ANSC*6240_S02 – W20 (provisional code)

Systems Biology: Integration of Structural and Functional Genomic data & -OMICS technologies

Course Outline, Winter 2020

Instructor
Angela Canovas
Animal Bioscience, office #125
acanovas@uoguelph.ca
(519) 824-4129 ext. 56295

Lectures
Wednesday , 14:30 -17:20, ANNU #101
Theory lectures
Practical lecture / Computer Labs

Credits: 0.50
Academic level: Graduate

Course Description

The course is designed to provide a framework to effectively use of new –OMICS technologies combining structural and functional genomic data into systems biology to understand the nature of complex phenotypes in animals. Using systems biology, the students will learn how to integrate the genomic information to develop testable hypothesis in relation to the regulation of gene networks associated with economically important traits in livestock.

Topics of the course include:

Block 1.- Genomics & fine mapping of positional candidate markers

- Genome-Wide Association Studies (GWAS). Identification of candidate genes/variants using results obtained from GWAS, whole genome sequencing, exome, etc. Summary of the basis of association studies using genomic data. Additionally, the course will introduce the concepts of variant identification and classification including SNP, INDELs and splice variants

- Fine mapping and identification of positional candidate genes. Identification of genes mapped in candidate regions identified by association studies using genomic annotation data provided by public repositories (Ensembl, NCBI, UCSC, etc.). Identification of haplotype blocks and linkage disequilibrium pattern analysis.

- QTL mapping of positional candidate genes. Identification of QTLs mapped in regions where candidate genes were identified using the public QTLdb (https://www.animalgenome.org/cgi-bin/QTLdb/).

Block 2.- Genetical Genomics to identify functional candidate genes using GWAS and other genomic data.

- Identification of functional candidate genes among statistical and positional candidates using functional genomic approach.

- Enrichment analysis for Gene Ontology terms, metabolic pathways and gene network analysis. Definition of GO terms categories and structure. To introduce data bases for metabolic pathway information.

- Prediction of gene networks to identify groups of related genes.

Block 3.- Transcriptomics analysis using RNA-Sequencing technology

- Definition of transcriptomics and RNA-Seq technology.

- RNA-Sequencing analysis using different software and packages. Quality control analysis. Identification of differentially expressed genes using RNA-Sequencing data. Prediction of key regulatory genes and transcription factor.

- SNP Discovery and variant detection using RNA-Sequencing. Differences between variants identification using genomic and transcriptomic data. Pipelines for structural variant identification using RNA-Seq data. Annotation and functional impact prediction and consequences.

Block 4.- Functional Genomics

- Functional Genomics analysis to identify the key regulatory genes and transcription factors using different software/ approaches. GO enrichment, metabolic pathway and gene network analysis using gene expression data.

-Introduction to Metagenomics, Metabolomics and other –OMICS technologies analysis.

Block 5.- Systems biology

- Importance of other –OMICS approaches, such as proteomics, metabolomics, metagenomics, in the identification of candidate genes for complex traits.

- Impact of data integration over accuracy of selection. Integration of multi-omics using results from different tissues, breeds and even organisms.

- Interpretation of the biological meaning of statistical results obtained by multi-OMIC and functional genomics approaches.

- Transition from theoretical assumptions to applicability.

- The use of genomics tools to accelerate the genetic improvement in livestock. Improving selection efficiency by combining functional and structural genomic data with the estimation of molecular breeding values in livestock species.

Computer Labs

- Computer Lab 1: Fine mapping of positional candidate markers: Gene and Genome regions annotation using Biomart. Haplotype blocks and LD analysis using haploview.

- Computer Lab 2: Practical exercises / data analysis using different software, package and approaches such as BLAS2GO, CateGOrizer and AmiGO software for Gene Ontology and enrichment analysis; DAVID, BLAST2GO, Reactome software for metabolic pathways analysis; STRING, GeneMania and NetworkAnalyst software for gene network analysis. Identification of functional candidate genes (GO, metabolic pathways and Gene network) using associated software and R packages.

