albiniger*, through to compounds such as thioridazine, a first generation anti-psychotic and the synthetic thiazine dye, methylene-blue.

The cellular transcriptional response to viral infection has been assembled from multiple independent publicly available data sets and shown to harbour a consistent set of viral response genes that are transcribed in response to insult. This together with the extensive coverage of the drug associated transcriptome suggests a repurposing strategy where drugs tending to potentiate the cellular viral response are hypothetical candidate therapeutics. The results presented here largely validate this hypothesis. Two large drug driven transcription data sets were queried with the SISP viral response signature with the results that many drugs with established antiviral activities were high scoring candidates. This observation on the one hand offers a novel approach to identify drugs with no previous experimental data on their antiviral potential for trial, but also, reveals an intriguing unifying principle or at least effective biomarker underlying a disparate set of established antiviral agents.

That among the drugs identified by this approach six have already been shown, by others, to have potent anti-SARS-CoV-2 activities (Choy et al., 2020; Jeon et al., 2020), surely lends some validity of using transcriptional profiling as a useful way to rapidly identify effective treatment for the current pandemic and potentially for other viral pandemics that will inflict mankind in the future. While some of these "actives" may not readily lend themselves as treatments for Covid-19, such as the emetic, ementine, and ouabain, others, niclosamide and digoxin may. Given this high "hit rate" of true actives from the small number of our data set tested, it now warrants the rapid assessment of other candidates by those with access to SARS-CoV-2 assays with the view to testing the most suitable in clinical trials as a matter of urgency.

References

- Amanat, F., and Krammer, F. (2020). SARS-CoV-2 Vaccines: Status Report. *Immunity* 52, 583-589.
- Andersen, P.I., Krpina, K., Ianevski, A., Shtaida, N., Jo, E., Yang, J., Koit, S., Tenson, T., Hukkanen, V., Anthonen, M.W., *et al.* (2019). Novel Antiviral Activities of Obatoclox, Emetine, Niclosamide, Brequinar, and Homoharringtonine. *Viruses* 11.
- Barrett, T., Troup, D.B., Wilhite, S.E., Ledoux, P., Rudnev, D., Evangelista, C., Kim, I.F., Soboleva, A., Tomashevsky, M., and Edgar, R. (2007). NCBI GEO: mining tens of millions of expression profiles--database and tools update. *Nucleic Acids Res* 35, D760-765.
- Brickelmaier, M., Lugovskoy, A., Kartikeyan, R., Reviriego-Mendoza, M.M., Allaire, N., Simon, K., Frisque, R.J., and Gorelik, L. (2009). Identification and characterization of mefloquine efficacy against JC virus in vitro. *Antimicrob Agents Chemother* 53, 1840-1849.
- Carbon, S., Ireland, A., Mungall, C.J., Shu, S., Marshall, B., Lewis, S., Ami, G.O.H., and Web Presence Working, G. (2009). AmiGO: online access to ontology and annotation data. *Bioinformatics* 25, 288-289.
- Chamoun-Emanuelli, A.M., Pecheur, E.I., Simeon, R.L., Huang, D., Cremer, P.S., and Chen, Z. (2013). Phenothiazines inhibit hepatitis C virus entry, likely by increasing the fluidity of cholesterol-rich membranes. *Antimicrob Agents Chemother* 57, 2571-2581.
- Cheung, Y.Y., Chen, K.C., Chen, H., Seng, E.K., and Chu, J.J. (2014). Antiviral activity of lanatoside C against dengue virus infection. *Antiviral Res* 111, 93-99.
- Choy, K.T., Wong, A.Y., Kaewpreedee, P., Sia, S.F., Chen, D., Hui, K.P.Y., Chu, D.K.W., Chan, M.C.W., Cheung, P.P., Huang, X., *et al.* (2020). Remdesivir, lopinavir, emetine, and homoharringtonine inhibit SARS-CoV-2 replication in vitro. *Antiviral Res* 178, 104786.
- DeWald, L.E., Dyal, J., Sword, J.M., Torzewski, L., Zhou, H., Postnikova, E., Kollins, E., Alexander, I., Gross, R., Cong, Y., *et al.* (2018). The Calcium Channel Blocker Bepridil Demonstrates Efficacy in the Murine Model of Marburg Virus Disease. *J Infect Dis* 218, S588-S591.
- Diamantopoulos, P.T., Michael, M., Benopoulou, O., Bazanis, E., Tzeletas, G., Meletis, J., Vayopoulos, G., and Viniou, N.A. (2012). Antiretroviral activity of 5-azacytidine during treatment of a HTLV-1 positive myelodysplastic syndrome with autoimmune manifestations. *Virol J* 9, 1.
- Dong, H.J., Wang, Z.H., Meng, W., Li, C.C., Hu, Y.X., Zhou, L., and Wang, X.J. (2018). The Natural Compound Homoharringtonine Presents Broad Antiviral Activity In Vitro and In Vivo. *Viruses* 10.
- Fabregat, A., Sidiropoulos, K., Viteri, G., Forner, O., Marin-Garcia, P., Arnau, V., D'Eustachio, P., Stein, L., and Hermjakob, H. (2017). Reactome pathway analysis: a high-performance in-memory approach. *BMC Bioinformatics* 18, 142.
- Gabrielsen, B., Monath, T.P., Huggins, J.W., Kefauver, D.F., Pettit, G.R., Groszek, G., Hollingshead, M., Kirsi, J.J., Shannon, W.M., Schubert, E.M., *et al.* (1992). Antiviral (RNA) activity of selected Amaryllidaceae isoquinoline constituents and synthesis of related substances. *J Nat Prod* 55, 1569-1581.
- Golub, T.R., Slonim, D.K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J.P., Coller, H., Loh, M.L., Downing, J.R., Caligiuri, M.A., *et al.* (1999). Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 286, 531-537.
- 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., *et al.* (2020). A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature*.
- Grover, A., Agrawal, V., Shandilya, A., Bisaria, V.S., and Sundar, D. (2011). Non-nucleosidic inhibition of Herpes simplex virus DNA polymerase: mechanistic insights into the anti-herpetic mode of action of herbal drug withaferin A. *BMC Bioinformatics* 12 Suppl 13, S22.
- Hughes, T.R., Marton, M.J., Jones, A.R., Roberts, C.J., Stoughton, R., Armour, C.D., Bennett, H.A., Coffey, E., Dai, H., He, Y.D., *et al.* (2000). Functional discovery via a compendium of expression profiles. *Cell* 102, 109-126.

