Planning for informatics in your grant applications

Samuel Volchenboum, MD, PhD

August 1, 2017

THE UNIVERSITY OF CHICAGO MEDICINE & BIOLOGICAL SCIENCES | CENTER FOR RESEARCH INFORMATICS

At the end of this talk, you will…

- Know what parts of a grant need informatics consideration
- Understand how important it is to seek help early
- Feel comfortable reaching out to CRI and asking for help

The idealized process…

- Have an idea
- Get preliminary data
- Write a proposal
- Get funding
- Do work
- Repeat

A more realistic process

- Have an idea or an extension of current work
- Apply for grant using old preliminary data
- Get award for new work
- Figure out how to actually do (and pay for) the work

Scenario #1 - The sequencer

- Researcher gets a pilot grant to study colon cancer patients using ChIP-Seq
- Pilot grant only for cost of sequencing
- No provisions made for analysis and interpretation

Scenario #2 - The multi-center trial

- Researcher gets U grant for testing a new survey tool at 30 cooperative sites
- Grant has no provisions for any research informatics support

Scenario #3 - The Big Data™ user

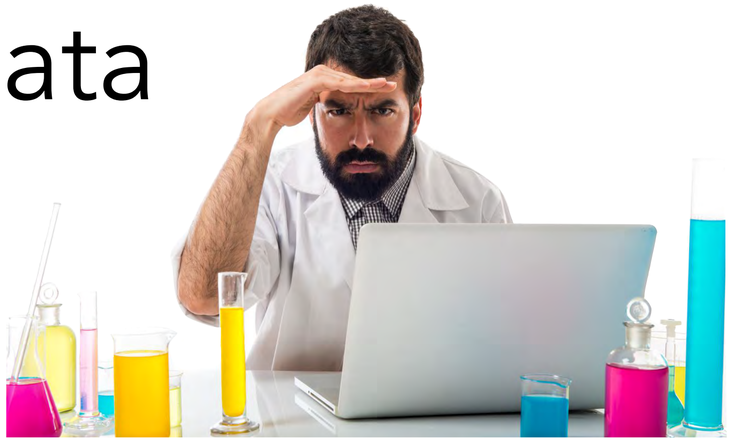
- Researcher gets funding to sequence 1000 whole genomes
- Gets funding for sequencing but then needs 20TB of storage space
- No grant provisions for storage or backup

Scenario #4 - The simulator

- Researcher gets funding to design, perform, and test molecular simulations on millions of drug-target combinations
- Requires millions of hours of HPC usage
- No funding for HPC

Scenario #5 - The analyzer

- Funding secured for pulling a large comprehensive data set from the data warehouse to perform disease modeling
- Data is pulled and given to research team but there is no one to analyze the data

