



**Integrative Biology
Systems Biology**

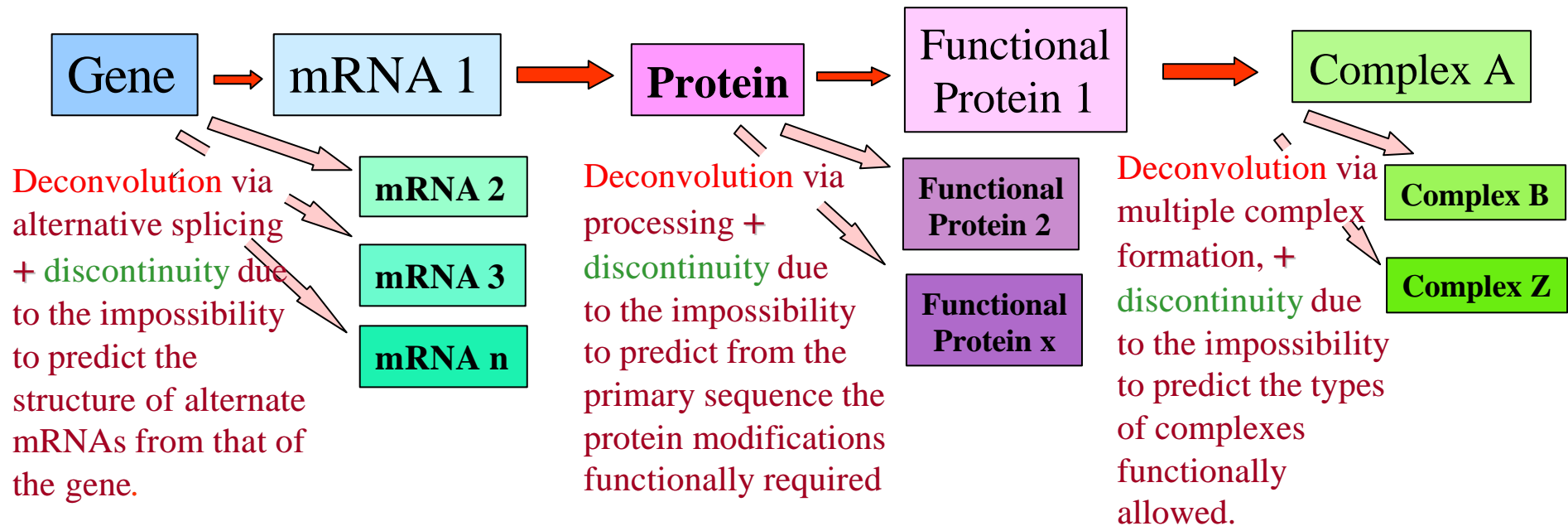
**Complex Systems Analyses & Biological Models
for
Therapeutics Engineering**

**Dr.françois iris Chairman & CSO Bio-Modeling Systems: info@bmsystems.net
Dr. François Xavier Directeur de Recherche CNRS, responsable Club Biotech ECRIN**

From genes to physiological functions

Four series of **deconvolutions** and **discontinuities** :

One gene = several different physiological functions

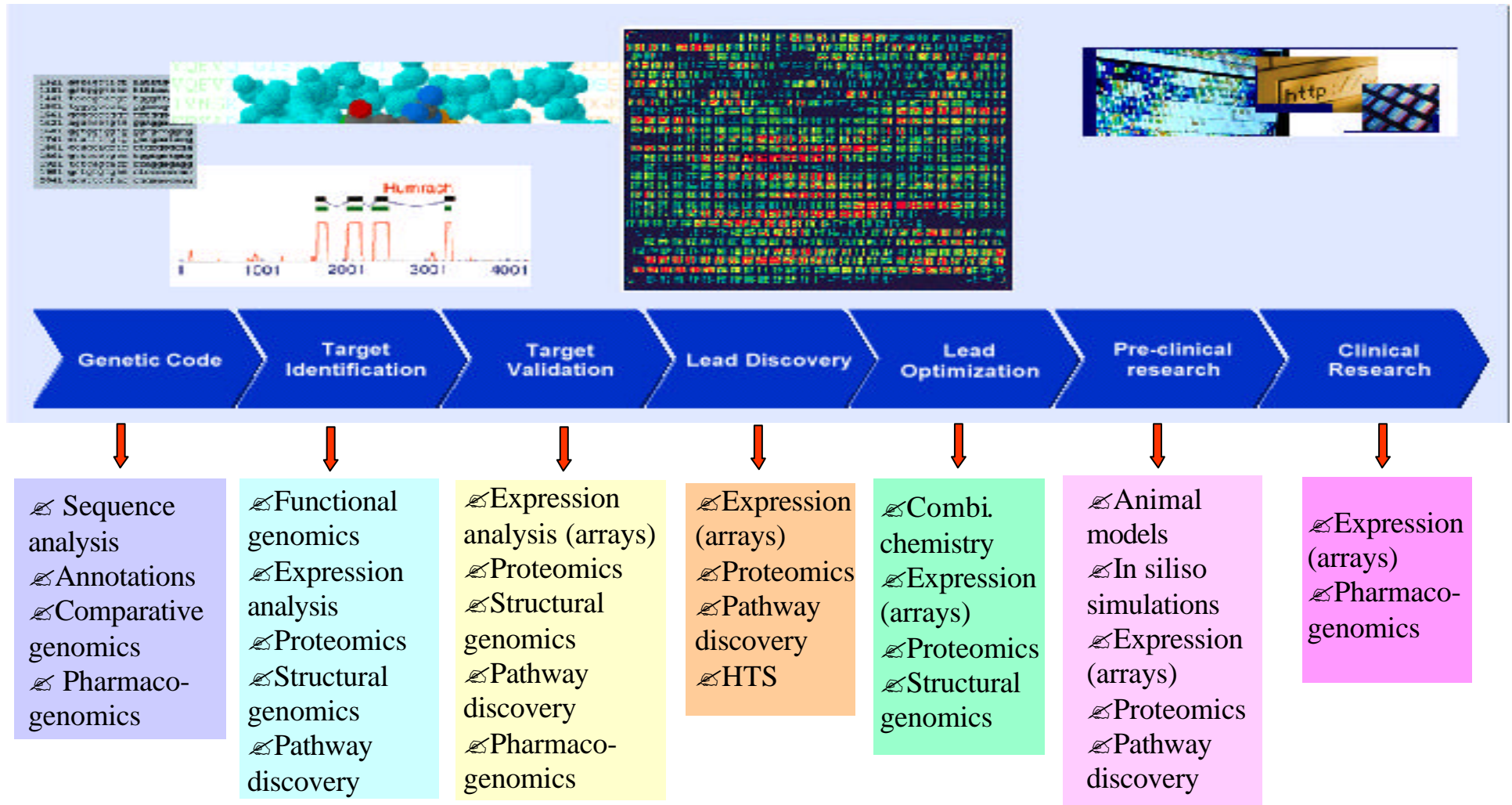


A non-linear integrative system.

Characteristics of complex systems :

- ✍ Interactions between components;
- ✍ Non-linear functional links;
- ✍ Open-type system;
- ✍ Multiplicity of local contexts;
- ✍ Multi-functional and context-dependent components
- ✍ Internal deconvolutions + discontinuities = functional state

From complex systems to Drugs :



A process generating a flood of information

The necessity for an integrative systemic approach

- Enables **collection and organization** of accumulated information.
- Encompasses the **totality of the known components** involved.
- Puts emphasis on components' **interaction and interdependence**.

 **Generates pertinent knowledge about the addressed biological problem.**

Information




Knowledge



What type of information?

ALL information pertaining to a process, a fact or an observation directly and /or indirectly attached to the problem being analysed .



