D4.2

An interactive online atlas of interconnected network maps based on the NaviCell platform

Project number	826121
Project acronym	iPC
Project title	individualizedPaediatricCure: Cloud-based virtual-patient models for precision paediatric oncology
Start date of the project	1st January, 2019
Duration	53 months
Programme	H2020-SC1-DTH-2018-1

Deliverable type	Website
Deliverable reference number	SC1-DTH-07-826121 / D4.2/ 1.0
Work package contributing to the deliverable	WP4
Due date	May 2021 – M29
Actual submission date	7th June, 2021

Responsible organisation	Institut Curie
Editor	Andrei Zinovyev
Dissemination level	PU
Revision	1.0

Abstract	This website allows the creation of network maps that can be superimposed with data, regrouped in an online atlas.
Keywords	Maps, Network, Data visualization, Web

The project iPC has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 826121.

Editor

Julio Saez-Rodriguez (UKL-HD)

Rosa Hernansaiz (UKL-HD)

Contributors (ordered according to beneficiary numbers)

Vincent Noël (CURIE)

Inna Kuperstein (CURIE)

Jean-Marie Ravel (CURIE)

Andrei Zinovyev (CURIE)

Disclaimer

Executive Summary

In this deliverable we developed NaviCell 3.0 web-server: complete and automated web-based infrastructure for hosting molecular maps, maps of patient similarity networks and multi-omics datasets.

The NaviCell platform supports molecular maps navigation and exploration using Google maps™ engine. The logic of navigation as scrolling and zooming; features as markers, pop-up bubbles and zoom bar are adapted from the Google map.

NaviCell's semantic zooming feature provides possibility for map exploring from detailed toward a top-level view achieved by gradual exclusion of details while zooming out.

NaviCell includes a powerful module for data visualization. Users can integrate and visualize different types of "omics" data on the NaviCell maps. There is also a Python API to automate tasks and communicate with the NaviCell web server.

The extended NaviCell 3.0 web-server allows users to create their own maps, superimposed with data, regrouped and share with the community, save the produced annotated maps. This activity is possible from the webpage interface, but also in order to create a large number of maps (e.g., from network-based omics data analysis), we developed a python client which allows to create maps in a programmatic fashion.

To illustrate these new functionalities, we generated community-based consensus networks inferred on 4 ipc multi-omics datasets from three paediatric tumors, Ewing sarcoma, medulloblastoma and neuroblastoma. These networks were reduced to top weighted edges and further functionally modularized. The resulting 200 modular maps were thus automatically uploaded to NaviCell 3.0. Finally, RNA expression data has been visualized in the context of the module maps using different visualization modes available in NaviCell, allowing to grasp a pattern in the gene regulation in these maps. In addition, we demonstrate how NaviCell platform can be used for online browsing of maps of Patient Similarity Networks produced from mutli-omics data analysis (iPC Deliverable 4.1).

This new NaviCell 3.0 web-server is freely available at https://navicell.curie.fr. Several step-by-step tutorials of the NaviCell 3.0 web-server usage are provided at https://github.com/sysbio-curie/NaviCell/tree/master/tutorials.

Table of Content

List of Figures

List of Tables

Chapter 1 Introduction

With the goal to allow users to create their own maps, and share with the community, the NaviCell platform has been extended. The basis of this project is thus the NaviCell platform developed earlier, which is described in the first part of this chapter, where we introduce the notion of a molecular map, principles of map construction, thus browsing of these molecular maps, and finally integrating and visualising data in the context of the maps.

Then we present our solution to allow users to upload their own maps and datasets for visualisation of the data on top of their maps. We also demonstrate how to save the produced annotated maps and how to retrieve them later, thanks to newly added functionalities of the NaviCell 3.0 web-server.

The new NaviCell 3.0 web-server is available at https://navicell.curie.fr. Several step-by-step tutorials of the NaviCell 3.0 web-server usage are provided at https://github.com/sysbio-curie/NaviCell/tree/master/tutorials.

Chapter 2 NaviCell interface: Google Maps-based tool for browsing large molecular maps and projecting multi-omics datasets

2.1 Description

NaviCell[1] is a web tool for exploring large maps of molecular interactions created in CellDesigner (http://celldesigner.org). In particular, this tool is currently used for navigation of the Atlas of Cancer Signalling Networks (ACSN)(https://acsn.curie.fr)[2]. The NaviCell tool is characterized by the unique combination of three essential features: efficient map navigation based on Google maps engine, semantic zooming for viewing different levels of details on the map and data plotting directly on top of the map[3].

2.2 Principles of knowledge formalisation and molecular map construction

Similar to geographic maps, representation of biological knowledge as a diagram allows to show complex processes inside a living cell in a visual and insightful way. It can help for systematic representation and formalization of molecular information distributed in thousands of papers. An additional advantage of representing biological processes in a graphical form is catching multiple cross-talks between components of different cell processes, which makes molecular maps also didactic tools.

Drawing individual biochemical or molecular mechanisms precludes clear representation of cross-regulations between biological processes. The alternative solution is creating a seamless map of biological mechanisms covering multiple cell processes at one canvas.

Finally, there are maps inferred directly from omics data, representing yet a different type of relations between molecules, thus providing the interaction networks. Depending on the type of data from which the networks are inferred, there are gene interactions, gene-drug interactions, protein-protein interactions network, etc. This type of map is far less detailed at the level of basic biochemical processes, but can be rather enriched by the number of participating molecules and interactions between them.

With the aim to create molecular maps compatible with multiple analytical tools, common rules of map drawing and standard graphical syntax are developed and consistently applied. The current solution suggested in the field is Systems Biology Graphical Notation (SBGN) syntax, which is compatible with many pathway drawing and analytical tools. In addition, to enable cross-compatibility, several common pathway exchange formats were suggested such as BioPAX, SBML, PSI-MI etc.