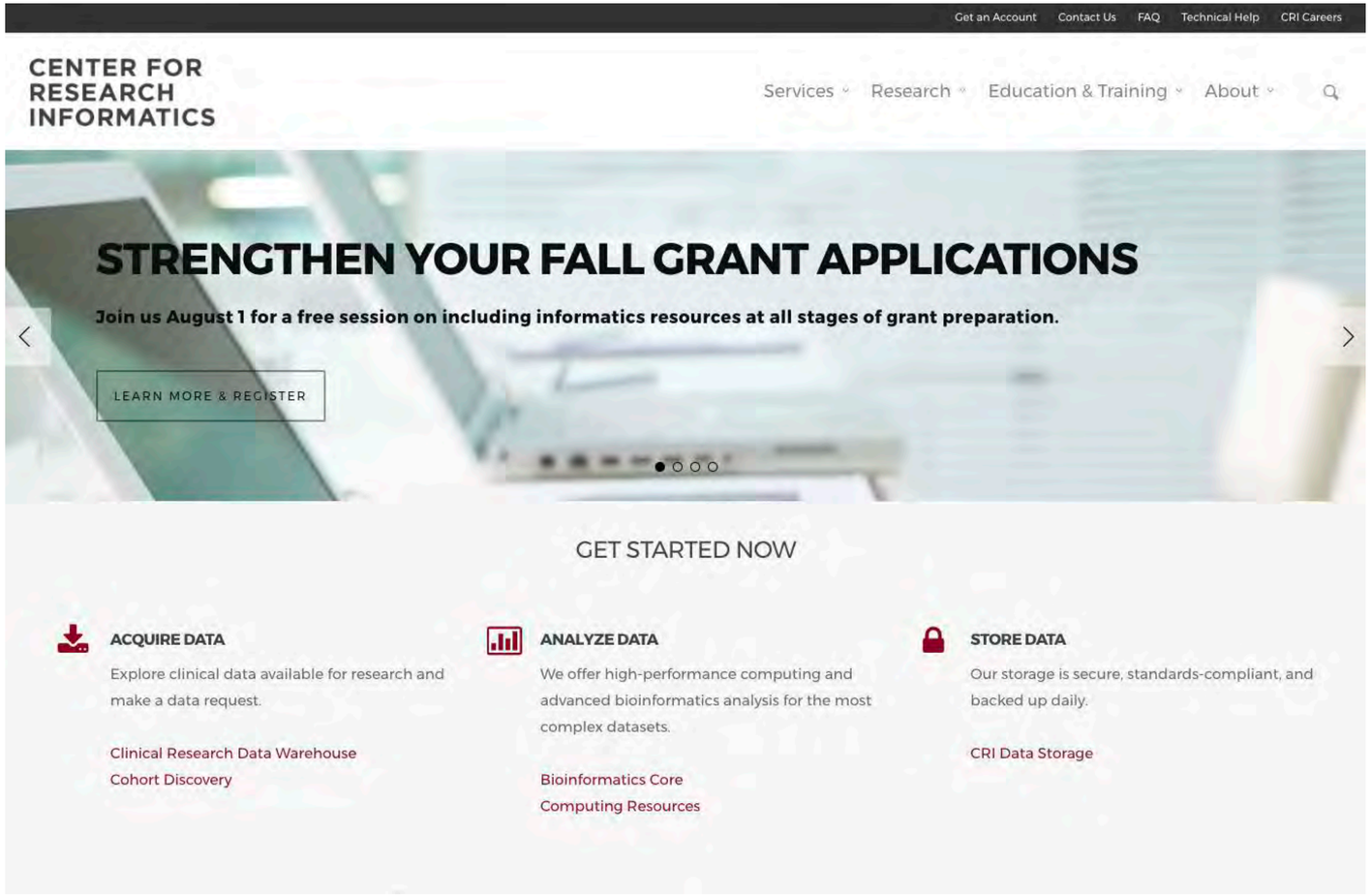


There are many opportunities to consider informatics resources

The best time is when you're just thinking about a project or writing about it.

Getting informatics help

http://cri.uchicago.edu



slv@uchicago.edu

support@rt.cri.uchicago.edu

Common to all proposals

- IRB writing and positioning
- Contracts, data use agreements
- Data storage, movement, backup
- Letters of support
- Facilities and resources documentation
- Data governance and stewardship
- Data sharing / software dissemination

IRB writing / positioning

- CRI has extensive experience in writing IRB protocols and shepherding them through the process
- Many of the issues have already been encountered for other proposals
- Engage the CRI early on in the process

Contracts and data use agreements

- Sharing data outside the BSD requires an agreement
- Contracts may be needed for IP, data use, etc.
- Monthly meeting with CRI, OCR, IRB, legal, and security to discuss and address these issues **proactively**

Data storage, movement, backup

- CRI has extensive storage and backup capabilities
- Every investigator gets 2TB storage and backup for “free” as a lab share
- More extensive data usage needs to have a budget

A word about storage

These aren't good places to store your data.