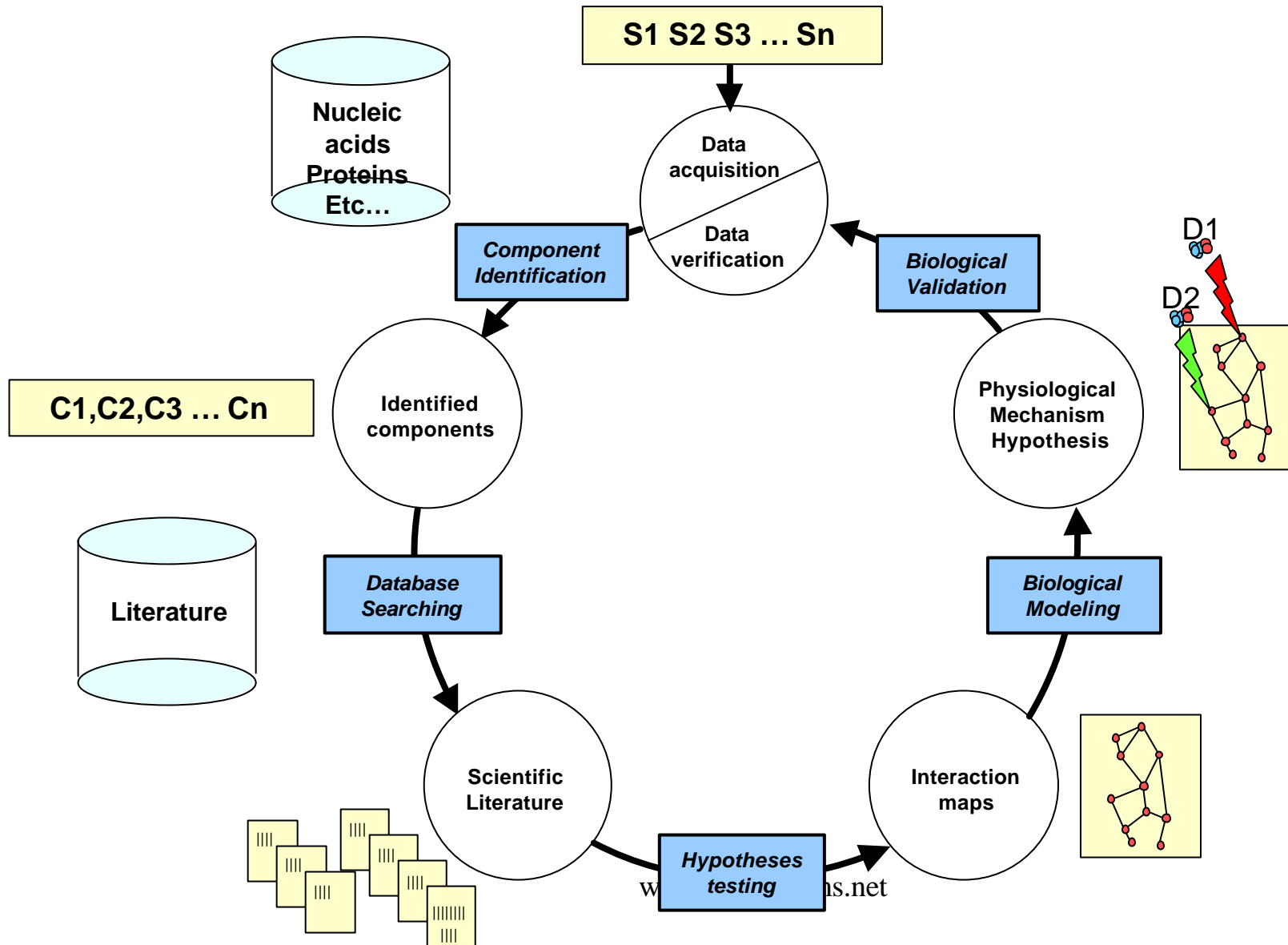
In the bio-medical realm, the relevant information touches on every biological domain, from anatomy to molecular biology and structural biophysics, spanning a multitude of functional contexts, corresponding to an enormous complexity.

The available information is necessarily

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

Generating a Physiological Model

From Information to Knowledge



What is an integrated physiological model?

It is a detailed map of the cellular mechanisms associated with a pathological state, allowing direct identification of

 the pharmacological targets, and

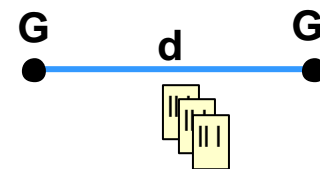
 the types of molecules

expected to produce defined therapeutic effects.

The 1st Generation: *Bio-Graph*TM

« *Graph* » indexation strategies of functional entities

- The whole of Medline
- All biological components

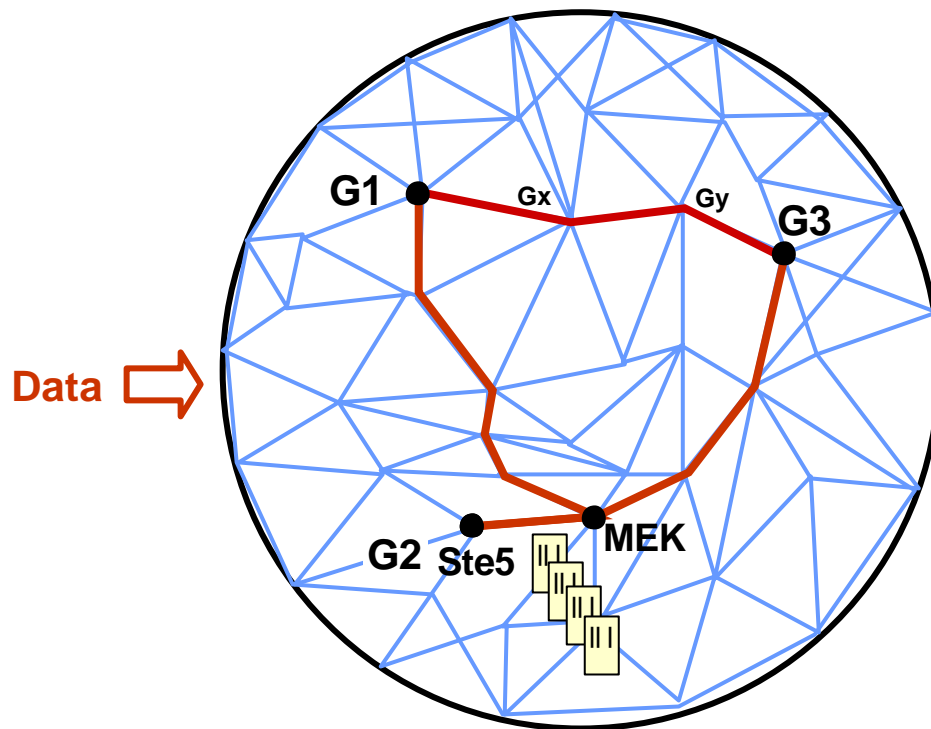


- Distance of co-occurrence :

