



King Abdullah University of
Science and Technology



Systems View of Biological Organisms: A Computational Approach

Presented by: Dr. Anna Tramontano

Medical Faculty

Sapienza University, Italy

Computational genomics and KAUST



Aims:

Understanding the (human) genome and deciphering how DNA controls the biology of a species at the molecular level

New or improved therapeutic targets

Better therapies, vaccines and diagnostic tools for genetic and infectious diseases, especially those of interest for the region (Brucellosis, Q fever, Malaria, Lysosomal storage disease, aminoacidemias and unique genetic syndromes such as Al-Aqeel Sewairi and Sanjad-Sakati, etc.)

In silico testing of compounds

La Sapienza and KAUST



Aims:

Development of common interdisciplinary curricula

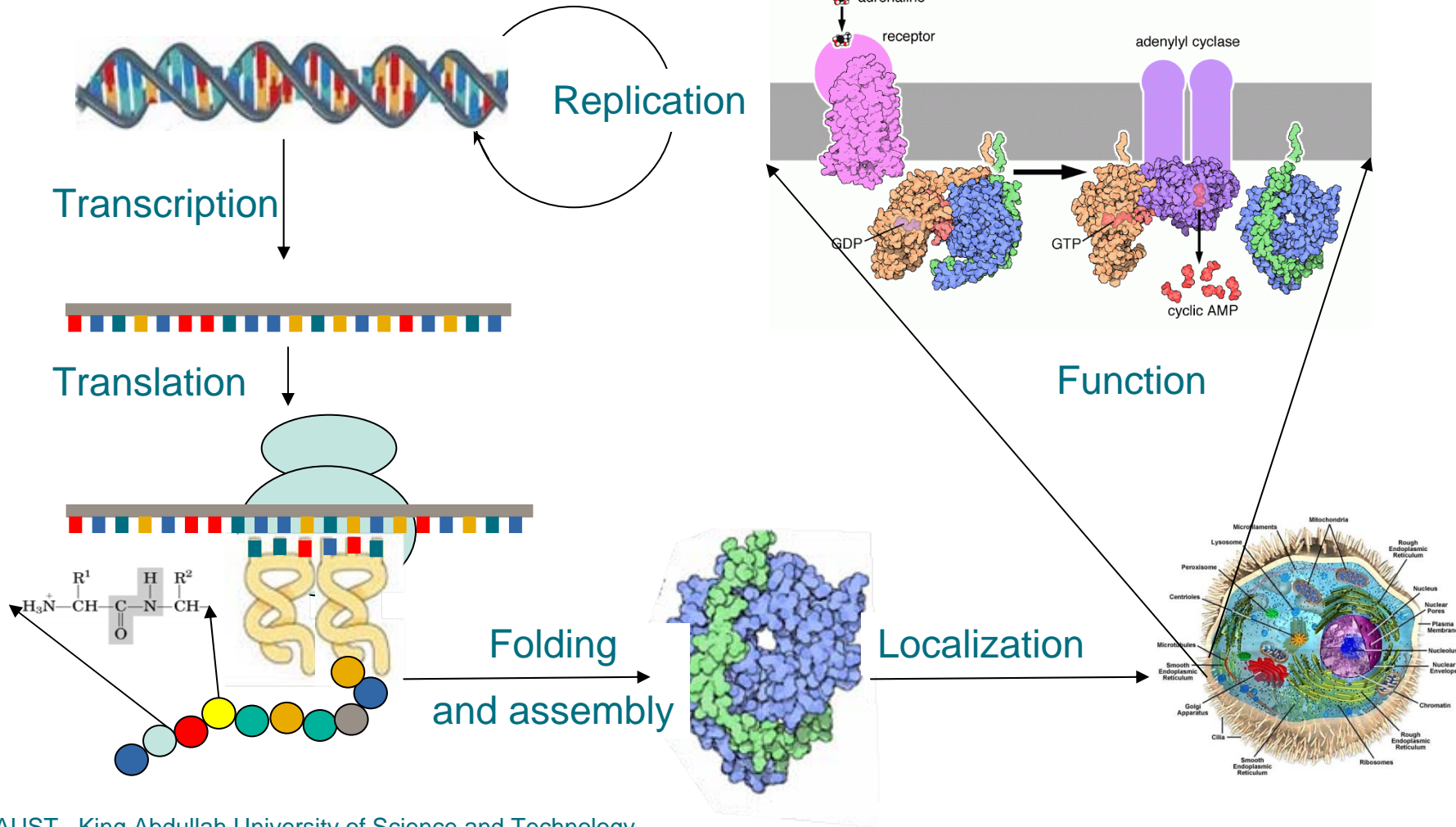
Exchange visits of undergraduate and post-docs, joint supervision of students