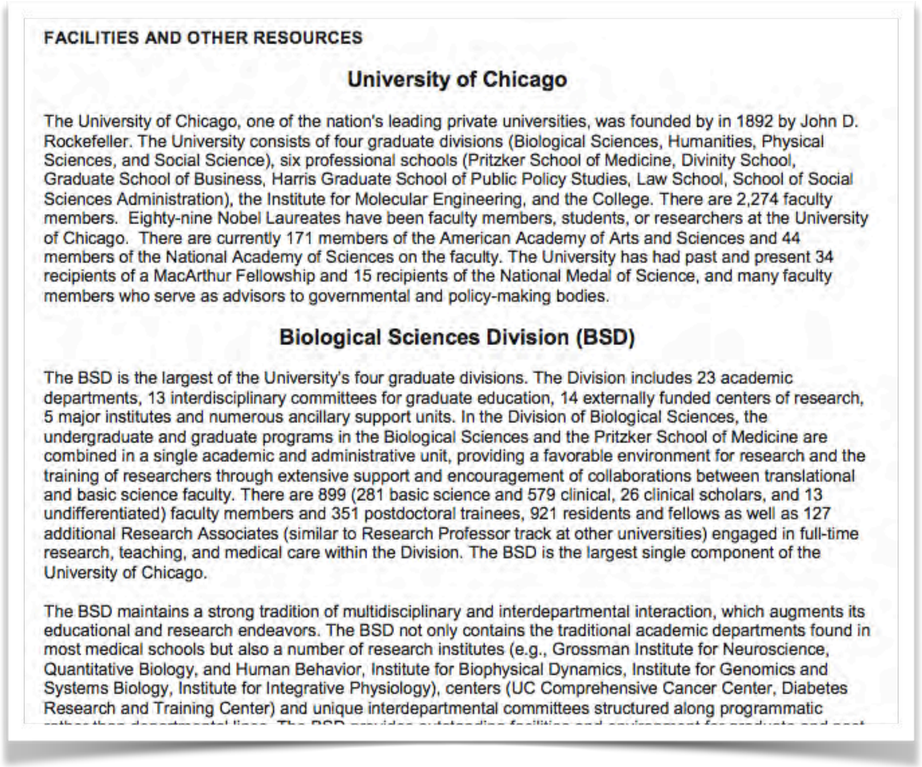
Why?

- Not HIPAA compliant
- Insecure
- No redundant backup
- Little chance of recovery if loss

Letters of support

- General letter from CRI
- Specific support for project from CRI leadership
- Contact the CRI director service line director for any LoS issues
- Do this early. A draft is always appreciated.

Facilities and resources pages



FACILITIES AND OTHER RESOURCES

University of Chicago

The University of Chicago, one of the nation's leading private universities, was founded by in 1892 by John D. Rockefeller. The University consists of four graduate divisions (Biological Sciences, Humanities, Physical Sciences, and Social Science), six professional schools (Pritzker School of Medicine, Divinity School, Graduate School of Business, Harris Graduate School of Public Policy Studies, Law School, School of Social Sciences Administration), the Institute for Molecular Engineering, and the College. There are 2,274 faculty members. Eighty-nine Nobel Laureates have been faculty members, students, or researchers at the University of Chicago. There are currently 171 members of the American Academy of Arts and Sciences and 44 members of the National Academy of Sciences on the faculty. The University has had past and present 34 recipients of a MacArthur Fellowship and 15 recipients of the National Medal of Science, and many faculty members who serve as advisors to governmental and policy-making bodies.

Biological Sciences Division (BSD)

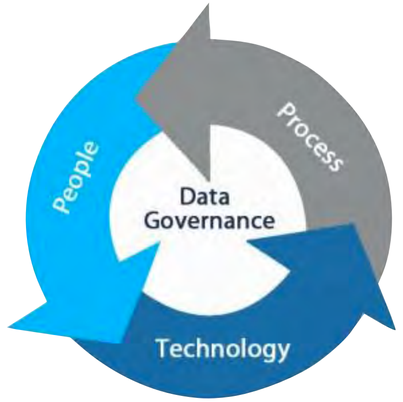
The BSD is the largest of the University's four graduate divisions. The Division includes 23 academic departments, 13 interdisciplinary committees for graduate education, 14 externally funded centers of research, 5 major institutes and numerous ancillary support units. In the Division of Biological Sciences, the undergraduate and graduate programs in the Biological Sciences and the Pritzker School of Medicine are combined in a single academic and administrative unit, providing a favorable environment for research and the training of researchers through extensive support and encouragement of collaborations between translational and basic science faculty. There are 899 (281 basic science and 579 clinical, 26 clinical scholars, and 13 undifferentiated) faculty members and 351 postdoctoral trainees, 921 residents and fellows as well as 127 additional Research Associates (similar to Research Professor track at other universities) engaged in full-time research, teaching, and medical care within the Division. The BSD is the largest single component of the University of Chicago.