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

- Functional relationships

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

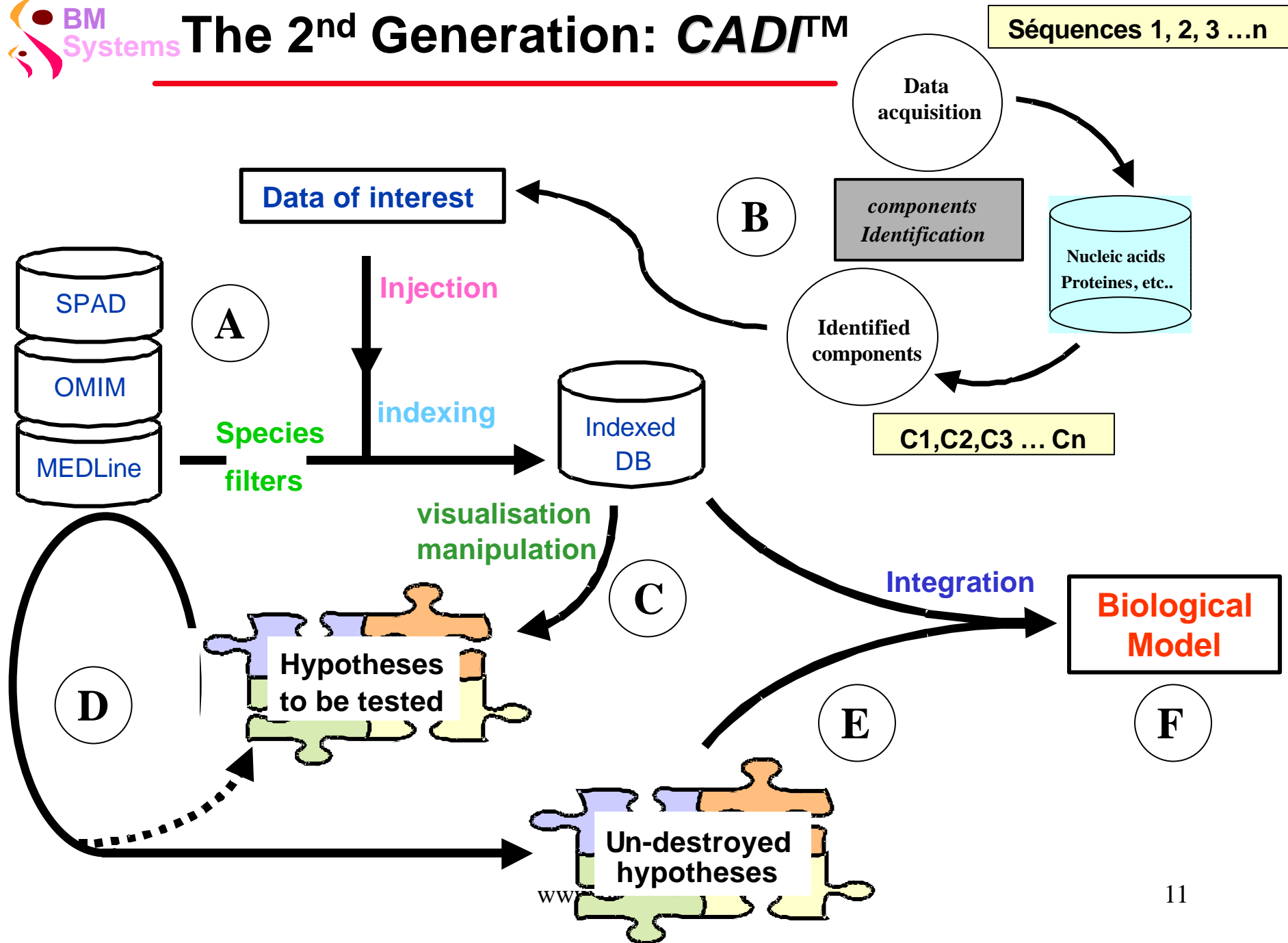


Ste5 simultaneously binds **MEK**

- **Sub-graph = the shortest paths linking components**

Problems associated to Graphs

- ✍ Enormous curation effort required prior to indexing.
 - ✍ Numerous internal incoherencies within the graphs.
- ✍ Extremely difficult to contextualise in a coherent manner.
 - ✍ The injection of 65 transcripts generates a multitude of potential paths (over 1000) within a graph.
- ✍ Few qualitative criteria applicable upstream to indexed data.
 - ✍ The vast majority (80+%) of the paths generated are fallacious.
- ✍ Data belonging different domains are extremely difficult to co-index.
 - ✍ Co-indexing transcriptome & proteome data generates aberrations.
- ✍ Modelling capacities limited to single cell-type + data injection.
 - ✍ Cannot approach systemic problems.



Tumour Progression: *MCF-7* vs *MCF-7 ras*

The very first model of a complex human pathology ever published

MCF-7: Breast epithelial cell

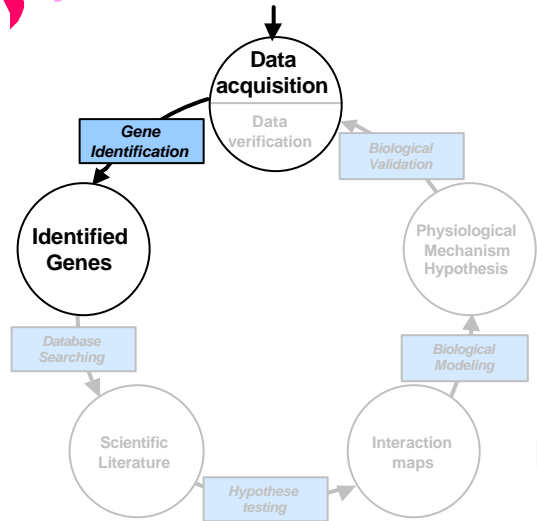
- Latently tumoral
- Hormono-dependent

MCF-7 ras: MCF-7 transfected with constitutively activated h-ras

- Aggressively invasive
- Hormono-sensitivity but no dependency

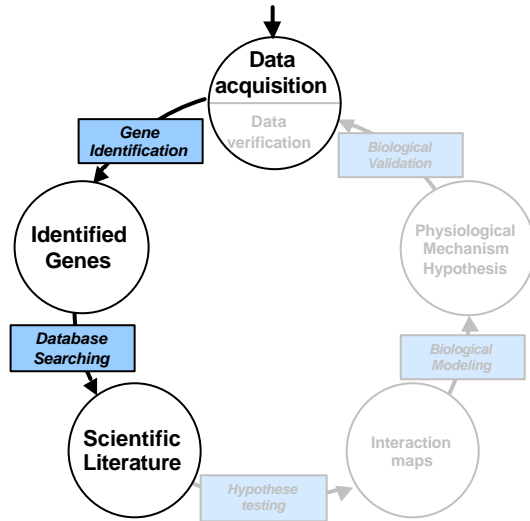
Differentially Expressed Genes

■ Over-expressed
■ Under-expressed

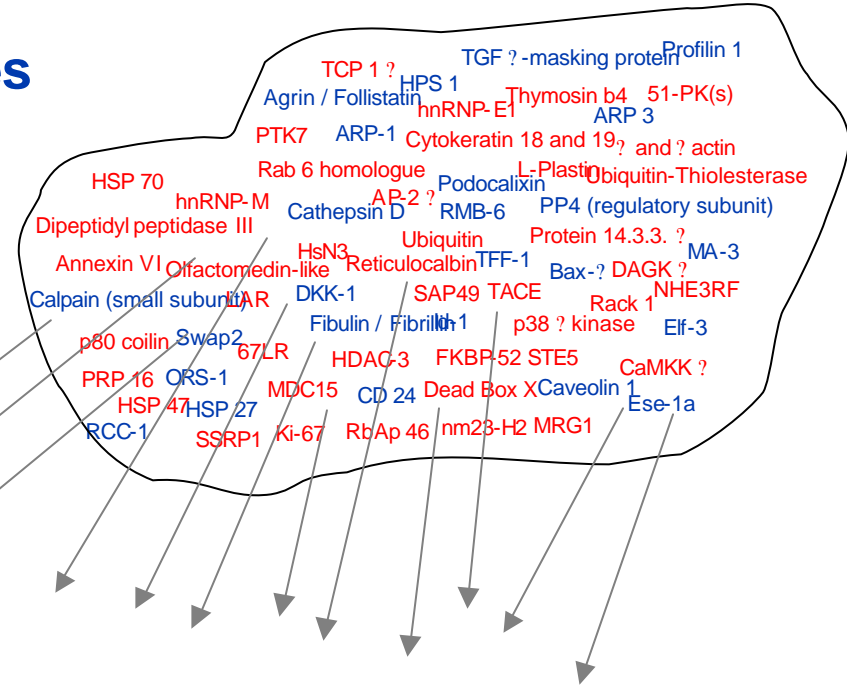


TCP 1 ? TGF ?-masking protein Profilin 1
Agrin / Follistatin HPS 1 51-PK(s)
PTK7 ARP-1 hnRNP-E1 Thymosin b4 ARP 3
Rab 6 homologue Cytokeratin 18 and 19 ? and ? actin
HSP 70 L-Plastin Ubiquitin-Thiolesterase
Dipeptidyl peptidase III hnRNP-M Cathepsin D AP-2 ? Podocalixin RMB-6 PP4 (regulatory subunit)
Annexin VI Olfactomedin-like HsN3 Reticulocalbin Ubiquitin Protein 14.3.3. ? TFF-1 MA-3
Calpain (small subunit) DKK-1 SAP49 TACE Bax-? DAGK ?
p80 coilin Swap2 Fibulin / Fibrillin Id-1 p38 ? kinase Rack 1 NHE3RF
PRP 16 ORS-1 67LR HDAC-3 FKBP-52 STE5 CaMKK ?
HSP 47 HSP 27 MDC15 CD 24 Dead Box X Caveolin 1 Ese-1a
RCC-1 SSRP1 Ki-67 RbAp 46 nm23-H2 MRG1