Jablonowski, H. (2005). [Antiviral therapy: when to begin, which goals should be set, what must be kept in mind? Tips for nonspecialists]. *MMW Fortschr Med* 147 Spec No 1, 10-11.

Jeon, S., Ko, M., Lee, J., Choi, I., Byun, S.Y., Park, S., Shum, D., and Kim, S. (2020). Identification of antiviral drug candidates against SARS-CoV-2 from FDA-approved drugs. *Antimicrob Agents Chemother*.

Jin, Z., Zhao, Y., Sun, Y., Zhang, B., Wang, H., Wu, Y., Zhu, Y., Zhu, C., Hu, T., Du, X., *et al.* (2020). Structural basis for the inhibition of SARS-CoV-2 main protease by antineoplastic drug carmofur. *Nat Struct Mol Biol*.

Jurgeit, A., McDowell, R., Moese, S., Meldrum, E., Schwendener, R., and Greber, U.F. (2012). Niclosamide is a proton carrier and targets acidic endosomes with broad antiviral effects. *PLoS Pathog* 8, e1002976.

Ko, W.C., Rolain, J.M., Lee, N.Y., Chen, P.L., Huang, C.T., Lee, P.I., and Hsueh, P.R. (2020). Arguments in favour of remdesivir for treating SARS-CoV-2 infections. *Int J Antimicrob Agents* 55, 105933.

Kumar, N., Sharma, N.R., Ly, H., Parslow, T.G., and Liang, Y. (2011). Receptor tyrosine kinase inhibitors that block replication of influenza A and other viruses. *Antimicrob Agents Chemother* 55, 5553-5559.

Lama, Z., Gaudin, Y., Blondel, D., and Lagaudriere-Gesbert, C. (2019). Kinase inhibitors tyrphostin 9 and rottlerin block early steps of rabies virus cycle. *Antiviral Res* 168, 51-60.

Lamb, J., Crawford, E.D., Peck, D., Modell, J.W., Blat, I.C., Wrobel, M.J., Lerner, J., Brunet, J.P., Subramanian, A., Ross, K.N., *et al.* (2006). The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* 313, 1929-1935.

Lan, J., Ge, J., Yu, J., Shan, S., Zhou, H., Fan, S., Zhang, Q., Shi, X., Wang, Q., Zhang, L., *et al.* (2020). Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature*.

Lee, T.I., and Young, R.A. (2013). Transcriptional regulation and its misregulation in disease. *Cell* 152, 1237-1251.

Liang, Y., de Wispelaere, M., Carocci, M., Liu, Q., Wang, J., Yang, P.L., and Gray, N.S. (2017). Structure-Activity Relationship Study of QL47: A Broad-Spectrum Antiviral Agent. *ACS Med Chem Lett* 8, 344-349.

Marton, M.J., DeRisi, J.L., Bennett, H.A., Iyer, V.R., Meyer, M.R., Roberts, C.J., Stoughton, R., Burchard, J., Slade, D., Dai, H., *et al.* (1998). Drug target validation and identification of secondary drug target effects using DNA microarrays. *Nat Med* 4, 1293-1301.

Michelini, F.M., Alche, L.E., and Bueno, C.A. (2018). Virucidal, antiviral and immunomodulatory activities of beta-escin and Aesculus hippocastanum extract. *J Pharm Pharmacol* 70, 1561-1571.