One of the most common tools for molecular diagram construction is CellDesigner, which is also useful for manually-curated maps construction and also for visual representation of data-inferred maps. The CellDesigner diagram editor is using standard Systems Biology Graphical Notation

(SBGN) syntax and is based on Systems Biology Markup Language (SBML) for further computational modelling of the map. The visual symbols used in CellDesigner cover such molecular entities as proteins, genes, RNAs, antisense RNAs, simple molecules, ions, drugs, phenotypes, complexes. Edges on the map represent biochemical reactions or reaction regulations including post-translational modifications, translation, transcription, complex formation or dissociation, transport, degradation, etc. Reaction regulations are catalysis, inhibition, modulation, trigger, and physical stimulation. It is also possible to depict cell compartments such as cytosol, nucleus, mitochondria, etc. See http://celldesigner.org for Cell Designer tool guide and http://www.sbgn.org for SBGN syntax explanation.

Finally, common entity annotation format and consistent integration of stable identifiers (ID) for map entities are essential for compatibility of maps with other tools. It also facilitates integration of data into maps and cross-curation of maps by specialists. Cell Designer tool is very flexible, it provides a possibility to introduce molecule's stable identifiers as gene, protein, drug, small molecule names. In addition, the map creators can add essential details on the entities and their interaction in a form of free text.

The maps created in Cell designer thus can be easily integrated into the NaviCell platform for further browsing and data visualisation.

2.3 Browsing maps with Google Map

Google Maps engine allows us to browse large images, in our case maps of signalling networks, and interact with them. NaviCell uses the engine to easily look at the complete map, and zoom in to see more details.

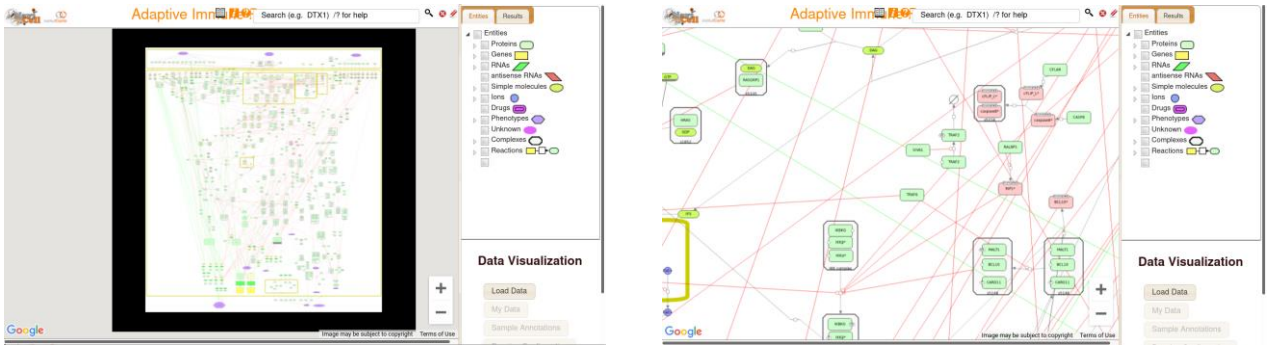

Figure 1. Views of a NaviCell map.

On the left, we see the complete map. On the right, we see a zoomed-in portion of the same map.

All nodes in the map are clickable, showing an extra information about this species, as well as links to several databases where this species is included. To build these pages, NaviCell relies on the annotation present in the CellDesigner file.

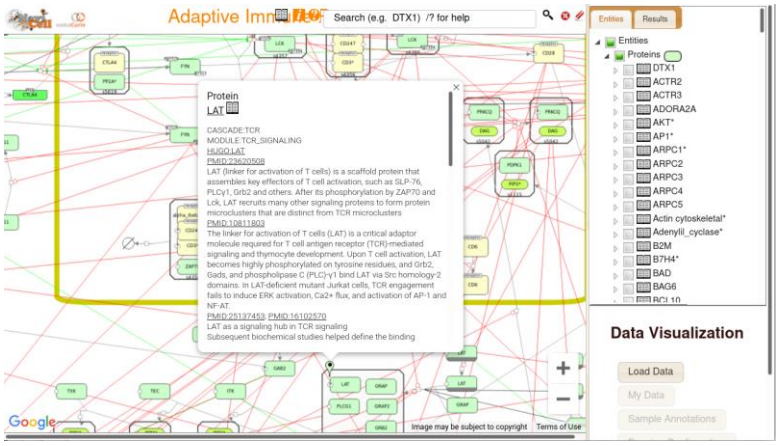

Figure 2. View of the tooltip of a map entity.

This allows to get more information on the entity, and to show links to access various public databases.

2.4 Plotting data on top of maps

An important usage of NaviCell is to visualize data superimposed on the map. This provided the users with visual information on changes in molecule's levels and also fluctuations in the molecular mechanisms where the molecules are participating, so-called pathway activities. In order to visualise data in the context of the maps, users can load data, and then use this data in the plot configuration. NaviCell data visualization web service is able to process several types of 'omics' data.

The different biological data types are mapped into several internal represen-tations that determine what methods of data visualization can be applied. For instance, a mRNA expression data matrix is associated with a 'continuous' numerical internal representation. Thus, if the user chooses to display this data with a heatmap, mapping to a color gradient will be applied by default, with a possibility to modify the default settings. On the other hand, when a matrix with discrete copy-number data is loaded, it is associated with a 'discrete ordered' internal representation for which a specific color palette is applied for visualization, with a distinct color associated to each discrete copy-number state.

The input format for data sets is standard tab-delimited text files, with rows representing genes, their products or metabolites and columns representing samples (or experiments, or time points). Genes in the first column should be labeled by their standard HUGO (HGNC) gene symbols that will be associated with the different entities on the pathway map. Users can also upload sample annotations (simple tab-delimited text files) to specify how samples can be organized into meaningful groups (e.g. disease versus control).

