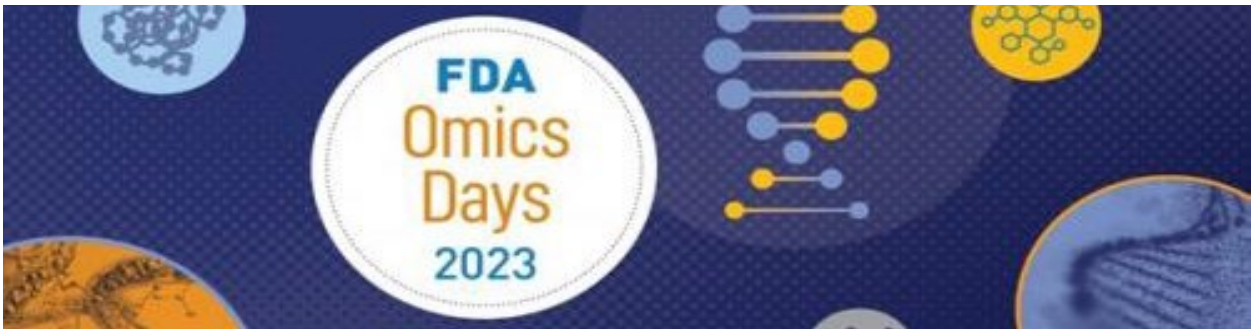

Abstract Booklet

Odd numbered posters on Tuesday, October 10, 2023, and even numbered posters on Wednesday, October 11, 2023. Poster numbers with an asterisk will not be presented in-person.

Contents

FDA OneHealth

25: FDA and NIST collaboration to evaluate assays and control materials for characterizing animal biotechnology products generated by genome editing

Patricia Kiesler, NIST; Natalia Kolmakova, NIST; Sierra Miller, NIST; Ayah Shevchenko, NIST; Stella S. Lee, FDA/CVM; Alexis L. Norris, FDA/CVM; Mayumi F. Miller, FDA/CVM; Adam L. Moyer, FDA/CVM; Samantha Maragh, NIST
maria.kiesler@nist.gov

Abstract

Genome editing technology has revolutionized the ability to make targeted changes to an animal's genome (intentional genomic alterations or IGAs), offering exciting promise for the development of animal biotechnology products that address animal and public health needs. Characterization of these IGAs is an important part of the regulatory process to ensure that the intended edit is made to the animal and to identify any unintended changes. However, there are currently no validated measurements and standards for characterizing unintended genomic alterations in animals. To address these needs, FDA CVM has established a collaboration with the U.S. National Institute of Standards and Technology (NIST) that will generate resources including standardized measurements for characterizing both intended and unintended alterations in animal biotechnology products resulting from genome editing. These resources will provide animal biotechnology product developers and FDA regulators with example characterization approaches that they could use as part of the development and regulatory process for IGAs in animals as well as for validating methods, materials and/or data. Here, we present preliminary outcomes of this NIST-FDA CVM collaboration. NIST qualified a commercially available pig cell line and its DNA as potential control materials. The cell line was characterized for genomic stability prior to editing, as well as for sequence before and after genome editing. Four CRISPR/Cas9 editing assays, including two newly developed by NIST, were evaluated using purified pig DNA and the pig cell line. Off-target sites identified from three in silico predictors and an existing biochemical assay that detects genomic positions cleaved by genome editing reagents, called CHANGE-seq, were also compared. A subset of the off-target sites identified by in silico predictors were also identified by CHANGE-seq and further analyzed for evidence of off-target editing in the edited pig cells. Additionally, the CHANGE-seq assay was evaluated for reproducibility and performed similarly on pig genomic DNA as compared to human genomic DNA. Experimental design, protocols, datasets, and measurements that NIST generated will be published and made accessible to animal biotechnology product developers and the public. Future work will focus on similar qualifications of potential bovine control materials and genome editing assays.

Novel aspect

FDA CVM is working collaboratively on a project, headed by the U.S. National Institute of Standards and Technology, that aims to provide new resources for researchers and companies creating innovative animal biotechnology products by using genome editing to alter the genome of animals.

27: RIPS: routine intuitive pathogen surveillance with whole genome sequence data

Tim Muruvanda, FDA/CFSAN; Arthur Pightling; FDA/CFSAN; James Pettengill, FDA/CFSAN; Hugh Rand, FDA/CFSAN
Tim.Muruvanda@fda.hhs.gov

Abstract

Background: The large Pathogen Detection (PD) database maintained by the NCBI includes more than one million bacterial genomes collected under various whole genome sequence (WGS) surveillance programs. The PD database is monitored for emerging outbreaks by finding matches between new and previously submitted data originating from foods, environments, and patients. Here, we present an accessible tool to perform routine WGS surveillance called RIPS–Routine Intuitive Pathogen Surveillance.

Purpose: RIPS makes routine surveillance of WGS data accessible to a wide range of end-users, democratizing the detection of emerging outbreaks while they are small and reducing the numbers of illnesses. It provides a means for uniform and objective application of criteria that define emerging outbreaks, as well as measuring the performance of the criteria which can be adjusted.

Methods: RIPS was written with the Shiny package in Posit. It downloads Rapid Reports data from the PD database. Results of a preliminary analysis that compares newly submitted sequence data to all other sequences in the database are further filtered using metrics, such as the numbers of sequences for bacteria collected from patients (clinical isolates) within a stated timeframe and the genomic distances between clinical and nonclinical isolates. Information is displayed in an easy-to-understand format, alongside additional information such as serotypes and geography.

Results: Comparisons with the other methodologies indicates RIPS is fast and accurate. An example surveillance run for Salmonella enterica for July 21, 2022 finds six clusters of interest (at least seven clinical isolates submitted in the prior 90 days with no more than seven allele differences from a nonclinical isolate). One cluster PDS000113505 includes a sequence from S.enterica isolated from a sample that was collected by FDA and seven clinical isolates that are part of an outbreak (now closed). RIPS allows users to quickly review these clusters in ten minutes and determine which isolates warrant further attention.

Conclusion: By simplifying surveillance, it encourages the use of WGS for pathogen surveillance, in foodborne pathogens and in other areas furthering FDA's One Health mission.

Novel aspect

DNA sequence data from foodborne bacteria are monitored for emerging outbreaks. However, these activities are largely restricted to those with computational skills. We present, an accessible tool for routine surveillance (the Routine Intuitive Pathogen Surveillance [RIPS] app), and provide an example of its use for detecting a Salmonella enterica outbreak.

