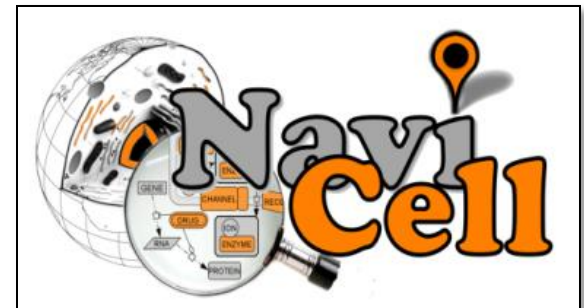
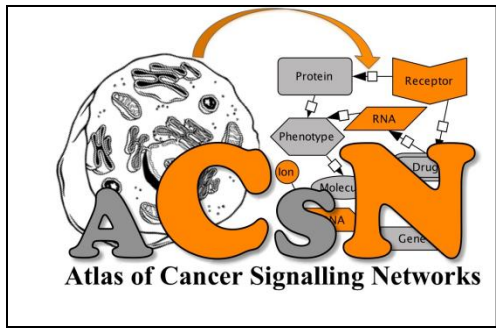

Atlas of Cancer Signaling Network and NaviCell



Computational Systems Biology of Cancer

U900 **Institut Curie**/INSERM/Ecole des Mines Paristech
Paris, France

Computational Systems Biology of Cancer

group at Institut Curie <http://sysbio.curie.fr>

Directions:

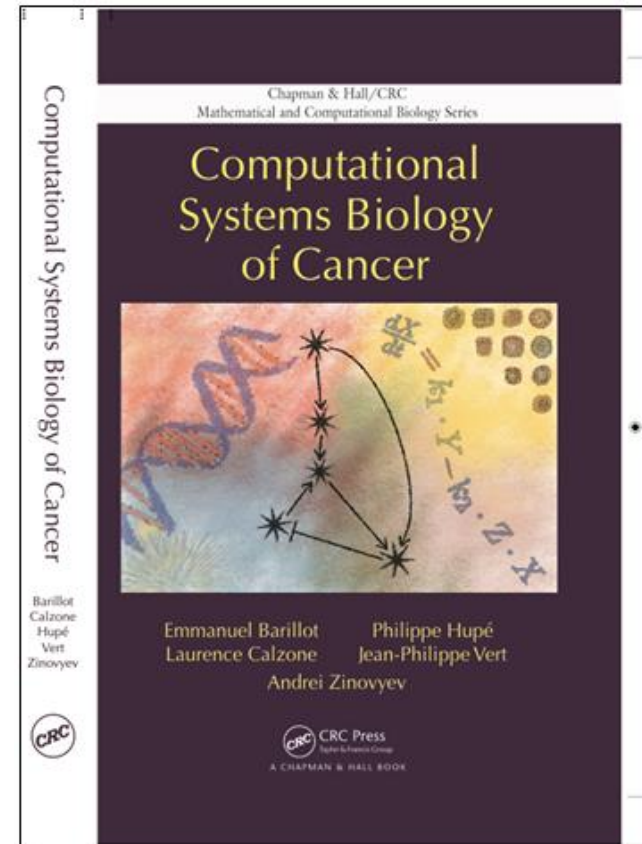
- 1) Omics data analysis using **biological networks**
- 2) Statistical **analysis of multi-level omics data**
- 3) **Mathematical modeling** of networks involved in tumor growth, interaction with microenvironment, metastases
- 4) Methods and **software development** for systems biology

Dynamic international and multidisciplinary environment

Numerous collaborative projects on concrete questions cancer biology and cancer treatment

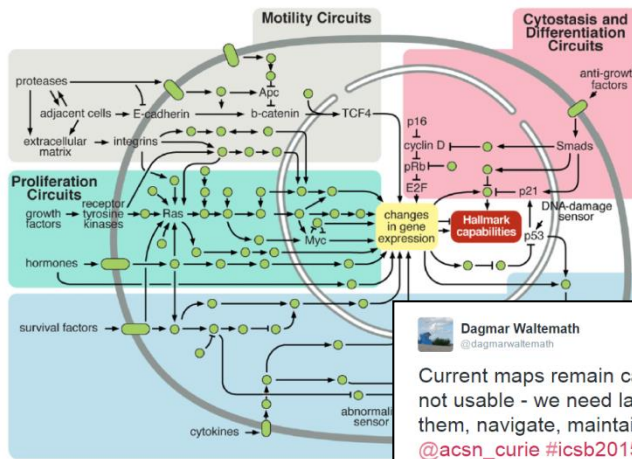
Access to original large-scale data from application of latest technologies

We are HIRING postdocs!



Atlas of cancer signaling network: project inspired by the authors of hallmarks of cancer

Hanahan and Weinberg, 2001



institut_curie
@institut_curie

Following

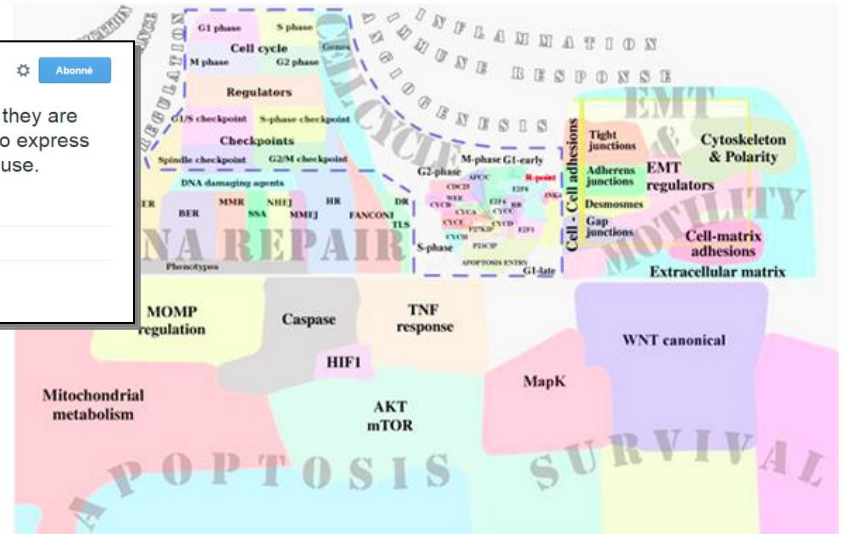
Researchers from #InstitutCurie launch the first @googlemaps of #Cancer Signalling ►
goo.gl/GT1JZL

Dagmar Waltemath
@dagmarwaltemath

Current maps remain cartoons - they are not usable - we need language to express them, navigate, maintain, and reuse.
[@acsn_curie](#) #icsb2015

RETWEETS 5 1

05:22 - 23 nov 2015



Oncogenesis

Journal home > Archive > July 20 2015 > Full text

- Journal home
- Latest articles
- Archive
- Browse by subject
- News

Original Article

Citation: *Oncogenesis* (2015) 4, e160; doi:10.1038/oncsis.2015.011
Published online 20 July 2015

Atlas of Cancer Signalling Network biology resource for integrative analysis of cancer data with **Google Maps**

OPEN

I Kuperstein^{1,2,3}, E Bonnet^{1,2,3}, H-A Nguyen^{1,2,3}, D Cohen^{1,2,3}, E Viara⁴, L Grieco^{1,2,3,5,6,7}, S Fourquet^{1,2,3}, L Calzone^{1,2,3}, C Russo^{1,2,3}, M Kondratova^{1,2,3}, M Dutreix^{1,8,9}, E Barillot^{1,2,3} and A Zinovyev^{1,2,3}

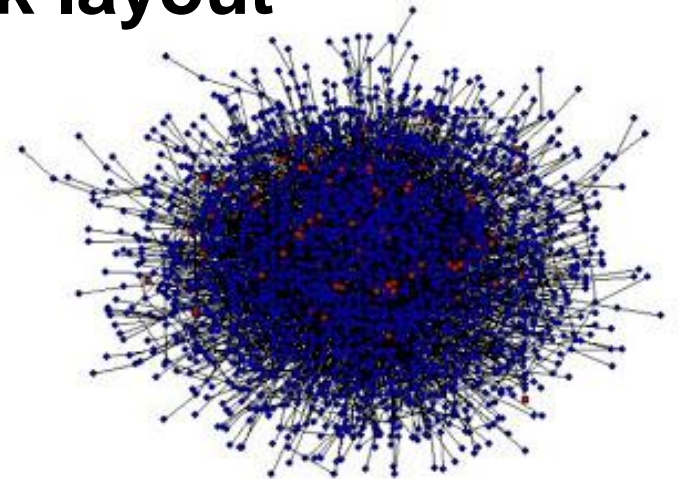
¹Institut Curie, Paris, France

RETWEETS 16 FAVOURITES 13



Biological network and Biological network map

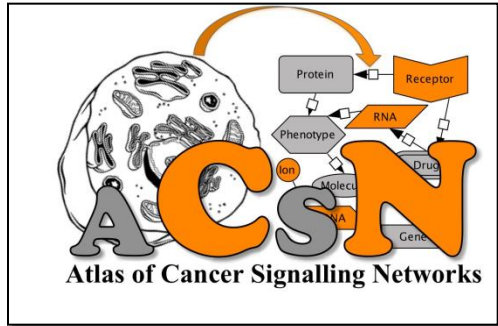
- **Network** – set of connections between biological entities
- **Network map** – graphical representation of the network, layout + decorations
- Problem of meaningful **network layout**



Atlas of Cancer Signaling Network and NaviCell

<http://acsn.curie.fr> <http://navicell.curie.fr>

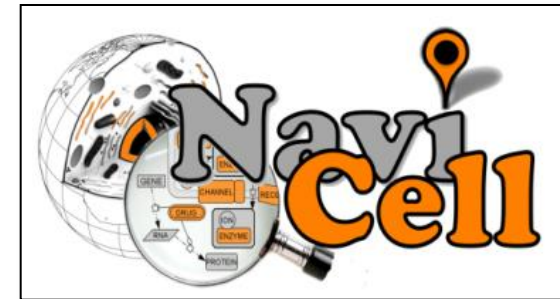
- **ACSN** – “world map” of cancer biology