Mitchell, H.D., Eisfeld, A.J., Sims, A.C., McDermott, J.E., Matzke, M.M., Webb-Robertson, B.J., Tilton, S.C., Tchitchek, N., Josset, L., Li, C., *et al.* (2013). A network integration approach to predict conserved regulators related to pathogenicity of influenza and SARS-CoV respiratory viruses. *PLoS One* 8, e69374.

Mukherjee, S., Weiner, W.S., Schroeder, C.E., Simpson, D.S., Hanson, A.M., Sweeney, N.L., Marvin, R.K., Ndjomou, J., Kolli, R., Isailovic, D., *et al.* (2014). Ebselen inhibits hepatitis C virus NS3 helicase binding to nucleic acid and prevents viral replication. *ACS Chem Biol* 9, 2393-2403.

Nio, Y., Sasai, M., Akahori, Y., Okamura, H., Hasegawa, H., Oshima, M., Watashi, K., Wakita, T., Ryo, A., Tanaka, Y., *et al.* (2019). Bardoxolone methyl as a novel potent antiviral agent against hepatitis B and C viruses in human hepatocyte cell culture systems. *Antiviral Res* 169, 104537.

Pietschmann, T. (2017). Clinically Approved Ion Channel Inhibitors Close Gates for Hepatitis C Virus and Open Doors for Drug Repurposing in Infectious Viral Diseases. *J Virol* 91.

Porcari, A.R., Ptak, R.G., Borysko, K.Z., Breitenbach, J.M., Drach, J.C., and Townsend, L.B. (2003). Synthesis and antiviral activity of 2-substituted analogs of tricyridine. *Nucleosides Nucleotides Nucleic Acids* 22, 2171-2193.

Premecz, G., Markovits, A., and Foldes, I. (1985). Antiviral activity of phorbol myristate acetate and possible relationships with interferon action. *FEBS Lett* 180, 300-302.

Quintana, V.M., Selisko, B., Brunetti, J.E., Eydoux, C., Guillemot, J.C., Canard, B., Damonte, E.B., Julander, J.G., and Castilla, V. (2020). Antiviral activity of the natural alkaloid anisomycin against dengue and Zika viruses. *Antiviral Res* 176, 104749.

Radloff, R.J., Deck, L.M., Royer, R.E., and Vander Jagt, D.L. (1986). Antiviral activities of gossypol and its derivatives against herpes simplex virus type II. *Pharmacol Res Commun* *18*, 1063-1073.

Ramabhadran, T.V., and Thach, R.E. (1980). Specificity of protein synthesis inhibitors in the inhibition of encephalomyocarditis virus replication. *J Virol* *34*, 293-296.

Schwartz, O., Marechal, V., Friguet, B., Arenzana-Seisdedos, F., and Heard, J.M. (1998). Antiviral activity of the proteasome on incoming human immunodeficiency virus type 1. *J Virol* *72*, 3845-3850.

Shen, L., Niu, J., Wang, C., Huang, B., Wang, W., Zhu, N., Deng, Y., Wang, H., Ye, F., Cen, S., *et al.* (2019). High-Throughput Screening and Identification of Potent Broad-Spectrum Inhibitors of Coronaviruses. *J Virol* *93*.

Simanjuntak, Y., Liang, J.J., Lee, Y.L., and Lin, Y.L. (2015). Repurposing of prochlorperazine for use against dengue virus infection. *J Infect Dis* *211*, 394-404.

Subramanian, A., Narayan, R., Corsello, S.M., Peck, D.D., Natoli, T.E., Lu, X., Gould, J., Davis, J.F., Tubelli, A.A., Asiedu, J.K., *et al.* (2017). A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. *Cell* *171*, 1437-1452 e1417.

Sun, Q., Wu, R., Cai, S., Lin, Y., Sellers, L., Sakamoto, K., He, B., and Peterson, B.R. (2011). Synthesis and biological evaluation of analogues of AKT (protein kinase B) inhibitor-IV. *J Med Chem* *54*, 1126-1139.

Torriani, G., Trofimenko, E., Mayor, J., Fedeli, C., Moreno, H., Michel, S., Heulot, M., Chevalier, N., Zimmer, G., Shrestha, N., *et al.* (2019). Identification of Clotrimazole Derivatives as Specific Inhibitors of Arenavirus Fusion. *J Virol* *93*.

Walf-Vorderwulbecke, V., Pearce, K., Brooks, T., Hubank, M., van den Heuvel-Eibrink, M.M., Zwaan, C.M., Adams, S., Edwards, D., Bartram, J., Samarasinghe, S., *et al.* (2018). Targeting acute myeloid leukemia by drug-induced c-MYB degradation. *Leukemia* *32*, 882-889.

Wang, P., Li, L.F., Wang, Q.Y., Shang, L.Q., Shi, P.Y., and Yin, Z. (2014). Anti-dengue-virus activity and structure-activity relationship studies of lycorine derivatives. *ChemMedChem* *9*, 1522-1533.

Webster, R.J., Giles, K.M., Price, K.J., Zhang, P.M., Mattick, J.S., and Leedman, P.J. (2009). Regulation of epidermal growth factor receptor signaling in human cancer cells by microRNA-7. *J Biol Chem* *284*, 5731-5741.

