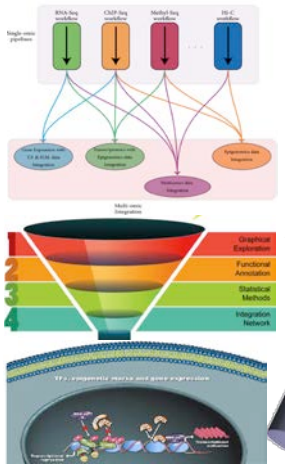
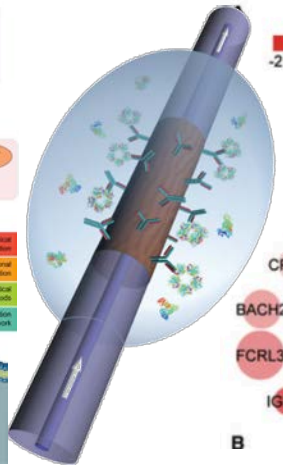
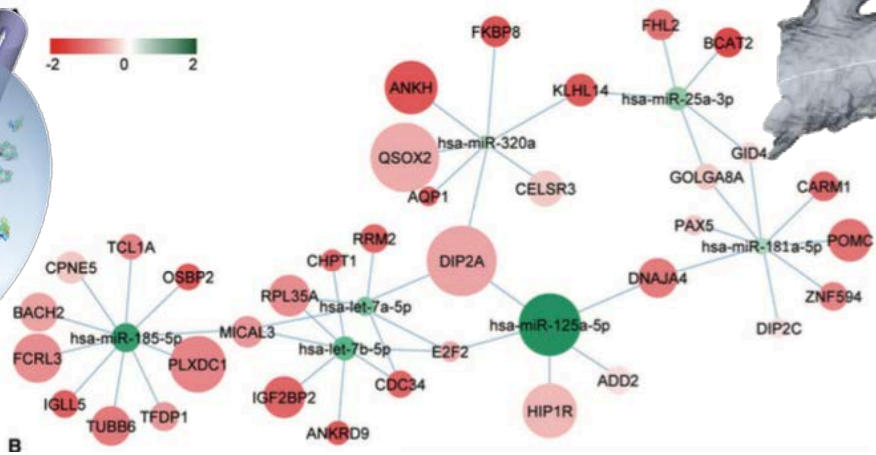
Project Area Biotechnology

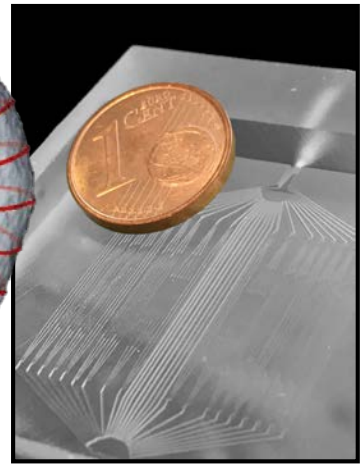
cnr DITET
Consiglio Nazionale delle Ricerche
Engineering, ICT and Technologies for Energy and Trasportation Department

Objective: Development of novel methodologies, algorithms, software tools, hardware devices, processes and technologies for clinical, medical and biotechnological applications.









Scientific Results:

AP Biotechnology is divided into 6 pillars fully in line with the European priorities:

- **Analysis, management and integration of biological big data**
- **Development of network-based approaches towards precision medicine**
- **Biophotonics**
- **Bioprocess and tissue engineering**
- **Mathematical software tools for bioengineering and biology**
- **Bio-sensors and Bio-inspired systems**

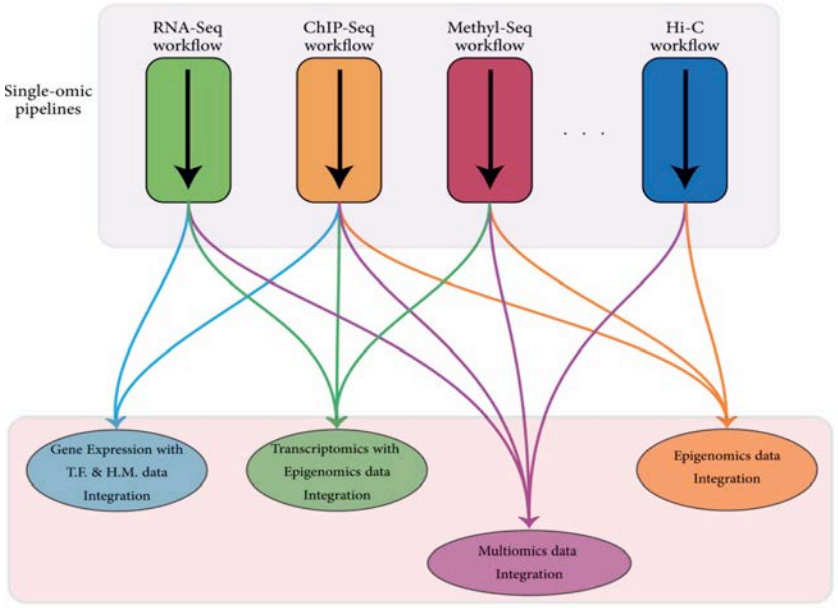
1. Biological big data

1. New techniques of analysis on big data to cope with the increasing volume of genomic and proteomic data.
    1. Biological database management systems
    2. Knowledge of biochemical mechanisms and processes that are the base of the life and death of cells.
2. Statistical machine learning algorithms
    1. Origin of cellular mechanisms
    2. Disruption under pathological conditions,
    3. Novel therapeutic targets and strategies.
3. Data analysis and predictive models
    1. Identification of the interactions among essential molecules
    2. Genetic and chronic diseases with a high economic impact.
4. FAIR (findable, accessible, interoperable, reproducible) scientific and technological research.

Biological big omics

Contribution

- **Development of novel statistical learning algorithms**
- **Multi-omics data analysis**
- **Single-cell data analysis**

