



Integrative Analyses

Complex Systems & Predictive Models.

Dr. François Iris CSO BMSystems

Integrative Biology Training Program Cochin Institute 2015

UE Cellular Signaling

<http://cochin.inserm.fr/institute/institute-presentation>

You received this document for your personal use.

This non-public document is under copyright law and cannot be used for other purposes without BMSystems' permission.

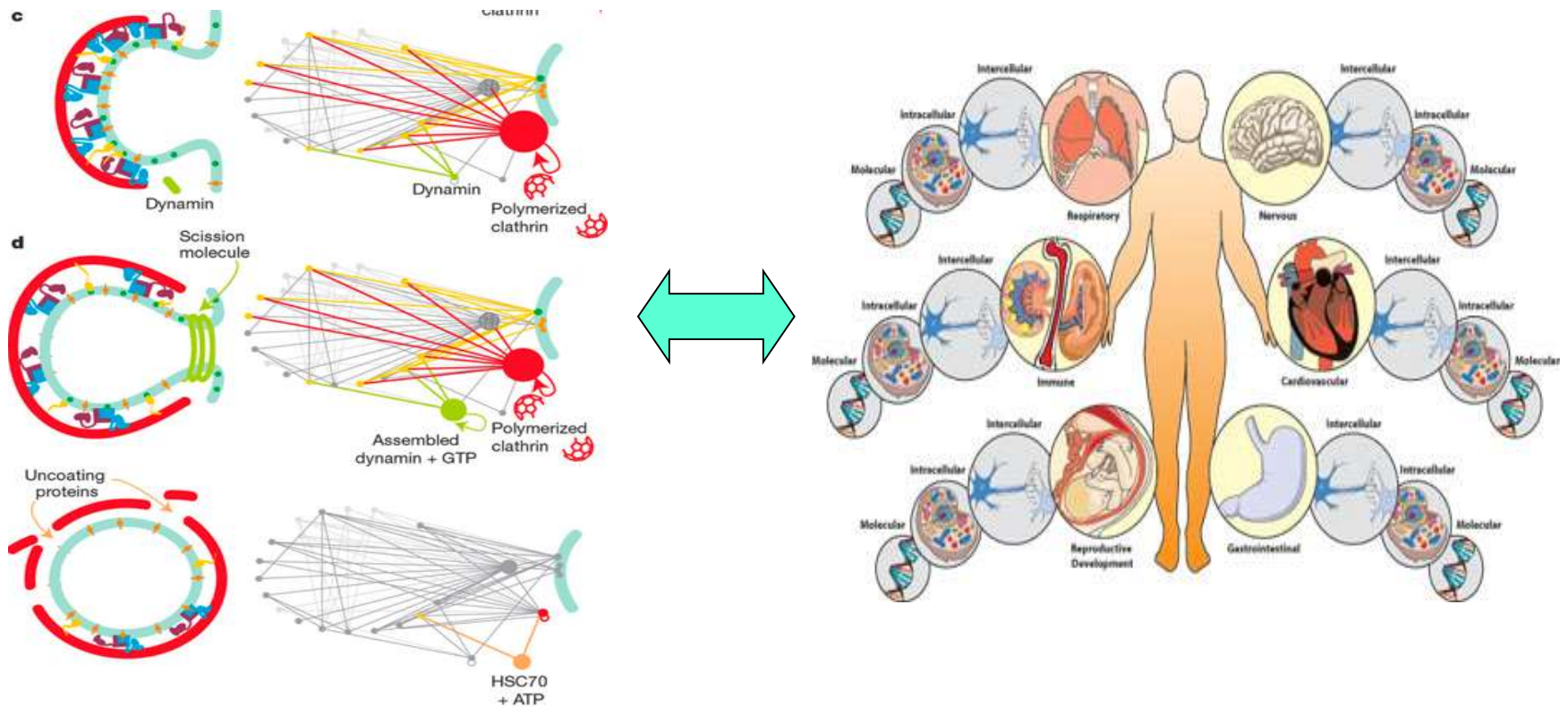
Agenda of the session

- What is Systems Biology Pages 3-6
- Example #1 CNS: Creutzfeldt-Jakob disease physiopathology deciphered Pages 7-27
- Modelling living systems? The 3 issues/problems/limits Pages 28-39
- Classical Systems Biology concepts, tools and issues. Pages 40-51
- Another way of thinking to address the issues. Pages 52-57
- CADI Discovery concept and workflow deciphered. Pages 58-65
- Example #2 Cancer: The two approaches outputs compared. Pages 66-73
- Example #3 Infectious diseases: Multiple Resistance Bacteria's solution. Pages 74-81

This training session is designed to be presented to professional students. The missing oral comments, PPT animations, and the “funny moppet's show part” context may be more difficult to understand for non-french readers. This is why we are open to organize specific training sessions or webinars.

What is "Systems Biology"?

It is an analytical approach that utilises biological data to deduce interaction networks and biological functions that are then integrated to construct models predictive of the physiological consequences resulting from defined functional alternation of these networks



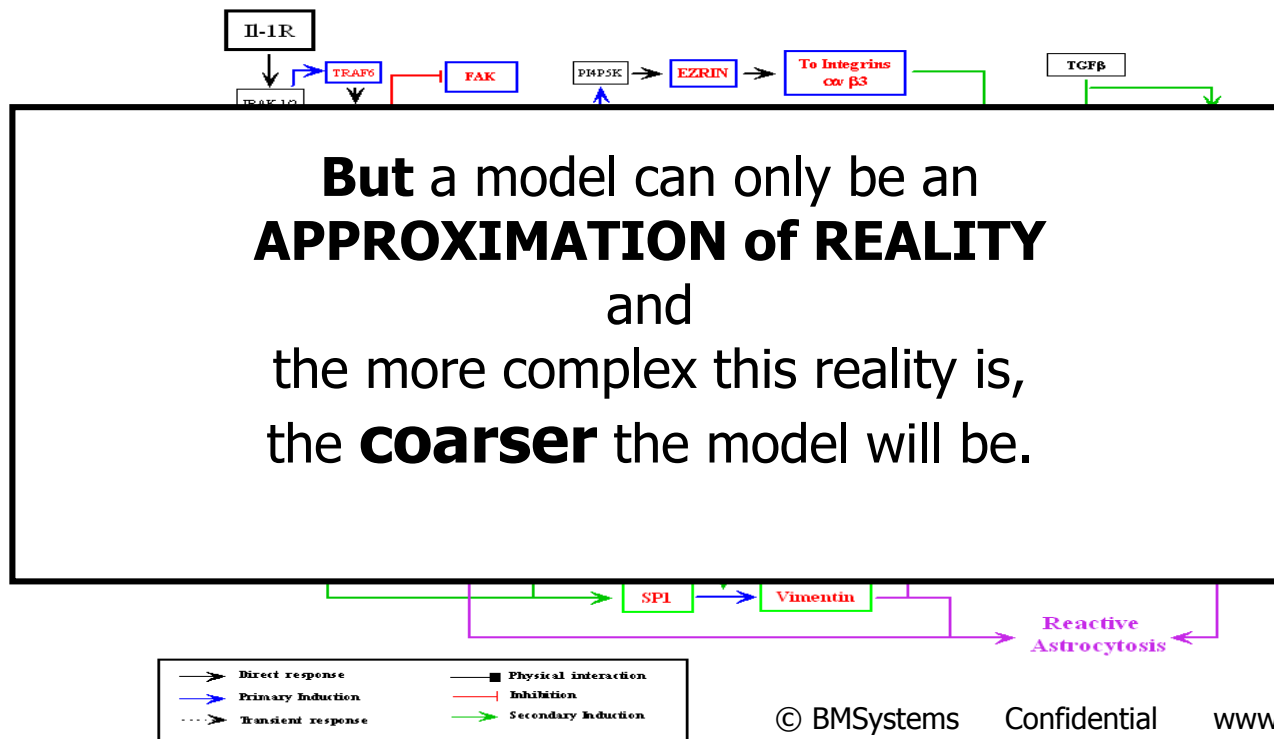
Systems biology isn't domain-dependent but is **information**-dependent.

What is an Integrated Predictive Model?

It is a detailed map of the mechanisms associated with a **defined biological state**, allowing the direct identification of:

- the key mechanisms, and
- the forms of interactions

required to produce the **physiological events** driving this state.



What is the use of an integrated predictive model?

It predicts which components/mechanisms should be targeted

- When,
- Where,
- How, and
- Why

To produce **defined effects** upon a given biological state.

An understanding of the mechanisms directly leads to:

- 1) Identification of the most relevant biomarkers (positives AND negatives);*
- 2) The discovery of potential means of therapeutic intervention.*

But there are NO magic wands!

As long as it is not biologically validated,
a model merely remains
an exercise in science-fiction.

Moreover

A model only is an **assistance** to thoughts
and NOT a substitute for thoughts!

By definition, a *model*, even biologically “validated”, cannot, by any means, be regarded as “correct” in the absolute.

But, how is a “model” constructed and
what does it look like?

Example # 1

In vivo mechanisms of Creutzfeldt-Jakob Disease (CJD) pathogenesis & progression.

Initial model constructed from:
Scientific publications only.

(Simultaneous modelling in neurons, glial cells, astrocytes & vasculature)

Validated in vivo (molecular & functional levels) in:

- the mouse;
- the hamster.

Academic collaboration BM-Systems / CEA (Fontenay-aux-Roses)

This model addresses an organ composed of a multitude of poorly characterised cell populations in constant communications.

This constitutes a truly complex system the model of which was tested in vivo.

CJD: A few facts.

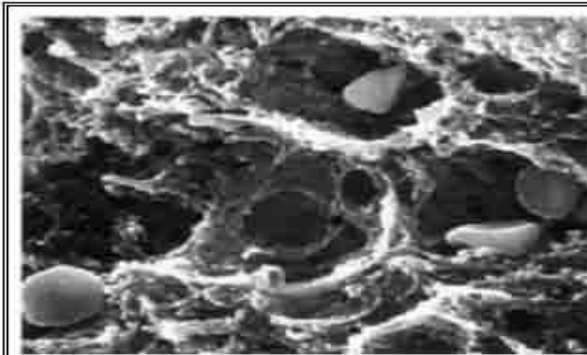
- Progressive neurodegenerative disease, invariably fatal.
- Long asymptomatic incubation phase (30 years + in man)
- Short clinical phase (death within 6 months - 2 years)
- Pathological agent: abnormally folded form of the PrP protein
- Pathological mechanisms : **Unknown**
- Mode of propagation within the CNS: **Unknown**
- Clinical progression mechanisms : **Unknown**
- Therapeutic or preventive means: **Unknown**
- Pathogenesis Biomarkers: **Clinical symptoms + CNS spongiosis** (post-mortem)

Histopathological lesions confined to the central nervous system :

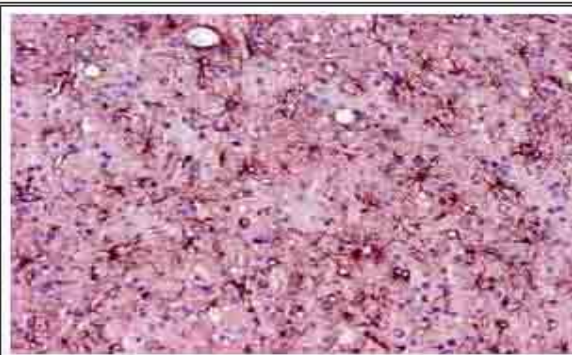
- Massive neuronal death
- Spongiosis
- Reactive Astrogliosis
- Inconstant amyloïd plaques

Tripartite implication:

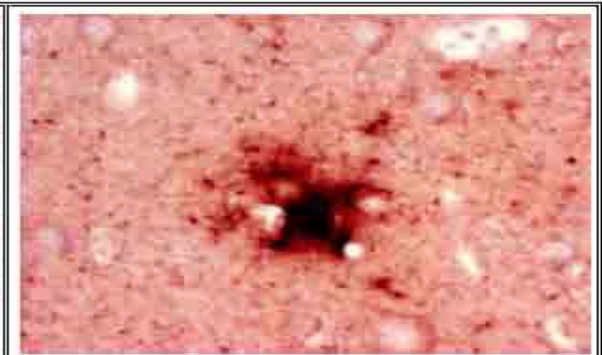
- The neurons
- The astrocytes
- The microglial cells



Spongiose

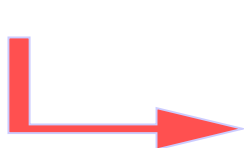


Astrogliose



Plaque amyloïde

Limits of the exclusively experimental approach



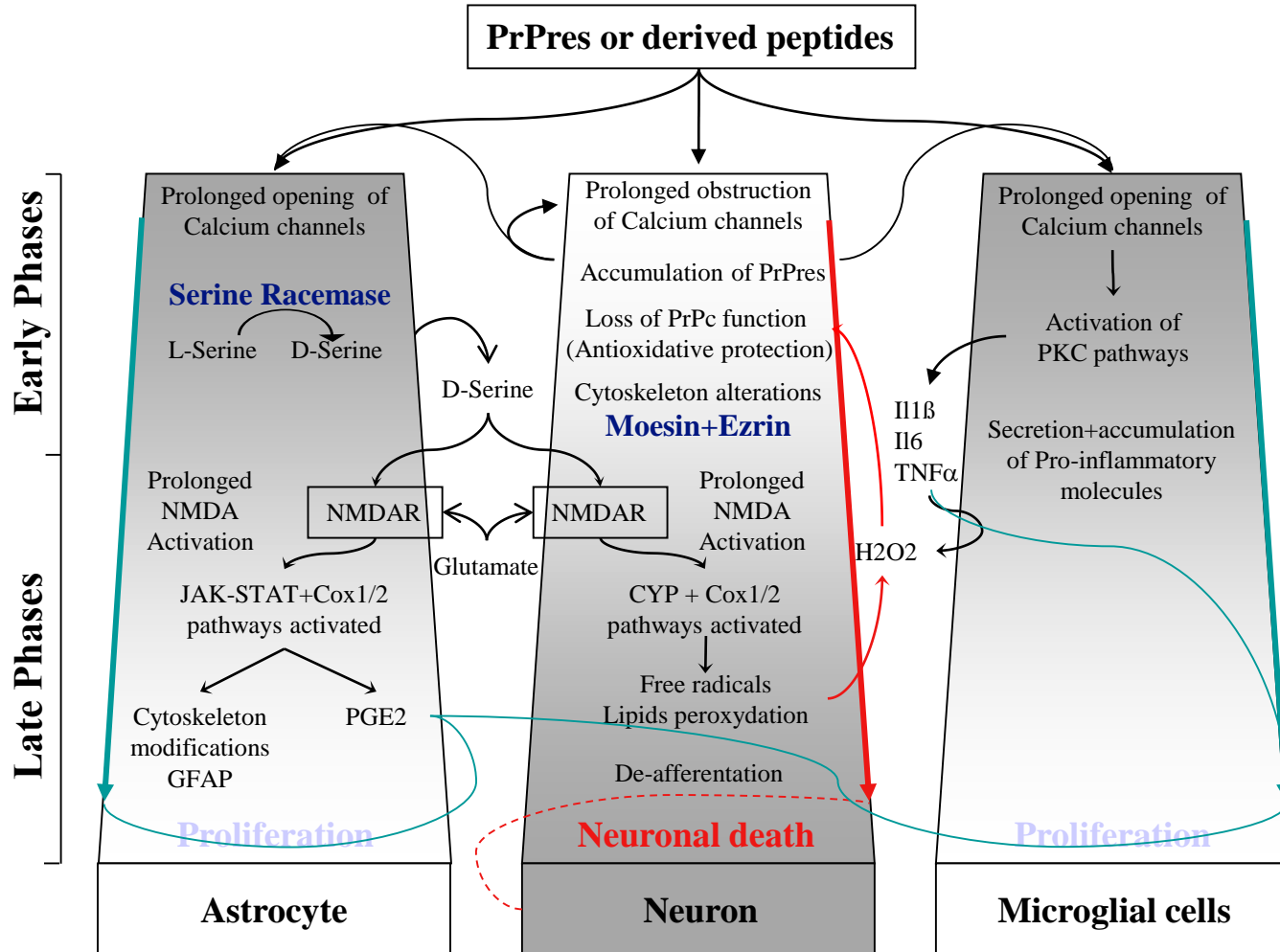
« Step by step » Research

Exhaustive bibliographic analyses impossible

(october 2008 : Prion = 12.095 articles, Brain = 1 894.937 articles, Metabolism = 5.912.125 articles)

Creutzfeldt-Jakob Disease.

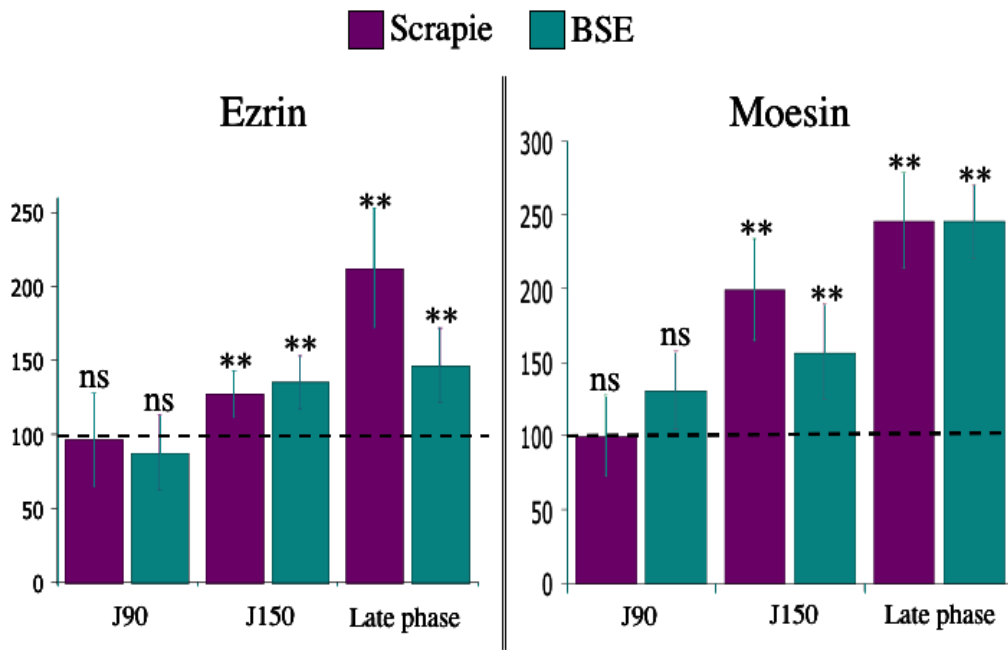
Phase 1: Initial model.



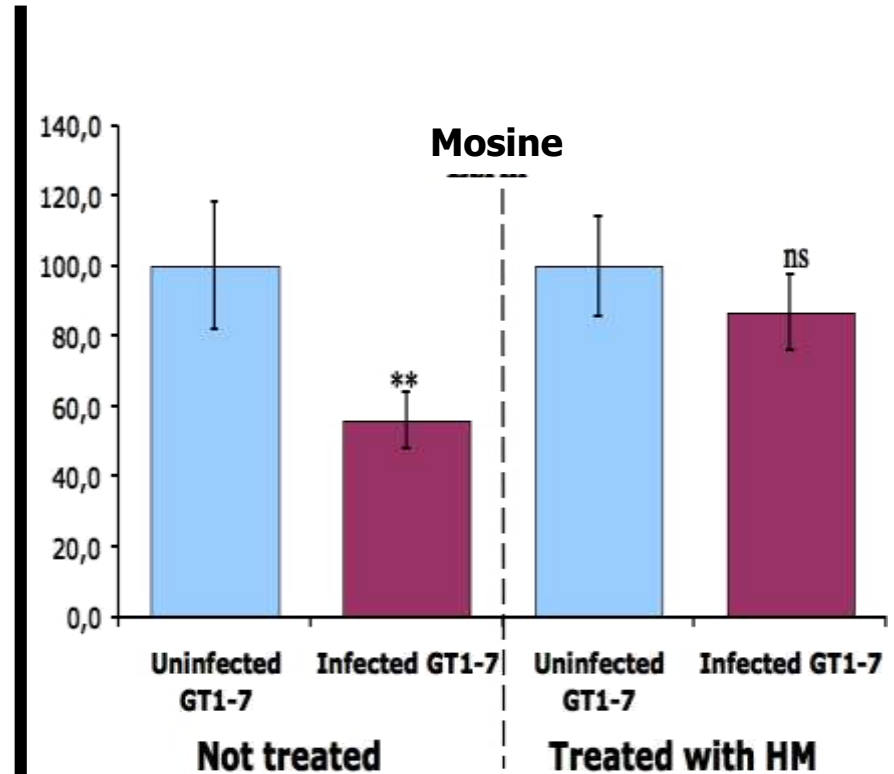
Objectives

- **Three targets:**
 - Ezrine
 - Moesine
 - Serine Racemase

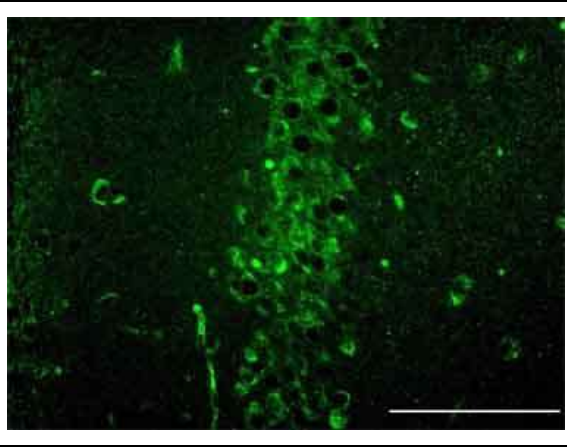
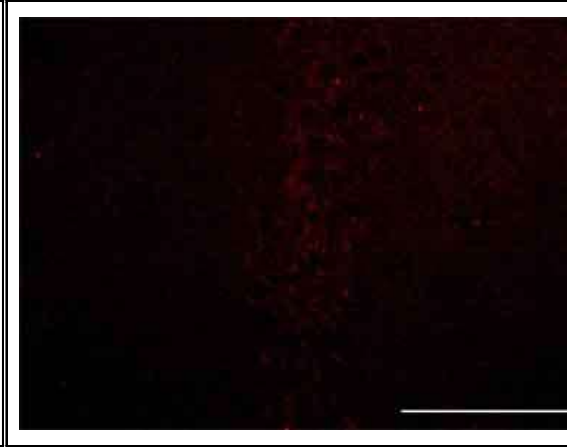
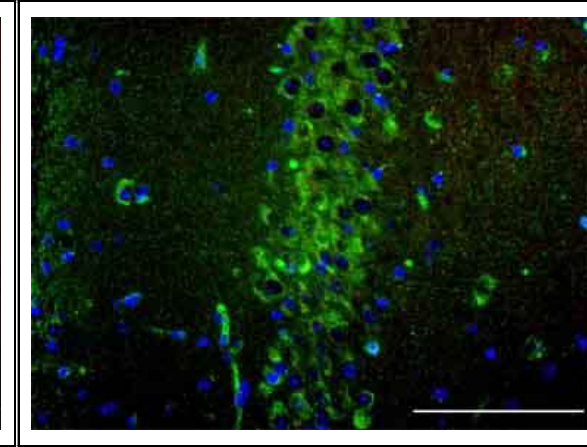
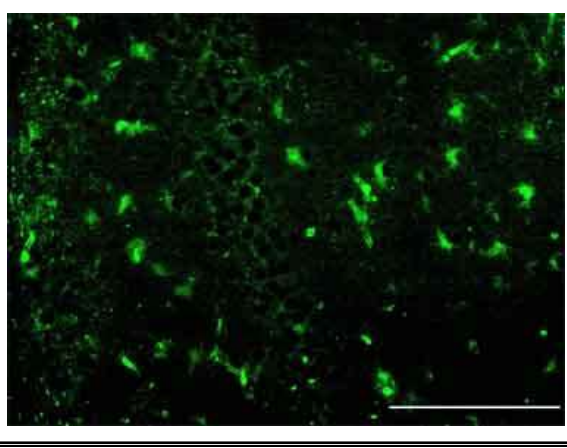
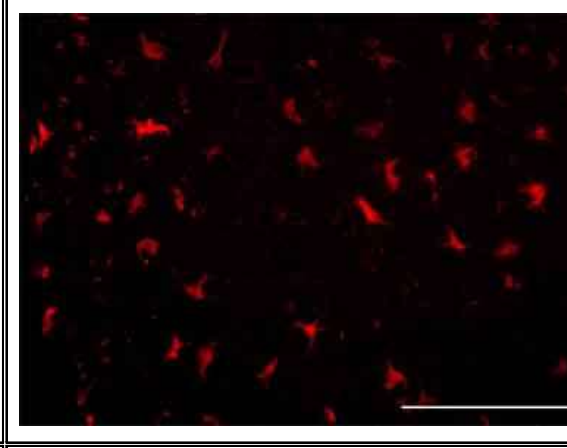
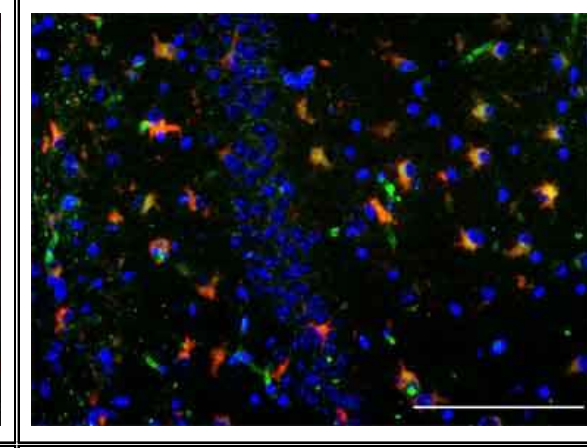
- **Two complementary approaches:**
 - Patterns of gene expression
 - Localisation in tissues & cells