Wei, G., Twomey, D., Lamb, J., Schlis, K., Agarwal, J., Stam, R.W., Opferman, J.T., Sallan, S.E., den Boer, M.L., Pieters, R., *et al.* (2006). Gene expression-based chemical genomics identifies rapamycin as a modulator of MCL1 and glucocorticoid resistance. *Cancer Cell* *10*, 331-342.

Williams, G. (2012). A searchable cross-platform gene expression database reveals connections between drug treatments and disease. *BMC Genomics* *13*, 12.

Williams, G. (2013). SPIEDw: a searchable platform-independent expression database web tool. *BMC Genomics* *14*, 765.

Williams, G., Gatt, A., Clarke, E., Corcoran, J., Doherty, P., Chambers, D., and Ballard, C. (2019). Drug repurposing for Alzheimer's disease based on transcriptional profiling of human iPSC-derived cortical neurons. *Transl Psychiatry* *9*, 220.

Wong, R.W., Balachandran, A., Ostrowski, M.A., and Cochrane, A. (2013). Digoxin suppresses HIV-1 replication by altering viral RNA processing. *PLoS Pathog* *9*, e1003241.

Xu, J., Shi, P.Y., Li, H., and Zhou, J. (2020). Broad Spectrum Antiviral Agent Niclosamide and Its Therapeutic Potential. *ACS Infect Dis* *6*, 909-915.

Yang, S., Xu, M., Lee, E.M., Gorshkov, K., Shiryaev, S.A., He, S., Sun, W., Cheng, Y.S., Hu, X., Tharappel, A.M., *et al.* (2018). Emetine inhibits Zika and Ebola virus infections through two molecular mechanisms: inhibiting viral replication and decreasing viral entry. *Cell Discov* *4*, 31.

Yoshikawa, T., Hill, T.E., Yoshikawa, N., Popov, V.L., Galindo, C.L., Garner, H.R., Peters, C.J., and Tseng, C.T. (2010). Dynamic innate immune responses of human bronchial epithelial cells to severe acute respiratory syndrome-associated coronavirus infection. *PLoS One* *5*, e8729.

Yu, J.S., Tseng, C.K., Lin, C.K., Hsu, Y.C., Wu, Y.H., Hsieh, C.L., and Lee, J.C. (2017). Celastrol inhibits dengue virus replication via up-regulating type I interferon and downstream interferon-stimulated responses. *Antiviral Res* *137*, 49-57.

Zhang, D., Ciciriello, F., Anjos, S.M., Carissimo, A., Liao, J., Carlile, G.W., Balghi, H., Robert, R., Luini, A., Hanrahan, J.W., *et al.* (2012). Ouabain Mimics Low Temperature Rescue of F508del-CFTR in Cystic Fibrosis Epithelial Cells. *Front Pharmacol* 3, 176.

Tables

	GSE147507-CALU3	GSE147507-A549	GSE148729-CALU3-S1	GSE148729-CALU3-S2	GSE17400-24h	GSE17400-28h	GSE37827-24-30h	GSE37827-36-54h	GSE37827-60-72h	GSE47963-24-48h(2)	GSE47963-60-96h(2)	GSE47963-24-48h(3)	GSE47963-24-48h(4)	GSE47963-60-96h(4)	GSE48142-24-36h	GSE48142-48-72h
GSE147507-NHBE	0.42	0.23	0.44	0.39	0.24	0.27	0.38	0.36	0.43	0.37	0.52	0.18	0.34	0.57	0.41	0.39
GSE147507-CALU3		0.73	0.82	0.84	0.35	0.55	0.55	0.65	0.6	0.18	0.46	0.38	0.21	0.49	0.63	0.66
GSE147507-A549			0.38	0.32	0.08	0.07	0.17	0.2	0.1	-0.2	0.03	0.11	-0.2	-0.1	0.2	0.16
GSE148729-CALU3-S1				0.93	0.37	0.61	0.66	0.72	0.7	0.33	0.51	0.42	0.32	0.56	0.74	0.74
GSE148729-CALU3-S2					0.4	0.67	0.58	0.7	0.71	0.3	0.52	0.35	0.34	0.58	0.72	0.75
GSE17400-24h						0.67	0.4	0.39	0.34	0.22	0.23	0.22	0.29	0.32	0.44	0.34
GSE17400-28h							0.5	0.62	0.65	0.2	0.49	0.29	0.31	0.55	0.64	0.62
GSE37827-24-30h								0.86	0.78	0.44	0.58	0.64	0.14	0.6	0.83	0.73
GSE37827-36-54h									0.93	0.34	0.59	0.46	0.29	0.66	0.84	0.86
GSE37827-60-72h										0.3	0.6	0.39	0.24	0.66	0.76	0.86
GSE47963-24-48h(2)											0.79	0.72	0.89	0.83	0.47	0.27
GSE47963-60-96h(2)												0.8	0.69	0.92	0.67	0.64
GSE47963-24-48h(3)													0.73	0.76	0.52	0.4
GSE47963-24-48h(4)														0.84	0.33	0.18
GSE47963-60-96h(4)															0.74	0.69
GSE48142-24-36h																0.86



