Open PHACTS

Deliverable 6.5.2

First ENSO Researchathon completed and output communicated to the project

Prepared by USC, Janssen
Approved by USC, Janssen, Lilly, GSK, RSC

February 2015
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: 60 months

Nature of the Deliverable	
Report	**x**
Prototype	
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:

GSK – GlaxoSmithKline – **Coordinator**
UNIVIE – Universität Wien – **Managing Entity of IMI JU funding**
DTU – Technical University of Denmark
UHAM – University of Hamburg, Center for Bioinformatics
BIT – BioSolveIT GmbH
PSMAR – Consorci Mar Parc de Salut de Barcelona
FIMIM – Fundacio Institut Mar d'Investigacions Mediques
UPF – Universitat Pompeu Fabra
LUMC – Leiden University Medical Centre
RSC – Royal Society of Chemistry
RSCWW – RSC World Wide Ltd
VUA – Stichting VU-VUMC
CNIO – Centro Nacional de Investigaciones Oncológicas
UNIMAN – University of Manchester
UM – Universiteit Maastricht
ACK – ACKnowledge
USC – Universidade de Santiago de Compostela
UBO – Rheinische Friedrich-Wilhelms-Universität Bonn
AZ – AstraZeneca AB
Pfizer – Pfizer Limited
Esteve – Laboratorios del Dr. Esteve, S.A.
Novartis – Novartis Pharma AG
ME – Merck
HLU – H. Lundbeck A/S
Lilly – Eli Lilly and Company Limited
NBIC – Stichting Netherlands Bioinformatics Centre
SIB – Swiss Institute of Bioinformatics
CD – ConnectedDiscovery
EMBL-EBI – European Molecular Biology Laboratory
Janssen – Janssen Pharmaceutica NV
OGL – OpenLink Group Ltd
OPF – The Open PHACTS Foundation
ALM – Laboratorios Almirall S.A.
SciBite – SciBite Limited

1 INTRODUCTION

The researchathon/workshop was held in Santiago de Compostela on February 16th and 17th 2015. 51 people from different pharmaceutical companies and academic institutions attended the workshop (see annex I for a full list of attendees). The title of the workshop was "**Understanding the knowledge management needs of phenotypic screening"**.

2 AIM AND EXPECTED OUTCOME

The Open PHACTS phenotypic screening workshop focused on current informatics bottlenecks of cell-based phenotypic screening and target deconvolution. The aim was to review the present computational approaches to phenotypic screens and to identify where gaps exist.
The expected outcome from this workshop was to obtain a collection of use cases relevant to phenotypic screening to be used to guide the further development of the Open PHACTS Discovery Platform.

3 WORKSHOP DYNAMICS

The agenda of the workshop is reported as annex II.
The first session (February 16th, Monday) focused on the identification of current computational bottlenecks in the field of phenotypic screening. Key contributors in the field of phenotypic screening briefly presented their approaches and needs. All presentations can be found on our project website: http://www.openphacts.org/news-and-events/news-archive/2015/382-open-phacts-phenotypic-screening-workshop.

Contributors/Company name
Javier Burgueño (Esteve)
Ola Engkvist (AstraZeneca)
Edgar Jacoby (Johnson&Johnson)
Arsenio Nueda (Almirall)
María Jesús Blanco/Marta Piñeiro (Eli Lilly)
Ceara Rea (GlaxoSmithKline)
Bryn Williams-Jones (Open PHACTS Foundation)