Organization of joint specialized workshops at KAUST and at “La Sapienza” on emerging issues and methodologies in computational genomics

Specialized hands-on courses at KAUST on new developments in computational genomics and proteomics

Liaisons with worldwide initiatives in computational biology

The system

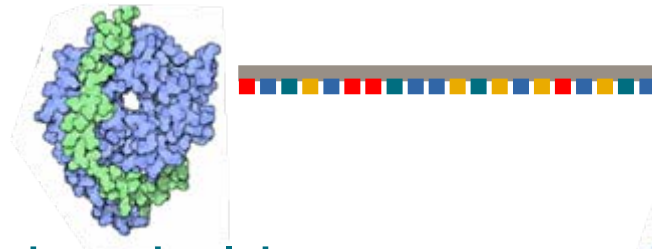


The system

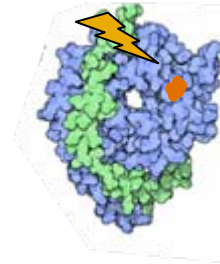


As in every engineering project we need to find out:

The parts list



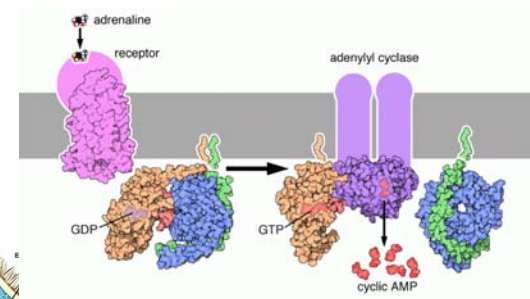
Their tolerance threshold



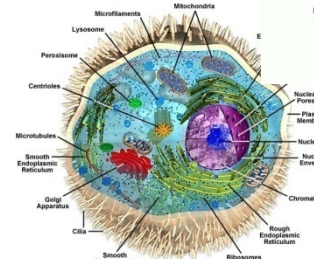
The assembly instructions



The dynamic behavior



The simulation of its functioning



The parts list: genes and proteins



CATACACATGATGACGATAATATAGATAGATAGATA...GATGACGATAATATAGATAGATAGATA ACTAATTGACGATGAC

The human genome contains ~3,000,000,000 nucleotides: about 3% codes for proteins:



Coding regions are not contiguous:

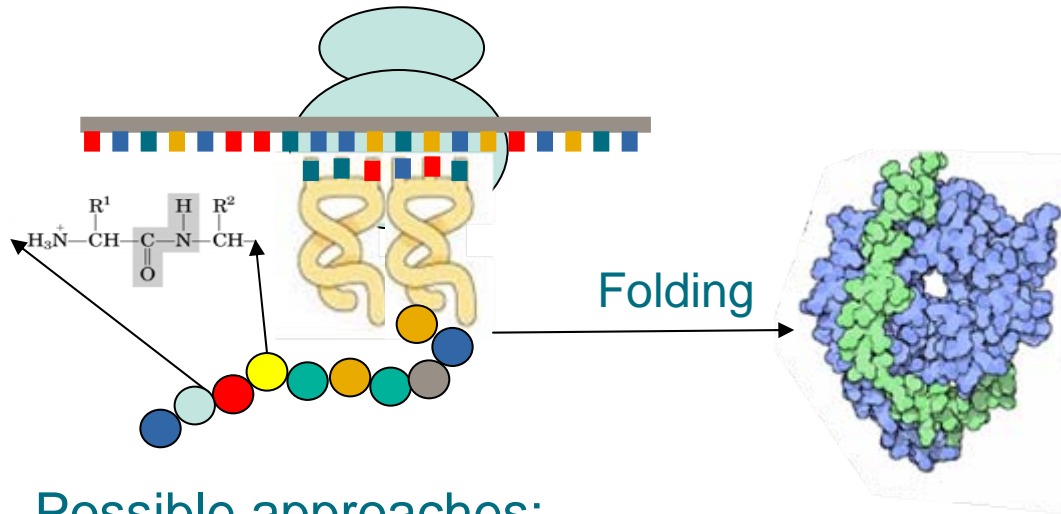


ovalbumin

Very weak signal

Present methods for gene finding achieve about 70% accuracy

The parts list: genes and proteins



Probably the hardest problem in biology

Possible approaches:

Experimental (x-ray diffraction or NMR) - Time and labor consuming

Physics based methods - Not suitable: proteins are only marginally stable and their stability is due to an enormous number of very weak interactions

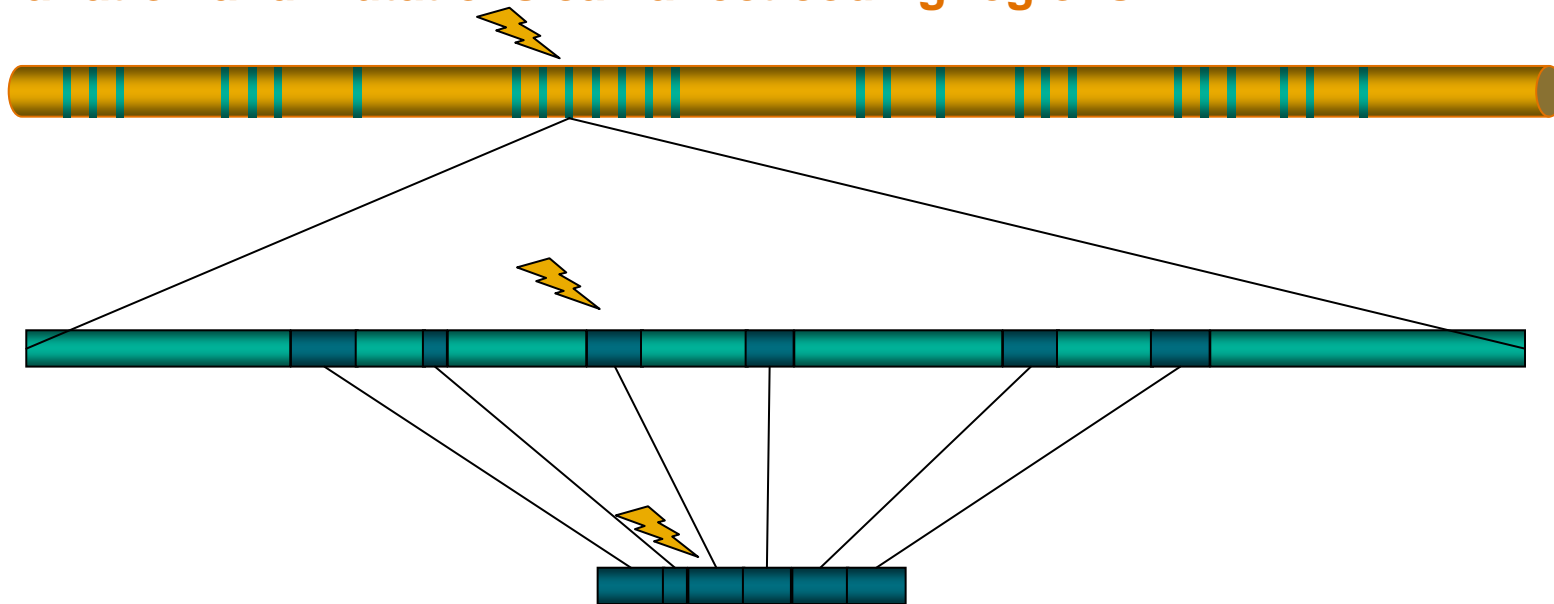
Hybrid physics and knowledge based methods work in some but not all cases

