BMC Bioinformatics

RESEARCH

Open Access

Metabolic and signalling network maps integration: application to cross-talk studies and omics data analysis in cancer

Nicolas Sompairac[1,2,3,4], Jennifer Modamio[5], Emmanuel Barillot[1,2,3,4], Ronan M. T. Fleming[5], Andrei Zinovyev[1,2,3,4] and Inna Kuperstein[1,2,3,4*]

Abstract

Background: The interplay between metabolic processes and signalling pathways remains poorly understood. Global, detailed and comprehensive reconstructions of human metabolism and signalling pathways exist in the form of molecular maps, but they have never been integrated together. We aim at filling in this gap by integrating of both signalling and metabolic pathways allowing a visual exploration of multi-level omics data and study of cross-regulatory circuits between these processes in health and in disease.

Results: We combined two comprehensive manually curated network maps. Atlas of Cancer Signalling Network (ACSN), containing mechanisms frequently implicated in cancer; and ReconMap 2.0, a comprehensive reconstruction of human metabolic network. We linked ACSN and ReconMap 2.0 maps via common players and represented the two maps as interconnected layers using the NaviCell platform for maps exploration (https://navicell.curie.fr/pages/maps_ReconMap%202.html). In addition, proteins catalysing metabolic reactions in ReconMap 2.0 were not previously visually represented on the map canvas. This precluded visualisation of omics data in the context of ReconMap 2.0. We suggested a solution for displaying protein nodes on the ReconMap 2.0 map in the vicinity of the corresponding reaction or process nodes. This permits multi-omics data visualisation in the context of both map layers. Exploration and shuttling between the two map layers is possible using Google Maps-like features of NaviCell. The integrated networks ACSN-ReconMap 2.0 are accessible online and allows data visualisation through various modes such as markers, heat maps, bar-plots, glyphs and map staining. The integrated networks were applied for comparison of immunoreactive and proliferative ovarian cancer subtypes using transcriptomic, copy number and mutation multi-omics data. A certain number of metabolic and signalling processes specifically deregulated in each of the ovarian cancer sub-types were identified.

Conclusions: As knowledge evolves and new omics data becomes more heterogeneous, gathering together existing domains of biology under common platforms is essential. We believe that an integrated ACSN-ReconMap 2.0 networks will help in understanding various disease mechanisms and discovery of new interactions at the intersection of cell signalling and metabolism. In addition, the successful integration of metabolic and signalling networks allows broader systems biology approach application for data interpretation and retrieval of intervention points to tackle simultaneously the key players coordinating signalling and metabolism in human diseases.

Keywords: Signalling, Metabolism, Networks, Comprehensive map, Systems biology, Cancer, Multi-level omics data, Data visualisation

* Correspondence: inna.kuperstein@curie.fr
[1]Institut Curie, 26 rue d'Ulm, F-75005 Paris, France
[2]Inserm, U900, F-75005 Paris, France
Full list of author information is available at the end of the article

Sompairac *et al. BMC Bioinformatics* 2019, **20**(Suppl 4):140

Page 2 of 14

Background

There is still a gap in understanding the coordination between metabolic functions and signalling pathways in mammalian cells. Metabolic processes and cell signalling pathways contain a large number of molecular species together with their complex relationships. No single mind can accurately account for all these molecular interactions whilst drawing conclusions from a process of descriptive thought. To tackle the complexity of these multi-molecular interactions networks, a systems biology approach is needed. In addition, there is a high number of omics data such as transcriptome, proteome, metabolome, etc. accumulated for many human diseases as age-related disorders (e.g. neurodegeneration or cancer). Modelling and interpretation of these data combining metabolic and signalling networks together can help to decipher the mechanisms responsible for deregulations in human disorders by considering a broader range of molecular processes types.

Much of the produced high-throughput molecular data in many medical and biological applications remain under-explored due to the lack of insightful methods for data representation in the context of formally represented biological knowledge. Carefully designed maps of complex molecular mechanisms such as the whole-cell reconstructions of human metabolism in ReconMap 2.0 [1, 2] or the global reconstruction of cell signalling of cancer in ACSN [3] potentially provide ways to better exploit existing and new multi-omics data, by overlaying it on top of large molecular maps.

ACSN is a resource and a web-based environment that contains a collection of interconnected signalling network maps (https://acsn.curie.fr). Cell signalling mechanisms are depicted on the maps at the level of biochemical interactions, forming a large network of 4600 reactions covering 1821 proteins and 564 genes and connecting several major cellular processes [3]. ACSN is composed of 5 interconnected maps of major biological processes implicated in cancer. The maps are further divided into functional modules that represent signalling pathways collectively responsible for the execution of a particular process. In total, there are 52 functional modules in the ACSN resource (See Table 1 for terms definition). Each of these modules can be visualised in the context of the global ACSN map or accessed as individual maps. The Atlas is a "geographic-like" interactive "world map" of molecular interactions. ACSN is supported by NaviCell platform for easy map navigation and its annotations 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. Finally, NaviCell includes a powerful module for data visualisation. Users can map and visualise different types of "omics" data on the NaviCell maps [4, 5].

The manually curated genome-scale reconstruction Recon2.04 is a representation of the human metabolism. It accounts for 1733 enzyme-encoding genes associated to 7440 reactions which are distributed in 100 subsystems, referring to metabolic pathways. Furthermore, Recon2.04 accounts with 2626 unique metabolites distributed over eight cellular compartments [2]. Subsequently, to visualise the resource, a comprehensive metabolic map termed ReconMap 2.0 was generated from the Recon2.04 resource [1]. In the ReconMap 2.0 reactions (hyper-edges) were manually laid out using the biochemical network editor CellDesigner [6]. ReconMap 2.0 is currently distributed in a Systems Biology Graphical Notation (SBGN) compliant format and its content is also accessible via a web interface (https://vmh.uni.lu/#reconmap). All major human metabolic pathways are considered and represented as a seamless network where different pathways are interconnected via common molecules. There are 96 subsystems on the ReconMap 2.0, each of them representing a specific metabolic pathway (See Table 1 for terms definition).

By integrating these resources together, it will be possible to elucidate the crosstalk between metabolic and signalling networks. In addition, the integrated networks, provided in a common graphical language and available in standard exchange formats, makes them accessible for multiple systems biology tools. It opens an opportunity to model coordination between signalling pathways and metabolism using various systems biology approaches. Among others, there are several methods for multi-level omics data analysis in the context of the biological network maps that allow defining "hot" areas in molecular mechanisms and point to key regulators in physiological or in pathological situations [7–9] and beyond.

General workflow for integration of ACSN and ReconMap 2.0 networks

With the aim to integrate signalling and metabolic networks there is a need to find common players (proteins) that participate in the regulation of metabolic processes and simultaneously involved in signal transduction pathways. Thus, the networks can be interconnected via these common players. In addition, some solution for visualisation of proteins participating in the catalytic process in ReconMap 2.0 should be provided, since there is no such representation up to date.

The rationale behind the proposed methodology is to take an advantage of the CellDesigner SBML format for networks representation and develop a robust automated algorithm for an efficient finding of coordinates for new entities avoiding an overlap with existing elements and visualising these entities in the vicinity of the corresponding reactions they regulate. The integrated networks can