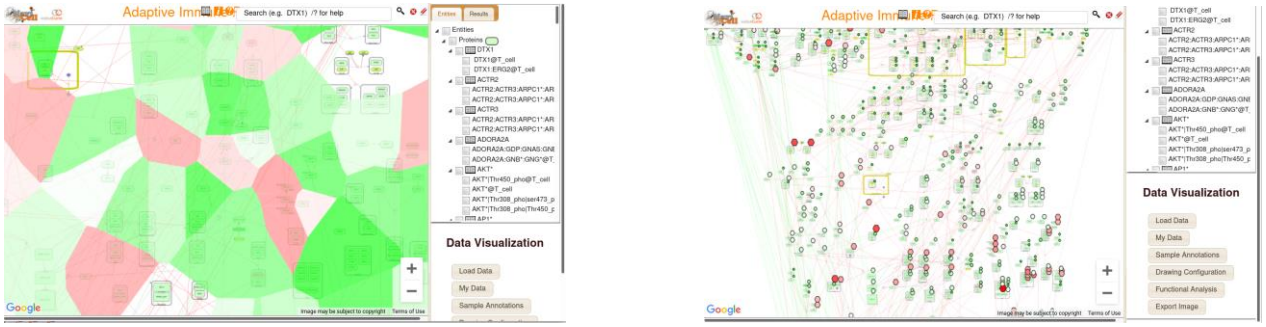

Figure 3. Different ways of plotting data in a NaviCell map.

On the left, voronoi cells are displayed with a color to indicate their under/over-expression. On the right, the same under/over-expression is displayed next to the node, in the form of a glyph.

2.5 NaviCell API

The web service is implemented at two interconnected functional levels. The first level corresponds to the interactive, dialog-based use of the web interface to load and visualize data. This level is implemented as a specific set of JavaScript functions linked to the menu located in the right hand panel of the NaviCell web interface. The second level corresponds to a programmatic usage of the service, which allows users to write the code that will communicate with the NaviCell server in order to automate all the visualization operations. This functional level is implemented as a RESTful web service, using standard web protocol HTTP operations and data encoding (JSON) to perform all the necessary operations.

The NaviCell API allows users to control a web browsing session using an API, available in several programming languages such as Python and R, and Java. These APIs will facilitate the integration of NaviCell Web Service in other applications. Using this API, users can programmatically load data on a NaviCell map and use it to plot data on top of it, highlight species, etc.

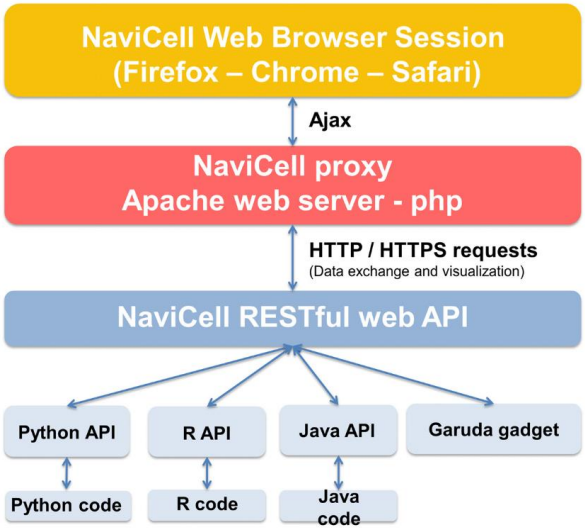

Figure 4. NaviCell API Structure.

Various API are available to communicate with the NaviCell proxy. They allow to open sessions, opening a map in a web browser and performing various actions on the map.

2.6 cd2sbgnml: bidirectional conversion between CellDesigner and SBGN formats

Currently, the NaviCell tool is tightly coupled to the use of CellDesigner format. CellDesigner is a well-established biological map editor used in many large-scale scientific efforts. However, the interoperability between the Systems Biology Graphical Notation (SBGN) Markup Language (SBGN-ML) and the CellDesigner's proprietary Systems Biology Markup Language (SBML) extension formats remains a challenge due to the proprietary extensions used in CellDesigner files.

In the scope of the work on developing a network-based map hosting platform for iPC project, we relaxed this limitation of NaviCell and opened it to the use of existing systems biology standards for storing network maps. We developed a library named cd2sbgnml and an associated web service for

bidirectional conversion between CellDesigner's proprietary SBML extension and SBGN-ML formats. The cd2sbgnml converter was successfully used for the translation of comprehensive large-scale diagrams such as the RECON Human Metabolic network and the complete Atlas of Cancer Signalling Network, from the CellDesigner file format into SBGN-ML.

The cd2sbgnml conversion library and the web service were developed in Java, and distributed under the GNU Lesser General Public License v3.0. The sources along with a set of examples are available on GitHub (https://github.com/sbgn/cd2sbgnml and https://github.com/sbgn/cd2sbgnml-webservice, respectively).

A report on the functionality and development of cd2sbgnml has been published as Bioinformatics Application Note[4].

Chapter 3 NaviCell 3.0 web-server: complete and automated web-based infrastructure for hosting molecular maps and multi-omics datasets

This work brings a new level to the NaviCell platform, called NaviCell 3.0 web server. Users can now easily create new maps, and make them available to the community. A metadata database allows searching for specific species within available maps. The NaviCell web server also makes use of the NaviCell proxy to create browsing sessions, and save them.

