Open PHACTS

Deliverable 5.3.5

First release of Polypharmacology browser, GARField

The Open PHACTS Poly-pharmacology Browsers

Prepared by DTU
Approved by DTU, PSMAR, Janssen, UNIVIE, AZ

September 2013
Version 1.0

Project title: An open, integrated and sustainable chemistry, biology and pharmacology knowledge resource for drug discovery
Instrument: IMI JU
Contract no: 115191

Start date: 01 March 2011
Duration: 3 years

Nature of the Deliverable	
Report	
Prototype	**x**
Other	
Dissemination level	
Public dissemination level	**x**
For internal use only	

Definitions

- Partners of the Open PHACTS Consortium are referred to herein according to the following codes:

  Pfizer – Pfizer limited – **Coordinator**
  UNIVIE – Universität Wien – **Managing entity of IMI JU funding**
  DTU – Technical University of Denmark – DTU
  UHAM – University of Hamburg, Center for Bioinformatics
  BIT – BioSolveIT GmbH
  PSMAR – Consorci Mar Parc de Salut de Barcelona
  LUMC – Leiden University Medical Centre
  RSC – Royal Society of Chemistry
  VUA – Vrije Universiteit Amsterdam
  CNIO – Spanish National Cancer Research Centre
  UNIMAN – University of Manchester
  UM – University of Maastricht
  ACK – ACKnowledge
  USC – University of Santiago de Compostela
  UBO – Rheinische Friedrich-Wilhelms-Universität Bonn
  AZ – AstraZeneca
  GSK – GlaxoSmithKline
  Esteve – Laboratorios del Dr. Esteve, S.A.
  Novartis – Novartis
  ME – Merck Serono
  HLU – H. Lundbeck A/S
  E.Lilly – Eli Lilly
  NBIC – Stichting Netherlands Bioinformatics Centre
  SIB – Swiss Institute of Bioinformatics
  ConnDisc – Connected Discovery
  EBI – European Bioinformatics Institute
  Janssen – Janssen Pharmaceutica
  OGL – OpenLink Software

- **Grant Agreement**: The agreement signed between the beneficiaries and the IMI JU for the undertaking of the Open PHACTS project.
- **Project**: The sum of all activities carried out in the framework of the Grant Agreement.
- **Work plan**: Schedule of tasks, deliverables, efforts, dates and responsibilities corresponding to the work to be carried, out as specified in the Grant Agreement.
- **Consortium**: The Open PHACTS Consortium composed of the above-mentioned legal entities.
- **Project Agreement**: Agreement concluded amongst Open PHACTS participants for the implementation of the Grant Agreement. Such an agreement shall not affect the parties' obligations to the Community and/or to one another arising from the Grant Agreement.

1 Introduction

The most commonly used drug discovery paradigm, i.e. one drug for one target, is based on the underlying assumption of selective bioactivity, and has been the main driving force in medicinal chemistry optimization when developing small molecule therapeutics. Recent advances in chemical biology and systems biology have shown that most drugs interact with multiple targets and the pharmacological profile of a drug is not as reductionist as we believed. Therefore, the increasing amount of available data on ligand-target interactions brings on the need to develop new interactive tools for data integration and mining that can facilitate knowledge extraction.

Open PHACTS has recently created and released the first open access pharmacological space, applying semantic web technologies, with the aim to facilitate open innovation in drug discovery research. Among the pharmacological information integrated in the Open PHACTS discovery platform, bioactivities data for a large set of compounds is available. To explore, to analyze and to visualize such pharmacological data, interactive applications (eApps) have been developed and integrated with the Open PHACTS platform release. For example, ChemBioNavigator allows plotting the physicochemical properties of molecular groups against each other. With, PharmaTrek, the user can navigate within the pharmacological space content of Open PHACTS. More information about these eApps can be found at this link: http://www.openphacts.org/

With the release of an updated version of the Open PHACTS platform for the end of 2013, integration of new eApps is also planned. Among them, GARField (Graph Activitiy Relationship Visualization Field) is a tool aiming to predict the pharmacological profiling of a compound using several and diverse metrics (structure-based ligand and sequence-based target). GARField is an integrative and interactive web-based tool, that allow to visualize the bioactivities annotated to a compound as well as to provide potential new bioactivities not yet tested.

