RNA-Seq libraries in three hours. The percent mRNA mapped for enriched samples was reduced from 12% to 3% when compared to the non-enriched RNA sample. miRBase analysis of the enriched sRNA samples increased the percentage of sRNA mapped reads from 54% to 74% with a two-fold increase in the number of known and unique miRs.

51

Quality Control and Pre-Qualification of NGS Libraries Made from Clinical Samples

John Langmore

Rubicon Genomics

We have compared sequencing metrics from different types of clinical samples and different methods of making NGS libraries, for purposes of quality control of the samples and sample preps. We have performed the metrics using the HiSeq and MiSeq, however the same QC metrics could be measured on other platforms. By choosing metrics that can be measured from small amounts of data (e.g., 300,000 reads), we can measure the quality of the clinical samples and NGS libraries in a highly multiplexed format, before spending considerable time and money for downstream processes such as sequence enrichment and NGS. More predictive than measurement of insert size and concentration, these metrics predict whether the amount and quality of genomic DNA, as well as the sample preparation method is sufficiently efficient to generate high coverage from deep sequencing. Sequencing results from sonicated human and bacterial DNA, as well as human maternal plasma show that different sample prep methods yield libraries of very different diversity, uniformity of coverage and background. Library quality can also vary considerably in different lots of reagents and also different number of amplification cycles.

By optimizing the efficiency of making genomic DNA into sequencing libraries fewer reads are necessary to achieve reproducible, high quality results, especially from limiting amounts of plasma, FFPE tissue, chromatin immune precipitates, and single cells.

52

Towards Clinical Grade Genomes with Joint Bayesian Variant Identification

Francisco De La Vega, Brian Hilbush Len Trigg, Richard Littin, John G. Cleary

Real Time Genomics, Inc.

The precipitous fall in the cost of sequencing spurred by innovations in high-throughout sequencing (HTS) is bringing the use of genome sequencing closer to the clinic. An important question yet to be answered is whether current HTS protocols provide data that meets clinical standards of quality. While false positives (FP) can be evaluated experimentally, false negatives are more difficult to assess due to the lack of an established gold standard. Sequencing of family pedigrees already enormously simplifies the identification of highly penetrant disease genes. However, joint analysis of family members raw data could also provide a significant boost in variant calling accuracy because related individuals share haplotype blocks. Here we present our Joint Bayesian Calling (JBC) method for pedigrees and show it reduces false positives & negatives and improves accuracy of identified variants in trios and larger pedigrees. Our approach reduces Mendelian errors in trios to 0.1% compared to 2% in singleton calling, and improves specificity of *de novo* variant identification by reducing FP 50%. We demonstrate JBC scalability to large pedigrees by analyzing sequencing data of a large CEPH pedigree where the genomes of 17 individuals where sequenced to ~50X. As more pedigree members are added accuracy improves and we are also capable of imputing genotypes of missing subjects. Our approach is able to inform trade-offs between depth of coverage and number of family members for research study planning, and coupled with a proprietary fast read mapping algorithm is able to analyze a full depth WGS trio in less than a day (hours for exomes) in commodity hardware. We believe these advances will be crucial for the adoption of genomes & exomes in clinical settings.

53

Which RNA-Seq Processing Algorithm Should I Pick? A Comparison of RNA-Seq Pipelines Based on Experimental Design, Sample Properties and Sequencing Technology

Stuart Levine, Zachary Banks, Vincent Butty

Massachusetts Institute of Technology, Cambridge, MA

In the past few years, the dramatic increase in sequencing throughput has been paralleled by a proliferation of bioinformatic tools designed to analyze and quantitate next generation sequencing data. In the case of RNA-Seq, a number of analysis packages have been developed, but whether some might be better suited to the analysis of samples within different ranges of read coverage or RNA quality has not been extensively studied. Here, we evaluated head-to-head eight different packages on single-end- and paired-end-RNA-Seq data at different read coverages. Gene expression and alternative splicing were quantitated and compared across the different methods using standard accuracy and precision metrics. We demonstrate critical differences in how the different algorithms are able to handle different read lengths and read coverage levels and how these changes impact gene expression and isoform usage values.

FLOW CYTOMETRY

54

Different Sorts for Different Folks: The Importance of Technological Diversity in a Cell Sorting Facility

Monica DeLay[1], A. Nicole White[2], Edith Janssen[2], George Babcock[3], Christopher A. Worth[4], Sherry Thornton[2]

[1]*Cincinnati Children's Hospital,* [2]*Research Flow Cytometry Core, Division of Rheumatology and Division of Cellular and Molecular Immunology, Cincinnati Children's Hospital, Cincinnati, OH,* [3]*Flow Cytometry Lab, Shriner's Hospitals for Children-Cincinnati, Cincinnati, OH,* [4]*Stem Cell Institute at James Graham Brown Cancer Center, University of Louisville, Louisville, KY*

Flow cytometry core facilities have the challenge of offering appropriate technologies to diverse client needs. Therefore, there is no one-size-fits-all technology that can handle any and all cell types or applications. For cell sorting, this becomes even more of a challenge when a variety of cells need to be purified for an even wider variety of downstream applications. There are two types of technologies currently available for droplet based cell sorting; cuvette and jet-in-air. While there are advantages and disadvantages to both technologies, no cell sorter design can fit all needs, especially when the cell type to be sorted is sensitive to manipulations.