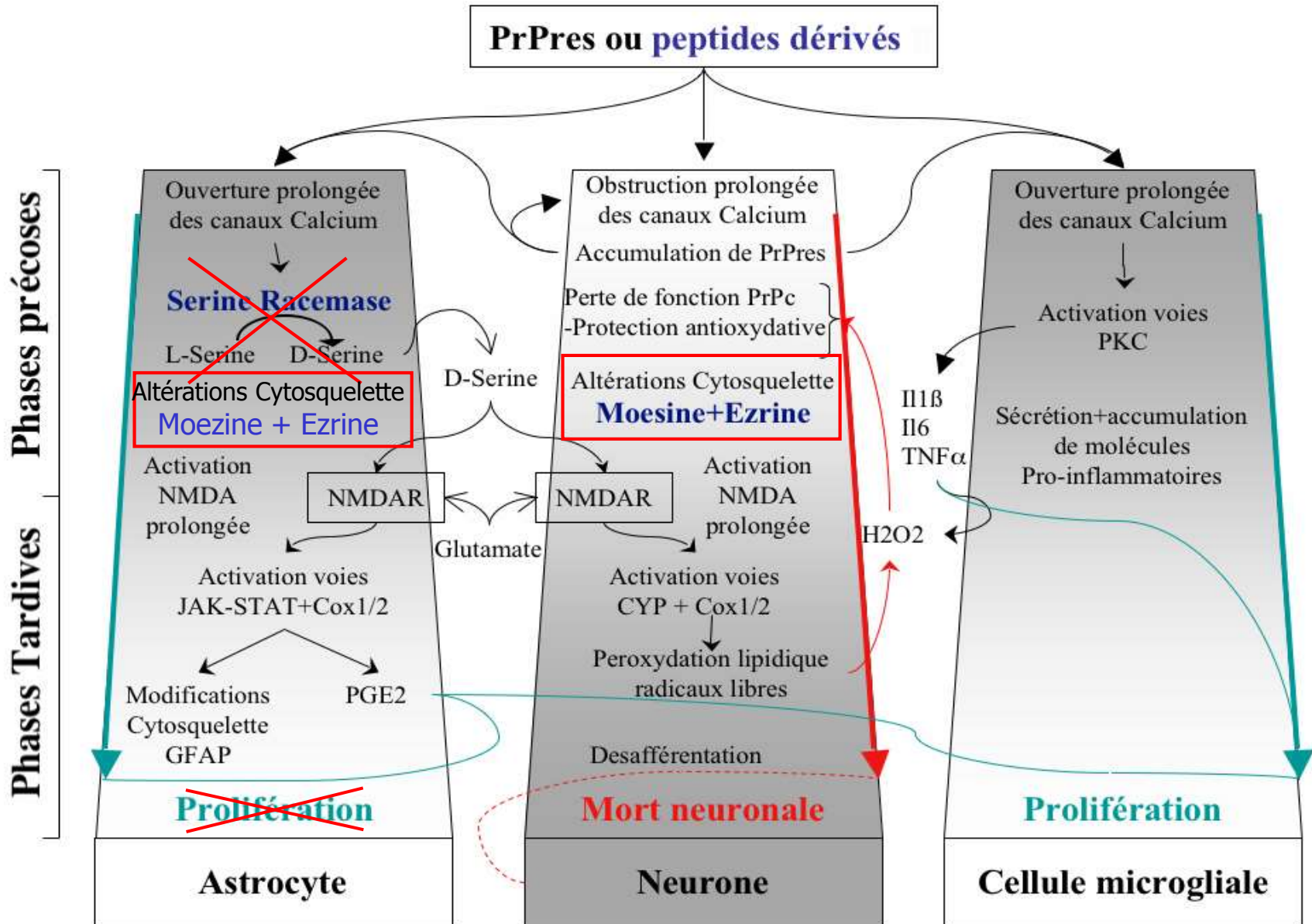
In vivo Ezrin and Moesin levels are increased in the whole brain during the course of PrPres infection & propagation.



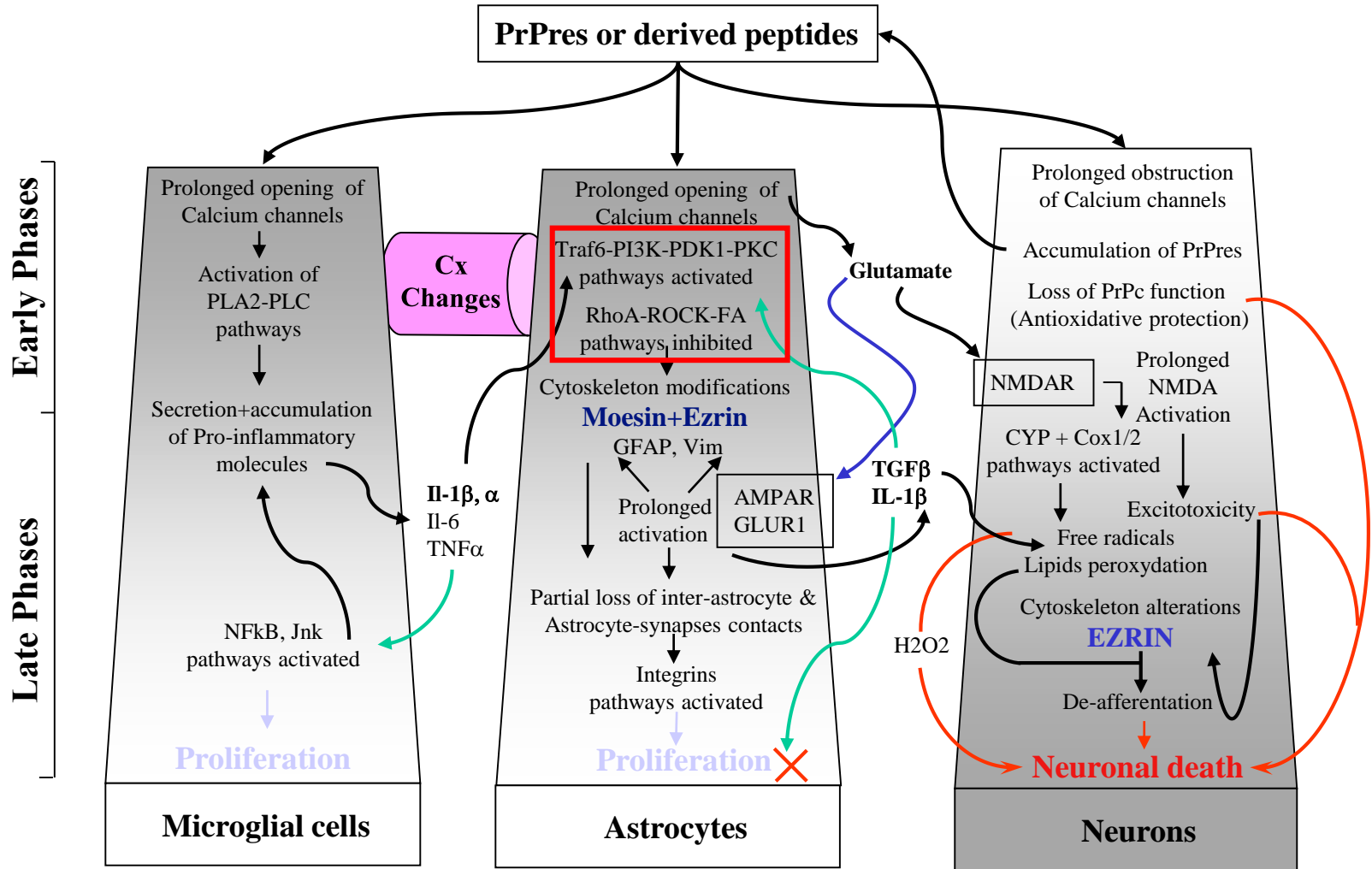
Decrease in Ezrin levels specifically associated with PrPres infection (infected GT1-7 neuronal cells cultures).

Control Mouse			
Infected Mouse			
Moésine Alexa 488 nm	Vimentine Alexa 568 nm	Superposition Moésine-Vimentine-dapi	

Validation outcome.



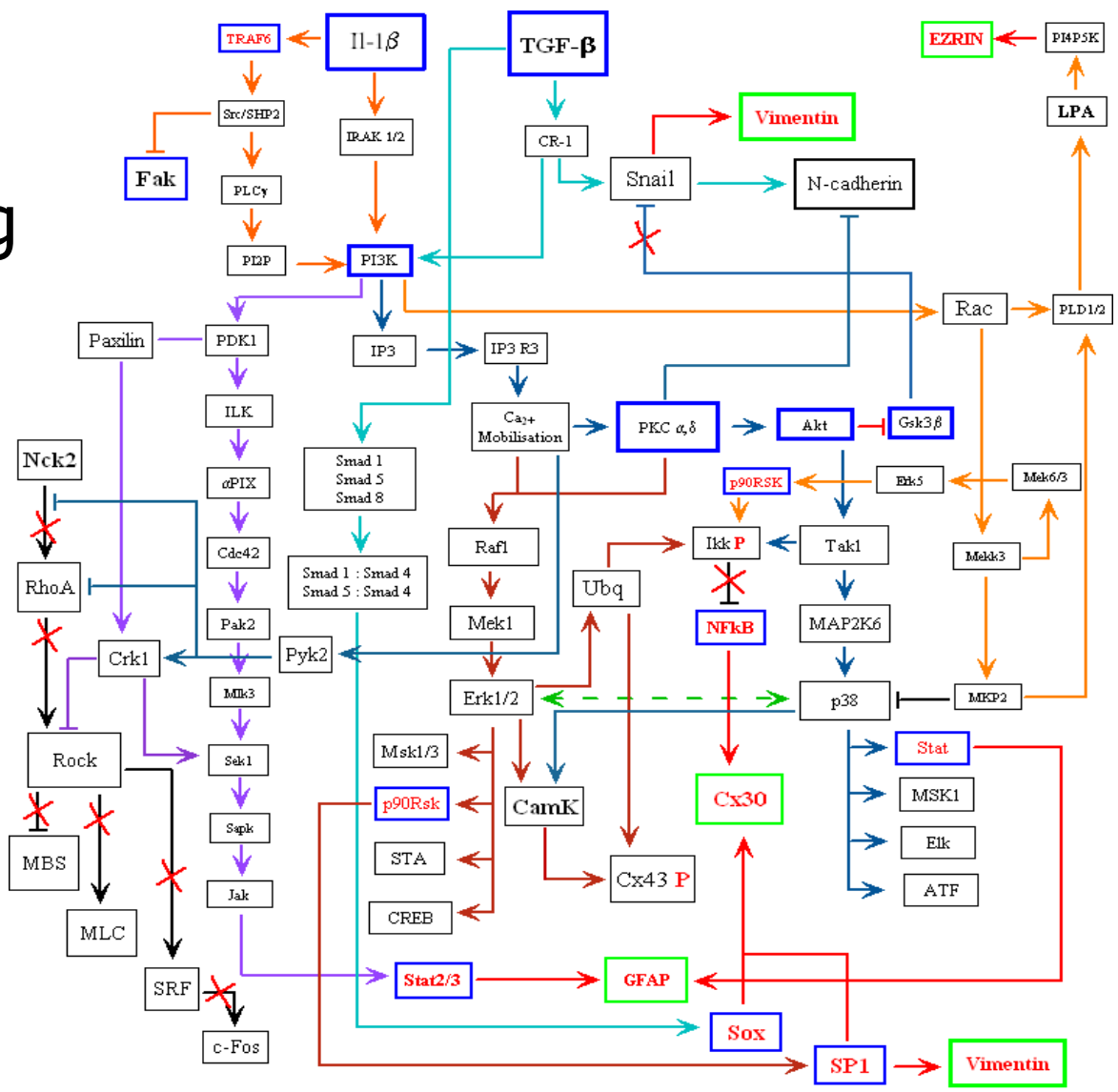
Phase 2: Revised model



The main driving mechanism.

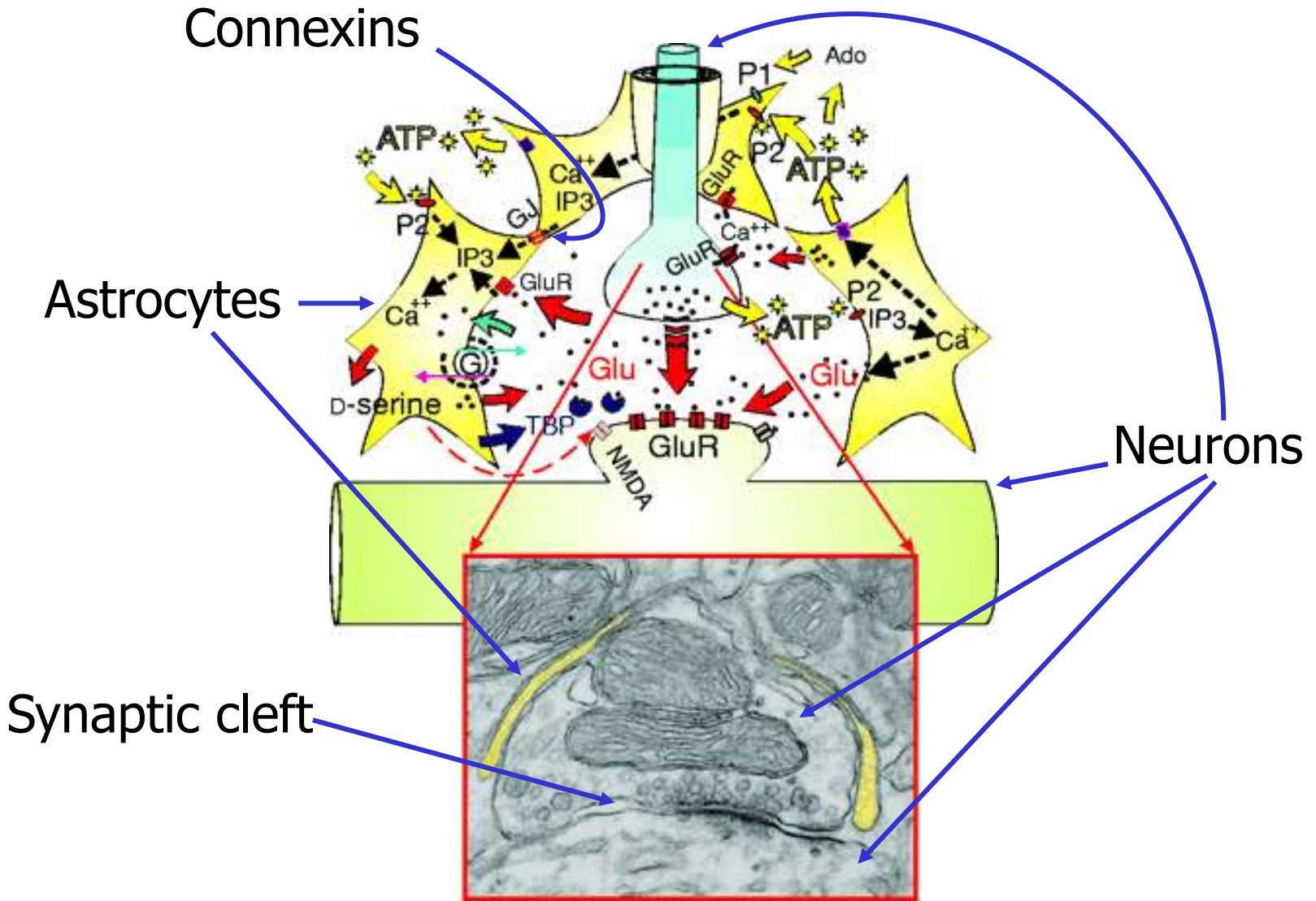
Il-1 β & TGF β -mediated signalling in hippocampus astrocytes

The pathways through which chronic neuronal stress signalling and concurrent glial pro-inflammatory responses lead to reactive astrocyte activation (GFAP + Vim) associated with cytoskeleton reorganisation (ezrin) associated with cytoskeleton reorganisation (ezrin). This leads to a major switch in Cx targeting & distribution, resulting in the formation of an astrocytic syncytium with massively altered diffusive properties and neurotrophic functions.



TGF β -driven response	\rightarrow (light blue arrow)
Il-1 β -driven response	\rightarrow (orange arrow)
TGF β response Il-1 β -modulated	\rightarrow (dark blue arrow)
Il-1 β response TGF β -modulated	\rightarrow (purple arrow)
Balanced Il-1 β -TGF β responses	\rightarrow (red arrow)
Equilibrating cross-talk	\rightarrow (dashed green arrow)
Key end-point responses	\rightarrow (red arrow with box)

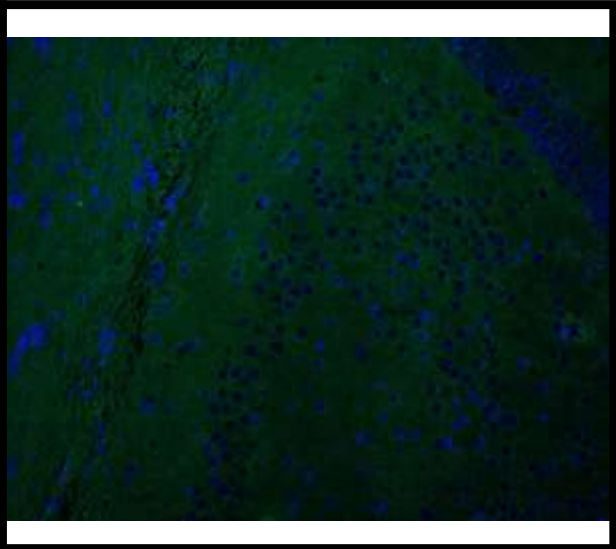
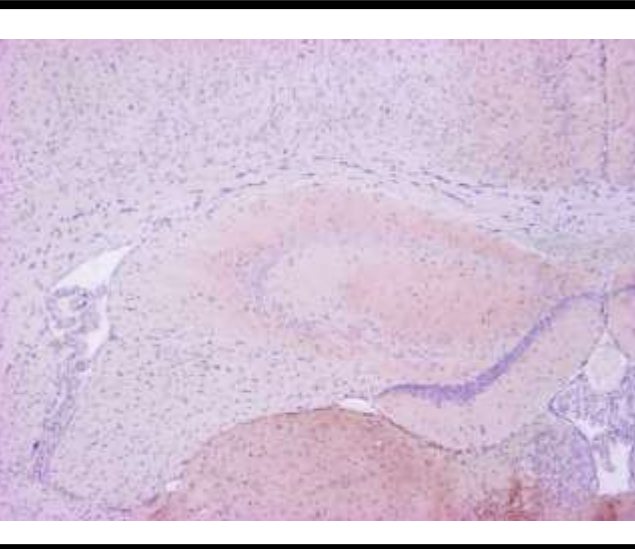
Synapses, astrocytes & connexins.



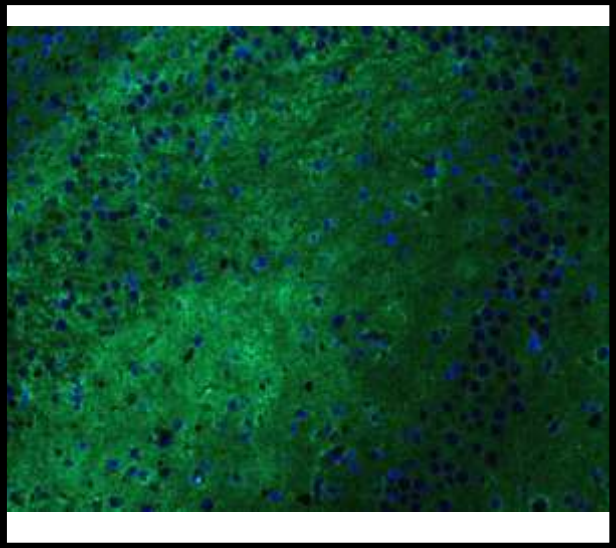
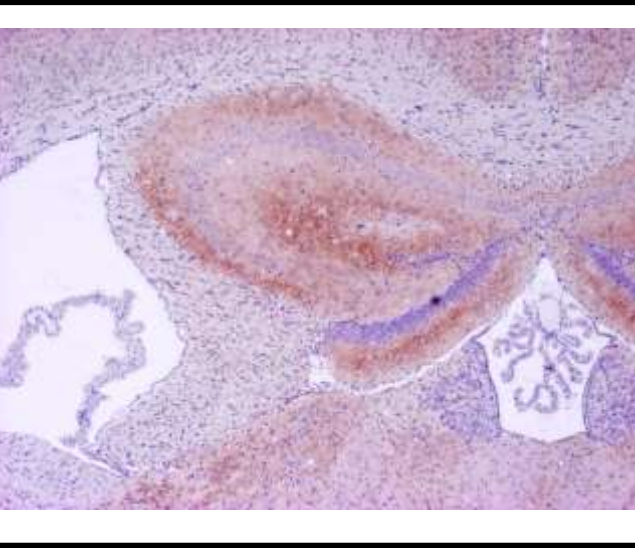
Immunohistochemical Evaluation

Connexins

Non Infected mice



Infected mice



Half brain Connexin 30

Hippocampus Connexin 30

Hippocampus (CA3) Connexin 30

Evaluation of intercellular diffusion processes associated with modifications of glial connexions.

Healthy hamster

Stereotaxic Injection of 500nl fluorescent traceur < 1000d



280µm



Scrapies hamster

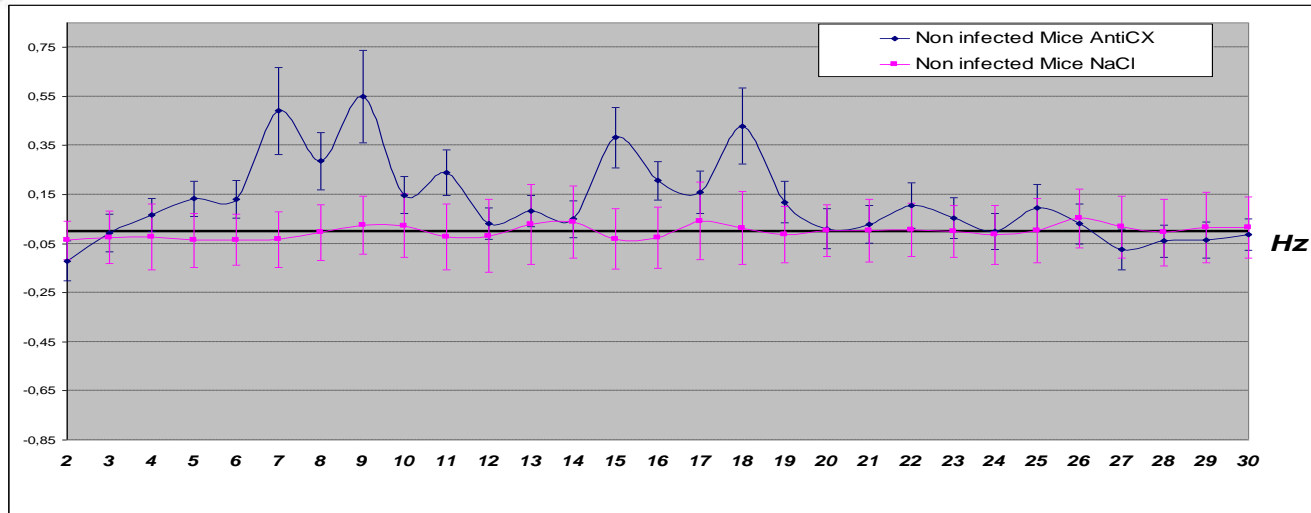
Stereotaxic Injection of 500nl fluorescent traceur < 1000d



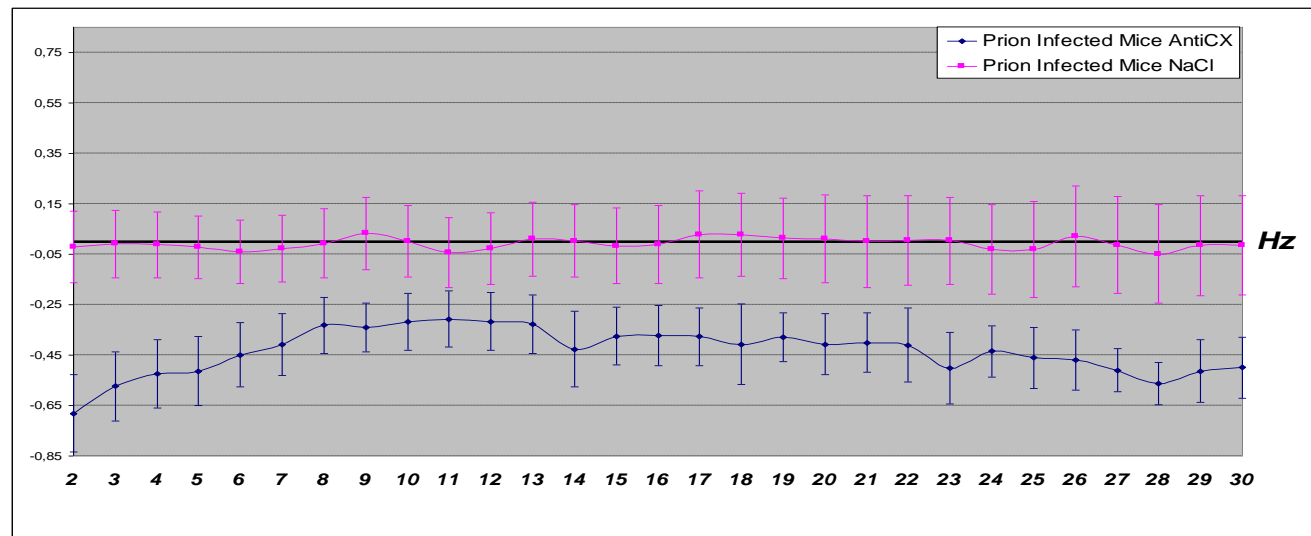
1080µm



Functionnal alterations linked with modifications of glial connections.