Sompairac *et al. BMC Bioinformatics* 2019, **20**(Suppl 4):140

Page 3 of 14

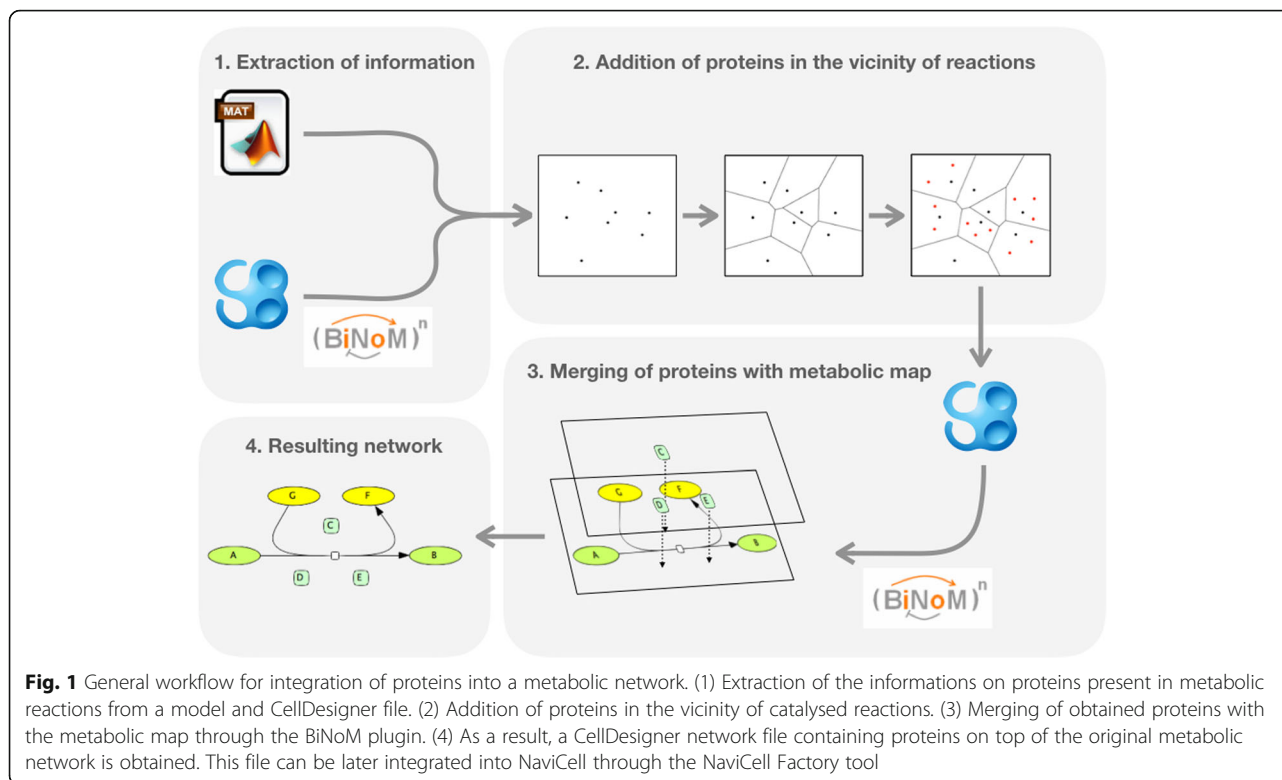**Table 1** Term definitions used in the paper

Graphical standards and exchange formats

XML format	Markup language that defines a set of rules for encoding documents used to store and transport data by describing the content in terms of what data is being described.
SBGN	Systems Biology Graphical Notation (SBGN) is a standard graphical syntax for representation of biological processes and interactions. SBGN is compatible with multiple pathway drawing and analytical tools, http://sbgn.github.io/sbgn/
SBML	Systems Biology Markup Language (SBML) is a representation format, based on XML, for communicating and storing computational models of biological processes. It is a free and open standard language with widespread software support, http://sbml.org
Standard identifier (ID)	Community-accepted nomenclature for scientific naming of biomolecules as genes, proteins, chemicals, drugs etc. The sources for standard IDs are repositories as UNIPROT, CHEB, HUGO, http://identifiers.org
Data and models exchange formats	Standard formats for data and models to facilitate networks and software intercompatibility. There are two major standard networks exchange formats, BIOPAX for complex networks and SIF for simple binary interactions. The CellDesigner xml format is a commonly-used exchange format compatible with multiple network analysis tools.

Signalling and metabolic network map

Map	Diagram of detailed molecular interactions with meaningful layout reflecting a certain biological process, which is graphically represented in CellDesigner tool.
Map module (in ACSN)	Part of the map representing a sequence of molecular interactions responsible for execution of a particular function.
Metabolic pathway Subsystem (in ReconMap 2.0)	Set of reactions forming a metabolic function. Set of metabolic reactions associated to (representing) a specific metabolic pathway.
Map node	Graphical representation of a molecule on the map.
Map entity and alias	Unique representation of a molecule on the map. As each molecule can be present multiple times on a map, each individual representation is called an alias of the entity. This definition corresponds to the CellDesigner feature.

Networks merging procedure

Voronoi cell	Individual shape allocated to a seed

Table 1 Term definitions used in the paper *(Continued)*

	from the Voronoi method. Each cell is a space that contains only the seed and can be used to generate points inside without overlapping with close seeds.
Voronoi tessellation	A partitioning of a plane into regions based on distance to points in a specific subset of the plane. That set of points (called seeds) is specified beforehand, and for each seed there is a corresponding region consisting of all points closer to that seed than to any other. These regions are called Voronoi cells. In our case, each seed is a molecule or a reaction's central glyph.
Centroid	Barycenter of a cluster.
Merging function of BiNoM	Function allowing taking two or more CellDesigner maps and merging them in one unique map. This function modifies each entities' id and alias but keeps the name, coordinates and notes.

NaviCell

Semantic zoom	A mechanism providing several map views with different levels of details depiction achieved by gradual exclusion of details while zooming out. It simplifies navigation through large maps of molecular interactions by providing several levels of details, resembling navigation through geographical maps. Exploring the map from a detailed toward a top-level view is achieved by gradual exclusion and modification (simplification and abstraction) of details. One of the main principles of semantic zooming is in that every detail which is shown on the map at a current zoom level, should be readable.
Marker	Symbol indicating location of chosen objects on the map; adapted from Google maps.
Pop-up bubble	Small window that opens by clicking on marker. Contains short description and hyperlinks related to the marked entity.
Annotation post	Detailed map entity annotation created in CellDesigner by map manager. The annotation is converted to Annotation post and displayed in the associated blog by NaviCell.

be provided as interconnected layers supported by NaviCell platform for navigation and data integration.

The suggested methodology is applied for ACSN and ReconMap 2.0 resources integration. However, this is a generic method applicable for integration of different types of networks prepared in CellDesigner SBML format (Fig. 1). In the following sections of the paper, we

Sompairac *et al. BMC Bioinformatics* 2019, **20**(Suppl 4):140

Page 4 of 14

Fig. 1 General workflow for integration of proteins into a metabolic network. (1) Extraction of the informations on proteins present in metabolic reactions from a model and CellDesigner file. (2) Addition of proteins in the vicinity of catalysed reactions. (3) Merging of obtained proteins with the metabolic map through the BiNoM plugin. (4) As a result, a CellDesigner network file containing proteins on top of the original metabolic network is obtained. This file can be later integrated into NaviCell through the NaviCell Factory tool

explain the challenges and describe how each step mentioned in the workflow was addressed.

The workflow in the Section 2 includes the following major steps (see Table 1 for terms definition):

- Identification of common proteins between ACSN and ReconMap 2.0 networks
- Finding metabolic and molecular processes crosstalk between ACSN and ReconMap 2.0
- Displaying protein nodes on the ReconMap 2.0 map
- ACSN-ReconMap 2.0 networks integration and visualisation using NaviCell

Materials and methods
Step-by-step procedure for network integration
Identification of common proteins between ACSN and ReconMap 2.0 networks

ACSN and ReconMap 2.0 maps contain information on proteins implicated in the regulation of reactions. First, the systematic use of the common identifiers as standard protein names (HUGO) for all proteins in both resources was verified and inconsistencies corrected. Thus, the proteins found in both resources ACSN and Recon-Map 2.0, were compared, quantified and visualised. We detected 252 proteins in common between the two networks (Additional file 1).

Displaying protein nodes on the ReconMap 2.0 map

ACSN and ReconMap 2.0 are both used as visual objects for exploration of processes as well as for data integration and visualisation in the context of the maps. After identification of the cross-talks between the two resources, it is important to ensure that all components of the maps are represented in a visual manner suitable for meaningful visualisation of omics data.

Due to the different nature of the networks, protein nodes are explicitly visualised on the ACSN map. However, in the ReconMap 2.0 the Standard Names (Identifiers) of proteins regulating metabolic reactions are included into the reaction annotations, but not represented visually on the map canvas. This precludes visualisation of omics data in the context of ReconMap 2.0 map. We developed a procedure for displaying the protein nodes on the ReconMap 2.0 map in the vicinity of the corresponding reaction edges, that now permits a multi-omics data visualisation in the context of both ACSN and ReconMap 2.0 layers.

Extraction of information regarding reactions and implicated genes in the metabolic network

- *Recuperation of the information from the Recon2.04 model*

    ReconMap 2.0 is the graphical representation of the literature-based genome-scale metabolic

Sompairac *et al. BMC Bioinformatics* 2019, **20**(Suppl 4):140

Page 5 of 14

reconstruction Recon2.04, which is freely available at (https://vmh.uni.lu/#downloadview). It is stored as a MatLab ".mat" file which contains a direct link between metabolic reactions and gene Entrez, specified by gene-rules. Therefore, it is possible to generate a direct protein-reaction association based on the gene codifying for the protein. As ACSN uses HUGO Standard Identifiers, Entrez IDs in ReconMap 2.0 were first converted to HUGO.