- Computer Lab 3: Predicting Functional Consequences of SNPs using variant effect predictor.

- Computer Lab 4: Integration data – Systems Biology. Integration of 4 groups of data from the previous results obtained from Computer Lab 1, 2 and 3 such as list of genes with significant SNP from GWAS (Computer Lab 1), list of genes involved in significant metabolic pathways (Computer Lab 2), list of genes connected in gene networks (Computer Lab 2) and list of genes with SNP with functional consequences (Computer Lab 3)

- Computer Lab 5: Practical exercises. Integration of structural and functional data from multi-omics using results from different tissues, breeds and species. To perform the Ven diagram with the Integration of 5 groups of data from the results / inputs provided from transcriptomics analysis using RNA-Sequencing technology including the list of differentially expressed genes between healthy and sick sheep from two different breeds of sheep, list of SNP significant associated with health in sheep from a GWAS study, list of Ensembl IDs for the genes associated with the intestinal immune network for IgA production pathway in sheep and list Genomic Regions also associated with health in sheep

Course objectives

To provide learning opportunities in:

- Integration of structural and functional genomic data into a Systems Biology approach.
- Transcriptomics and other –OMICS analysis using different software and packages.
- Functional genomics analysis using different software and packages.
- Methods for optimizing genetic improvement strategies considering genomic and quantitative information.
- "-OMICS" technologies and methodologies to accelerate the genetic improvement in livestock.
- Improving selection efficiency by combining functional and structural genomic data with the estimation of molecular breeding values in livestock species.
- Graduate-level writing and oral presentation skills.

Materials

- Lecture slides and other materials will be posted on CourseLink.
- Manuals of the different software used in the course will be posted on Courselink
- R Manual, R codes and/or R packages will be provided and posted on Courselink.
- Suggested reading material (examples and references used in the course).

Barrett, J. C., Fry, B., Maller, J. D. M. J., & Daly, M. J. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics, 21(2), 263-265. 2004.

SNP discovery in the bovine milk transcriptome using RNA-Seq technology. Cánovas A, Rincon G, Islas-Trejo A, Wickramasinghe S, Medrano JF. Mamm Genome. 21(11-12):592-8. doi: 10.1007/s00335-010-9297-z. 2010.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., ... & Pachter, L. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nature protocols, 7(3), 562-578. 2012.

Franceschini, A. STRINGdb Package Vignette. Nucleic Acids Res. 2013.

RNA sequencing to study gene expression and single nucleotide polymorphism variation associated with citrate content in cow milk. Cánovas A, Rincón G, Islas-Trejo A, Jimenez-Flores R, Laubscher A, Medrano JF. J Dairy Sci. 96(4):2637-2648. doi: 10.3168/jds.2012-6213. 2013.

Marjoram, P., Zubair, A., & Nuzhdin, S. V. Post-GWAS: where next? More samples, more SNPs or more biology?. Heredity, 112(1), 79-88. 2014.

Review: RNA-Seq applications in Livestock. Wickramasinghe, S., Cánovas, A., Rincon, G., & Medrano, J. F. Livestock Science, 166, 206-216, doi: 10.1016/j.livsci.2014.06.015. 2014.

Comparison of five different RNA sources to examine the lactating bovine mammary gland transcriptome using RNA-Sequencing. Cánovas A, Rincón G, Bevilacqua C, Islas-Trejo A, Brenaut P, Hovey RC, Boutinaud M, Morgenthaler C, VanKlompenberg MK, Martin P, Medrano JF. Sci Rep. 4:5297. doi: 10.1038/srep05297. 2014.

