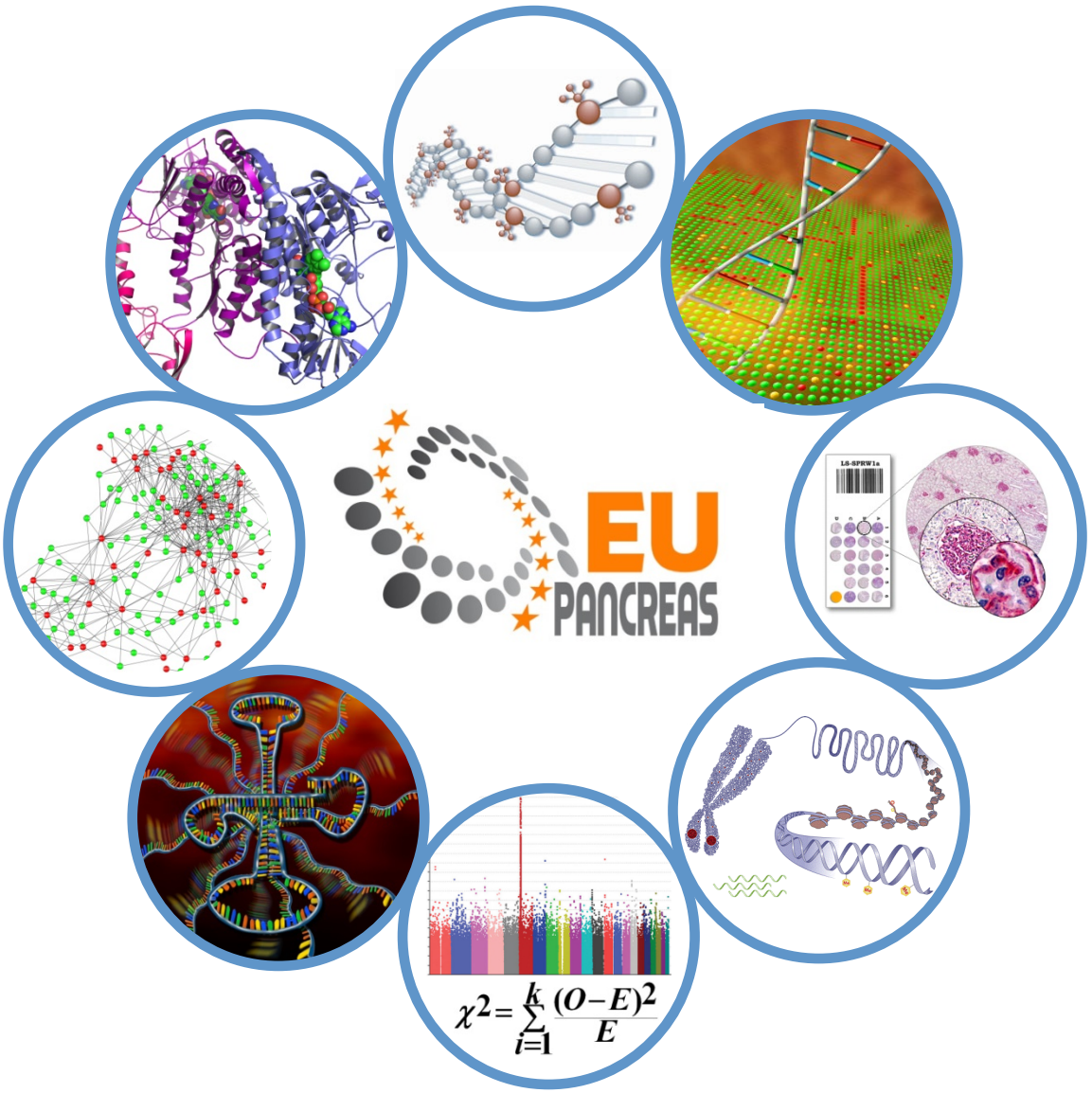
COST Action BM1203 – WG2 Training School