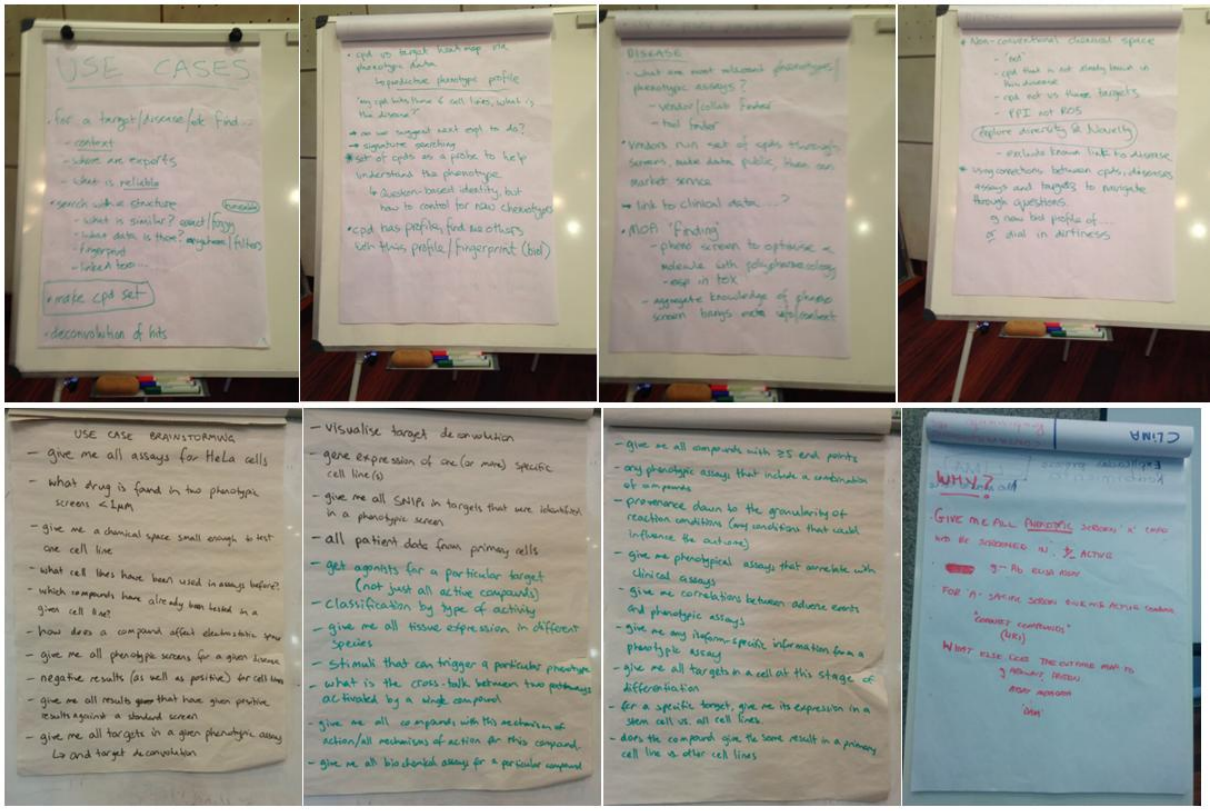
After the contributors presentations, four challengers were invited to show how they envision these bottlenecks and the role that the Open PHACTS Discovery Platform should play in helping to overcome these bottlenecks.

Challengers
Petr Bartunek (EU Open Screen)
Iván Cornella (Merck Research Laboratories) (Remote Participation)
Martin Daffertshofer (PerkinElmer)
David Wild (Data2Discovery Inc)

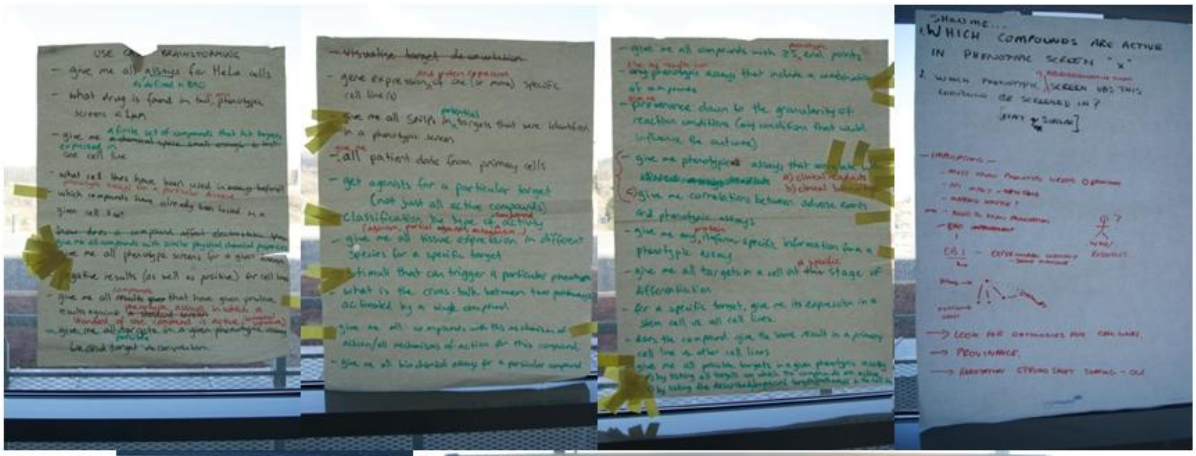
The presentations were followed by a lively discussion among contributors, challengers and attendees.