the largest map
represented in
CellDesigner format

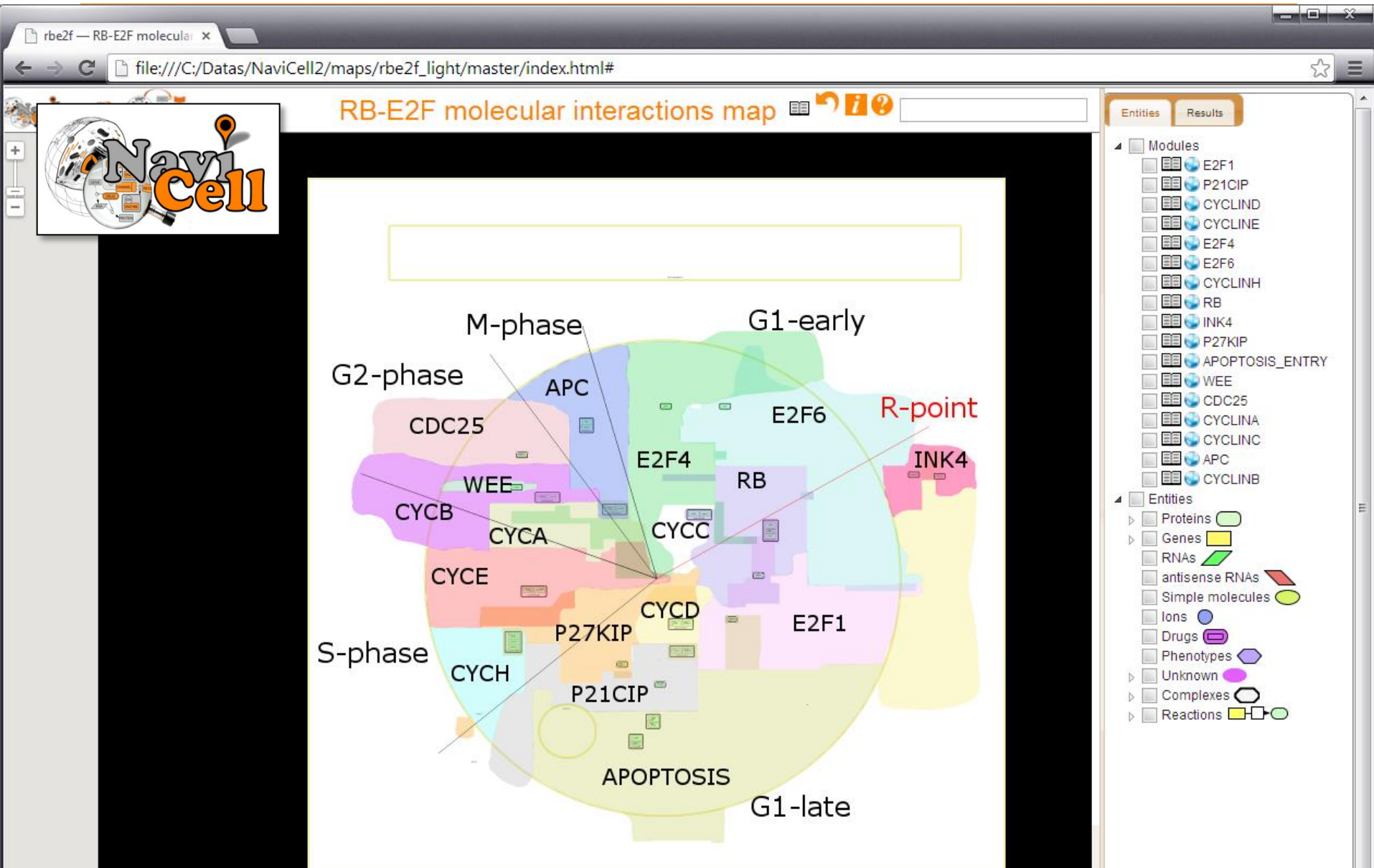
(4826 reactions, 2371 proteins,
5975 chemical species)

- **NaviCell** – Google Maps-based engine to browse VERY LARGE biological network maps and visualizing data on top of them



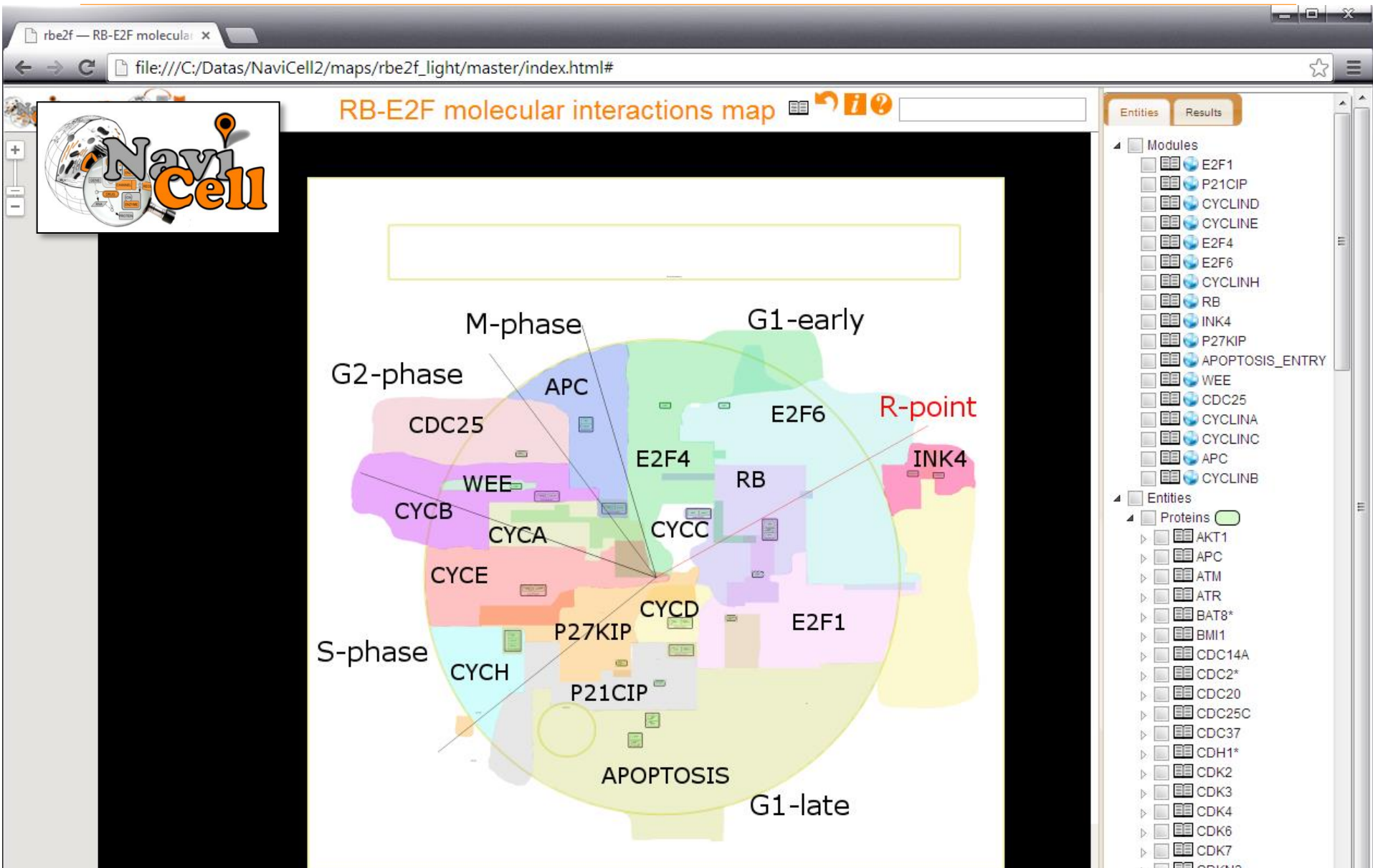
Example: Google maps of cell cycle

(Calzone et al, MSB, 2008; Kuperstein et al, BMC Sys Bio, 2013; Bonnet et al, NAR, 2015)



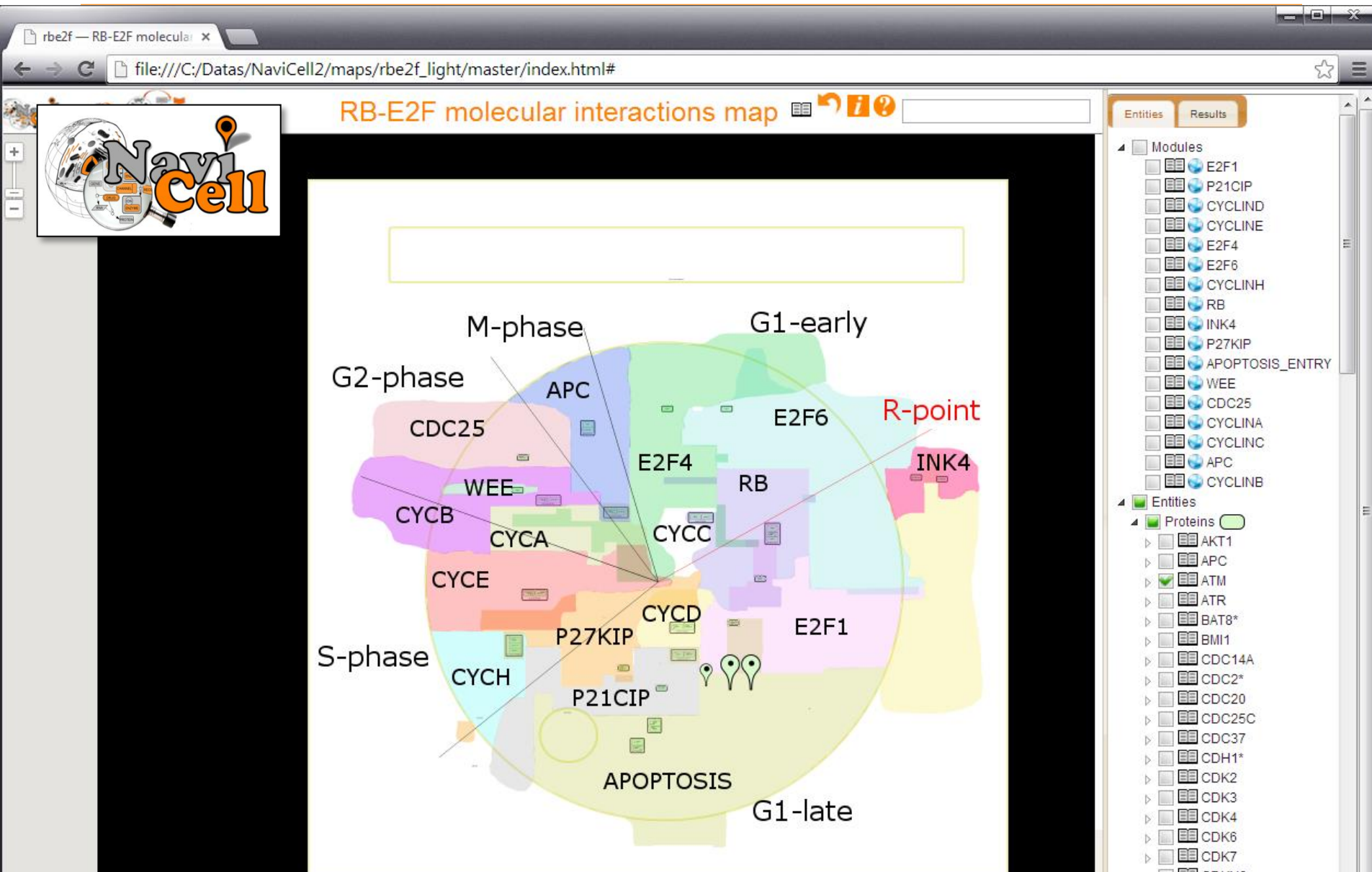
Example: Google maps of cell cycle

(Calzone et al, MSB, 2008; Kuperstein et al, BMC Sys Bio, 2013; Bonnet et al, NAR, 2015)



Example: Google maps of cell cycle

(Calzone et al, MSB, 2008; Kuperstein et al, BMC Sys Bio, 2013; Bonnet et al, NAR, 2015)



Semantic zoom in NaviCell

(Calzone et al, MSB, 2008; Kuperstein et al, BMC Sys Bio, 2013; Bonnet et al, NAR, 2015)

file:///C:/Datas/NaviCell2/maps/rbe2f_light/master/index.html#

RB-E2F molecular interactions map

Protein

ATM

Identifiers
Ataxia telangiectasia mutated
[HUGO:ATM, HGNC:795, ENTREZ:472, UNIPROT:Q13315](#)