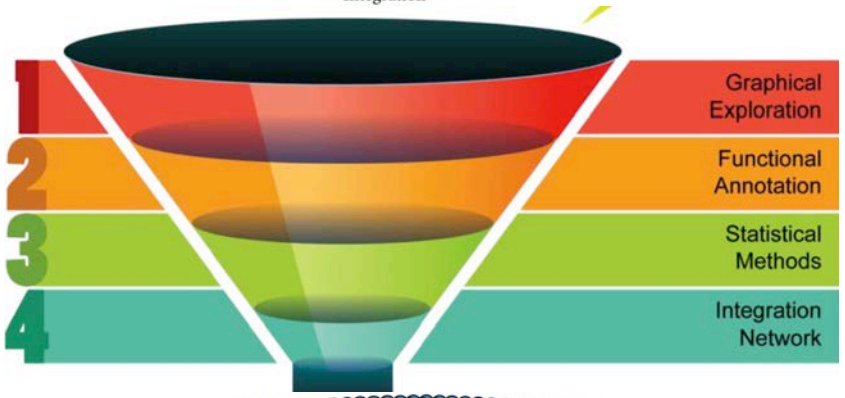


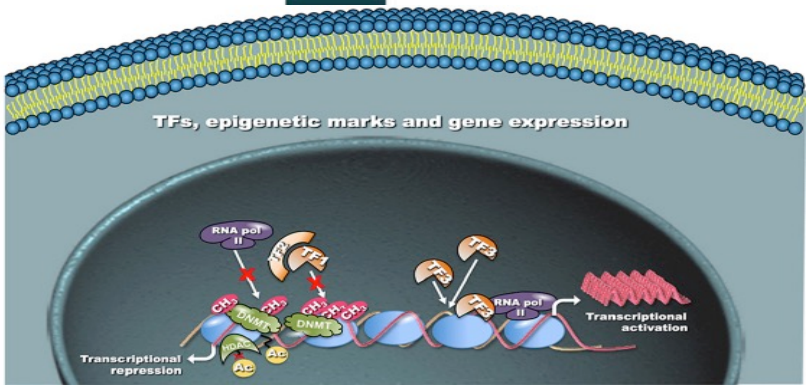
IMPACT and CHALLENGES

- Ultra-high dimensional data reduction
- Multi-omics data integration strategies
- Single-cell resolution

<u>In order to</u>

- Biomarkers identification
- Understand cellular mechanisms and their disruption under pathological conditions,
- Identify novel therapeutic targets and strategies
- Study epi-drugs mechamisms





Biological big data classification

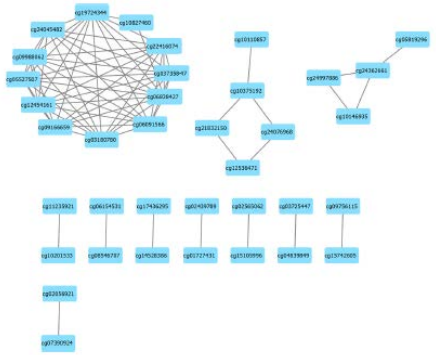
Typical problem in Bioinformatics:

- More than **1000 samples** (patients), **hundred of thousands features** (genes, sites, clinical variables, proteins)
- **Aim**: distinguish healthy vs diseased samples
- Not addressable by s.o.a. machine learning
- **Big Data solutions**

Data Set	Number of Samples	Number of Features
BRCA	897	485,512
THCA	571	485,512
KIRP	321	485,512

Big Biological Data Classifier (BIGBIOCL)

- Able to compute **multiple human readable classification models**
- **Based on Big Data technologies: Hadoop, Spark (mlib), and Cloud Computing**
- **BIGBIOCL** extracts multiple classification models by adopting a **feature elimination technique** and by **iterating the classification procedure**
- Available at: https://github.com/fcproj/BIGBIOCL and **published in**



Contents lists available at ScienceDirect

Big Data Research

www.elsevier.com/locate/bdr

Classification of Large DNA Methylation Datasets for Identifying Cancer Drivers ☆

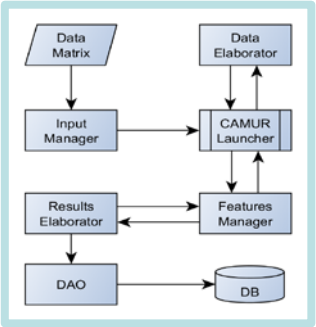
Fabrizio Celli [a], Fabio Cumbo [a,b], Emanuel Weitschek [c,a,*]

Biological big data classification

- **Classifier with Alternative and Multiple Rule-based models (CAMUR)**
- New method for classifying RNA-seq case-control samples, which is able to compute **multiple human readable classification models**
- **Aims of CAMUR:**

  1) To **classify** RNA-seq experiments
  2) To **extract** several **alternative** and **equivalent rule-based models**, which represent relevant sets of genes related to the case and control samples

- **CAMUR** extracts multiple classification models by adopting a **feature elimination technique** and by **iterating the classification procedure**



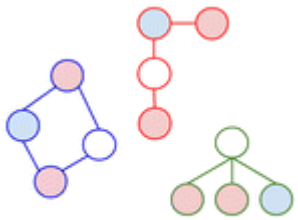
- **Prerequisite: Gene expression normalization**
  (RPKM = Reads Per Kilobase per Million Mapped)
- Available at: http://dmb.iasi.cnr.it/camur.php

Citation: V. Cestarelli, G. Fiscon, G. Felici, P. Bertolazzi, E. Weitschek: *CAMUR: Knowledge extraction from RNA-seq cancer data through equivalent classification rules*. **Bioinformatics**, 32(5): 697-704, 2016.

Data analysis and predictive models

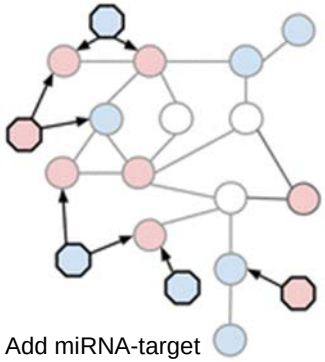
Study of the interactions of molecular components and their impact on many pathologies.

Pathway analysis



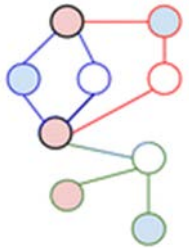
Find significantly altered pathways from miRNA expression data

MicroRNA regulation



Add miRNA-target interactions to PPI network

Network building



Merge all pathways in one network

For example:
study of the impact of miRNA molecules on mRNA repression, protein translation inhibition and their effects on cancer disease.

Impact and Challenges

Impact

Services and Tools

MiRNATIP: miRNA-Target interaction prediction.

MirTissue: A web service that characterize the type of miRNA-target interaction in specific tissues.

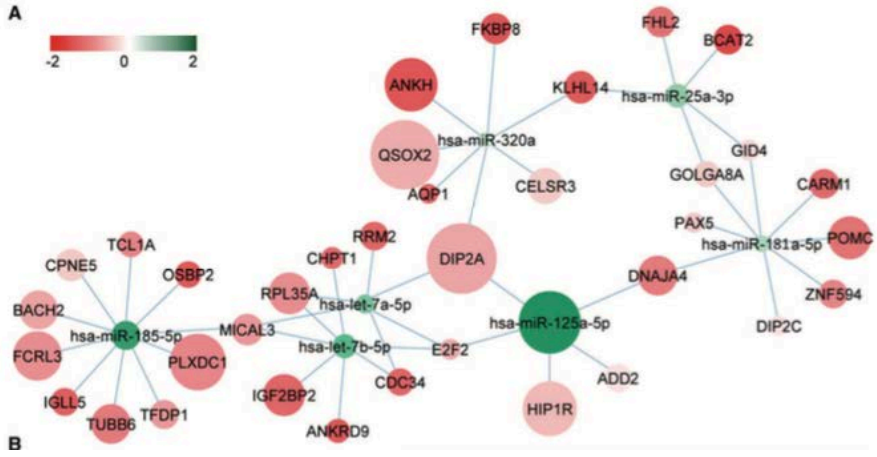
Contribution

Understanding of the mechanisms of development for pathologies like cancer diseases.

Development of non invasive tools for early and cheap diagnosis of pathologies connected with miRNA-mRNA interaction mechanisms.