To facilitate further discussions a breakout session followed as the next item on the agenda. Three groups were created (chaired by Christine Chichester, Ola Engkvist and Bryn Williams-Jones, see the groups composition in Annex III) whose aim was to debate the different opinions presented in the morning session and discuss the relevance and technical feasibility of different solutions.

In these breakout sessions several use cases were already identified as well as the technical challenges for solving them and compiled in Flipcharts.



The second session (February 17th, Tuesday) was focused on generating a list of specific use cases related to the field of phenotypic screening, prioritized according to the needs of researchers. For this, the same three groups from the breakout sessions worked on the use cases from the previous day and 31 use cases were proposed and captured on flipcharts. All attendees voted for their preferred use cases on the flipcharts in order to prioritize them.



4 OUTCOME

1.- A list with the use cases selected by the attendees as challenging in the knowledge frontier / computational technologies / data management in phenotypic screening and target deconvolution. The list with these use cases is included in Annex IV.

2.- Five top prioritized use cases:

i. I have a compound and a phenotypic screening readout. Show me the public evidence to connect my compound and readout to targets pathways, compounds and diseases. (26 votes)
ii. For a phenotypic assay, obtain the pathways involved and find experiments on these targets. Link the results to gene expression, compounds and disease (for further clarification see annex V) (12 votes)
iii. Give me all phenotypic screens for a specific cell line for a given disease (11 votes)
iv. Give me phenotypic assays that correlate with i) clinical readouts, ii) clinical biomarkers, iii) adverse events (10 votes)
v. Give me all targets in a given phenotypic assay i) by listing all targets on which the compounds are active, ii) by listing the described/expressed targets/pathways in the cell line. (8 votes)

Once the use cases were prioritized it was agreed, in the use case-driven working way of Open PHACTS, to create working groups “ad hoc” in order to study their technical feasibility, looking for data sources, bioassay ontologies and calls needed to solve them. The analysis of the technical feasibility will prioritize the selected use cases for implementation into the Open PHACTS Discovery Platform.