It is important to stress that this approach is based on a simplified assumption that if a protein is associated to a metabolic reaction in ReconMap 2.0, it may have a role in catalysis of the reactions. However, it is clear that the biological regulation is much more sophisticated than this basic assumption. For example, there are many protein complexes collectively regulating propagation of metabolic reaction and only part of them are actual enzymes that execute the catalysis, whereas others are co-factors of regulatory sub-units. Moreover, the activation states of proteins that is often regulated by post-translational modifications are also not taken into account in this simplified approach.

– *Recuperation of entities positions in ReconMap 2.0 from the XML network file*

In the graphical representation of reactions in CellDesigner, each reaction contains a central glyph in the form of a square. This glyph is normally used to allocate the position of the markers (see Table 1 for terms definition). However, its location is not explicitly saved in the network XML file. A specific function of NaviCell factory can calculate the coordinates of these glyphs and extract them in a separated file. These coordinates can be later used as a reference positions to assign protein nodes position in the ReconMap 2.0 map canvas.

Automated calculation of proteins coordinates in vicinity of corresponding reactions at ReconMap 2.0 network

– *Computing Voronoi cells for all elements*

By using the Voronoi method, each element of the network (molecules, reaction glyphs, etc.) is associated to a Voronoi cell. This method guarantees the lack of overlapping elements with already existing entities in the network when adding new proteins (Fig. 2).

– *Creation of randomly distributed points inside each reaction's Voronoi cell*

When each entity has a cell assigned, cells of reactions' central glyphs are utilised. Each cell has a certain number of points assigned randomly inside the cell. For our purpose, 100 points were deemed sufficient (Fig. 2).

– *Application of K-means algorithm to create K clusters*

Each reaction has a certain number of proteins implicated in its catalysis. Using the information from the model, the K-means algorithm was applied to identify the number of cluster centres corresponding to the number of protein nodes (Fig. 2).

– *Assigning protein positions using centroids coordinates of each cluster*

After the protein clusters are found, their centroids (see Table 1 for terms definition) are calculated and saved as the coordinated of the proteins tied to the specific reaction as catalysts (Fig. 2).

Conversion of obtained coordinates into a standard format (SBML)

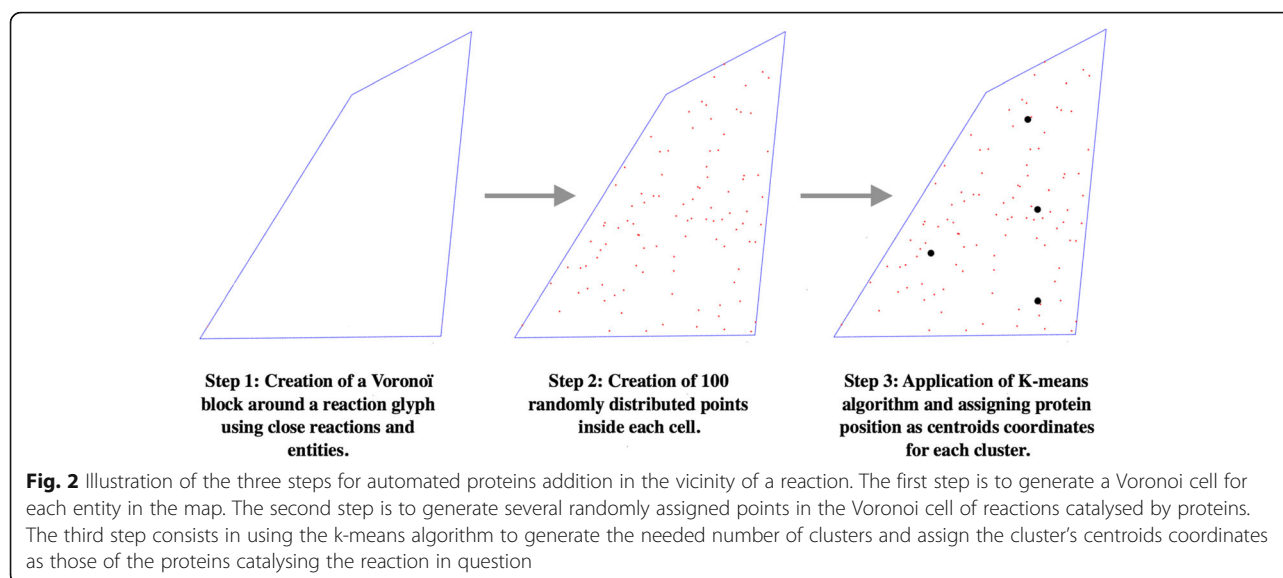– *Saving protein positions in a BiNoM Reaction Format*

Following the previous steps, a file in the BiNoM Reaction Format is obtained, containing the name of the proteins as well as their coordinates and sizes. This simple file will then be converted to a standard CellDesigner SBML format to be compatible with the original metabolic network. As CellDesigner allows the manipulation of "aliases" (multiple copies of the same entity); each protein with the same name present multiple times will have an apostrophe attached to its name based on the number of its repetition within the network.

– *Conversion of BiNoM Reaction Format into a CellDesigner map*

Using a custom python script, information stored in the BiNoM Reaction Format is transformed into a XML file following the SBML format. This file will contain each protein names, IDs, alias IDs, coordinates and type. As for now, only the manipulation of simple proteins is available.

– *Merging of the ReconMap 2.0 and Proteins maps using BiNoM merging function.*

Once the file containing proteins to add to the metabolic map is obtained, as they both are in the same SBML format, it is possible to merge them by using a function of the BiNoM plugin. This function allows transforming two or more separated maps into one unique map. This final merged map will be transformed into the NaviCell environment using the NaviCell Factory package (https://github.com/sysbio-curie/NaviCell).

Sompairac *et al. BMC Bioinformatics* 2019, **20**(Suppl 4):140

Page 6 of 14

Fig. 2 Illustration of the three steps for automated proteins addition in the vicinity of a reaction. The first step is to generate a Voronoi cell for each entity in the map. The second step is to generate several randomly assigned points in the Voronoi cell of reactions catalysed by proteins. The third step consists in using the k-means algorithm to generate the needed number of clusters and assign the cluster's centroids coordinates as those of the proteins catalysing the reaction in question

Thus, proteins implicated in the catalysis of a reaction can be seen in the vicinity of the corresponding reactions (Additional file 2, Fig. A). It is important to note that in some cases, reactions are regulated by many proteins, for example in the case of protein families, and the resulting configuration of protein nodes can be very dense (Additional file 2, Fig. B). This aspect can be improved by grouping protein families and visualising them together as a single generic entity. However, it is not always relevant to group all protein sharing a similar name by "family", since different family members might fulfil distinct or even opposite function, leading to a misinterpretation of the omics data in the context of the maps. Therefore, each protein was kept as a unique and independent entity.

Thanks to this method, 1.550 proteins were allocated in the ReconMap 2.0 canvas associated to more than 7.500 aliases. The algorithm for assigning proteins' coordinates is robust and its computation time is also scalable as the generation of the 7.500 allocation points is resolved in a matter of seconds.

ACSN and ReconMap 2.0 merging Once the protein positions file has been generated, it was converted to a CellDesigner [10, 11] XML format through a custom python script (https://github.com/sysbio-curie/CellDesigner_networks_map_integration_procedure). This script allows to obtain a file in XML format following the standard of CellDesigner's SBML. This 'map' contains only proteins in the positions they should belong on the final metabolic map. This file was then merged with the ReconMap 2.0 network by using an existing merging function of BiNoM [12, 13] to obtain the final network containing the original ReconMap 2.0 as well as the proteins in the vicinity of reactions they catalyse.

Tools, data source and code accessibility
Maps generation tool
CellDesigner [10, 11] is a tool used for the construction of both networks and its standard notation allowed the integration and linking across these maps. Both maps are available in a XML format, thus facilitating their automated manipulation.