Data analysis and predictive models

Contribution: integrated approach to identify circulating markers (miRNAs), target genes (mRNAs) and functional pathways associated with Pediatric Multiple Sclerosis.
Impact & Challenges: The interactions between the significant miRNAs and their targets uncovered genes related to immunological functions, as well as genes involved in autophagy-related processes and ATPase activity.



OXFORD

Human Molecular Genetics, 2018, Vol. 27, No. 1 66–79

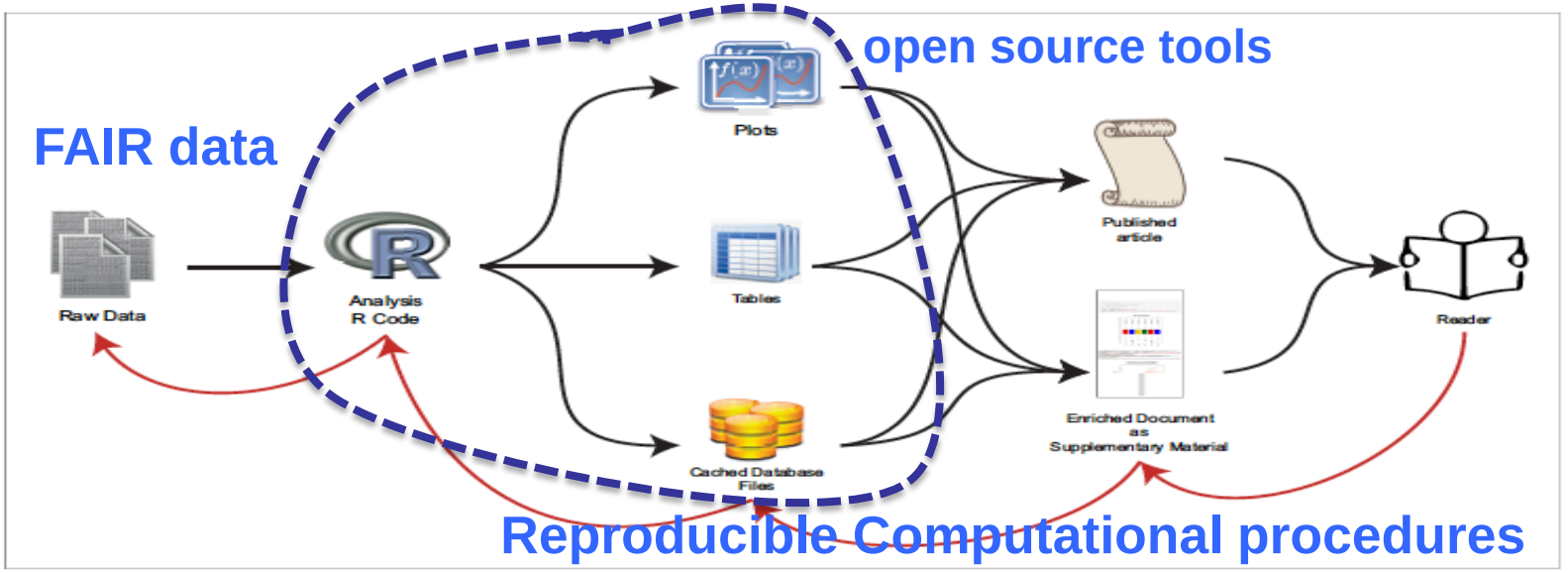
doi: 10.1093/hmg/ddx385
Advance Access Publication Date: 26 October 2017
Original Article

ORIGINAL ARTICLE

Combined microRNA and mRNA expression analysis in pediatric multiple sclerosis: an integrated approach to uncover novel pathogenic mechanisms of the disease

FAIR

- **Developmemt of open source tools for Reproducible Computational Research in the spirit of FAIR principles**



IMPACT and CHALLENGES

- Findable, accessible, interoperable, and re-usable data
- Reproducible computational procedures and graphical software interfaces

<u>In order to</u>

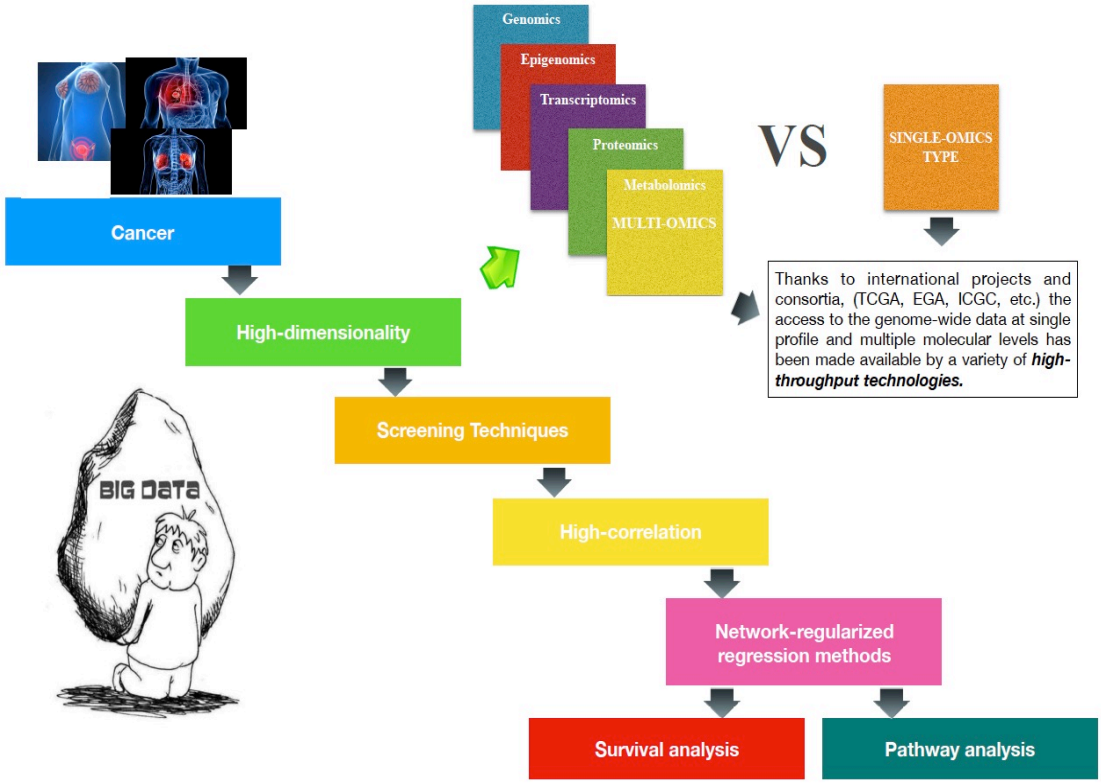
- Improve quality and transparency of findings
- Increase knowledge transfer
- Provide high level and high performance computational platforms

2. Network-based approach

- Precision medicine for a new era of health care delivery and treatment.
  - Understanding of the underlying mechanisms of high-impact diseases is allowing scientists to develop new drugs, targeted therapies, and preventive strategies.
  - Causes of high-impact diseases, including but not limited to human cancer, thus supporting the planning of healthcare services such as clinical trials and disease prevention.
- More precise prognoses to facilitate the subsequent clinical management of patients at risk of disease.
- Predictive “whole cell” computational models to exploited in personalized medicine,
  - Specific drug administration therapies tested on single cells or populations of virtual cells,
  - Adaptation of the individualized therapy to the environment.

