Published in final edited form as:

J Microbiol Methods. 2016 July ; 126: 30–34. doi:10.1016/j.mimet.2016.04.013.

Development of a web-based tool for automated processing and cataloging of a unique combinatorial drug screen

Alex G. Dalecki¹ and Frank Wolschendorf¹

¹Department of Medicine, Division of Infectious Diseases, University of Alabama at Birmingham, 845 19th St S, Birmingham, AL 35294

Abstract

Facing totally resistant bacteria, traditional drug discovery efforts have proven to be of limited use in replenishing our depleted arsenal of therapeutic antibiotics. Recently, the natural anti-bacterial properties of metal ions in synergy with metal-coordinating ligands have shown potential for generating new molecule candidates with potential therapeutic downstream applications. We recently developed a novel combinatorial screening approach to identify compounds with copper-dependent anti-bacterial properties. Through a parallel screening technique, the assay distinguishes between copper-dependent and independent activities against *Mycobacterium tuberculosis* with hits being defined as compounds with copper-dependent activities. These activities must then be linked to a compound master list to process and analyze the data and to identify the hit molecules, a labor intensive and mistake-prone analysis. Here, we describe a software program built to automate this analysis in order to streamline our workflow significantly. We conducted a small, 1440 compound screen against *Mycobacterium tuberculosis* and used it as an example framework to build and optimize the software. Though specifically adapted to our own needs, it can be readily expanded for any small- to medium-throughput screening effort, parallel or conventional. Further, by virtue of the underlying Linux server, it can be easily adapted for chemoinformatic analysis of screens through packages such as OpenBabel. Overall, this setup represents an easy-to-use solution for streamlining processing and analysis of biological screening data, as well as offering a scaffold for ready functionality expansion.

Keywords

Combinatorial screen; High throughput screen; Tuberculosis; Copper; Chemoinformatics

1 INTRODUCTION

Antibiotic resistance in pathogenic bacteria has become a significant clinical problem, threatening to not only increase direct mortality from infections but also hamper our ability

CORRESPONDING AUTHOR: Frank Wolschendorf (fwolsche@uab.edu).
(dalecki@uab.edu)

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

to perform crucial procedures reliant on treating infections, such as organ transplants and immunosuppressive cancer therapy (1-3). Compounding this, many pharmaceutical companies have divested themselves of drug discovery programs in response to the extreme cost of development and the uncertain return on investment, given the aforementioned propensity to lose effectiveness due to bacterial resistance (4-6). New screening mechanisms are sorely needed to uncover much needed antibiotics.

Recently, we described a novel approach to drug screening in which we specifically identify compounds acting in concert with copper ions, more closely mimicking *in vivo* conditions, as well as offering an ability to more deeply screen existing compound libraries, defraying the cost of screening studies when compared to the alternative option of synthesizing additional compounds (7-10). However, this screen, which directly compares the activities of compounds in a copper-activating medium (Cu+) against activity in a copper-deficient medium (Cu-), requires two assay plates run in parallel, and time consuming analysis to identify promising hit compounds (Figure 1). While a traditional drug screen uses a Boolean analysis of growth or activity values to determine hits, we required a non-trivial algorithm to impartially compare the given values and categorize them as copper-dependent hits, copper-independent hits, or not hits. Manual analysis with spreadsheet software, though possible, was laborious and prone to human error or bias. Thus, we sought to build an automated system on a Linux server, accessible through a web interface. This system would solicit user-generated files, parse them according to customizable parameters, manipulate and store the contained data within a MySQL database, generate reports of calculated hits, and link those reports to compound identities within a master library.

Here, we introduce that system, its structure, and an example analysis of a small, 1440 compound subset screened against *Mycobacterium tuberculosis*. Though specifically designed to tackle our unique parallel and combinatorial needs, this type of database solution can be adapted to any combinatorial screening strategy, as well as more traditional, single-plate analyses. The source code has been made available online (<https://github.com/adalecki/drugscreen>), along with a small script that automates the download and setup of a virtual machine with a demonstration dataset.