Further details about GARField features are described below.

2 Release of GARField to the community

GARField currently allows for navigating on the contents of ChEMBL ((version 14) and the OPS database, one of the largest public repositories of chemical structures annotated with pharmacological data. With GARField, users can make a query for compound(s) or target(s) and get bioactivities annotation and prediction within a graphical interface.

GARField was recently under testing through the testing@openphacts.org team with success. Some feedback was also suggested in order to optimize some implementations on which we are working on before to send an invitation to the whole OPS consortium for further testing.

2.1 Technical details

The application is located at http://potentia.cbs.dtu.dk/services/garfield/ and needs a login and a password to be accessed. It will be open to the public domain (without login) with the agreement of the OPS consortium.
GARFIELD is implemented as a web application using ExtJS (version 4.1.1) for the front-end interface. The queries to gather the information from the different databases are developed using javascript. Basically, our webserver access the RDF triple store of ChEMBL v14 and the OPS explorer through an application programming interface (API) provided by the OPS system. Then the chemical structures are converted internally into fingerprint in order to do the similarity search from the ligand-based. Similarly, the BLAST method was implemented internally for the protein sequence similarity search.

2.2 Use case scenario

2.2.1 Compound search

Currently, within GARFIELD, a user can search compound(s) from its name, synonyms, SMILES or drawing a structure. For example, searching for the antidepressant “paroxetine”, the user can get a 2D view of the chemical structure and some physicochemical information related to the compound (Figure 1). If the user is interesting to get information for others antidepressant like “venlafaxine” and “imipramine”, the user can search for these compounds and by clicking to the “add” button, the three compounds will be included to the “compound collection” field. By clicking to the “Show in Heatmap”, the bioactivities annotated to the three compounds will be depicted (Figure 2). The red circle is representing strong activity, the blue circle, weak activity, the black circle is associated to another type of activity (as a default the heatmap shows Ki activities, but the user can select another activity type like IC50, EC50, ED50…Finally the grey circle is for compounds where no bioactivities is associated. Some circles can show a mix of 2 colors (for example a majority of blue inner and a minority of red outer). This is the result of several measurements for the same chemical-protein interaction and with the same activity type.