Healthy mice



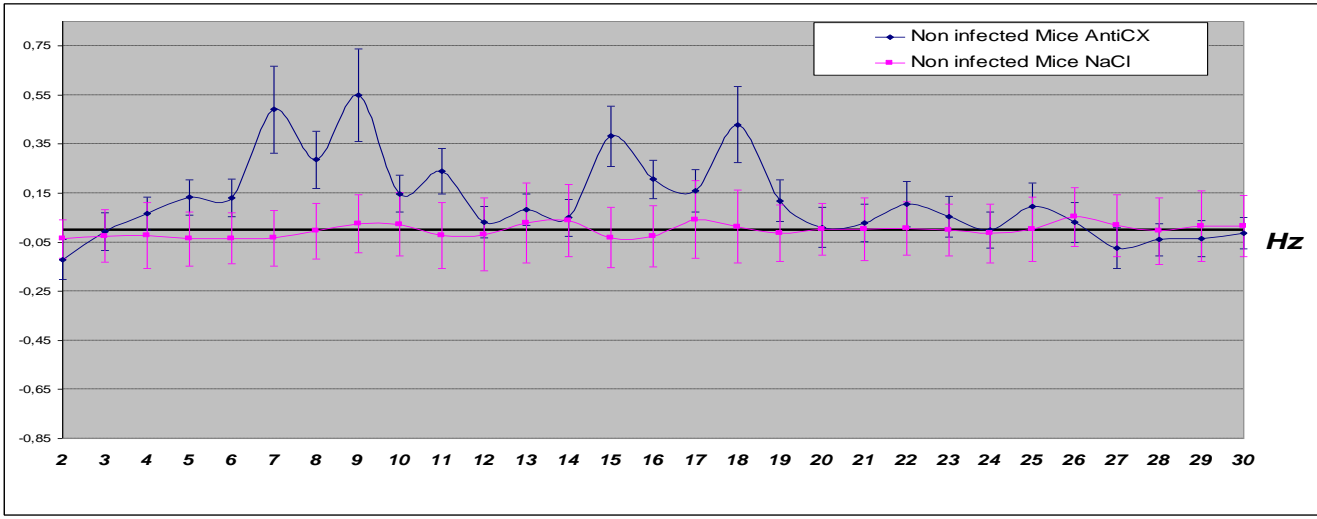
Scrapies mice
(asymptomatic phase).

■ NaCl

■ Anticonnexin

- 1- PrPres of external origin interact with neuronal PrPc, producing endogenous PrPres.
- 2- PrPres fragments (endogenous origin) induce auto-infection waves of ever increasing amplitude, leading to internal neuronal dysfunction, chronic neuronal stress signalling & local microglial activation.
- 3- The activated microglia secretes IL-1b, inducing TGFb secretion in astrocytes.
- 4- Paracrine and autocrine responses of astrocytes to IL-1b & TGFb induces local astrocytes activation, together with a radical shift in Cx expression and trafficking.
- 5- This induces the local constitution of a 3D activated astrocyte sheets within which Ca²⁺ waves propagate in an aberrant manner, leading to local failure of neuro-supportive activities.
- 6- Distant activation foci slowly merge, forming a continuously growing activated astrocyte sheet constitutively secreting pro-inflammatory and cytotoxic molecules while disseminating the infective agent (slow process).
- 7- Ever more uncoupled neuronal activities exacerbates neuronal stress signalling, reinforcing glia-mediated astrocyte activation & creating a vicious circle of highly toxic responses. These slow processes correspond to approx 2/3ds of the asymptomatic phase.
- 8- Healthy neurons and glial cells are killed at ever increasing frequencies (spongiosis) through bystander effects: clinical symptoms appear (failure of internal functional redundancies, abnormal, ever deteriorating EEG patterns) followed by death within a few months.

One of the roles of connexins is to dampen neuronal synchronisation.

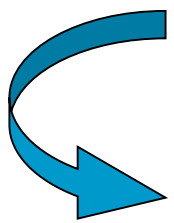


Healthy mice

■ NaCl

■ Anticonnexin

In healthy animals, pharmacological blockade of Cx activity results in quantitative EEG patterns closely resembling an epileptic crisis (frequency range-specific hyper-synchronisations).



This CEA/BMSystems collaborative research in CNS (psychiatric and neurological disorders) led to the co-owned patent WO201029131 (Use of anti-connexin agents for modulating the therapeutic effect of psychotropic drugs). Sept, 2008.

The net results.

CJD is not a neurological disease *stricto sensus*.

It is a disease that primarily affects astrocytes structures and functions which, over time, lethally affects healthy glial & neuronal cells through « bystander effects », leading to widespread CNS disorganisation (spongiosis) and functional failure.

But this model also provides an understanding of key mechanisms associated with psychiatric & neurology disorders.

An entirely new approach for their effective treatment was designed, tested in vivo and validated.



Patent covering novel therapeutics for psychiatry & neurodegenerative disorder (CEA/BMSystems).

This CEA/BMSystems [collaborative research in CNS](#) (psychiatric and neurological disorders) led to the co-owned patent [WO201029131](#) with a worldwide exclusive license to [Theranexus](#) CEA's spin-off currently in Phase II.

Neither of which have much to do with CJD *per se*...

This work received a Bio-IT World « Best Practices » award from the Cambridge HealthTech Institute (USA).



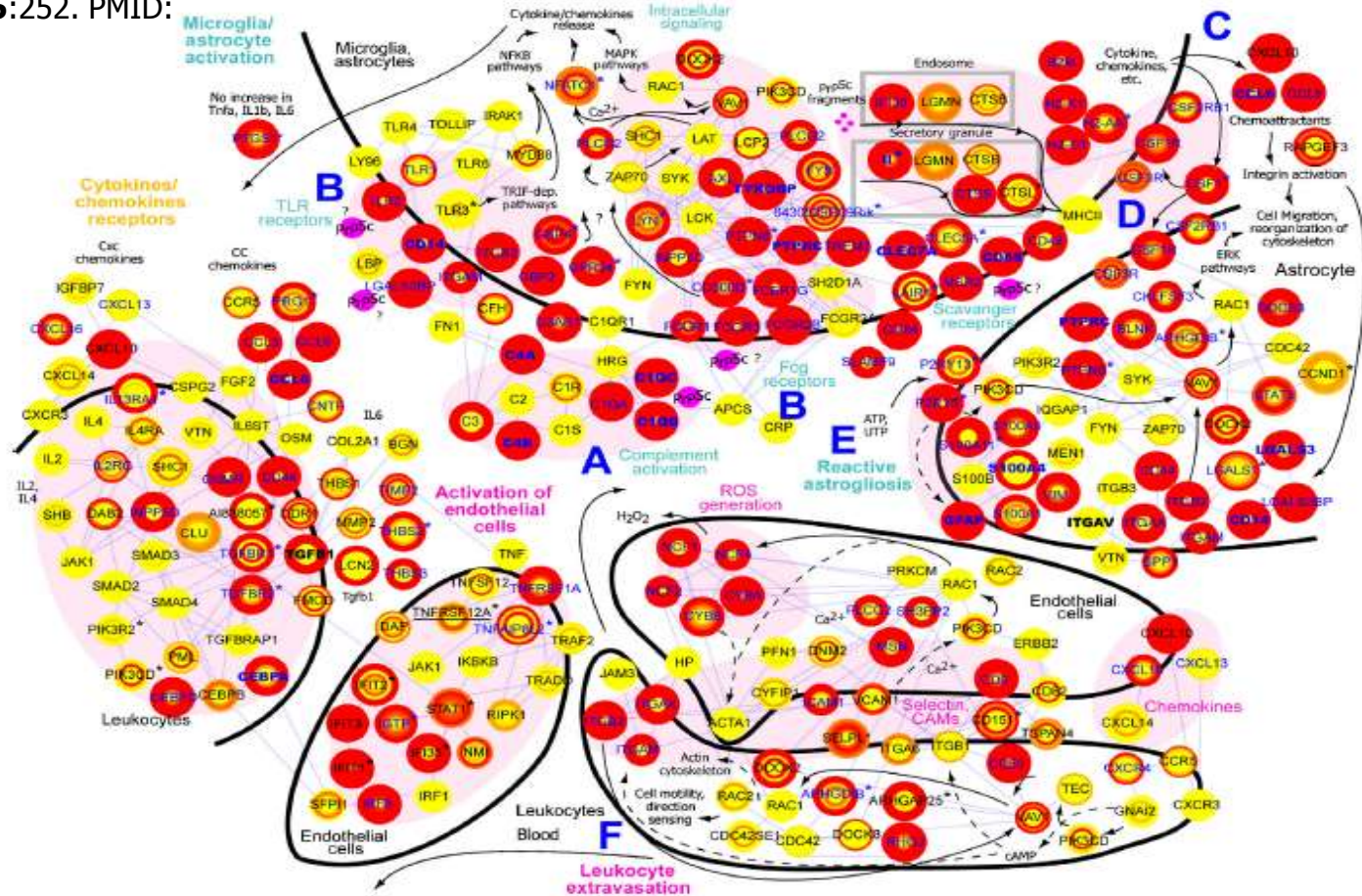
AND

Was selected as 1 of the 3 pan-European « state of the art examples of systems biology approaches of benefit to medicine » by the European Commission's DG Research, Directorate of Health (June 2010).

Analytical approaches: what must be avoided at all costs.

This Bayesian model of CJD was published in 2009 (Hwang D et al. Mol Syst Biol. 5:252. PMID: 19308092)

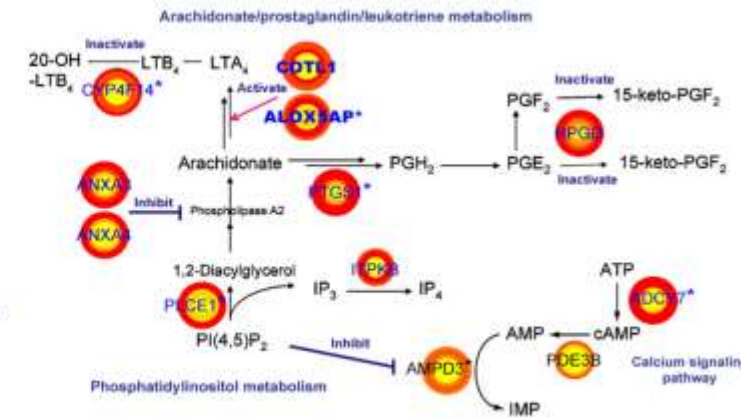
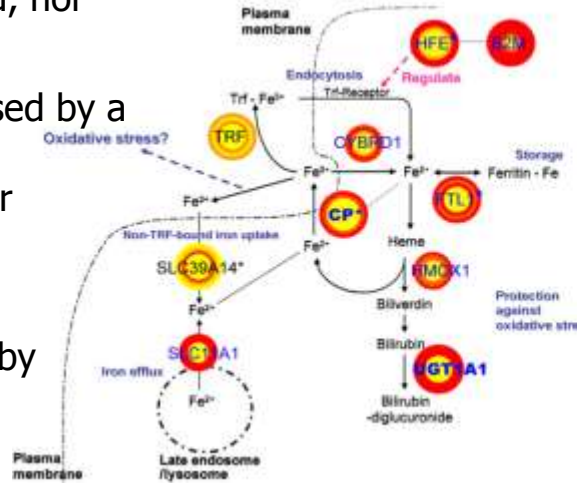
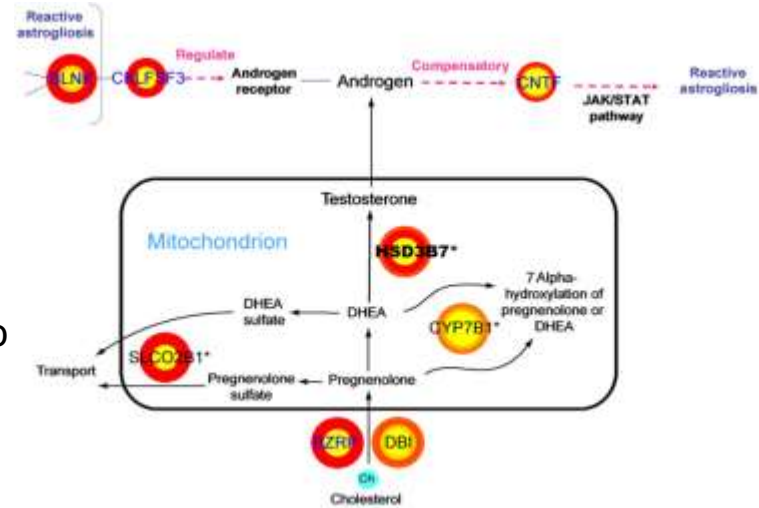
Constructed from a list of 333 shared DEGs together with protein-protein interactions information from novel targeted experiments and public databases, it describes the networks that are potentially involved in the activation of microglia and astrocytes.



Although it certainly does lead to sub-networks that do make some sense,

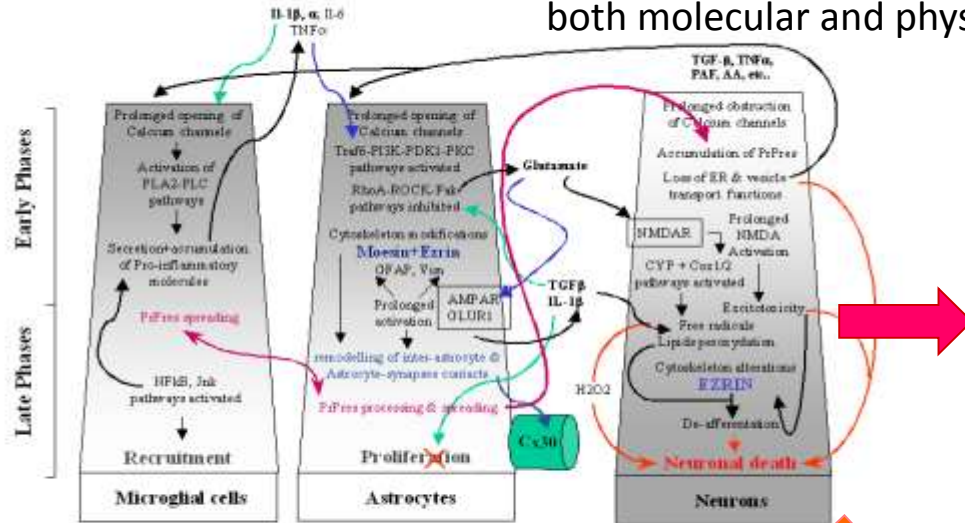
These cannot

- 1) be distinguished from Alzheimer's Disease, also characterised by astrogliosis, nor
- 2) begin to explain why neurons are killed in the CNS and never in the spinal cord, nor
- 3) why the disease is characterised by a spongiosis absent in other neurodegenerative disorders, nor
- 4) why the disease progresses silently for many years followed by a sudden very rapid clinical progression.

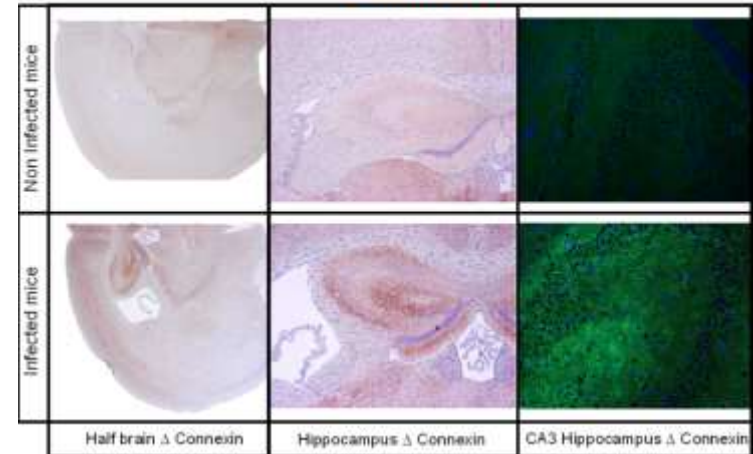


As compared to the heuristic model of CJD finalised in early 2008

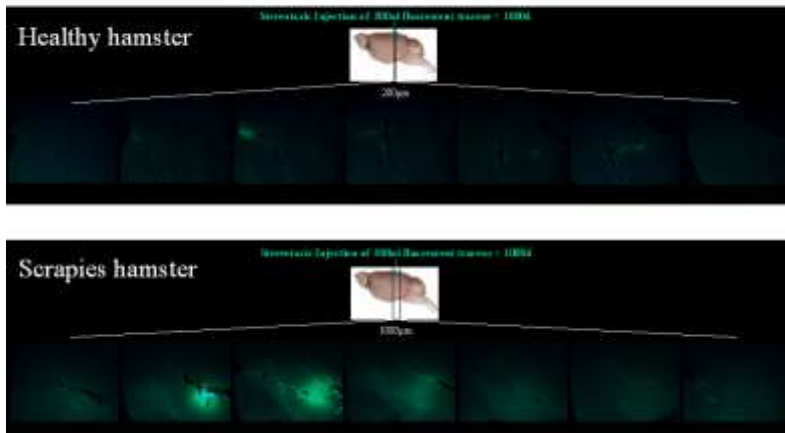
Which predicts and explains the pathological mechanisms at both molecular and physiological levels.



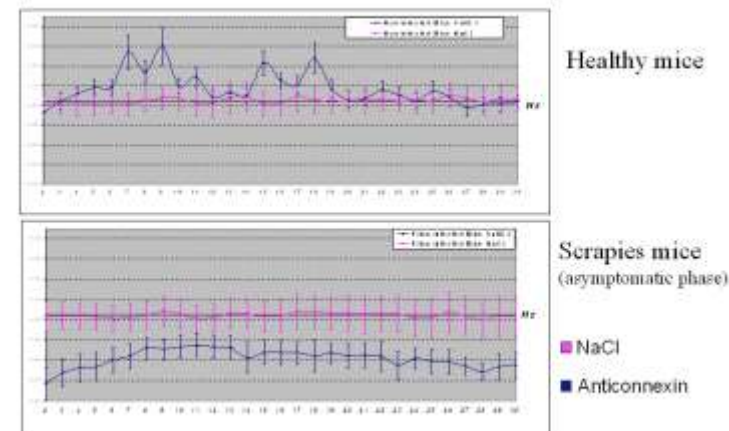
Immunohistochemical Evaluation Connexins



Functional modifications of glial connexions



Functional EEG alterations linked with modifications of glial connections.



Source:
BMSys & CEA

Modelling living systems?

How?

Modelling living systems...

is first and foremost a matter of integrating huge masses of information.

But there are first THREE problems & ONE paradox to resolve.

Problem 1. The nature of information

All that goes by the name « *Information* » is not necessarily **Useful** and/or **Utilisable!**

In the bio-medical realm, the information available is **ALWAYS:**

- **incomplete**, to an unknown extent;
- **biased**, to an unknown extent; and
- **erroneous**, to an unknown extent.

This has very serious consequences!



e.

The same place under 2 different names (one better known than the other)?

The same residing entities?

The same functional characteristics?

IMPORTANT TO UNDERSTAND: For non-french citizens: In France the "Palais de l'Élysée" is the house of the President of the State and [Theatre de Guignol](#) is a popular french muppet show for children and adults

The same place?



Well, there are
apparent
similarities.

The same residing entities, passed & present?



There too, there are
apparent similarities....

*TO UNDERSTAND: The muppets of our political
leaders in the [Guignols de l'info](#) a famous
muppet show on TV*

The same functional characteristics?



Yet again, more
apparent
similarities....

What happened?

The initial information is incomplete and its presentation led to interpretations;



These interpretations induced biased searches;

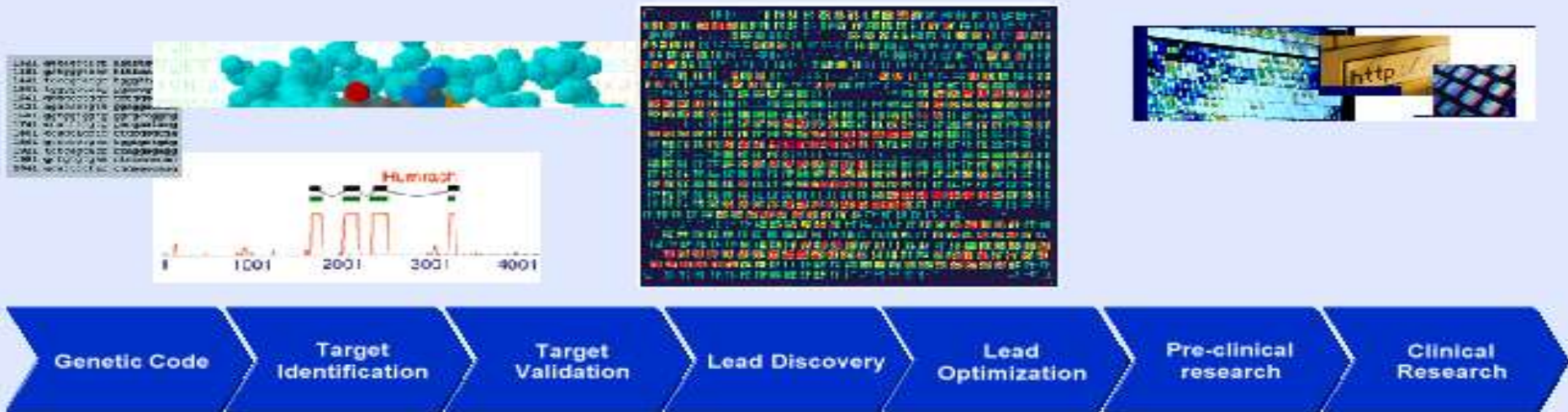


And these bias led to **erroneous conclusions**, making the distinction between appearances and reality ever more difficult to discern.



We have the very same problem in biology!

This takes particular significance in drugs development



- ❖ Sequence analysis
- ❖ Annotations
- ❖ Comparative genomics
- ❖ Pharmacogenomics

- ❖ Functional genomics
- ❖ Expression analysis
- ❖ Proteomics
- ❖ Structural genomics
- ❖ Pathway discovery

- ❖ Expression analysis (arrays)
- ❖ Proteomics
- ❖ Structural genomics
- ❖ Pathway discovery
- ❖ Pharmacogenomics

- ❖ Expression (arrays)
- ❖ Proteomics
- ❖ Pathway discovery
- ❖ HTS

- ❖ Combi. chemistry
- ❖ Expression (arrays)
- ❖ Proteomics
- ❖ Structural genomics

- ❖ Animal models
- ❖ In silico simulations
- ❖ Expression (arrays)
- ❖ Proteomics
- ❖ Pathway discovery

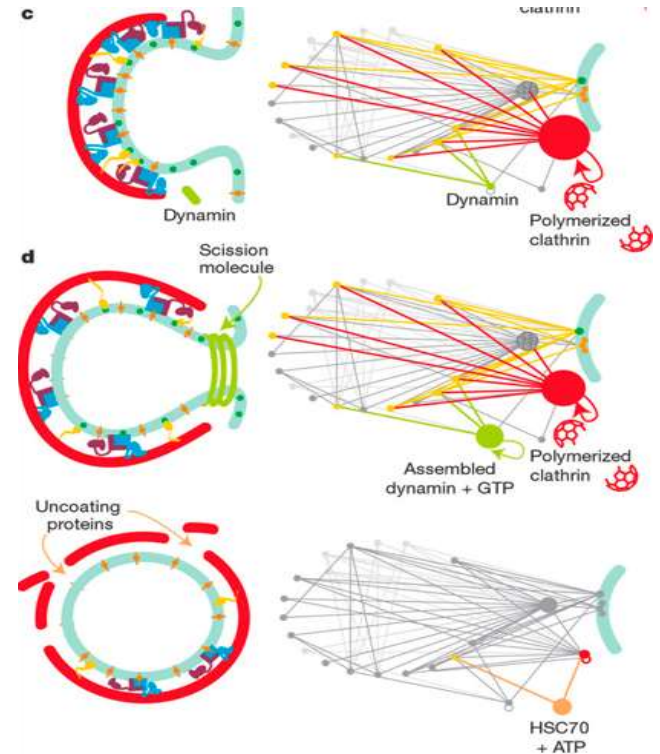
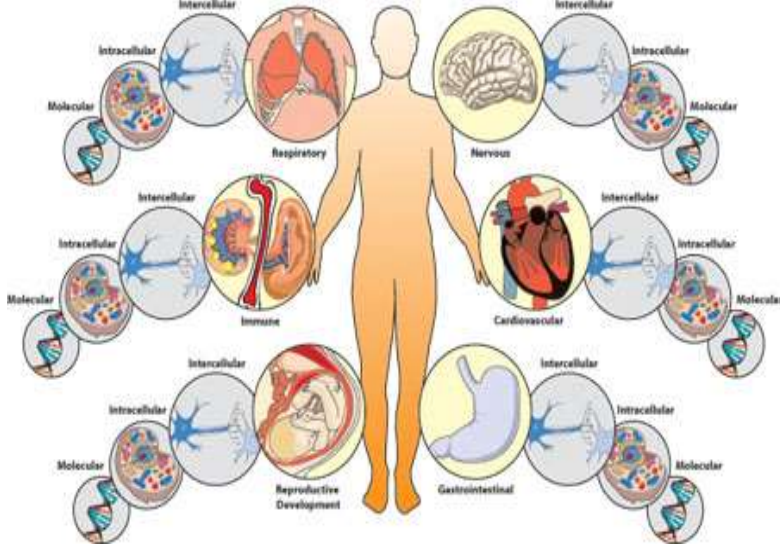
- ❖ Expression (arrays)
- ❖ Pharmacogenomics

A process generating a flood of heterogeneous information

Therapeutic success.