The World of Interactions Around Us



Program and Abstract Book

Antwerp University – Belgium

27 - 29 April 2016

Sponsors

CSCDA

EU PANCREAS

COST BM1204 - An integrated European platform for pancreas cancer research: from basic science to clinical and public health interventions for a rare disease

SBS BVS

Belgian Statistical Society

Universiteit Antwerpen

GIGA

BIO

Welcome

CSCDA

Dear participant,

It is our great pleasure to welcome you to this training school, as an activity of the COST Action BM1204 on pancreas cancer. This Action aims to create a unique European platform to facilitate the collaboration of a broad range of European and international PDAC multidisciplinary research groups while integrating knowledge and experience in a multidisciplinary way "from cell to society", promoting the application of uniform study tools and protocols, fostering their optimal use by early-stage researchers, enhancing the mobility and training of researchers and permeating the society with the results produced by the Action.

In the Action, Working Group 2 deals with "multi-source data integration". When developing methodologies for omics integration, improved power and insights can be obtained by better acknowledging the complex relational structures between omics data types but also within each omics dataset. As INTERACTIONS refer to a particular type of data structure it is essential to understand what they mean in different contexts (e.g., biological and statistical) and to be aware of the methods that can best detect them. This is what the current workshop is all about…

We are also grateful to the support of the Belgian Statistical Society. The Society's main mission is to contribute to scientific progress in statistics by promoting co-operation between Belgian statisticians and to help the general public to get better understanding of the place of statistics in the modern world. Determining the place of statistics in modern personalized/omics medicine requires an open and flexible mind, one we will adopt throughout the entire training school.

We wish you a fruitful meeting that will boost dialogue and create novel opportunities and collaborations to tackle the challenges of complex disease *INTERACTIONS*.

Sincerely,

Kristel Van Steen and Kristel Sleegers
(Presidents of the Organizing Committee)

Committees

CSCDA

Presidents

Kristel Sleegers	Professor	University of Antwerp, Belgium
Kristel Van Steen	Professor	University of Liège, Belgium

Organizing committee

Nuria Malats	Professor	Spanish National Cancer Research Centre, Spain
Joost Weyler	Professor	University of Antwerp, Belgium
Sanne Berwaerts	Administrator	University of Antwerp, Belgium
María Esther Molina	PhD	Spanish National Cancer Research Centre, Spain

Local assistants

Kridsadakorn Chaichoompu	PhD Student	University of Liège, Belgium
Ahmed Debit	PhD Student	University of Liège, Belgium

Program

CSCDA

Day 1 (27 April 2016) - Different faces of interactions – a primer

Time	Speakers	Room	Topic description
8:00-8:30	Registration		
8:30-10:30	Stijn Vansteelandt	C101	An introduction to statistical, causal and epistatic interaction
10:30-11:00	Coffee Break		
11:00-13:00	Maggie Wang and Kristel Van Steen	C101	Practical considerations in genome-wide interaction studies
	Maggie Wang		W-test
	Kristel Van Steen		MB MDR
13:00-14:00	Lunch		
14:00-16:00	Karsten Borgwardt lab	C101	DynMachine learning to uncover biological interactions
	Damian Roqueiro		Testability and Tarone's method to correct for multiple hypothesis testing
	Laetitia Papaxanthos		Significant interval search with correction for confounders
	Anja Gumpinger		Interactions in a graph via significant subgraph mining
16:00-16:30	Coffee Break		
16:30-18:30			Discussion

Day 2 (28 April 2016) - Complex interactions

Time	Speakers	Room	Topic description
8:00-8:30	Registration		
8:30-10:30	Heather Cordell	R010	Gene-gene and gene-environment interactions (with practicum)
10:30-11:00	Coffee Break		
11:00-13:00	Christina Kiel	K201	Quantitative systems analysis of protein interaction networks, 3D structural information and protein design to quantify the effect of disease mutations
13:00-14:00	Lunch		
14:00-16:00	Patrick Van Dijck	K201	Molecular interactions between pathogenic fungi, bacteria and their host
16:00-16:30	Coffee Break		