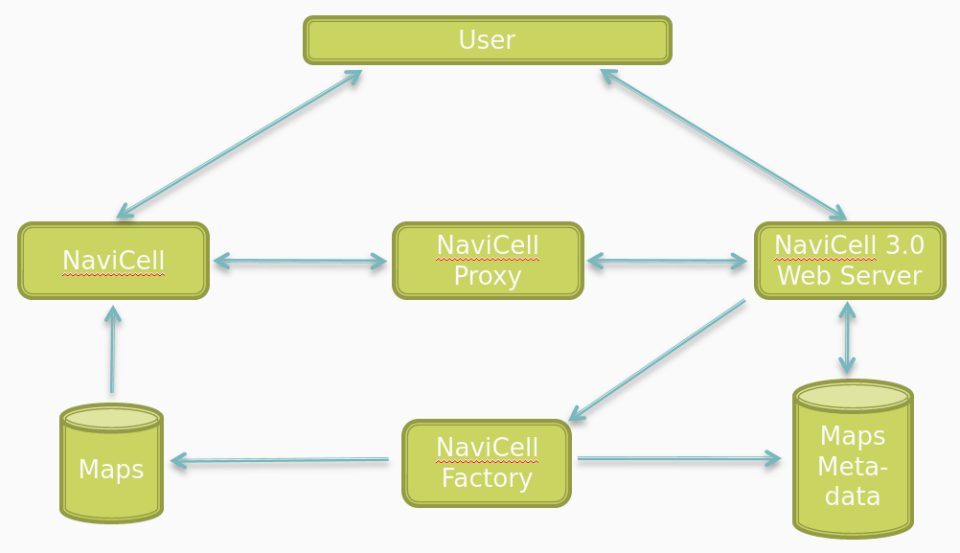

Figure 5. Workflow of NaviCell 3.0 web server.

3.1 Visualizing public maps

The newly developed NaviCell 3.0 web server contains a collection of public maps, which are available in the Maps tab. There users can browse the maps, open a session, or have a quick view at the map information. These public maps are associated with tags, which allow the users to quickly filter all the available maps by tags.

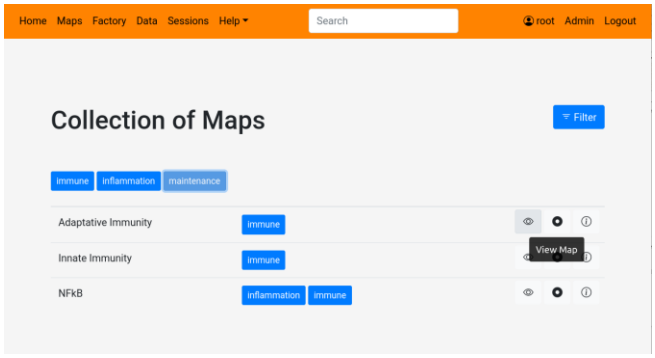

Figure 6. List of public maps.

On top, the list of tags used in the public maps. They can be inactivated to filter those maps out of the list.

3.2 Create Map

The maps can be created by going to the Factory tab, which is only available to registered users. In this page, users have access to the list of their maps, and have the possibility to turn them public. They also have access to the map creation form.

This form first asks for the name of the map, then for a CellDesigner file representing the map. To generate the image of the map, users have the possibility to use an automatic rendering system. Using this automatic rendering system, maps are visually represented using an automatic layout. Alternatively, previously-designed manual layout of the maps can be preserved and used for map visalisation. In this case, the users are requested to provide the PNG image of the maps with the predefined layout. Finally, a list of tags can be added to the maps, which will help to search for the map across all maps displayed in the NaviCell public collection.

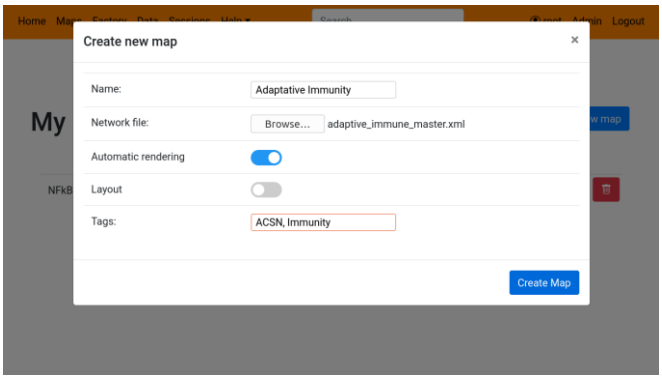

Figure 7. Creating a new map

3.3 NaviCell sessions

NaviCell 3.0 web server now provides a novel mechanism to use the NaviCell API to create new sessions, save them, and share them as working scripts allowing other users to reproduce the results of analysis performed within NaviCell. Users can start a new session by clicking on the round button in the list of maps.

When starting a new session, a new page loads which is dedicated to this session, a new window opens with the NaviCell map chosen. From the map window, users can already browse the map. Every action performed on this window will be recorded. At the same time, in the session window, the list of commands will appear progressively. At any moment, users can save this session to resume it later. From this page, a new dataset can be loaded into the map, and will be available if the session is later resumed.

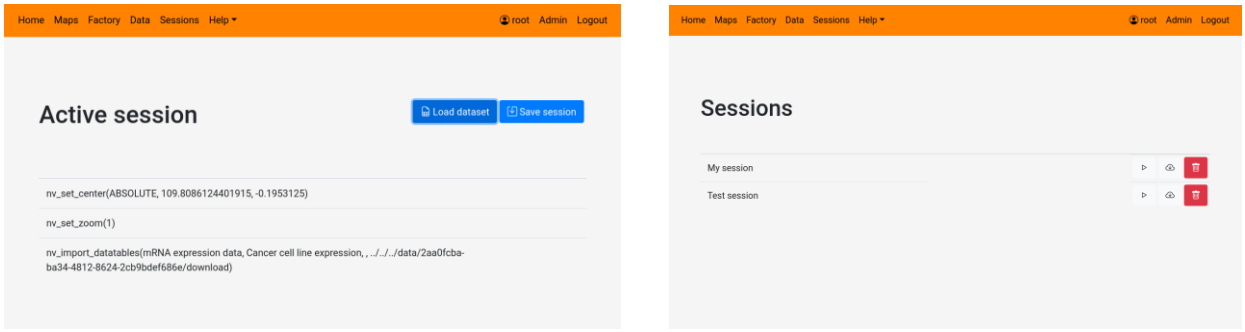

Figure 8. NaviCell sessions.

On the left, the page of an active session, showing the command history. Users can save their session, or load data into the session. On the right, the list of saved sessions, which allows to export, resume, or delete a session.

3.4 Upload new dataset

Registered users of NaviCell can store datasets, which allows them to quickly access it when loading a dataset in a session. Users also have the possibility to turn these datasets public.

To upload a new dataset, users must give it a name, and describe the type of datasets. The file can be a local file, or a public url.