The BSD maintains a strong tradition of multidisciplinary and interdepartmental interaction, which augments its educational and research endeavors. The BSD not only contains the traditional academic departments found in most medical schools but also a number of research institutes (e.g., Grossman Institute for Neuroscience, Quantitative Biology, and Human Behavior, Institute for Biophysical Dynamics, Institute for Genomics and Systems Biology, Institute for Integrative Physiology), centers (UC Comprehensive Cancer Center, Diabetes Research and Training Center) and unique interdepartmental committees structured along programmatic

CRI has boilerplate language for grants

Data governance and stewardship

- Grant readers are now looking for documentation of data governance procedures
- CRI can help document these procedures for your proposal



Examples of data governance considerations

- Why controls access to data?
- How is security documented?
- Will people have encrypted laptops?
- Is the storage HIPAA compliant?
- Are data being backed up regularly?
- How is data being moved securely between researchers?

Failure to address these questions adequately can doom a proposal.

Data sharing plan

- Data sharing
  - Discussion of how data will be deposited in common repositories and shared
- Software dissemination
  - How will software be shared?
  - What kind of license will be used?
- CRI will help with this

Bioinformatics considerations

- Methods and study design
- Budget planning for data generation
- Grant writing - preliminary data, methods, research plan
- Data storage, movement, backup
- Analysis and interpretation
- Integration of multiple data sources
- Manuscript preparation and submission

Bioinformatics - Methods and study design

- What kind of analysis?
  RNA-Seq? ChIP-Seq? WGS? WES?
- What depth of coverage?
- Power calculations: How many samples?
  Technical replicates? Biological replicates?

Bioinformatics - Budget planning for data generation

- How many chips? What cost to run?
- How about sample collection and preparation?
- CRI can help broker this process

Bioinformatics - Grant writing

- CRI can help with all phases of grant writing
  - Background
  - Preliminary data
  - Methods
  - Research plan

Bioinformatics - Data storage, movement, backup

- How much storage is needed?
- How will data be transferred between investigators?
- Is data being redundantly backed up?
- CRI can help ensure that all phases are secure

Bioinformatics - Analysis and interpretation

- Best to involve a bioinformatician from the start
- Partnership is key for a successful collaboration
- Project time is charged on an hourly basis or through dedicated time on grants
- Co-authorship is expected, where appropriate

Publications

2015

- Qin D, Huang L, Wlodaver A, Andrade J, Staley J. Sequencing of lariat termini in S. cerevisiae reveals 5' splice site, branch points, and novel splicing events. *RNA*. 2015 Dec 8. doi: 10.1261/rna.052829.115. [Epub ahead of print]
- Luke JJ, Spranger S, Bao R, Andrade J, Gajewski TF. The genetic landscape of the T cell non-inflamed tumor microenvironment in human solid tumors. *J Immunother Cancer*. 2015 Nov 4;3(Suppl 2):97. doi: 10.1186/2051-1426-3-S2-P97.
- Zha Y, Spranger S, Hernandez KM, Li Y, Bao R, Alexieff P, et al. Density of immunogenic antigens does not explain presence or absence of the T cell-inflamed tumor microenvironment in metastatic melanoma. *J Immunother Cancer*. 2015 Nov 4;3(Suppl 2):425 doi: 10.1186/2051-1426-3-S2-P425
- Singh P, Brock CO, Volden PA, Hernandez K, Skor M, Kocherginsky M, et al. Glucocorticoid receptor ChIP-sequencing of subcutaneous fat reveals modulation of inflammatory pathways. *Obesity*. 2015 Nov;23(11):2286-93. doi: 10.1002/oby.21251.
- Sasaki MM, Skol AD, Hungate EA, Bao R, Huang L, Kahn SA, et al. Whole-exome Sequence Analysis Implicates Rare Il17REL Variants in Familial and Sporadic Inflammatory Bowel Disease. *Inflamm Bowel Dis*. 2015 Oct 16. [Epub ahead of print]
- Kadri S, Zhen CJ, Wurst MN, Long BC, Jiang Z, Wang YL, et al. Amplicon Indel Hunter: A Novel Bioinformatics Tool to Detect Large Somatic Insertion/Deletion Mutations in Amplicon-Based Next-Generation Sequencing Data. *J Mol Diagn*. 2015 Aug 28. doi: 10.1016/j.jmoldx.2015.06.005. [Epub ahead of print]
- Hernandez KM, Understan
10.1111/nph.13607. Epub 201
- Zhong R, Bao R, Faber PW,
Aug 20. doi: 10.1158/0008-
- **Publication Highlight**: Bac
Exome Germline and Som
- Sasaki MM, Skol AD, Bao R
for Familial Nasopharynge
- Ernst LM, Rand CM, Bao R,
Cord Samples with Pathol
- **Publication Highlight**: Spr
9;523(7559):231-5. doi: 10.10
- Volchenboum SL, Andrade
prognostic importance of
10.1002/cjp2.9.