The tolerance threshold: effect of mutations

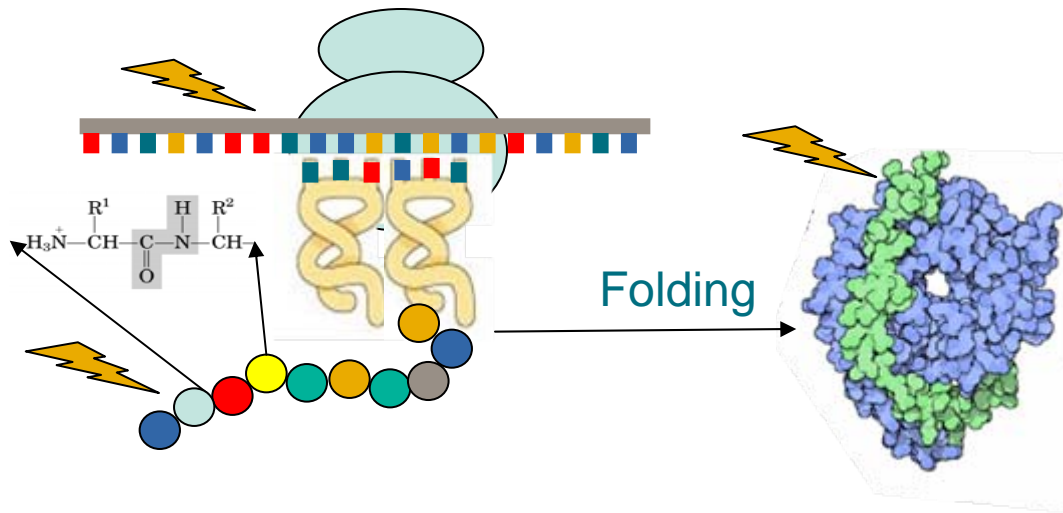


ATATAGATAGATAGATA... Human genomes differ by about 0.1% ...GATGACGATA

Human genomes differ by about 0.1% from each other and can have mutations
Variation and mutations can affect coding regions:



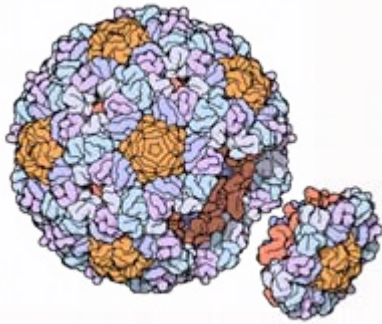
The tolerance threshold: effect of mutations



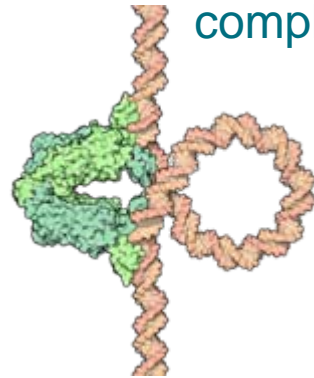
A mutation can affect the structure and function of the encoded protein. A variation might be neutral, pathological or affect the susceptibility to diseases.

Predicting the functional effect of genetic variations is essential for correct diagnosis of genetic diseases and/or predispositions
No reliable method available as yet.