Modules
MODULE:APOPTOSIS_ENTRY
MODULE:CYCLINB
MODULE:E2F1

References

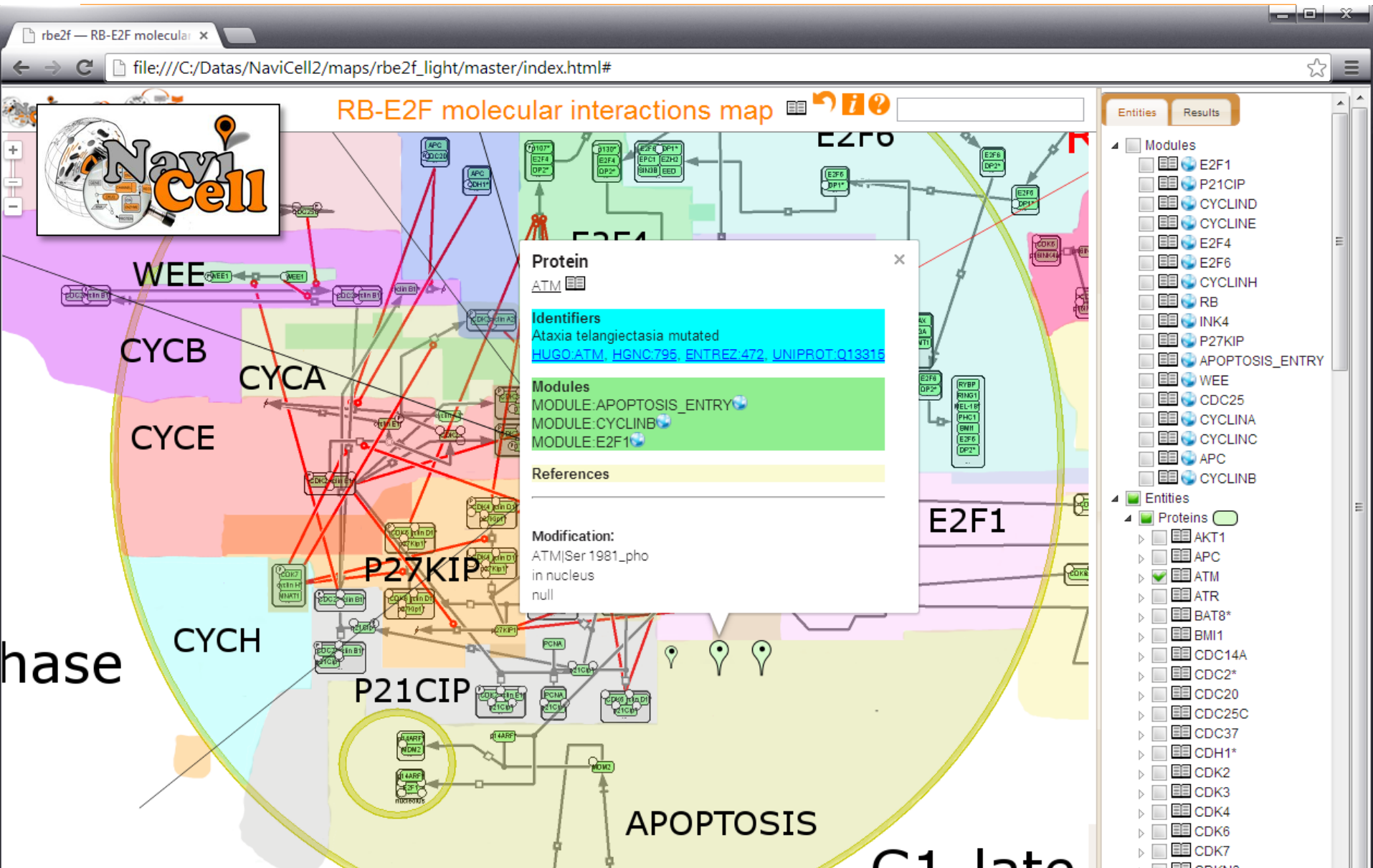
Modification:
ATM|Ser 1981_pho
in nucleus
null

Entities

- Proteins
 - AKT1
 - APC
 - ATM
 - ATR
 - BAT8*
 - BMI1
 - CDC14A
 - CDC2*
 - CDC20
 - CDC25C
 - CDC37
 - CDH1*
 - CDK2
 - CDK3
 - CDK4
 - CDK6
 - CDK7

Semantic zoom in NaviCell

(Calzone et al, MSB, 2008; Kuperstein et al, BMC Sys Bio, 2013; Bonnet et al, NAR, 2015)



Semantic zoom in NaviCell

(Calzone et al, MSB, 2008; Kuperstein et al, BMC Sys Bio, 2013; Bonnet et al, NAR, 2015)

The screenshot displays the NaviCell web application interface. The main window shows a complex network of molecular interactions for RB-E2F. A central popup window provides detailed information for the protein ATM:

- Protein:** ATM
- Identifiers:** Ataxia telangiectasia mutated
[HUGO:ATM](#), [HGNC:795](#), [ENTREZ:472](#), [UNIPROT:Q13315](#)
- Modules:** MODULE-APOPTOSIS_ENTRY, MODULE-CYCLINB, MODULE-E2F1
- References:** (empty)
- Modification:** ATM|Ser 1981_pho in nucleus null

The background map is titled "RB-E2F molecular interactions map" and features a NaviCell logo in the top left. The map is color-coded and shows numerous protein nodes and their interactions. A sidebar on the right lists "Entities" and "Results" under "Modules" and "Entities". The "Entities" list includes proteins such as AKT1, APC, ATM (checked), ATR, BAT8*, BMI1, CDC14A, CDC2*, CDC20, CDC25C, CDC37, CDH1*, CDK2, CDK3, CDK4, CDK6, and CDK7.

Module maps with simpler layout

(Calzone et al, MSB, 2008; Kuperstein et al, BMC Sys Bio, 2013; Bonnet et al, NAR, 2015)

The screenshot displays a web browser window showing a NaviCell module map for the E2F1 module. The browser tabs indicate the current page is 'rbe2f — E2F1 module'. The address bar shows the file path: 'file:///C:/Datas/NaviCell2/maps/rbe2f_light/E2F1/index.html'. The main content area is titled 'E2F1 module' and displays a complex network diagram of protein interactions within a cell, divided into 'nucleus' and 'cytosol'. The diagram features various protein complexes and molecules, with interactions indicated by lines and arrows. A legend on the right side of the browser window lists entities such as Proteins, Genes, RNAs, and Reactions, with corresponding icons. The NaviCell logo is visible in the top left corner of the browser window.

E2F1 module

Search (e.g. ATM)

Entities Results

- Entities
 - Proteins
 - ATM
 - ATM|Ser 1981_pho@nucleu
 - ATR
 - CDK2
 - CHEK2
 - CREBBP
 - DP1*
 - E2F1
 - EP300
 - HDAC1
 - PCAF*
 - TFIIH*
 - TOPBP1
 - cyclin A2*
 - p14ARF*
 - pRB*
 - Genes
 - RNAs
 - antisense RNAs
 - Simple molecules
 - Ions
 - Drugs
 - Phenotypes
 - Unknown
 - Complexes
 - Reactions

Keyword search in NaviCell

(Calzone et al, MSB, 2008; Kuperstein et al, BMC Sys Bio, 2013)

The screenshot displays the NaviCell web application interface. The browser address bar shows the URL: `file:///C:/Datas/NaviCell2/maps/rbe2f_light/master/index.html#`. The page title is "RB-E2F molecular interactions map". A search bar contains the keyword "melanoma". A red arrow points to the search bar.

The main visualization is a circular cell cycle diagram divided into phases: M-phase, G1-early, G1-late, S-phase, and G2-phase. Key proteins and genes are labeled within these phases, including APC, CDC25, WEE, CYCB, CYCA, CYCE, CYCH, P27KIP, P21, CYCC, RB, E2F4, E2F6, E2F1, and INK4. A red line labeled "R-point" is shown near the INK4 protein.

On the right side, a sidebar titled "Entities Results" shows 12 elements matching the keyword "melanoma". The results are categorized as follows:

- Entities
 - Proteins
 - p14ARF*
 - p14ARF*@nucleus
 - MDM2:p14ARF*@nucleus
 - E2F1:p14ARF*@nucleus
 - E2F1:p14ARF*@nucleole
 - MDM2:p14ARF*@nucleole
 - DP1*:p14ARF*@nucleole
 - DP1*:p14ARF*@nucleus
 - p16INK4a*
 - p16INK4a*@Cytosol
 - CDK6:p16INK4a*@Cytosol
 - CDK4:p16INK4a*@Cytosol
 - Genes
 - MCAM
 - gMCAM@Gene compartme
 - p14ARF*
 - gp14ARF*@Gene compartr

Keyword search in NaviCell

(Calzone et al, MSB, 2008; Kuperstein et al, BMC Sys Bio, 2013; Bonnet et al, NAR, 2015)

The screenshot displays the NaviCell web interface. The browser address bar shows the URL: `file:///C:/Datas/NaviCell2/maps/rbe2f_light/master/index.html#`. The page title is "RB-E2F molecular interactions map". A search bar contains the keyword "apoptosis". A red arrow points to the search bar.

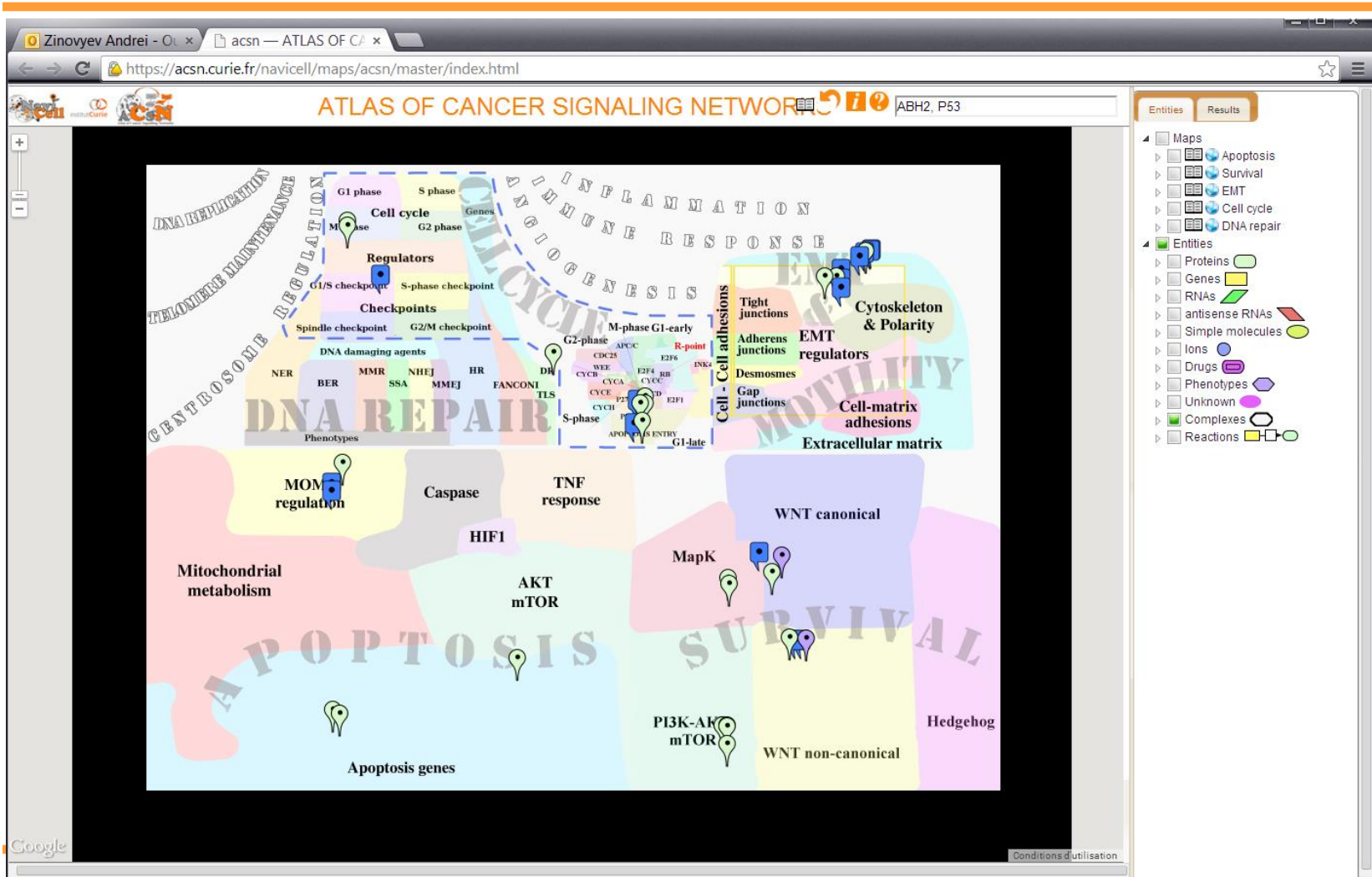
The main visualization is a circular map of the cell cycle, divided into phases: M-phase, G1-early, G1-late, S-phase, G2-phase, and R-point. Various proteins are labeled on the map, including APC, CDC25, WEE, CYCB, CYCA, CYCE, CYCH, P27KIP, P21, IP, CYCC, RB, E2F4, E2F6, E2F1, and INK4. A cluster of green location pins is labeled "APOPTOSIS" at the bottom of the map.