Map entity annotation with NaviCell format
The annotation panel followed the NaviCell annotation format of each entity and reaction of the maps includes sections 'Identifiers', 'Maps_Modules', 'References' and 'Confidence' as detailed in [3]. 'Identifiers' section provides standard identifiers and links to the corresponding entity descriptions in HGNC, UniProt, Entrez, SBO, GeneCards and cross-references in REACTOME, KEGG, Wiki Pathways and other databases. 'Maps_Modules' section includes tags of modules in ACSN and metabolic pathways in RecoMap 2, in which the entity is implicated. 'References' section contains links to related publications. Each entity annotation is represented as a post with extended information on the entity.

Generation of NaviCell map with NaviCell factory
NaviCell Factory (https://github.com/sysbio-curie/NaviCell) is a package allowing to convert a CellDesigner map annotated in the NaviCell format into NaviCell Google Maps-based environment. This result in a set of HTML pages with integrated JavaScript code that can be launched in a web browser for online use. HUGO identifiers in the annotation form allow using NaviCell tool for visualisation of omics data [5].

The detailed guidelines for the usage of NaviCell factory, embedded in the BiNoM Cytoscape plugin, is

Sompairac *et al. BMC Bioinformatics* 2019, **20**(Suppl 4):140

Page 7 of 14

provided at https://navicell.curie.fr/doc/NaviCellMapper AdminGuide.pdf .

BiNoM
BiNoM (https://binom.curie.fr/) [12, 13] is a Cytoscape plugin, developed to facilitate the manipulation of biological networks represented in standard systems biology formats (SBML, SBGN, BioPAX) and to carry out studies on the network structure. BiNoM provides the user with a complete interface for the analysis of biological networks in Cytoscape environment.

Maps navigation via NaviCell platform
ACSN-ReconMap 2.0 interconnected maps are navigable in NaviCell online platform (https://navicell.curie.fr/). NaviCell uses Google Maps and semantic zooming to browse large biological network maps and allows shuttling between the two layers of interconnected resource ACSN-ReconMap 2.0 via common player (proteins).

Omics visualisation using NaviCell web service tool
NaviCell Web Service [5] is a tool for network-based visualisation of 'omics' that allows to overlay on maps different types of the molecular data. The tool provides standard heatmaps, barplots and glyphs modes of data display on the maps. In addition, the map staining technique allows to project on the map and grasp large-scale trends in numerical values (such as the whole transcriptome). The web service provides a server mode, which allows automating visualisation tasks and retrieving data from maps via RESTful (standard HTTP) calls.

Multi-omics data source
The transcriptome, copy number variations and mutation frequencies for ovarian cancer datasets were obtained from the TCGA repository that was originally generated in [14]. The definition of ovarian cancer sub-types are used accordingly to this study [14].

The data is also available at:

https://navicell.curie.fr/pages/nav_web_service.html
https://acsn.curie.fr/downloads.html
https://vmh.uni.lu/#downloadview

Maps accessibility
ReconMap 2.0 in NaviCell format interconnected to ACSN is available at: https://navicell.curie.fr/pages/maps_ReconMap%202.html

The ACSN resource is accessible via https://acsn.curie.fr

The ReconMap 2.0 is accessible via https://vmh.uni.lu/#reconmap

Code accessibility
The code and procedures used for the integration of both networks is accessible on GitHub (https://github.com/sysbio-curie/CellDesigner_networks_map_integration_procedure).
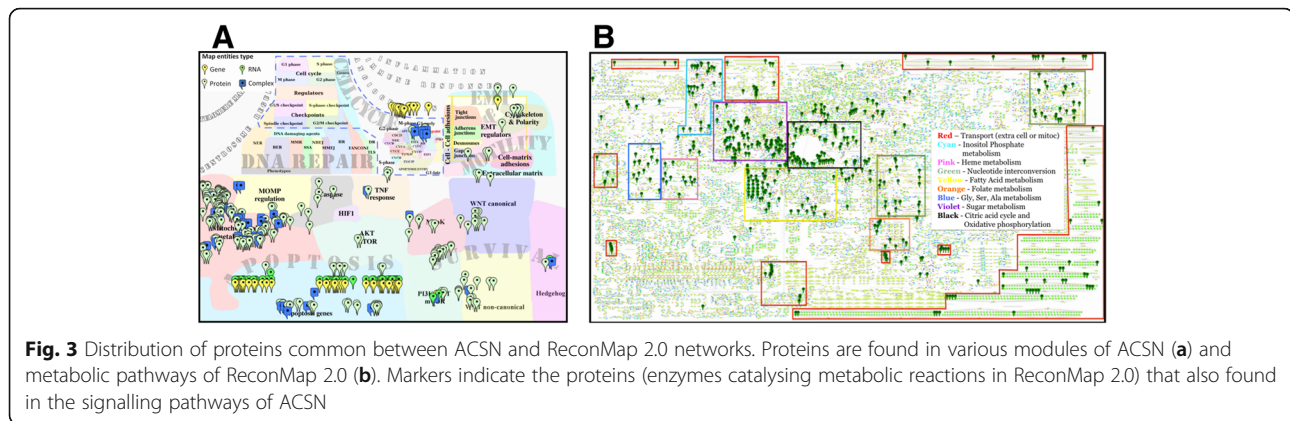
Results
Finding metabolic and molecular processes crosstalk between ACSN and ReconMap 2.0
ACSN and ReconMap 2.0 networks have a particular hierarchical structure. ACSN is divided into functional modules, whereas ReconMap 2.0 is divided into subsystems. Each of these structures are a subset of processes from the global network, implicated in regulation and execution of a specific molecular or metabolic pathway respectively (See Table 1 for terms definition). To address the question which metabolic processes are connected to which signalling mechanisms, the enrichment analysis of ACSN modules and Recon 2.0 subsystems was performed using the 252 common proteins (Additional file 1). The composition of ACSN modules and ReconMap 2.0 subsystems are provided as gene sets in Additional files 3 and 4 in Gene Matrix Transposed (GMT) file format and the enrichment was calculated using a hypergeometric test on the union of proteins of both maps (Additional file 5). The analysis demonstrated that proteins shared between the two maps are implicated in 22 modules of ACSN and in 51 subsystems of ReconMap 2.0 (Fig. 3 and Additional file 6).

Majority of proteins on both resources are participating in catalysis of biochemical or metabolic reactions. The information for the protein-reaction association is encoded in the network structure and in the annotations on the CellDesigner XML files. Information for reactions in each ACSN module and ReconMap 2.0 subsystem was retrieved and quantified. The number of reactions in ReconMap 2.0 subsystems regulated by proteins from ACSN modules is shown in Additional file 7.