Genomics, Transcriptomics, Metagenomics

05*: Evaluating Renal Pathology in Post-COVID-19 Human Autopsy Tissues

Elysia Masters, FDA/NCTR; Mallikarjun Bidarimath, FDA/NCTR; Jennifer Hanks, TPA Inc.; Vaunita Parihar, Emory University; Amy Inselman, FDA/NCTR; Jessica Hawes, FDA/NCTR; Laura Schnackenberg, FDA/NCT; Kelly Mercer, FDA/NCTR

Elysia.Masters@fda.hhs.gov

Abstract

Coronavirus disease 2019 (COVID-19), caused by SARS-CoV-2, emerged in late 2019 and continues to circulate globally. Acute kidney injury has been reported in as high as 57% of patients hospitalized with COVID-19. It is currently unknown if the prolonged renal effects following COVID-19 are due to systemic immune responses, direct kidney viral infection, drug toxicity, glomeruli structural changes, or a combination thereof. Thus, the objective of this work was to evaluate kidney tissue samples obtained from autopsies of individuals who died with COVID-19 to explore possible mechanisms for COVID-19-associated renal nephropathy. Formalin-fixed paraffin embedded tissues were evaluated using histopathology evaluations, immunohistochemistry (IHC), immunofluorescence (IF), spatial glycan analysis by matrix-assisted laser desorption/ionization imaging mass spectrometry (MALDI-IMS), and spatial transcriptomics. Postmortem kidney autopsy samples were collected from COVID-19 positive cases (N=9) (Emory Research University Hospital, Atlanta, GA). Histopathology evaluation of COVID-19 autopsy tissues revealed regions of inflammatory infiltrates, glomerulosclerosis and interstitial fibrosis. IHC studies identified CD8+ T cells, CD4+ T cells, CD163+ monocytes, and M2 macrophages within regions of kidney injury. Quantification of immune cell staining showed a significant increase of CD8+ T cells in the renal medulla versus the cortex. Select regions of renal tubules also stained positive for SARS-CoV-2, which was co-stained with hACE2, the receptor for SARS-CoV-2 entry. MALDI-IMS analysis identified glycans associated with regions of damaged medulla and immune cell infiltration. Finally, preliminary spatial transcriptomics analysis identified gene expression signatures associated with renal cortex and medulla regions.Taken together, immune cell infiltrate, renal structural changes and possibly direct viral infection were all identified in COVID-19 kidney autopsy samples, suggesting a combination of mechanisms that may contribute to COVID-19-associated AKI. Continued analysis of spatial transcriptomics data will identify gene expression differences between COVID-19+ and healthy control tissues.

Novel aspect

In this work we combine traditional histologic methods for tissue evaluation including histopathology, immunohistochemistry and immunofluorescent staining, with novel -omics methodologies. This project applies both matrix-assisted laser desorption/ionization imaging mass spectrometry (MALDI-IMS) and spatial transcriptomics to explore and discover novel mechanisms for COVID-19-associated renal nephropathy.

12: Comparison of target amplicon sequencing using the MiSeq and GridION next generation sequencing platforms for detection of foodborne pathogens

Isha Patel, FDA/CFSAN; Mark Mammel, FDA/CFSAN; Jayanthi Gangiredla, FDA/CFSAN
Isha.Patel@fda.hhs.gov

Abstract

Introduction: The use of Whole Genome Sequencing for pathogen detection has increased the accuracy and reduced the time for traceback and source attribution in foodborne outbreaks. However, low level contaminants may go undetected due to challenges in isolating them from food matrices or due to a high background of other microbial flora. We have previously shown using the Illumina platform that a custom targeted amplicon sequencing (TAS) panel offers increased sensitivity and specificity for pathogen detection in "spike-in" experiments. The turnaround time from isolation of DNA to obtaining data is two days. Nanopore GridION platform offers a sequencing approach that enables direct near real-time sequencing thus saving time to get results.

Methods: We used NIST Microbial Pathogen DNA Standards (RM 8376), which consists of 20 individual gDNA components from 19 bacterial strains and 1 human reference genome. Specifically, we used Listeria monocytogenes ATCC 19115, Escherichia coli BAA 2309 and Salmonella enterica ATCC 12324 to compare the limit of detection of the pathogens using three serial dilutions (100 ng, 10 ng and 1 ng). Respective library preparation methods were employed for the MiSeq and GridION sequencers. GalaxyTrakr and BLAST matching of the amplicons was used for data analysis.

Results: Preliminary results show that both sequencing platforms using their respective analysis pipelines detect specific pathogens and their associated virulence genes at similar levels. However, the GridION offers a near real-time sequencing advantage resulting in obtaining data in less than an hour.

Conclusion: This study shows proof of concept that targeted sequencing to detect pathogens yields similar results with both GridION and MiSeq; however, GridION is faster, providing data in real time.

Novel aspect

The objective of this work is to provide a rapid and sensitive method using targeted detection of low-level pathogen contamination in complex samples and thereby positively impacting the utility of metagenomics as a rapid screening method for pathogen detection.

14: Characterization of Soil and Lettuce Resistomes from Harvest Through Storage in Modified Atmosphere Packaging

Taylor K.S. Richter, FDA/CFSAN; Mark K. Mammel, FDA/CFSAN; Ivan Simko, USDA/ARS; Maria T Brandl, USDA/ARS; Susan R. Leonard, FDA/CFSAN
Taylor.Richter@fda.hhs.gov

Abstract

Background: It has been demonstrated that soil microbial communities contribute to the lettuce phyllosphere microbiome. Leafy vegetables carrying bacteria that harbor antimicrobial resistance genes (ARGs) may provide a pathway for those genes to enter the human microbiome and metal and biocide resistance genes (MRGs and BRGs) may contribute to co-selection of ARGs depending on the selection pressure encountered.

Purpose: In this work, the distribution of resistomes associated with lettuce from field through cold storage in modified atmosphere packaging (MAP) was investigated. ARGs, MRGs, and BRGs in soil and lettuce phyllosphere microbiomes were profiled and compared to determine possible transfer between soil and lettuce as well as persistence during lettuce processing and storage.

Methodology: Shotgun metagenomic sequencing was performed on a total of 225 samples consisting of surface soil, harvested lettuce heads, processed lettuce (cut and washed), and processed lettuce cold-stored in MAP for five different harvests in Salinas, California. Sequencing was performed on an Illumina NextSeq platform generating paired-end 150 bp reads, and classification of resistance determinants in the sequence datasets was accomplished using MEGARes 2.0.

Results: Overall, a higher number of resistance genes per million reads was found in processed lettuce after storage (9.8) compared to processed lettuce before storage (0.74) (Wilcoxon Rank Sum, $P<0.001$). MRGs were prevalent in both soil and lettuce, with copper resistance particularly high in lettuce. BRGs were most frequent in processed lettuce before storage. In the five separate harvests, between three and 59 different resistance gene alleles were identified in both soil and at least one lettuce sample. Omitting genes conferring both drug and biocide resistance, and including all samples, the greatest percentage of ARGs belonged to the beta-lactam class (30%) followed by aminoglycosides (6.7%), and specific allele sequences from both classes were observed in common between soil and lettuce.