On the right side, a sidebar shows the search results for "apoptosis". The sidebar has tabs for "Entities" and "Results". The "Results" tab is active, showing 95 elements matching the keyword. The results are organized into a tree structure:

- 95 elements matching "apoptosis"
 - Modules
 - E2F1
 - APOPTOSIS_ENTRY
 - Entities
 - Proteins
 - AKT1
 - AKT1@nucleus
 - ATM
 - ATM|Ser 1981_pho@nuc
 - ATM@nucleus
 - ATM|Ser 1981_pho:NBS
 - ATR
 - ATR@nucleus
 - CHEK1
 - CHEK1|S317_pho|S345
 - CHEK1@nucleus
 - CHEK2
 - CHEK2|Thr68_pho@nuc
 - CHEK2@nucleus
 - DP1*
 - DP1*:p14ARF*@nucleol
 - DP1*:p14ARF*@nucleus
 - E2F1
 - E2F1|Ser403_pho@nucl
 - E2F1|pho@nucleus
 - E2F1|pho|Lys_ace@nuc
 - E2F1@nucleus
 - DP1*:E2F1:SW|SNF*:pF
 - DP1*:E2F1|Lys_ace:PC/
 - DP1*:E2F1:HDAC1:SUV:
 - DP1*:E2F1|Lys_ace|Ser:
 - DP1*:E2F1|Lys_ace|Ser:
 - E2F1:p14ARF*@nucleus
 - DP1*:pho:E2F1|pho|Lys_
 - DP1*:E2F1|Lys_ace|Ser:

Atlas of Cancer Signaling Network

5 “continents” and 52 “countries”



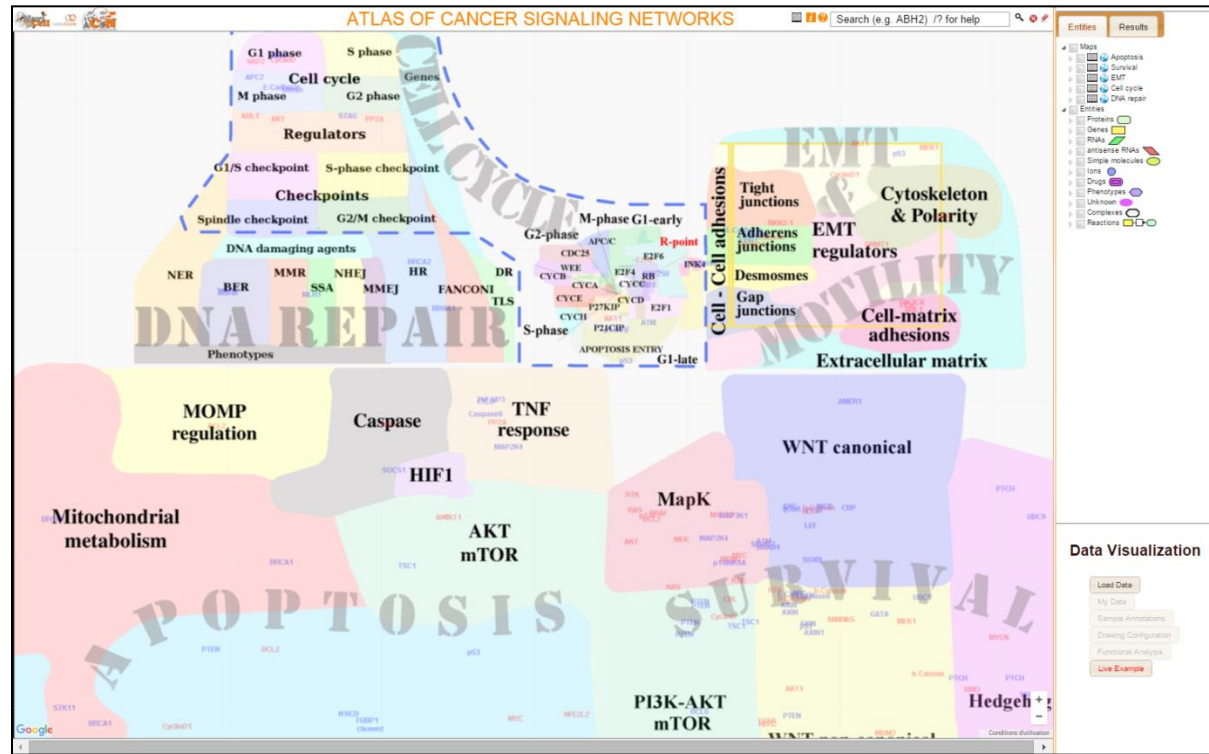
ACSN content today and tomorrow

Today

- Cell Cycle, DNA Repair, Apoptosis, Cell Survival, EMT and Cell Motility

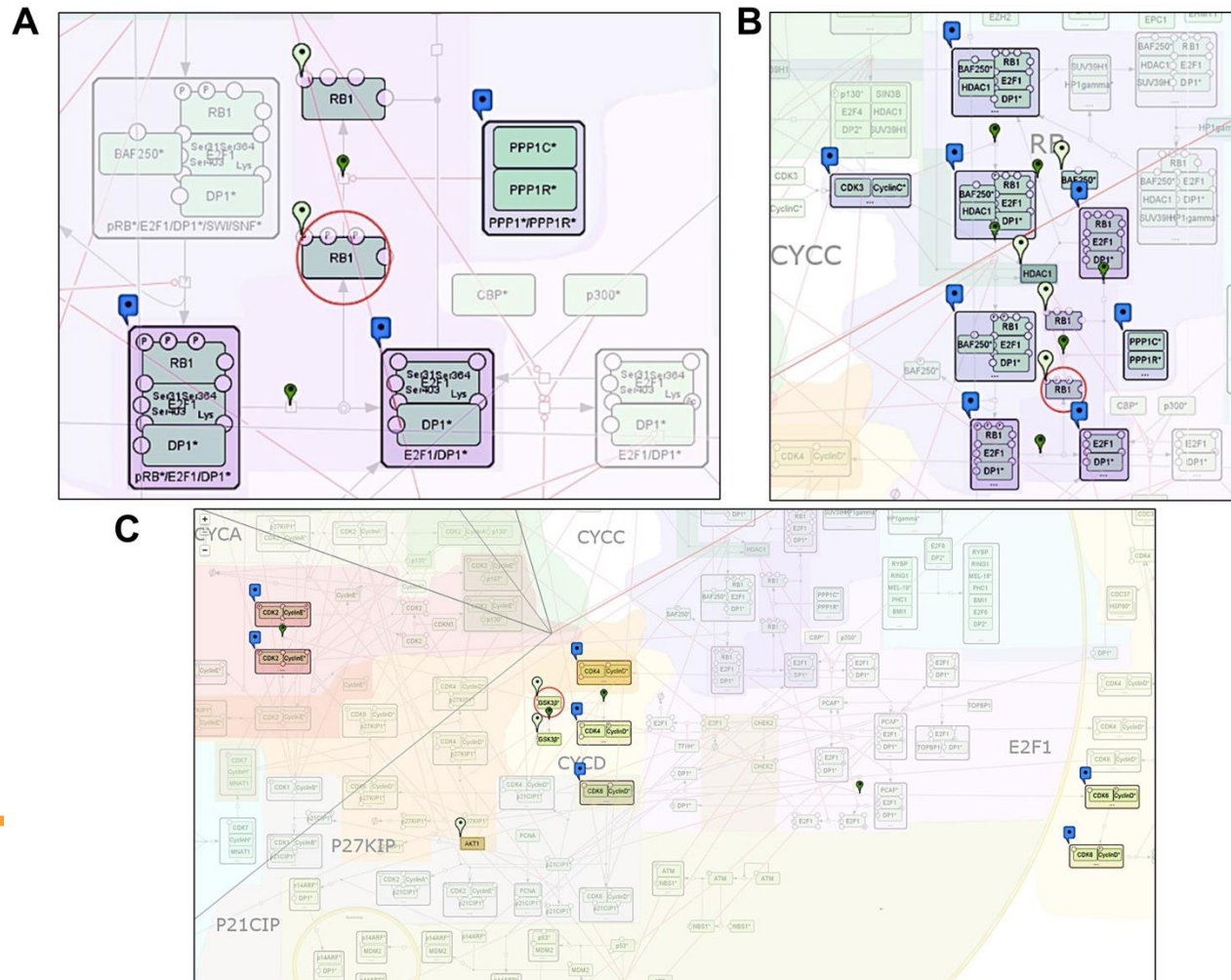
Tomorrow

- Detailed Regulated Cell Death map instead of Apoptosis
- Role of innate and adaptive immunity in cancer
- Angiogenesis in cancer
- Telomeres maintenance
- DNA replication



ACSN-NaviCell new functionality

- Highlighting part of the map
- Exploring protein/species network neighborhood



ACSN-NaviCell new functionality

Confidence scores for reactions and complexes

- **Reference score** (number of publications in the annotation)
- **Functional proximity score** (average distance between products and reactants or protein complex components in a PPI graph)

A

Reaction known_transition_omitted e_re1010

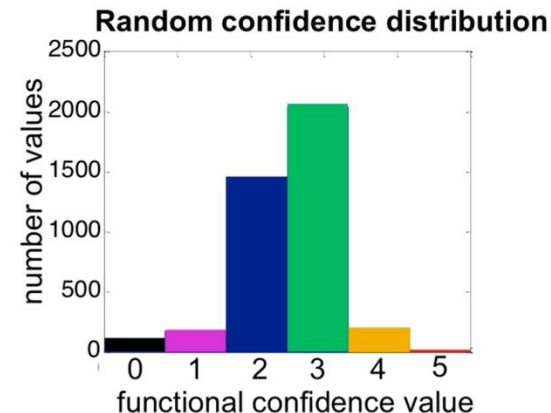
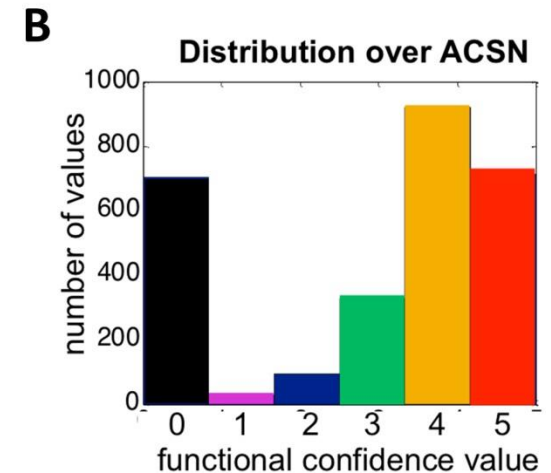
beta-Catenin*@Cytoplasm → _beta_-Catenin*|pho@Cytoplasm

Reaction regulators:
Catalysis

1. gamma-SRC*@Cytoplasm

References
[PMID:10593980](#)
[PMID:12123611](#)
Phosphorylation of CTNNB1 at Y654 by SRC prevents the interaction between Cadherin and CTNNB1. This phosphorylation therefore reduces adhesive function.