Table 1. The SARS associated profiles tend to correlate across multiple independent sample sets. The first five profiles are based on two independent RNAseq data series for SARS treatment of three cell lines (human bronchial epithelial cell line (CALU3), human alveolar basal epithelial cells (A549), human bronchial/tracheal epithelial cells (NHBE)). The last twelve profiles are generated from four microarray series, with GSE17400, GSE37827 and GSE48142 defined on a CALU3 cell line and GSE47963 on a human airway epithelium cell line (HAE). Profiles were defined based on grouping treatment and control samples corresponding to at least 24 hours of virus treatment. The colour scale indicates the correlation Z score and the values are the Pearson correlation coefficients across the genes regulated by the given profiles.

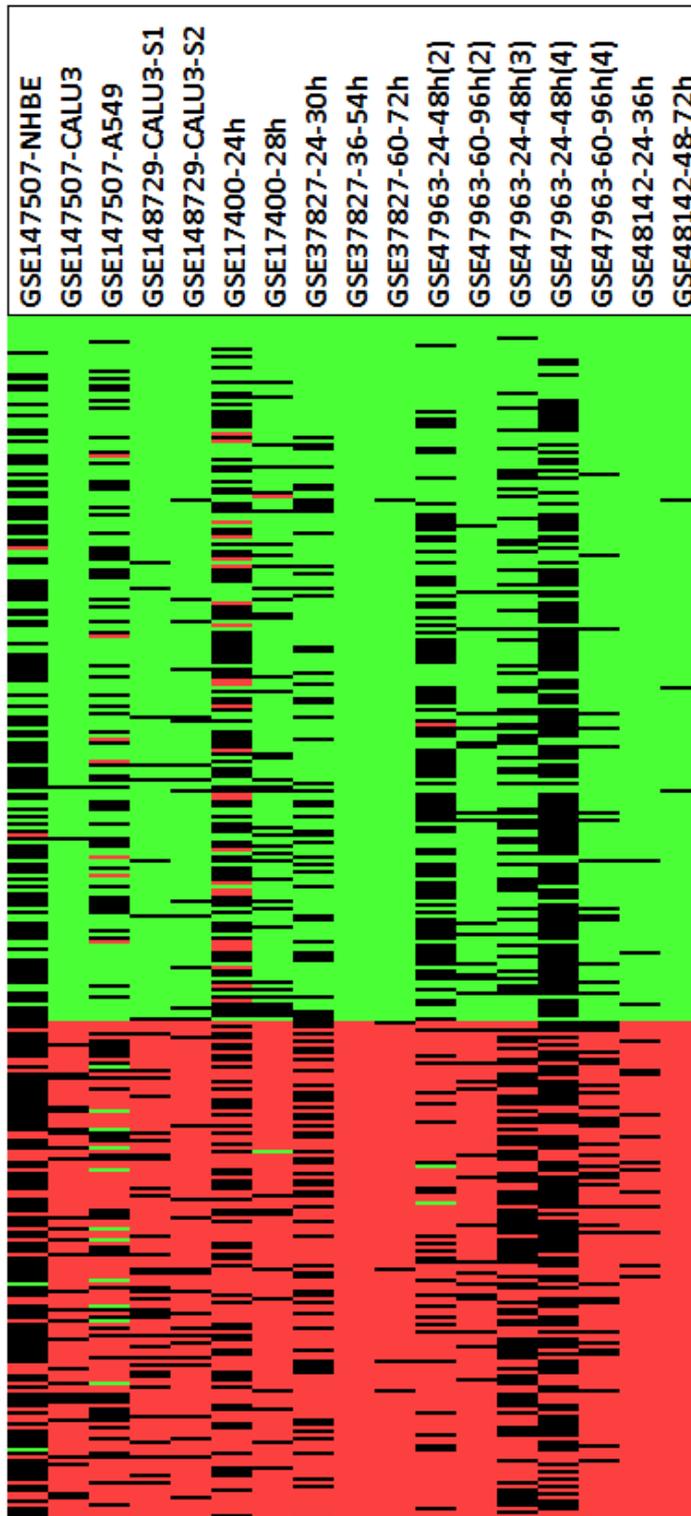


Table 2. The up and down regulation status of each gene in the SARS signature profile across the 17 SARS profiles. The signature profile is defined by there being at least 2/3 agreement across the profiles. The up-regulated set is highly enriched by viral response pathways. The down regulated set tends to be enriched by cell cycle pathways, see Table 3.