Conclusion: These results provide insight into the transfer of antimicrobial resistance genes from soil to the lettuce phyllosphere and ready-to-eat packaged lettuce, thus their spread

Novel aspect

This work examines the transfer of resistance genes from the soil microbiome to the lettuce microbiome and its persistence during lettuce processing and storage, providing insight into a possible route for antimicrobial resistance genes to enter into the human food chain.

15: Whole Genome Metagenomics as a Tool for Probiotics Analysis

Carmen Tartera, FDA/CFSAN; Jayanthi Gangiredla, FDA/CFSAN; Mark Mammel, FDA/CFSAN; Isha Patel, FDA/CFSAN
Carmen.Tartera@fda.hhs.gov

Abstract

Background: Many dietary products available to the consumer, such as dietary supplements contain live microorganisms purported to confer a human health benefit. There are reports of label discrepancies of commercial probiotics where species are often misclassified or absent and contamination by microorganisms not listed on the label.

Methodology: In this study, DNA was purified from 123 probiotic products and metagenomic sequencing was performed in an Illumina MiSeq platform. The generated sequence was used to identify microbial constituents with unique species-specific signatures, based on a novel in-house K-mer database. In parallel, using a culture-dependent approach, the microbial contents of the products were grown for single colony isolation followed by WGS to create a genome sequence database of beneficial microbes. Additionally, Resistance Gene Identifier and Virulence Finder bioinformatic tools were used to identify the presence of antibiotic resistance and virulence factors genes.

Results: Results showed that from the 123 products evaluated, 80 products (60%) were compliant, however 43 products (35%) contain non-listed species, or species that were absent. Antibiotic resistance genes were identified in four products and virulence factors in two products. A challenge in the metagenomic approach is the detection of microbes present in low numbers (constituents, contaminants, and pathogens) in the presence of high numbers of intentionally added species like Lactobacillus and Bifidobacterium. To circumvent this problem two approaches were evaluated: the use of specific phages and/or purified bacteriophage-lysins to reduce the product's indigenous microorganisms to improve the detection of low-level microbial constituents and using target-amplicon sequencing. Results showed that both approaches increased the sensitivity of detection of E. coli that were spiked at low levels.

Conclusion: These studies will lead to a better understanding of the quality of live microbial supplements available to the consumers and provide an analytical pathway to detect harmful pathogens present in low amounts.

Novel aspect

The goal of this study is to provide a better understanding of the quality of probiotics products, the accuracy of their label, and the ability to identify potential contaminants and low-level constituents using Whole Genome Sequencing (WGS) metagenomics.

16: Utility of miROmics for Identification of Circulating Pharmacodynamic Biomarkers of IFNβ-1a Biologics

Mai Mehanna, FDA/CDER; Lakshmi Manasa Chekka, FDA/CDER; Deepti Samarth, FDA/CDER; Erica Decker, FDA/CDER; Esraa Mohamed, FDA/CDER; Yan Guo, FDA/CDER; James L Weaver, FDA/CDER; Sarah Schrieber, FDA/CDER; Jeffry Florian, FDA/CDER; Yow-Ming Wang, FDA/CDER; David Strauss, FDA/CDER; Paula Hyland, FDA/CDER
Mai.Mehanna@fda.hhs.gov

Abstract

Background: The U.S. Food and Drug Administration is conducting research to identify pharmacodynamic (PD) biomarkers to support the demonstration of biosimilarity. These PD biomarkers can streamline development programs by negating the need for comparative clinical studies with efficacy endpoint(s).

Purpose: To evaluate the utility of miRNA profiling (miRomics) and develop an analytical framework for identifying potential circulating PD biomarkers of IFNβ-1a and pegIFNβ-1a products.

Methodology: A pilot study was conducted using plasma samples from 36 healthy subjects from a placebo-controlled randomized single dose clinical study with IFNβ-1a and pegIFNβ-1a. Using miRNA-sequencing, we measured miRNAs at baseline/pre-treatment in all subjects, at 9 timepoints over 6 days in the IFNβ-1a group (n=11 [30µg]), and at 11 timepoints over 13 days in the pegIFNβ-1a group (n=11[125µg]) and placebo-specific groups (n=6 each). We identified 108 mature miRNAs (with a minimum of 10 read counts in at least 50% of samples). We conducted linear-mixed effect models regressing the normalized count changes from baseline with treatment*time interaction. miRNAs with false discovery rate (FDR)-corrected p-values<0.1 were considered differentially expressed. Analysis was conducted in R (v4.1.2). DIANA-miRPath v3.0 was used for functional characterization of miRNA biomarkers.

Results: We identified 11 and 13 differentially expressed miRNAs over treatment and time by IFNβ-1a and pegIFNβ-1a, respectively, compared to placebo. hsa-miR-223-3p and hsa-miR-21-5p were common for both products. Importantly, hsa-miR-223-3p regulates Mx1 and STAT1 which are proposed individual candidate PD biomarkers for IFNβ-1a and pegIFNβ-1a and are also involved in IFNβ-1a signaling. Functional analysis of top miRNAs identified 24 overlapping pathways for both products including Hepatitis B and Hippo signaling.

Conclusion: Using miRomics, we identified two plasma miRNAs as potential PD biomarkers of IFN-β1a biologics for further investigation to support biosimilar development programs.

Novel aspect

In this study, we used miRNA profiling to identify potential pharmacoynamic biomarkers of IFNβ-1a biologics to support the demonstration of biosimilarity.

18: PrecisionFDA Truth Challenge V2: Using Crowdsourcing to Benchmark Variant Calling Innovation

Ezekiel Maier, Booz Allen Hamilton; Elaine Johanson, FDA/ODT; Nathan Olson, NIST; Justin Wagner, NIST; Jennifer McDaniel, NIST; Justin Zook, NIST; Samuel Westreich, DNAnexus; Omar Serang, DNAnexus; Anish Prasanna, Booz Allen Hamilton
Adrienne.Phifer@fda.hhs.gov

Abstract

Background: In 2016, the first Genome In A Bottle (GIAB)-precisionFDA Truth Challenge occurred and GIAD has since expanded benchmarking methods, including covering more challenging regions of the genome. The Truth Challenge V2 built on the Truth Challenge to test novel expanded variant calling benchmarks to better discern mutations from false positives and false negatives.

Purpose: The Truth Challenge V2 was conducted from May to June 2020, to assess state-of-the-art variant calling in difficult-to-map regions and the Major Histocompatibility Complex.

Methodology: Participants generated variant calls as Variant Call Format (VCF) files for sequencing data HG002, HG003, and HG004 originally given as FASTQ files. Sequencing data were provided from Illumina, Pacific Biosciences, and Oxford Nanopore Technologies, at 35X, and 50X coverage, respectively. The variant calls were generated against the GRCh38 version of the human reference genome. Submissions were evaluated based on the harmonic mean of parents' F1 scores for combined SNVs and INDELSs.