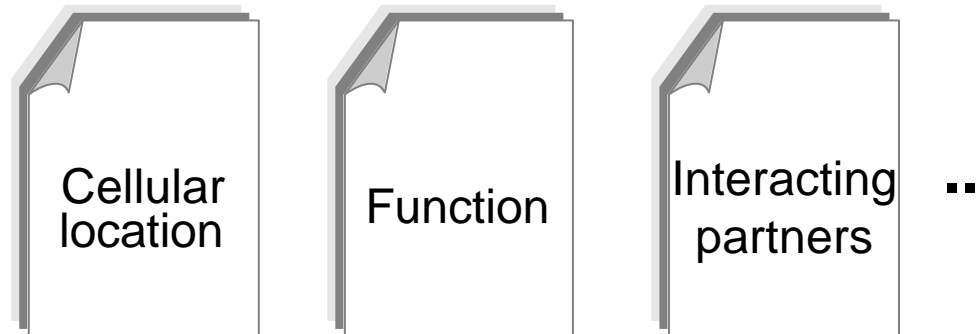
Information Retrieval and Extraction

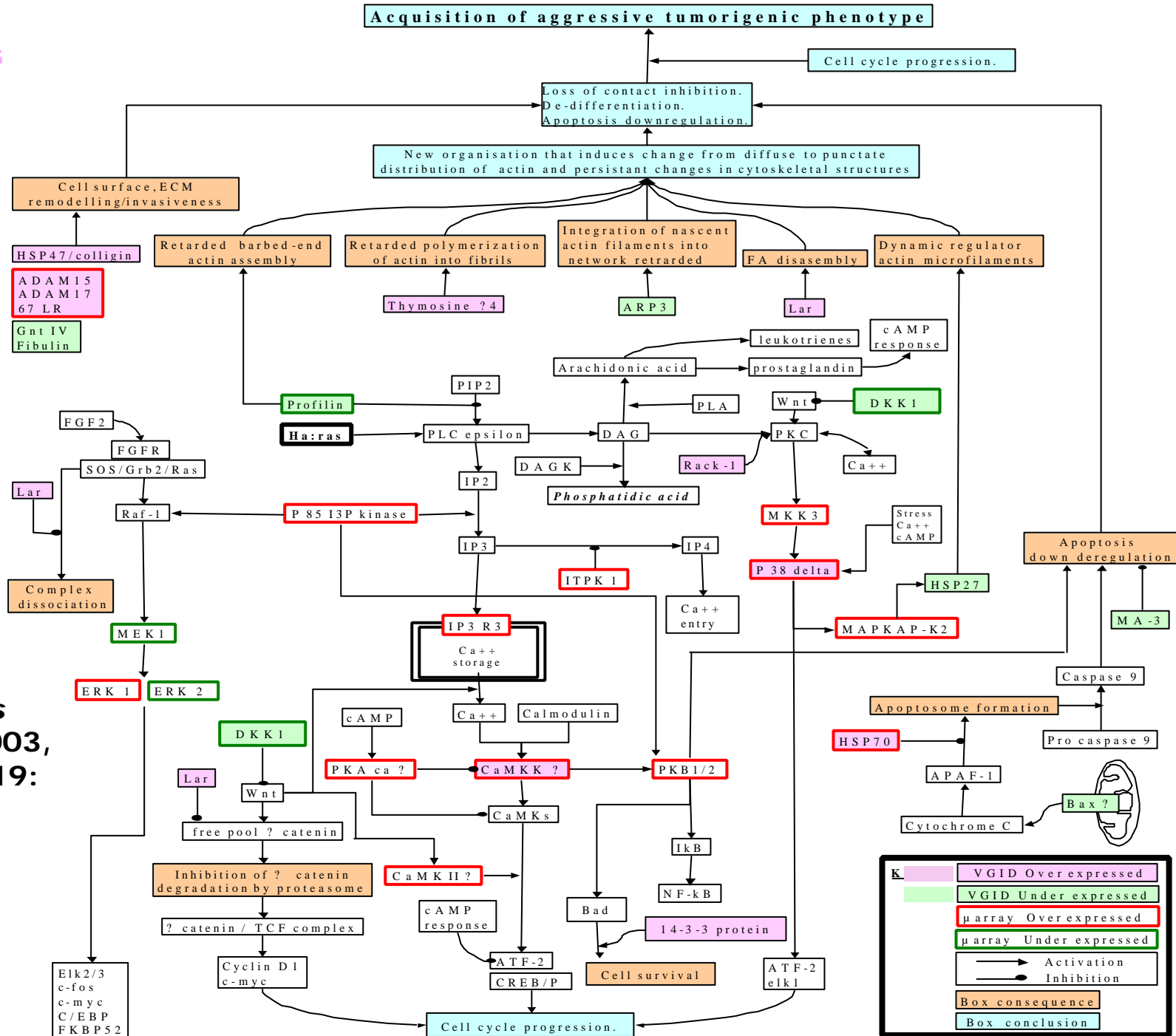


Set of genes



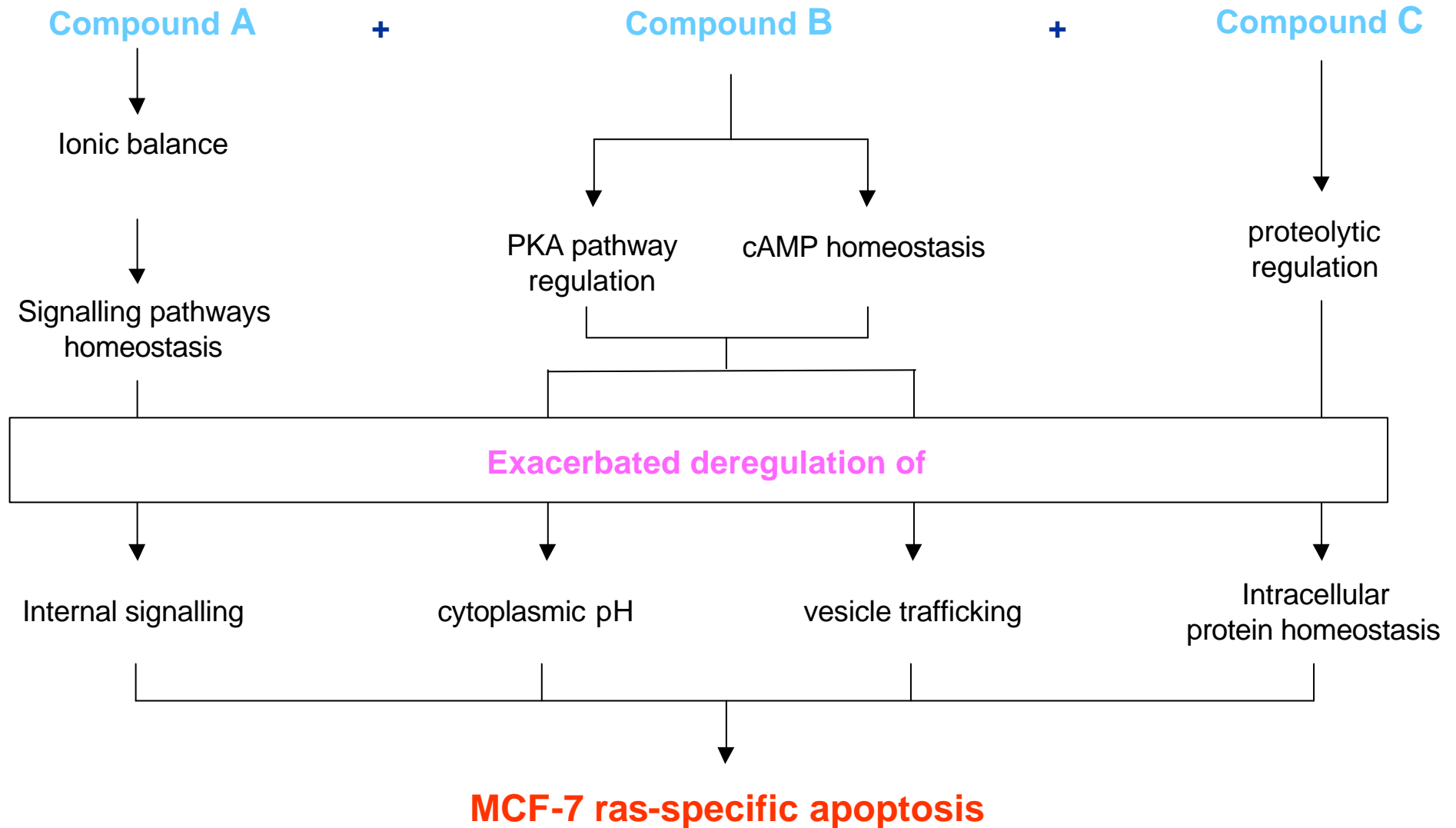
Set of papers





Published in:
**Nucleic Acids
Research, 2003,
Vol. 31, No. 19:
5789-5804**

Anti MCF-7 ras intervention using sub-optimal doses (nM)



Validation of the breast cancer model

The theoretical model made three types of predictions:

-A) the cellular mechanisms.