Day 2 (28 April 2016) - Complex interactions

Time	Speakers	Room	Topic description
16:30-18:30	Participants	K201	Case-studies from participants
	Floris Schoeters		A high-throughput Candida albicans two hybrid system
	François Van Lishout		I have found a statistical interaction – what next?
	Taesung Park		Unified multifactor dimensionality reduction analysis for gene-gene interaction analysis
	Yongkang Kim		Gene-Gene Interaction Analysis for Bipolar Disorder
> 18:30	Social event		

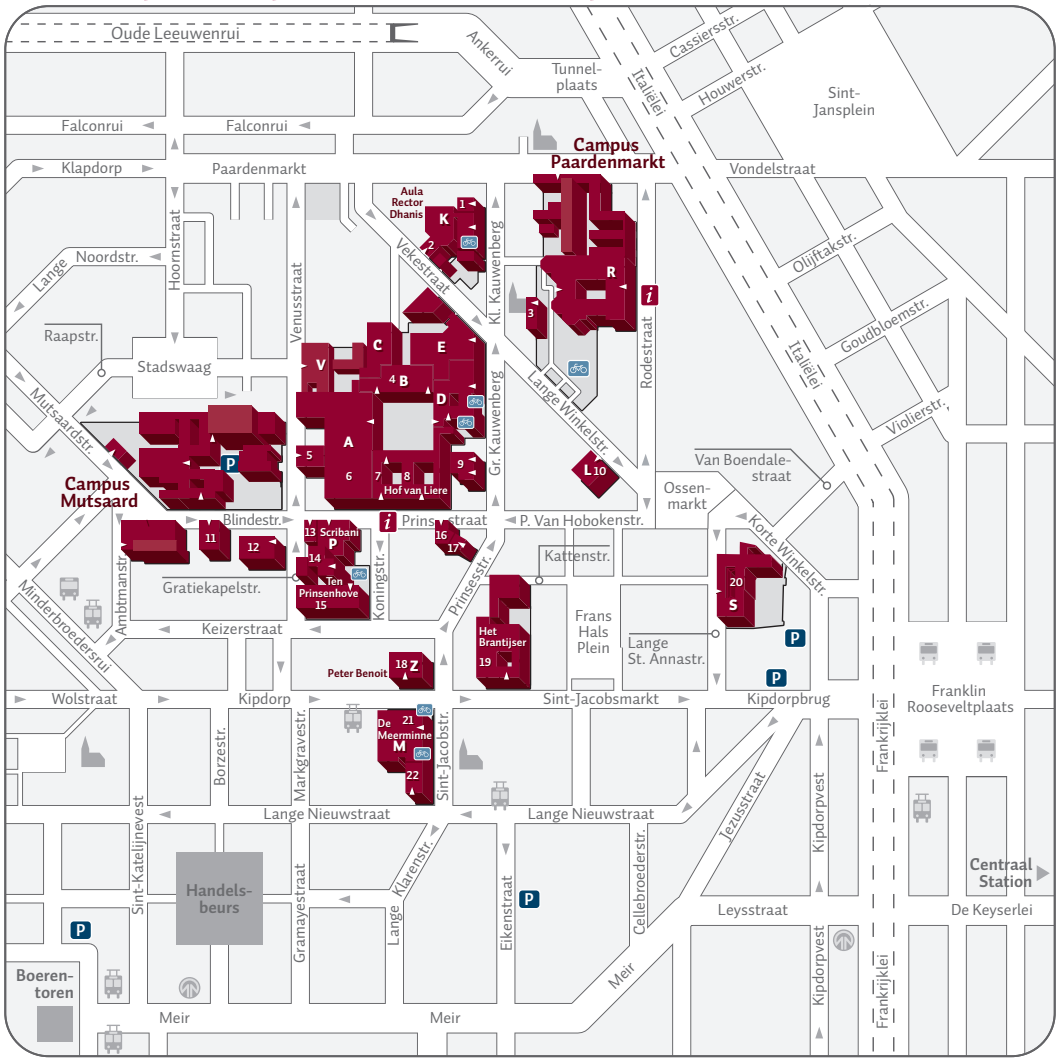
Day 3 (29 April 2016) - Analytic tools – an example