Results: From the 64 submissions, top performers came from Sentieon, Roche Sequencing Solutions, The Genomics Team in Google Health, DRAGEN, Seven Bridges Genomics, The UCSC CGL and Google Health, and Wang Genomics Lab. The top performing submissions combined all 3 technologies. The performance of each submission varied across stratifications, in which the best-performing multi-technology call sets had similar performances overall. 90% of submissions for long-read-only used deep-learning-based methods. The short-read submissions with the best performance used statistical variant-calling algorithms with graph reference. The addition of deep learning and machine learning have advanced variant calling by enabling faster adoption of new sequencing technologies.

Conclusion: Several submissions developed and applied new approaches to integrate multiple technologies and combine data to outperform individual technology. With these submissions outperforming the previous Truth Challenge submissions, it indicates that improvements in variant calling were made. This challenge ignited the development and widespread sharing of various new bioinformatics methods for multiple technologies.

Novel aspect

The Truth Challenge V2 was conducted to better distinguish mutations from false positives and false negatives with a novel, expanded variant calling benchmark. Achievements from the Challenge included new advances in variant calling methods and innovative method combinations.

20: Assessment of plasmids for relating the 2020 Salmonella enterica serovar Newport onion outbreak to farms implicated by the outbreak investigation

Seth Commichaux, FDA/CFSAN; Hugh Rand, FDA/CFSAN; Kiran Javkar, University of Maryland/Department of Computer Science; Erin K. Molloy, University of Maryland/Department of Computer Science; James B. Pettengill, FDA/CFSAN; Arthur Pightling, FDA/CFSAN; Maria Hoffmann, FDA/CFSAN; Mihai Pop, University of Maryland/Department of Computer Science; Victor Jayeola, FDA/CFSAN; Steven Foley, FDA/NCTR; Yan Luo, FDA/CFSAN
Seth.Commichaux@fda.hhs.gov

Abstract

The Salmonella Newport red onion outbreak of 2020 was the largest foodborne outbreak of Salmonella in over a decade. The epidemiological investigation suggested two farms as the likely source of contamination. However, single nucleotide polymorphism (SNP) analysis of the whole genome sequencing (WGS) data showed that none of the Salmonella isolates collected from the farm regions were linked to the clinical isolates. Here, we explored an alternative method for analyzing the WGS data driven by the hypothesis that if the outbreak strain had come from the farm regions, then the clinical isolates would disproportionately contain plasmids found in isolates from the farm regions due to horizontal transfer. We applied SNP, pangenome, phylogenetic, and statistical analyses to assess if plasmids could relate the outbreak strain to the microbiome of the implicated farms. SNP analysis confirmed that the clinical isolates formed a nearly-clonal clade with evidence for ancestry in California. The clinical clade had a large core genome (4,399 genes) and a large, sparsely distributed accessory genome (2,577 genes, ≥64% on plasmids). Over 20 plasmid types occurred in the clinical clade, more than were found in the literature for Salmonella Newport. A small number of plasmids, 14 from 13 clinical isolates and 17 from 8 farm isolates, were found to be highly similar (>95% identical)—indicating they might be related by horizontal transfer. Phylogenetic analysis was unable to determine the geographic origin, isolation source, or time of transfer of the plasmids, likely due to their promiscuous and transient nature. However, our resampling analysis suggested that observing a similar number and combination of highly similar plasmids in random samples of environmental Salmonella enterica within NCBI Pathogen Detection database was unlikely. Horizontally transferred plasmids supported a connection between clinical isolates and the farms implicated as the source of the outbreak. Our case study suggests that such analyses might add a new dimension to source tracking investigations, but highlights the need for detailed and accurate metadata, more extensive environmental sampling, and a better understanding of plasmid molecular evolution.

Novel aspect

Plasmids were used to relate clinical isolates to the microbiome of the farms implicated as the source of the outbreak.

21: Global transcriptomic analyses of Salmonella enterica serovar Agona reveals the mechanism of survival on low moisture food

Sultana Solaiman, FDA/CFSAN; Ellie Meeks, University of Maryland/JIFSAN; Ian Hines, FDA/CFSAN; Jie Zheng, FDA/CFSAN; Maria Hoffmann, FDA/CFSAN
Sultana.Solaiman@fda.hhs.gov

Abstract

Background: Salmonella enterica serovar Agona caused recurrent multistate outbreaks associated with cereal between 1998 and 2008, highlighting the persistence of Salmonella over time in food processing facilities. Understanding the survivability and physiology of this pathogen in low moisture food (LMF) and low-moisture environments is necessary for developing future intervention strategies.

Purpose: To determine the molecular mechanism of survival of S. Agona strain in LMF.

Methodology: Salmonella Agona strain (CFSAN 000477) linked to the two outbreaks of salmonellosis separated by ten years was selected for the study. 100µl suspension with a concentration of ~1010 cfu/ml was inoculated in 10g of rice cereals. Three replications of inoculated cereals were subjected to desiccation stress (aw ≤ 0.25) for 4h at room temperature (25^{o}C). Inoculated cereal was sampled for 4 timepoints post inoculation. Cells were separated from the food matrix for RNA extraction. RNA sequencing was performed at Center for Food Safety and Applied Nutrition, FDA, College Park, MD using NextSeq 2000 platform. Read counts were generated with Salmon v1.9.0. downstream analysis was done in R and KEGG mapper.

Result: There were 1120 differentially expressed genes (DEGs) of S. Agona in response to desiccation stress (Padj > 0.01, |log2FoldChange| >1 where 647 were downregulated and 473 were upregulated. Functional analysis of downregulated DEGs showed most of the genes were associated with metabolic pathways followed by translation, suggesting slower growth in the population. Among the upregulated genes, kdp and ccm operon and tisB were also directly associated with growth and metabolism. kdp has the regulatory function to reduce replication, ccm involves reducing the consumption of molecular oxygen and tisB is associated persister formation.

Conclusion: The data suggests that Salmonella Agona survive in extremely low moisture food by conserving its energy, lowering metabolism and reducing replication. This bacterium can survive in low moisture environment for long period of time which emphasizes the importance of proper sanitization in low moisture food processing industry.

Novel aspect

SNP analysis showed similar genetic makeup between 2 outbreak strains of S. Agona, suggesting there might be single organism that survived for 10 years in the same facility. This study has identified the major determinants and their expression level that S. Agona employ to survive on cereal/extremely low moisture food.

22: Altered transcriptome and chromatin dynamics following short- and long-term Zika infections

Aaron Scholl, FDA/CBER; Bingjie Li, FDA/CBER; Sandip De, FDA/CBER
Aaron.Scholl@fda.hhs.gov

Abstract

Background: The ability to synthetically reprogram human somatic cells into induced pluripotent stem (iPS) cells provides a near-limitless source of ES-like cells. However, the conventional reprogramming process faces challenges due to the original epigenetic memory of somatic cells, which may have been altered by invading pathogens like viruses. Viruses, such as SARS-CoV-2, can manipulate transcriptional regulation through methods like histone mimicry and cleavage. With rising global temperatures and increased vector (ticks and mosquitoes) populations, vector-borne diseases like flaviviruses are increasing. In the present study, we use Zika as a model flavivirus to study changes in transcriptomes and chromatin dynamics and how this can influence host cells characteristics, an area that has received limited attention.