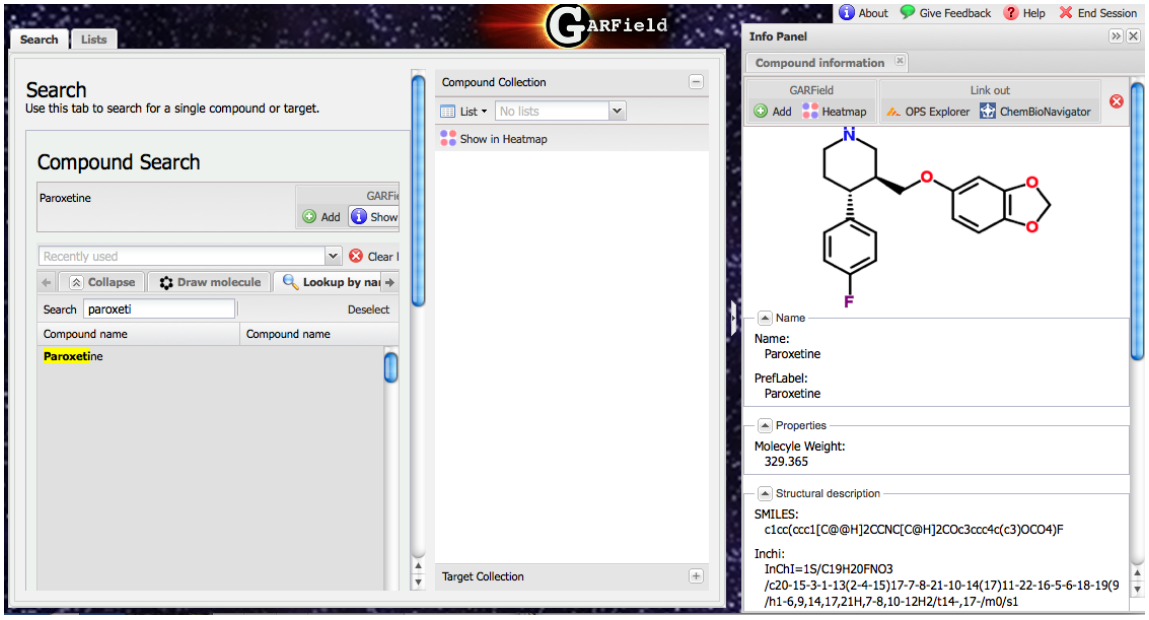

Figure 1: “Compound search” query

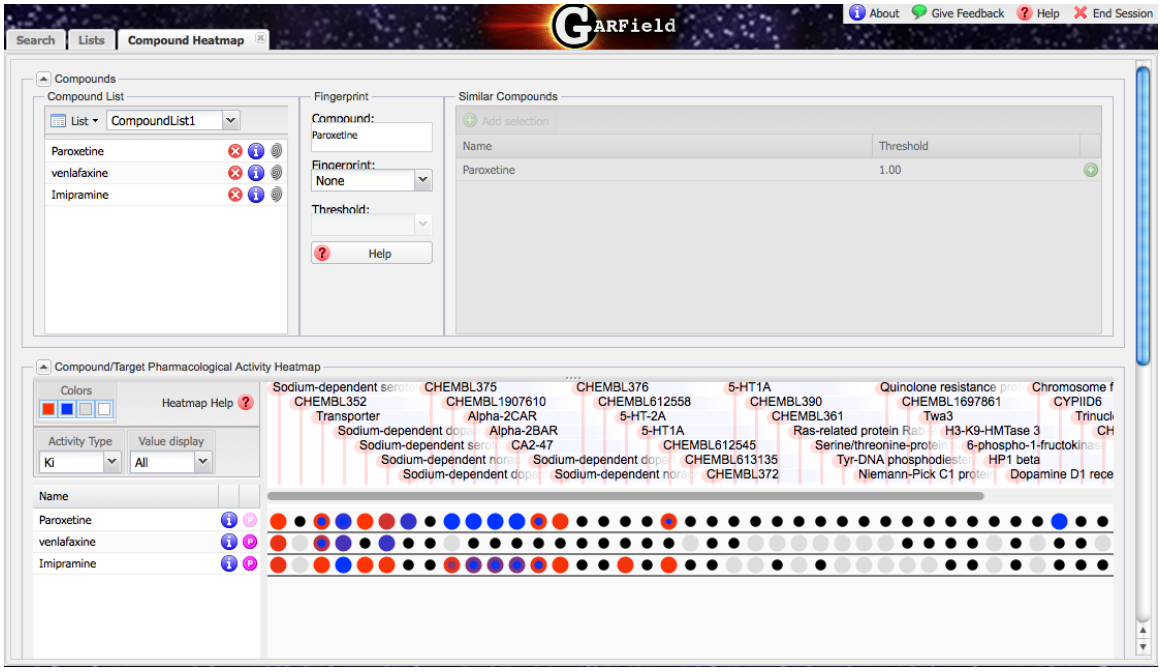

Figure 2: Heatmap representation of the bioactivities annotated to 3 compounds.

Again, clicking to the compound name or to the target name will pop up information for the chemical or protein of interest.

To enrich the chemical-protein annotation, the user might be interested to look into chemicals with a similar structure. To do so, the user has to select the compound of interest, the fingerprint type (MACCS, FP2…) and the threshold of similarity (based on the Tanimoto

coefficient). An example with paroxetine and using the MACCs fingerprint is shown in Figure 3. A list of compounds that fulfill the request will show up in a table. Then, the user can select compounds and include the bioactivities information associated into the heatmap. If the user is interested to have the heatmap for the full list of similar compounds, it is possible to do that directly from the “compound search” panel for one compound at the time.