Time	Speakers	Room	Topic description
9:00-10:30	Giorgio Russolillo	K201	A tour in components-based path modeling
10:30-11:00	Coffee Break		
11:00-12:30	Giorgio Russolillo	K201	A tour in components-based path modeling
12:30-13:30	Lunch		
13:30-15:00	Giorgio Russolillo	K201	A tour in components-based path modeling
15:00-15:30	Coffee Break		
15:30-17:00	Giorgio Russolillo	K201	A tour in components-based path modeling

Location

CSCDA

Stadscampus - Campus Paardenmarkt - Campus Mutsaard



Onthaal

- A **Gebouw A** - Prinsstraat 13
- B **Gebouw B** - Prinsstraat 13
- C **Gebouw C** - Prinsstraat 13
- D **Gebouw D** - Grote Kauwenberg 18
- E **Agora / Sporthal** - Grote Kauwenberg 2
- K **Aula Rector Dhanis** - Kleine Kauwenberg 14
- L **Gebouw L** - Lange Winkelstraat 40
- M **De Meerminne** - Sint-Jacobstraat 2
- P **Scribani** - Prinsstraat 10
- R **Gebouw R** - Rodestraat 14
- S **Grauwzusters** - Lange Sint-Annastraat 7
- V **Gebouw V** - Venusstraat 23
- Z **Peter Benoit** - Kipdorp 61

1. **Linguapolis** - Kleine Kauwenberg 12
2. **Centraal magazijn** - Vekestraat 33
3. **Annexe** - Rodestraat 14
4. **Steunpunt Goederenstromen** - Prinsstraat 13
5. **IOIW / Moreel Consulent / Preventiedienst** - Venusstraat 35
6. **Bibliotheek** - Prinsstraat 13
7. **Hof van Liere** - Prinsstraat 13
8. **Universiteitsclub** - Prinsstraat 13b
9. **Ruusbroecgenootschap** - Grote Kauwenberg 32-34
10. **L. P. Boon Documentatiecentrum / Studiecentrum H. Claus / CLiPS / IJS** - Lange Winkelstraat 40
11. **Studiecentrum Open Universiteit/Centrum WeST/ Centre for Law and Cosmopolitan Values** - Blindestraat 14
12. **IOIW / Dienst Internationale Samenwerking** - Gratiekapelstraat 10
13. **Linguapolis** - Prinsstraat 8
14. **Studentenhome Ten Prinsenhove** - Koningstraat 8
15. **Restaurant Ten Prinsenhove / Labotheek** - Koningstraat 8
16. **Zomaar een dak** - Prinsstraat 32
17. **Universitas** - Prinsesstraat 16
18. **ITMMA** - Kipdorp 59
19. **CNO / CVA / AMS - Het Brantijser** - Sint-Jacobsmarkt 13
20. **IOB/Onderwijsadministratie, Studieadvies en studentenbegeleiding/ Steunpunt Buitenlands Beleid** - Lange Sint-Annastraat 7
21. **OASeS** - Sint-Jacobstraat 2
22. **CeMIS / Steunpunt Gelijkekansenbeleid** - Lange Nieuwstraat 55

Universiteit Antwerpen

Stijn Vansteelandt (DAY 1, 27 April 2016)

Stijn Vansteelandt graduated as Master in Mathematics at Ghent University in 1998, and obtained a PhD in Mathematics (Statistics) in 2002 at the same university. After postdoctoral research at the Department of Biostatistics of the Harvard School of Public Health, he returned to Ghent University in 2004, where he is now Full Professor in the Department of Applied Mathematics, Computer Science and Statistics. He is furthermore Honorary Professor in the Department of Epidemiology and Population Health at the London School of Public Health.