Purpose: 1. Explore transcriptomic and epigenetic alterations 2. Understand how these changes can modify host cell characteristics 3. Detect host RNA and epigenetic modifications that could serve as indicators for both present and past flavivirus infections.

Methodology: SK-N-SH (neuronal) and U937 (monocyte) cells are sensitive and resistant cells to ZIKV, respectively. These cells were infected with ZIKV-MR766 strain. Following short- and long-term infections, the cells were collected at different time points to investigate the transcriptomic changes. Histones and associated modifications were also studied via western blot (WB) and immunofluorescence assay. Finally, we plan to study epigenetic landscape and three-dimensional genome organizational alterations using CUT&Tag and capture Hi-C assays.

Results: As anticipated, SK-N-SH cells exhibited higher viral load compared to U937 cells, evident from TCID50, qRT-PCR, and WB analyses. This led to increased cytopathic effects and reduced cell viability in SK-N-SH cells. We did not observe any such effect in U937 cells post-ZIKV infection. Notably, H3 and H4 histone levels rose in U937 cells after 3 days of ZIKV infection. Similarly, ZIKV-infected Vero cells displayed heightened H4 levels at 1-day post-infection. Furthermore, histone modification marks associated with active (H3K4me3, K3K36me3, H4K20me1) and repressed (H3K27me3, H3K9me3, H4K20me3) chromatin were elevated in U937 cells. We're currently analyzing RNA-seq data to delve deeper into these findings.

Conclusion: Our data suggest histones, especially H4, may play a role in defending against ZIKV. We aim to validate this hypothesis in upcoming research.

Novel aspect

Transcriptome and epigenome analysis found elevated total and modified histones in flavivirus-infected cells, indicating a possible antiviral role. We aim to explore if these modified histones affect cell traits and serve as indicators for past and present flavivirus infections.

23: Metagenomics and Targeted Capture Next Generation Sequencing to Detect and Identify Insects in Foods

Monica Pava-Ripoll, FDA/CFSAN; Mark K. Mammel, FDA/CFSAN; Elizabeth Reed, FDA/CFSAN; Martine Ferguson, FDA/CFSAN; Padmini Ramachandran, FDA/CFSAN
Monica.Pava-Ripoll@fda.hhs.gov

Abstract

Background: Insects are capable of adulterating food products. Current methods to analyze for insect contaminants involve their extraction from the food matrix followed by microscopic examination to quantify and identify extracted insect fragments. Both quantification and identification are relevant for regulatory purposes as they provide information on the level and origin of the insect adulteration (field pests versus storage pests). However, the taxonomical identification of the insect fragments using microscopic examination is time-consuming and does not always provide species level identification. Advances in metagenomics have enabled the development of new techniques to detect insect adulterants more effectively. Target enrichment by hybridization-based capture is becoming an effective method to increase sensitivity for low level detection of target DNA in a complex food background. When used in conjunction with reliable databases, it also provides accurate taxonomical identification of target organisms.

Purpose: Evaluate next generation sequencing (NGS) metagenomics and targeted capture using insect-specific baits to increase the sensitivity of insect detection and identification in foods.

Methodology: A targeted capture kit containing baits from about 2200 insect species was custom designed. Insect fragments from the Indian meal moth, Plodia interpunctella, were used to spike wheat flour at 8 levels (0, 1, 2.5, 5, 7.5, 10, 100, and 1000 ppm), five replicates per spiking level (n=40). Genomic DNA was extracted from 10g of samples and then fragmented using the KAPA HyperPlus Kit. Pre- and post-target capture libraries were amplified and sequenced on the Illumina Miseq system. Sequencing data were analyzed with the MitoK-mer pipeline to estimate the relative abundance of sequence reads and the four-parameter logistic analysis was used to calculate limit of detection (LOD) in pre- and post-targeted capture data.

Results: The relative abundance of spiked insect fragments ranged from 0.1%±0.07 to 62%±2.5 and from 16%±1.4 to 99%±0.05 in pre- and post-targeted capture samples, respectively. Targeted capture decreased sequence reads from the food background and increased the LOD by 3.2-fold, from 30.7 to 9.7 ppm.

Conclusion: Metagenomics and targeted capture NGS using insect-specific baits increased the sensitivity of detection of low levels of insect fragments in food samples and provided accurate taxonomical insect identification.

Novel aspect

A novel approach for the detection of insect contaminants in food samples is being developed. This approach combines targeted hybridization-based capture with next generation sequencing (NGS) metagenomics, thus enabling a high degree of sensitivity and accurate taxonomical identification of contaminants from arthropod origin.

26: Assessing viral variant detection algorithms to improve characterization of gene therapy products and biologics

Yanfei Zhou, MilliporeSigma; Bradley Hassan, MilliporeSigma; Alexandra Bridgeland, MilliporeSigma; William Dolan, Milliporesigma; Amber Overgard, MilliporeSigma
hiral.desai@milliporesigma.com

Abstract

Identity testing is an evolving requirement of regulatory agencies to characterize biologics intended for human therapeutic products such as vaccines and gene therapy. The use of Next generation sequencing (NGS) is increasingly used in the biosafety testing sphere to characterize and confirm the identity of viral seed stocks and expression vectors used in vaccine or gene therapy products . The most common approaches for gene therapy involve using a viral vector to deliver genetic material to patient cells. It is important to characterize viral vectors to ensure they do not contain variants which can negatively impact patient outcomes. Downstream bioinformatics analysis must capture true positive variants and limit spurious results to establish sequence identity and purity of the expression vectors. Advancements in best practices and standards for viral variant detection is critical to ensure the safety of patients and meet the expectations of regulatory guidance.

For identity testing, a diverse range of computational approaches is utilized to analyze 12 open-source and commercial variant callers to examine viral variant calling. We evaluate 12 variant callers including bcfTools, mutect2, varscan2, freebayes, deepvariant, vardict for sensitivity and specificity of viral variant detection. We use 189 in silico and synthetic viral data representing different variants including snps, indels, large insertions and large deletions to conduct studies to determine false positives, false negatives, and recommend optimal methods for viral variant detection. By improving the sensitivity and specificity for detecting variants within biologics we can improve the confidence in the results of clinical trials which can benefit clinical outcomes for patients requiring novel approaches to treating genetic conditions or cancer.

Results demonstrate that NGS can detect 1% variant frequencies, integrating multiple variant calling applications is essential for detecting different types of variants, and marking duplicate reads arising from PCR amplification can improve analysis turnaround time while maintaining sensitivity and specificity. This presentation will also demonstrate how utilizing paired-end inward oriented reads for analysis can improve confidence in variant calling by depleting false positive variants.