UP-GENES

PATHWAY	PROB
INTERFERON SIGNALING	6.72E-27
INTERFERON ALPHA BETA SIGNALING	1.27E-26
CYTOKINE SIGNALING IN IMMUNE SYSTEM	4.61E-26
IMMUNE SYSTEM	4.51E-22
INTERFERON GAMMA SIGNALING	3.53E-14
RIG I MDA5 MEDIATED INDUCTION OF IFN ALPHA BETA PATHWAYS	5.48E-13
INNATE IMMUNE SYSTEM	7.14E-11
NEGATIVE REGULATORS OF RIG I MDA5 SIGNALING	6.36E-10
BMAL1 CLOCK NPAS2 ACTIVATES CIRCADIAN EXPRESSION	1.91E-06
TRAF6 MEDIATED NFKB ACTIVATION	2.20E-06
ANTIVIRAL MECHANISM BY IFN STIMULATED GENES	5.41E-06
TRIF MEDIATED TLR3 SIGNALING	1.08E-05
TRAF6 MEDIATED IRF7 ACTIVATION	1.43E-05
CIRCADIAN CLOCK	1.73E-05
ACTIVATED TLR4 SIGNALLING	5.00E-05
NFKB AND MAP KINASES ACTIVATION MEDIATED BY TLR4 SIGNALING REPERTOIRE	9.54E-05
SIGNALING BY ILS	1.32E-04
TOLL RECEPTOR CASCADES	2.19E-04
TRAF6 MEDIATED INDUCTION OF NFKB AND MAP KINASES UPON TLR7 8 OR 9 ACTIVATION	1.18E-03
MYD88 MAL CASCADE INITIATED ON PLASMA MEMBRANE	1.66E-03
ADAPTIVE IMMUNE SYSTEM	3.20E-03
GENERIC TRANSCRIPTION PATHWAY	1.10E-02
CLASS I MHC MEDIATED ANTIGEN PROCESSING PRESENTATION	4.08E-02

DOWN-GENES

PATHWAY	PROB
LOSS OF NLP FROM MITOTIC CENTROSOMES	6.63E-05
RECRUITMENT OF MITOTIC CENTROSOME PROTEINS AND COMPLEXES	1.16E-04
MITOTIC G2 G2 M PHASES	3.13E-04
CELL CYCLE MITOTIC	5.43E-04
CELL CYCLE	3.21E-03
NGF SIGNALLING VIA TRKA FROM THE PLASMA MEMBRANE	3.35E-03
SIGNALLING BY NGF	5.22E-03
ADAPTIVE IMMUNE SYSTEM	1.50E-02

Table 3. The SARS signature profile pathway analysis. The signature has relatively more up-regulated genes (192 v 134 down-regulated) and this set is highly enriched with viral response pathways. The down-regulated set is enriched by cell cycle pathways.

UP-GENES

GENE ONTOLOGY CLASS	PROB
defence response to virus	3.70E-32
type I interferon signalling pathway	1.89E-26
response to virus	9.38E-24
positive regulation of transcription from RNA polymerase II promoter	4.61E-21
negative regulation of viral genome replication	5.72E-17
interferon-gamma-mediated signalling pathway	5.85E-12
negative regulation of type I interferon production	1.58E-11
RNA polymerase II core promoter proximal region sequence-specific DNA binding	1.84E-11
transcription factor activity, sequence-specific DNA binding	4.82E-11
inflammatory response	5.15E-11
transcription factor activity, RNA polymerase II core promoter proximal region sequence-specific binding	8.44E-11
innate immune response	2.44E-10
transcription from RNA polymerase II promoter	1.16E-09
apoptotic process	2.06E-09
regulation of transcription from RNA polymerase II promoter	2.52E-09
response to lipopolysaccharide	5.87E-09
double-stranded RNA binding	1.25E-08
positive regulation of NF-kappaB transcription factor activity	1.56E-08
positive regulation of transcription, DNA-templated	1.64E-08
negative regulation of transcription from RNA polymerase II promoter	3.50E-08
immune response	7.12E-08
response to cAMP	9.11E-08
nucleoplasm	9.25E-08
positive regulation of interferon-alpha production	9.25E-08

DOWN-GENES

GENE ONTOLOGY CLASS	PROB
ciliary transition zone	2.31E-07
extracellular exosome	1.37E-06
centrosome	2.96E-06
vesicle	1.76E-05
ciliary basal body	3.67E-05
G2/M transition of mitotic cell cycle	6.65E-05
axoneme	1.90E-04
membrane	5.42E-04
endoplasmic reticulum	6.37E-04
ubiquitin protein ligase binding	7.79E-04
mitochondrion	8.71E-04
ER to Golgi vesicle-mediated transport	9.57E-04

Table 4. The SARS signature profile gene ontology enrichment analysis. The ontologies enriched in the SARS signature profile largely correspond to the pathways shown in Table 3.