Confidence
★★★★☆ REF=2 FUNC=5



ACSN-NaviCell new functionality

- Built-in module enrichment analysis

tps://acsncurie.fr/navicell/maps/acsncurie/master/index.html



Functional Analysis

Select an Analysis: Gene Enrichment Analysis

Gene List: COMP1

P-value Threshold: 0.05

Background Set: Select an item
Whole Genome

Correction for Multiple Testing

Cancel
Execute

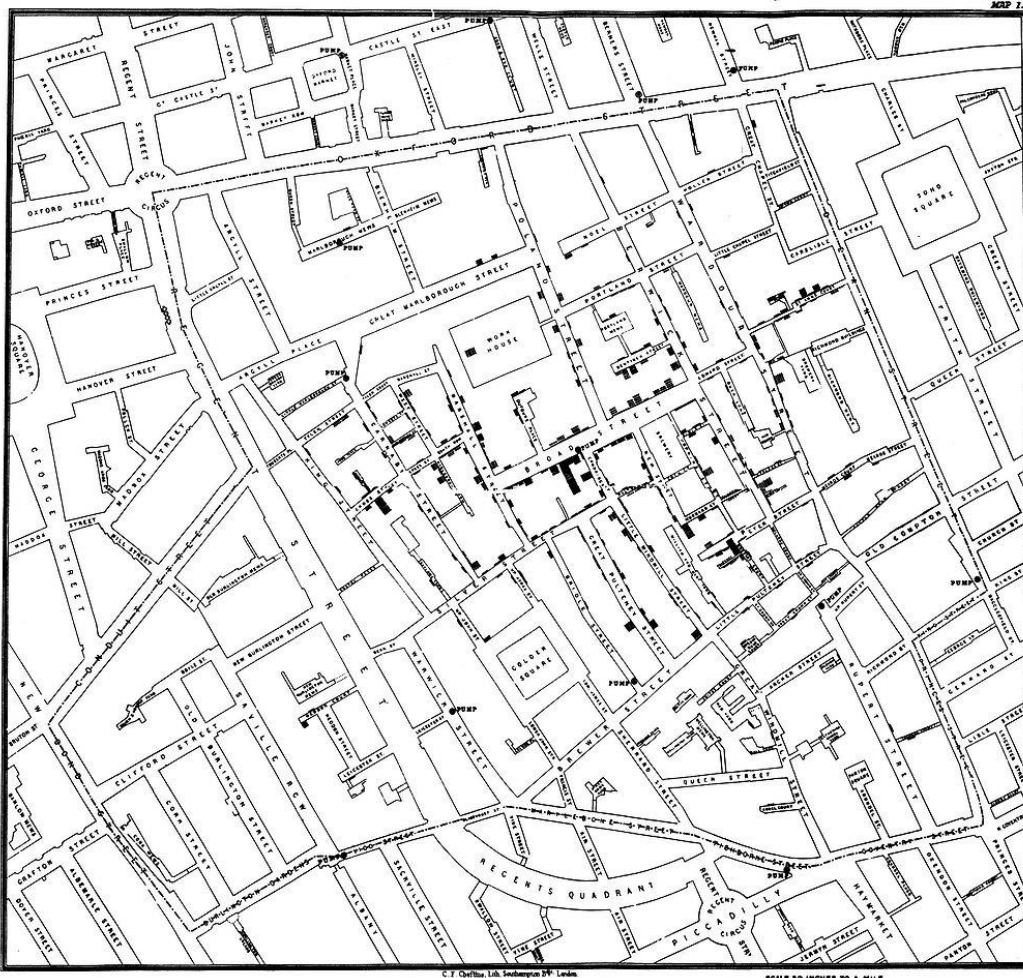


Gene Enrichment Analysis

Module	Module size	Nb genes in module	p-value (corrected)	
cellcycle:master	256	39	2.1e-19	NDC80 RRM2 TYMS RRM1 C UHRF1
dnarepair:master	346	42	2.9e-17	AURKB MAD2 CDC25C RFC3 POLQ RPA3
dnarepair:S_CC_PHASE	112	22	8.8e-13	CDK1 CCNB1 C GINS1 RPA3
dnarepair:SPINDLE_CHECKPOINT	51	13	2.7e-09	AURKB MAD2
dnarepair:FANCONI	82	14	1.6e-07	UBE2T FANCI
dnarepair:M_CC_PHASE	43	9	5.7e-06	MAD2L1 CDK1
cellcycle:CYCLINB	19	6	1.9e-05	CDK1 CCNB1 C
cellcycle:CDC25	12	5	2.1e-05	CDK1 CCNB1 C
dnarepair:BER	76	10	1.2e-04	CLSPN FEN1 R
dnarepair:S_PHASE_CHECKPOINT	65	9	1.8e-04	CLSPN FANCD
dnarepair:HR	82	10	2.3e-04	FANCI FANCD
cellcycle:APC	24	5	8.3e-04	CDK1 CCNB1 C

«Geographic» method

First application of geographic method

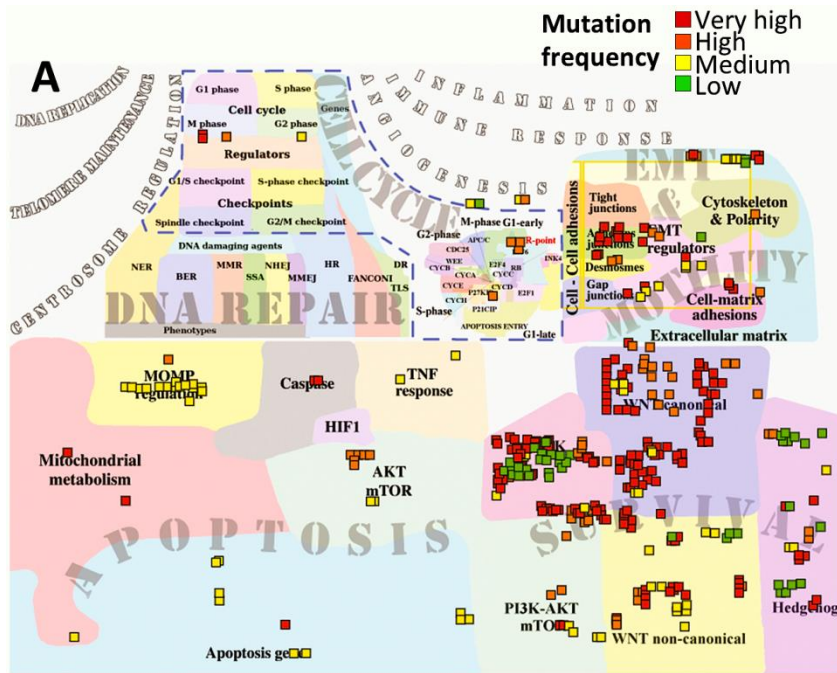


Original map by John Snow showing the clusters of cholera cases in the London epidemic of 1854. The pump is located at the intersection of Broad Street and Cambridge Street.

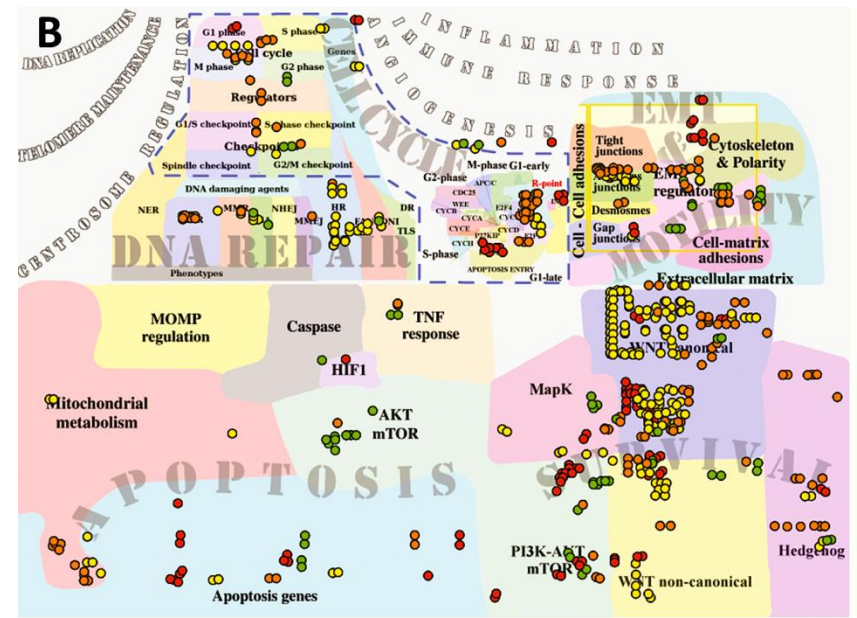
Data visualization with NaviCell 2.0

NaviCell Web Service (Bonnet et al, 2015, NAR)

Cancer driver genes in ACSN



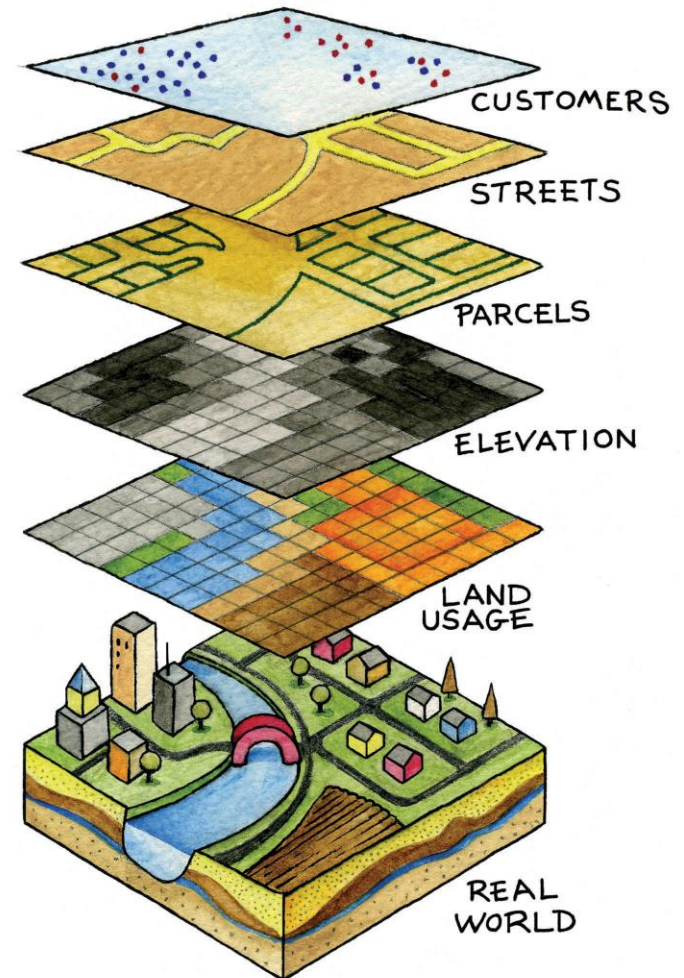
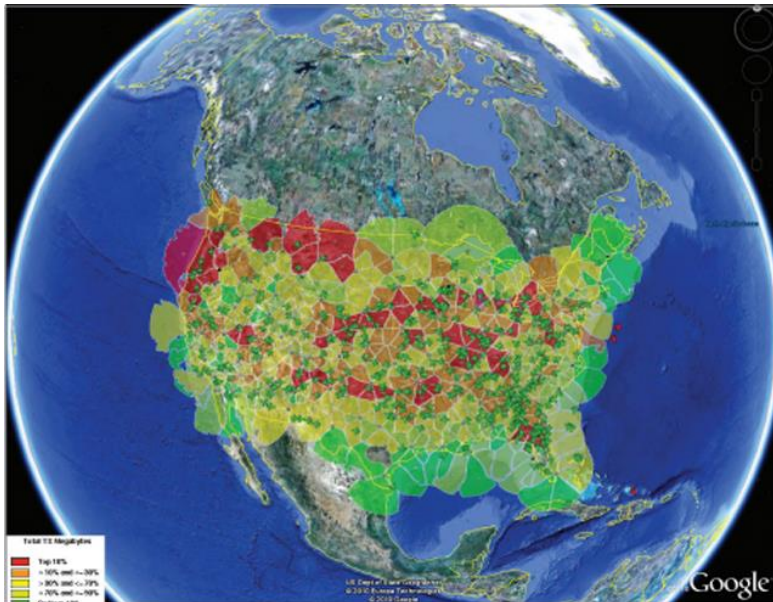
Oncogenes