Since ACSN is the resource focused on processes implicated in cancer; as expected, the ACSN modules containing mechanisms related to Mitochondria metabolism and Apoptosis have many shared proteins with ReconMap 2.0. Interestingly, ACSN modules related to cell survival processes as WNT-non-canonical pathway and modules related to the invasion and mobility processes as EMT regulators are also enriched by proteins regulating metabolic pathways of ReconMap 2.0. On the ReconMap 2.0, the shared proteins are enriched in energy-providing processes such as the citric acid cycle (Krebs cycle) and oxidative phosphorylation and in processes related to pentose phosphate pathway, fructose and mannose metabolism, glycolysis and gluconeogenesis. Finally, the subsystem inositol phosphate metabolism is enriched by the shared proteins (Fig. 3 and Additional file 5).
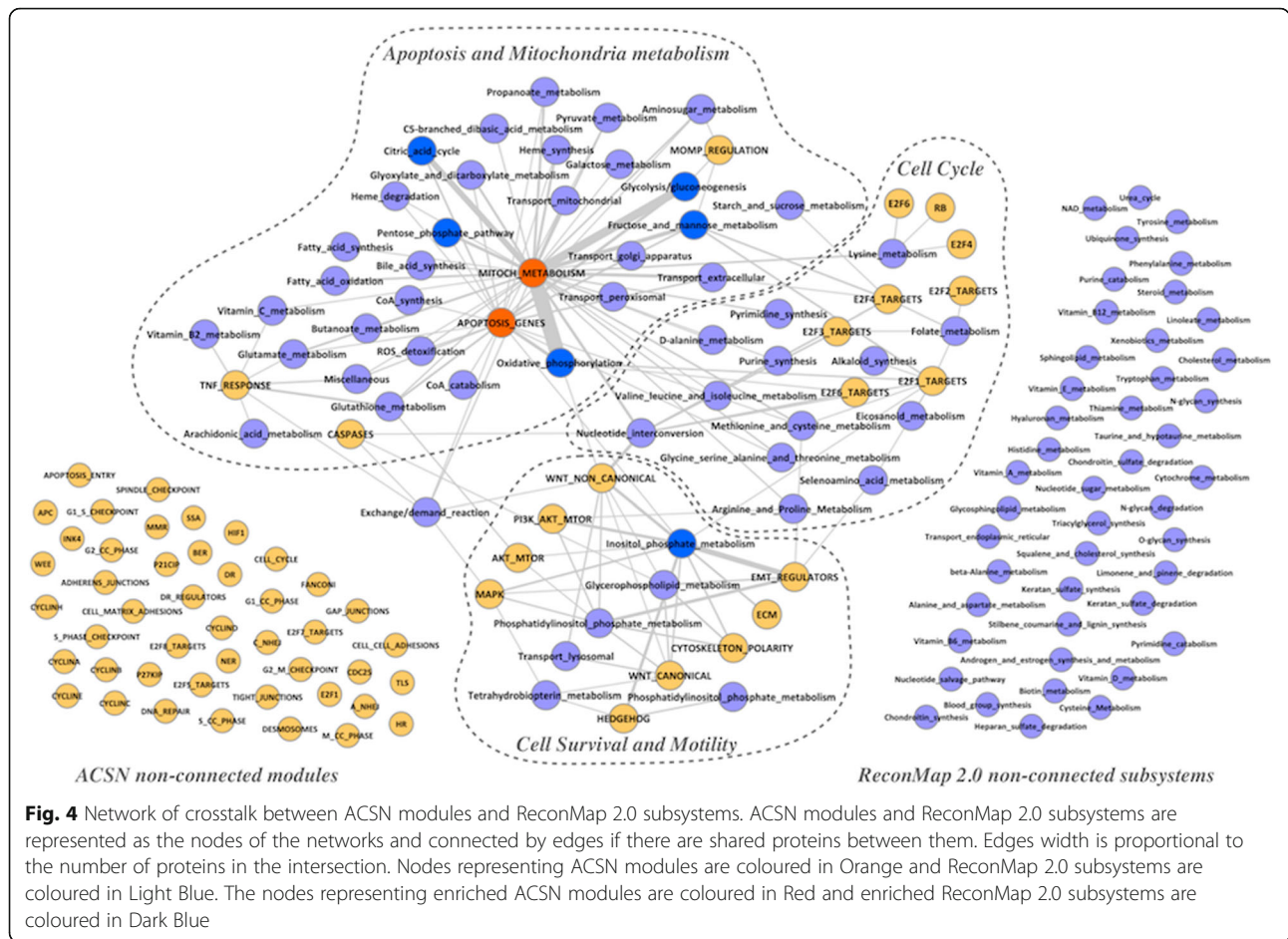
Sompairac *et al. BMC Bioinformatics* 2019, **20**(Suppl 4):140

Page 8 of 14

Fig. 3 Distribution of proteins common between ACSN and ReconMap 2.0 networks. Proteins are found in various modules of ACSN (**a**) and metabolic pathways of ReconMap 2.0 (**b**). Markers indicate the proteins (enzymes catalysing metabolic reactions in ReconMap 2.0) that also found in the signalling pathways of ACSN

The same trend is observed at the level of the reactions regulation, finding proteins from apoptosis, cell cycle and mitochondrial processes implicated in catalysis of a large number of reactions in the ReconMap 2.0 is expected as these pathways are known to be closely related to the metabolism. However, it is interesting to note that some less intuitive connections between ReconMap 2.0 subsystems and ACSN modules were retrieved. For example, proteins implicated in cell survival modules of ACSN as hedgehog, MAPK, PI3K-AKT-mTOR and WNT regulate reactions in the inositol phosphate metabolism subsystem from ReconMap 2.0. In addition, the proteins from cell migration-related and epithelial-to-mesenchymal transition (EMT)-related processes are involved in the regulation on the reactions in five different phospholipids and amino acid metabolic pathways, indicating most probably that the invasion process requires very active metabolism in migrating cancer cells. These connections are less obvious and may help to highlight interesting relations between signalling and metabolic processes in cancer (Additional file 7).

By extracting information about crosstalk between ACSN modules and ReconMap 2.0 subsystems, it was possible to generate a network where nodes represent ACSN modules and ReconMap 2.0 subsystems connected by edges if they shared common proteins (Fig. 4, Additional files 5 and 8). The obtained network contains one large connected component and also a number of modules and subsystems that are not connected to each other.

The interconnections between many signalling and metabolic processes on the network form communities. It was possible to identify three major communities each containing ACSN modules interconnected with ReconMap 2.0 subsystems, that we called 'Apoptosis and Mitochondrial Metabolism', 'Cell Cycle' and 'Cell Survival and Motility' according to their biological functions (Fig. 4). The nodes Mitochondria Metabolism and Apoptosis Genes, two modules of ACSN, are shown to be enriched with common

proteins and connected to 36 and 24 subsystems of Recon-Map 2.0 respectively. A total of 20 metabolic subsystems were found grouped forming a big community related to 'Apoptosis and Mitochondrial Metabolism' This result is not surprising since within the mitochondrion we find main cellular pathways such as citric acid cycle, oxidative phosphorylation and fatty acid oxidation. Furthermore, mitochondrion is a key organelle regulating cell death via two ways. On one hand, the lack of ATP which is mainly produced in the mitochondria via oxidative phosphorylation will lead to cell death through necrosis [15] and on the other hand, a variety of signalling apoptotic processes are linked to mitochondria [16], such as for example the family Bcl-2 which regulates apoptosis through mitochondrial permeability [15, 17]. The inositol phosphate metabolism subsystem is linked to 11 modules of ACSN, most of them being part of the 'Cell Survival and Motility' community. In accordance with this result, inositol phosphate functions act as second messengers for a variety of extracellular signals. Their effect in cell motility is due to the interaction of cellular membranes with proteins of the cytoskeleton [18]. Furthermore, members of the inositol phosphate metabolism pathway regulate the phosphatidylinositol-3-kinase (PI3K)/AKT signalling pathway, therefore cellular quiescence, proliferation, cancer, and longevity processes [19]. Interestingly enough, the subsystems responsible for nucleotides synthesis and metabolism related to the 'Cell Cycle' community are actually crosstalking with all three communities. Amino acids (aa) are not only used as a main energy source via oxidation and integration within the citric acid cycle in the mitochondrion; but also, they play a key role in several signalling pathways. Amino acids deprivation leads to cellular death through apoptosis and autophagy [20]. Moreover, their relation with cell motility has been also previously reported [21]. In addition, their association with cell proliferation seems to be correlated with the differential synthesis of proteins at different stages of the cell cycle [22, 23]. This observation demonstrates the central role of nucleotide metabolism in the major cell processes.

Sompairac *et al. BMC Bioinformatics* 2019, **20**(Suppl 4):140

Page 9 of 14

Fig. 4 Network of crosstalk between ACSN modules and ReconMap 2.0 subsystems. ACSN modules and ReconMap 2.0 subsystems are represented as the nodes of the networks and connected by edges if there are shared proteins between them. Edges width is proportional to the number of proteins in the intersection. Nodes representing ACSN modules are coloured in Orange and ReconMap 2.0 subsystems are coloured in Light Blue. The nodes representing enriched ACSN modules are coloured in Red and enriched ReconMap 2.0 subsystems are coloured in Dark Blue

The list of signalling modules and metabolic sub-systems that do not intersect in the current versions of both resources suggests performing literature mining with the aim to figure out whether a potential coordination between these processes is documented.

NaviCell representation of ACNS-ReconMap 2.0 resource

Due to their corresponding size and technical limitations, the two maps couldn't be merged into one single seamless map. The cross-linking via shared proteins was performed and the two maps were represented as interconnected layers using NaviCell web-based platform, allowing to shuttle between the maps by clicking on a common entity (see next paragraph).

Both maps were preserved with their original layout so that their correspondent relevance of the visual organisation was not lost. Furthermore, this allows users to have an easier view and understanding of the whole system. Moreover, entities annotations from ReconMap 2.0 have been recuperated from the original map and transferring to a NaviCell annotation format (see Materials and Methods), allowing to link entities to corresponding databases (Fig. 5).

Navigation of ACSN-ReconMap 2.0 resource using NaviCell platform

ACSN and ReconMap 2.0 maps were linked via the common players. Now that proteins had been displayed on top of the ReconMap 2.0 network canvas, those common with ACSN will be used as links to shuttle through both maps. Exploration and shuttling between the two map layers is possible using the NaviCell Google Maps-like features [5]. By clicking on a protein existing in both maps, a window with annotations will appear and a 'globe' icon will be clickable in the list seen in the 'Maps_and_Modules' section. A new window will be opened and the corresponding protein will be shown on this called map.