The success of a therapeutic approach largely arises from the coherent manipulation of a physiological system as a whole

and not from that of a target in a molecular context.

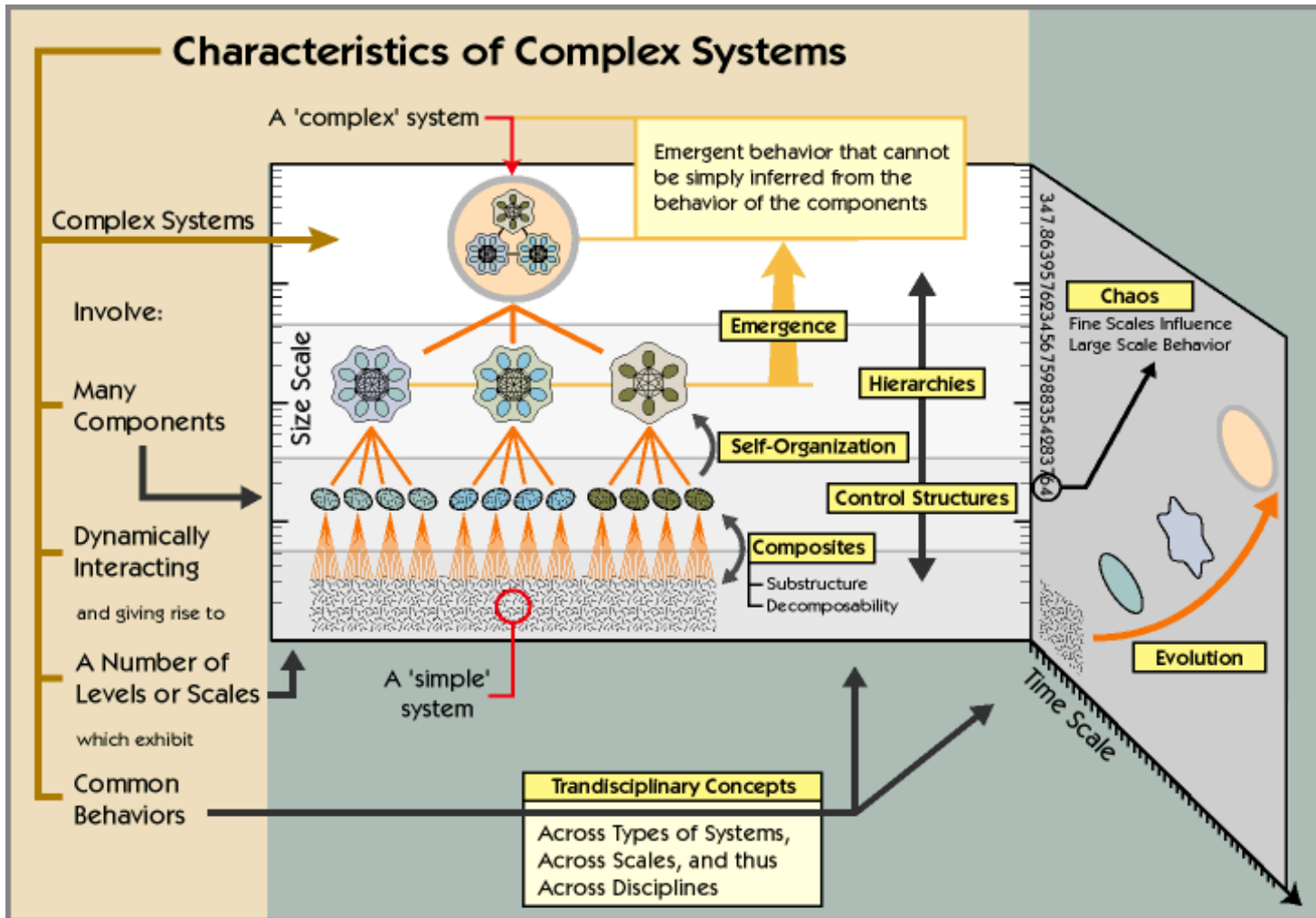


Therefore, medical problems are best approached from a systems standpoint.

However, the available information is always heterogeneous AND incomplete....

Problem 2: The nature of the system.

Living Systems: the properties of ensembles cannot be deduced from those of components!



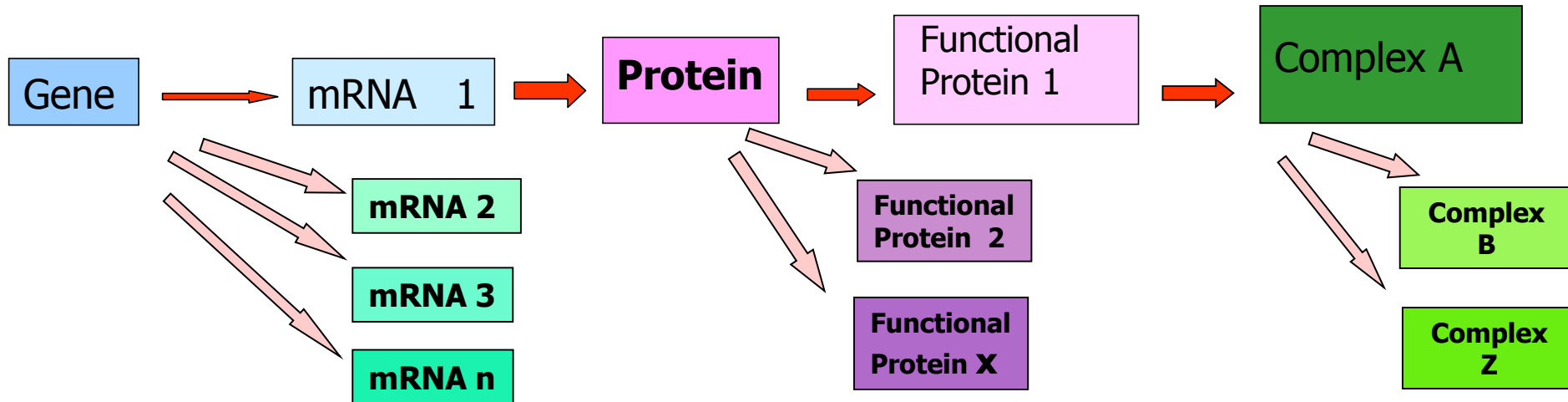
Multiple scalar levels concurrently present.

Problem 3. the behaviours of the system.

From genes to physiological functions:

Three series of **deconvolutions** and **discontinuities**

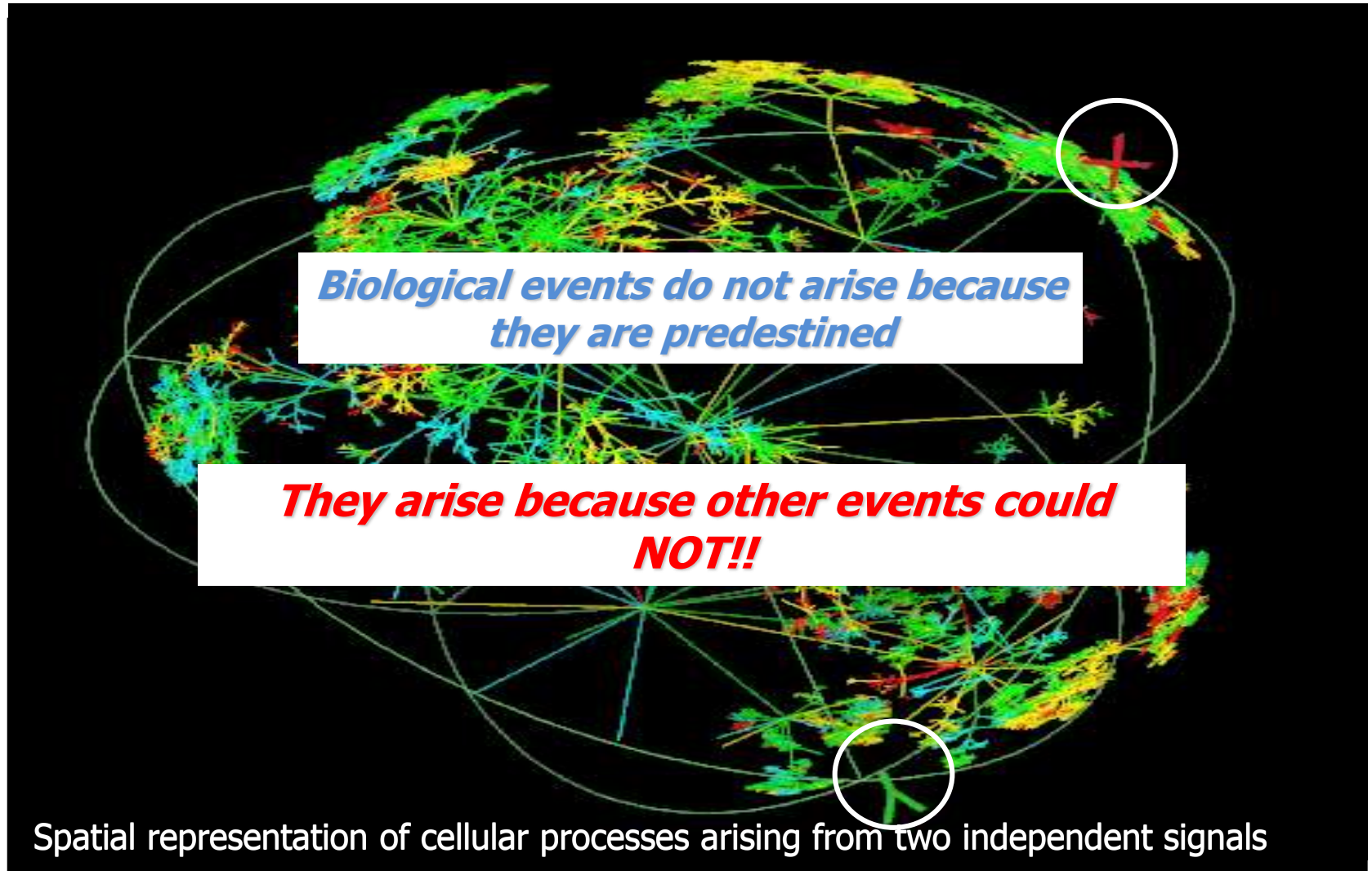
One gene = several different physiological functions



A complex, non-linear & integrative system.
(27000 genes → over 1 000 000 protein functions)

The behaviours of the system.

At a functional level, what abortes does matter!



The biological effects result from the differential integration of ALL these events.

The paradox that must be resolved.

If you dream to create the first operational model of a bird...

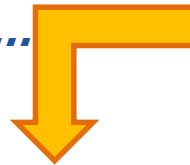


... a "basic" living Complex system that not only flies...

Be sure to use the appropriate modeling concepts & tools. If not...

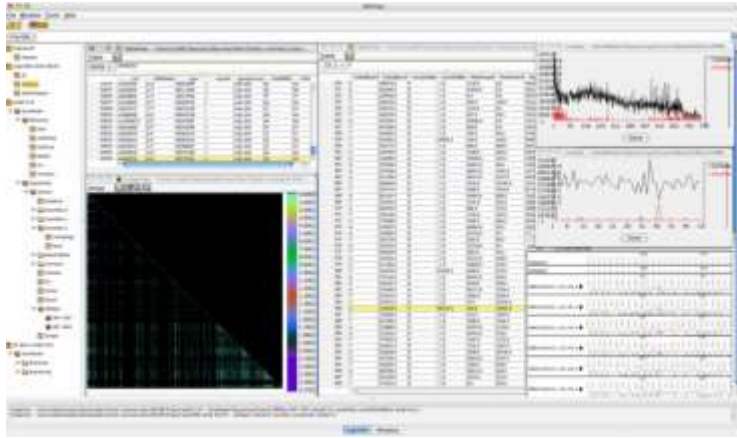


...you'll get a Complicated "Cartesian" system. It does fly, but...

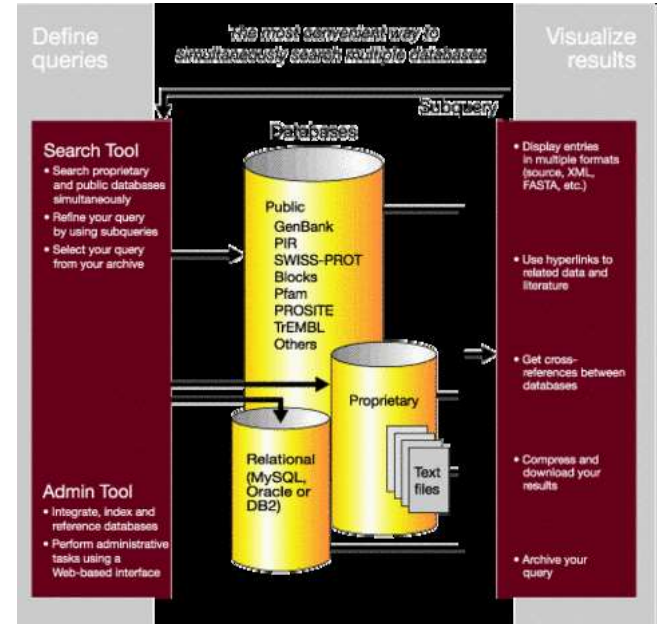
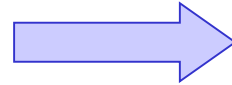


According to this "model": 1) birds carry highly explosive eggs; 2) when a bird flies, its rear end becomes extremely hot; 3) to land safely, a bird requires three legs; etc...

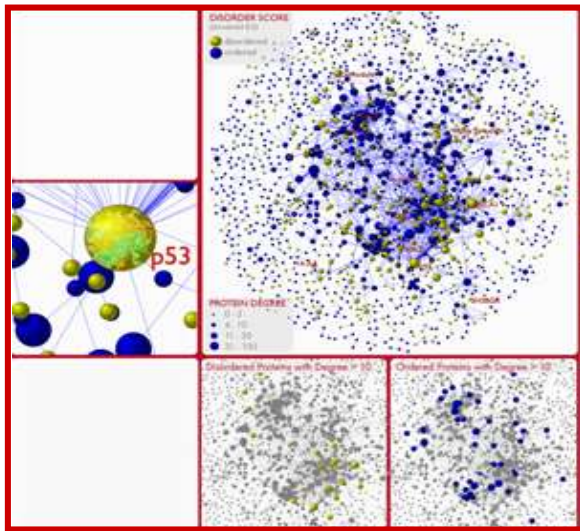
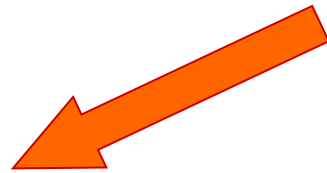
Yet, to analyse the data & build a model we must all utilise similar tools.



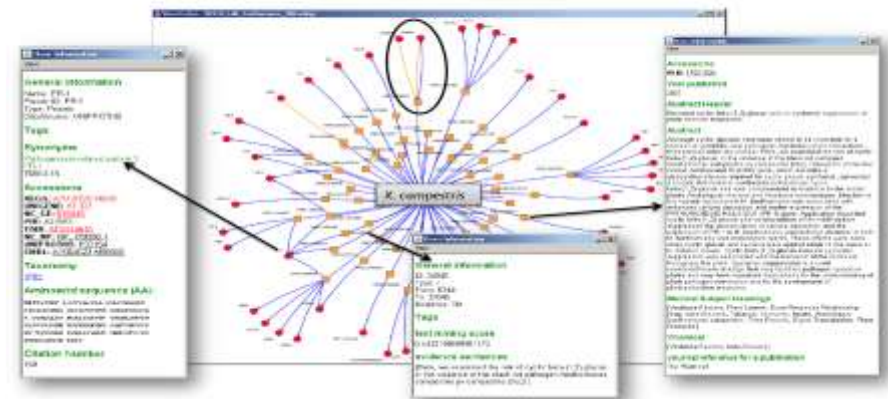
Data flat-file



Analytical WEB resources



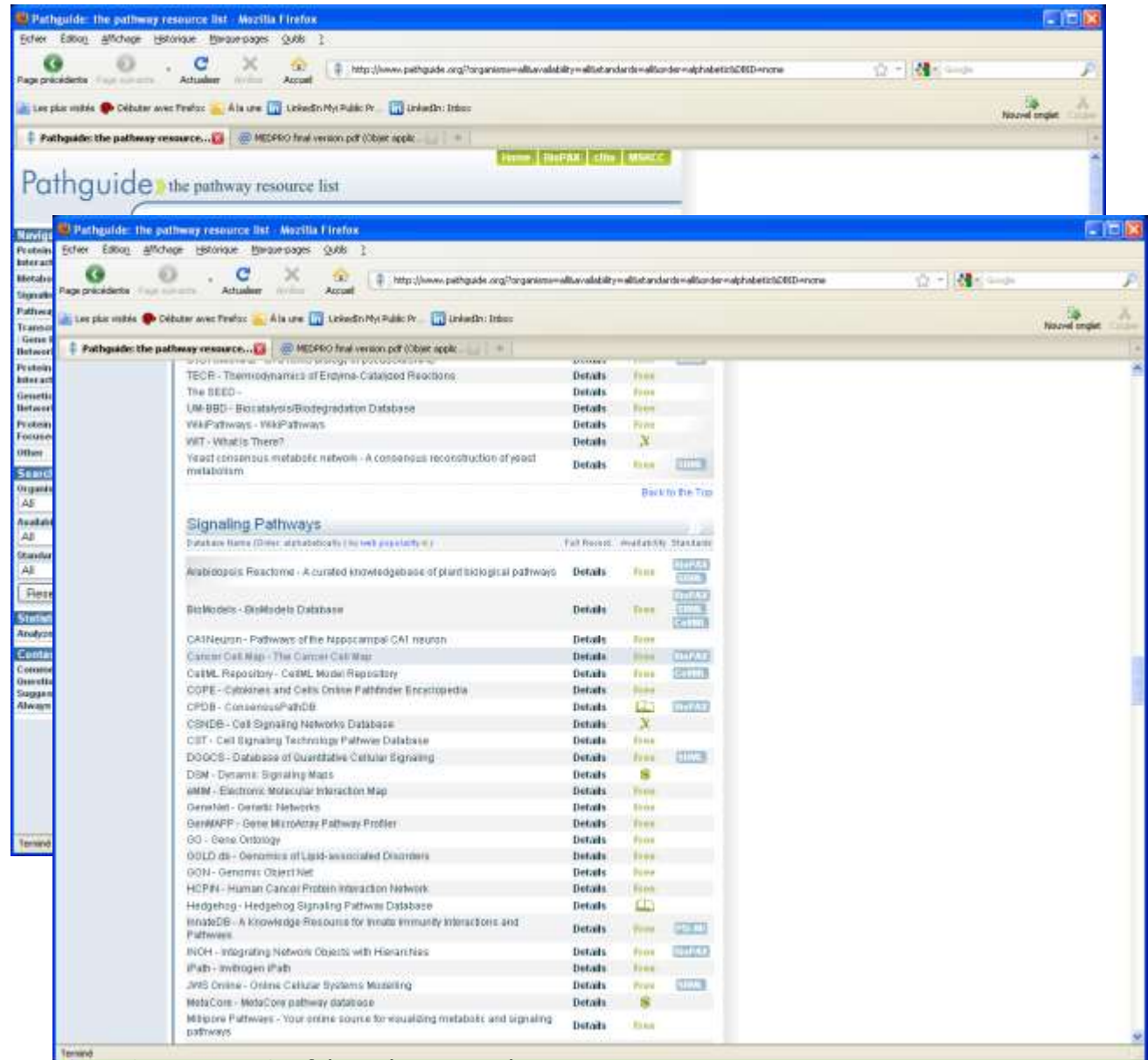
Visualisation tools



Networks construction

Currently, some **317** web resources (databases) provide access to

- Thousands of pathways and networks, documenting
- Millions of interactions between proteins, genes, small molecules..



(<http://pathguide.org/>)

In practical terms, it means extracting as much information as possible from archived records.

UniProt > UniProtKB

Search Blast * Align * Retrieve ID Mapping *

Search in: Protein Knowledgebase (UniProtKB) Query: ACBP4 Search

★ Reviewed, UniProtKB/Swiss-Prot **Q9MA65** (ACBP4_ARATH)
 Last modified March 2, 2010. Version 58. [History...](#)

Ontologies

Keywords

Biological process	Transport
Cellular component	Cytoplasm
Coding sequence diversity	Alternative splicing
Domain	Coiled coil Kelch repeat Repeat
Ligand	Lipid-binding
PTM	Phosphoprotein
Technical term	Complete proteome

Gene Ontology (GO)

Biological process	lipid transport [Set 4] Inferred from direct assay. Source: TAIR
	response to ethylene stimulus [Set 6] Inferred from expression pattern. Source: UniProtKB
	response to jasmonic acid stimulus [Set 6] Inferred from expression pattern. Source: UniProtKB
	response to light stimulus [Set 10] Traceable author statement. Source: UniProtKB

References

Show large scale references ▶

[4] "ACBP4 and ACBP5, novel Arabidopsis acyl-CoA-binding proteins with kelch motifs that bind oleoyl-CoA." Leung K.-C., Li H.-Y., Mishra G., Chye M.-L. *Plant Mol. Biol.* 55:297-309(2004) [PubMed: 15504682] [Abstract] [Cited for:](#) FUNCTION, MUTAGENESIS OF GLY-24, LEU-25, SER-28, LEU-45, TYR-48, GLN-52, LYS-74 AND PHE-83.

[5] "Ethylene- and pathogen-inducible Arabidopsis acyl-CoA-binding protein 4 interacts with an ethylene response." Li H.-Y., Xiao S., Chye M.-L. *J. Exp. Bot.* 59:3997-4006(2008) [PubMed: 18836139] [Abstract] [Cited for:](#) SUBCELLULAR LOCATION, INDUCTION BY ETHYLENE, JASMONATE AND BOTRYTIS CINEREA, TISSUE

[7] "Arabidopsis acyl-CoA-binding proteins ACBP4 and ACBP5 are subcellularly localized to the cytosol and ACB." Xiao S., Li H.-Y., Zhang J.P., Chan S.W., Chye M.-L. *Plant Mol. Biol.* 68:571-683(2009) [PubMed: 18773301] [Abstract] [Cited for:](#) FUNCTION, SUBCELLULAR LOCATION.

<http://www.uniprot.org>

PubMedID: [16797716](#)

Title: New and future **migraine** therapy.

Abstract:

Modern neuroscience advanced our understanding of putative **migraine** mechanisms, which led to improved therapeutics. Indeed, mechanism-based **acute migraine** therapy gained **steam** in the early 1990s after the introduction of the triptans (**5-HT_{1D}** agonists). Post-triptans, novel targets such as **calcitonin gene-related receptor (CGR)** antagonists, inhibitors of excitatory glutamatergic receptors, and **nitric oxide synthase (NOS)** inhibitors are leading the pack in this exploding field of discovery research. In contrast, novel therapeutic targets for **migraine** prevention are lacking despite a hugely unmet need. To date, **migraine** prophylactic drugs are advanced based on expanded indications for already approved pharmaceuticals (e.g., **topiramate**, **valproate**, **propranolol**, and **timolol**). An improved understanding of the predisposition to an attack, genomic discoveries, valid and reliable biomarkers and surrogates, and predictive preclinical models likely will unravel the neuronal substrates for central **hyperexcitability** and nociceptive dysmodulation, hopefully leading us to better mechanism-based targets for prevention, and ultimately yielding drugs with optimal therapeutic ratios or indices.

Legend:

GENE OR PROTEIN METABOLITE BACTERIA ORGAN SYMPTOM or DISEASE
 PHENOMENON PROCEDURE INDICATOR
 Acronym

Journal: *Pharmacol. Ther.* 2006;112(1):199-212

Author(s): [Ramadan_NM](#), [Buchanan_TM](#)

Mesh Heading(s): [Humans](#), [Migraine Disorders](#), [Migraine Disorders -- drug therapy](#), [Migraine Disorders -- prevention & control](#)

This data mining approach pre-supposes that

- There are well defined rules for expressing relationships between components and they are practically always obeyed by authors.