Tumour suppressors

ACSN: Towards “Geographic” information system (GIS) for molecular biology

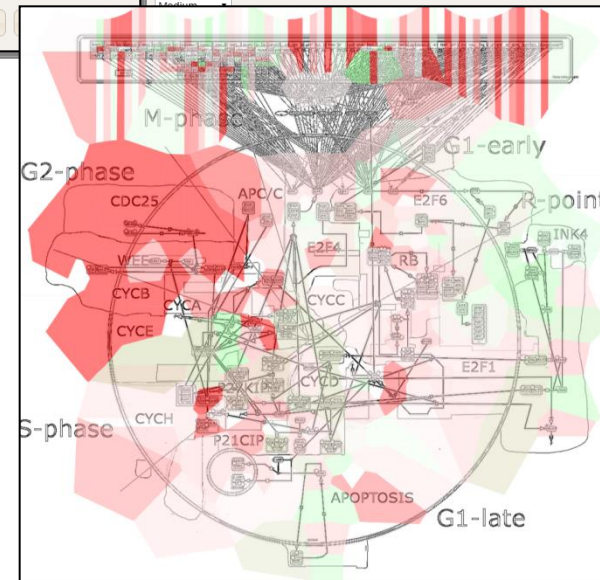
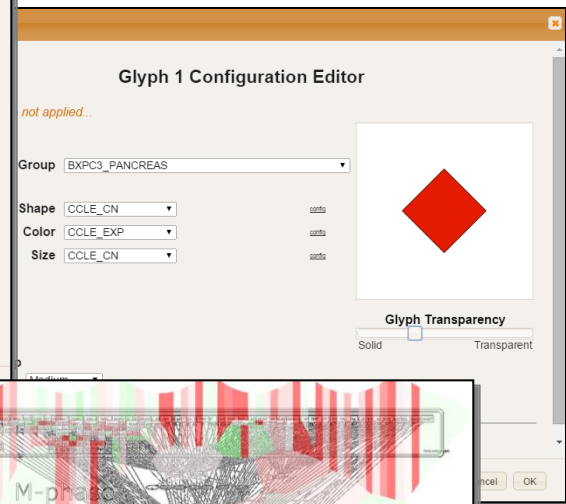
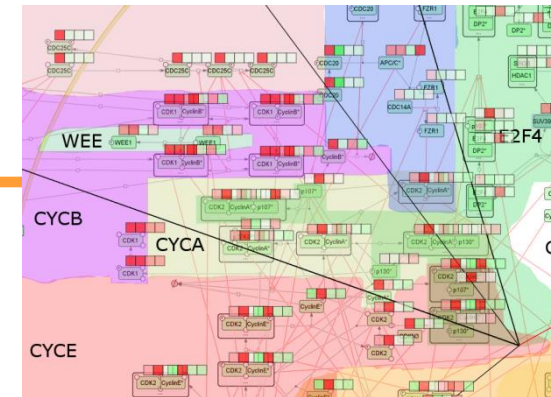
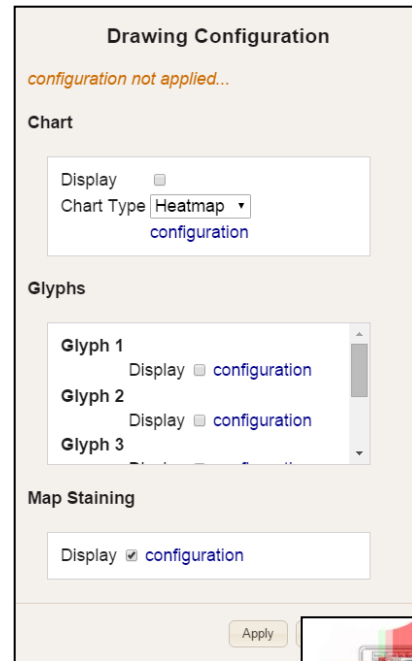
- Duality: an object exists in geographical space and in data space at the same time



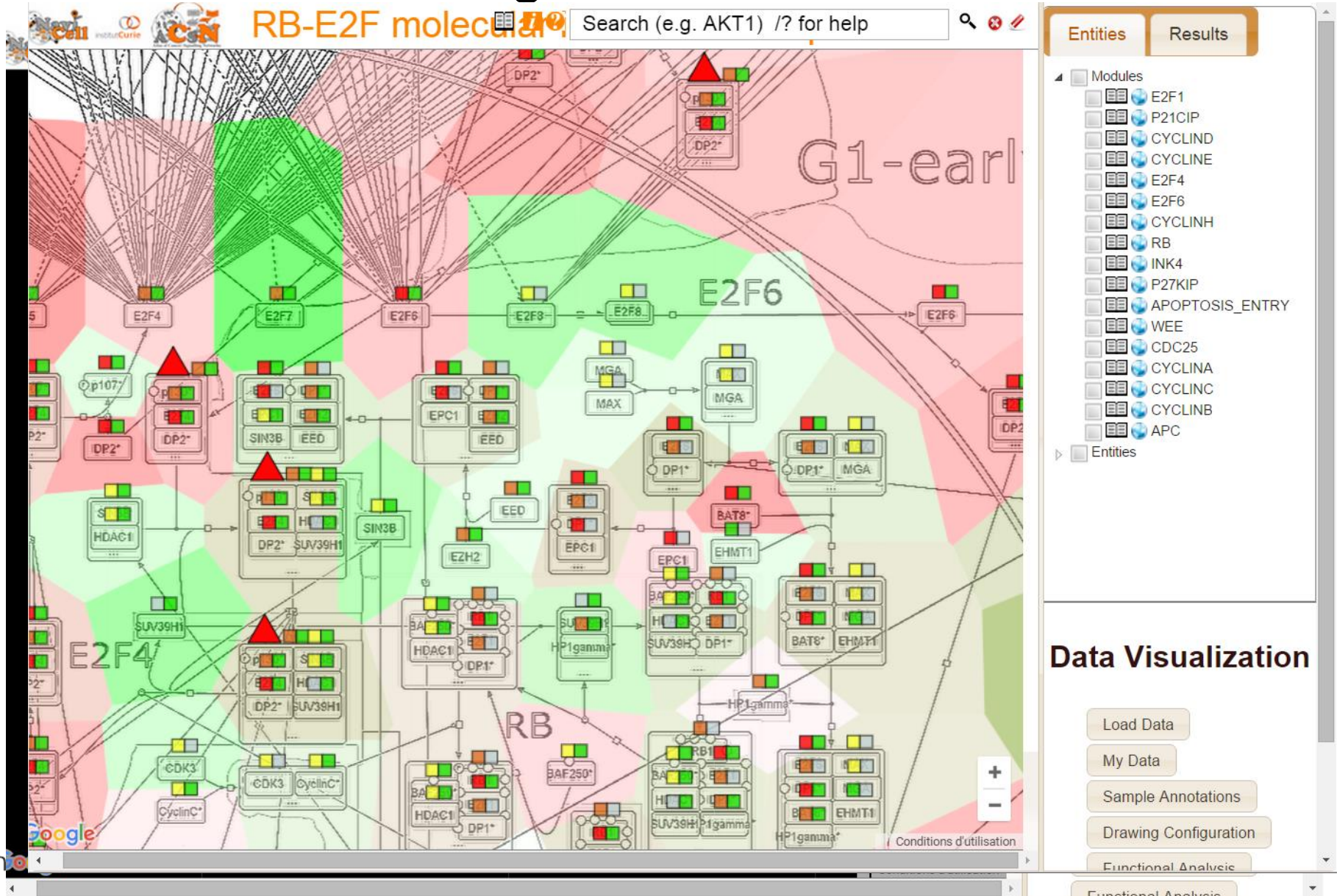
From <https://www.cachecounty.org/gis/>

Principles of multi-layer omics data visualization in ACSN and NaviCell

- **Intrinsic data types**
 - lists of objects
 - numerical object/sample tables (discrete and continuous)
 - categorical tables object/sample tables
 - sample annotation
- **Data visualization channels**
 - Charts: Barplots and heatmaps
 - Glyphs
 - Using map background (map staining)
- **Assigning a data channel to specific omics data**
i.e. Expression in map staining, mutations in glyphs
- **Aggregating measurements into groups accordingly to sample annotations (i.e. clinical data)**

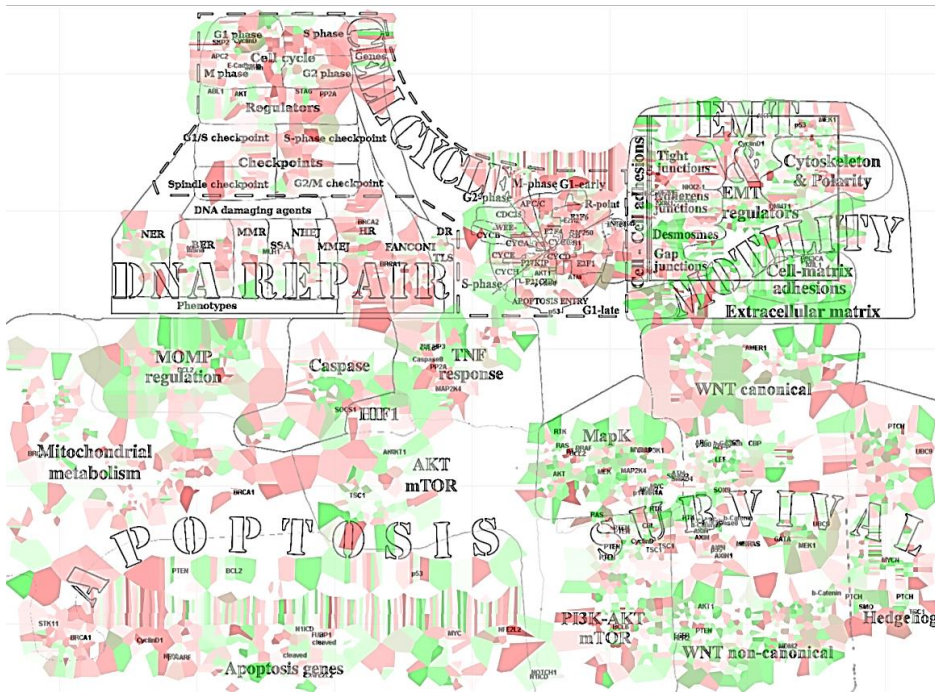


Complex interactive data visualization scenes in NaviCell: combining data visualization channels