Network-based approaches

- **Advanced statistical approaches for more precise prognoses to facilitate the subsequent clinical management of patients at risk of disease**
- **Network-based regression approaches**
- **Survival analysis**



IMPACT and CHALLENGES

- Data integration and network based approaches

In order to

- Identify potential biomarkers
- Better identify patients at risk if a cancer disease
- Improve precision medicine approaches

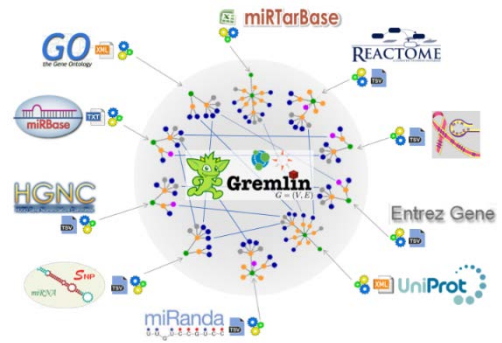
Data integration for enhanced analysis

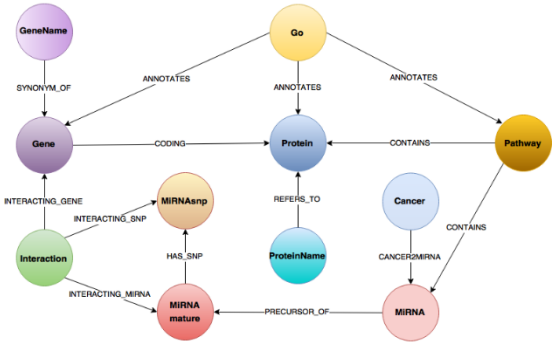
Impact

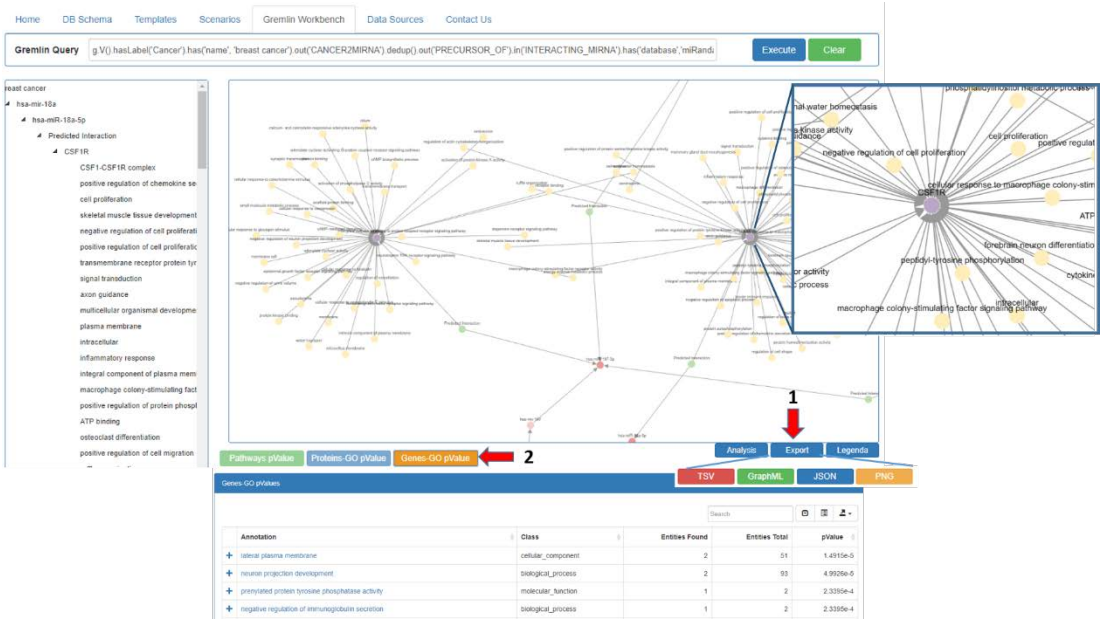
BioGraph: an online service and an integrated graph database that collects and links heterogeneous bioinformatics resources.

Challenges and Future Work

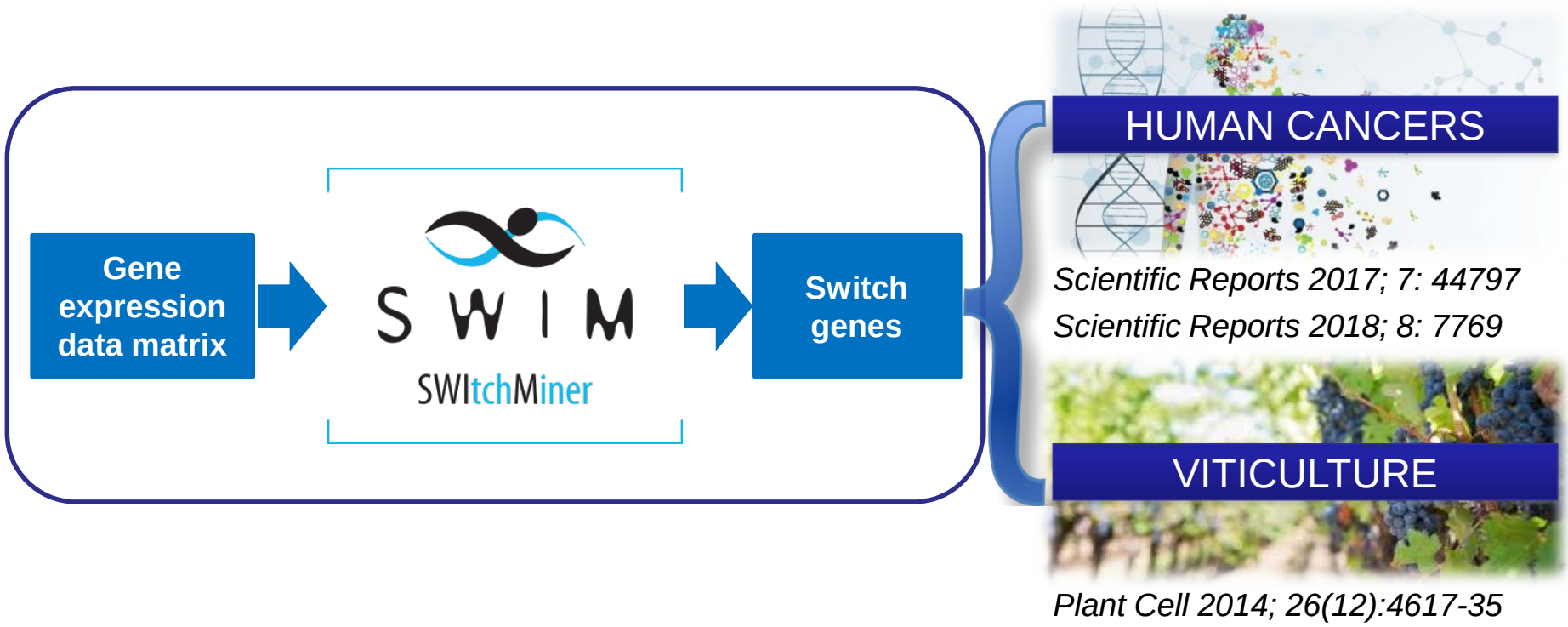
Allowing the integration of large databases in order to obtain a fast querying and clear overall vision of many connected biological phenomena







