Computational and Translational Methods for Cancer Genomics

A workshop of the UNIBO-UCSD Cooperation Project

BiGeA and FaBiT Departments - UNIBO
Department of Medicine - UCSD

Meeting Organizers
Yana Bromberg, Rutgers University (USA)
Emidio Capriotti, University of Bologna (Italy)
Hannah Carter, University of California, San Diego (USA)

Special thanks for the organization and support
Rita Casadio
Pier Luigi Martelli
Castrense Savojardo
Paola Turina

Registration Desk
Davide Baldazzi
Emina Merdan

Invited collaborators
Piero Fariselli, University of Padova (Italy)
Jairo Rocha, University of Balearic Islands (Spain)

Speakers
Fatima Al-Shahrour, CNIO, Madrid (Spain)
Hannah Carter, University of California, San Diego (USA)
Pietro Lio', University of Cambridge (UK)
Roberta Maestro, CRO, Aviano (Italy)
Christine Nardini, Karolinska Institute (Sweden)
Fabio Vandin, University of Padova (Italy)

Funding
International Cooperation Project, University of Bologna

Computational and Translational Methods for Cancer Genomics

Tuesday May 29, 2018

Sala Ulisse, Accademia delle Scienze – Via Zamboni 31, Bologna

Session I: Translational approaches
Chair: Yana Bromberg, Rutgers University

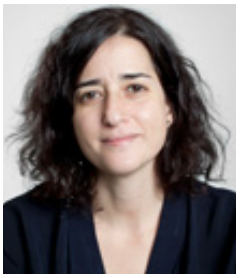
10:00 – 10:10	Opening
10:10 – 11:00	***Christine Nardini***, Karolinska Institute (Sweden) Epigenomics and cancer – methylation as a marker in oncology.
11:00 – 11:20	Coffee Break
11:20 – 12:10	***Fatima Al-Shahrour***, CNIO, Madrid (Spain). Identifying druggable cancer dependencies for personalized medicine.
12:10 – 13:00	***Roberta Maestro***, CRO, Aviano Italy Genomics of soft tissue sarcomas.

Session II: Computational methods and tools
Chair: Emidio Capriotti, University of Bologna

14:30 – 15:20	***Fabio Vandin***, University of Padova (Italy) Computational methods to discover significant mutations in cancer genomes.
15:20 – 16:10	***Pietro Lio'***, University of Cambridge (UK) Methodologies for data integration in cancer.
16:10 – 16:30	Coffee Break
16:30 – 17:20	***Hannah Carter***, UC San Diego (USA) Are tumors predictable? Inherited immune variation constrains tumor evolution.
17:20 – 17:30	Closing Remarks

Fatima Al-Shahrour

Bioinformatics Unit
Spanish National Cancer Research Centre (Spain)
email: falshahrour@cnio.es
www: https://goo.gl/sY4uHA



Identifying druggable cancer dependencies for personalized medicine

The paradigm of personalized medicine is the identification of the appropriate drug for the right patient, using molecular profiles. In Oncology, it is well established that the anticancer drugs are effective in only a small subset of patients. Moreover, many of the new targeted therapies inhibit specific proteins, and they are only effective in tumors that are genetically altered. Consequently, the success of personalized treatment depends on each individual molecular profile, which a priori can be considered as very heterogeneous.

Here, we present new computational approaches based on the analysis and integration of genomic data (mutations, copy number variations or gene expression levels), with functional data (protein essentiality) and pharmacological data. These methods aim to identify those vulnerable molecular alterations that drive tumor progression and could be druggable based on the patient's molecular profile, and propose an individualized therapeutic strategy to guide clinical decision making for cancer patients.

Hannah Carter

Department of Medicine
University of California San Diego (USA)
email: hkcarter@ucsd.edu
www: https://profiles.ucsd.edu/hannah.carter

Are tumors predictable?
Inherited immune variation constrains tumor evolution

Significant insights into tumorigenesis have been gained by characterizing the extensive somatic alterations that arise during cancer and uncovering rare inherited mutations that lead to early onset cancer syndromes. However, little is understood about the role of genetic background in 'sporadic' adulthood cancers. We have demonstrated that somatic evolution of a tumor is influenced by inherited polymorphisms. The genomic region encoding the Major Histocompatibility Complex Class (MHC) is one of the most variable regions in the human population. MHC molecules expose peptide fragments on the cell surface, allowing T-Cell elimination of cells contaminated by foreign peptide. Although this system has evolved as a defense against microbial and viral agents, MHC can also trigger elimination of cells harboring mutant peptides (neoantigens) in cancer. Each individual carries multiple MHC alleles that define the set of peptides that can be effectively presented for immune surveillance. We hypothesized that individual variation in MHC could create personal gaps in immune surveillance, generating individual-specific susceptibility for cells to acquire specific oncogenic mutations. To test this hypothesis, we developed residue-centric patient presentation scores for MHC class I and II molecules and applied them to 1,018 recurrent oncogenic mutations in 9,176 cancer patients. This analysis uncovered a clear signature of immune selection acting on tumors; cancer-causing mutations were more likely to be observed when a patient's genotype-based scores suggested poor MHC-based presentation. Mutations that were poorly presented by most patients were more likely to reach high frequency among tumors. Individual coverage of driver mutations by MHC-based presentation was also found to be a determinant of age at diagnosis. Thus the landscape of oncogenic mutations observed in clinically diagnosed tumors is shaped by MHC genotype-restricted immunoediting during tumor formation, and individual MHC genotype provides information about the mutations likely to emerge in tumors that develop later in life.