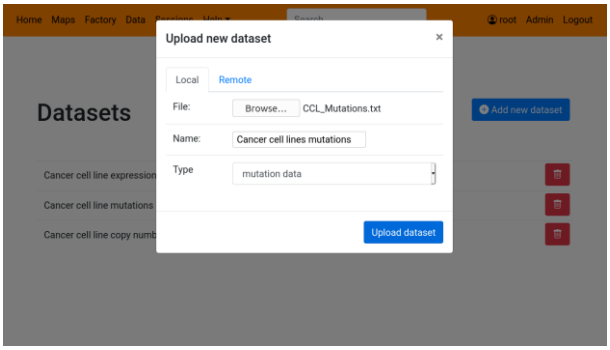

Figure 9. The form to upload a new dataset to the user's account.

A local file can be provided, as well as a remote URL.

3.5 Search elements in public maps

NaviCell allows to search for elements across all public maps.

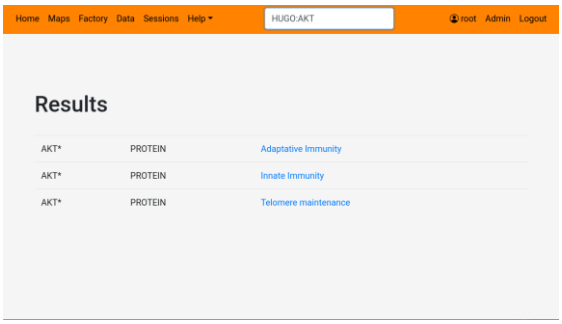

Figure 10. Search in public maps.

Search can be done with a name, or a HUGO identifier. Here, we show a search for AKT, which returns results in multiple maps. The link will open a new map and highlight the search element.

3.6 Map creation from Python

To create a large number of maps (e.g., from network-based omics data analysis), we developed a python client which allows you to create a map in a programmatic fashion. This allows us to create maps from a single CellDesigner file, or to use pre-generated images.

```
1  from client import Client
2
3  MAPS_FOLDER = "maps_src/navicell_src/"
4
5  c = Client('root', '******', url="http://navicell.curie.fr")
6
7  # mTOR signalling network
8
9  map = c.uploadLayeredMap(
10     'mTOR signalling network',
11     MAPS_FOLDER + "mtor_src/master.xml",
12     [(MAPS_FOLDER + "mtor_src/master-%d.png" % i) for i in range(4)],
13     tags=["NaviCell Collection", "mTOR", "Process"]
14
15 )
16
17 c.publishMap(map)
```

Figure 11. Example of python script uploading a map.

Example of python script uploading a map of Ewing's sarcoma signalling network on the NaviCell web server. This allows us to automatize the uploading process, and handle a large amount of maps.

Chapter 4 Using iPC module network maps with NaviCell 3.0

4.1 iPC data-driven molecular networks

Consensus molecular networks inferred from pediatric cancer type specific networks in iPC Deliverable 4.1 were retrieved for visualization of cancer specific modules. Briefly, the main idea below the construction of consensus networks is the fact that the performance of inference methods varies strongly, with a different method performing best in each setting, and that integrated predictions across inference methods, and resulting community-based consensus networks, are more robust[5].

Thus, networks of statistical associations between molecular profiles were computed by aggregating multiple methods. Specifically, COSIFER, a package to infer molecular networks from expression data using state-of-the-art consensus approaches, has been employed. Consensus networks were constructed using SUMMA (Unsupervised Evaluation and Weighted Aggregation of Ranked Predictions) (Ahsen et al., 2018) approach, an ensemble learning algorithm, completely unsupervised, that estimates the AUROC (Area under the Receiver Operating Characteristic) of each single inference method and associates a weight to each method proportional to its estimated AUROC.

The resulting consensus networks used genes as nodes and edge intensities were the normalized weighted sum of intensities predicted from the different methods.

4.2 Use of iPC data

Consensus networks were inferred on 4 multi-omics datasets from three paediatric tumors, namely one from Ewing sarcoma, two medulloblastoma and one neuroblastoma. The Ewing sarcoma data consists of gene expression profiles[6]. The medulloblastoma datasets come from two different cohorts of patients. The first, described in detail in reference 7 consists of gene expression, methylation, proteomic and phospho-proteomic profiles (with missing values); the second, described in detail in reference 8, consists of gene expression and methylation profiles. The neuroblastoma dataset[9] consists of microarray-based comparative genomic hybridization (aCGH), gene expression and methylation.

4.3 Module identification and functional enrichment analysis

SUMMA consensus networks were generally dense and required a significant amount of disk space to store them (for example the SUMMA network obtained at the gene expression level of the medulloblastoma dataset from reference 10 was 6.7Gb). To deal with this problem, we set an edge weight threshold to reduce the data type-specific consensus networks, so that they contain only top weighted edges and are amenable for further network-based analyses.

The reduced consensus networks at each level of cellular organization have been retrieved from nextcloud (https://data.ipc-project.bsc.es/apps/files/?dir=/D4.1&fileid=13268). Module identification has been performed through MCL cluster function in Cytoscape on each of these networks. Different values for the inflation parameter have been tested to achieve the greater number of modules with more than 10 genes. When no modules could be found, a cutoff on edge weight has been

superimposed heuristically. In Table 1 we present the number of modules identified at each molecular level for each one of the molecular networks inferred in the context of the iPC deliverable 4.1. Functional enrichment of the modules has been performed using the gprofiler2 package in R.

The inferred network modules were used as features of multi-omics datasets, with a score computed using the first principal component (eigengene score). Using these features, Patient Similarity Networks were constructed representing the similarity between individual patients' multiomics profiles.

The detailed description of the data-driven network and module inference can be found in the text of Deliverable 4.1.