Stijn Vansteelandt is a leading expert in causal inference: a fast-growing field within statistics, which focuses on the development of statistical methods for inferring the causal effect of an exposure on an outcome from experimental and observational data under minimal and well-understood assumptions. He has authored over 100 peer-reviewed publications in international journals on a variety of topics in biostatistics, epidemiology and medicine, such as the analysis of longitudinal and clustered data, missing data, mediation and moderation/interaction, instrumental variables, family-based genetic association studies, analysis of outcome-dependent samples and phylogenetic inference. He is Co-Editor of Biometrics, the leading flagship journal of the International Biometrics Society, and has previously served as Associate Editor for the journals Biometrics, Biostatistics, Epidemiology, Epidemiologic Methods and the Journal of Causal Inference.

Maggie Wang (DAY 1, 27 April 2016)

Maggie Wang obtained her PhD degree from the Hong Kong University of Science and Technology in Statistics in 2011, and in the same year, joined Chinese University of Hong Kong as a faculty member. Dr. Wang is the Assistant Director of the Master program in Epidemiology and Biostatistics in the School of Public Health and Primary Care, CUHK. She is the principle investigator of the Hong Kong RGC-GRF grant, National Science Foundation of China (NSFC) General Project and Youth Project. Her major research area is developing bioinformatics methods for variable screening and clinical applications. She also collaborate with Shenzhen CDC and Zhejiang CDC in infectious disease viral evolution and transmission modelling.

Kristel Van Steen (DAY 1, 27 April 2016)

Kristel Van Steen received a Master of Science degree in 1991 and a Doctoral degree in 1996 both in mathematics from Ghent University (Belgium). She also received a Master of Science degree in Biostatistics in 1999 and a second Doctoral degree in Biomedical Science in 2005 from Hasselt University (Belgium). She worked on family-based genetic association tests during a post-doc of two years (2003-2005) at the Harvard School of Public health (Harvard University, Boston, USA). During her stay in the US she acquired Biotechnology Project Management skills at the Harvard Extension School and the Massachusetts Biotechnology Council. Currently, Kristel Van Steen is an Associate Professor in bioinformatics / statistical genetics and is leading the BIO3 (Biostatistics, Biomedicine, Bioinformatics) group. This group is part of the academic research center GIGA (Interdisciplinary Group of Applied Proteomic) at the University of Liège, in particular the Medical Genomics GIGA-R thematic research unit, of which she is Director. In addition, she holds honorary guest (professor) positions at the Belgian Center of Human Genetics at the Catholic University of Leuven (KUL) and at the Center of Medical Genetics at Ghent University (UG). Kristel's interest in systems genetics lies in developing and applying methods to detect gene x gene and gene x environment interactions, in unifying biological and statistical networks, as well as in omics-stratified precision medicine.

Karsten Borgwardt lab (DAY 1, 27 April 2016)

Damian, Laetitia and Anja are members of the Machine Learning and Computational Biology (MLCB) lab lead by Prof. Dr. Karsten Borgwardt at ETH Zürich.

Damian received an engineer diploma (Dipl.-Ing.) in Information Systems from the National University of Technology (Universidad Tecnólogica Nacional) in Buenos Aires, Argentina. He continued his graduate studies in the United States where he obtained an M.Sc. in Computer Science and a Ph.D. in Bioinformatics from the University of Illinois at Chicago. After completion of his doctoral studies in September of 2013 he joined the MLCB lab to work on applying machine learning techniques in the creation of probabilistic models to better understand the association between specific diseases and the genetic markup of individuals who are affected by those diseases. His current work centers around the topic of significant subgraph mining and the development of methods to perform multi-trait genome-wide association studies (GWASs).