The model predicted the expression patterns of 13 key genes associated with the physiological changes revealed during the model-building process.

These predictions were independently tested, using RNA-chip technologies, at Hospital Tenon.

-B) the therapeutic targets

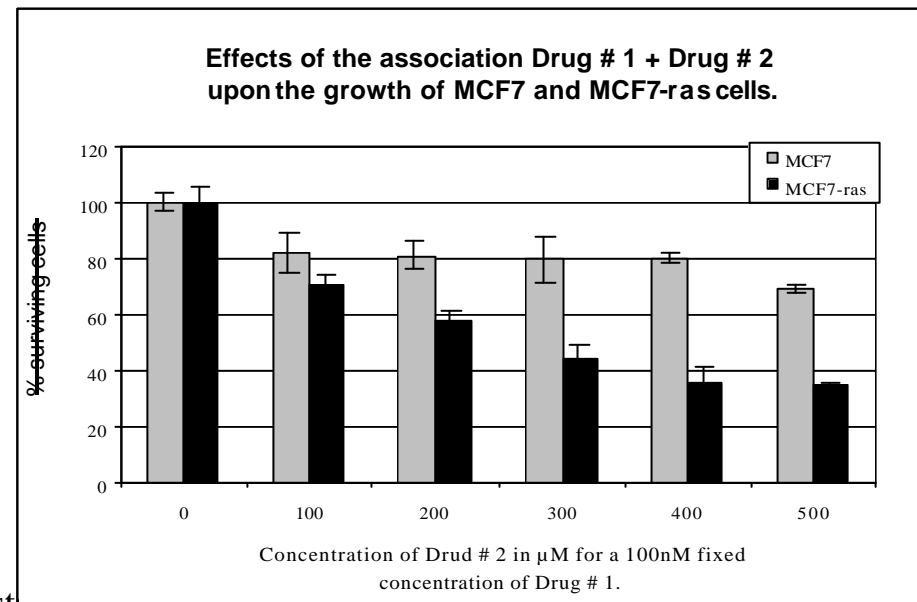
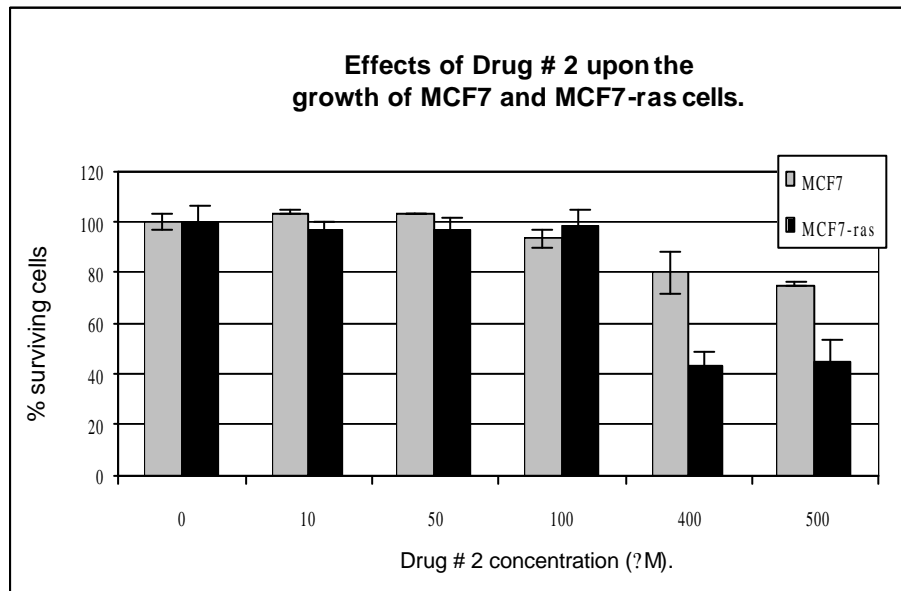
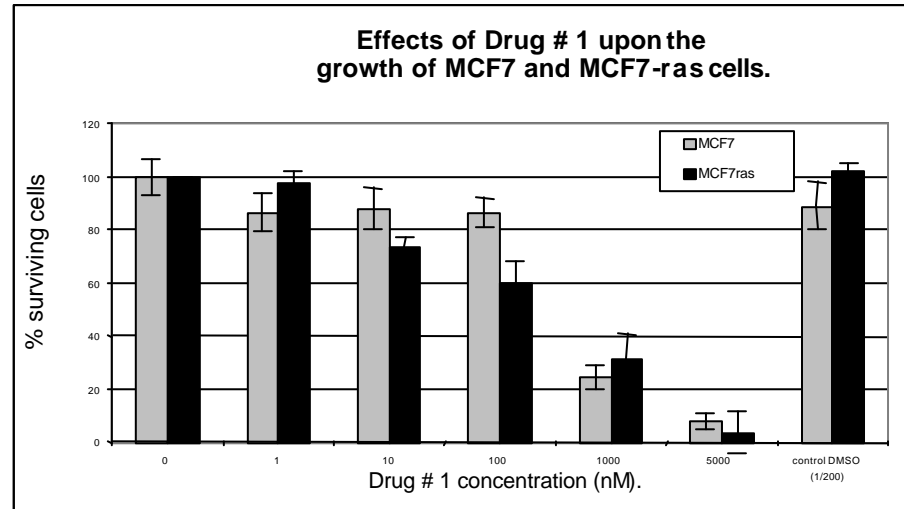
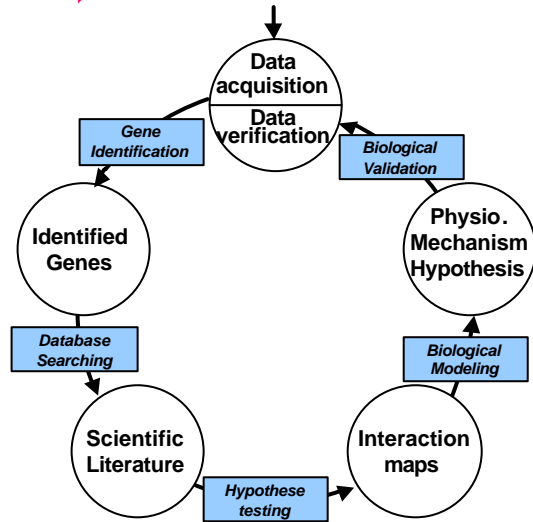
The model indicated three different cellular processes as being key to the maintenance of the hormone-sensitive malignant state. In each case clearly defined protein targets (isoforms level) were identified.

-C) the types of therapeutic interventions required

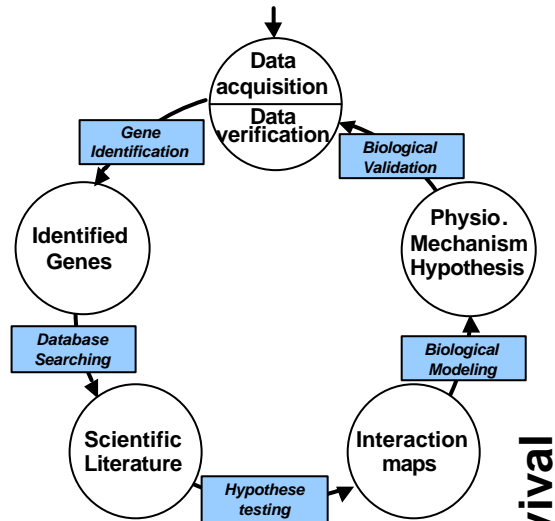
The model indicated three different molecules which, in combination and at sub-optimal concentrations, would have the required effects on the protein targets of cancer cells, leaving non-cancer cells largely unaffected.

These predictions were directly and independently tested on the cells by cancer specialists at Hospital Avicenne (Prof.M.Crépin), Hospital Tenon (Prof.R.Lacave) and the CEPH Institute (Prof.L.Cazes) in Paris.