Novel aspect

NGS is increasingly being used to characterize and confirm the identity of viral vectors used for therapeutics. Standards for viral variant detection must be leveraged by analyzing and optimizing bioinformatics applications for viral variant analysis.

31: Temporal changes in Shiga-Toxin producing Escherichia coli (STEC) O121 transcriptome during storage in bleached flour

Ian_Hines, FDA/CFSAN; Emily_Nguyen, University of Maryland/JIFSAN; Ellie_Meeks, University of Maryland/JIFSAN; Sultana_Solaiman, FDA/CFSAN; Elizabeth_Reed, FDA/CFSAN; Maria_Hoffmann, FDA/CFSAN; Jie_Zheng, FDA/CFSAN
Ian.Hines@fda.hhs.gov

Abstract

Introduction: Persistence of Shiga-toxin producing Escherichia coli (STEC) within low-moisture foods, such as flour, has led to several STEC-related outbreaks associated with consumption of raw or undercooked flour products once considered microbiologically safe. The molecular mechanisms by which STEC can survive in flour are not well understood.

Purpose: This study sought to develop an effective sample treatment and RNA extraction method for polysaccharide-based matrices like flour; and to profile STEC transcriptome changes during storage in flour.

Methods: All-purpose bleached flour was inoculated with STEC O121 at 109 CFU/g flour. Flour samples were collected at nine time points within the first 48 hours post inoculation and treated with different centrifugation-filtration schemes. The RNA extraction method was optimized, and, after total RNA extraction of all samples, libraries were prepared using the Illumina Stranded Total RNA Prep kit and sequenced on the Illumina NextSeq2000 platform. The raw reads were quantified using salmon and analyzed with DeSeq2 to identify temporal changes in the transcriptome profile upon flour inoculation and storage.

Results: Cellular loss and residual flour in the filtrates were successfully minimized by electing to utilize a 1 µm pore-size membrane filter following quick centrifugation, which consistently led to isolation of high-quality RNA. Preliminary analysis revealed several genes were differentially expressed in all flour samples relative to a pure culture control. Approximately 1200 genes were expressed significantly differently (Padj < 0.05) between culture control and flour inoculation. Expression levels for temporally associated genes, mainly stress response and metabolic genes, had a major spike in the rate of expression change at inoculation. However, the levels appeared to stabilize after five hours of storage at which point few genes were differentially expressed relative to the prior timepoint.

Significance: Transcriptomic profiling of STEC persistence within flour will increase the understanding of STEC survival in low-moisture environments and help develop new mitigation strategies.

Novel aspect

Transcriptomic profiling of STEC persistence within flour will increase the understanding of STEC survival in low-moisture environments and help develop new mitigation strategies.

32: A Machine Learning Approach for Identifying Variables Associated with Risk of Developing Neutralizing Antidrug Antibodies to Factor VIII

Atul Rawal, FDA/CBER; Christopher Kidchob, FDA/CBER; Jiayi Ou, FDA/CBER; Zuben E. Sauna, FDA/CBER; Osman N. Yogurtcu, FDA/CBER ; Hong Yang, FDA/CBER
Atul.Rawal@fda.hhs.gov

Abstract

A key unmet need in the management of hemophilia A (HA) is the lack of clinically validated markers that are associated with the development of neutralizing antibodies to Factor VIII (FVIII) (commonly referred to as inhibitors). This study aimed to identify relevant biomarkers for FVIII inhibition using Machine Learning (ML) and Explainable AI (XAI) using the My Life Our Future (MLOF) research repository. The dataset includes biologically relevant variables such as age, race, sex, ethnicity, and the variants in the F8 gene. In addition, we previously carried out Human Leucocyte Antigen Class II (HLA-II) typing on samples obtained from the MLOF repository. Using this information, we derived other patient-specific biologically and genetically important variables. These included identifying the number of foreign FVIII derived peptides, based on the alignment of the endogenous FVIII and infused drug sequences, and the foreign-peptide HLA-II molecule binding affinity calculated using NetMHCIIpan. The data were processed and trained with multiple ML classification models to identify the top performing models. The top performing model was then chosen to apply XAI via SHAP, (SHapley Additive exPlanations) to identify the variables critical for the prediction of FVIII inhibitor development in a hemophilia A patient. Using XAI we provide a robust and ranked identification of variables that could be predictive for developing inhibitors to FVIII drugs in hemophilia A patients. These variables could be validated as biomarkers and used in making clinical decisions and during drug development. The top five variables for predicting inhibitor development based on SHAP values are: (i) the baseline activity of the FVIII protein, (ii) mean affinity of all foreign peptides for HLA DRB 3, 4, & 5 alleles, (iii) mean affinity of all foreign peptides for HLA DRB1 alleles), (iv) the minimum affinity among all foreign peptides for HLA DRB1 alleles, and (v) F8 mutation type.

Novel aspect

Use of AI/ML methods to predict inhibibitor development in patients with Hemophilia A. Explainable AI (XAI) is utilized to highlight the variables that are predictive in inhibitor development.

38: Single cell sequencing of brain sequestered CD8+ T cells during experimental cerebral malaria

Miranda S. Oakley, FDA/CDER; Micah Phillip, Emory University; Victoria Majam, FDA/CBER; Hong Zheng, FDA/CBER; Mark A. KuKuruga, FDA/CBER; Gregory K. Tharp, Emory University; Sanjai Kumar, FDA/CBER
Miranda.Oakley@fda.hhs.gov

Abstract

Background: Brain-sequestered CD8+ T cells play a prominent role in the pathogenesis of experimental cerebral malaria (ECM) in mice. A recent study has also shown that CD8+ T cells target the cerebrovasculature in children with cerebral malaria caused by Plasmodium falciparum. CD8+ T cells accumulate in the brain during the effector phase of a P. berghei ANKA (Pb-A) infection in mice and promote pathogenesis by inducing apoptosis of endothelial cells of the blood brain barrier.

Purpose: The molecular events associated with this CD8+ mediated pathogenesis remain poorly understood. To better characterize how brain sequestered CD8+ T cells exert their pathogenic effect, we performed single cell sequencing of the transcriptome of CD8+ T cells isolated from perfused brain tissue of Pb-A infected moribund and non-moribund C57BL/6 mice and uninfected C57BL/6 mice.

Methodology: A single cell whole transcriptome library was prepared from RNA isolated from approximately 50,000 brain sequestered leukocytes. Approximately 2-2.5 billion clusters were then paired-end sequenced on a NovaSeq S4 PE100 flow cell. Bioinformatic analysis was subsequently performed.