In addition, the exploration of the ACSN-ReconMap 2.0 resource is facilitated by the semantic zooming principle of NaviCell platform. As navigating large geographical maps, the semantic zooming on the molecular networks consists in hiding low-level invisible details at a less detailed level of zoom with simultaneous transforming and changing the scale of the representation of the essential objects by creating their abstractions.

This principle can be used for browsing large comprehensive maps of molecular mechanisms such as ACSN
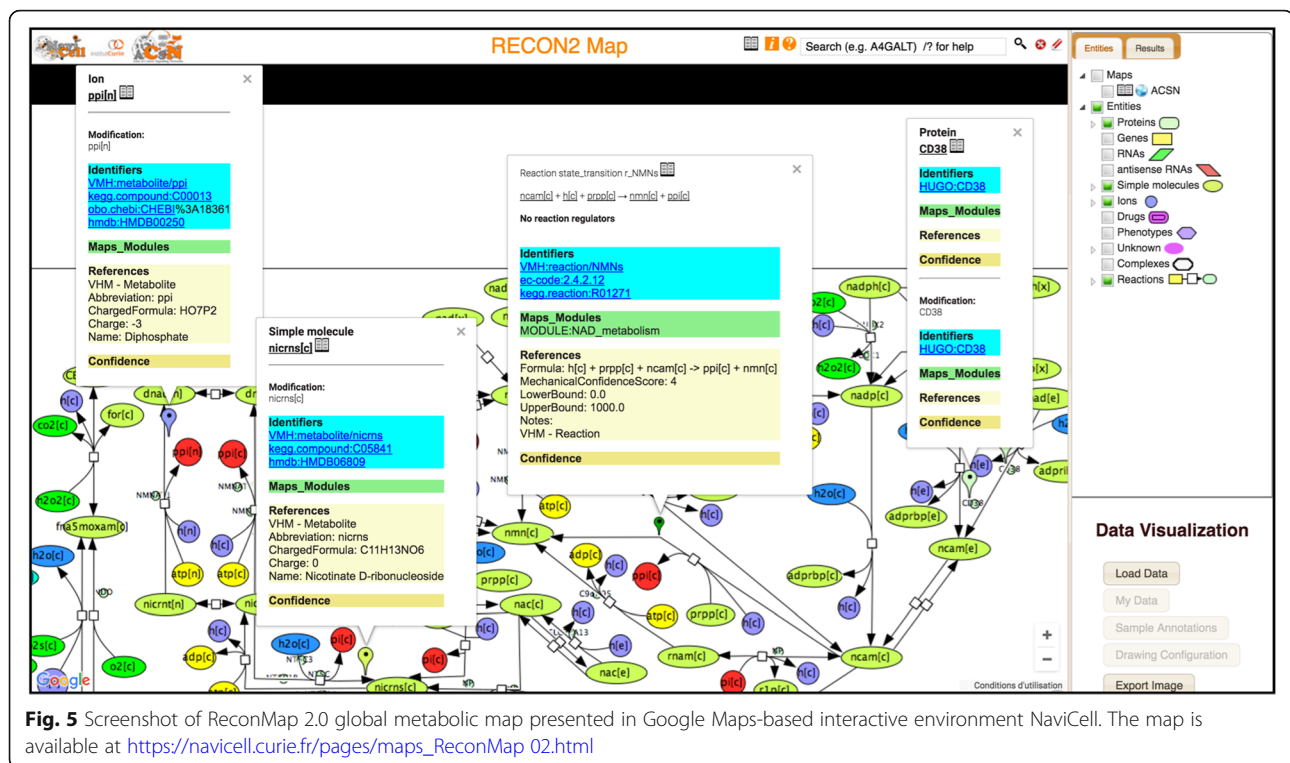
Sompairac *et al. BMC Bioinformatics* 2019, **20**(Suppl 4):140

Page 10 of 14

Fig. 5 Screenshot of ReconMap 2.0 global metabolic map presented in Google Maps-based interactive environment NaviCell. The map is available at https://navicell.curie.fr/pages/maps_ReconMap 02.html

[3] and ReconMap 2.0 [1] thanks to the existing open code of Google Maps API.

Visualisation of cancer multi-omics data in the context of integrated ACSN-ReconMap 2.0 resource

The interconnected ACSN-ReconMap 2.0 resource was applied for visualisation of multi-omic data representing ovarian cancer subtypes. The transcriptomic, copy number and mutation data from ovarian cancer study [5] from TCGA resource was used for the visualisation in the context of the ACSN-ReconMap 2.0 resource using Navicell Web Service toolbox [5], generating molecular portraits of immunoreactive and proliferative ovarian cancer subtype. Here below we demonstrate several examples from the molecular portraits and discuss the possible biological significance.

First, we demonstrate that ReconMap 2.0 with displayed protein nodes on the map canvas and provided in NaviCell platform, is now applicable for meaningful multi-omic data visualisation.

As shown in Fig. 6, the expression of keratan metabolism pathway regulators is high in the immunoreactive subtype, where most of the pathway-related genes are mutated and undergo more frequently copy number alterations (Fig. 6a). The opposite picture is seen for the Proliferative subtype of ovarian cancer, with mostly under-expressed genes being mutated (Fig. 6b).

It is known that the Nucleotide metabolism plays an important role in cancer development. The molecular portrait of this metabolic pathway is very different comparing to the keratan metabolism pathway. The regulators of the nucleotide transport pathways are under-expressed in the Immunoreactive subtype (Additional file 9, Fig. A) comparing to the in proliferative subtype (Additional file 9, Fig. B) of ovarian cancer. This observation is in coherence with the fact that highly proliferative cells would consume higher amounts of nucleotides for their growth.

Further, for more interpretable visualisation of cross-talks between both maps, only the 252 genes common between ACSN and ReconMap 2.0 were used for visualisations of data.

We performed the analysis of the regulation of 252 proteins shared between ReconMap 2.0 and ACSN and retrieved their corresponding implications in the function modules for both maps, comparing two ovarian cancer subtypes as in the previous example. Production of energy is a crucial mechanism necessary for the development of cancer cells, therefore it is not surprising to find significant changes in the regulation of the Energy metabolism module between two ovarian cancer subtypes, especially profound in the Krebs cycle, glycolysis and gluconeogenesis mechanisms (Additional file 10). In the immunoreactive subtype, genes implicated in the Krebs cycle are over-expressed while those involved in the glucose metabolism are under-expressed (Additional file 10, Fig. A). The opposite is found in proliferative cells with over-expressed genes in the glucose metabolism being mutated (Additional file 10, Fig. B).
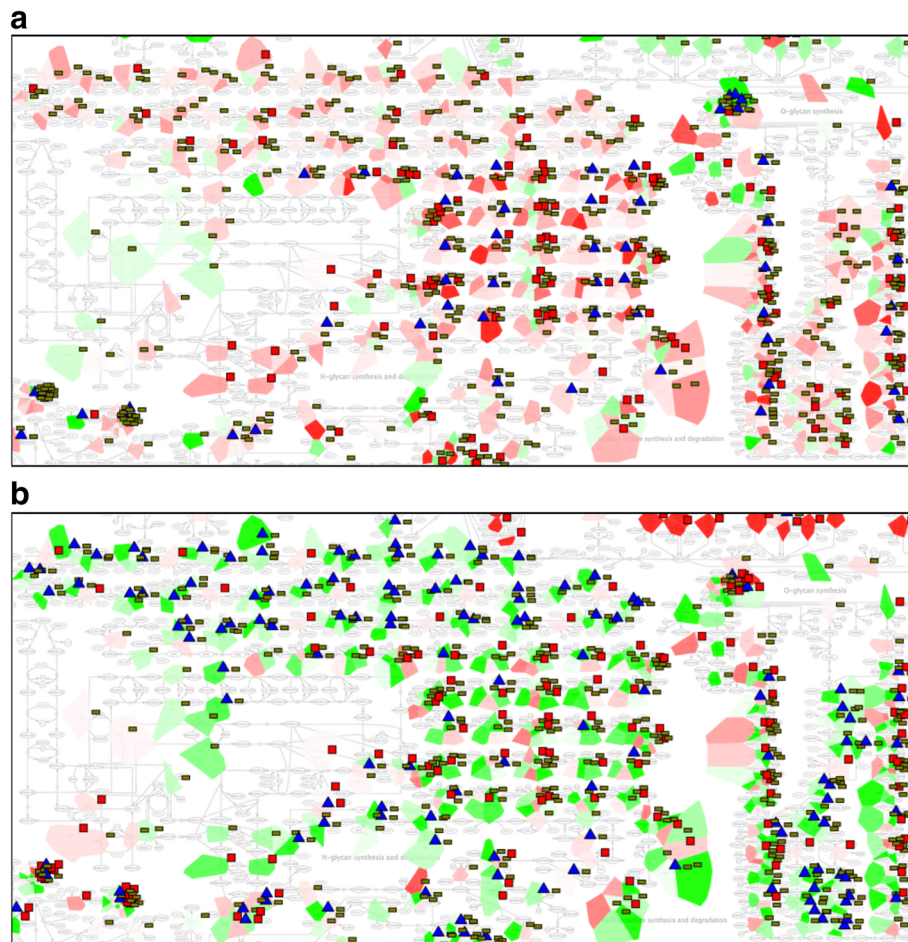
Sompairac *et al. BMC Bioinformatics* 2019, **20**(Suppl 4):140

