**Statistical bioinformatics and genomics approaches to cancer classification, diagnosis, prognosis, prediction and biomarker discovery**

Vladimir A. Kuznetsov

Bioinformatics Institute/A\*STAR, Singapore

vladimirk@bii.a-star.edu.sg

How to identify complex yet meaningful biological phenomena, molecular pathways/patterns, as well as make clinically useful and statistical-based predictions from high-dimensional, noisy, poorly annotated, incomplete and varied massive omics data sets?

Herein, I will present results from our recent studies in the fields of computational genomics, statistical bioinformatics, disease classification and clinical biomarkers discovery for ovarian and breast cancers. I will address the above-mentioned problems and proposed solutions from the perspectives and emphasis of industry alignment and clinical implementation. Also, I will demonstrate how the fields of cancer genome data analysis, disease classification and predictions of perspective clinical biomarkers are shifting from data-driven molecular feature selection, classification and pattern recognition to big data integrative analysis, disease/survival model–based approaches and the personalized risk assessment of disease recurrence.

Progress in these areas allows us (i) to acquire a better understanding of the molecular and cellular basis of various cancers, (ii) to provide quantitative and reproducible characterization of the enormous patho-biological heterogeneity within and between cancers, (iii) to identify clinically relevant and economically reasonable 'combined biomarkers' for early detection of the cancers and their sub-typing, (iv) to stratify patients according to individual risk of the disease recurrence (v) to initiate the development of new assays, and (vi) to specify parameters and unbiased systems for a more efficient and effective decision-making process in the clinics.