Virus treatment related SPIED results for SARS signature profile query

RANK	correl	ln(prob)	DESCRIPTION	SERIES	SAMPLES
1	1	2.19E-194	part of SARS signature query	GSE37827	GSM928604
2	0.96	5.26E-152	part of SARS signature query	GSE17400	GSM432393
3	1	1.40E-131	part of SARS signature query	GSE47960	GSM1163345
4	1	3.09E-123	part of SARS signature query	GSE47962	GSM1163583
5	1	6.64E-123	part of SARS signature query	GSE48142	GSM1169654
6	0.99	2.48E-111	part of SARS signature query	GSE47961	GSM1163525
7	0.96	3.36E-109	CALU-3 treated with H5N1	GSE43204	GSM1058275
8	0.93	3.38E-103	CALU-3 treated with H1N1	GSE40844	GSM1003210
9	0.92	3.39E-95	CALU-3 treated with H1N1	GSE37571	GSM921929
15	0.97	1.54E-77	HTBE cells traeted with H1N1	GSE39200	GSM957993
16	0.92	3.87E-81	HBE cells infected with H1N1	GSE19392	GSM528676
19	0.95	1.05E-76	A549 cells infected with HPIV1	GSE12664	GSM317901
20	1	1.39E-71	lung tissue from mice infected with SARS MA15	GSE49262	GSM1196080
21	0.92	1.74E-77	lung tissue from mice infected with SARS MA15	GSE50000	GSM1211771
22	0.88	1.26E-80	HIV infected human monocytes	GSE30536	GSM757658
23	0.94	1.42E-75	murine norovirus infected macrophages	GSE50093	GSM1214141
25	0.92	8.46E-77	human airway epithelia traeted with Rhinovirus	GSE51392	GSM1244644
28	0.9	9.56E-75	MRC5 cells treated with sendai virus	GSE38016	GSM932393
40	0.88	1.45E-71	lung tissue from H1N1 infection of mouse	GSE37569	GSM921844
42	0.95	1.26E-65	lung tissue from SARS nsp16 infected mice	GSE49263	GSM1196129

Table 5. The top ranked virus treatment related SPIED profiles against the SARS signature profile. The top ranked hits are from the series populating the query itself. In the next 42 hits 14 correspond to virus treatments. This establishes the signature as a robust quantitative descriptor of the gene expression changes upon viral infection. The signature is a broad-brush descriptor sharing gene sets with other virus types.

CMAP

RANK	COMPOUND	STATISTIC
1	emetine	(129,11,9.50E-25)
2	thioridazine	(137,16,8.49E-23)
3	cephaeline	(142,19,2.97E-22)
4	terfenadine	(99,7,2.91E-20)
5	puromycin	(114,14,2.37E-19)
6	prenylamine	(90,9,1.20E-17)
7	suloctidil	(90,7,1.52E-17)
8	dioquinol	(78,5,4.73E-17)
9	astemizole	(91,11,2.86E-16)
10	bepridil	(83,10,1.52E-15)
11	prochlorperazine	(95,16,2.22E-14)
12	mefloquine	(83,12,1.33E-13)
13	perhexiline	(66,7,8.05E-13)
14	doperastine	(83,15,1.92E-12)
15	felodipine	(72,9,2.82E-12)
16	beta-escin	(70,10,4.35E-12)
17	oxybutynin	(56,5,4.95E-12)
18	alexidine	(57,5,6.62E-12)
19	pyrvinium	(82,14,6.97E-12)
20	rottlerin	(93,22,1.48E-11)
21	anisomycin	(119,25,1.77E-11)
22	ebesen	(72,14,2.67E-11)
23	lycorine	(80,13,4.30E-11)
24	mebendazole	(94,24,6.40E-11)
25	niclosamide	(71,13,1.07E-10)
26	trifluoperazine	(95,24,1.15E-10)
27	clotrimazole	(57,7,1.20E-10)
28	digoxin	(99,26,1.23E-10)
29	azacyclonol	(65,4,2.66E-10)
30	MG-262	(92,26,3.76E-10)
31	pimozide	(61,10,4.99E-10)
32	piperone	(56,7,1.05E-9)
33	8-azaguanine	(85,26,1.09E-9)
34	calmidazolium	(82,14,1.72E-9)
35	nocodazole	(60,11,1.80E-9)
36	thapsigargin	(90,24,5.22E-9)
37	dorgiline	(52,8,5.27E-9)
38	gossypol	(65,14,6.50E-9)
39	ouabain	(88,26,8.03E-9)
40	15-delta_prostagla	(90,28,9.27E-9)
41	fendiline	(62,13,1.07E-8)
42	hexestrol	(74,19,1.83E-8)
43	cicloheximide	(83,19,1.83E-8)
44	fluphenazine	(70,16,3.31E-8)
45	lanatoside C	(105,35,4.14E-8)