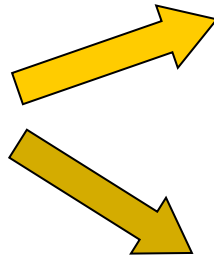
- **Distance of co-occurrence :**

$$\log \left(\frac{N_{\text{phr}}(G) \cdot N_{\text{phr}}(G')}{N_{\text{phr}}(G \text{ et } G')^2} \right)$$

- **Functional relationships**

$$N_{\text{phr}}(G_1, G_2, R)$$

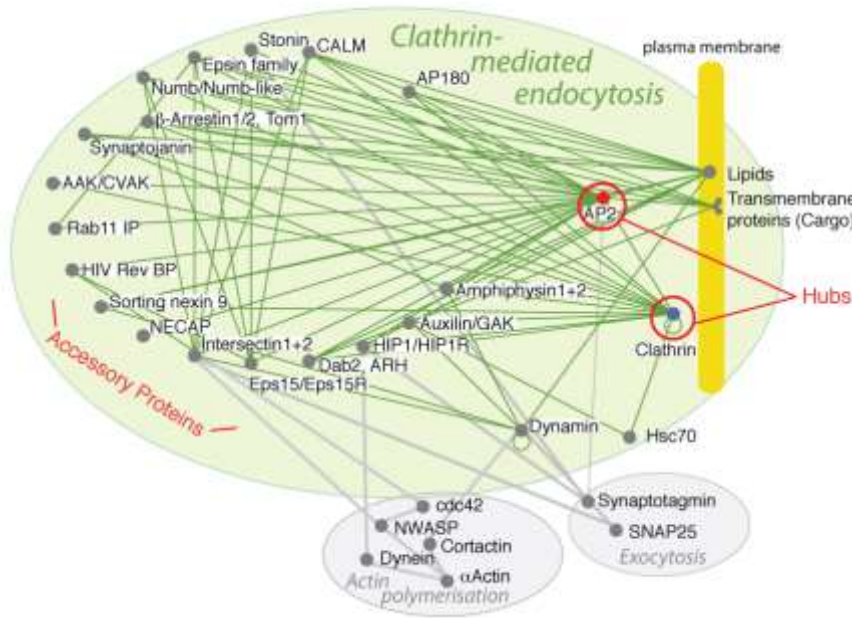
Hence the systematic application
of
mining algorithms
such as



to all 'omics data, from transcriptomics to proteomics & metabolomics

This generates vast amounts of spurious correlations requiring human intervention (curating).

It also leads to a highly misleading vision of protein interactions & networks.



How can single hub proteins bind so many different partners?

The problem is largely non-existent and resides in the construction and the representation of protein interaction networks within data-bases.

Proteins derived from a single gene, even if different, are clustered in maps into a single node.

This leads to the impression that a single protein binds to a very large number of partners.

In reality, it does not.

Protein networks reflect confusions involving combinations of functional plasticity addressing a same protein together with distinct physiological roles of different proteins encoded by one gene.

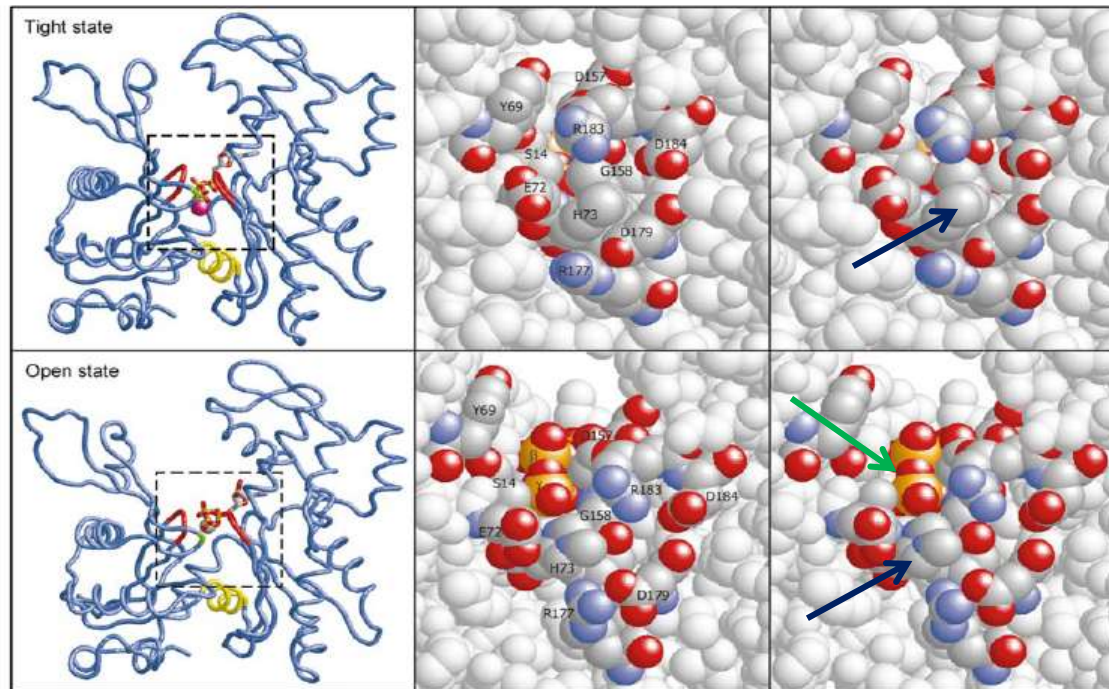
One protein = Functional plasticity

Post-transcriptional modification of Actin

Actin His 73 methylation = ATP exchange rate increased, ATP hydrolysis and phosphate release prior to and independent of filament formation;



Polymer formation delayed.

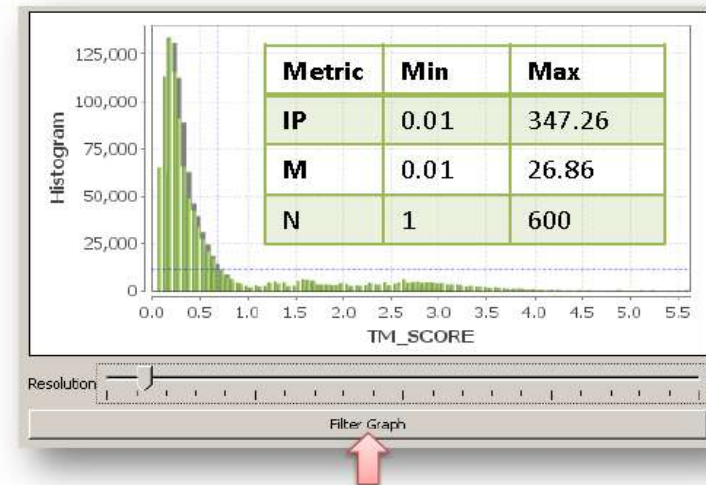
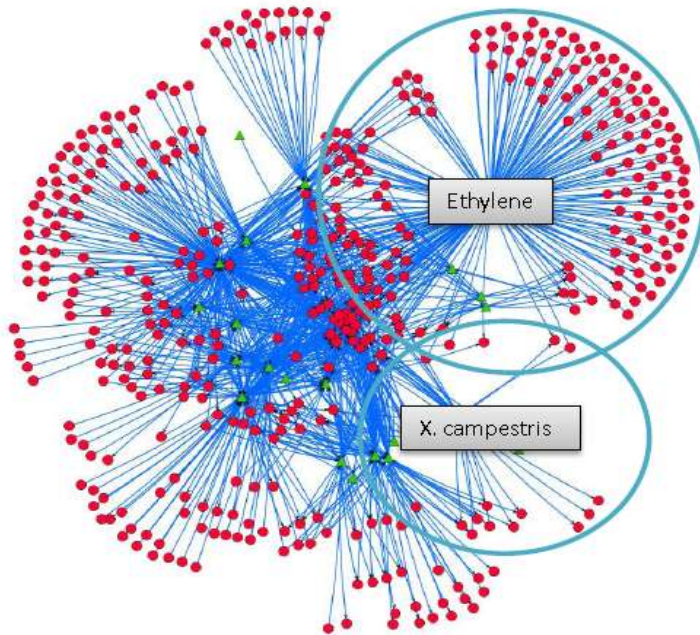


Cytoskeleton dynamics significantly altered.

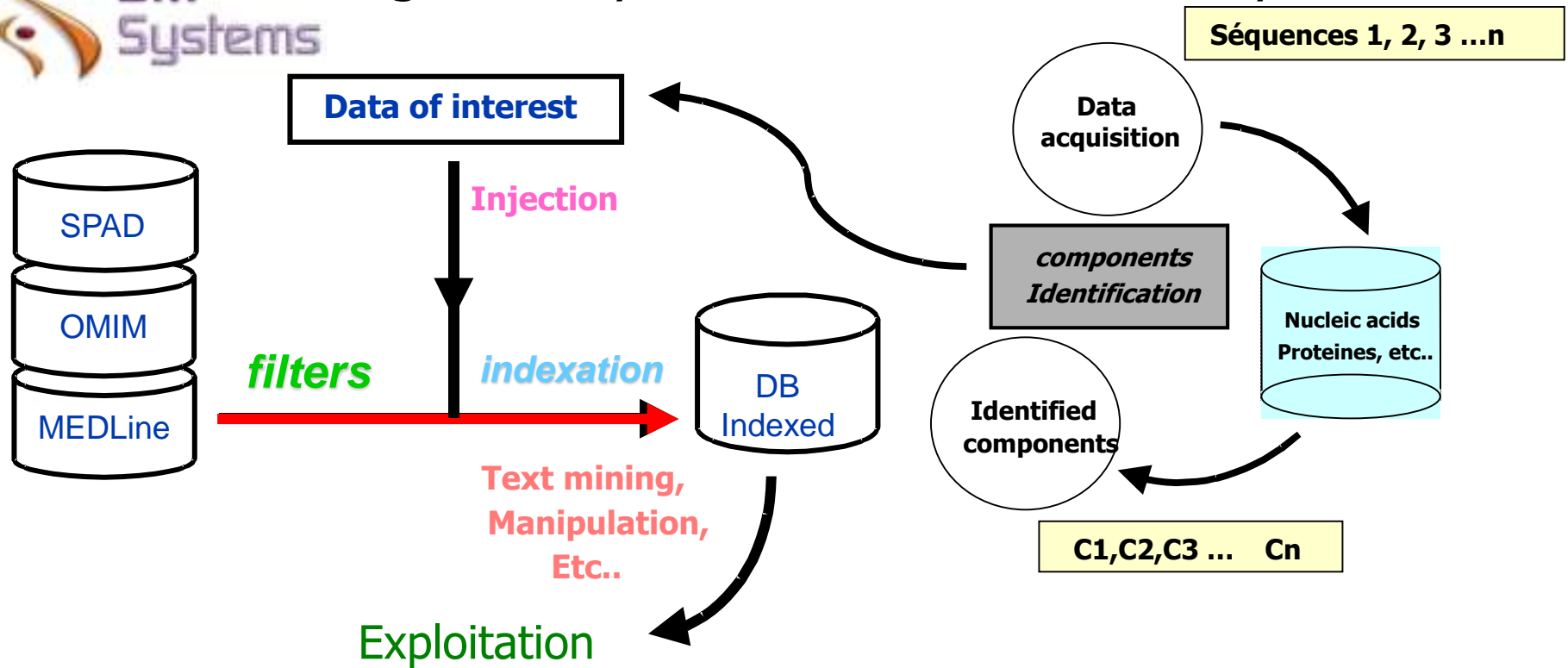
D CD44 glycosylation patterns = drastically different effects on regulation of immune cells.

And the more complex the organism being analysed,
the worse it gets.

Protein-Stress Association Network



- 3145 proteins linked to 32 stresses by 10777 relations
- On average
 - each protein associated with 3.4 stresses
 - each stress associated with 337 proteins
- Filtering associations based on three scoring metrics IP, M and N
- Which metric and cut-off are most suited for filtering noise?



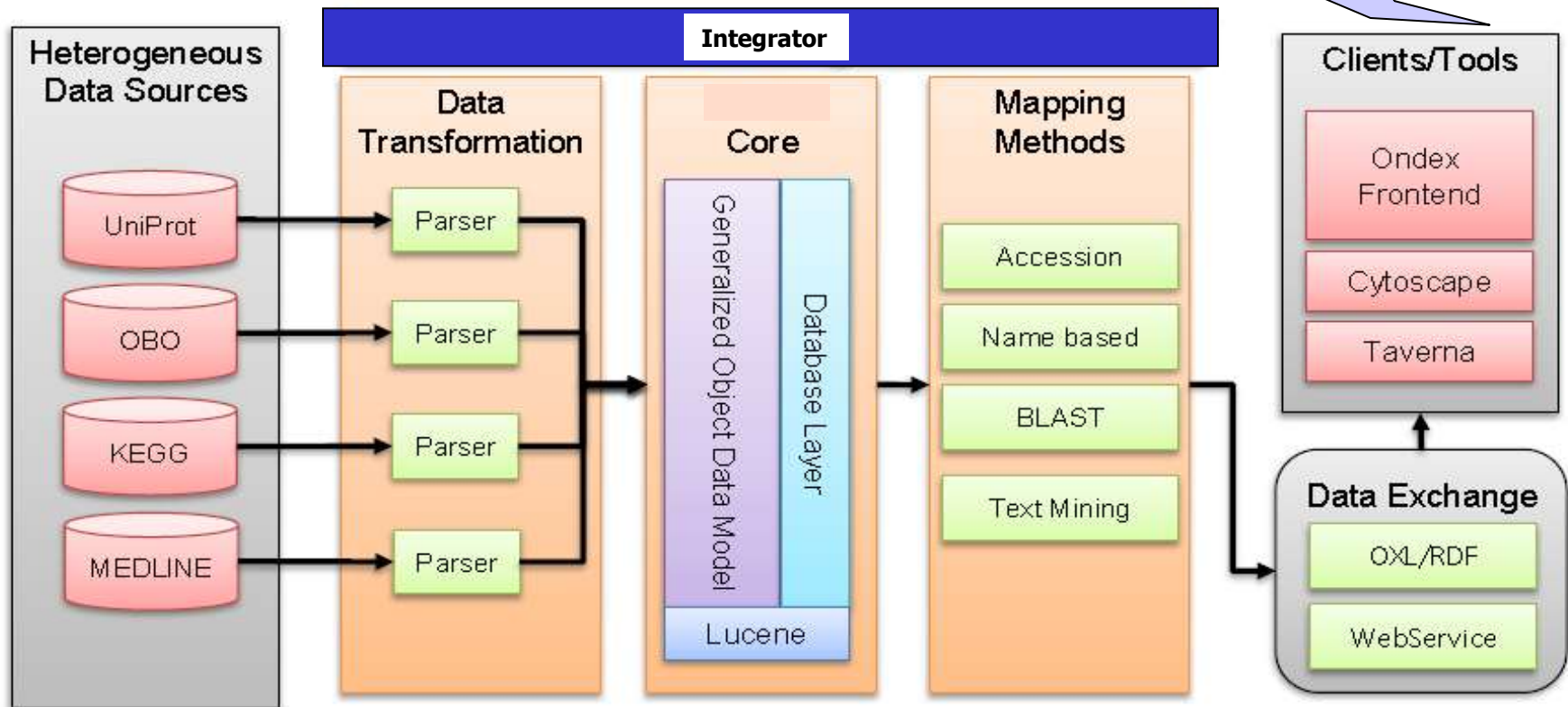
What constitutes a **GOOD** filter?

What constitutes a **GOOD** indexation strategy?

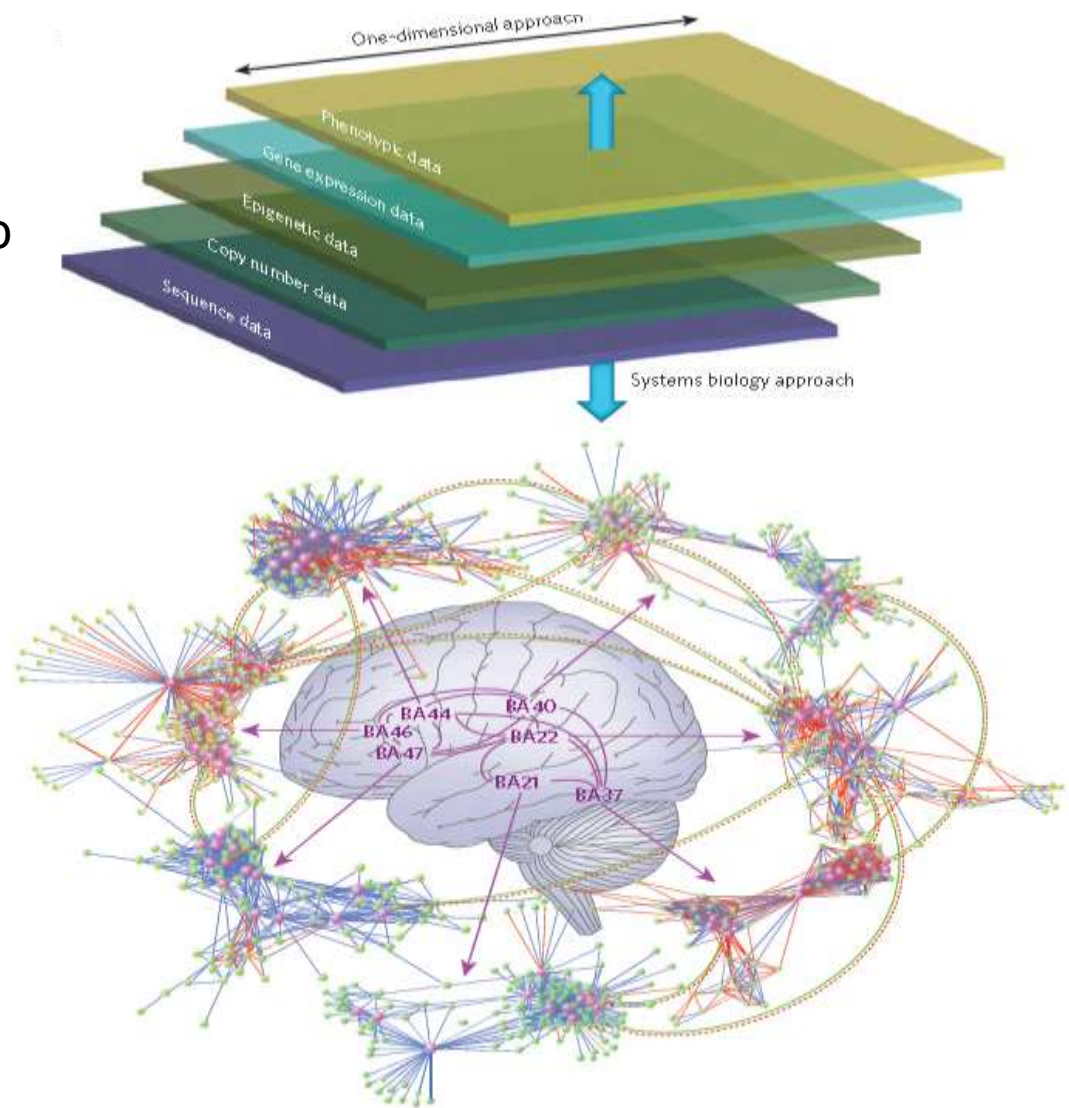
ALL information entered into the DB is **ALWAYS**
biased, incomplet, erroneus, etc...

Accumulation
of inconsistencies

As a result, this generates coherence issues affecting the entire analytical chain of events, irrespective of the visualisation and manipulation tools finally used!



As a consequence, when it comes to really complex systems and poorly defined physiological processes, such as those associated with functions within the CNS and disorders thereof, Bayesian approaches lead to nearly intractable difficulties.



Mathematical models are remarkable validation/fine-tuning tools when applied to well defined processes.

They are inadequate discovery tools when applied to poorly understood multicellular processes.

The challenge is clearly not a question of technologies only!

The paradox that must be resolved.

If you dream to create the first operational bird model...



... a "basic" living Complex system that not only flies...

Be sure to use the appropriate modeling concepts & tools. If not...



...you get a Complicated "Cartesian" system. It does fly, but...

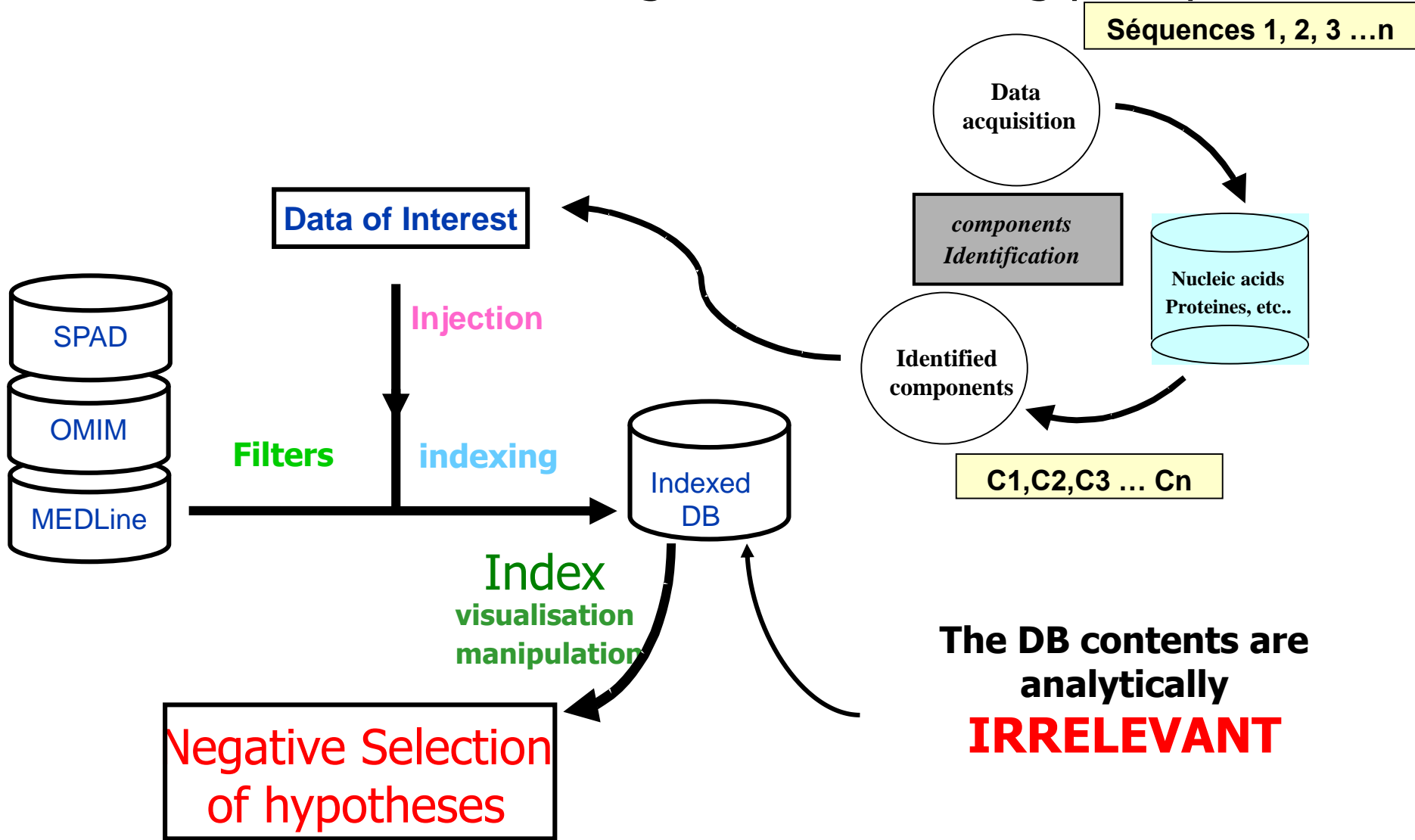


How to escape the paradox whereby we have NO OTHER CHOICE but utilise tools that CANNOT enable the goal to be reached?



By changing our intellectual approach!

Another way of thinking: The *CADI*TM Integration & Modelling principles.



This constitutes a heuristic approach!

The differences between « heuristic » and « mathematical » approaches.

Heuristics:

A problems solving approach evaluating each step in a process, searching for satisfactory solutions rather than for optimal solutions, using all available qualitative information instead of quantitative information.

Thus,

Heuristic modelling starts from accumulated information to produce a model capable of describing the mechanisms that generated the observed outcome / data and predict their modifications associated with a different outcome;

It plays the role of an architect

While

mathematical (Bayesian) modelling starts from quantitative data to produce models capable of reiterating this data and predict the outcome of a different experimental paradigm.


It plays the role of an engineer

and

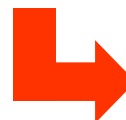
Far from being incompatible, these two approaches can be complementary.

Bayesian and Heuristic approaches can be complementary, provided they are harnessed in the proper order.

Bayesian approaches are of limited usefulness when applied to poorly defined multicellular physiological systems because they cannot efficiently reveal & define the functional states within such a system (cross-talks alterations, etc...).