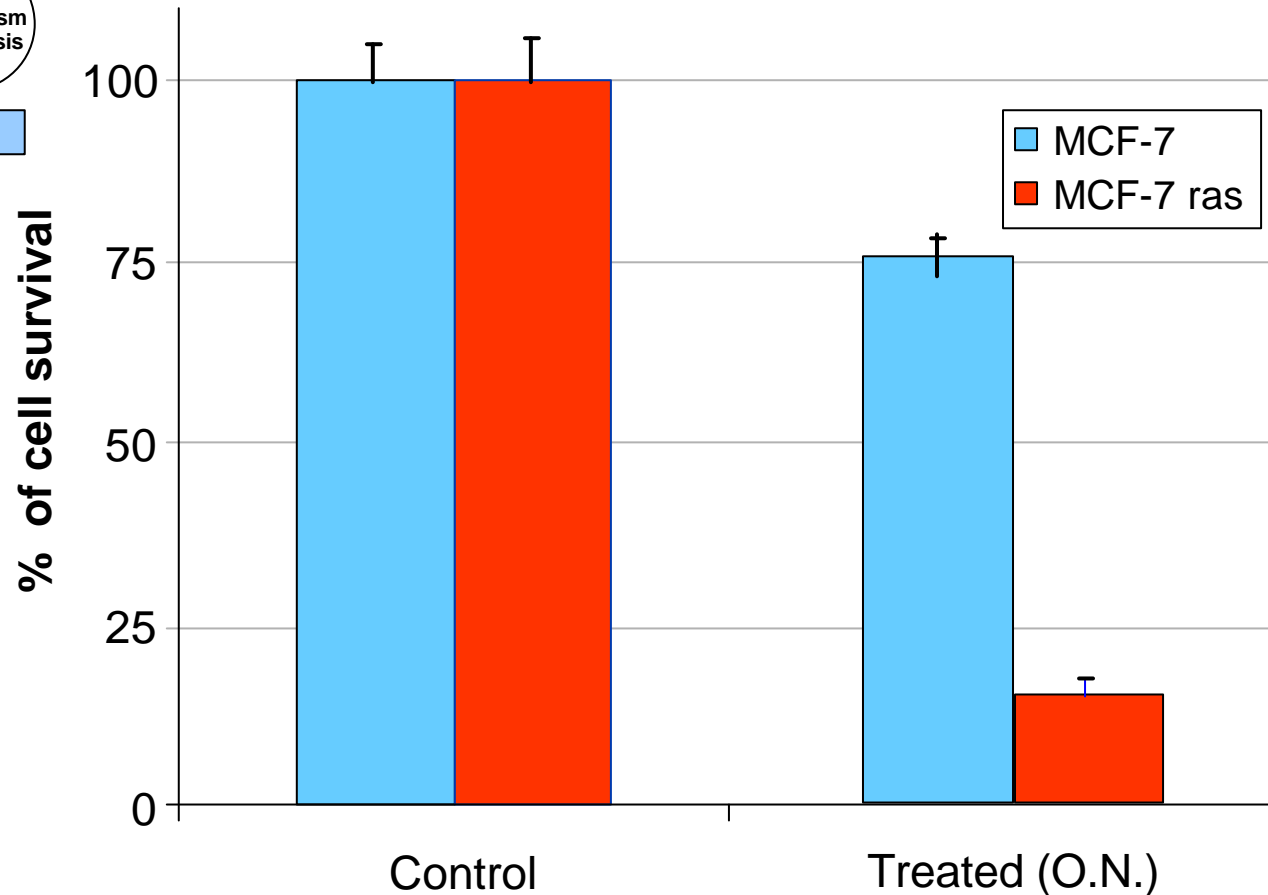
Biological Validation



Biological Validation



Effects of the association Drug # 1 + Drug # 2 + Drug 3 upon the growth of MCF7 and MCF7-ras cells.



CADI™ generated models:

To date, several physiological models have been generated using this analytical procedure :

- ✍ **Ras-dependent breast cancer * : Validated**
- ✍ **Creutzfeld-Jakob Disease (CJD): Undergoing tests (CEA)**
- ✍ **Chronic Fatigue Syndrome (CFS): Partly validated**
- ✍ **FGF4-driven fibroblast differentiation* : Untested**
- ✍ **Hypercholesterolemia: Untested**
- ✍ **Eradication of Multi-Resistant bacteria: Untested**
- ✍ **Hepatitis C: Untested**
- ✍ **Etc...**

* = Models constructed using differential expression data as inputs.

All other models were constructed using pathophysiological / medical information as inputs

What can a model achieve ?

- **Find a new application, or**
- **Rescue an existing molecule...**

- ✍ Modelise the physiological mechanisms in which failure causes of the past become therapeutic opportunities (« curare» effect);
- ✍ Utilise bio-medical information to identify the pathologies where these mechanisms are most strongly involved;
- ✍ Modelise the most pertinent pathology in order to identify:
 - ✍ the key processes;
 - ✍ the modulatory elements, and
 - ✍ the complementary targets.
- ✍ Inject into the model epidemiological and pharmacological information so as to identify:
 - ✍ the probable side effects, and
 - ✍ the appropriate exclusion / inclusion criteria.

...which implies investing in new clinical phases!

The practical consequences of a model :

A model allows to determine, for a given molecule:

- How to utilise the molecule;
- What is its mechanism of action and its drug interaction potential;
- How to modulate its effects (*required by the FDA*);
- What are its indication and /or exclusion criteria; etc....

This, however, cannot be applied to all molecules
(such as antibodies and peptides for example).

A model also allows to identify:

- ✍ Development programmes that should be abandoned;
- ✍ Mechanisms presenting therapeutic opportunities;
- ✍ Possibilities for alternative development, etc....

Unsolved problems and limiting factors

- ✍ **Automated** information retrieval & indexing
- ✍ Incompatibilities between the numerous existing tools.
- ✍ « Carry-all » data bases.
- ✍ Accumulation of erroneous data un-detectable *a priori*
- ✍ Accumulation of un-utilisable data (too slanted)
- ✍ Impossibility to integrate non-textual information
- ✍ Etc...

The necessity for a 3^d generation

Current biological modelling approaches and tools are restricted to information under text format only.

A huge body of information under the form of images, charts, diagrams, cryptic notation, etc...remains largely un-exploitable.

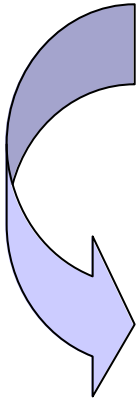


**It is necessary to develop
format-independent approaches.**

Evident in early 2002, this required in-depth scientific evaluation.

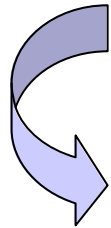
It is at this point that the first meeting with Françoise Xavier happened.

2002: Frequent scientific meetings between Françoise Xavier (CNRS / ECRIN) & François Iris (ECP / BM-Systems).



**Exploring the fundamental basis
of « Integrative Biology »
and its evolution potential.**

Mid 2003: Strategic planning. Basis on which public research and industry could cooperate to accelerate the development of « integrative Biology ».



The ECRIN plate-form identified as the most appropriate structure (*experience, independence, guaranties of quality & confidentiality, etc..*)



Create and facilitate
Public Research - Private Enterprise
upstream collaborations
for innovation and development