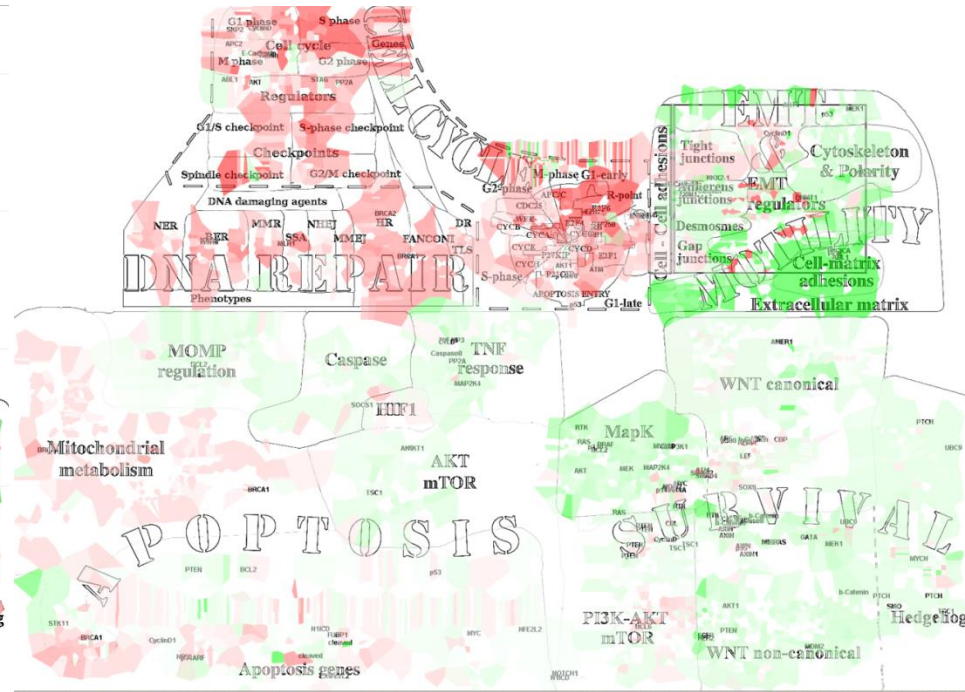


Pathway-based data abstraction in ACSN

“network smoothing” of the data



protein-based map staining



module-based map staining

NaviCell and ACSN as a web-server with API



= Browser+JavaScript+
Google Maps API



HTTP query
in R or Python
or Java

download map with code



manipulate NaviCell

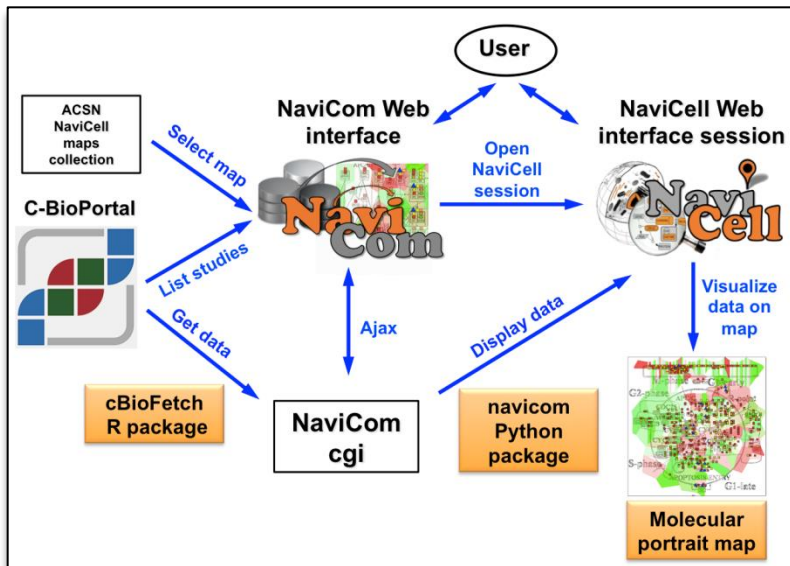
send command

send data



NaviCom portal: one click bridge between TCGA/cBioPortal and ACSN

<http://navicom.curie.fr>



NaviCom portal

[navicom vignette](#)

Welcome to NaviCom portal, a link between [cBioPortal](#) database and [NaviCell](#) web service. Select a study from which you want to fetch data, the map on which you want it to be displayed and the type of display you want to see. NaviCom uses the display function defined in the [navicom](#) python package. Note that downloading data from [cBioPortal](#) and exporting them to NaviCell can take some time, depending on the speed of your connection and your computer.

Study:
Bladder Cancer (MSKCC, J Clin Oncol 2013) ?

ACSN Map:
Cell cycle map or Alternative map URL (ex: http) ?

Display mode:
Complete display ?

Display configuration
Color for lowest values: 00FF00
Color for highest values: FF0000
Color for zero (if present): FFFFFFFF

Perform data visualisation Download cBioPortal data

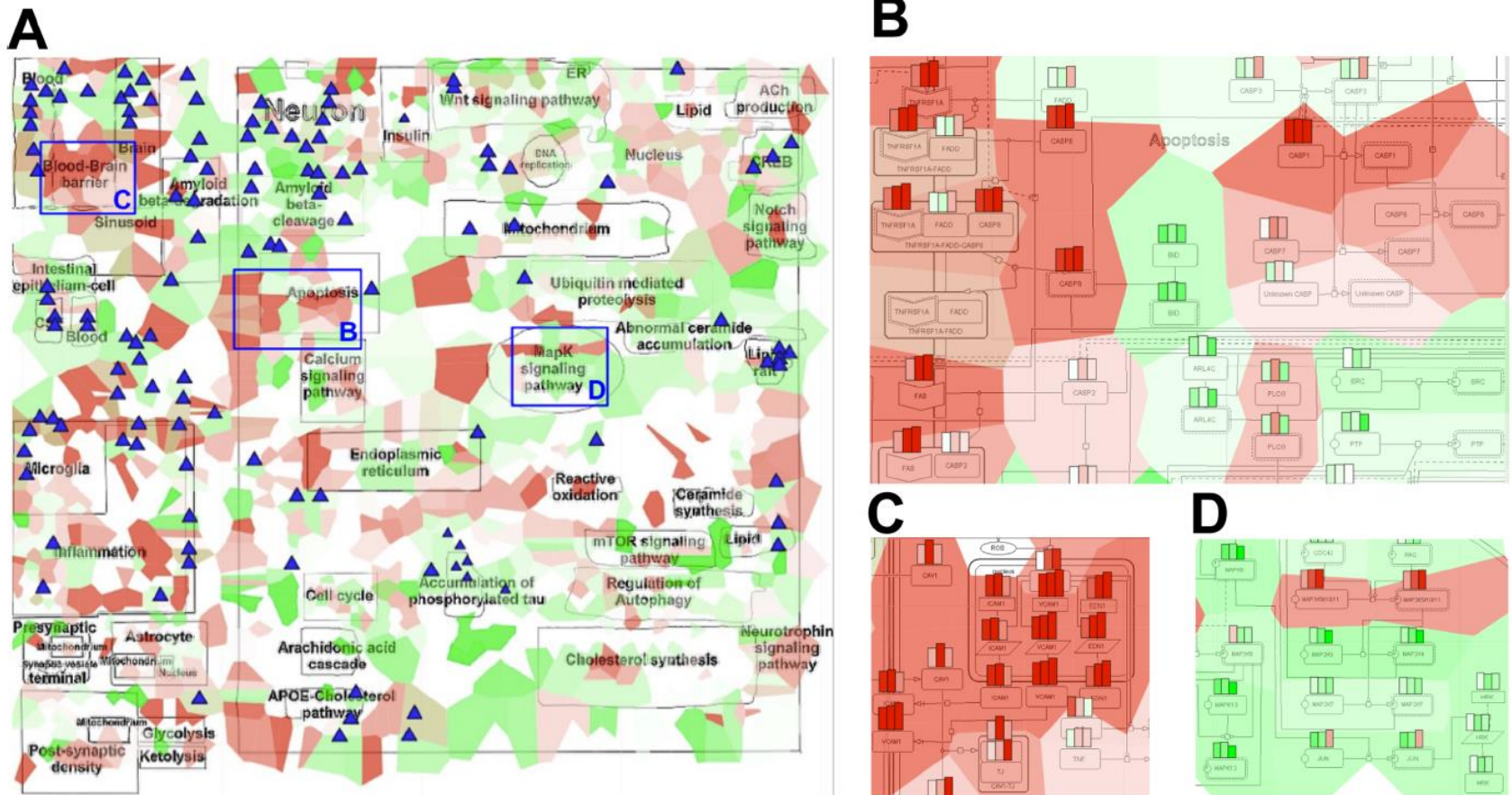
NaviCell beyond ACSN

<http://navicell.curie.fr>

- NaviCell can be used to
 - Browse any CellDesigner map not from ACSN
 - Browse any non-CellDesigner network map (i.e. any biological network that can be imported to Cytoscape)
 - Browse any graph, not necessary biological network
- The benefits are
 - Web-based environment, clickable interactive content
 - All functionality of NaviCell (semantic zoom, search, highlighting, data visualization toolbox)

Using Alzheimer disease CellDesigner map

https://navicell.curie.fr/pages/maps_alzheimer.html



(Bonnet et al, NAR, 2015)

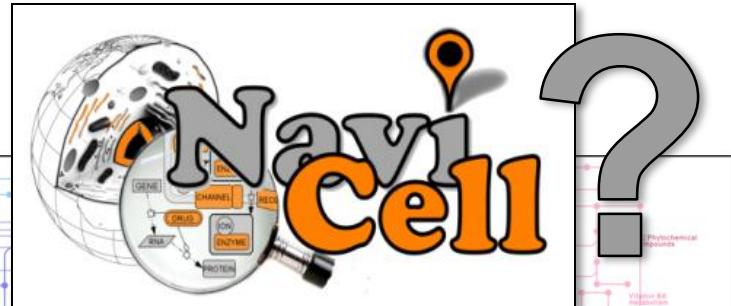
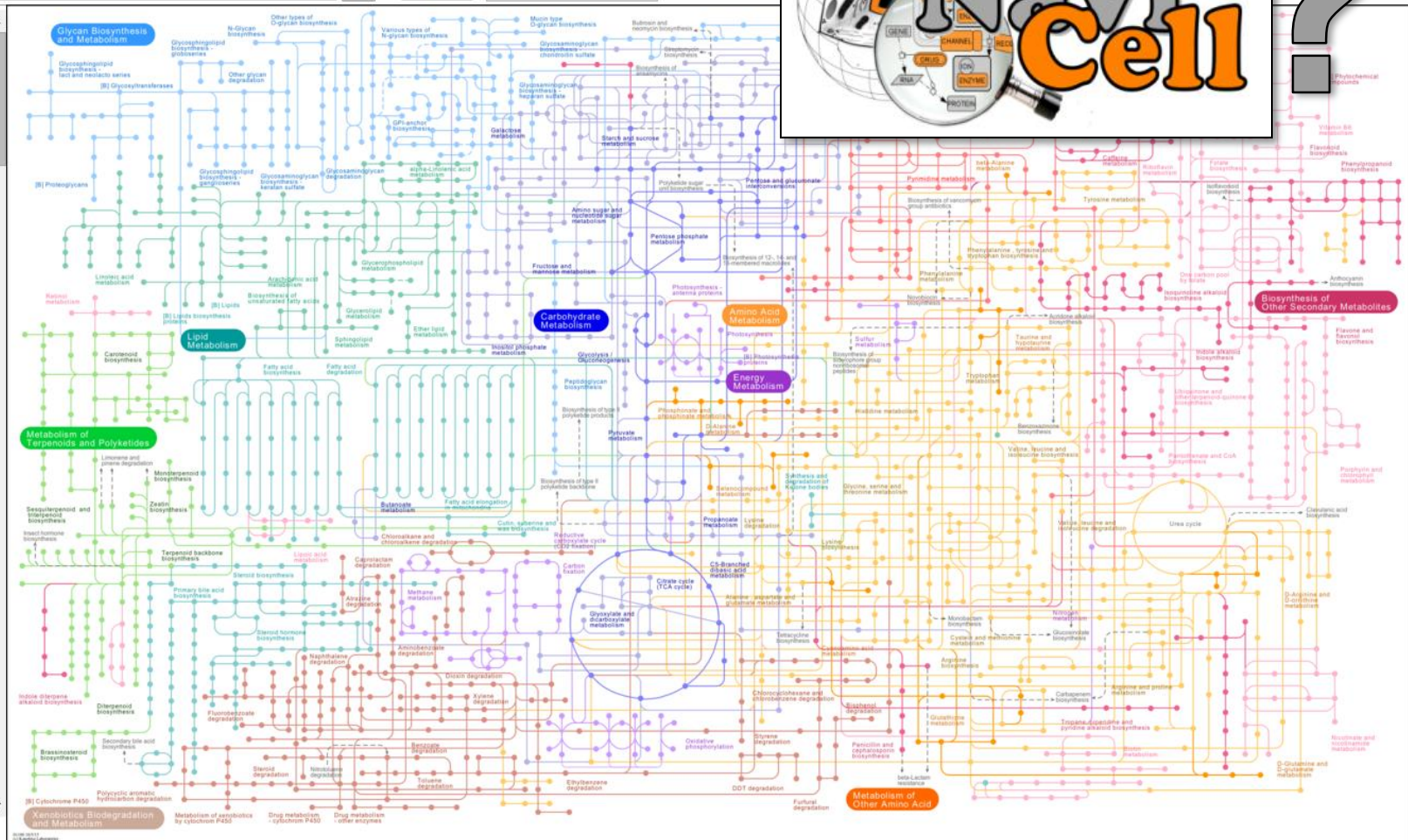
NaviCell as an interface for other pathway databases (eg, KEGG)