2017

- Kadri S, Lee J, Fitzpatrick C, Galanina N, Sukhanova M, Venkataraman G, et al. Clonal evolution underlying leukemia progression and Richter transformation in patients with ibrutinib-relapsed CLL. *Blood Advances*. 2017 May 9. doi: 10.1182/bloodadvances.2016003632.
- Saloura V, Fatima A, Zewde M, Kiyotani K, Brisson RJ, Park JH, et al. Characterization of the tumor T-cell receptor repertoire and immune microenvironment in patients with locoregionally advanced squamous cell carcinoma of the head and neck. *Clin Cancer Res*. 2017 April 25. doi: 10.1158/1078-0432.CCR-17-0103.
- Kach J, Long TM, Selman P, Tonsing-Carter EY, Bacalao MA, Lastra RR, et al. Selective glucocorticoid receptor modulators (SGRMs) delay castrate-resistant prostate cancer growth. *Mol Cancer Ther*. 2017 Apr 20. doi: 10.1158/1535-7163.MCT-16-0923. [Epub ahead of print]
- Wang LJ, Chan WC, Chou MC, Chou WJ, Lee MJ, Lee SY, et al. Polymorphisms of STS gene and SULT2A1 gene and neurosteroid levels in Han Chinese boys with attention-deficit/hyperactivity disorder: an exploratory investigation. *Sci Rep*. 2017 Apr 3. doi: 10.1038/srep45595.
- Pitroda S, Bao R, Andrade J, Weichselbaum RR, Connell PP. Low Recombination Proficiency Score (RPS) predicts heightened sensitivity to DNA-damaging chemotherapy in breast cancer. *Clin Cancer Res*. 2017 Mar 24. doi: 10.1158/1078-0432.CCR-16-2845
- Cortese R, Gileles-Hillel A, Khalyfa A, Almendros I, Akbarpour M, Khalyfa AA, et al. Aorta macrophage inflammatory and epigenetic changes in a murine model of obstructive sleep apnea: Potential role of CD36. *Sci. Rep*. 2017 Feb 27. doi: 10.1038/srep43648
- Li Y, Andrade J. DEApp: an interactive web interface for differential expression analysis of next generation sequence data. *Source Code for Biology and Medicine*. 2017 Feb 3. doi: 10.1186/s13029-017-0063-4
- Deb DK, Bao R, Li YC. Critical role of the cAMP-PKA pathway in hyperglycemia-induced epigenetic activation of fibrogenic program in the kidney. *FASEB J*. 2017 Feb 1. doi: 10.1096/fj.201601116R. [Epub ahead of print]
- Khalyfa A, Cortese R, Qiao Z, Ye H, Bao R, Andrade J, et al. Late gestational intermittent hypoxia induces metabolic and epigenetic changes in male adult offspring mice. *J Physiology*. 2017 Jan 15. doi: 10.1113/JP273570. [Epub ahead of print]