Multi-tissue omics analyses reveal molecular regulatory networks for puberty in composite beef cattle. Cánovas A, Reverter A, DeAtley KL, Ashley RL, Colgrave ML, Fortes MR, Islas-Trejo A, Lehnert S, Porto-Neto L, Rincón G, Silver GA, Snelling WM, Medrano JF, Thomas MG. PLoS One. 9(7):e102551. doi: 10.1371/journal.pone.0102551. 2014.

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience, 4(1), 7. 2015.

Smedley, D., Haider, S., Durinck, S., Pandini, L., Provero, P., Allen, J., & Bardou, P. The BioMart community portal: an innovative alternative to large, centralized data repositories. Nucleic acids research, 43(W1), W589-W598. 2015.

Transcriptome analyses identify five transcription factors differentially expressed in the hypothalamus of post- versus prepubertal Brahman heifers. Fortes MR, Nguyen LT, Weller MM, Cánovas A, Islas-Trejo A, Porto-Neto LR, Reverter A, Lehnert SA, Boe-Hansen GB, Thomas MG, Medrano JF, Moore SS. J Anim Sci. 94(9):3693-3702. doi: 10.2527/jas.2016-0471. 2016.

Hu, Z. L., Park, C. A., & Reecy, J. M. Developmental progress and current status of the Animal QTLdb. Nucleic acids research, 44(D1), D827-D833. 2016.

RNA-seq based detection of differentially expressed genes in the skeletal muscle of Duroc pigs with distinct lipid profiles. Cardoso TF, Cánovas A, Canela-Xandri O, González-Prendes R, Amills M, Quintanilla R. Sci Rep. 7:40005. doi: 10.1038/srep40005. 2017.

Joint QTL mapping and gene expression analysis identify positional candidate genes influencing pork quality traits. González-Prendes R, Quintanilla R, Cánovas A, Manunza A, Figueiredo Cardoso T, Jordana J, Noguera JL, Pena RN, Amills M. Sci Rep. 7:39830. doi: 10.1038/srep39830. 2017.

Global differential gene expression in the pituitary gland and the ovaries of pre- and postpubertal Brahman heifers. Nguyen LT, Reverter A, Cánovas A, Venus B, Islas-Trejo A, Porto-Neto LR, Lehnert SA, Medrano JF, Moore SS, Fortes MR. J Anim Sci. 95(2):599-615. doi: 10.2527/jas.2016.0921. 2017.

SNP detection using RNA-sequences of candidate genes associated with puberty in cattle. Dias MM, Cánovas A, Mantilla-Rojas C, Riley DG, Luna-Nevarez P, Coleman SJ, Speidel SE, Enns RM, Islas-Trejo A, Medrano JF, Moore SS, Fortes MR, Nguyen LT, Venus B, Diaz IS, Souza FR, Fonseca LF, Baldi F, Albuquerque LG, Thomas MG, Oliveira HN. Genet Mol Res. 16(1). doi: 10.4238/gmr16019522. 2017.

Coursework

Grades will be based on:

1. Report project based on the computer Labs (75%).

The structure of the course combines both theory and practical (Computer Lab) lectures. Based on the topics and section covered in the course, the students will perform five computer Labs.

The students will prepare a draft report showing their progress in the data analysis for each of the five computer Labs. The instructor will review and grade them.

The Computer Lab reports are due at the end of every section (specific deadline will be provided at the beginning of the course). Each computer Lab will count the 15% of the final grade (15% grade x 5 Computer Lab = 75%). All reports can be submitted via the Dropbox on CourseLink.

2. Paper review and oral presentation (25%)

The paper review is designed to provide an opportunity to discuss course topics as they relate to current research in Systems Biology. Student will find a research paper to review, which must be approved by the instructor. Each student will present their review in a short (8-10 min) presentation during the class toward the end of the semester (date to be determined). Grades will be based on students' own presentation and participation in discussing others' papers. Titles and citations of two potential papers should be submitted to the Dropbox by the end of week 7 of the course (beginning of March). Paper review and oral presentation will count the 25% of the final grade.