2 MATERIALS AND METHODS

2.1 Software

This software package was built inside an Ubuntu 14.04 LTS virtual operating system utilizing Oracle VM Virtualbox 4.3.18. We incorporated nginx 1.7.1 as a web server platform, MariaDB 10.0 as a database, and PHP-FPM 5.5.9 as an interface with PHP, the primary programming language used. Bootstrap v3.2.0 was used as a developmental framework. The program is accessible through a standard HTML web browser with server level password protection. Drug screens were conducted in 96 well plates, according to published protocols (7-9), and read using a Cytation 3 plate reader (BioTek) and included Gen5 2.05 or 2.06 software.

2.2 Data Preparation

Our compound library is organized by plate number and well coordinates: each stock 96 well plate has 80 wells with a unique compound (with the remaining 16 wells used as assay controls) that are transferred to their corresponding well coordinates on the assay plates. As such, knowledge of the original plate number is crucial to linking hits back to their compound ID, and the easiest way to pass that information to the program is through the file name. After reading all assay plates, the user builds CSV files within the Gen5 export feature, using well coordinates and raw read values as the two fields, and the plate number as the first integer within the filename (additional information, such as presence or absence of copper, may be included for human legibility, but is ignored by the program). Each individual plate is exported as a separate CSV file, uploaded in the next step.

2.3 User Interface

The program operates with two fundamental parts: The HTML-based web interface, used for soliciting information and displaying results, and the MySQL database on a Linux virtual server, used for processing, storing, and sorting data (Figure 2). From the main page, the user visits the “Upload File” page, where the program solicits identifying information (Figure 3). Elements on this page include: Separate upload buttons for substance positive (e.g. +Cu) and substance negative (e.g. -Cu), allowing simultaneous entrance of both types of plates; an organism selection, to put the plates in the correct table; a substance selection, such as copper (or any other agent for compound selection), to put the plates in the correct table; an expected blank value, in order to automatically disregard any blank that may have been contaminated; a plate layout, allowing users to alter their plates depending on actual location of controls; and finally a comment feature, by which a user can record notes about plates, such as specific assay conditions or known problems.

2.4 Program Structure

The database structure is reliant on three separate tables for each unique screen: a plate table, in which presence of substance positive and substance negative plates are tracked (upload of only one of the two required plates would result in calculation errors), as well as plate-wide values and notes for ease of later calculations; a control table, in which the control values (positives, negatives, and blanks) are tracked for each plate; and a sample table, containing the actual compound data. Every combination of organism and substance dynamically generates unique tables within the database upon first file entry, in the style of [organism]_[substance]_[table] (e.g., “Mycobacterium_tuberculosis_copper_samples.”). The multiple table structure allows easy tracking of relevant data in a variety of applications, such as general hit reporting, Z' Factor calculation, chemical substructure searching, and so on. This also facilitates inclusion of multiple combinations of screens all within one single database.

Upon entrance of the desired plates, users may generate a hit report, or a list of all compounds present in the database satisfying desired criteria (Figure 4). Reports are generated individually for a specified pair of organism and substance tested, allowing this algorithm and storage scheme to simultaneously track and compare multiple and varied drug screens without confusing results. Each tested compound is passed through an algorithm that

corrects the value with the plate-specific blank, normalizes to respective positive controls, and compares the computed values for both plates before assigning a hit category to each compound. Compounds are returned as primary (strong) hits, secondary or tertiary (weak) hits, independent (with regard to presence of substance) hits, inverse (in which the tested substance neutralized the effect of the compound) hits, or not a hit. Desired categories are subsequently included in the prepared report. Additionally, after the report is generated, users may download a CSV file of the report for permanent reference.

Integration with other services, such as the open source OpenBabel chemical toolbox (11), is accomplished through the Linux backend. By passing search criteria from web input to direct server commands, and then piping resulting output to the web interface, an end user can avoid command-line programming queries. As each compound's chemical Simplified Molecular Input Line Entry Specification (SMILES) is included in the database, OpenBabel can query those chemical structures for desired substructures, only returning compounds meeting specified criteria. The same toolbox is also used for dynamically displaying structures on screen (Figure 5B), removing the need for cross-referencing a supplier's website for quick, visual identification. Different capabilities, such as further utilization of OpenBabel (e.g. calculation of molecular properties, clustering analysis), inclusion of larger datasets like the ChEMBL database (12), or integration of statistical packages like R, can be readily built into this same web/server crosstalk.