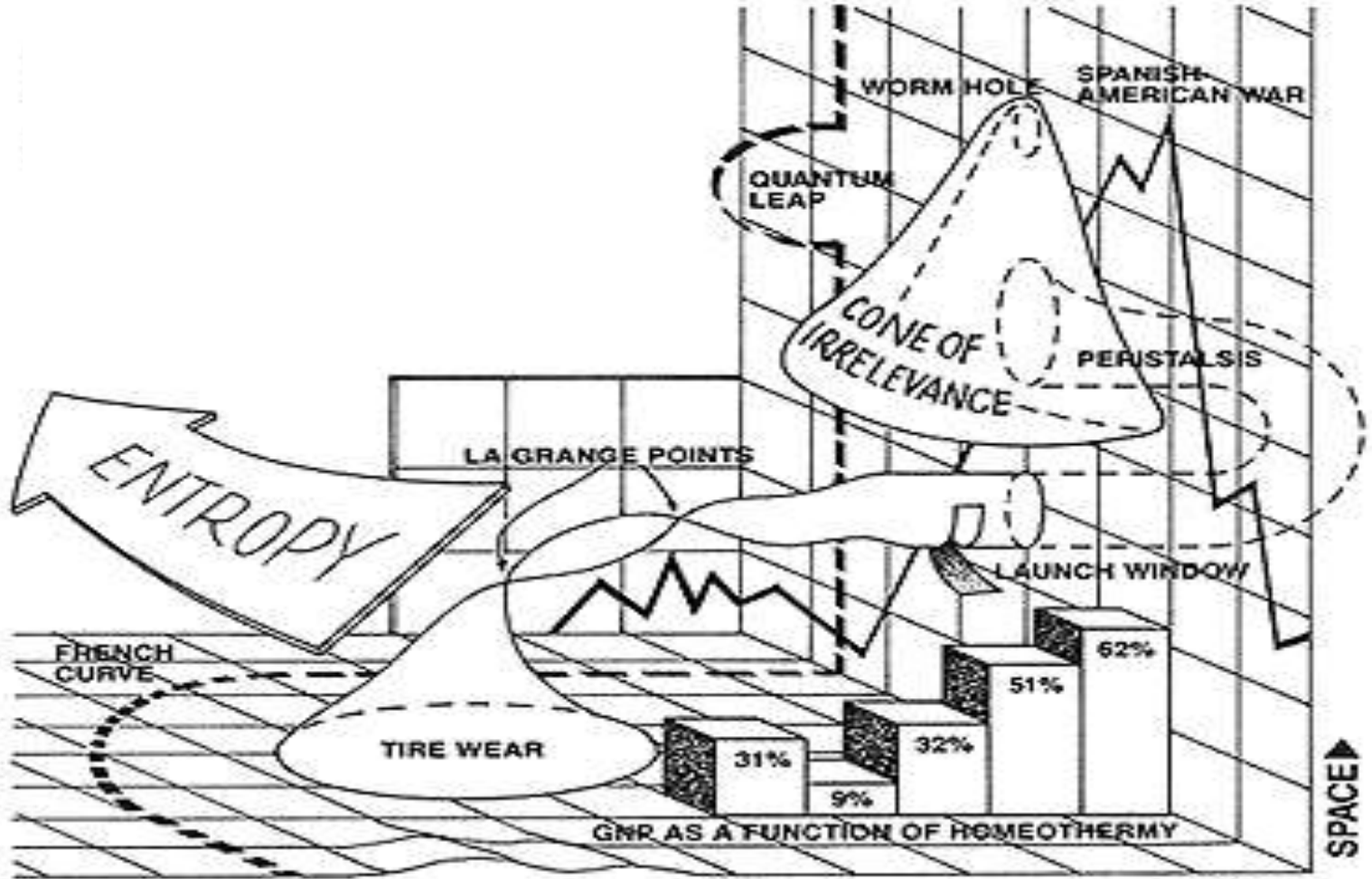
 But heuristic approaches are very efficient at doing precisely this.

Heuristic models are of limited usefulness when addressing the dynamics of defined complex physiological pathways structures and cross-talks because they are not open to mathematical manipulations.

 But Bayesian models are very efficient at doing precisely this.

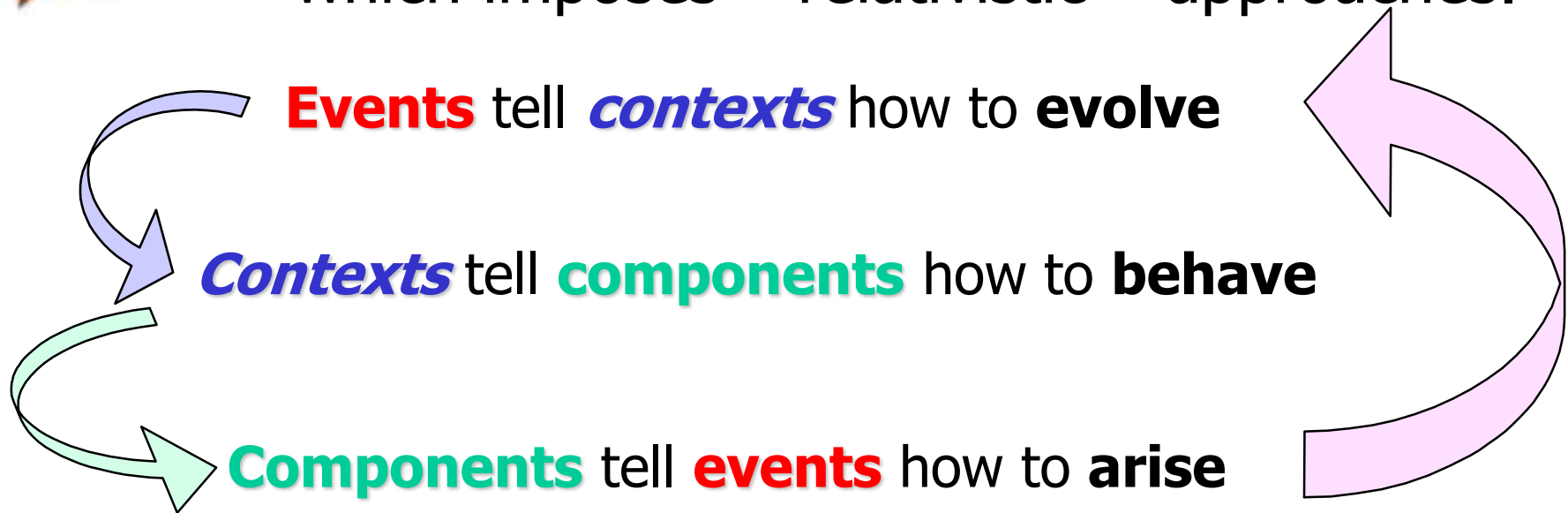
To efficiently address the translation of systems biology to clinical & medical interventions (dominated by patient's data heterogeneity and largely unstructured documents), ways to achieve synergy between Heuristic and Bayesian approaches can be effectively designed.

The heuristic analytical process must follow a « relativistic » approach.



Within this framework, *Non-linearity*, *Irrelevance*, *Wear*, *Relative weights* & *Contexts* are key concepts.

Why? Because of a very simple set of rules which imposes « relativistic » approaches.



*Analyses in terms of biological components and functions are now **IRRELEVANT**.*

EVENT-DRIVEN (relativistic) analytical approaches become necessary.

This, in turn, imposes analytical procedures based upon the negative selection of working hypotheses.

Why “negative selection” of working hypotheses?

“While it is not always possible to demonstrate that a statement is true, it is always possible to demonstrate it to be false” Karl Popper, 1963.

Mathematical approaches are based on “positive selection”: it is assumed that every data-set/statement is actually valid.

Yet, “an estimated 85% of current published research findings are false or exaggerated”

J.P.A Joannidis, 2014 [PLoS Med. 11(10): e1001747]; F. Prinz et al., 2011 [Nat Rev Drug Discov. 10 (9):712]

Positive selection becomes a killer!

How to identify what is NOT false and/or exaggerated?

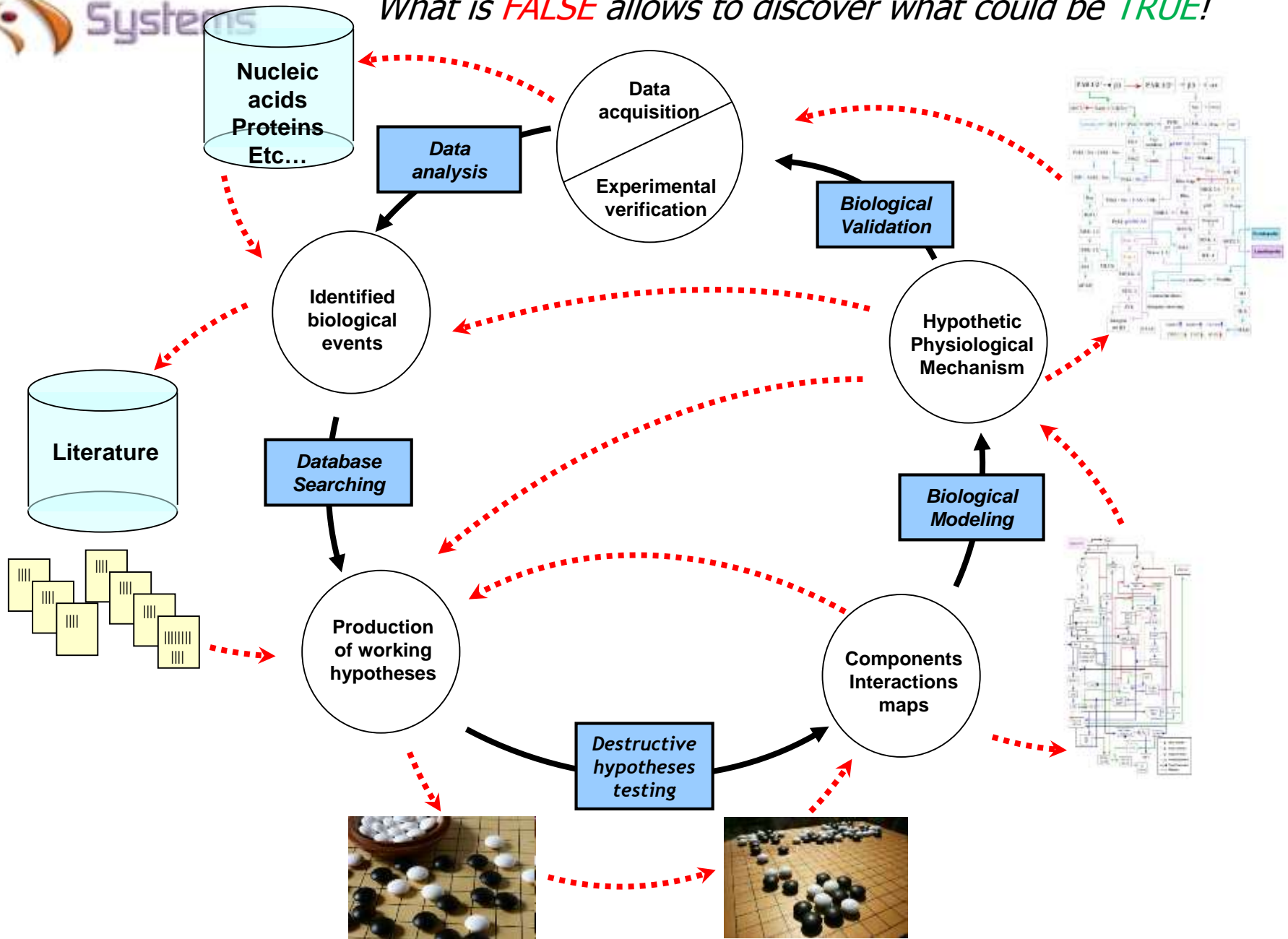
By doing every thing possible to destroy working hypotheses!

Only hypotheses that resist destruction are worth retaining.

Hence, what is demonstrated “False” can now be used to discover what could be “True”.

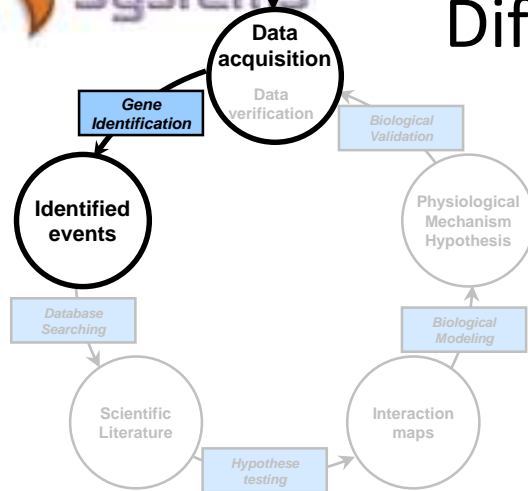
The **CADI**TM Integration workflow

What is **FALSE** allows to discover what could be **TRUE**!



MCF-7/MCF-7 ras Differentially Expressed Genes

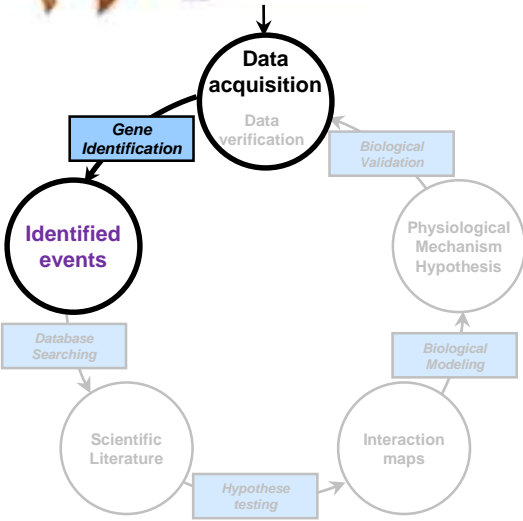
■ Over-expressed
■ Under-expressed
over 600 genes



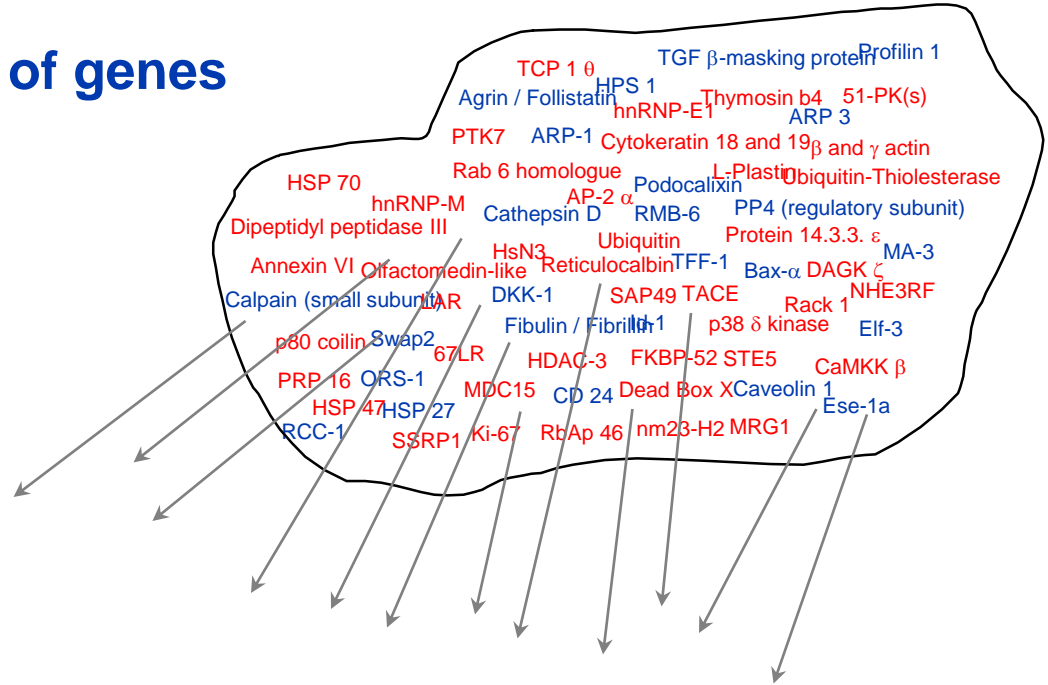
TCP 1 θ TGF β -masking protein Profilin 1
Agrin / Follistatin HPS 1 Thymosin b4 51-PK(s)
hnRNP-E1 ARP 3
PTK7 ARP-1 Cytokeratin 18 and 19 β and γ actin
Rab 6 homologue L-Plastin Ubiquitin-Thiolesterase
HSP 70 hnRNP-M Podocalixin
Dipeptidyl peptidase III AP-2 α RMB-6 PP4 (regulatory subunit)
Annexin VI Olfactomedin-like HsN3 Ubiquitin Protein 14.3.3. ϵ
Calpain (small subunit) LAR DKK-1 SAP49 TFF-1 Bax- α DAGK ζ MA-3
p80 coilin Swap2 67LR Fibulin / Fibrillin Id-1 p38 δ kinase Elf-3
PRP 16 ORS-1 MDC15 HDAC-3 FKBP-52 STE5 CaMKK β
HSP 47 CD 24 Dead Box X Caveolin 1
RCC-1 HSP 27 Ki-67 RbAp 46 nm23-H2 MRG1 Ese-1a
SSRP1

Gadal F. *et al.* (2003) *Nucleic Acids Research* 19 (31): 5789-5804

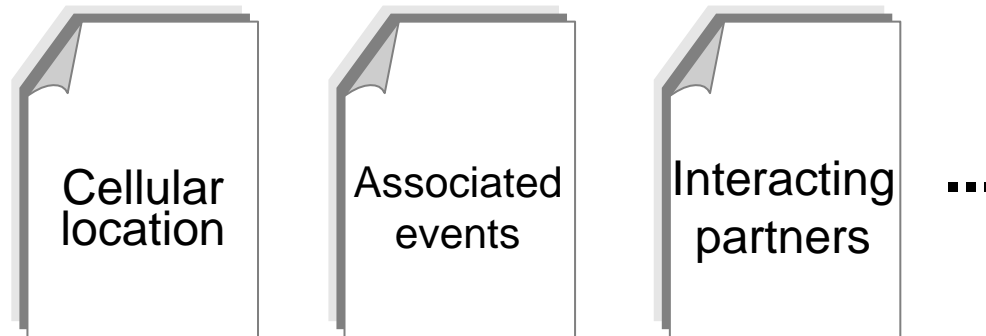
Information Retrieval & Extraction



Set of genes



Set of papers



Information MEDLINE

text only

UI - 98057316

TI - The human 37-kDa laminin receptor precursor interacts with the prion protein in eukaryotic cells.

AU - Rieger R

AU - Edenhofer F

AU - Lasmezas CI

AU - Weiss S

LA - eng

MH - Actins/metabolism

MH - Animal

MH - Binding Sites

MH - COS Cells

MH - Cell Line

MH - Eukaryotic Cells

MH - Hamsters

MH - Human

MH - Mice

MH - Mice, Inbred C57BL

MH - PrPSc Proteins/*metabolism

MH - ProteiPrecursors/chemistry/*metabolism

MH - Rabbits

MH - Laminin/chemistry/genetics/*metabolism

MH - Saccharomyces cerevisiae/metabolism

MH - Spodoptera/cytology

SO - Nat Med 1997 Dec;3(12):1383-8.

AB - Prions are thought to consist of infectious proteins that cause transmissible spongiform encephalopathies.

According to overwhelming evidence, the pathogenic prion protein PrPSc converts its host encoded isoform PrPC into insoluble aggregates of PrPSc, concomitant with pathological modifications (for review, see refs. 1-3).

Although the physiological role of PrPC is poorly understood, studies with PrP knockout mice demonstrated that PrPC is required for the development of prion diseases. Using the yeast two-hybrid technology in *Saccharomyces cerevisiae*, we identified the 37-kDa laminin receptor precursor (LRP) as interacting with the cellular prion protein PrPC. Mapping analysis of the LRP-PrP interaction site in *S. cerevisiae* revealed that PrP and laminin share the same binding domain (amino acids 161 to 180) on LRP. The LRP-PrP interaction was confirmed in vivo in insect (*Sf9*) and mammalian cells (COS-7). The LRP level was increased in scrapie-infected murine N2a cells and in brain and spleen of scrapie-infected mice. In contrast, the LRP concentration was not significantly altered in these organs from mice infected with the bovine spongiform encephalopathic agent (BSE), which have a lower PrPSc accumulation. LRP levels, however, were dramatically increased in brain and pancreas, slightly increased in the spleen and not altered in the liver of scrapie-infected hamsters. ...

Information MEDLINE

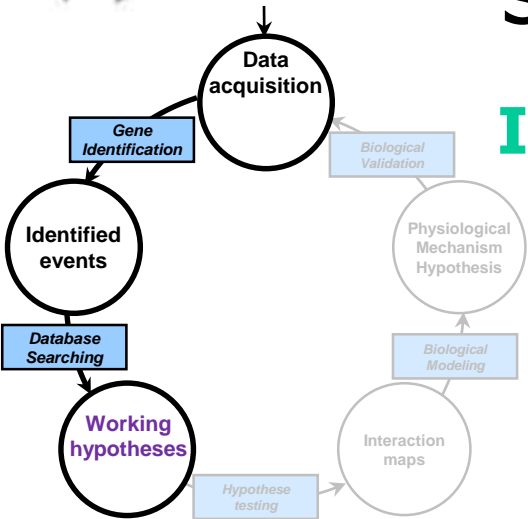
text only

UI - 98057316
TI - The human 37-kDa **laminin receptor precursor** **interacts** with the **prion** protein in eukaryotic cells.
AU - Rieger R
AU - Edenhofer F
AU - Lasmezas CI
AU - Weiss S
LA - eng
MH - Actins/metabolism
MH - Animal
MH - Binding Sites
MH - COS Cells
MH - Cell Line
MH - Eukaryotic Cells
MH - Hamsters
MH - Human
MH - Mice
MH - Mice, Inbred C57BL
MH - PrPSc Proteins/*metabolism
MH - ProteiPrecursors/chemistry/*metabolism
MH - Rabbits
MH - Laminin/chemistry/genetics/*metabolism
MH - Saccharomyces cerevisiae/metabolism
MH - Spodoptera/cytology
SO - Nat Med 1997 Dec;3(12):1383-8.

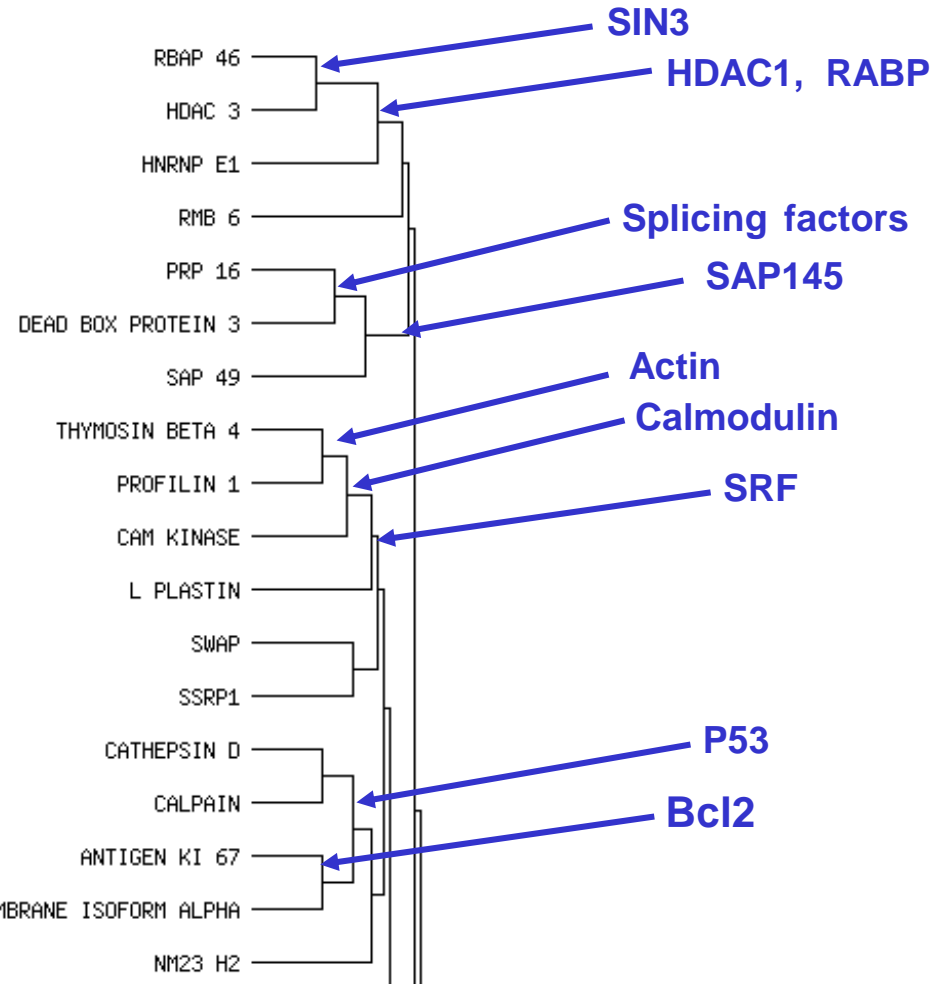
AB - **Prions** are thought to consist of infectious proteins that cause transmissible spongiform encephalopathies. According to overwhelming evidence, the pathogenic prion protein **PrPSc** converts its host encoded isoform **PrPC** into insoluble aggregates of **PrPSc**, concomitant with pathological modifications (for review, see refs. 1-3). Although the physiological role of **PrPC** is poorly understood, studies with **PrP** knockout mice demonstrated that **PrPC** is required for the development of prion diseases. Using the yeast two-hybrid technology in *Saccharomyces cerevisiae*, we identified the 37-kDa **laminin receptor precursor (LRP)** as **interacting** with the cellular prion protein **PrPC**. Mapping analysis of the **LRP PrP interaction** site in *S. cerevisiae* revealed that **PrP** and **laminin** share the same **binding** domain (amino acids 161 to 180) on **LRP**. The **LRP PrP interaction** was confirmed in vivo in insect (Sf9) and mammalian cells (COS-7). The **LRP** level was increased in scrapie-infected murine N2a cells and in brain and spleen of scrapie-infected mice. In contrast, the **LRP** concentration was not significantly altered in these organs from mice infected with the bovine spongiform encephalopathic agent (BSE), which have a lower **PrPSc** accumulation. **LRP** levels, however, were dramatically increased in brain and pancreas, slightly increased in the spleen and not altered in the liver of scrapie-infected hamsters. ...

Generating hypotheses & systematic destructive testing.