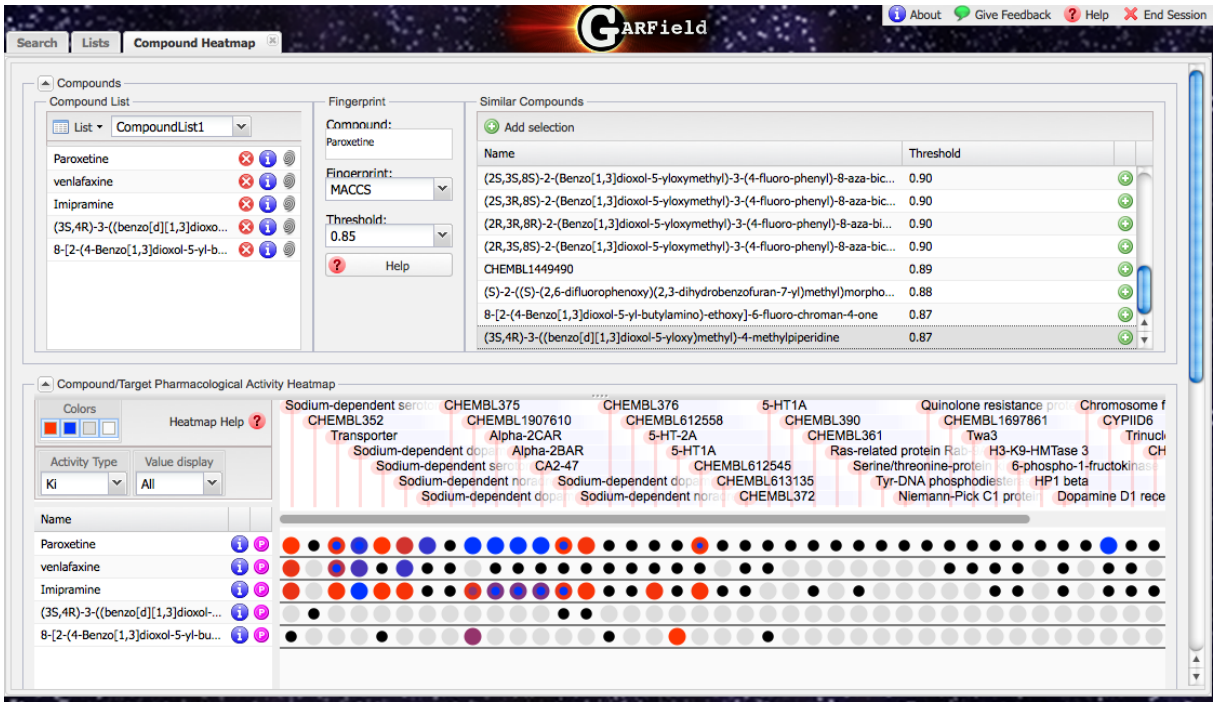

Figure 3: Similarity search with paroxetine and using the MACCS fingerprint (Threshold of 0.85)

Another prediction tool is the Similarity Ensemble Approach (SEA) developed by Keiser et al. and that has been implemented internally and integrated to GARField. In the “Activity Type” field, the user can select SEA and then the proteins predicted to interact with the compounds will be depicted following the color spectrum from red (high probability) to blue (low probability) (Figure 4)