3 RESULTS AND DISCUSSION

The example assay used here is characterized by a strong Z' factor of 0.89 (7), a standard measure of assay robustness and quality assessment. As anything above a Z' factor of 0.5 is considered excellent (13), the *M. tuberculosis* screen readily served as an example dataset for developing this software. Instead of defining hits as those with inhibition more than three standard deviations from the sample mean, we opted for a more stringent definition requiring 90% inhibition to be considered a primary hit, as statistical significance is often not indicative of biological significance (14). These primary hits feature 90% inhibition or greater in the copper-positive plate, while the copper-negative plate had at least thirty percentage points more growth. Secondary and tertiary hits have weaker gradations of inhibition.

On our pilot screen of 1440 compounds against *M. tuberculosis*, we found 16 total independent hits, that is, compounds that inhibited *M. tuberculosis* growth to less than 30% in both parallel plates, regardless of copper content. These independent hits represent what a standard, non-combinatorial drug screen would uncover, a mere 1.1% of the tested library. Our novel combinatorial screen revealed 49 primary hits, in which only the copper-positive plate experienced inhibition of less than 10% growth, while the corresponding well on the copper-negative plate had at least 30 percentage points more growth (Figure 4). This represents a more than threefold increase in the number of discovered hits when compared to a traditional drug screen. Further, there were 27 secondary and 13 tertiary hits, defined as inhibition in the copper-positive plate between 10%-20% and between 20%-30%, respectively. Though these compounds would not necessarily be prioritized for further

development, they can provide valuable information when identifying a lead series from common molecular substructures.

The software package here has proven far superior to other analysis methods, such as Excel calculation templates. Though a spreadsheet analysis could accomplish many of the same goals, human manipulation invites user error and bias. Additionally, use of a software algorithm dramatically decreases time spent analyzing data; while manual analysis previously required approximately one hour for every five pairs of plates, the server reduces computational time to under five minutes, most of that spent generating export files from within the plate reader software.

While the incarnation of the server presented here is fairly basic in its scope, the database allows easy future expansion of any imaginable analysis. Automated Z' factor calculations can be implemented to track plate quality upon upload; specific compound identifiers can be queried to view results over multiple screens, potentially discovering broad spectrum results; or the HTML front end can interface with chemoinformatics programs on the server, such as OpenBabel (11), allowing interactive molecular fingerprint searching. Only a basic working knowledge of PHP, HTML, and MySQL is required for implementation, putting such an analysis program well within the reach of most labs.

4 CONCLUSION

In response to the looming antibiotic crisis, many academic and independent labs have begun to develop low- and medium-throughput screening assays, often with the goal of eventually optimizing and passing them to industrial level high-throughput screening partners. Unfortunately, while professional screening centers can afford complex data analysis, small- to mid-level research labs often have no reasonable access to complex software packages and solutions. The software described here represents a free and reasonable alternative to proprietary options. Centralized and customized access facilitates comparison across separate screening campaigns, as well as ready integration with other software packages.

While we have used the program specifically to examine copper-dependent activities, the basic combinatorial algorithm is expandable to any treatments with potentially synergistic effects. The software package's design readily permits many separate screens to be tracked by the database simultaneously, categorized both by organism and combinatorial treatment. These combinatorial treatments could study treatments such as media-dependent compound inhibitors (*e.g.*, copper-dependent inhibitors), antimicrobial interactions (*e.g.*, β -lactam inhibitors, efficacy of antibiotic cocktails), transposon library analysis (*e.g.*, strain survival in altered growth conditions), or disease-specific effects (*e.g.*, cancer-specific inhibitors).

The software has been made available online (<https://github.com/adalecki/drugscreen>), along with a sample database, and five pairs of screening plates for demonstration purposes. Through use of the included bash setup file (<https://github.com/adalecki/drugscreen/blob/master/setup.sh>), a demo virtual machine can quickly be configured for an Ubuntu/Nginx/MariaDB/PHP LAMP stack, with the sample database included.