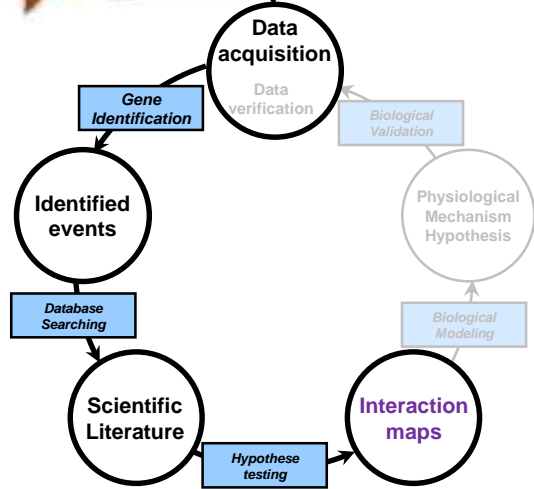
Index visualisation & manipulation



RNA Splicing & expression



Apoptosis & cell survival

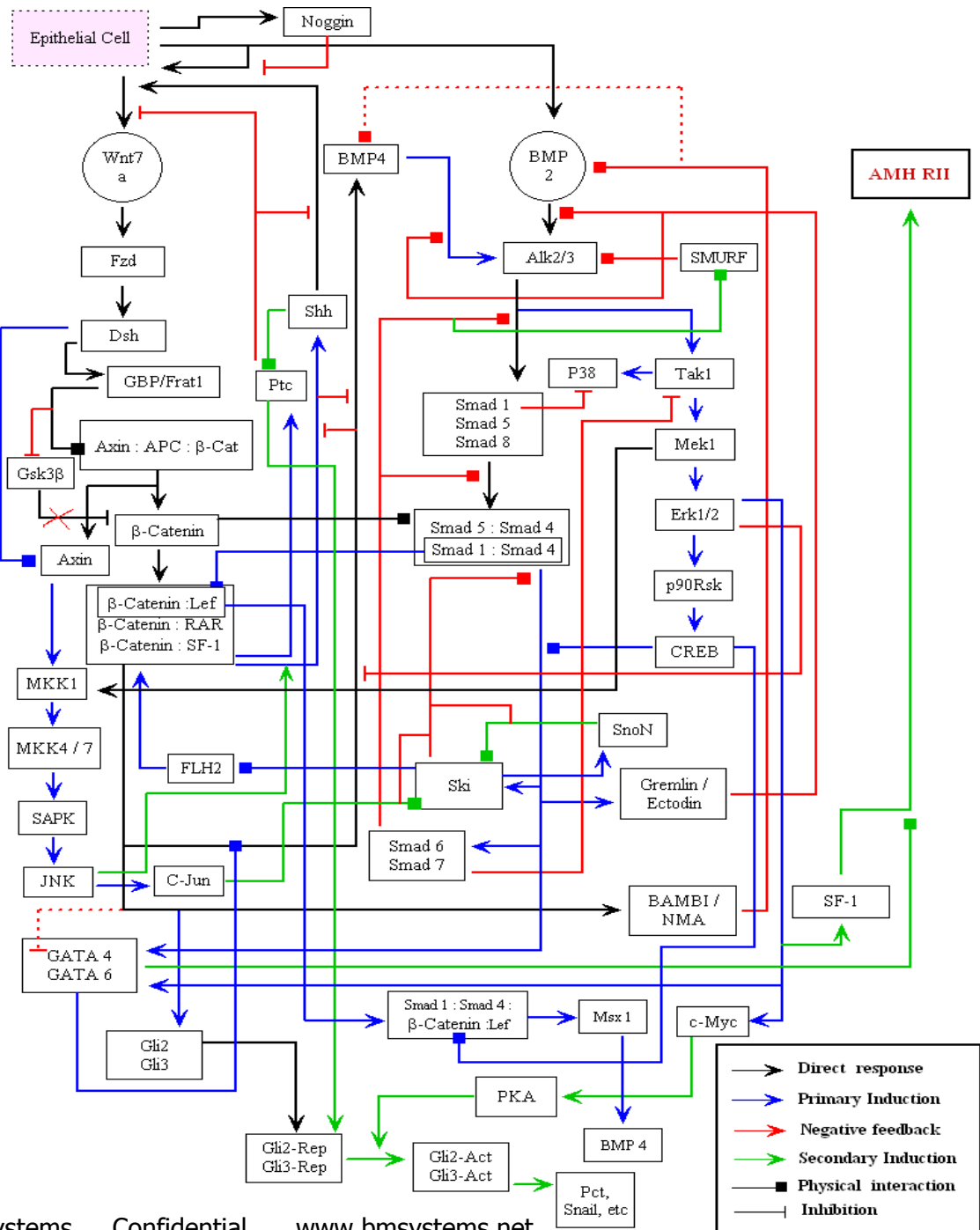


Interaction maps

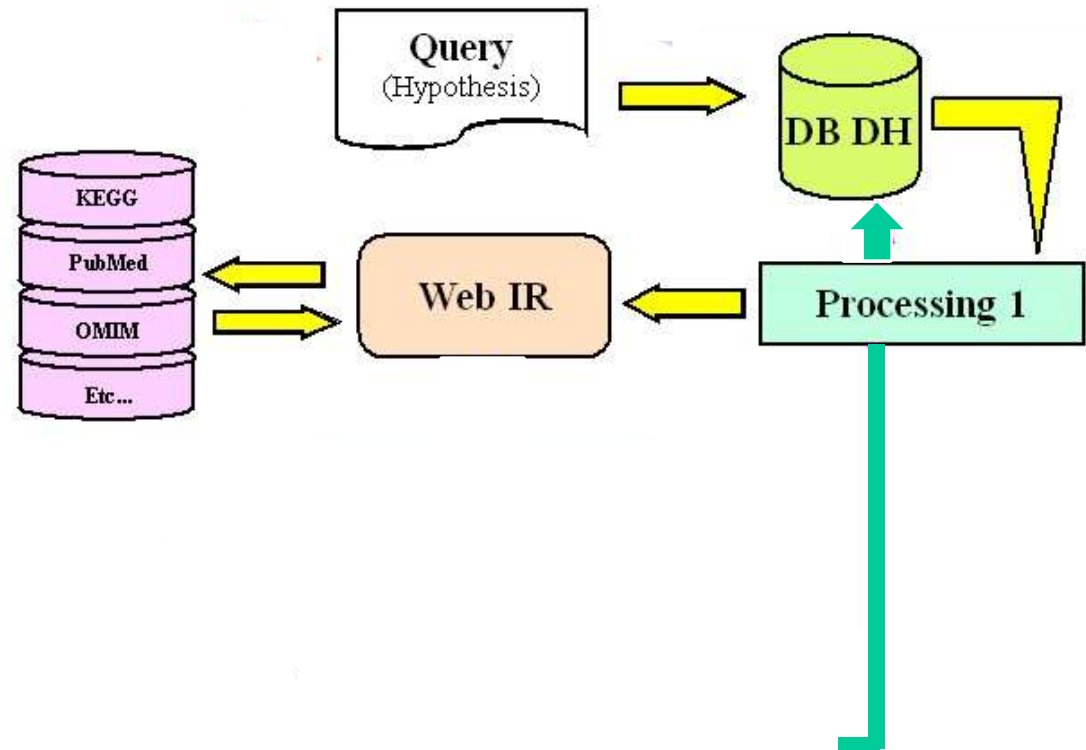
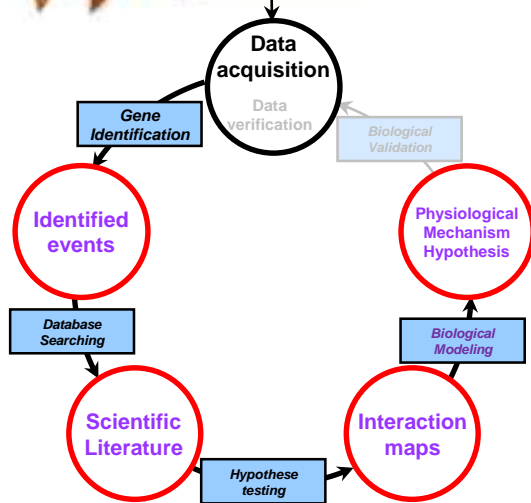
describe the pathways that have **become functional** and those that have **become forbidden** in response to local conditions imposed by the activation of defined biological mechanisms.

Specific biological events do not occur because they are fated to.

They occur because other events could not!



The CADI™ Integration & Modelling Process.



This iterative process does three things:

- It largely resolves the coherence issues attached to the classical approach;
- It reveals hitherto unknown mechanisms/processes, and
- It allows the translation of systems biology to clinical & medical interventions.

A concrete example #2.

The ADAM-15 RGD peptide

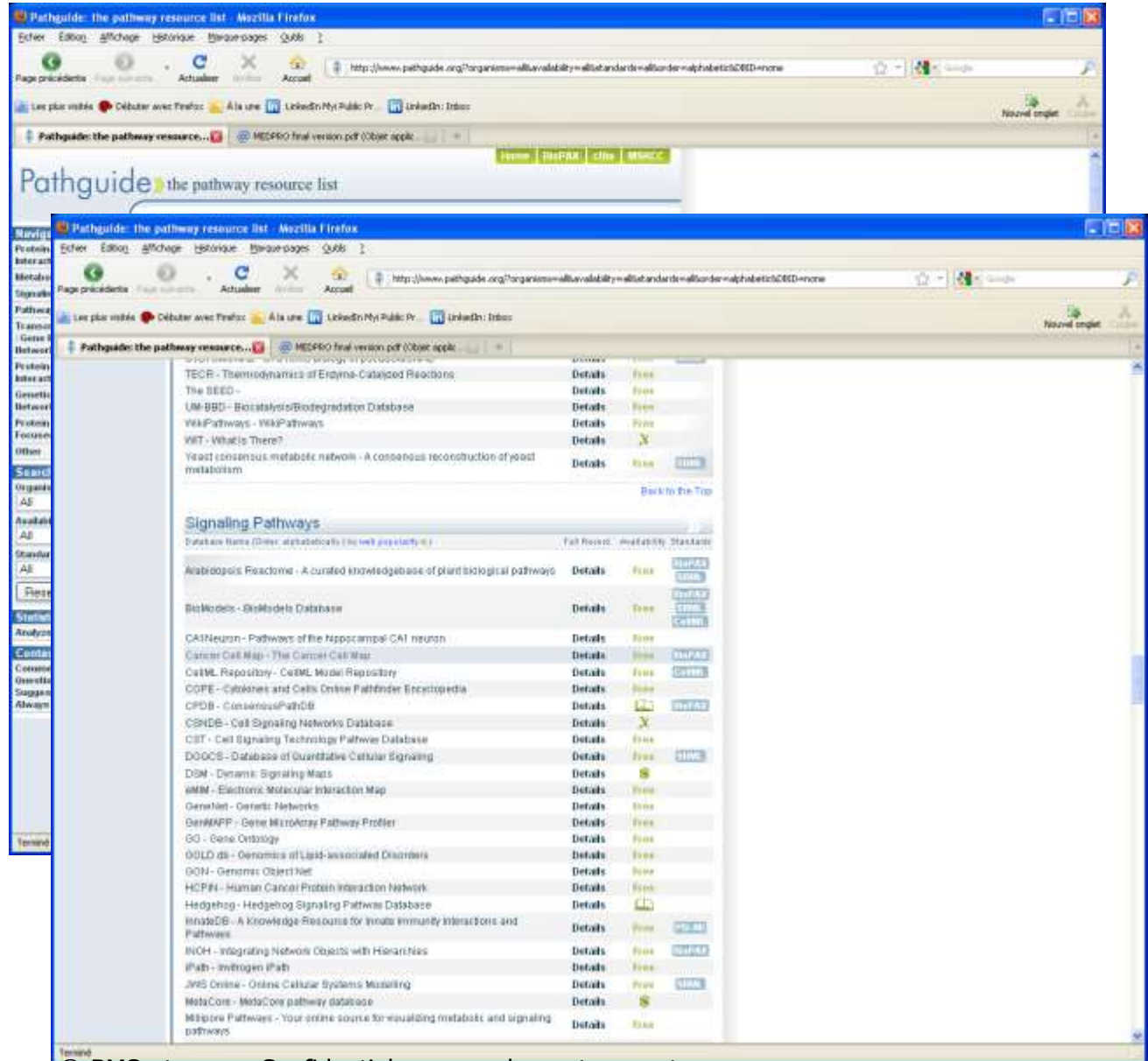
- ❖ 10 amino acids fragment (GCGGRGDGGC) derived from the metaloprotease Adam 15.
- ❖ Induces cytotostasis (dormancy) in very aggressive metastatic breast cancer cells both *in vitro* and *in vivo*.
- ❖ The target of this peptide appears to be integrin-containing structures (cell adhesion) but not integrins alone (anti-integrins antibodies have no such effects).
 - What is (are) the target (s) of this peptide?
 - What mechanisms does it affect? and
 - Why does it induce cytotostasis?

This model addresses the mode of action of a potential treatment.

Collaboration BMSystems-INSERM U 553 (Hôpital St Louis)

The classical Bayesian attack.

Given the problem, one may wish to start by working on the integrins-associated mechanisms of metastasis, using already partly corroborated interaction networks issued from multiple gene expression analyses (affimetrix chips quantitative data) on which one might superimpose newly generated data.



You start by first locating the most appropriate database through the « Pathguide » web resource.

In the present case « Cancer Cell Map » which regroups the results from several hundred quantitative gene expression studies.

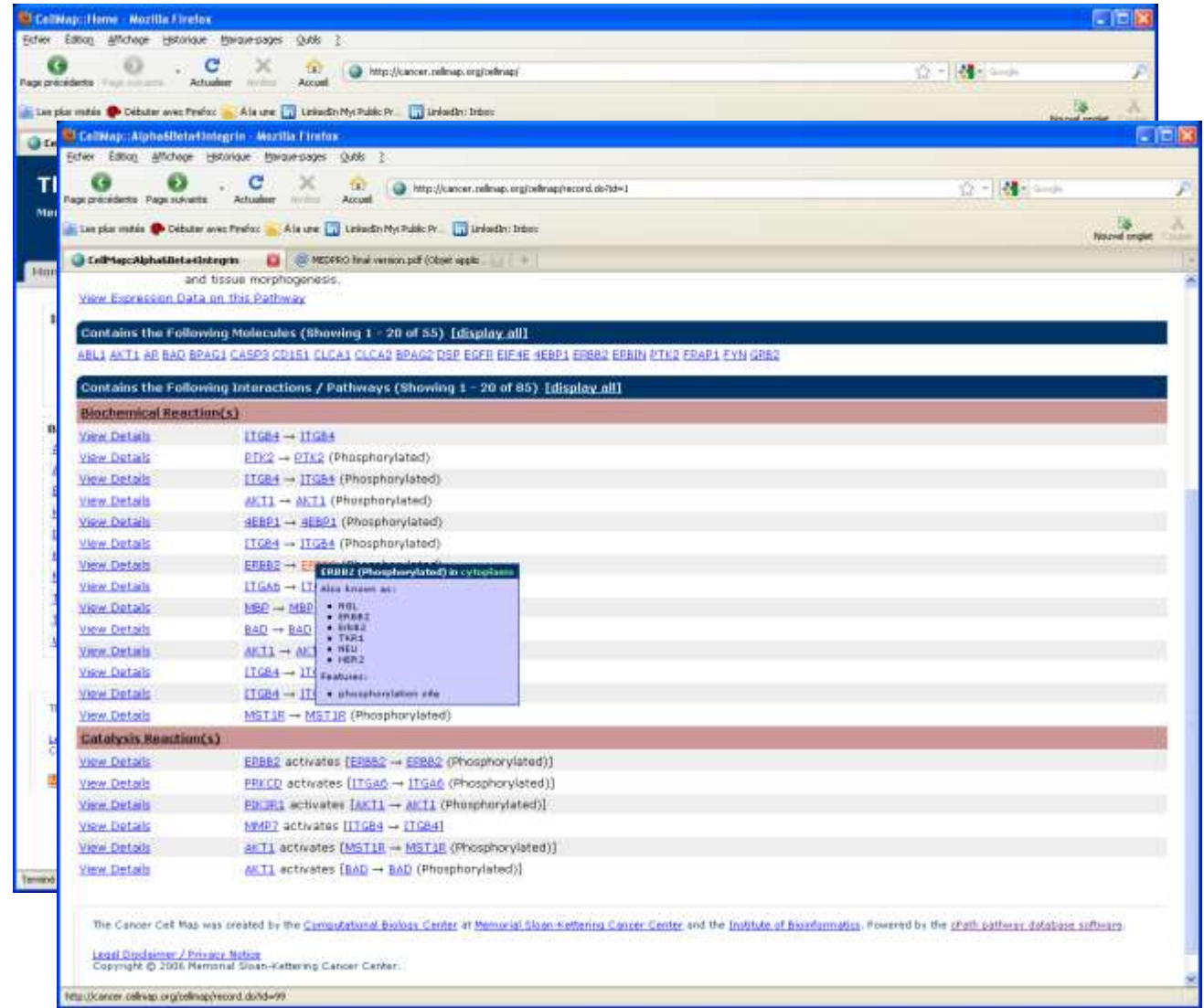
So, you open the web link to « Cancer Cell Map »

And you find that you have access to 10 pathways, including the integrins $\alpha6\beta4$ pathway which turns-out to be of primary interest.

You open the link to this pathway

And you are faced with a tabulation of biochemical and catalysis reactions that gives all the « known » components and their aliases but no idea of actual flow.

You then realise that, to really exploit this resource, you first need to load and activate the « Cytoscape » analytical tool (one of the 61 different freely accessible tools currently available).



The screenshot shows a web browser displaying the 'Cancer Cell Map' website. The main content area is titled 'CellMap: Alpha6beta4Integrin' and lists various biochemical and catalysis reactions. A tooltip is visible over the reaction 'E2F3 (Phosphorylated) in cytoplasm', listing aliases: RBL, ERB2, ERB3, TSK1, RLU, and HER2. The page also includes a list of molecules and interactions, and a footer with copyright information for Memorial Sloan-Kettering Cancer Center.

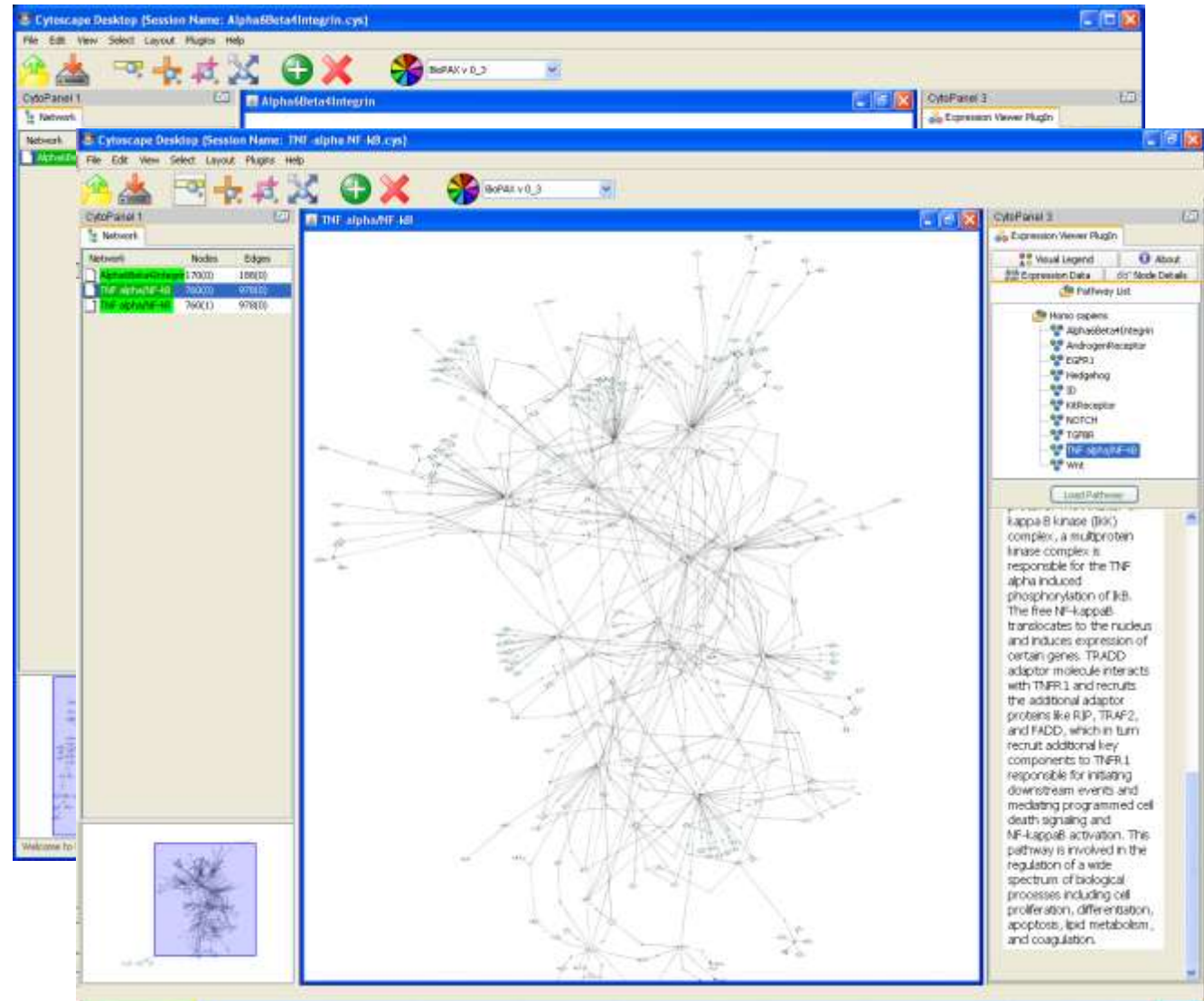
You install that tool, and you finally gain access to a spatial representation of the pathway, the links and biological interactions between its components.

Having started to explore the links within this diagram, you suddenly realise that this does not make any distinction between the various processes in which the $\alpha6\beta4$ system could be involved, such as cell proliferation, differentiation, apoptosis, migration and tissue morphogenesis.

To make some sense of all this, you need to attach a closely connected pathway which, together with the $\alpha6\beta4$ system, may help to better distinguish the major components of metastasis. In the present case, the TNF α /NF-kB pathway.

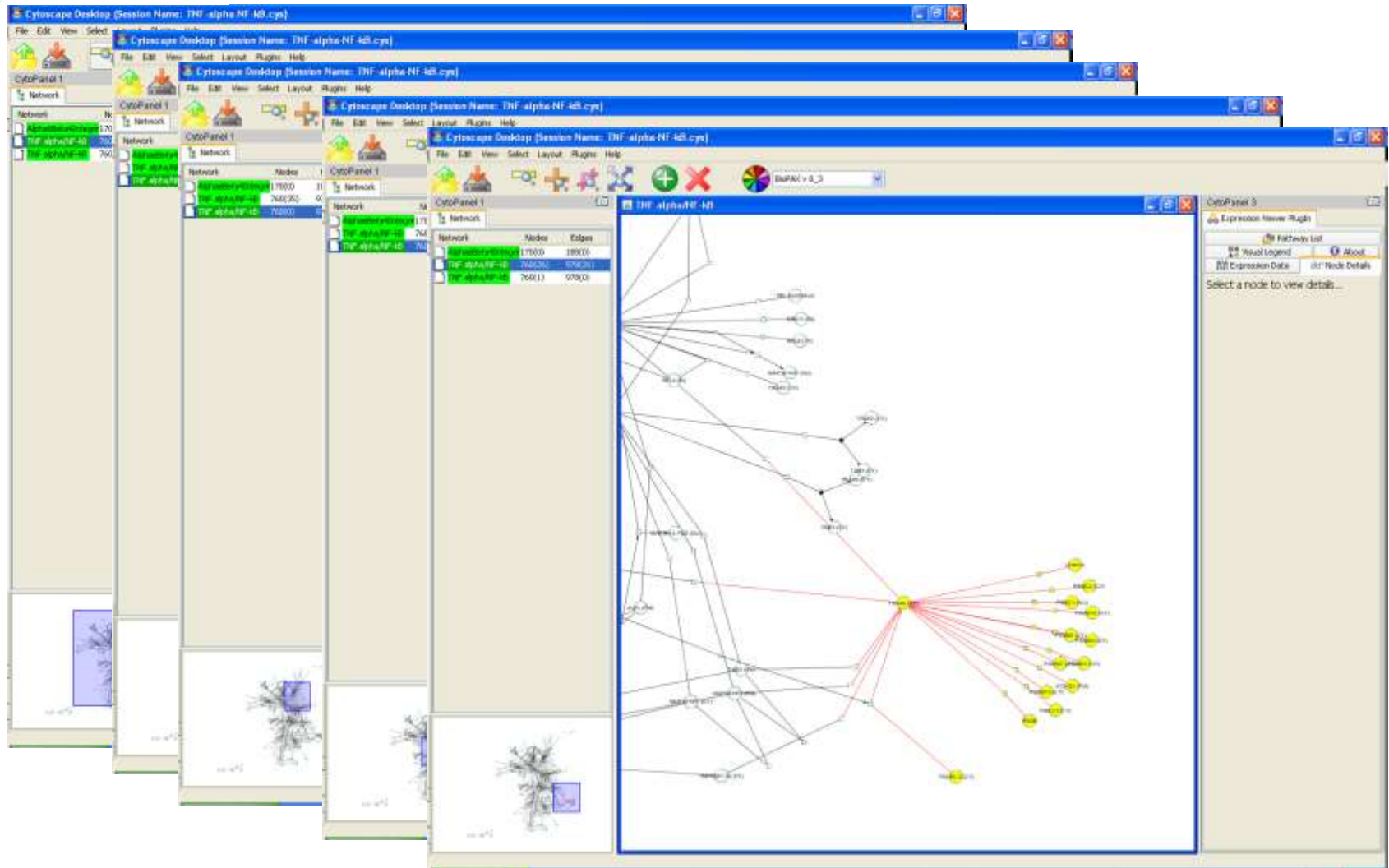
You open this link, and....

SHOCK!