Key genes identification



SWIM is a software able to identify key genes in a network of interactions of various sorts by defining appropriate roles to genes according to their local/global positioning in the overall network. SWIM represents a solution of excellence that can improve the knowledge of the cellular events crucial for carcinogenesis and may unveil many potential prognostic and novel therapeutic targets. Using SWIM could provide important clues that will stimulate research activities into the causes of high-impact diseases, including but not limited to human cancer, thus supporting the planning of healthcare services such as clinical trials and disease prevention.

Whole cell models

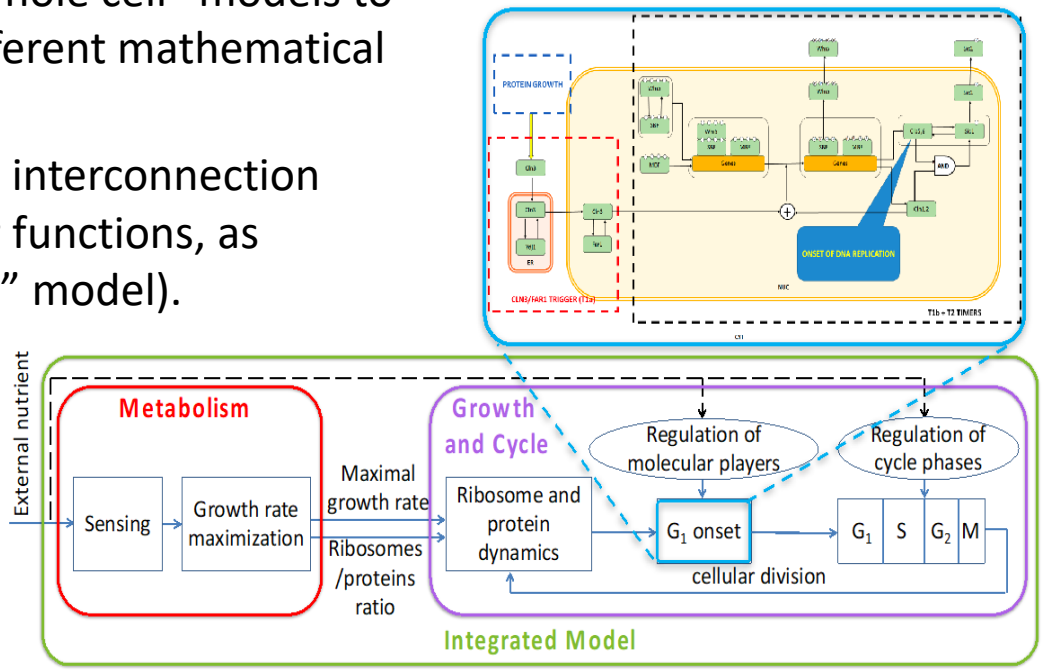
Activities:

❖Development of multi-scale, dynamical “whole cell” models to explain complex phenotypes, combining different mathematical approaches and large amount of data.

❖Development a hierarchical “coarse-grain” interconnection scheme that integrates the essential cellular functions, as metabolism, growth and cell cycle ("scaffold” model).

❖ Development of molecularly detailed models of the cell functions and integration within the "scaffold” model (probing of the plugged-in functions in the context of a cycling cell).

❖ Formulation of a detailed mathematical model of the Whi5 phosphorylation mechanisms describing the G1/S transition (Palumbo et al., Nat. Commun. 2016).



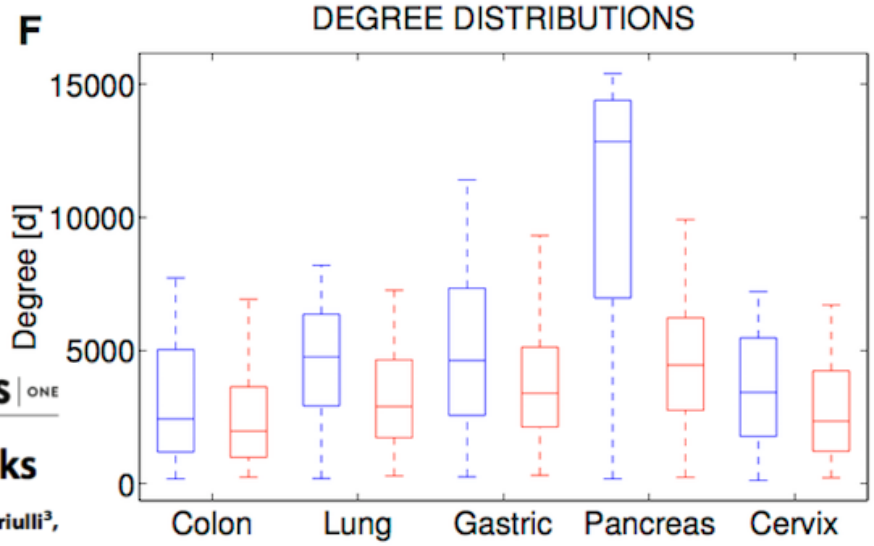
Nature Communications (2016), 7:11372

Network characterization

Development and application of methods in disease transcriptional networks alterations:

- Key drivers identification of the onset and progression of complex diseases,
- Normal cell biology and its pathologic dysregulation.

The loss of connectivity among differentially connected genes is a common topological trait of cancer networks, and unveils novel candidate cancer genes.



OPEN ACCESS Freely available online
PLOS ONE

Loss of Connectivity in Cancer Co-Expression Networks

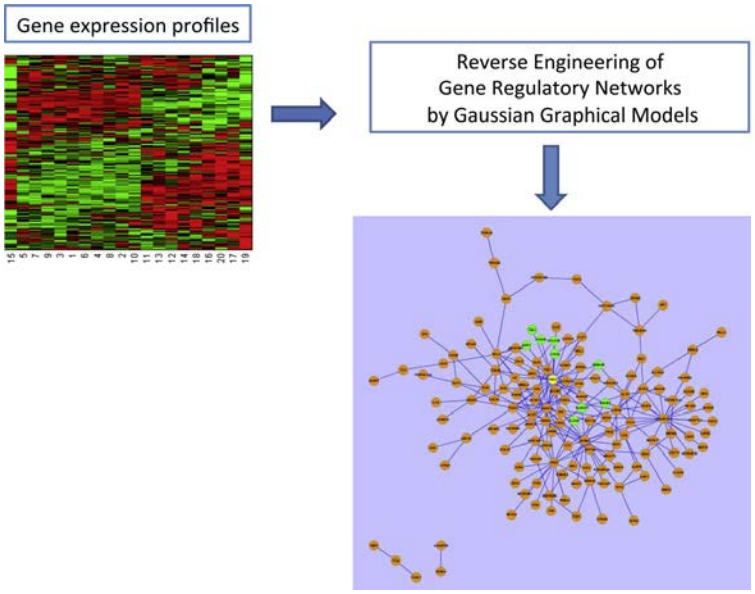
Roberto Anglani[1☯], Teresa M. Creanza[1,2☯], Vania C. Liuzzi[1], Ada Piepoli[3], Anna Panza[3], Angelo Andriulli[3], Nicola Ancona[1*]

Gene regulatory networks

Inference of gene networks from expression data:

- The theoretical structure and the constructive methodology for large scale undirected graphical models when p >> n,
- The study of performance on simulated data sets and application to two biological case studies.