Page 11 of 14

Fig. 6 Ovarian cancer multi-omics data visualisation on ReconMap 2.0: zoomed on keratan sulfate synthesis and degradation metabolic pathway. Two ovarian cancer subtypes are compared: Immunoreactive (**a**), Proliferative (**b**). Patches using the map staining function represent the average expression level (underexpressed in green and over-expressed in red). Barplots indicate the copy number state (red means at least 2 copy number). Glyphs shown as blue triangles are viewed near genes possessing mutations

The genes implicated in the inositol phosphate metabolism also show opposite regulation patterns between the two studied groups (Additional file 11). This process is known to be dysregulated in cancer and has an impact on cell proliferation and migration [18]. Interpretation of the results in the context of a map can shade light on mechanisms governing these perturbations.

The same type of visualisation has been performed on the ACSN map, demonstrating three deregulated modules: mitochondrial metabolism (Additional file 12), WNT canonical (Additional file 13) and WNT non-canonical (Additional file 14). These 3 modules are indeed the most enriched in common genes between signalling and metabolic pathways [24].

Interestingly, in WNT canonical and non-canonical modules, the whole cascades seem to be implicated and differ greatly between the two ovarian cancer subtypes. Genes participating in these pathways have been found

participating also in the inositol phosphate metabolism (Figs. 3 and 4 and Additional file 5). This new information is especially valuable because these pathways were not yet shown as related to metabolic processes.

In addition, the data visualisation demonstrated that TNF response factors in ACSN are linked to the vitamin B2 metabolism in ReconMap 2.0 map and show differential regulation between the two subtypes of ovarian cancer (Figs. 3 and 4 and Additional file 5).

Discussion and conclusions

A systems biology approach involving integration of signalling and metabolic networks permits characterisation of cross-links between the two types of molecular mechanisms in different conditions (e.g. healthy and disease). The integrated ACSN-ReconMap 2.0 resource provided under the NaviCell platform opens an opportunity for a full exploitation of multi-omics data using visualisation

Sompairac *et al. BMC Bioinformatics* 2019, **20**(Suppl 4):140

Page 12 of 14

features of NaviCell [4]. NaviCell allows to visualise and analyse data based on hierarchical structure of ACSN modules and ReconMap 2.0 subsystems respectively, evaluating 'activity' of modules and subsystems thanks to the map staining techniques [5].

ACSN and ReconMap resources are constantly extended with new discoveries in the corresponding fields. Future modifications of these networks will be maintained and the workflow described in this manuscript will be reused to allow updates of the integrated ACSN-ReconMap 2.0 resource.

The developed networks integration methodology and the suggested workflow is a generic mechanism and can be easily applied for integration of other comprehensive maps. The robustness of the method, computational speed and memory usage allows it to be used on any computer with Python and Java installed. Scripts used are open-source and accessible on GitHub (https://github.com/sysbio-curie/CellDesigner_networks_map_integration_procedure).

In this manuscript, we have shown that the merging of metabolic and signalling networks can be achieved and it provides many possibilities for data analyses and comprehension of implicated processes across both maps. In addition, the integrated resource allows to find gaps in connectively between signalling and metabolic processes and suggests exploration of potential links. The integrated ACSN-ReconMap resource will help in further elucidating the crosstalk between of the metabolic and signalling processes and understanding what are the key coordination players in cancer and other human disease.

We will further develop the ACSN-ReconMap resource and integrate into an open software platform together with tools as ROMA [25], COBRA [26], etc. for multi-scale data analysis at morphological, subsystem, reaction and atomic scales. The platform will allow metabolic networks modelling under the regulation of signalling processes with the aim to predict disease status and beyond.

Additional files

Additional file 1: List of common proteins. List of the 252 proteins found in common between ACSN and ReconMap 2.0 maps (available at https://navicell.curie.fr/pages/maps_ReconMap%202.html). (TXT 1 kb)

Additional file 2: Protein nodes display on the ReconMap 2.0 canvas. (A) Dispersed distribution of protein nodes around reactions, (B) Case of dense protein nodes distribution around a reaction, implicating 46 proteins. Protein nodes are represented by pale green circles. (TIFF 357 kb)

Additional file 3: ACSN GMT. Gene sets composing ACSN modules (available at https://navicell.curie.fr/pages/maps_ReconMap%202.html). (GMT 31 kb)

Additional file 4: ReconMap2 GMT. Gene sets composing ReconMap 2.0 subsystems (available at https://navicell.curie.fr/pages/maps_ReconMap%202.html). (GMT 21 kb)

Additional file 5: Crosstalk between ACSN modules and ReconMap 2.0 metabolic pathways. Using the composition of each module and

subsystem (Additional files 3 and 4) and the list of the 252 shared proteins between both maps, modules and subsystems having common proteins were linked together. Thus, a crosstalk between both maps has been uncovered with the number of shared proteins showed here. *Modules and subsystems in bold are the result of an Enrichment Analysis of the 252 common proteins between ACSN and ReconMap 2. (DOCX 16 kb)

Additional file 6: Enrichment analysis of common proteins between ACSN modules and ReconMap 2.0 metabolic pathways. Enrichment Analysis of the 252 common proteins between ACSN and ReconMap 2.0 was performed using the Survival Function Method on the union of proteins of both maps. The *p*-value is obtained through a standard Hypergeometric test on modules containing at least 10 proteins using a Bonferroni adjustment for the multiple testing correction. (DOCX 14 kb)

Additional file 7: Regulations of reactions in ReconMap 2.0 subsystems by proteins found in ACSN modules. The information encoded in the annotations of reactions and entities on the CellDesigner XML files was analysed and the correspondence between shared proteins in ACSN modules to regulated reactions in ReconMap 2.0 subsystems was retrieved and quantified. Number of ReconMap 2.0 reaction regulated by a subset of shared proteins is shown. (DOCX 17 kb)

Additional file 8: Crosstalks_network. Network of crosstalks based on shared proteins between modules of ACSN and subsystems of ReconMap 2.0 in TXT format. This file contains Source nodes, Target nodes, Interaction type and the number of intersection proteins corresponding to each edge. The file was used for visualisation of the crosstalks in Fig. 4 (available at https://navicell.curie.fr/pages/maps_ReconMap%202.html). (TXT 9 kb)

Additional file 9: Ovarian cancer multi-omics data visualisation on ReconMap 2.0: zoomed on nucleotide transport metabolic pathway. Two ovarian cancer subtypes are compared: Immunoreactive (top), Proliferative (bottom) Patches using the map staining function represent the average expression level of the corresponding genes (under-expressed in green and over-expressed in red). Barplots indicate the copy number state (red means at least 2 copy number). Glyphs shown as blue triangles are viewed near genes possessing mutations. (TIFF 3250 kb)

Additional file 10: Ovarian cancer multi-omics data visualisation on ReconMap 2.0: zoomed on energy metabolism pathways. Two subtypes are present: Immunoreactive (top), Proliferative (bottom). Patches using the map staining function represent the average expression level of the corresponding genes (under-expressed in green and over-expressed in red). Barplots indicate the copy number state (red means at least 2 copy number). Glyphs shown as blue triangles are viewed near genes possessing mutations. (TIFF 3412 kb)