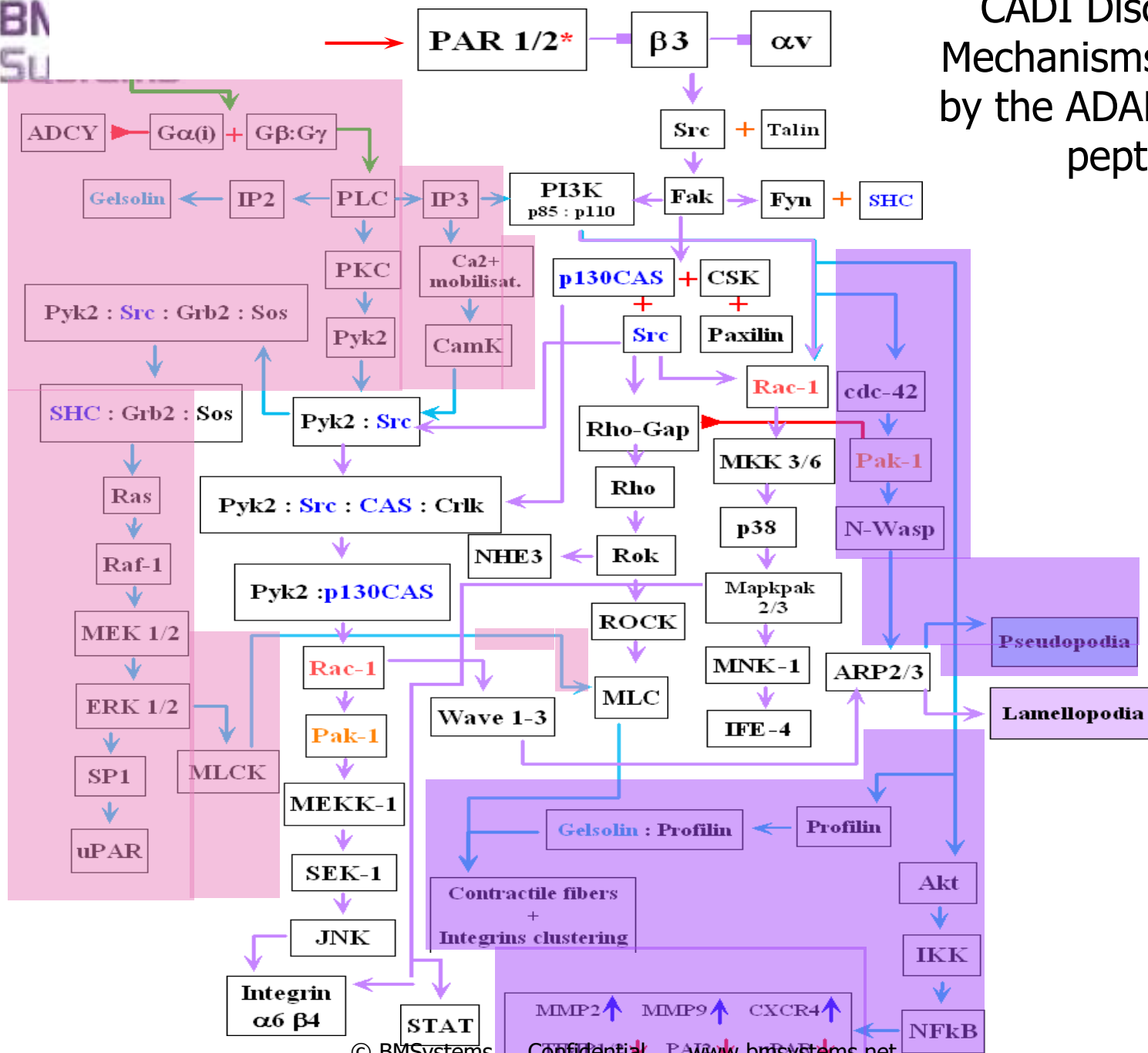
You are confronted with a « star map »!...

Sure, you can navigate in the « star map ».
You can even zoom into specific regions..



But since this map, just as the previous one, does not distinguish between cell proliferation, differentiation, motility and apoptosis, you are now even further away from a solution to your problem than you were with the first map alone!

CADI Discovery's Mechanisms targeted by the ADAM-15 RGD peptide



Practical outcome.

(Validated *in vitro* & *in vivo* by Inserm U 553)

Treatment with the ADAM-15 RGD peptide massively triggers the integrins-associated, β -arrestin-dependent endocytotic mechanisms.

This effectively abolishes integrins-dependent cytokinesis mechanisms while inducing cytostasis in cells where these mechanisms are being implemented.

But this does NOT solely address metastasis, where integrin-dependent mechanisms are inappropriately activated.

It will also address and inhibit physiologically coherent integrins-dependent cytokinesis mechanisms, such as fibroblasts migration during wound healing, or repair of damaged capillaries.

Thus, the potential therapeutic benefits of ADAM-15 RGD (inhibition of integrin-dependent metastasis) are annihilated by direct, highly undesirable effects (failure of wound healing and thus serious risks of gangrene; failure of vascular repair and thus serious risks of multiple organ failures; etc.).

In spite of high efficacy, the therapeutic potential of the ADAM-15 RGD peptide is extremely low.  ***Reconsider therapeutic development.***

Example #3:

Understanding the co-evolutionary interplays between bacteria and bacteriophages leads to the discovery of the means whereby **undefined** multi-resistant bacterial pathogens can be efficiently controlled.

The questions (French Defence)

- How to rapidly (less than 30 min) and efficiently **detect** the presence of any given **LIVE** bacterial pathogen?
- How to rapidly and efficiently **destroy** any **unknown** bacterial pathogen or emerging strain **without using**

A) Antibiotics: too many resistant strains, and very rapid resistance acquisition.

B) Vaccines: much too slow to act, and small strain variations often lead to inefficacy.

In other words, what is required is a “detector-killer”.

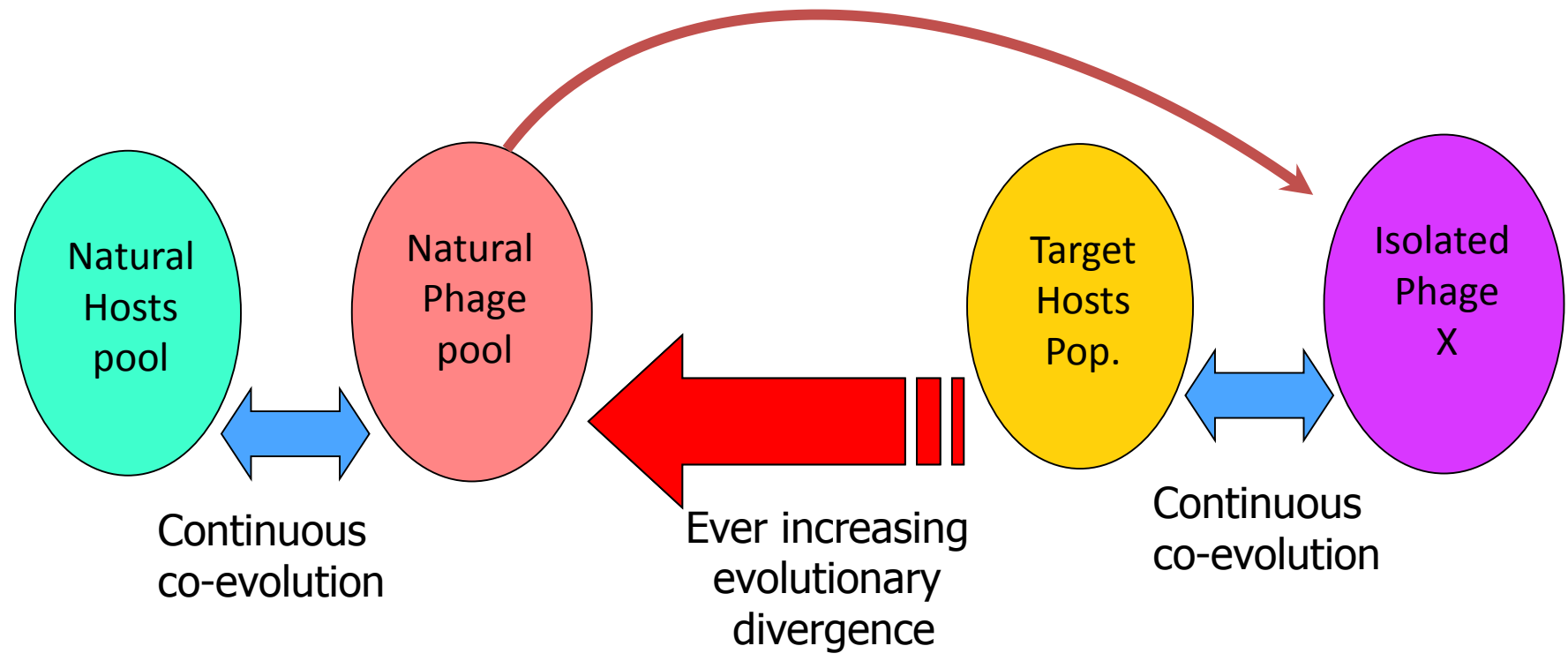
The first apparent answer

Bacteriophages, the natural predators of bacteria, could present the best potential to act as detectors-killers.

- Many are very host-specific,
- They only replicate in LIVE bacterial cells,
- Many kill the cells in which they replicate,
- As the phage progeny population increases that of the target diminishes (in a « closed » environment, few targets, if any, should escape), and
- They are extremely numerous and varied (they probably represent the most numerous « life forms » on the planet).

BUT the matter is NOT as simple as it first appears!

Co-evolution versus unidirectional predatory pressure.



+
Ever decreasing chances to
find a new efficient lytic phage.

Bacteria have existed for nearly 4 BILLION Years. They have so far resisted to EVERYTHING. And it is certainly NOT for lack of phages!

The model-derived solution.

- ***What, in essence, is the problem?***

The bacterial targets will try anything to escape predation and we have no idea what will be the successful strategy. Furthermore, this strategy is likely to vary between locations (populations) for a same target.

- ***What do we need to achieve?***

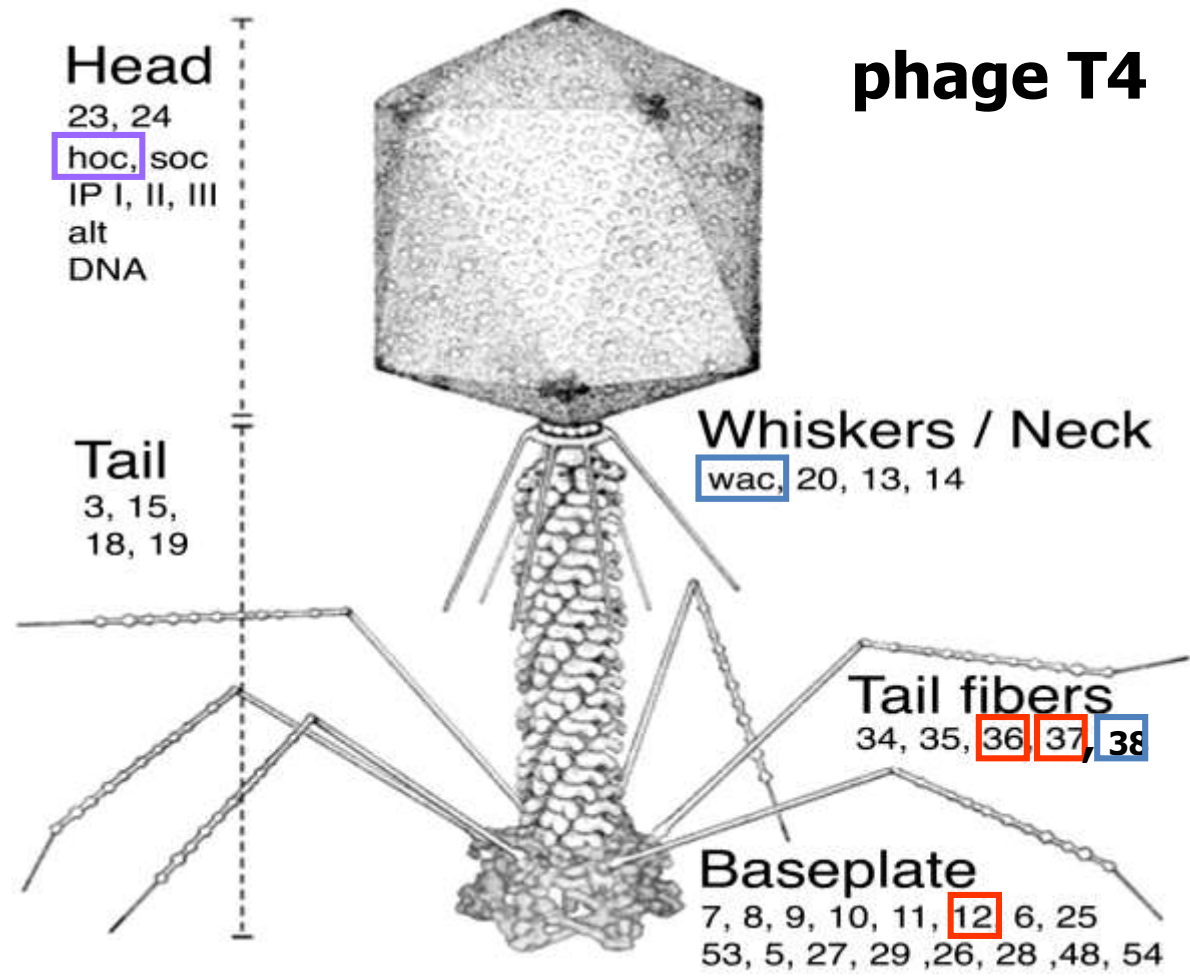
We must be capable of always preceding the targets escape strategies, no matter what they could be.

- ***The best-fit solution*** (model-derived):
we MUST adopt a stochastic approach.

It becomes necessary to

- abandon all idea of « natural phage pools » and,
- stochastically engineer phage banks in order to produce particles capable of targeting anything and everything while maintaining their capacity to replicate in the face of targets evasion attempts.

The problems:



How to modify any of these proteins in N different regions, at X different sites, in Z different manners, all this simultaneously and then recombine the multitude of variants generated in a population of obligate lytic phages?

The technological answers.

Three proprietary technologies (invented at BMSystems) allowing the production of stochastically engineered phage banks.

TAPETM (P) (WO 2008/093009):

A technology allowing to rapidly & simultaneously introduce defined densities of random mutations in any number of selected regions within a gene while conserving intact any number of defined coding domains in this same gene.

Applicable to any known coding sequence.

RipHTM (P) (WO 2009/090081):

A technology allowing to reversibly inactivate the genome of an obligate lytic phage within its host and carry out high efficiency homologous recombinations targeting multiple genes simultaneously without adversely affecting the host bacteria and the replicative capacity of the phage.

Ab-ACCUSTM (P) (WO 2008/093010):

A recombination technology allowing the rapid & efficient production of lytic phage banks in which every individual differs from all others for any number of selected genes or other sequences.

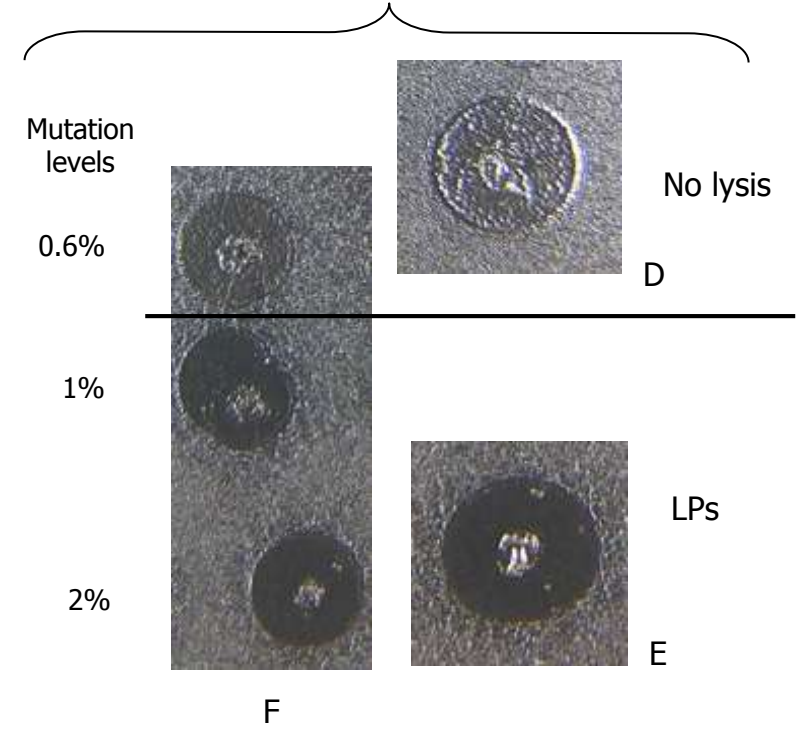
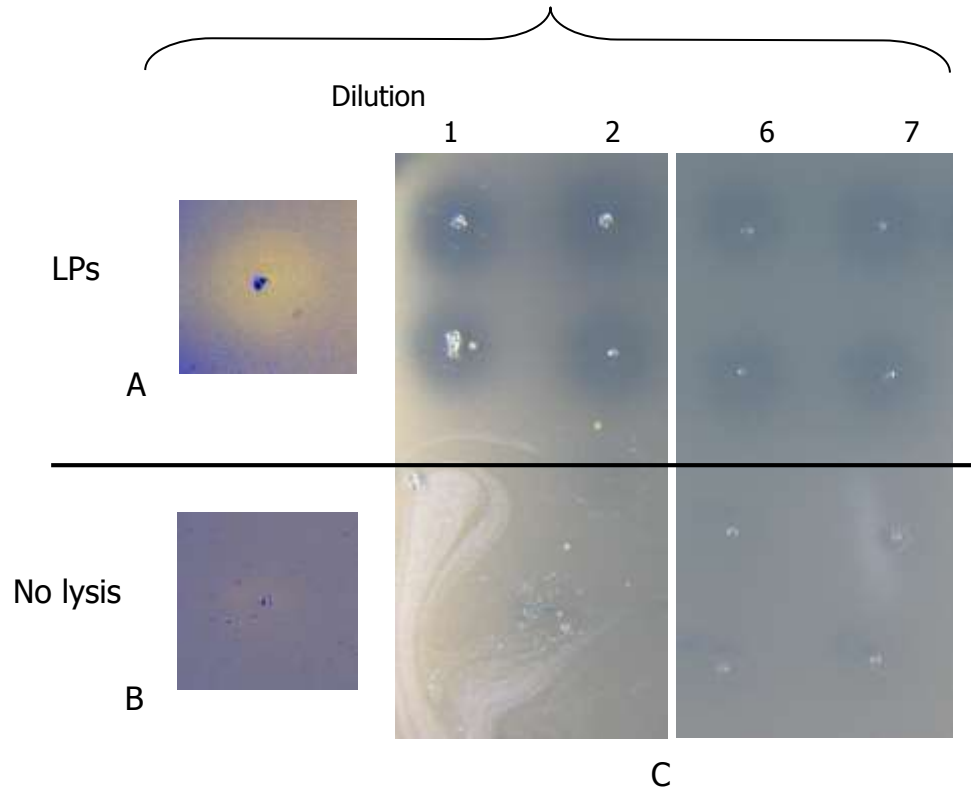
Applicable to any phage and to any known sequence.

The results.

While T4 is specific to a narrow range of *E. coli* strains,
 An engineered T4 bank contains variants capable of detecting and killing gram⁻
 bacteria far removed from *E. coli*.

Yersinia sp.

Pseudomonas sp.



Pouillot F, Blois H & Iris F (2010); Biosecur Bioterror. **8** (2): 155-169.

Practical consequences

(besides high-profile publications)

- Three technological patents with broad applications,
- Creation and financing of a Bio-Pharma company (Pherecydes Pharma) specialised in biodefense & biosecurity (40% owned by BM-Systems),



- Research program with French civil and defence health-services,
- Discussions with US food industry firms,
- Contract with anti-bodies producer & discussions with enzymes producers.

All this in less than 3 years.

From our point of view, under the street light is definitely not the right place to search!

But how far can we go?

Reconstruct, from differential quantitative proteomics data (300 + proteins), the brain mechanisms associated with chronic anxiety (forms of synaptic plasticity deregulation + metabolic & mitochondrial adaptations + differential effects upon signal versus noise discrimination + differential effects upon neuronal networks + effects upon learning capacity + etc...)

Iris F, Filiou M & Turck CW (2014): *Am J Psychiat Neurosci*, **2**: 25-42.
Online access: ISSN: 2330-426X.

Or, discover a novel treatment for Parkinson's disease that should be directly applicable to human patients without prior animal trials

Two long known, entirely non-toxic drugs given at different times (no possibilities of adverse drugs interactions).

Or, decipher sex-specific embryological phenomena that have resisted elucidation for well over 60 years.

Iris F. *et al.* (2009); *Med Sci.* 25: 608-616.

Etc...

For more information: Publications to date

Publications in prestigious peer-reviewed journals: (click on the links to get the pdf)

- [2014, CNS Psychiatry publication](#): American Journal of Psychiatry and Neuroscience. Second publications with the Max Planck Institute of Psychiatry in Munich: Differential proteomics analyses reveal anxiety-associated molecular and cellular mechanisms in cingulate cortex synapses. The first output of the DECIUS CNS research program.
- [2012, CNS NEURODEGENERATIVE & PSYCHIATRY](#): PharmacoPsychiatry publishes the first review describing a productive vision of Systems Medicine that will change R&D organization and interactions between clinicians & researchers & reveals how the world's first explanation of the mechanisms of the Creutzfeldt-Jakob disease led to the discovery of a truly innovative psychiatric treatment.
- [2011, CNS PSYCHIATRY](#): Pharmaco Psychiatry publication: Proteome-Based Pathway Modelling of Psychiatric Disorders. Publication with The max Planck Institute of Psychiatry in Munich
- [2010, INFECTIOUS DISEASES](#): Biosecurity and Bioterrorism: Biodefense Strategy, Practice, and Science :Genetically Engineered Virulent Phage Banks in the Detection and Control of Emergent Pathogenic Bacteria. Publication with Pherecydes-Pharma.
- [2009, TISSUE DIFFERENTIATION](#): Médecine & Sciences: Müllerian duct regression explanation. Integrative systems biology & experimental Biology. Publication with CNRS experimental data.
- [2005, CANCER](#): Journal of molecular Endocrinology: Integrative analysis of gene expression patterns predicts specific modulations of defined cell functions by estrogen and Tamoxifen in MCF7 breast cancer cells. Publication in collaboration with INSERM unit 553.
- [2003, CANCER](#): Nucleic Acids Research: Integrated transcriptome analysis of the cellular mechanisms associated with H-ras-dependent malignant transformation of the human breast epithelial MCF7 cell line. Publication in collaboration with INSERM unit 553. World first. First in-silico model of a complex human disease validated in-vitro and published.

Collaboration to scientific reference books:

- [2014, Dermatology Cosmetics](#). The first reference book on “Computational Biophysics of the Skin” edited by Prof. Bernard Querleux , scientific chairperson of the International Society for Biophysics and Imaging of the Skin
- [2011: Phage Nano Technology](#) book published by [Valery Petrenko](#). Chapter 8: Genetically Engineered Virulent Phage Banks for the Detection and Control of Bacterial Biosecurity Threats.
- [2008, CNS](#): Biomarkers for Psychiatric Disorders. (Ref. ISBN: 978-0-387-79250-7, November 2008). Dr. François Iris, is the author of the Integrative Biology chapter of the book. The editor, Prof. Christoph W. Turck, is head of the Proteomics and Biomarkers branch at the Max Planck Institute for Psychiatry

Nevertheless, it MUST be remembered that

A model is an ASSISTANCE to thoughts

and certainly NOT a substitute for thoughts.

Our collaborative R&D programs & their outputs

This list excludes our contractual research programs with our clients



CEA : **"Creutzfeld-Jacob Disease CJD"** World's first in vivo validation of the mechanisms of Creutzfeldt-Jakob disease pathogenesis & progression. US, EU & French Awards; Awards (2009 and 2010) . CEA SEPIA department.

Successfully completed; **1 publication**.



CEA: **CNS disorders**. Collaborative research program that led to a [novel therapeutic strategy](#) for the treatment of psychiatric and neurological disorders. Copatent [WO/2010/029131](#)– **Use of anti-connexin agents for modulating the therapeutic effect of psychotropic drugs**. September, 2008 **CEA/BMSystems**,

[Pherecydes-Pharma](#) **BMSystems' spin-off created in 2006, novel M.R. anti-bacterial nano-agents biotherapies 3 patents**. Two indications: [Multi-resistant Skin infections](#) and osteo-articular infections.



Max Planck Institute (Munich): **Project "Chronic Anxiety"**.

Successfully completed; **3 publications** & a **Reference Book "Biomarkers for Psychiatric disorders"** chapter 19.



INSERM: **3 Projects "Tumoral Progression"; "Therapeutic Resistance"; "RGD 15 & Metastasis"**.

All 3 successfully completed, **3 publications**.



CNRS: **Project "Müllerian Regression"** Tissue differentiation

Successfully completed, **1 publication**.



Foundation FondaMental: **Project "Bipolar Disorders & Schizophrenia"**.

Immuno-inflammatory hypothesis. On going, **1 publication pending**



L'OREAL Arkema, Rhodia/Solvay ARD : **"Synthons" Government funded feasibility Program at IAR cluster Industrial Biotech**

Feasibility study Completed 16 molecules evaluated, **2 strains built, 1 program with 1 patent (industrial partner only)**

- Skin Homeostasis: **Reference book "Computational Biophysics of Skin"** chapter 15 with Dr. Querleux (L'Oréal)



Centre of excellence in Epigenetics IISER Pune India: **Project "Etiology & Epigenetic for metabolic disorders"**

Etiology & Epigenetic for metabolic disorders, on going **1 publication pending**

Questions