www.ecrin.asso.fr

Networks animator

22 people
(Full-time equivalents)
including 14 employees

Activities in 7 Fields

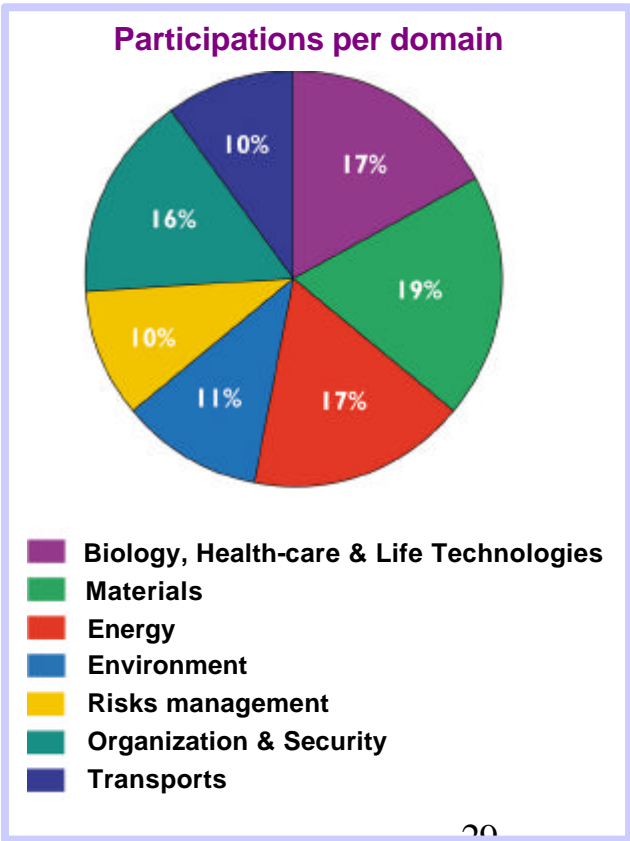
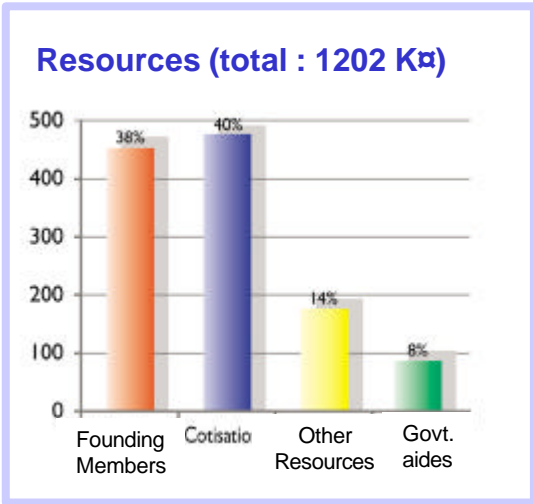
Members and partners

4371 participations

- ≈ 84 enterprises
- ≈ 58 research & transfer public organisations
- ≈ 5 out-posts including 4 technopoles
- ≈ 9 « grandes écoles »
- ≈ 83 laboratories in 24 Institutes and 34 universities

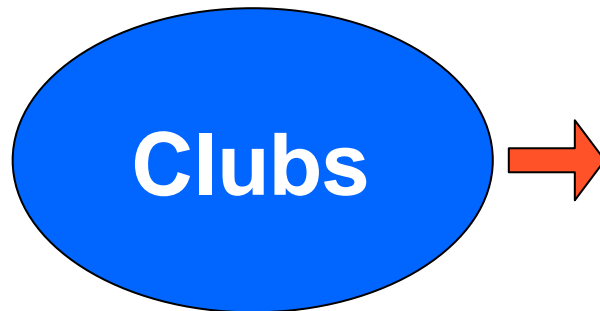
2002 - 2003

Budget



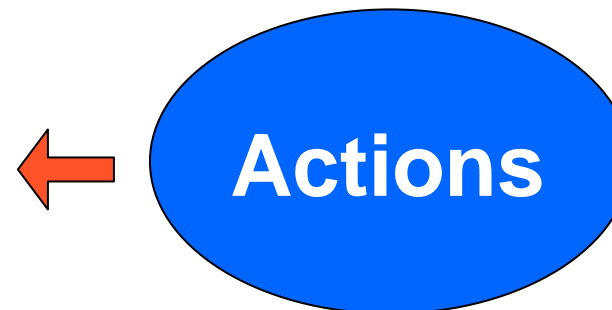
15 Clubs & 13 Actions

Permanent structures



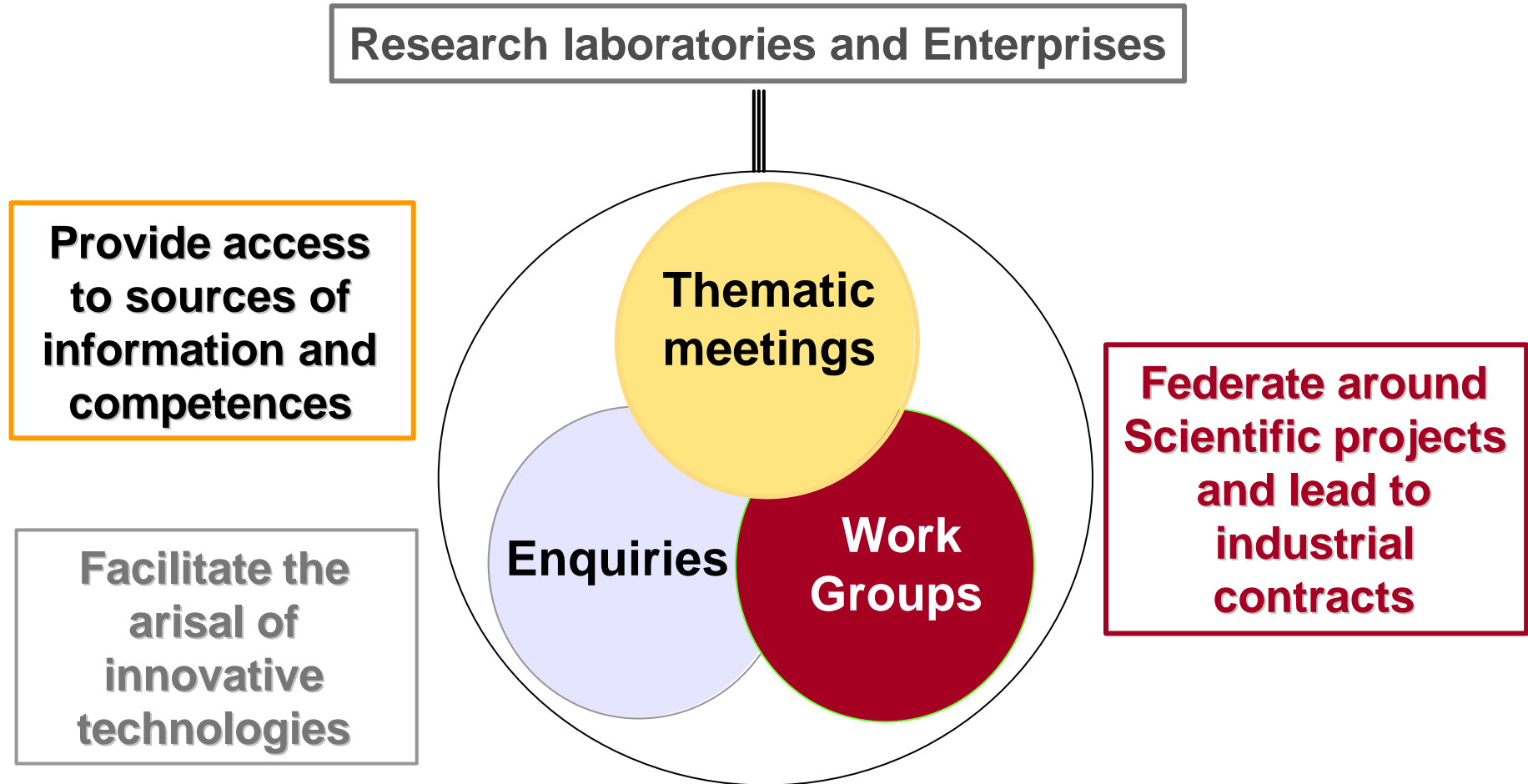
- ✍ Experts Networks
- ✍ High priority themes
- ✍ Scientific Intelligence