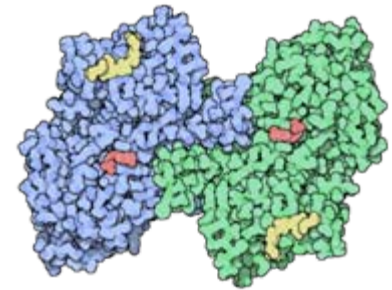
The assembly instructions



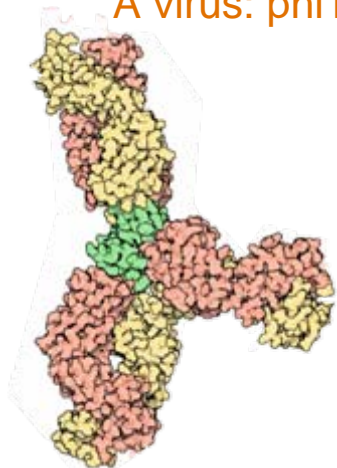
A virus: phi174



A regulator: the lac repressor



An enzyme: glycogen phosphorylase



Antibodies recognizing their antigen

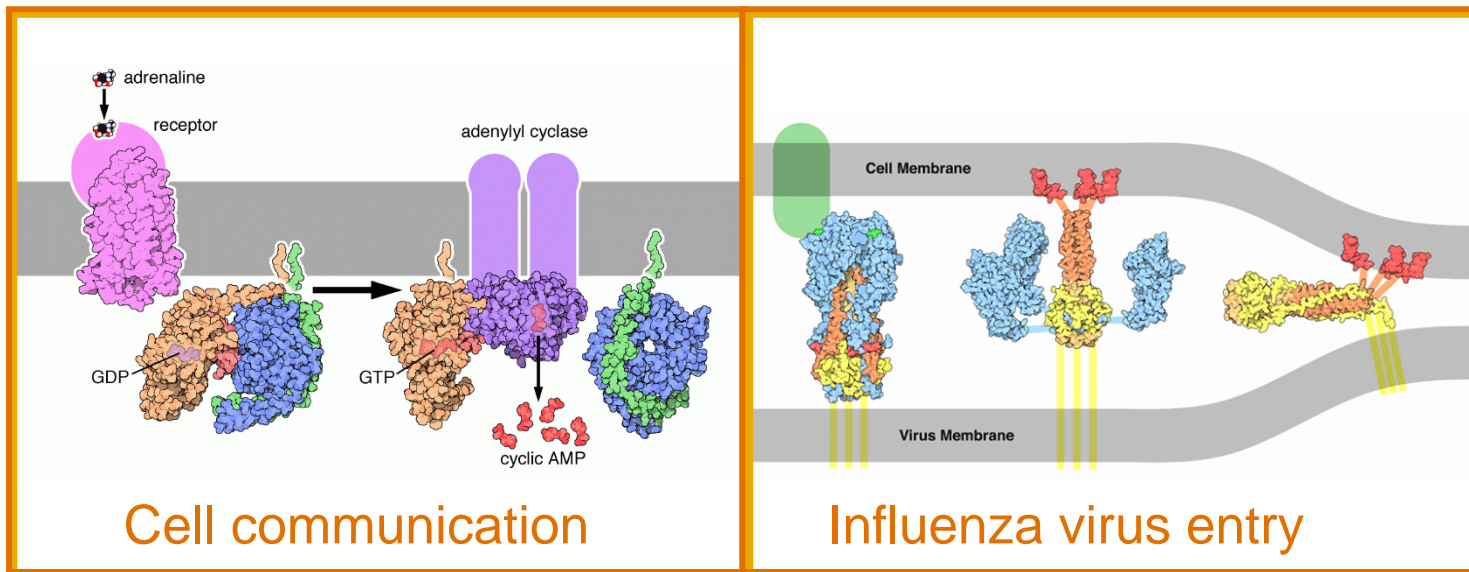
Protein recognize each other, nucleic acids and small organic molecules via shape complementarities and chemical interactions.

Present methods for predicting the interaction mode work sporadically and only when the experimental structures are available. Understanding protein-protein, protein-DNA and protein-small molecule interactions is a prerequisite for drug design.