ACKNOWLEDGEMENT

The authors would like to thank Dr. Seung Park and Dr. Elliot Lefkowitz for programmatic advice during development of this software package and manuscript. This research was supported by NIH grant R01-AI104952, awarded to FW, and made possible in part by the UAB Center for Clinical and Translational Science Grant Number UL1TR001417 from the National Center for Advancing Translational Sciences (NCATS) of the National Institutes of Health (NIH).

LITERATURE CITED

1. Mehta KC, Dargad RR, Borade DM, Swami OC. Burden of Antibiotic Resistance in Common Infectious Diseases: Role of Antibiotic Combination Therapy. *Journal of Clinical and Diagnostic Research : JCDR*. 2014; 8:ME05–ME08. [PubMed: 25121020]
2. Livermore DM. Has the era of untreatable infections arrived? *J Antimicrob Chemother* 64 Suppl. 2009; 1:i29–36.
3. Livermore DM. Discovery research: the scientific challenge of finding new antibiotics. *J Antimicrob Chemother*. 2011; 66:1941–1944. [PubMed: 21700626]
4. Shlaes DM, Sahm D, Opiela C, Spellberg B. The FDA Reboot of Antibiotic Development. *Antimicrobial Agents and Chemotherapy*. 2013; 57:4605–4607. [PubMed: 23896479]
5. America IDSo. Combating Antimicrobial Resistance: Policy Recommendations to Save Lives. *Clinical Infectious Diseases*. 2011; 52:S397–S428. [PubMed: 21474585]
6. Pearson H. Antibiotic faces uncertain future. *Nature*. 2006; 441:260–261. [PubMed: 16710378]
7. Speer A, Shrestha TB, Bossmann SH, Basaraba RJ, Harber GJ, Michalek SM, Niederweis M, Kutsch O, Wolschendorf F. Copper-boosting compounds: a novel concept for antimycobacterial drug discovery. *Antimicrob Agents Chemother*. 2013; 57:1089–1091. [PubMed: 23254420]
8. Dalecki AG, Haeili M, Shah S, Speer A, Niederweis M, Kutsch O, Wolschendorf F. Disulfiram and copper ions kill *Mycobacterium tuberculosis* in a synergistic manner. *Antimicrobial Agents and Chemotherapy*. 2015 doi:10.1128/aac.00692-15.
9. Haeili M, Moore C, Davis CJC, Cochran JB, Shah S, Shrestha TB, Zhang Y, Bossmann SH, Benjamin WH, Kutsch O, Wolschendorf F. Copper complexation screen reveals compounds with potent antibiotic properties against methicillin-resistant *Staphylococcus aureus*. *Antimicrobial Agents and Chemotherapy*. 2014; 58:3727–3736. [PubMed: 24752262]
10. Dalecki AG, Malalasekera AP, Schaaf K, Kutsch O, Bossmann SH, Wolschendorf F. Combinatorial phenotypic screen uncovers unrecognized family of extended thiourea inhibitors with copper-dependent anti-staphylococcal activity. *Metallomics*. 2016 doi:10.1039/C6MT00003G.
11. O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR. Open Babel: An open chemical toolbox. *Journal of Cheminformatics*. 2011; 3:33–33. [PubMed: 21982300]
12. Bento AP, Gaulton A, Hersey A, Bellis LJ, Chambers J, Davies M, Krüger FA, Light Y, Mak L, McGlinchey S, Nowotka M, Papadatos G, Santos R, Overington JP. The ChEMBL bioactivity database: an update. *Nucleic acids research*. 2014; 42:D1083–1090. [PubMed: 24214965]
13. Zhang J-H, Chung TDY, Oldenburg KR. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *Journal of Biomolecular Screening*. 1999; 4:67–73. [PubMed: 10838414]
14. Wasserstein RL, Lazar NA. The ASA's statement on p-values: context, process, and purpose. *The American Statistician*. 2016 doi:10.1080/00031305.2016.1154108:00-00.