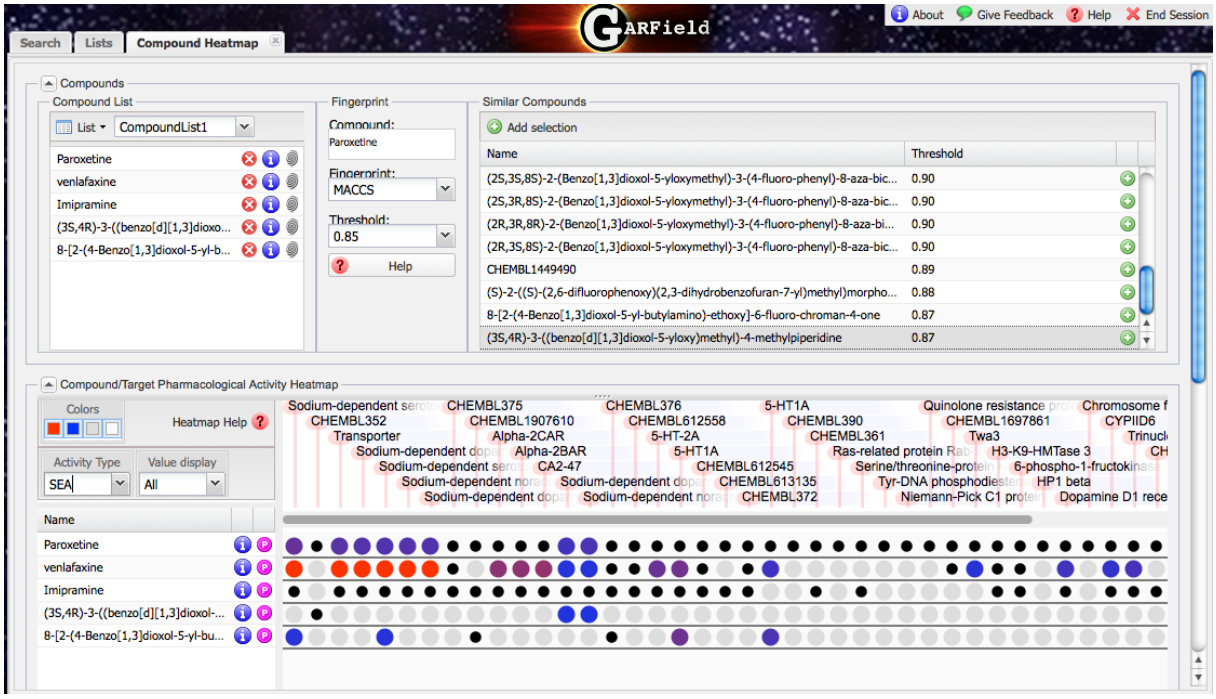

Figure 4: Prediction using the Similar Ensemble Approach (SEA).

2.2.2 Target search

Similarly to compound search, users can look for protein(s) on which compounds have been tested. The user can find a target by “look up by name” or looking on the ChEMBL Protein Target Tree (Browse target tree). With the “Add” button, it is possible to include several targets and to look on compounds tested experimentally on the set of targets selected (Figure 5)

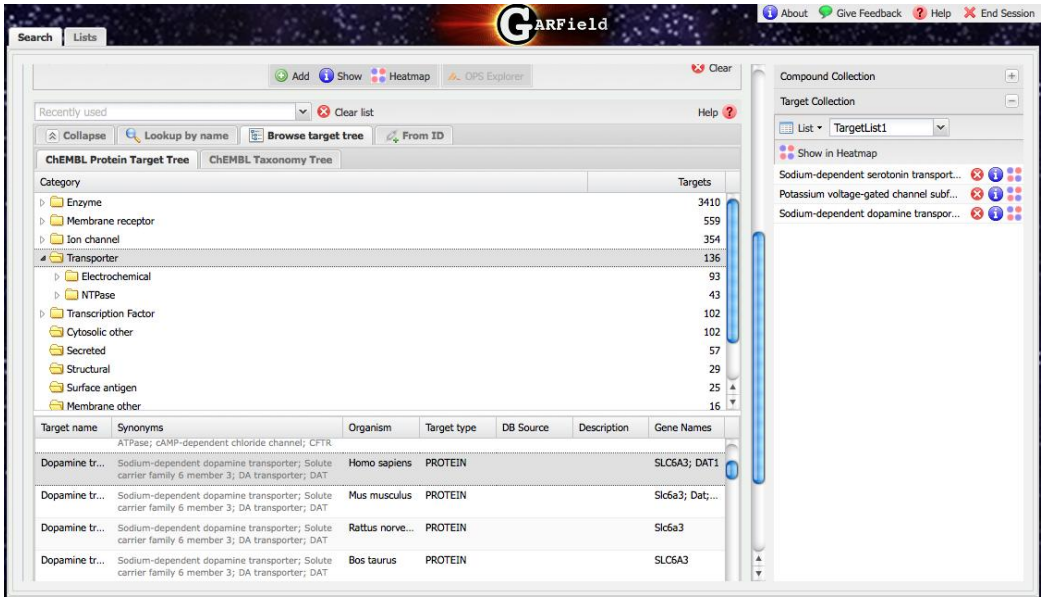

Figure 5: Screenshot of how to select three proteins (sodium dependent serotonin transporter, potassium voltage-gated channel and sodium-dependent dopamine transporter

By clicking on heatmap, the compounds with activity on the first protein selected are depicted. If the compounds have been also assessed on the two others proteins, it will be show up with a color sphere. Finally, the user can choose the sequence similarity search using BLAST and assess if the compounds has been tested on others proteins (Figure 6).