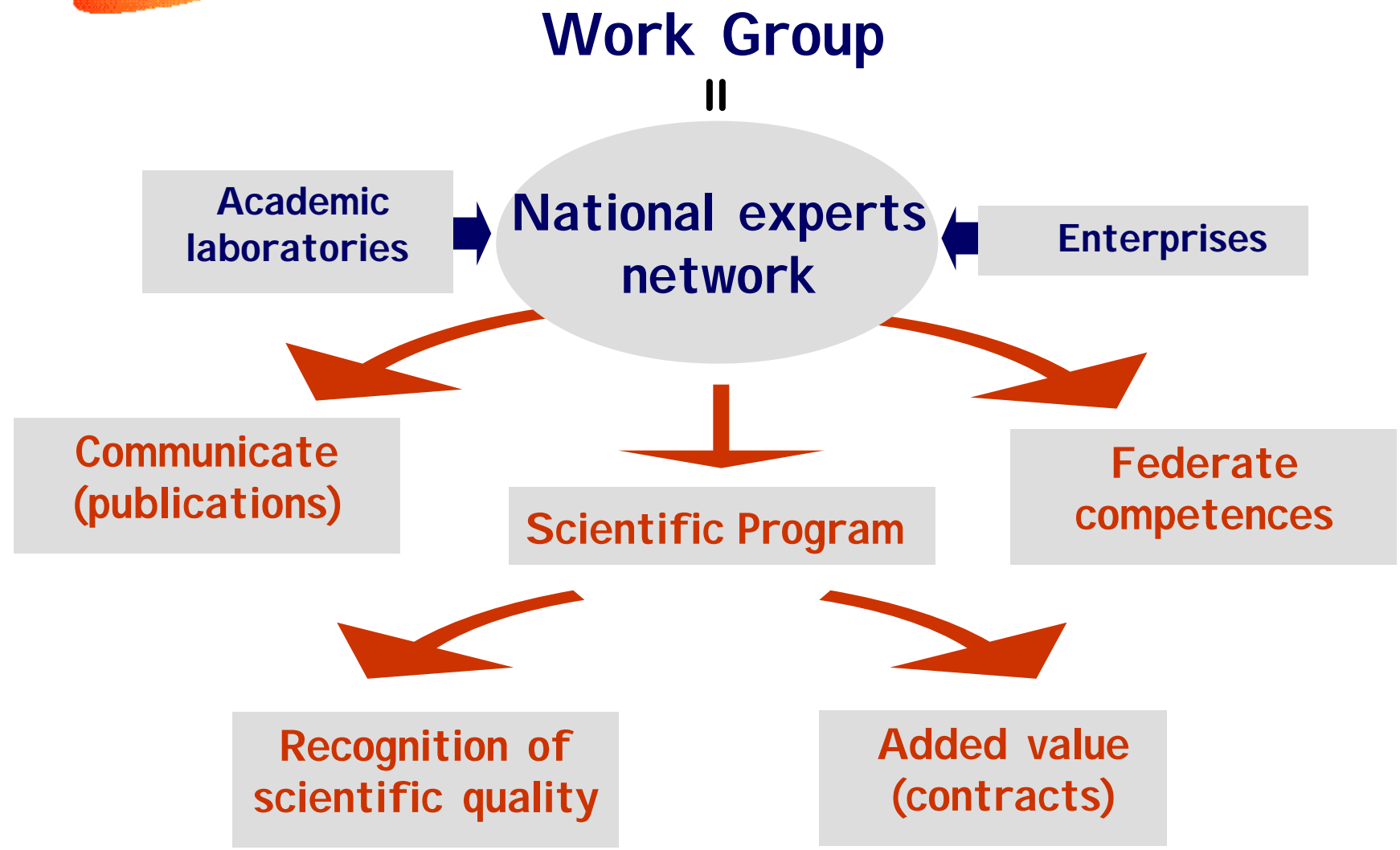
Flexible structures

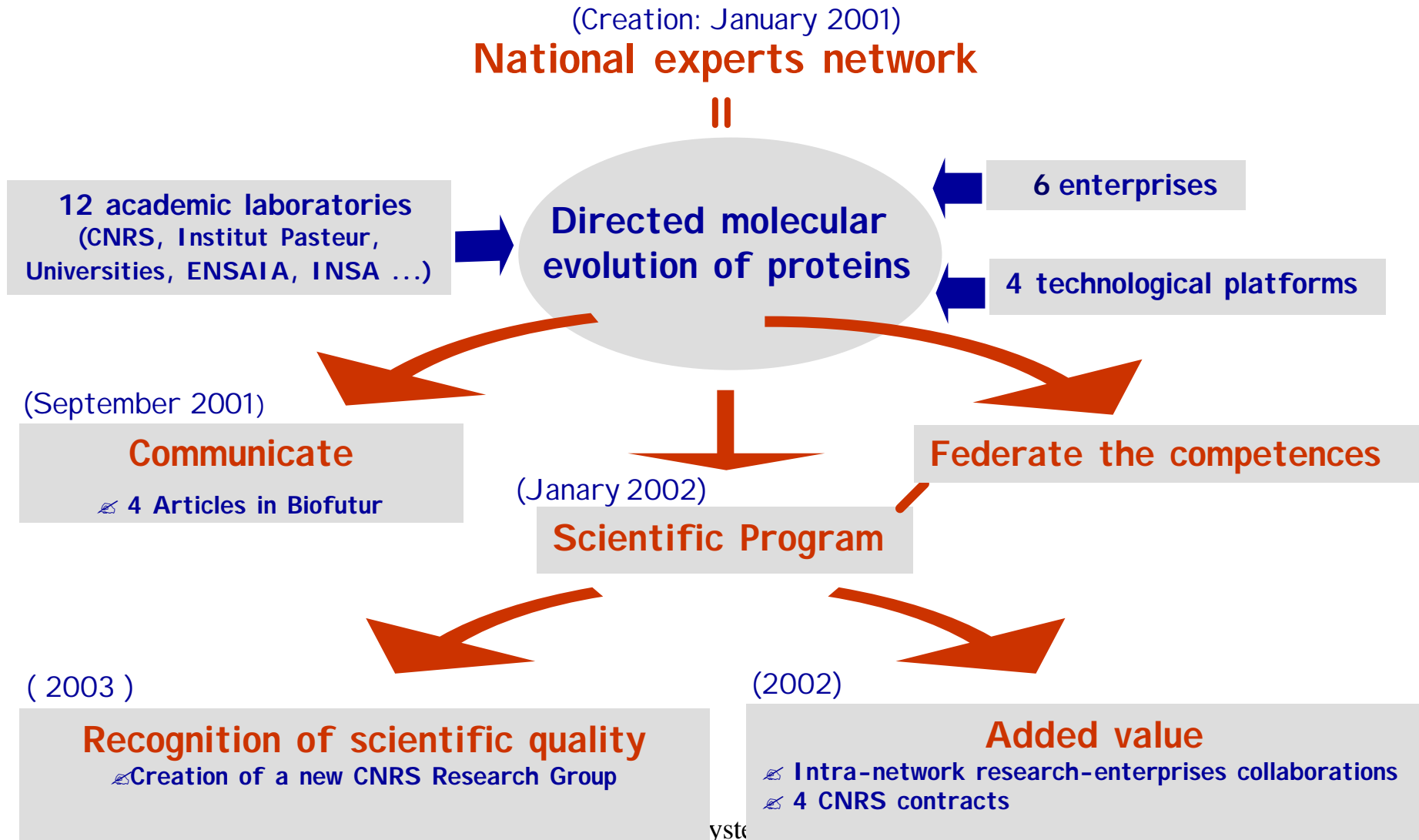


- ✍ Set duration
- ✍ Multi- disciplinary approaches
- ✍ Specific individual or collective demand.

Its functional mode





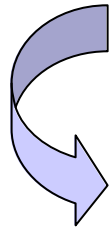


- **Independent and neutral entity**, federating private enterprises with representatives of public research organisations around a scientific program dedicated to a common goal.
- **Implement pluridisciplinary network**, involving academic research and private enterprises, thus creating synergistic interactions between scientific disciplines, technological approaches and public-private world views.



Fosters *creativity* and *innovation* and facilitates industrial development.

Mid 2003: Strategic planning. Basis on which public research and industry could cooperate to accelerate the development of « integrative Biology ».



The ECRIN plate-form identified as the most appropriate structure (*experience, independence, guaranties of quality & confidentiality, etc..*)

Early 2003: Evidence of real industrial interest for the innovative growth opportunities represented by « Integrative Biology ».

Late 2003: Creation of the « Integrative Biology » pole at ECRIN with the participation of 6 large Industrial & 1 specialised Biotech Companies.

Jan 2004: First work meeting of the « Physiological Modelling » Group at ECRIN. Next meeting scheduled for April.

Feb 2004: Inception of the Second Work-Group within the « Integrative Biology » pole at ECRIN.

March 2004: Creation of the scientific consortium attached to the « Physiological Modelling » Group.



ÉTHIQUE et DÉONTOLOGIE

INNOVATION ET TECHNOLOGIE DANS LE DOMAINE DU VIVANT

FORUM

**PLACÉ SOUS LE PARRAINAGE DU MINISTÈRE DE LA RECHERCHE
ET SOUS LA PRÉSIDENCE DE GUY PAILLOTIN, PRÉSIDENT D'ECRIN**

1^{ER} DÉCEMBRE 2004 AU MINISTÈRE DE LA RECHERCHE

Si vous êtes intéressé(e) par ce Forum laissez vos coordonnées au stand ECRIN