The dynamic behavior



Protein folding and assembly, but also flexibility, are the basis of most biological mechanisms. For example:

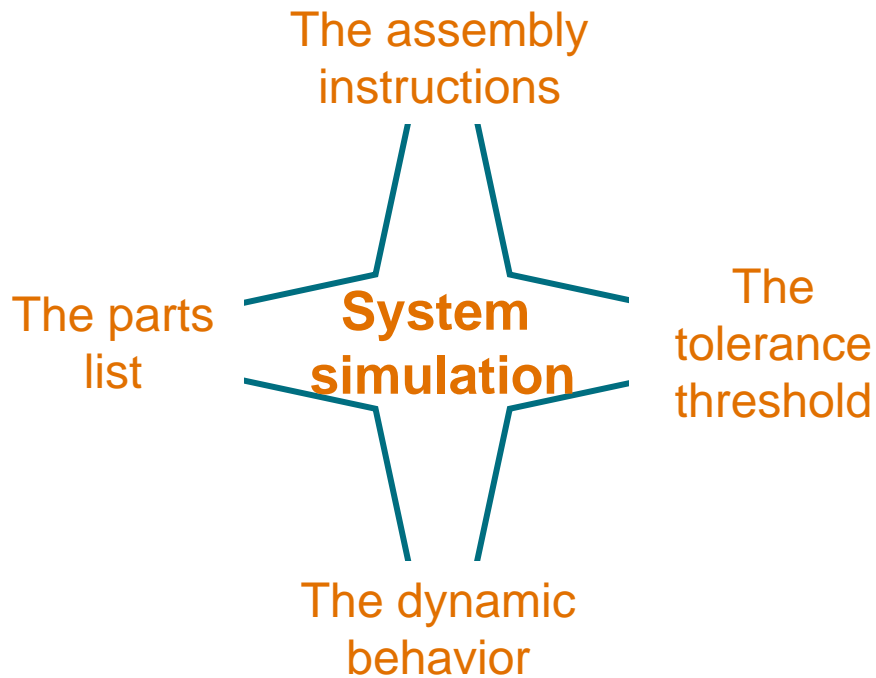


Predicting the effect of flexibility and interactions on protein structures is important for understanding how they work and how we can interfere with their mechanism. This requires extensive computer power.

A system view



If we understand a system, we should be able to simulate its response to stimuli.



Can we simulate the behavior of cell compartments, cells and even organs *in silico*?

We might be able to understand and predict the individual responses to therapy and immunization as well as adverse side effects from prescribed medicines

Our strategy



An interdisciplinary approach combining physics, biology, computer science, engineering and chemistry:

- Analysis and integration of large sets of experimental and computationally produced data
- Prediction of the three-dimensional structure of all putative gene products
- Evaluation of each predicted structure to identify real genes
- Prediction of the mode of interactions between gene products
- Development of structure based methods for assessing the effect of mutations
- Development of mathematical models and stochastic simulations of cell systems

Contribution to Saudi Arabia and KAUST



Human health and Biotechnology:

Gene finding: new or improved therapeutic targets, discovery of novel biochemical reactions

Structure prediction: drug and vaccine design, design of new catalysts

Mutation analysis: early diagnosis of genetic diseases, optimization of biocatalysts

Systems simulation: *in silico* testing of compounds, biotechnology safety

Education:

Student exchange, development of joint interdisciplinary projects and curricula, organization of outreach activities, scientific workshops and exposure to a truly interdisciplinary environment

The research environment



A truly interdisciplinary group composed of:

Physicists
Engineers
Biologists
Biotechnologists
Computer scientists
Chemists

Ongoing collaborations within Sapienza University:

Biochemistry and Structural biology
Immunology
Molecular and Cell biology
Genetics
Oncology
Statistics
Computer science

Ongoing international collaborations:

EBI, Cambridge, UK
CBS, Denmark, DK
CNIO, Madrid, Spain
UC Davis, Davis, USA
U. Maryland, Rockville, USA
U. Montreal, Montreal, CA
U. Oxford, Oxford, UK
U. Penn, Philadelphia, USA

Funding:

KAUST
European Union
Human Frontier Program
Pasteur Institute
Italian Ministry of Research
Italian Cancer Research Foundation