Results: A large cluster of brain sequestered CD8+ T cells defined by abundant expression of the cd3d, cd3e, cd8B1, cd8a, gzmb, and gzmk genes was identified. Interestingly, this cluster of CD8+ T cells consists of eight subsets and in-depth analysis is being performed to define these subsets. Importantly, to understand which genes and clusters associate with the symptoms of cerebral malaria, we determined the set of genes that are differentially expressed and also examined changes in cluster size in moribund versus non-moribund mice. This same analysis was also performed in non-moribund versus uninfected mice to better understand which genes and clusters associate with P. berghei ANKA infection. Lastly, the distribution of expression of known biomarkers of cerebral malaria pathogenesis such as IFN-y and TNF-a is being determined.

Conclusion: To our knowledge, this is the first study to perform single cell sequencing on brain sequestered CD8+ T cells during experimental cerebral malaria. Future experiments are being designed to validate our results by flow cytometry.

Novel aspect

Our single cell sequencing has resulted in the identification of novel molecular biomarkers and unique cellular subsets of CD8+ T cell mediated pathogenesis of cerebral malaria.

40: The Transcriptome Landscape of 3D-cultured Placental Trophoblasts Reveals Activation of TLR2 and TLR3/7 in Response to Trypanosoma cruzi

Erica Silberstein, FDA/CBER; Charles C. Chung, FDA/CBER, Alain Debrabant, FDA/CBER
Erica.Silberstein@fda.hhs.gov

Abstract

Background: Vertical transmission (VT) of the blood-borne protozoan Trypanosoma cruzi (T. cruzi) has become a global health problem accounting for 22% of new cases of Chagas disease (CD). Congenital infection is now considered the main route of CD spread in non-endemic countries, including the U.S., where the estimated number of T. cruzi-infected women of childbearing age is 40,000 and no routine disease testing is implemented during pregnancy. The mechanisms of T. cruzi VT remain poorly understood. Congenital infection may occur when bloodstream parasites interact with the placenta syncytiotrophoblasts (SYNs). We have previously shown that 3D-grown SYNs are highly resistant to T. cruzi infection. In this work, we used RNA sequencing and whole transcriptome analysis to explore the immunological signatures that drive SYNs' infection control.

Methodology: Syncytiotrophoblasts were cultured on microcarrier beads in a rotating bioreactor. 3D-cultured SYNs were exposed to T. cruzi. RNA sequencing followed by transcriptome analysis was performed to identify differentially expressed genes (DEGs) between 3D and 2D SYNs, unexposed or exposed to T. cruzi. TLR2 antibody blockade and siRNA silencing of TLR3 and TLR7 studies were conducted to elucidate the role of Toll Like Receptors (TLRs) in SYNs' control of parasite infection.

Results: When we compare the transcriptome of 3D-grown SYNs to 2D SYNs, we found that the largest category of DEGs are associated with inflammation and innate immunity functions. Quantitative RT-PCR evaluation of selected DEGs, together with detection of cytokines and chemokines in 3D SYNs culture supernatants, confirmed the transcriptome data. Several genes implicated in the Toll-like receptors signaling pathways were upregulated in 3D-grown SYNs. In fact, TLR2 blockade and TLR3/7 knockdown stimulated T. cruzi growth, suggesting that these molecules play a significant role in the host cell response to infection. Ingenuity Pathway Analysis of DEGs predicted the activation of pathways such as S100 protein family, pathogen induced cytokine storm, wound healing, HIF1a signaling and phagosome formation after T. cruzi exposure. Our findings indicate that SYNs resist infection by constitutively expressing pro-inflammatory molecules and modulating multiple defense mechanisms that interfere with the parasite's intracellular life cycle, contributing to parasite killing and infection control.

Novel aspect

Babies born to T. cruzi-infected mothers are at risk of developing severe Chagas disease Using an in-vitro 3D model of the human placenta and through transcriptome analysis, we identified multiple defense mechanisms and factors involved in parasite infection control. These studies may help evaluate new therapeutics to reduce congenital CD.

42*: Dual scRNA-Seq Analysis Reveals Rare and Uncommon Parasitized Cell Populations in Chronic L. Donovani Infection

Thalia Pacheco-Fernandez FDA/CBER;, Konstantinos Karagiannis, Moderna; Sreenivas Gannavaram FDA/CBER; Chaitenya Verma Wexner, The Ohio State University; Parna Bhattacharya, FDA/CDRH; Hira L. Nakhasi, FDA/CBER; Abhay R. Satoskar, The Ohio State University
Thalia.PachecoFernandez@fda.hhs.gov

Abstract

Background: Leishmaniasis is a parasitic disease affecting more than 90 counties. Visceral leishmaniasis (VL) is the systemic manifestation of this disease caused by the parasite Leishmania donovani. VL can be lethal in up to 95% of the cases if left untreated, while distinct rates of relapse are observed in endemic regions. The mechanism of relapse is still unknown, although some studies suggest that low numbers of parasites remain dormant in the host until the immune system becomes permissive for the disease to reemerge.

Purpose: Although phagocytic cells are documented targets of Leishmania parasites, it is unclear whether other cell types can be infected. Our goal is to use identify and annotate all infected cells in the spleen and bone marrow to detect even the rarest cell populations that function as parasite hosts.

Methodology: Bone marrow and spleen from chonically infected mice were dissected to obtain sibgle cell suspensions. Unbiased scRNA-seq was used to simultaneously analyze host cell and Leishmania donovani transcriptomes to identify and annotate parasitized cells in spleen and bone marrow in chronically infected mice.

Results: Our dual-scRNA-seq methodology allowed the detection of heterogenous parasitized populations. In the spleen, monocytes and macrophages are the dominant parasitized cells, while megakaryocytes, basophils and NK cells are found unexpectedly infected. In the bone marrow, the Hematopoietic Stem Cells (HSCs) expressing phagocytic receptors FcγR and CD93 are the main parasitized cells. Additionally, we also detected parasitized cycling basal cells, eosinophils, and macrophages in chronically infected mice. Flow cytometric analysis confirmed the presence of parasitized HSCs.

Conclusion: Our unbiased dual scRNA-seq method identifies rare, parasitized cells, potentially implicated in pathogenesis, persistence, and protective immunity without prior knowledge of potentially infected cells.

Novel aspect

The use of sc-RNA-seq from both the parasite and the host genome and integrate it to idenify rarely parasitized cells.