2016

- Spranger S, Luke JJ, Bao R, Zha Y, Hernandez KM, Gakewski AP, et al. Density of immunogenic antigens does not explain the presence or absence of the T-cell-inflamed tumor microenvironment in melanoma. *P Natl Acad Sci USA*. 2016 Nov 11. 10.1113/JP273570. [Epub ahead of print]
- Applebaum MA, Jha AR, Kao C, Hernandez KM, DeWane G, Salwen HR, et al. Integrative genomics reveals hypoxia inducible genes that are associated with a poor prognosis in neuroblastoma patients. *Oncotarget*. 2016 Oct 17. doi:10.18632/oncotarget.12713
- Eckert MA, Pan S, Hernandez KM, Loth RM, Andrade J, Volchenboum S, et al. Genomics of Ovarian Cancer Progression Reveals Diverse Metastatic Trajectories Including Intraepithelial Metastasis to the Fallopian Tube. *Cancer Discov*. 2016 Oct 7. doi:10.1158/2159-8290.CD-16-0607
- Shah S, Ward JE, Bao R, Hall CR, Brockstein BE, Luke JJ. Clinical response of a patient to anti-PD-1 immunotherapy and the immune landscape of testicular germ cell tumors. *Cancer Immunol Res*. 2016 Sep 16. doi:10.1158/2326-6066.CIR-16-0087. [Epub ahead of print]
- Khalyfa A, Zhang C, Khalyfa AA, Foster GE, Beaudin AE, Andrade J, et al. Effect on Intermittent Hypoxia on Plasma Exosomal Micro RNA Signature and Endothelial Function in Healthy Adults. *Sleep*. 2016 Sep 9. pii: sp-00145-16. [Epub ahead of print]
- Almendros I, Khalyfa A, Trzepizur W, Gileles-Hillel A, Huang L, Akbarpour M, et al. Tumor Cell Malignant Properties Are Enhanced by Circulating Exosomes in Sleep Apnea. *Chest*. 2016 Aug 25. doi:10.1016/j.chest.2016.08.1438. [Epub ahead of print]
- Khalyfa A, Almendros I, Gileles-Hillel A, Akbarpour M, Trzepizur W, Mokhlesi Babak, et al. Circulating exosomes potentiate tumor malignant properties in a mouse model of chronic sleep fragmentation. *Oncotarget*. 2016 Jul 13.
- Gunasekharan VK, Li Y, Andrade J, Laimins LA. Post-Transcriptional Regulation of KLF4 by High-Risk Human Papillomaviruses Is Necessary for the Differentiation-Dependent Viral Life Cycle. *PLoS Pathogens*. 2016 Jul 7. doi:10.1371/journal.ppat.1005747
- McConnell SC, Hernandez KM, Wcisel DJ, Kettleborough RN, Stemple DL, Yoder JA, et al. Alternative haplotypes of antigen processing genes in zebrafish diverged early in vertebrate evolution. *P Natl Acad Sci USA*. 2016 Jun 23. doi:10.1073/pnas.1607602113
- Kuo HC, Hsieh KS, Ming-Huey Guo M, Weng KP, Ger LP, Chan WC, et al. Next-generation sequencing identifies micro-RNA-based biomarker panel for Kawasaki disease. *J Allergy Clin Immunol*. 2016 Jun 18. doi:10.1016/j.jaci.2016.04.050. [Epub ahead of print]
- **Publication Highlight**: The 1001 Genomes Consortium. 1,135 Genomes Reveal the Global Pattern of Polymorphism in Arabidopsis thaliana. *Cell*. 2016 Jun 9. doi:10.1016/j.cell.2016.05.063
- Sweis R, Spranger S, Bao R, Paner GP, Stadler WM, Steinberg G, et al. Molecular Drivers of the Non-T Cell-Inflamed Tumor Microenvironment in Urothelial Bladder Cancer. *Cancer Immunol Res*. 2016 May 17. doi: 10.1158/2326-6066.CIR-15-0274 [Epub ahead of print]
- Kuo H, Li S, Guo M, Huang Y, Yu H, Huang F, et al. Genome-Wide Association Study Identifies Novel Susceptibility Genes Associated with Coronary Artery Aneurysm Formation in Kawasaki Disease. *PLoS One*. 2016 May 12. doi:10.1371/journal.pone.0154943
- Khalyfa A, Kheirandish-Gozal L, Khalyfa AA, Philby MF, Alonso-Alvarez ML, Mohammadi M, et al. Circulating Plasma Extracellular Microvesicle miRNA Cargo and Endothelial Dysfunction in OSA Children. *Am J Respir Crit Care Med*. 2016 May 10. doi:10.1164/rccm.201602-0323OC [Epub ahead of print]
- Khalyfa A, Khalyfa AA, Akbarpour M, Connes P, Romana M, Lapping-Carr G, et al. Extracellular microvesicle microRNAs in children with sickle cell anaemia with divergent clinical phenotypes. *Brit J Haematol*. 2016 May 10. doi: 10.1111/bjh.14104. [Epub ahead of print]