Annex I – Participant list

No	Name	First Name	E-Mail	Affiliation
1	Ardao	Inés	ines.ardao@usc.es	USC
2	Bartunek	Petr	bartunek@img.cas.cz	EU Open
3	Blanco	Maria	blanco_maria@lilly.com	Lilly US
4	Brea	Pepo	pepo.brea@usc.es	USC
5	Burgueno	Javier	jburgueno@esteve.es	Esteve
6	Cadavid	María Isabel	mariaisabel.cadavid@usc.es	USC
7	Castro	Marián	marian.castro@usc.es	USC
8	Castro	Cristina	cristinacastrofdez@gmail.com	USC
9	Chichester	Christine	christine.chichester@isb-sib.ch	SIB
10	Digles	Daniela	daniela.digles@univie.ac.at	UNIVIE
11	Domíngez	Eduardo	eduardo.dominguez@usc.es	USC
12	Engkvist	Ola	Ola.Engkvist@astrazeneca.com	AZ
13	Evelo	Chris	chris.evelo@maastrichtuniversity.nl	UM
14	Franklin	Natalie	franklin_natalie_i@network.lilly.com	Lilly US
15	Furlong	Laura	lfurlong@imim.es	PSMAR
16	García	Ángel	angel.garcia@usc.es	USC
17	Garcia-Fuentes	Angel Daniel	garcia-fuentes_daniel@lilly.com	Lilly
18	González	Alejandro	alejandro.gonzalez@usc.es	USC
19	González-Couto	Eduardo	Eduardo.Gonzalez@PERKINELMER.COM	Perkin Elmer
20	Harguindey	Eduardo	harguindey_lopez_eduardo@lilly.com	Lilly
21	Van Vlijmen	Herman	hvvlijme@its.jnj.com	Janssen
22	Iglesias	Alba	alba.iglesias@usc.es	USC
23	Jacoby	Edgar	ejacoby@its.jnj.com	Janssen
24	Lage	Sonia	sonia.lage@usc.es	USC
25	Lallena	María Jose	lallena_maria_jose@lilly.com	Lilly
26	Lestón	Óscar	oscar.leston@usc.es	USC
27	Loza	Mabel	mabel.loza@usc.es	USC
28	Lumley	James Andrew	lumley_james_andrew@lilly.com	Lilly
29	Mallo	José Manuel	josemanuel.mallo@usc.es	USC
30	Marren	Derek	MARREN_DEREK@LILLY.COM	Lilly
31	Daffertshofer	Martin	Martin.Daffertshofer@PERKINELMER.COM	Perkin Elmer
32	McNeice	Kiera	mcneicek@rsc.org	RSC
33	Moreira	David	david.moreira@usc.es	USC
34	Nueda	Arsenio	arsenio.nueda@almirall.com	ALM
35	Pérez	Amparo	amparo.perez@usc.es	USC
36	Pineiro	Marta	pineiro-nunez_marta@lilly.com	Lilly US
37	Queralt	Nuria	nqueralt@imim.es	PSMAR
38	Rea	Ceara	ceara.3.rea@gsk.com	GSK
39	Rodriguez	Ricardo Julio	Ricardo.Julio.Rodriguez.Fernandez@sergas.es	IDIS
40	Roibás	Begona	begona.roibas@usc.es	USC
41	Romero	Luz	lromero@esteve.es	Esteve
42	Salgado	Leonardo	lsalgado@esteve.es	Esteve
43	Santamaría	José Manuel	josemanuel.santamaria@usc.es	USC

44	Segarra	Victor	victor.segarra@almirall.com	ALM
45	Solla	Marta	marta.solla@usc.es	USC
46	Senger	Stefan	stefan.x.senger@gsk.com	GSK
47	Torrecilla	Daniel	daniel.torrecilla@usc.es	USC
48	Torres	Fernando	torresandon@gmail.com	USC
49	Varela	María Jose	mariajose.varela@usc.es	USC
50	Wild	David	djwild@indiana.edu	Indiana University
51	Williams-Jones	Bryn	bryn@connecteddiscovery.com	CD/OPF

Annex II – Agenda

2nd Open PHACTS Workshop

"Understanding the knowledge management needs of phenotypic screening"

February 16th and 17th, 2015

February 16th, Monday
Venue: Centro de Estudos Avanzados (CEA)
Address: Parque de Vista Alegre, Rúa Salvadas s/n) 15705 Santiago de Compostela

08:30-09:00	**Reception and Welcome Coffee** Reception: Maribel Cadavid, Alba Iglesias, Sonia Lage, Oscar Lestón, Jose Manuel Mallo, Jose Manuel Santamaría and Marta Solla (*USC*)
09:00-09:30	**Workshop Welcome and Introduction** Chair: Mabel Loza Facilitators: • Kiera McNeice • Begoña Roibás
09:30-11:30	**Contributors' Presentations (listed in alphabetical order)** • Javier Burgueño *(ESTEVE)* • Ola Engkvist *(AstraZeneca)* • Edgar Jacoby *(Johnson & Johnson)* • Arsenio Nueda *(Almirall)* • María Jesús Blanco/ Marta Piñeiro *(Eli Lilly and Company)* • Ceara Rea *(GlaxoSmithKline)* • Bryn Williams-Jones *(Open PHACTS Foundation)* Chair: Derek Marren Facilitator: Kiera McNeice Assistants: Daniela Digles Alba Iglesias Sonia Lage
11:30-12:00	**Coffee Break**