Laetitia joined the MLCB lab in October of 2014 as a PhD candidate. She finished her studies at the École Normale Supérieure de la Rue d'Ulm in Paris and in the École des mines de Paris where she studied theoretical physics, computational physics and biotechnology. Her work at MLCB focuses on the development of significant pattern mining methods and on their applicability to key aspects of GWASs such as genetic heterogeneity and confounding factors. Currently, she is also developing new methods for promoter prediction.

Anja received her M.Sc. degree from the Department of Mathematics at Technische Universität München where she worked on time series analysis of gene expression data using transfer entropy. She joined the MLCB lab as a Ph.D. candidate in May of 2015. Her current research work focuses on significant subgraph mining, that is to say, on the identification of patterns -mainly interacting subgraph structures- in complex molecular graphs.

Heather Cordell (DAY 2, 28 April 2016)

Heather Cordell is Professor of Statistical Genetics and a Wellcome Senior Fellow in the Institute of Genetic Medicine at Newcastle University, UK. Heather obtained her undergraduate degree in Mathematics from Cambridge University followed by an MSc in Applied Statistics and DPhil (PhD) in Mathematical Genetics from Oxford University. She then spent three postdoctoral years at Case Western Reserve University in Cleveland, Ohio. From 2000-2004 Heather held a Wellcome Trust/JDRF Career Development Fellowship at the Department of Medical Genetics in Cambridge, and in October 2004 she took up a Wellcome Senior Fellowship at the same department. In 2006 Heather moved to Newcastle University to take up the newly-established Chair of Statistical Genetics. In addition to being involved in a number of applied studies, Heather's research interests include the development of methods for detecting linkage/association (including maternal and parent-of-origin effects) using family-based data, and the modelling of effects at multiple disease loci (including interaction effects) simultaneously. Heather was previously (2006-2012) a member of the Board of Directors of the International Genetic Epidemiology Society and acted as the 2010-2011 President of the Society.

Christina Kiel (DAY 2, 28 April 2016)

Christina Kiel studied Biochemistry at the Ruhr-University Bochum and obtained her PhD degree in the Structural Biology department of the Max-Planck-Institute for Molecular Physiology in 2003. She moved for her postdoctoral studies to the Department of Structural and Computational Biology at EMBL Heidelberg. Currently she is Staff Scientist in the Systems Biology Department at the Center for Genomic Regulation in Barcelona, where she is heading the signaling sub-team in the laboratory of Luis Serrano. Her major research area is the quantitative, systems and structural analysis of signaling pathways and protein interaction networks relevant in human diseases.

Patrick Van Dijck (DAY 2, 28 April 2016)

In our lab we are studying human fungal pathogens, such as *Candida albicans* and *Candida glabrata*. These species are commensal organisms in most people, but in immoncompromized patients, these pathogens can become deadly organisms. Each species has a set of virulence factors. The most important virulence factor in *C. albicans* is the yeast-to-hyphae morphogenetic switch and in *C. glabrata*, the expression of up to 70 different adhesion proteins.
Currently a major problem in hospitals with these species is that they easily form biofilms on different types of implants, such as catheters. A biofilm is a community of cells that protect themselves from the immune system and from antifungal drugs by producing an extracellular matrix. In such a biofilm cells need to attach to the substrate as well as to each other. We are investigating the proteins that are involved in these interactions. We use molecular biology and cell biology tools but also atomic force microscopy (AFM).