HIGHLIGHTS

- Small- to mid-size academic labs often lack software for screening analysis
- We detail an in-house server solution for analyzing combinatorial drug screens
- The software is easily integrated with statistic or chemoinformatic packages
- Source code and an automated script for setup of a virtual machine available online

Figure 1. General assay scheme and results

(A) Our combinatorial drug screen assay requires analyzing two plates with identical compounds in parallel, one with copper and one without (though any substance can be substituted). (B) After incubating for the requisite length of time, resazurin dye is added as a viability indicator: live cells will metabolize resazurin, which appears blue, to resorufin, a pink fluorescent dye. The fluorescence readings are corrected to blank values and normalized to positive growth wells. When comparing wells from the two plates, one of four possible outcomes results: First, growth may be uninhibited in both wells; this represents the majority of compounds, inactive whether or not copper is present. Second, the compound may block growth as a general inhibitor, whether or not copper is present. Third, copper may, in rare cases, neutralize the compound, leaving the copper-positive plate uninhibited and the copper-negative plate inhibited. Finally, we may find “hit” compounds, which do not inhibit growth in the absence of copper, but exhibit strong inhibitory effects in copper-positive conditions.

Figure 2. General program scheme

The program contains two main components: a Web-accessible user interface (UI) as the front end, and a Linux server hosting a MySQL database as the backend. Users upload prepared data files, along with identifying information, to the server, which extracts and processes the data. All data points are stored in the database until needed. After uploading all needed files, the user can query the database with specific instructions, *i.e.* the specific organism and substance they are interested in. The server then pulls all relevant data points from the stored database and passes them through an algorithm to determine which hit type (primary, secondary, tertiary, independent, inverse, or no hit) each linked pair is. A report is procedurally generated as the algorithm analyzes the database, and upon completion that report is displayed to the user through the Web UI interface. The user also has the option of downloading the report in a CSV format directly from the report screen.

Wolschendorf Lab Database Home Report Upload Files Add Screening Categories

Plate Entry

Substance Positive Plates Substance Negative Plates

Organism?

Substance?

Expected Blank? 1900

	1	2	3	4	5	6	7	8	9	10	11	12
A	Blank ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Positive ▾
B	Blank ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Positive ▾
C	Blank ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Positive ▾
D	Blank ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Positive ▾
E	Blank ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Negative ▾
F	Blank ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Negative ▾
G	Blank ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Negative ▾
H	Blank ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Sample ▾	Negative ▾

Please enter any desired notes about this set of plates below:

Submit

Figure 3. File upload screen with customization options

The file upload screen allows the user to specify all important information about the files being imported. In addition to organism and substance solicitation, users may specify expected blank values to automatically exclude unexpected blanks. Through use of a drop down menu representing each well, a user may specify the plate layout (designating blanks, samples, positive and negative controls, or excluded wells), though the form autopopulates with a default layout. Finally, a text entrance box allows the user to record any pertinent notes, such as mishaps with the assays or oddities observed about certain wells; those notes will appear in generated reports.

Wolschendorf Lab Database Home Report Upload Files Add Screening Categories

Mycobacterium_tuberculosis
Copper

Total compounds analysed: 400
12565, 12566, 12567, 12568, 12569

Entries without substance-positive plate: 0
Entries without substance-negative plate: 0