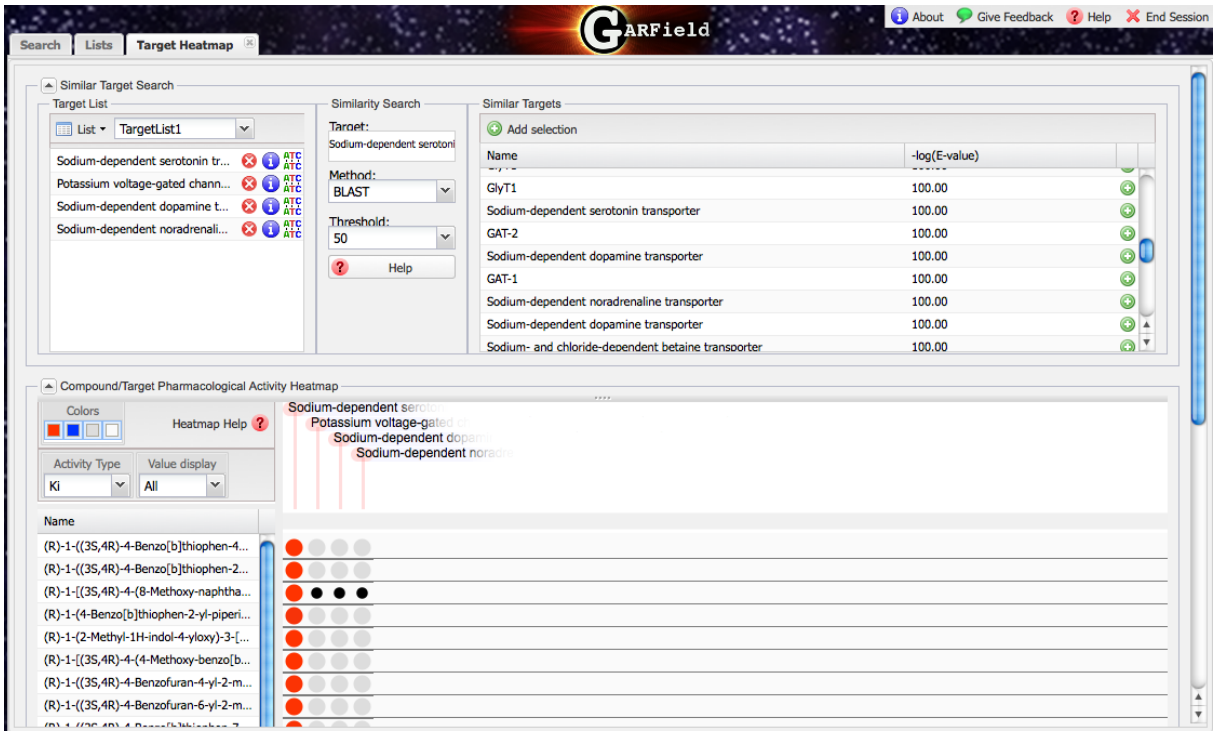

Figure 6: Chemicals with bioactivities for the three proteins selected by the user + 1 protein selected through the protein sequence similarity (based on BLAST)

2.2.3 Connection to others eApps

An interesting feature is the connection to the ChemBioNavigator and Open PHACTS Explorer eApps. So, any compounds reported in the GARField session can then be transferred to these two eApps for further information and analysis of the ligand.

A short introduction and walk-through the GARField web browser is accessible under the “Help” menu of GARField based on videos and snapshots. We encourage feedback for all aspects of our development of the exemplar. For instance, suggestion about missing features, alternative workflows, errors, etc