Pietro Lio'

Department of Computer Science
University of Cambridge (UK)
email: pietro.lio@cl.cam.ac.uk
web: https://www.cl.cam.ac.uk/~pl219

Methodologies for data integration in cancer

Breast cancer is one of the most common invasive tumors causing high mortality among women. It is characterized by high heterogeneity regarding its biological and clinical characteristics. Several high-throughput assays have been used to collect genome-wide information for many patients in large collaborative studies. This knowledge has improved our understanding of its biology and led to new methods of diagnosing and treating the disease. In particular, precision medicine has become a valid approach to obtain better insights into breast cancer biological mechanisms. A crucial component of precision medicine lies in identifying and using novel biomarkers that can be predictive for breast cancer patient prognosis on the basis of the molecular signature of the tumor sample. However, the high dimension and low sample size of data greatly increase the difficulty of cancer survival analysis demanding for the development of ad-hoc statistical methods In this talk I will describe approaches based on mathematical models, statistics, deep learning and logic. I will discuss pros and cons of these methods to help clinicians to provide more precise prognoses and to facilitate the subsequent clinical management of patients at risk of disease

Roberta Maestro

Oncogenetics and Functional Oncogenomics Unit
National Cancer Institute (Aviano, Italy)
email: rmaestro@cro.it
www: https://goo.gl/B1x5UZ

Genomics of soft tissue sarcomas

The Unit of Oncogenetics and Functional Oncogenomics at the CRO Aviano National Cancer Institute focuses its research activity on the identification and characterization of the molecular alterations implicated in cancer development and progression, with particular attention to soft tissue sarcomas. Sarcomas are a rare and highly heterogeneous group of mesenchymal malignancies for which current therapeutic strategies remain unsatisfactory.

The talk will describe how, through the use of different and complementary genomic and genetic approaches, the Unit addresses the biology of these tumors with the final goal to translate the new findings into tools for better management of sarcoma patients.

Christine Nardini

Division of Clinical Chemistry
Karolinska Institute (Sweden)
email: christine.nardini@ki.se
www: https://ki.se/en/people/chrnar

Epigenomics and cancer
methylation as a marker in oncology

Among the growing number of omics available to the investigation of molecular mechanisms, epigenomic variations are gaining importance. Among the epigenomic alterations that can be screened with high-throughput technologies, methylation offers a stable and promising source of information. In humans, DNA methylation consists of the covalent addition of a methyl group at the 5-carbon of the cytosine ring and occurs mainly in CpG dinucleotides. The distribution of CpG dinucleotides and their methylation status varies widely across the genome: in the bulk of the genome CpG dinucleotides are underrepresented and tend to be pervasively methylated, while regions of high CpG density, termed CpG islands, are often found at gene promoters' sites in a non-methylated status. The study of the distribution of changes in methylation have been found to be extremely relevant in the description of biological ageing and associated diseases, including cancer. In the following I will present two relevant exemplars of this application.

Fabio Vandin

Deprtment of Computer Science
University of Padova (Italy)
email: fabio.vandin@unipd.it
www: http://www.dei.unipd.it/~vandinfa/

Computational methods to discover significant mutations in cancer genomes

Cancer is a disease that is mostly driven by somatic mutations accumulating in the genome during an individual's lifetime. Recent advances in DNA sequencing technologies have enabled genome-wide measurements of these mutations in large cohorts of cancer patients. A major challenge in analyzing these data is to distinguish functional "driver" mutations responsible for cancer progression from "passenger", random mutations not related to the disease. Recent cancer sequencing studies have shown that somatic mutations are distributed over a large number of genes. This mutational heterogeneity is due in part to the fact that somatic mutations target pathways, or groups of genes, and that a mutation in any of dozens possible genes might be sufficient to perturb a pathway. While some of the cancer driver pathways are well characterized, many others are only approximately known.

I will describe algorithms for discovering cancer driver pathways using DNA mutation data from large cohorts of cancer samples. The first algorithm uses a heat diffusion process on graphs and a novel statistical test to identify subnetworks of a large gene interaction network that are mutated in a significant number of cancer samples. The second algorithm identifies subnetworks that have mutations associated with clinical variables, in particular survival time. I will illustrate applications of these algorithms to data from The Cancer Genome Atlas, a project that has characterized the genomes of thousands of samples from dozens of cancer types