Table 1. Modules identified at each molecular level for each one of the molecular networks inferred in the context of the iPC data

Tumor type	**Level**	**Number of modules (N>=10)**	**Reference**
Medulloblastoma **MB**	Gene expression	51	Forget et al. 2018
	Methylation	16	
	Proteomics	43	
	Phosphoproteomics	2	
Medulloblastoma **MB**	Gene expression	8	Cavalli et al. 2017
	Methylation	7	
Ewing sarcoma **ES**	Gene expression	57	Postel-Vinay et al. 2012
Neuroblastoma **NB**	Gene expression	38	Henrich et al. 2016
	Methylation	6	
	aCGH	1	

4.4 Automatic map generation from Cytoscape JSON format to CellDesigner file and rendered image

In order to facilitate creation of data-driven network maps, we developed a converter from .cyjs files (JSON Cytoscape format) to a pair of CellDesigner .xml and .png files, which can be uploaded to NaviCell 3.0 collection either manually or automatically (see below). A Jupyter notebook with an example of using the converter can be found at https://github.com/sysbio-curie/NaviCell/tree/master/CYJS2CD4_converter. This converter was used to generate a few hundreds of module maps, for each of the modules described above. The naming of the module maps followed the following format : <Cancer Type>_<Data type>_<module number>. The generated module maps can be used to visualize data from the corresponding datasets (see below).

4.5 Automatic map upload to NaviCell 3.0

In order to upload these newly generated maps, we used the simple python client described in 3.6. A Jupyter notebook with an example of using the automated upload function is available at https://github.com/sysbio-curie/NaviCell/tree/master/CYJS2CD4_converter. Using this code, we populated the NaviCell 3.0 database with several hundreds of data-driven module network maps, each of which can be used for network-based pediatric cancer omics data visualization.

4.6 Visualisation of pediatric cancer data in the context of iPC module network map

4.6.1 Visualization of molecular data on top of molecular entity networks

NaviCell Data visualisation tool allows to overlay different data types on the network or visualize the same data using different visualisation modes. Figure 12 shows an overlay of RNA expression data on the network, allowing to grasp a pattern in the gene regulation in these correlation networks.

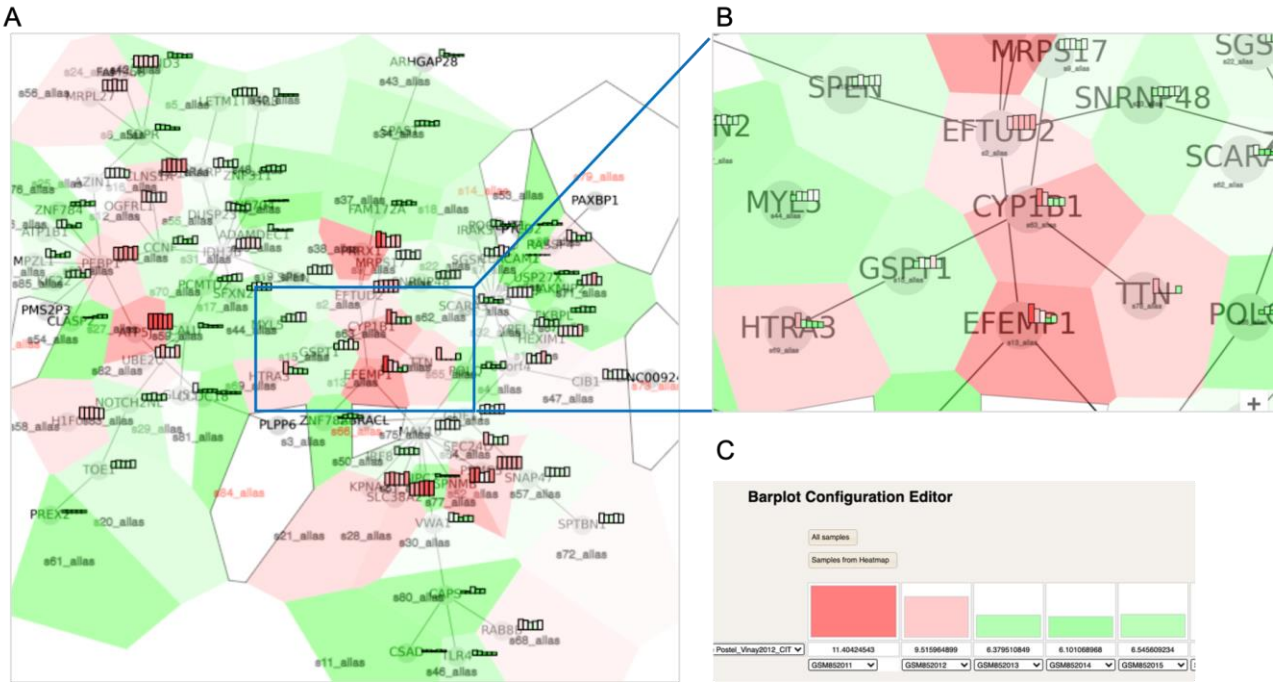

Figure 12. Visualisation of RNA expression data from Ewing sarcoma transcriptome

Visualisation of RNA expression data from Ewing sarcoma transcriptome Postel_Vinay2012_CIT dataset on ES_Postel_Vinay_Expression_module02 data-driven network. (A) Background coloring is done using “Map staining“ visualisation method, the color intensity corresponds to the expression values of each gene on the network. (B) Barplots represent expression values of corresponding genes in five samples from the dataset. (C) Barplot configuration.

4.6.2 Visualization of omics data on top of patient similarity networks (PSN)

The NaviCell platform can be used as an online data visualization front-end for molecular networks, but also for networks representing similarities between patients. In this case it becomes possible to visualize scores attributed to patients. Patient similarity networks computed in D4.1 can be used for the purpose of data visualization in particular datasets. In Figure 13 we provide an example of visualization of a score computed based on application of ROMA method to a module of genes identified in D4.1 for the multi-omics medulloblastoma dataset published in Forget et al, 2018. The multiomics patient similarity network computed for this dataset in Figure 13 has been automatically converted to NaviCell format and uploaded to NaviCell web server, as well as other PSNs.