Giorgio Russolillo (DAY 3, 29 April 2016)

Giorgio Russolillo is an Assistant Professor (Maître de Conferences) at the Conservatoire National des Arts et Métiers (CNAM) of Paris, where he teaches Statistics. He is a member of the Centre d'Etudes et de Recherche en Informatique et Communications (CEDRIC) laboratory. He earned his Doctorate in Statistics from the Università degli Studi di Napoli "Federico II" (Italy) in 2010 and joined CNAM in the same year as a postdoctoral fellow. He spent 2005-2006 at the Italian National Research Council as a biostatistician.
His research interests involve Multivariate Data Analysis models and methods applied to social and life sciences. Much of his current research focuses on Partial Least Squares (PLS) algorithms. In particular, he develops PLS-type algorithms for handling both quantitative and qualitative variables in component-based approaches to regularized regression and Structural Equation Modeling

Floris Schoeters KU Leuven, Belgium

Title: A high-throughput *Candida albicans* two hybrid (C2H) system
Abstract: By translating the CUG codon into serine rather than leucine *Candida albicans* does not follow the universal genetic code. This makes it hard to use the model system, *Saccharomyces cerevisiae*, as a host for two-hybrid experiments. To overcome this problem we have developed a candida two hybrid (C2H) system to screen for protein-protein interactions in *C. albicans*. With the availability of the orfeome collection (containing all the ORFs from *C. albicans* in gateway adapted vectors) is it now also possible to apply a high-throughput approach using the C2H system. In order to set up this high-throughput C2H system the two hybrid vectors had to be made gateway compatible and a mating strategy had to be optimized and tested.
By applying the gateway system, the mating strategy and using approximately 12000 potential protein-protein interactions to test the system we now have an optimized high-throughput C2H system.

Francois Van Lishout University of Liege, Liege, Belgium

Title: I have found a statistical interaction – what next?
Abstract: In genetic association studies for common complex diseases, single nucleotide polymorphisms (SNPs) are the most commonly used type of genetic markers. This is in part understood by their dense distribution across the genome and their low mutation rate. SNPs are also increasingly being used in large-scale DNA x DNA interaction (GWAI) studies. Irrespective of the analytic method chosen, good common practice should be maintained, especially since several issues remain unresolved. These issues include how to best deal with heterogeneity and population structure and how to make the translation from statistically significant findings to biological relevance of clinical impact. Here, we focus on actions that can be taken to put statistical SNP-based epistasis findings into their appropriate context, such that the way is paved for functional follow-ups.

Taesung Park Seoul National University, Seoul, Republic of Korea

Title: Unified multifactor dimensionality reduction analysis for gene-gene interaction analysis
Abstract: Epistasis is one of the clues to explaining the missing heritability of common complex traits. The multifactor dimensionality reduction (MDR) method has been popular for detecting interaction effects. However, there are several disadvantage of existing MDR approaches, such as lack of efficient way of evaluate significance of multi-locus models and a heavy burden of computation because of intensive permutation is needed. Further, MDR does not distinguish marginal effects from pure interaction effects. In this work, we propose a two-step unified model-based MDR approach (UM-MDR), in which, the significance of a multi-locus model, even a high-order model, can be easily obtained through a regression framework and a semiparametric correction procedure. The proposed UM-MDR approach is flexible in the sense of its ability to incorporate different types of trait and evaluate significances of existing MDR extensions. Simulation studies and an application to a real example of a genome-wide association study of Korean population are provided to demonstrate the utility of the proposed method. UM-MDR achieves the same power as MDR for most scenarios, and it outperforms MDR, especially when there are some SNPs having only marginal effects, which makes it difficult for the existing MDR approaches to detect the causal epistasis.

Yongkang Kim Seoul National University, Seoul, Republic of Korea

Title: Gene-Gene Interaction Analysis for Bipolar Disorder

Abstract: Bipolar disorder is well known mental diseases with high heritability. Although many clues show that bipolar disorder is caused by some genetic factors, missing heritability makes hard to explain genetic mechanism of bipolar disorder. In our recent study, we found that the interaction effect between two genes having strong main effects was significant for bipolar disorder. In this talk, we provide a more though investigation of the interaction effect between two genes via classical logistic regression model and multi-factor dimensionality reduction analysis.