There have been reports (published and anecdotal) of phenotypic and functional changes to dendritic cells after isolation using different techniques. In our core facility, DC populations that were sorted on the FACSAria II showed an increase in cell death and were found to be nonfunctional in an assay testing their ability to process cell associated antigen and stimulate proliferation of T cells. In contrast, cells sorted on a FACSVantage, MoFlo Legacy or MoFlo XDP were able to function in the same *in vitro* assay. We attribute this difference to the differences in fluid dynamics through the sample path that may damage the cells and decrease their ability to function. Here we present data comparing the functionality of DCs with each system. We conclude that it is best to have a diversity of technologies in a cell sorting facility to better meet the needs of all clients.

55

Evaluation of Aerosol Containment of BD FACSAria Cell Sorter in a Class II Bigneat "LABS" Robotics Enclosure

Lora Barsky[1], David Stanley[1], George Salazar[2], Stephen Pratt[3]

[1]*University of Southern California,* [2]*Biosero LLC,* [3]*BigNeat Ltd*

Steam in air cell sorting of unfixed human cell specimen poses a biosafety hazard to the occupants of the core facility as described by the International Society for Analytical Cytology Biosafety Standard for Sorting of Unfixed Cells. Assessment of aerosol containment with an Aerosol Management System (AMS) is a necessary validation for steam in air cell sorters commonly performed with the Glo Germ protocol as published by Perfetto etal. BD FACSAria cell sorters may be placed in the Baker BioProtect IV line of Biological Safety Cabinet with integrated AMS, which eliminates the purpose of the Whisper AMS peripheral device and provides both operator and sample sterility. This study evaluates a Bigneat LABS custom robotics enclosure designed to contain a BD FACSAria cell sorter as an alternative to Baker BioProtect in terms of its ability to contain and evacuate aerosols generated by the instrument. This enclosure provides Class II, type B containment offering operator protection and sample sterility to ISO 5 cleanroom rating. Aerosol containment was assessed using the KI discus test as described in the BS EN 12649:2000 standard to validate the integrity of the air curtain at the time of production and then again after site installation with BD FACSAria enclosed. We questioned if the Whisper AMS was necessary to contain aerosols within this enclosure and have assessed the AMS utility under stream failure using the GloGerm assay as described by Perfetto et al. The Class II Bigneat "LABS" Robotics Enclosure passed all testing performed to access air curtain integrity and containment of aerosols generated by the FACSAria cell sorter.

FRAGMENT ANALYSIS

56

The Advantages of Micro-Cal™ Auto-iTC200 for Secondary Screening in a Fragment-based Drug Discovery Campaign

Mark Arsenault[1], Natalia Markova[2] and Ronan O'Brien[2], Jessica Martinsson[3], Kenth Halberg[3], Martin Andersson[3], and Martin Welin[3]

[1]*GE Healthcare Bio-Sciences AB, Uppsala, Sweden,* [2]*GE Healthcare, Northampton, MA and* [3]*Sprint Bioscience AB, Stockholm*

Background: Recent developments in ITC instrumentation, namely MicroCal iTC200 and MicroCal Auto-iTC200 have led to an increase in the throughput and decrease in the protein consumption of the technique. In addition, there have been recent methodological Advancements 1 that have extended the affinity range that ITC can measure into the mM range. The combination of all these factors has made the technique ideal for fragment-based drug discovery (FBDD) campaigns. Method: This work outlines the role of MicroCal Auto-iTC200, in the fragment based drug discovery program of Sprint Bioscience, to identify and optimize potential drug candidates that will inhibit the activity of Vps34. This class phosphatidylinositol3-kinase is central to autophagy and has been shown to play an important role in resistance to cancer drugs2, 3. As such it has been identified as a target for therapeutic intervention. ITC is a generic assay without the need for assay development and as such the affinity of all 50 compounds was measured in less than three days after receiving the purified protein. Conclusion: This approach was fast and proved very successful for identifying fragments that co crystallized with the target protein. Of the14 compounds chosen, based on the ITC data, 12 formed crystals that could be used in the optimization process.

GENETIC VARIATION

57

Uniparental Disomy is Associated with Embryonal Rhabdomyosarcoma in Costello Syndrome and Nonsyndromic Patients: From Core-Side to Bed-Side and Back Again, Impact of Core Laboratories on Biomedical Research

Katia Sol-Church[1], Katherine M. Robbins [2], Deborah L. Stabley[2], Jennifer Holbrook[2], Karen W. Gripp[2]

[1]*Nemours Center for Pediatric Research, A.I. duPont Hospital for Children Wilmington, DE,* [2]*Division of Medical Genetics A.I. duPont Hospital for Children Wilmington, DE*

Embryonal rhabdomyosarcoma (ERMS) is the most prevalent pediatric soft tissue tumor and is characterized at the molecular level by loss of heterozygosity (LOH) at