LINCS

RANK	COMPOUND	STATISTIC	RANK	STATISTIC
1	homoharringtonine	(142,13,4.79E-24)	-	-
2	emetine	(134,6,2.63E-23)	1	(129,11,9.50E-25)
3	QL-XII-47	(150,19,2.88E-23)	-	-
4	cucurbitacin-i	(138,18,1.59E-22)	-	-
5	calmidazolium	(81,0,1.90E-21)	34	(82,14,1.72E-9)
6	narciclasine	(125,10,7.93E-21)	-	-
7	tricitribine	(108,8,1.54E-20)	-	-
8	ouabain	(141,18,9.76E-20)	39	(88,26,8.03E-9)
9	IKK-2-inhibitor-V	(146,27,1.29E-19)	-	-
10	niclosamide	(147,25,1.86E-19)	25	(71,13,1.07E-10)
11	methylene-blue	(90,6,4.48E-19)	-	-
12	AKT-inhibitor-IV	(94,6,6.27E-19)	-	-
13	bruceantin	(105,9,1.97E-18)	-	-
14	menadione	(117,14,2.49E-18)	107	(47,14,1.84E-4)
15	LDN-193189	(147,25,3.15E-18)	-	-
16	CHEMBL-399379	(83,2,4.12E-18)	-	-
17	manumycin-a	(139,27,9.26E-18)	-	-
18	BGT-226	(139,28,2.83E-17)	-	-
19	FCCP	(81,4,7.79E-17)	-	-
20	anisomycin	(110,13,2.57E-16)	21	(119,25,1.77E-11)
21	NSC-632839	(116,20,5.84E-16)	-	-
22	TAS-103	(115,10,8.70E-16)	-	-
23	digoxin	(126,21,9.88E-16)	28	(99,26,1.23E-10)
24	wortmannin	(130,30,3.31E-15)	-	-
25	phorbol-myristate-	(129,26,4.66E-15)	-	-
26	tyrphostin-A9	(75,6,4.77E-15)	-	-
27	azacitidine	(112,21,1.37E-14)	-	-
28	delanzomib	(105,19,1.45E-14)	-	-
29	celastrol	(135,33,1.88E-14)	-	-
30	avicin-g	(94,15,2.22E-14)	-	-
31	CGK-733	(86,11,3.10E-14)	-	-
32	oprozomib	(103,14,3.23E-14)	-	-
33	bardoxolone-meth	(106,24,1.42E-13)	-	-
34	suloctidil	(64,3,1.67E-13)	7	(90,7,1.52E-17)
35	cephaeline	(71,2,1.73E-13)	3	(142,19,2.97E-22)
36	BIX-01294	(108,23,3.27E-13)	-	-
37	F-1566-0341	(93,13,4.62E-13)	-	-
38	WYE-125132	(118,30,5.93E-13)	-	-
39	withaferin A	(133,36,6.92E-13)	-	-
40	MG-132	(116,27,7.33E-13)	-	-
41	rottlerin	(105,21,7.51E-13)	20	(93,22,1.48E-11)
42	JW-7-24-1	(104,23,7.56E-13)	-	-
43	buparlisib	(129,39,7.94E-13)	-	-
44	CD-1530	(58,2,8.18E-13)	-	-
45	bortezomib	(114,24,8.35E-13)	-	-

CMAP

Table 6. The results for the SISP query against the CMAP and LINCS data sets. Correlations are scored according to the exact Fisher test for the direction of regulation of shared genes. Candidate therapeutics are hypothesised to recapitulate or enhance the viral response mechanism of the host cell. The top 45 positively correlating compounds are shown for the CMAP and LINCS datasets. There is a degree of agreement between the two compound sets as can be seen from the CMAP column at the right of the LINCS results. The compounds highlighted in yellow have reported anti-viral activities:

emetine and the closely related cephaeline (Andersen et al., 2019; Choy et al., 2020; Yang et al., 2018), bepridil (DeWald et al., 2018), prochlorperazine (Simanjuntak et al., 2015), mefloquine (Brickelmaier et al., 2009), beta-escin (Michelini et al., 2018), pyrvinium (Shen et al., 2019), rottlerin (Lama et al., 2019), anisomycin (Quintana et al., 2020), ebselen (Mukherjee et al., 2014), lycorine (Wang et al., 2014), niclosamide (Jurgeit et al., 2012; Xu et al., 2020), trifluoperazine (Pietschmann, 2017), clotrimazole (Torriani et al., 2019), digoxin (Wong et al., 2013), calmidazolium (Jablonowski, 2005), gossypol (Radloff et al., 1986), Ouabain (Jeon et al., 2020), cycloheximide (Ramabhadran and Thach, 1980), fluphenazine (Chamoun-Emanuelli et al., 2013), lanatoside (Cheung et al., 2014). Within the top 45 LINCS list are found the following additional hits: Homoharringtonine or Omacetaxine (Andersen et al., 2019; Dong et al., 2018), QL-XII-47 a potent Burton's tyrosine kinase inhibitor (Liang et al., 2017), Narciclasine (Gabrielsen et al., 1992), Triciribine (Porcari et al., 2003), bruceantin, AKT-inhibitor-IV (Sun et al., 2011), phorbol myristate acetate (Premecz et al., 1985), tyrphostin A9 (Kumar et al., 2011), azacytidine (Diamantopoulos et al., 2012), celastrol (Yu et al., 2017), bardoxolone methyl (Nio et al., 2019), withaferin A (Grover et al., 2011), MG-132 a proteasome inhibitor (Schwartz et al., 1998).