12:00-13:00	**Roundtable Discussion among Contributors, Challengers and Attendees** **Challengers (listed in alphabetical order)** • Petr Bartunek *(EU-OPENSCREEN)* • Iván Cornella (Remote Participation) *(Merck Research Laboratories)* • Martin Daffertshofer *(PerkinElmer)* • David Wild *(Data2Discovery Inc.)* Chair: Edgar Jacoby Facilitator: Begoña Roibás Assistants: Sonia Lage Nuria Queralt
13:15-14:15	**Lunch**
14:15-15:45	**Breakout Sessions** Chairs: • Pepo Brea • Christine Chichester • Ola Engkvist • Bryn Williams-Jones Facilitator: Kiera McNeice
15:45-16:15	**Coffee Break**
16:15-17:15	**Conclusions and Wrap-up** Chair: Stefan Senger Facilitator: Begoña Roibás Assistants: Daniela Digles Alba Iglesias Sonia Lage
17:15-17:30	**Session Closure** Chairs: • Edgar Jacoby • Mabel Loza • Stefan Senger
20:30-22:30	**Dinner hosted by USC Team** Venue: Restaurante Enxebre Address: Praza Obradoiro 1, 15705 Santiago de Compostela

February 17th, Tuesday
Venue: CiMUS Research Center
Address: Avda. Barcelona s/n-15782 Santiago de Compostela

08:30-09:00	**Welcome Coffee**
09:00-10:00	**Use Case Brainstorming** Chairs: • Pepo Brea • Christine Chichester • Ola Engkvist • Bryn Williams-Jones Facilitator: Begoña Roibás
10:00-11:00	**Use Case Prioritisation** Chair: Stefan Senger Facilitator: Kiera McNeice Assistants: Sonia Lage Núria Queralt
11:00-11:15	**Coffee Break**
11:15-12:30	**Discussion of Next Steps** Chair: Edgar Jacoby Facilitator: Begoña Roibás Assistants: Cristina Castro Daniela Digles Alba Iglesias
12:30-13:30	**Lunch**
13:30-14:00	**Wrap-up Discussion and Conclusions** Chairs: • Edgar Jacoby • Mabel Loza • Stefan Senger Assistants: Cristina Castro Sonia Lage

For those interested, it would be the USC Team's pleasure to take Attendees for a visit of the Research Center and Screening Platform following the conclusion of the Workshop.

Workshop Personnel

- **María Jesús Blanco:** Director, Discovery Chemistry Research & Technologies, Eli-Lilly and Company
- *** Pepo Brea:** Senior Principal Scientist BioFarma USC Research Group
- **Javier Burgueño:** Director, Screening In Vitro and Assay Development, ESTEVE
- *** Christine Chichester:** Chief Information Officer, Open PHACTS; Senior Researcher, CALIPHO Group, SIB (Swiss Institute of Bioinformatics)
- **Cristina Castro:** Predoctoral Researcher, BioFarma USC Research Group
- **Martin Daffertshofer:** Senior Director Strategic Marketing, PerkinElmer
- **Daniela Digles:** Postdoc / Senior Lecturer, University of Vienna
- *** Gerhard Ecker:** University of Vienna
- *** Ola Engkvist:** Team Leader, Computational Chemistry, Chemistry Innovation Centre, Discovery Sciences, AstraZeneca R&D
- *** Edgar Jacoby:** Senior Principal Scientist, Molecular Informatics, Johnson & Johnson
- *** Alba Iglesias:** Predoctoral Researcher, BioFarma USC Research Group
- **Sonia Lage:** Predoctoral Researcher, BioFarma USC Research Group
- *** Mabel Loza:** BioFarma USC Research Group Coordinator
- *** Derek Marren:** Director, IT, Research Biology Systems, Eli Lilly and Company; Erl Wood Research IT Site Director
- *** Kiera McNeice:** eScience Support Specialist, Royal Society of Chemistry
- **Arsenio Nueda:** Head, Biological Reagents - Phenotypic Screening Biological Reagents Assay Development and Screening, Almirall
- **Nuria Queralt:** Biomedical Informatics Postdoctoral Researcher, Parc de Salut Mar.
- **Marta Piñeiro:** Director, Open Innovation Drug Discovery, Eli Lilly and Company
- **Ceara Rea:** Investigator, Computational Chemistry, GlaxoSmithKline
- *** Anika Robl:** Project Manager, IMI Open PHACTS Project
- *** Begoña Roibás:** Business Developer Manager, BioFarma USC Research Group
- *** Stefan Senger:** EFPIA coordinator, IMI Open PHACTS Project
- *** Bryn Williams-Jones:** CEO, Open PHACTS Foundation