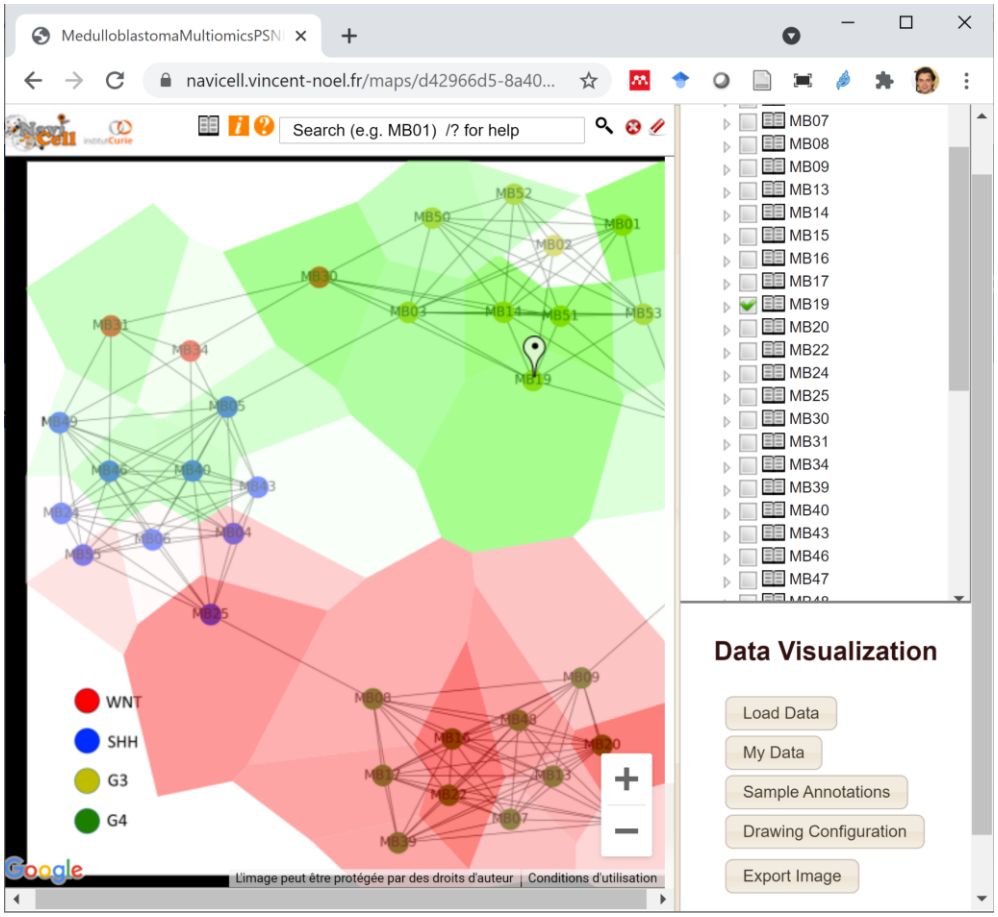

Figure 13. Visualisation of scores computed for a gene expression module.

Visualisation of scores computed for a gene expression module identified in the multiomics data analysis of medulloblastoma (details in D4.1) on top of a kNN-graph (k=10) representing similarities between medulloblastoma patients of four molecular subtypes.

The similarity between patients is computed using a composite metrics of module activities as described in D4.1. Red color highlights those patients where the module activity is high, and green those in which the module activity is low.

Chapter 5 Summary and Conclusion

In this deliverable, we have developed a new NaviCell 3.0 platform for hosting comprehensive maps of molecular interactions, provided in CellDesigner format. The hosting platform was fully re-worked compared to the previous version of the NaviCell web-site, and now can be fully automated via the use of programmatic API. The platform was adapted and extended towards the use of data-driven network maps which are produced in the framework of iPC project (D4.1), based on pediatric cancer multi-omics data. Therefore, the new platform represents an efficient, flexible and rich tool for network-based data visualization and analysis.

The NaviCell 3.0 platform has an advantage of complete automation of uploading maps and datasets, and configuring the data visualization scene. Therefore, it can be integrated with many other existing tools. At the moment NaviCell API is used as a front-end for visualizing the multiomics data and the results of their analysis in the packages such as ACSNMinerR (https://cran.r-project.org/web/packages/ACSNMineR/) and BIODICA (https://github.com/LabBandSB/BIODICA), the last one being exploited in iPC in the Task 3.1.

In the near future we plan to include the iPC cloud-based computing infrastructure, in close collaboration with the BSC iPC partner. The purpose will be to use NaviCell 3.0 as a tool for presenting the results of network-based data analysis pipeline developed in D4.1, and also a tool for facilitating interpreting the results of machine learning-based methods from WP3, by projecting them on top of appropriate molecular maps. A rich collection of such maps, part of them specific to cancer, is currently installed and can be immediately used at the new NaviCell 3.0 web platform in the scope of iPC project.

A publication on the newly developed NaviCell 3.0 web-based map hosting platform and examples of its use is in preparation.

Chapter 6 List of Abbreviations

Abbreviation	Translation
ACSN	Atlas of Cancer Signaling Network
SBML	Systems Biology Markup Language
SBGN	Systems Biology Graphical Notation
COSIFER	Consensus inference of molecular networks
API	Application Programming Interface
PNG	Portable Graphics Format
JSON	JavaScript Object Notation
HUGO	Human Genome Organisation

Chapter 7 Bibliography

[1] Kuperstein I, Cohen DP, Pook S, Viara E, Calzone L, Barillot E, Zinovyev A. NaviCell: a web-based environment for navigation, curation and maintenance of large molecular interaction maps. BMC Syst Biol. 2013 Oct 7;7:100. doi: 10.1186/1752-0509-7-100. PMID: 24099179; PMCID: PMC3851986.

[2] Kuperstein I, Bonnet E, Nguyen HA, Cohen D, Viara E, Grieco L, Fourquet S, Calzone L, Russo C, Kondratova M, Dutreix M, Barillot E, Zinovyev A. Atlas of Cancer Signalling Network: a systems biology resource for integrative analysis of cancer data with Google Maps. Oncogenesis. 2015 Jul 20;4(7):e160. doi: 10.1038/oncsis.2015.19.