Inverse covariance matrix reveals a cross talk between the isoprenoid biosynthesis pathways in Arabidopsis thaliana.



Journal of Biomedical Informatics 46 (2013) 894–904

Contents lists available at SciVerse ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

A comparative study of covariance selection models for the inference of gene regulatory networks ☆

Patrizia F. Stifanelli [a], Teresa M. Creanza [a], Roberto Anglani [a], Vania C. Liuzzi [a], Sayan Mukherjee [c], Francesco P. Schena [b], Nicola Ancona [a,*]

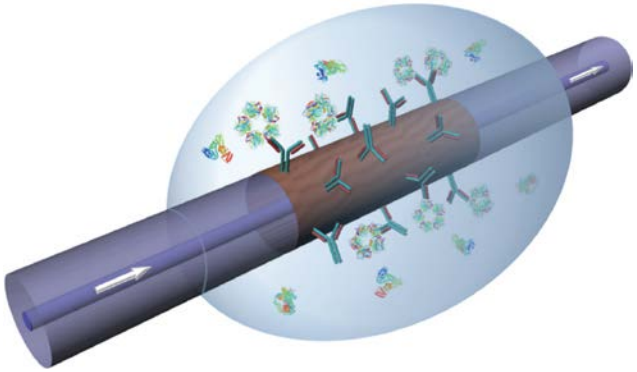
3. Biophotonics

- The development of devices capable to achieve single molecule detection, or with very low limit of detection:
  - Identification and monitoring of single molecules which can open new knowledge in the comprehension of the their physics and chemistry,
  - the measurement of analytes at very low concentrations can play a fundamental role in medicine in the detection of the onset of pathologies at a very early stage.
- The use of light- based devices can be used in a tailor-made surgical application:
  - Realization of precise and customized surgical cuts and suturing patterns;
  - Design of cardiovascular stents perfectly matching the patient's anatomy.
- All these approaches are at the basis of the precision and personalized medicine.

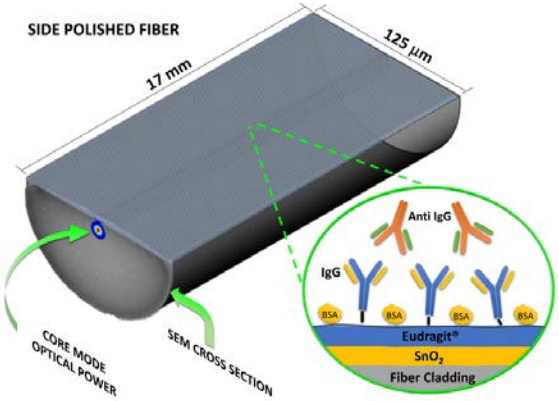
Biosensors

❖**Optical biosensors:**

Novel platforms based on optical resonators (e.g. microbubble resonators, long period gratings, lossy mode resonances) capable to achieve very low limit of detection



F.Chiavaioli, F.Baldini, S.Tombelli, C.Trono, A.Giannetti, "Biosensing with optical fiber gratings", Nanophotonics, 6 (4), 663- 679, 2017



F.Chiavaioli, P.Zubiate, I.Del Villar, C.R.Zamarreño, A.Giannetti, S.Tombelli, C.Trono, F.J.Arregui, I.R.Matias, F.Baldini, "Femto-molar Detection by Nanocoated Fiber Label-Free Biosensors", ACS Sensors, 3 (5), 936-943, 2018

IMPACT and CHALLENGES

Achieving very low limit of detection (theoretically down to single molecule detection) paves the way to the possibility of:

- **Detection, in medical field, of the onset of pathologies at a very early stage with a huge benefit for the patient**
- **Study of the behaviour of the single molecule**

Surgical & diagnostic techniques

- **Tailor-made surgeries: realization of precise and customized surgical approaches, perfectly matching the patient's anatomy and pathology requirements**
- **High precision microspectroscopy and imaging for early detection of deseases**

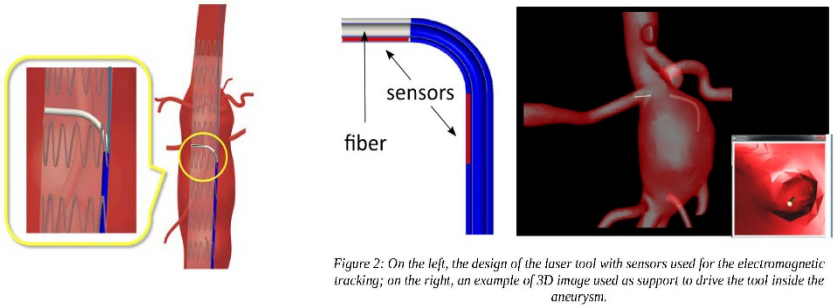


Figure 2: On the left, the design of the laser tool with sensors used for the electromagnetic tracking; on the right, an example of 3D image used as support to drive the tool inside the aneurysm.

Electromagnetic guided in-situ laser fenestration of endovascular endoprosthesis

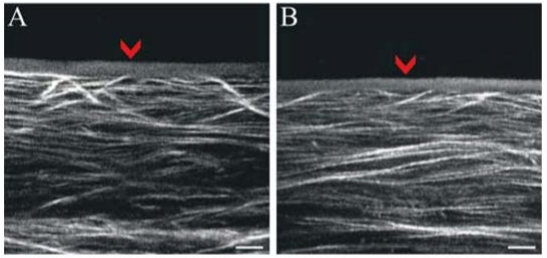
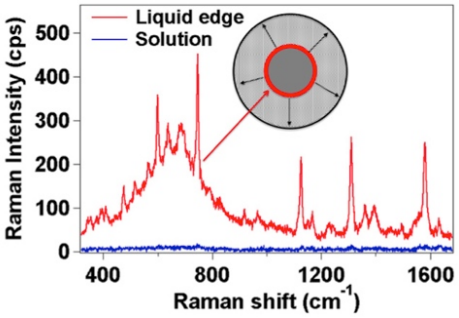


Figure 1 (**A**) SHG image acquired with sagittal optical sectioning of a healthy cornea and (**B**) of keratoconic cornea. The red arrows indicate Bowman's Membrane. Scale bars: 30 μm.