2014

- Xie T, Vigil J, MacCracken E, Gasparaitis A, Young J, Kang W, et al. Low-frequency stimulation of STN-DBS reduces aspiration and freezing of gait in patients with PD. *Neurology*. 2015 Jan 27;84(4):415-20. doi: 10.1212/WNL.0000000000001184. Epub 2014 Dec 24.
- Regnier SM, Kirkley AG, Ye H, El-Hashani E, Zhang X, Neel BA, et al. Dietary Exposure to the Endocrine Disruptor Tolylfluanid Promotes Global Metabolic Dysfunction in Male Mice. *Endocrinology*. 2015 Mar;156(3):896-910. doi: 10.1210/en.2014-1668. Epub 2014 Dec 23.
- Malcom JW, Hernandez KM, Likos R, Wayne T, Leibold MA, Juenger TE. Extensive cross-environment fitness variation lies along few axes of genetic variation in the model alga, Chlamydomonas reinhardtii. *New Phytol*. 2015 Jan;205(2):841-51. doi: 10.1111/nph.13063. Epub 2014 Sep 29.
- Lowry DB, Hernandez K, Taylor SH, Meyer E, Logan TL, Barry KW, et al. The genetics of divergence and reproductive isolation between ecotypes of Panicum hallii. *New Phytol*. 2015 Jan;205(1):402-14. doi: 10.1111/nph.13027. Epub 2014 Sep 23.
- Chen C, Zhang C, Cheng L, Reilly JL, Bishop JR, Sweeney JA, et al. Correlation between DNA methylation and gene expression in the brains of patients with bipolar disorder and schizophrenia. *Bipolar Disord*. 2014 Dec;16(8):790-9. doi: 10.1111/bdi.12255. Epub 2014 Sep 22.
- Bao R, Huang L, Andrade J, Tan W, Kibbe WA, Jiang H, et al. Review of current methods, applications, and data management for the bioinformatics analysis of whole exome sequencing. *Cancer Informatics*. 2014 Sep 21;13(Suppl 2):67-82. doi: 10.4137/CIN.S13779. eCollection 2014.
- Chen B, Moore TV, Li Z, Sperling AI, Zhang C, Andrade J, et al. Gata5 Deficiency Causes Airway Constrictor Hyperresponsiveness in Mice. *Am J Resp Cell Mol*. 2014 Apr;50(4):787-95. doi: 10.1165/rcmb.2013-0294OC.
- Fahrenbach J, Andrade J, McNally EM. The CO-Regulation Database (CORD): a tool to identify coordinately expressed genes. *PLoS One*. 2014 Mar 5;9(3):e90408. doi: 10.1371/journal.pone.0090408. eCollection 2014.
- Widau RC, Parekh A, Ranck MC, Golden DW, Kumar KA, Sood RF, et al. The RIG-I like receptor LGP2 protects tumor cells from ionizing radiation. *P Natl Acad Sci*. 2014 Jan 28;111(4):E484-91. doi: 10.1073/pnas.1323253111. Epub 2014 Jan 13.

Bioinformatics - Data integration

- Consider both phenotype and genotype data
- How will the clinical data be collected?
- Who is integrating these data into the analysis?
- CRI can get the clinical data **and** integrate it with the genomics information - this may require engaging the CRDW