[3] Bonnet E, Viara E, Kuperstein I, Calzone L, Cohen DP, Barillot E, Zinovyev A. NaviCell Web Service for network-based data visualization. Nucleic Acids Res. 2015 Jul 1;43(W1):W560-5. doi: 10.1093/nar/gkv450. Epub 2015 May 9. PMID: 25958393; PMCID: PMC4489283.

[4] Balaur I, Roy L, Mazein A, Karaca SG, Dogrusoz U, Barillot E, Zinovyev A. cd2sbgnml: bidirectional conversion between CellDesigner and SBGN formats. Bioinformatics. 2020 Apr 15;36(8):2620-2622. doi: 10.1093/bioinformatics/btz969. Erratum in: Bioinformatics. 2020 Dec 8;36(19):4975. PMID: 31904823.

[5] Marbach D, Costello JC, Küffner R, Vega NM, Prill RJ, Camacho DM, Allison KR; DREAM5 Consortium, Kellis M, Collins JJ, Stolovitzky G. Wisdom of crowds for robust gene network inference. Nat Methods. 2012 Jul 15;9(8):796-804. doi: 10.1038/nmeth.2016. PMID: 22796662; PMCID: PMC3512113.

[6] Mehmet Eren Ahsen, Robert M Vogel, Gustavo A Stolovitzky Unsupervised Evaluation and Weighted Aggregation of Ranked Classification Predictions Journal of Machine Learning Research, 20(166):1−40, 2019.

[7] Postel-Vinay S, Véron AS, Tirode F, Pierron G, Reynaud S, Kovar H, Oberlin O, Lapouble E, Ballet S, Lucchesi C, Kontny U, González-Neira A, Picci P, Alonso J, Patino-Garcia A, de Paillerets BB, Laud K, Dina C, Froguel P, Clavel-Chapelon F, Doz F, Michon J, Chanock SJ, Thomas G, Cox DG, Delattre O. Common variants near TARDBP and EGR2 are associated with susceptibility to Ewing sarcoma. Nat Genet. 2012 Feb 12;44(3):323-7. doi: 10.1038/ng.1085. PMID: 22327514.

[8] Forget A, Martignetti L, Puget S, Calzone L, Brabetz S, Picard D, Montagud A, Liva S, Sta A, Dingli F, Arras G, Rivera J, Loew D, Besnard A, Lacombe J, Pagès M, Varlet P, Dufour C, Yu H, Mercier AL, Indersie E, Chivet A, Leboucher S, Sieber L, Beccaria K, Gombert M, Meyer FD, Qin N, Bartl J, Chavez L, Okonechnikov K, Sharma T, Thatikonda V, Bourdeaut F, Pouponnot C, Ramaswamy V, Korshunov A, Borkhardt A, Reifenberger G, Poullet P, Taylor MD, Kool M, Pfister SM, Kawauchi D, Barillot E, Remke M, Ayrault O. Aberrant ERBB4-SRC Signaling as a Hallmark of Group 4 Medulloblastoma Revealed by Integrative Phosphoproteomic Profiling. Cancer Cell. 2018 Sep 10;34(3):379-395.e7. doi: 10.1016/j.ccell.2018.08.002. PMID: 30205043.

[9] Cavalli FMG, Remke M, Rampasek L, Peacock J, Shih DJH, Luu B, Garzia L, Torchia J, Nor C, Morrissy AS, Agnihotri S, Thompson YY, Kuzan-Fischer CM, Farooq H, Isaev K, Daniels C, Cho BK, Kim SK, Wang KC, Lee JY, Grajkowska WA, Perek-Polnik M, Vasiljevic A, Faure-Conter C, Jouvet A, Giannini C, Nageswara Rao AA, Li KKW, Ng HK, Eberhart CG, Pollack IF, Hamilton RL, Gillespie GY, Olson JM, Leary S, Weiss WA, Lach B, Chambless LB, Thompson RC, Cooper MK, Vibhakar R, Hauser P, van Veelen MC, Kros JM, French PJ, Ra YS, Kumabe T, López-Aguilar E, Zitterbart K, Sterba J, Finocchiaro G, Massimino M, Van Meir EG, Osuka S, Shofuda T, Klekner A, Zollo M, Leonard JR, Rubin JB, Jabado N, Albrecht S, Mora J, Van Meter TE, Jung S, Moore AS, Hallahan AR, Chan JA, Tirapelli DPC, Carlotti CG, Fouladi M, Pimentel J, Faria CC, Saad AG, Massimi L, Liau LM, Wheeler H, Nakamura H, Elbabaa SK, Perezpeña-Diazconti M, Chico Ponce de León F,

Robinson S, Zapotocky M, Lassaletta A, Huang A, Hawkins CE, Tabori U, Bouffet E, Bartels U, Dirks PB, Rutka JT, Bader GD, Reimand J, Goldenberg A, Ramaswamy V, Taylor MD. Intertumoral Heterogeneity within Medulloblastoma Subgroups. Cancer Cell. 2017 Jun 12;31(6):737-754.e6. doi: 10.1016/j.ccell.2017.05.005. PMID: 28609654; PMCID: PMC6163053.

[10] Henrich KO, Bender S, Saadati M, Dreidax D, Gartlgruber M, Shao C, Herrmann C, Wiesenfarth M, Parzonka M, Wehrmann L, Fischer M, Duffy DJ, Bell E, Torkov A, Schmezer P, Plass C, Höfer T, Benner A, Pfister SM, Westermann F. Integrative Genome-Scale Analysis Identifies Epigenetic Mechanisms of Transcriptional Deregulation in Unfavorable Neuroblastomas. Cancer Res. 2016 Sep 15;76(18):5523-37. doi: 10.1158/0008-5472.CAN-15-2507. Epub 2016 Sep 7. PMID: 27635046.