(* Indicates a member of the organising committee)

Annex III – Break out session groups

Arsenio Nueda	Bryn
Lumley James Andrew	Bryn
Marta Piñeiro	Bryn
Stefan Senger	Bryn
David Wild	Bryn
Edgar Jacoby	Bryn
Bryn Williams	Bryn
Nuria Queralt	Bryn
Daniela Digles	Bryn
Eduardo Harguindey	Christine
María Jesus Blanco	Christine
María José Lallena	Christine
Natalie Franklin	Christine
Leonardo Salgado	Christine
Ceara Rea	Christine
Herman van Vlijmen	Christine
Laura Furlong	Christine
Martin Daffertshofer	Christine
Chistine Chichester	Christine
Nuria de Pedro	Christine
Mabel Loza	Christine
Ola Engkvist	Ola
Daniel García	Ola
Derek Marren	Ola
Javier Burgueño	Ola
Luz Romero	Ola
Petr Bartunek	Ola
Eduardo Gonzalez	Ola
Chris Evelo	Ola
Pepo Brea	Ola
Fernando Torres	Ola
Kiera McNeice	
Begoña Roibás	

Annex IV

- (26) I have a compound and a phenotypic screening readout. Show me the public evidence to connect my compound and readout to target pathways, compounds and diseases
- (12) See Annex V
- (11) Give me all phenotypic screens for a specific cell line for a given disease
- (10) Give me phenotypic assays that correlate with 1. clinical readouts, 2. clinical biomarkers, 3. adverse events
- (8) Give me all targets in a given phenotypic assay 1. by listing all targets on which the compounds are active, 2. by listing the described/expressed targets/pathways in the cell line
- (3) Give me results for phenotypic assays that include a combination of compounds
- (3) Give me all SNPs in potential targets that were identified in a phenotypic screen
- (3) Find me the targets and pathways associated with this novel phenotypic assay given these readouts
- (2) Classification by type of compound activity (agonism, partial agonism, antagonism…)
- (2) Stimuli that can trigger a particular phenotype
- (2) Give me all compounds with this mechanism of action/all mechanisms of action for this compound
- (1) Give me any protein isoform-specific information from a phenotypic assay
- (1)Which compounds have already been tested in a given cell line?
- Give me all compounds that have given positive results against phenotypic assays in which a standard of care compound is active (biological signature)
- Give me all assays for HeLa cells
- What drug is found in two, or more, phenotypic screens < 1 µM
- Give a finite set of compounds that hit targets expressed in one cell line
- What cells have been used in a phenotypic assays for a particular disease?
- Negative results (as well as positive) for cell lines
- Gene and protein expression of one (or more) specific cell line
- Give me all patients data from primary cells
- Give me all tissue expression in different species for a specific target
- What is the cross-talk between to pathways activated by a single compound?
- Give me all biochemical assays for a given compound
- Give me all compounds with ≥ 5 phenotypic endpoints
- Give me provenance down to the granularity of reaction conditions (any condition that could influence the outcome)
- Give me all targets in a cell at a specific stage of differentiation
- For a specific target give me its expression in a stem cell vs. all cell lines
- Does this compound give the same result in a primary cell line vs. other cell lines?
- Which compounds are active in phenotypic screen “x”?
- Which phenotypic neurodegenerative disease screen was this compound screened in?

Annex V