Mercatelli R, et al, Three-dimensional mapping of the orientation of collagen corneal lamellae in healthy and keratoconic human corneas using SHG microscopy. J Biophotonics, doi: 10.1002/jbio.201600122 (2016)



Banchelli M, et al. Triggering molecular assembly at the mesoscale for advanced Raman detection of proteins in liquid. *Sci Rep,* 8, 1033, 2018

Optical detection of disease biomarkers by controlled evaporative dynamics of biological fluids

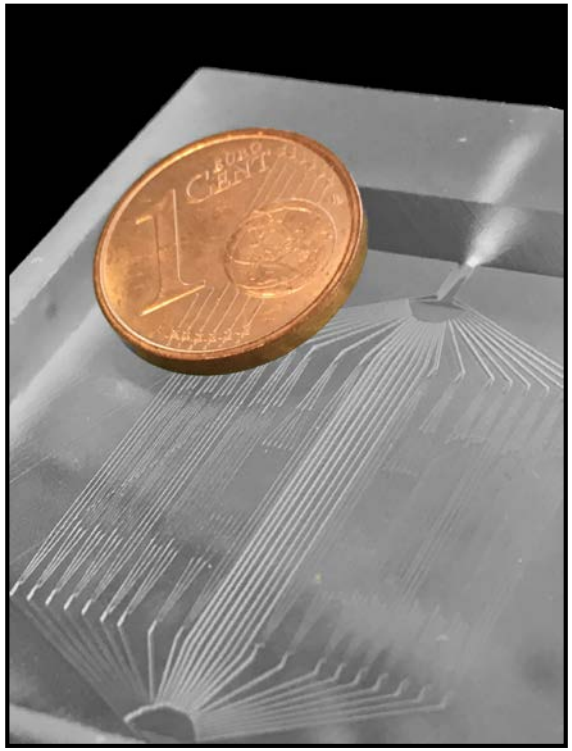
IMPACT and CHALLENGES

- **The use of minimally invasive surgical and diagnostic techniques enable the design and application of personalized and precision medicine.**

4. Bioprocess and Tissue engineering

Impact:

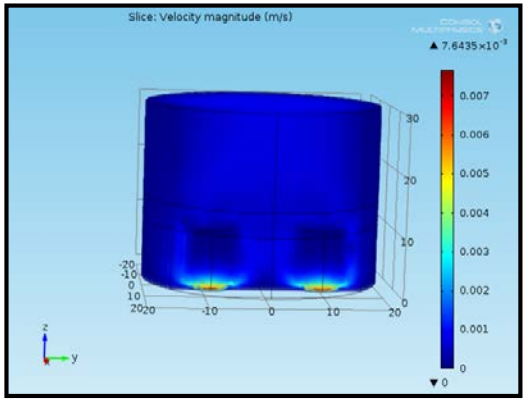
Tissue engineering approaches for more predictable R&D and/or in vitro screening platforms, by providing more physiologically relevant conditions than traditional cell culture and/or animal models, and by enabling the availability and effectiveness of personalized and predictive screening tools for therapies that will prevent or limit spread of diseases

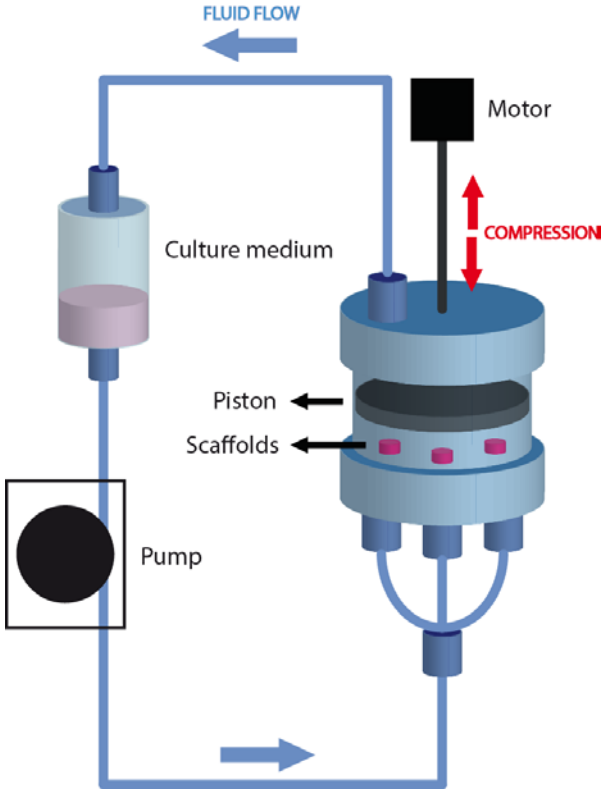


Micro-fluidic device for in vitro testing

Mechanical bioreactor for articular applications

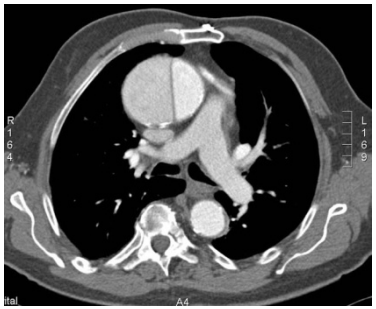
CFD simulation of stimuli applied in vitro to prothesis

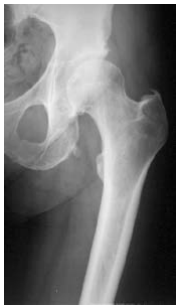


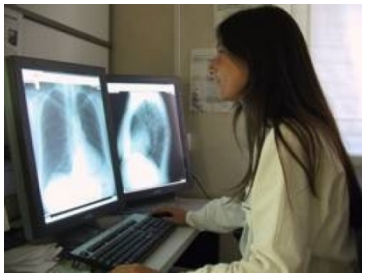


Objectives

Doctors' demands for personalized protheses or tool for pre-operative patient specific planning