Compound ID	SMILES	Plate ID	Coordinates	Hit Type	Pos Ctrl (+)	Pos Ctrl (-)	Neg Ctrl (+)	Neg Ctrl (-)	Norm'd (+)	Norm'd (-)
7523850	<chem>O=C(NC1SCC(=O)Nc1ccc(cc1)OC)C1Cc1ccc(cc1)Br</chem>	12565	C2	Primary	35312.000	27422.500	1296.000	27690.500	1.887	126.325
7529799	<chem>C1(C(C(C(C1F)F)F)F)OCc1cc(C(=O)Nc2ccc(N(CO)CC)cc2)cc1</chem>	12565	D7	Primary	35312.000	27422.500	1296.000	27690.500	2.219	122.603
7530499	<chem>C12(C(=O)NC3=NC(=O)C(C1)C(C2)C(C3)C)C</chem>	12565	H8	Primary	35312.000	27422.500	1296.000	27690.500	3.053	104.773
7528020	<chem>C(=S)N1Cc2c(CO)C(CCC2)NC1CCCC1</chem>	12566	H8	Primary	33639.000	27724.250	1336.250	27110.750	3.605	35.948
7533746	<chem>N1(C(NC1SCC(=O)C1ccc(cc1)Br)C1Cc1ccc(cc1)CCCC2)C</chem>	12567	C10	Primary	30860.500	26195.250	1330.250	26677.000	9.634	129.225
7535569	<chem>O=C(NC1SCC(=O)Nc1ccc(C(F)F)cc1)C1Cc1ccc(cc1)Br</chem>	12568	F8	Primary	32324.000	24811.750	1552.000	23713.000	6.780	136.541
7544138	<chem>N1(C(NC1COc1c(C)cccc1)SCC(=O)Nc1c(F)cccc1)C</chem>	12569	B11	Primary	31294.250	26113.250	1540.750	26233.750	5.815	107.644
7545149	<chem>C1(NC2C([NH1])CCCC2)SCc1c(Cc1cc1)C)C</chem>	12569	D10	Primary	31294.250	26113.250	1540.750	26233.750	6.721	126.710
7540710	<chem>O=C(NC1COc1ccc(cc1)C(C)C)SCC(=O)Nc1ccc(C(=O)OC)cc1</chem>	12569	E7	Primary	31294.250	26113.250	1540.750	26233.750	4.930	114.558
7530060	<chem>C1(NC2C(s1)ccc(cc2)C)NC(=O)C1ccc(OCCCC)ccc1</chem>	12565	B8	Secondary	35312.000	27422.500	1296.000	27690.500	14.936	121.497
7527272	<chem>C(=O)N(Cc1ccc(CNC(=O)CSc2cccc2)ccc1)CSc1cccc1</chem>	12565	G10	Secondary	35312.000	27422.500	1296.000	27690.500	14.451	75.643
7527792	<chem>N1(C(NC2c1cccc2)COc1c(C)cccc1)Cc1cccc1</chem>	12566	D8	Secondary	33639.000	27724.250	1336.250	27110.750	18.120	121.098

Figure 4. Assay results displayed as HTML table

The user may generate a detailed report after all desired data has been entered into the database. The top portion of the report lists the plates analyzed as well as presence of incomplete entries (in which only the substance-positive or substance-negative plate was entered into the database); incomplete entries are excluded from calculations. Every hit is listed with its unique Compound ID, Plate ID (both as unique lab identifiers here), SMILES chemical identifier, and plate coordinates, as well as the hit type and the corrected and normalized values used in the calculation. The positive and negative controls for both plates are also reported, to allow easy verification of results; if a user detects one plate with an inordinate number of hits, the user can check to ensure all assay controls were successful and within expected ranges. Finally, the total numbers of each hit type are displayed at the bottom, as well as a button to generate a CSV file from the displayed data. Only the beginning and end of the report screen are displayed, with the middle truncated for space.

A

Structure Query

[Search Instructions](#)

Type of Search?

☐ Whole Library

☒ Screened Library

☐ Hits in Screen

Structural Motif?

Organism?

Substance?

B

Wolschendorf Lab Database Home Report Upload Files Add Screening Categories

Mycobacterium_tuberculosis
Copper
Motif: c12ccccc1cccc2

Compound ID	SMILES	Structure
7523862	<chem>C(=O)(NCc1ccc(F)cc1)c1cc(COc2cc3c(cc2)cccc3)ccc1</chem>
7525207	<chem>c1(c2c(cc(cc2)Br)ccc1OC)COc1ccc(cc1)[N+](=O)[O-]</chem>

Figure 5. Chemoinformatic substructure searching

By integrating with software installed on the server, such as the open source chemistry toolbox OpenBabel, we are able to conduct substructure searches in an easy-to-use web environment. **(A)** Users select the data set to query (whole library, screened subset, or only hits) and, if applicable, the target screening campaign. Users also input a motif query based on the SMILES and SMARTS standards; for instance, the shown query (c12ccccc1cccc2) selects all fused benzene rings in a subset. **(B)** All results are then displayed in a table, complete with structures for easy reference.