Additional file 11: Ovarian cancer multi-omics data visualisation on ReconMap 2.0: zoomed on inositol phosphate metabolic pathway. Immunoreactive (A) and Proliferative (B) subtypes are compared. Patches using the map staining function represent the average expression level of the corresponding genes (under-expressed in green and over-expressed in red). Barplots indicate the copy number state (red means at least 2 copy number). Glyphs shown as blue triangles are viewed near genes possessing mutations. (TIFF 1338 kb)

Additional file 12: Ovarian cancer multi-omics data visualisation on ACSN: zoomed on mitochondrial metabolism signalling pathway. Immunoreactive (A) and Proliferative (B) subtypes are compared. Patches using the map staining function represent the average expression level of the corresponding genes (under-expressed in green and over-expressed in red). Barplots indicate the copy number state (red means at least 2 copy number). Glyphs shown as blue triangles are viewed near genes possessing mutations. (TIFF 935 kb)

Additional file 13: Ovarian cancer multi-omics data visualisation on ACSN: zoomed on WNT canonical signalling pathways. Immunoreactive (A) and Proliferative (B) subtypes are compared. Patches using the map staining function represent average the expression level of the corresponding genes (under-expressed in green and over-expressed in red). Barplots indicate the copy number state (red means at least 2 copy number). Glyphs shown as blue triangles are viewed near genes possessing mutations. (TIFF 763 kb)

Sompairac *et al. BMC Bioinformatics* 2019, **20**(Suppl 4):140

Page 13 of 14

Abbreviations
ACSN: Atlas of Cancer Signalling Network; ATP: Adenosine triphosphate; GMT: Gene Matrix Transposed file format; HGNC: HUGO Gene Nomenclature Committee; HTML: HyperText Markup Language; HTTP: HyperText Transfer Protocol; HUGO: Human Genome Organization; SBGN: Systems Biology Graphical Notation; SBML: Systems Biology Markup Language; TCGA: The Cancer Genome Atlas; XML: eXtensible Markup Language

Availability of data and materials
The open sources of multi-omics data used in this article are indicated.

About this supplement
This article has been published as part of *BMC Bioinformatics Volume 20 Supplement 4, 2019: Methods, tools and platforms for Personalized Medicine in the Big Data Era (NETTAB 2017).* The full contents of the supplement are available online at https://bmcbioinformatics.biomedcentral.com/articles/supplements/volume-20-supplement-4.

Authors' contributions
NS developed the method for integration of signalling and metabolic networks and wrote the paper; JM participated in the case studies generation and interpretation of the results; EB advised during the project and revised the paper; RMTF supervised the work performed in LCSB, advised during the project and critically revised and restructured the paper; AZ supervised the method development in IC, advised during the project and revised the paper; IK led the project and wrote the paper. Each author have read and approved the final manuscript.

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
[1]Institut Curie, 26 rue d'Ulm, F-75005 Paris, France. [2]Inserm, U900, F-75005 Paris, France. [3]Mines Paris Tech, F-77305 Fontainebleau cedex, France. [4]PSL Research University, F-75005 Paris, France. [5]Centre for Systems Biomedicine, University of Luxembourg, L-4367 Belvaux, Luxembourg.

References
1. Noronha A, Daníelsdóttir AD, Gawron P, Jóhannsson F, Jónsdóttir S, Jarlsson S, et al. ReconMap: an interactive visualization of human metabolism. Bioinformatics. 2017;33(4):605–7.
2. Thiele I, Swainston N, Fleming RMT, Hoppe A, Sahoo S, Aurich MK, et al. O. A community-driven global reconstruction of human metabolism. Nat Biotechnol. 2013;31(5):419–25.
3. Kuperstein I, Bonnet E, Nguyen HA, Cohen D, Viara E, Grieco L, et al. Atlas of Cancer Signalling network: a systems biology resource for integrative analysis of cancer data with Google maps. Oncogenesis. 2015;4.
4. Kuperstein I, Cohen DPA, 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;7.
5. Bonnet E, Viara E, Kuperstein I, Calzone L, Cohen DPA, Barillot E, Zinovyev A. NaviCell web service for network-based data visualization. Nucleic Acids Res. 2015;43(W1):W560–5.
6. Kitano H, Funahashi A, Matsuoka Y, Oda K. Using process diagrams for the graphical representation of biological networks. Nat Biotechnol. 2005.
7. Stoll G, Surdez D, Tirode F, Laud K, Barillot E, Zinovyev A, Delattre O. Systems biology of Ewing sarcoma: a network model of EWS-FLI1 effect on proliferation and apoptosis. Nucleic Acids Res. 2013;41(19):8853–71.
8. Biton A, Bernard-Pierrot I, Lou Y, Krucker C, Chapeaublanc E, Rubio-Pérez C, et al. Independent component analysis uncovers the landscape of the bladder tumor transcriptome and reveals insights into luminal and basal subtypes. Cell Rep. 2014;9(4):1235–45.
9. Jdey W, Thierry S, Russo C, Devun F, Al Abo M, Noguiez-Hellin P, et al. Drug-driven synthetic lethality: bypassing tumor cell genetics with a combination of AsiDNA and PARP inhibitors. Clin Cancer Res. 2017;23(4):1001–11.
10. Funahashi A, Morohashi M, Kitano H, Tanimura N. CellDesigner: a process diagram editor for gene-regulatory and biochemical networks. BIOSILICO. 2003;1(5):159–62.
11. Funahashi A, Matsuoka Y, Jouraku A, Morohashi M, Kikuchi N, Kitano H. CellDesigner 3.5: a versatile modeling tool for biochemical networks. Proc IEEE. 2008;96(8):1254–65.
12. Zinovyev A, Viara E, Calzone L, Barillot E. BiNoM: a Cytoscape plugin for manipulating and analyzing biological networks. Bioinformatics. 2008;24(6):876–7.
13. Bonnet E, Calzone L, Rovera D, Stoll G, Barillot E, Zinovyev A. BiNoM 2.0, a Cytoscape plugin for accessing and analyzing pathways using standard systems biology formats. BMC Syst Biol. 2013;7.
14. Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. Nature. 2011;474(7353):609–15.
15. Sassone J, Maraschi A, Sassone F, Silani V, Ciammola A. Defining the role of the Bcl-2 family proteins in Huntington's disease. Cell Death Dis. 2013;4(8):e772.
16. Green, John C, Reed DR, Green DR, Reed JC, Kluck RM, Bossy-wetzel E, Green DR, Newmeyer DD. Mitochondria and apoptosis. Science. 1998;281(5381):1309–12.
17. Wang C, Youle RJ. The role of mitochondria in apoptosis. Annu Rev Genet. 2009;43:95–118.
18. Tkachuk VA. Phosphoinositide metabolism and Ca2+ oscillation. Biochemistry. 1998;63(1):38–46.
19. Tan J, Yu C-Y, Wang Z-H, Chen H-Y, Guan J, Chen Y-X, Fang J-Y. Genetic variants in the inositol phosphate metabolism pathway and risk of different types of cancer. Sci Rep. 2015;5(1):8473.
20. Martinet W, De Meyer GRY, Herman AG, Kockx MM. Amino acid deprivation induces both apoptosis and autophagy in murine C2C12 muscle cells. Biotechnol Lett. 2005;27(16):1157–63.
21. Fu YM, Yu ZX, Lin H, Fu X, Meadows GG. Selective amino acid restriction differentially affects the motility and directionality of DU145 and PC3 prostate cancer cells. J Cell Physiol. 2008;217(1):184–93.
22. Bhargava PM, Allin EP, Montagnier L. Uptake of amino acids and thymidine during the first cell cycle of synchronized hamster cells. J Membrane Biol. 1976;26:1–17.
23. Meijer AJ, Dubbelhuis PF. Amino acid signalling and the integration of metabolism. Biochem Biophys Res Commun. 2004;313:397–403.
24. Pate KT, Stringari C, Sprowl-Tanio S, Wang K, TeSlaa T, Hoverter NP, et al. Wnt signaling directs a metabolic program of glycolysis and angiogenesis in colon cancer. EMBO J. 2014.

Sompairac *et al. BMC Bioinformatics* 2019, **20**(Suppl 4):140

Page 14 of 14

25. Martignetti L, Calzone L, Bonnet E, Barillot E, Zinovyev A. ROMA: Representation and quantification of module activity from target expression data. Front Genet 2016;7:18.
26. Schellenberger J, Que R, Fleming RMT, Thiele I, Orth JD, Feist AM, et al. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA toolbox v2.0. Nat Protoc. 2011;6(9):1290–307.