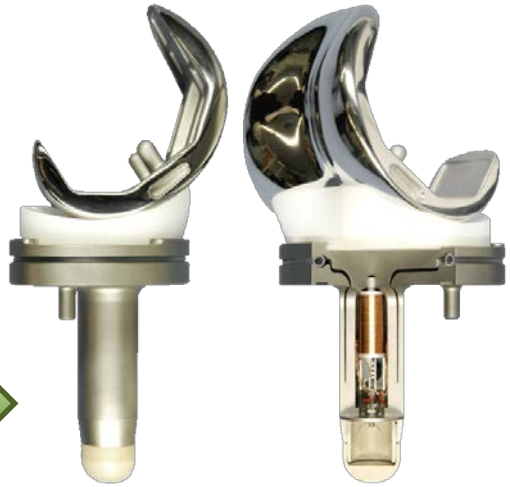




Fab @ Hospital

Production service within hospitals

Personalized protheses

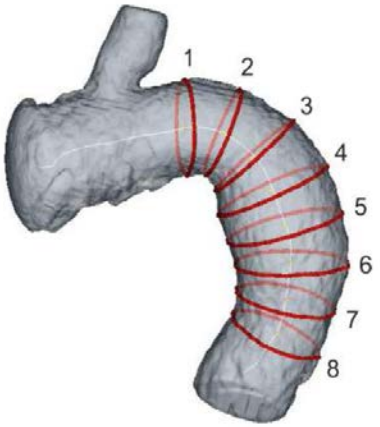


Pre-operative patient specific tool

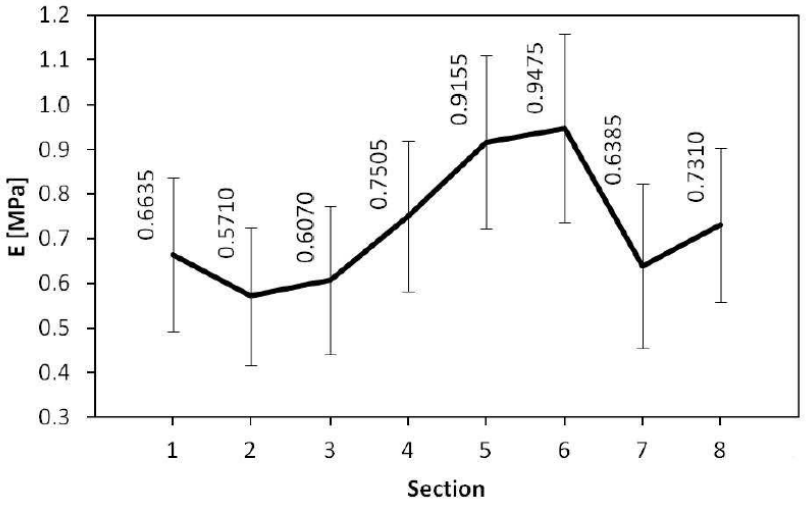
Anatomo-functional properties

Aim: to estimate the stiffness of a given aortic region and its spatial variations based on non-invasively acquired information.

- Aortic radius at 8 section measured through Computed Tomography Angiography (CTA) images along with cardiac cycle.
- Corresponding aortic pressures at each section and each time instant generated by means of an appropriate lumped parameter model to keep the approach non-invasive.
- Pressure and radius measurement at each section linked through a differential constitutive model of the vessel wall in which stiffness appears.
- Estimation of the aortic stiffness by means of a Bayesian approach applied to the differential constitutive equation.



Example of reconstructed radii for a given CTA image

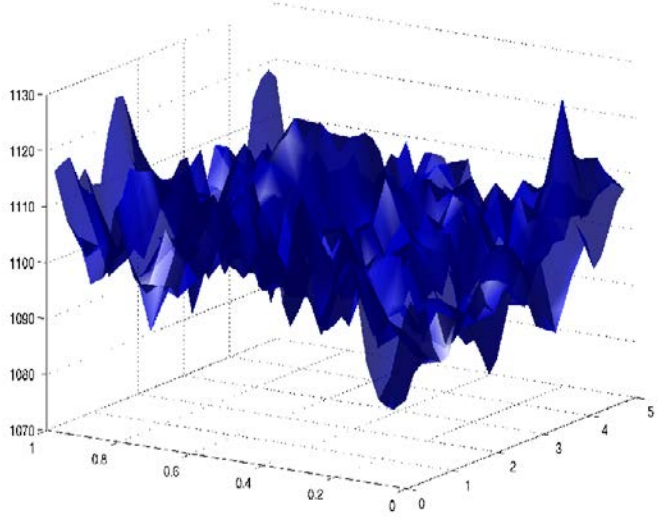


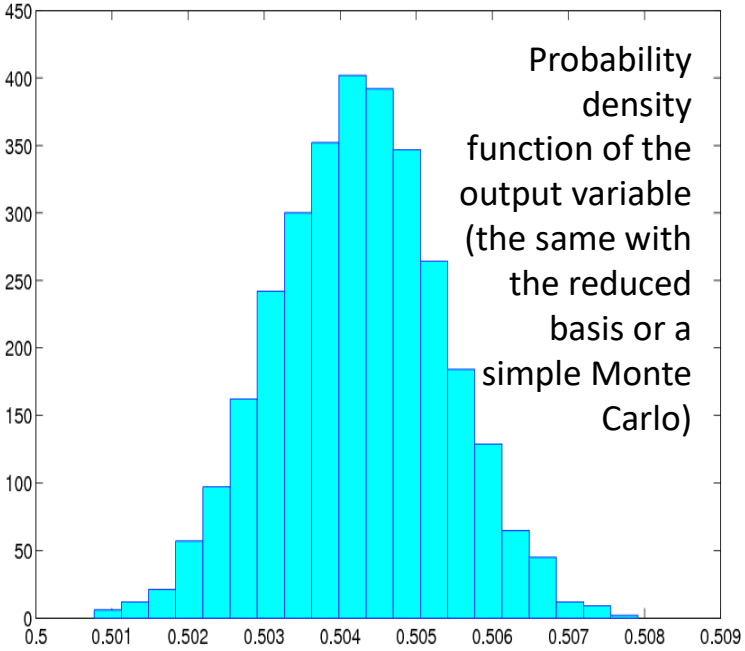
Results from a real clinical case

Stochastic FEM

Aim: estimate probability distributions of some outputs of interest given stochastic models for the PDE coefficients (input parameters) using *non-intrusive* techniques.

Test case: an elastic problem in the presence of an uncertain Young modulus over a 2D domain, e.g., it is represented as a Gaussian random field [figure on the





Challenges: Reduce the computational complexity of Monte Carlo simulation to investigate how this uncertain input influences an output variable of interest(e.g., expected or maximum displacement),

5. Mathematical software

- New innovative models are required to guarantee the growth of social and health conditions of citizens.
  - A fundamental enabler of medical services and healthcare, technology and industry.
- New statistical methods and approaches to elaborate accurate anatomical models from bio-images for personalized products.
- Improvements to existing technologies for personalized products.
  - Personalized medical products;
  - Innovative process combinations and manufacturing approaches;
  - New technological approaches for real products and medical scenarios.

6. Bio-sensors and Bio-inspired systems

- Combination of memory and processor in a single system for synapse-like electronic elements.
  - ❖ Developed devices and networks to mimic functions of the nervous system and brain,
  - ❖ Devices interfaced with nervous system for recovering damaged parts and to control prosthesis.
- Pharmaceutical preparation containers for the delivery to the diseased areas of the body or to the zones of risk.
  - ❖ Nanoengineered polymeric capsules filled with active drugs,
  - ❖ Incorporation of magnetic nanoparticles and receptors for targeted delivery.
- Organic biosensors and devices miniaturized and transferred onto biocompatible substrates for in-vivo biomedical applications.

Conclusions

- Biotechnology provides a constantly evolving scenario, where research has achieved significant scientific and technological progress using and integrating multidisciplinary skills.
- These advances have economic repercussions in health, agriculture and industry, and in support to the social challenges of H2020 program.
- These activities provide results with close ties with other design areas such as Healthcare and Well-being, Nanotechnology and Applied Mathematics and Low Carbon Technologies.
- The results attest a gradient ascending direction towards major achievements in terms of scientific contribution, ability to collaborate in national and international research projects, and young scientists training.