 Metabolic pathways - Reference pathway

[Pathway menu | Organism menu | Pathway entry | Hide module list | User data mapping | Image (png) file]

Reference pathway Go

- KEGG module**
- Energy metabolism
- Carbon fixation
 - M00165 Reductive pe
 - M00166 Reductive pe
 - M00167 Reductive pe
 - M00168 CAM (Crassu
 - M00169 CAM (Crassu
 - M00172 C4-dicarboxy
 - M00171 C4-dicarboxy
 - M00170 C4-dicarboxy
 - M00173 Reductive cit
 - M00579 Phosphate a
- Nitrogen metabolism
- M00175 Nitrogen fixa
- Methane metabolism
- M00567 Methanogen
 - M00174 Methane oxi
- Sulfur metabolism
- M00176 Assimilatory
 - M00596 Dissimilatory
 - M00595 Thiosulfate c
- Carbohydrate and lipid met
- Central carbohydrate m
 - M00001 Glycolysis (E
 - M00002 Glycolysis, cc
 - M00003 Gluconeogen
 - M00307 Pyruvate oxi
 - M00009 Citrate cycle
 - M00010 Citrate cycle,
 - M00011 Citrate cycle,
 - M00004 Pentose pho
 - M00006 Pentose pho
 - M00007 Pentose pho
 - M00580 Pentose pho
 - M00005 PRPP biosynt
 - M00008 Entner-Douc
 - M00308 Semi-phosph
 - M00633 Semi-phosph
 - M00309 Non-phosph



Acknowledgements

Present and past members of Computational Systems Biology of Cancer group (Institut Curie)

Inna Kuperstein

Eric Bonnet

Emmanuel Barillot

Laurence Calzone

David Cohen

Mathurin Dorel

Luca Grieco

Christophe Russo

Maria Kondratova

Simon Fourquet

Hien-Ahn Nguyen

Urszula Czerwinska

All experts contributed to map curation
(see full list at the web-site)

Sysra company

Eric Viara and Stuart Pook

INSERM U1021

Marie Dutriex

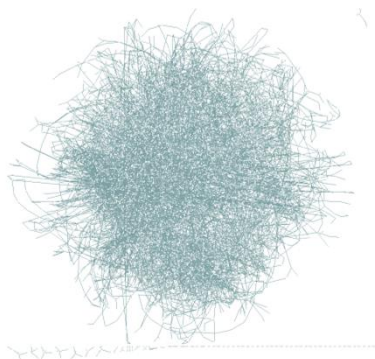
Agilent company

Thought Leader Award-2013

More than 15 projects including European ones, supported ACSN map construction in 2008-2015

ACSN

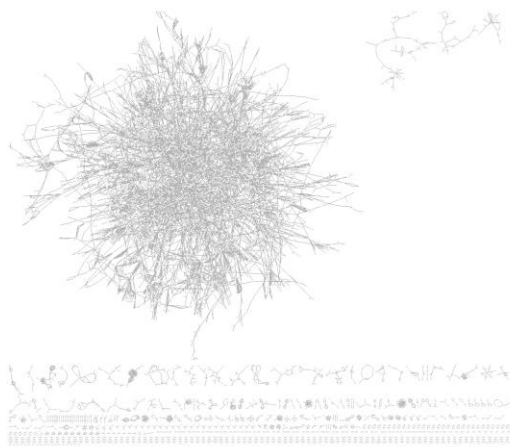
of distinct proteins = **1814**
of protein complexes = **1306**
Reactions per protein = **2.77**
Complexes per protein = **0.72**



LCC size = **5875** species/**4305** reactions
LCC Characteristic path length, directed = **43.6**
LCC Characteristic path length, undirected = **16.3**

NCI PID

of distinct proteins = **3399**
of protein complexes = **3904**
Reactions per protein = **2.67**
Complexes per protein = **1.15**



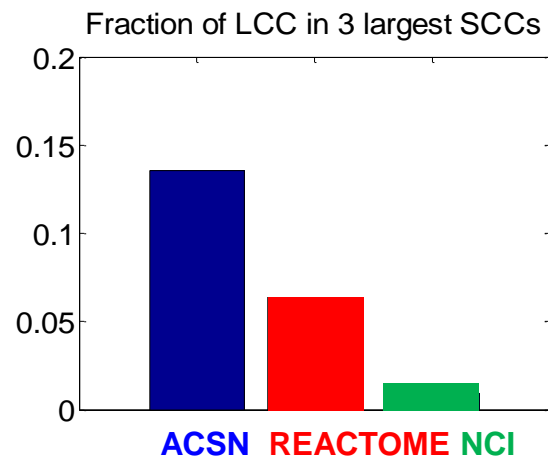
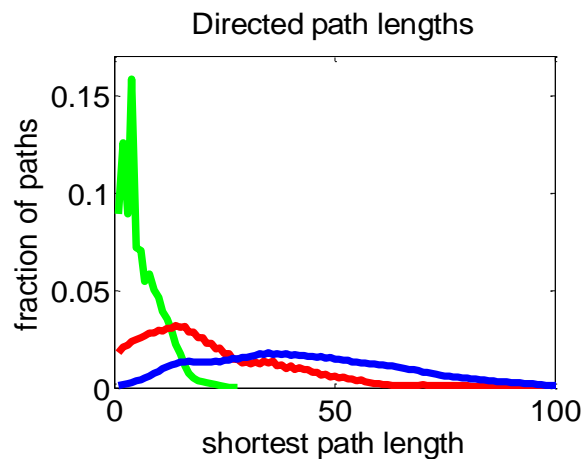
LCC size = **5038** species/**4501** reactions
LCC Characteristic path length, directed = **6.4**
LCC Characteristic path length, undirected = **16.8**

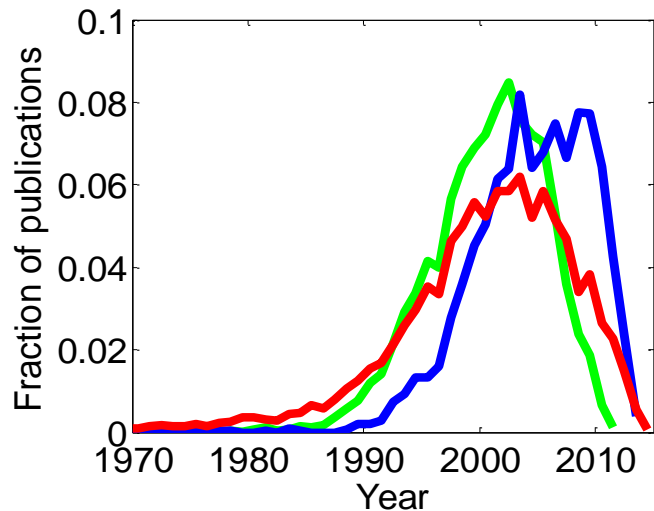
REACTOME

of distinct proteins = **8082**
of protein complexes = **6552**
Reactions per protein = **1.06**
Complexes per protein = **0.85**



LCC size = **5181** species/**3392** reactions
LCC Characteristic path length, directed = **23.6**
LCC Characteristic path length, undirected = **24.6**



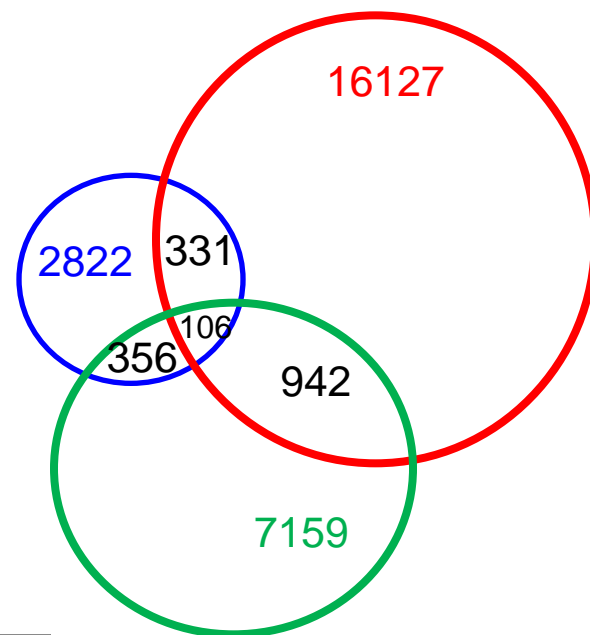


Median year of publication:

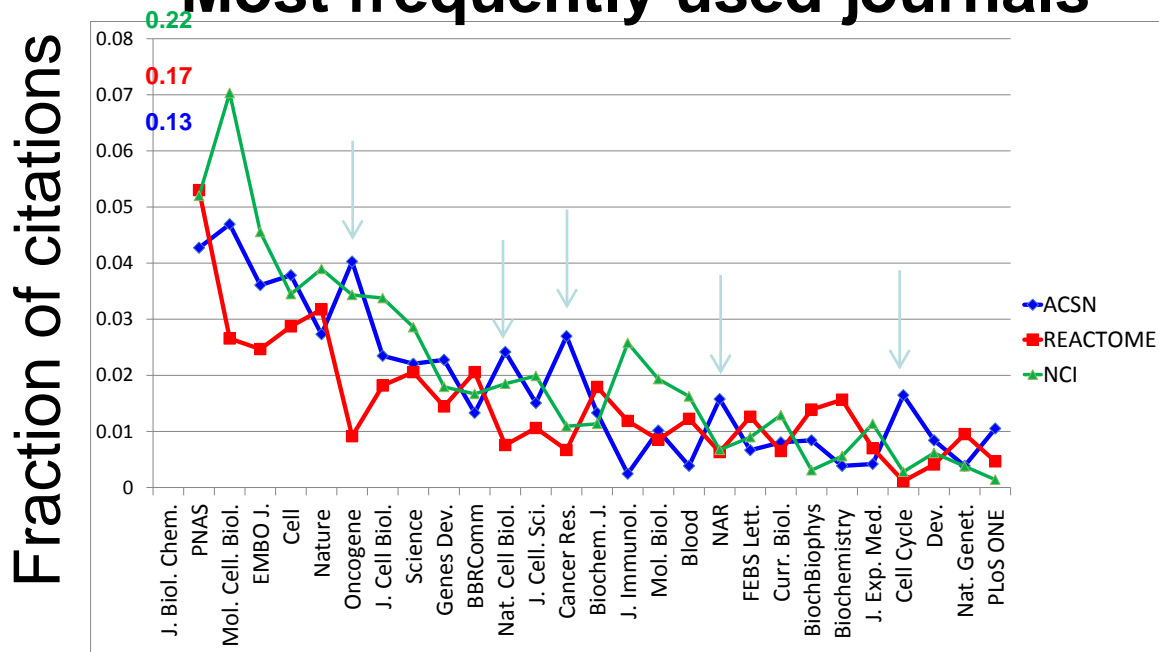
REACTOME=2002

NCI PID=2002

ACSN=2005



Most frequently used journals



Use of reviews:

REACTOME 17%

NCI PID 0.8%

ACSN 20%

Total number of journals used/
normalized by number of citations:

REACTOME 1198 / 0.074

NCI PID 334 / 0.047

ACSN 401 / 0.142