46*: CSP2: A Nextflow Pipeline for the Fast and Accurate Genetic Distance Estimation from Assembled Pathogen Genomes

Robert Literman, FDA/CFSAN; James Pettengill, FDA/CFSAN; Hugh Rand, FDA/CFSAN
Robert.Literman@fda.hhs.gov

Abstract

Accurate genetic distance estimation among pathogens facilitates precise responses to outbreak events and is critical for understanding pathogen movement throughout the supply chain. FDA currently employs the CFSAN SNP Pipeline to estimate genetic differences among pathogens, and this pipeline maps short-read, whole-genome sequencing data against a reference genome. While accurate, read mapping can be time-intensive and analytical runtimes will continue to grow as isolate clusters get larger. Here we present CSP2 (CFSAN SNP Pipeline 2), a Nextflow pipeline for genetic distance estimation that replaces read mapping (which can take minutes to hours) with MUmmer fast whole-genome alignment (which takes seconds) with comparable results. CSP2 can be run in two main modes: Reference Screener Mode and SNP Pipeline Mode. In Reference Screener Mode, the user can rapidly query incoming read or assembly data against a custom reference database (e.g., sets of lab control strains, or active outbreak strains). In SNP Pipeline Mode a reference isolate is either provided or inferred from the data, and CSP2 identifies variable sites, then outputs a sequence alignment and pairwise distances comparable to those in the NCBI Pathogen Detection database or generated through the CFSAN SNP Pipeline. Here we analyze clusters of Salmonella, E. coli, L. monocytogenes, and Cronobacter and find that CSP2 results are strongly correlated with those from the CFSAN SNP Pipeline but are generated in a fraction of time.

Novel aspect

Contrasted with read mapping-based strategies, applying Mummer fast whole genome alignment and sensible data filtering can reduce the time for estimation of genetic distances among isolates from hours to minutes, all while maintaining accuracy

47: Analysis of Various Cryopreservation Conditions for the Storage of Induced Pluripotent Stem Cells and other Cell-based Therapies

Shweta Kotian, FDA/CBER; Farhad Farjood, Neural Stem Cell Institute; Nathan Boles, Neural Stem Cell Institute; John T. Thomas, FDA/CBER; Brigitte L. Arduini, Neural Stem Cell Institute; Jeffrey Stern, Neural Stem Cell Institute; Sally Temple, Neural Stem Cell Institute; Malcolm Moos Jr., FDA/CBER
Shweta.Kotian@fda.hhs.gov

Abstract

Cellular therapies present immense therapeutic opportunities and challenges. Identifying critical quality attributes to assure product safety and effectiveness and preserving these during storage and transport to the point of care are vital considerations from a regulatory standpoint. We used relative fragile iPSCs to compare fixation methods- methanol or DSP [dithio-bis (succinimidyl propionate)] with conventional cryopreservation (with DMSO) at several time points. Cell viability was compromised in all methods compared with fresh cells. Single-cell RNA sequencing (sc-RNA seq) analysis revealed that all cell preservation methods altered transcriptomic profiles in comparison with fresh iPSCs. Fresh and cryopreserved cells had similar profiles whereas methanol and DSP fixation caused significant cell death and altered profiles. Methanol and DSP fixed cells also displayed increased mitochondrial genes compared with fresh and frozen iPSCs, an indicator of poor quality. These data suggest that neither fixation method is suitable for iPSCs, and cryopreservation outperforms the other methods. Therefore, we analyzed cryopreservation in more detail, comparing cryopreserved iPSCs and other cells such as retinal pigmented epithelium (RPE) cells with their fresh counterparts. Fresh RPE cells were compared with cells subjected to two cryopreservation timepoints: a) cryopreserved for 24 hours, thawed, and immediately processed for analysis; and b) cryopreserved for 24 hours, thawed, reconstituted in a proprietary recovery medium for 48 hours at 4 °C, then processed for analysis. While viability was reduced in the cells in recovery medium versus the fresh and frozen-thawed cells, other quality metrics were above average for all samples. Cells were analyzed by scRNA-seq, scATAC-seq, and multiome-seq. Initial data analysis revealed that there were significant differences in the transcriptomic profiles between fresh and cryopreserved RPEs, and in their chromatin profiles. Subsets within each cohort with variable profiles may help identify quality attributes relevant to the safety and efficacy of fresh versus cryopreserved RPE cells. We will include additional cryopreservation timepoints and protocols for iPSCs and RPE cells and other cell therapy products, to explore cellular responses to cryopreservation and reconstitution. Such a comprehensive analysis will help create a ‘best practices’ set of standards to evaluate and optimize storage and transportation of cell therapy

Novel aspect

A comprehensive analysis which will identify critical quality attributes to assure cell therapy product safety and efficacy during storage and transport to the point of care which are vital considerations from a regulatory standpoint.

Metabolomics, Proteomics

03: The Challenge of Plant Identification in Complex Mixtures: Closely Related Families, Large Proteomes, and Unsequenced Genomes

Melinda A. McFarland, FDA/CFSAN; Sara M. Handy, FDA/CFSAN; Elizabeth Hunter, FDA/CFSAN; Christine H. Parker, FDA/CFSAN; Ann M. Knolhoff, FDA/CFSAN
Melinda.McFarland@fda.hhs.gov

Abstract

Identification of unknown plant proteins is a formidable task due to the large number of species with few known protein sequences, complex genomes, and high sequence homology. While food safety efforts have long used mass spectrometry to identify contaminants in food and identification of proteins by LC-MS/MS has become routine, the interface of proteomics and multi-species food analysis remains remarkably complicated. Interest in plant-based protein-rich food continues to grow, but informatics solutions to facilitate the identification of multiple plant species in a complex mixture have not kept pace. This is especially true for identification of closely related plant species and unknown plant contaminants. The presented data will illustrate informatic challenges encountered in proteomics identification of closely related seeds, legumes, and toxic plant contaminants.

An analysis of pecan and walnut illustrates the challenge of differentiating closely related plant species when one has a sequenced genome and the other is sparsely sequenced. These results inform selection of protein sequences for more complex multi-species plant databases akin to metaproteomics. A subsequent analysis of an unknown and unsequenced plant contaminant in a complex mixture of other plants found in contaminated fortified cereal distributed as food aid in Uganda in 2019 is also presented. While small molecule analysis identified the toxin, the proteomics informatic workflow determined the toxin was likely introduced as part of a plant, the part of the plant, the plant's phylogenetic family, and that the contaminant was likely an unsequenced plant in a complex mixture of other plants. Genome skimming was used to identify the specific plant species.

The presented results were derived from LC-MS/MS data of tryptic digests from a variety of plant-based food samples. The resultant peak lists were searched against custom multi-species protein sequence databases using the Mascot search engine. Parsimonious lists of identified protein families, peptides, associated species, and comparisons across samples were generated with MassSieve software.

Novel aspect

Identification of unknown plants in complex mixtures by proteomics.

04*: Proteomic Profiling Reveals Antibiotic Resistance Mechanisms in Staphylococcus epidermidis Biofilms under Tigecycline Pressure

Kidon Sung, FDA/NCTR; Miseon Park, FDA/NCTR; Jungwhan Chon, Kyungin Women's University, Korea; Ohgew Kweon, FDA/NCTR; Saeed Khan, FDA/NCTR; Andrew Shen, FDA/NCTR; Angel Paredes, FDA/NCTR
Kidon.Sung@fda.hhs.gov

Abstract

Staphylococcus epidermidis is a leading cause of biofilm-associated infections on implanted medical devices. During the treatment of an infection, bacterial cells inside biofilms may be exposed to sublethal concentrations of the antimicrobial agents. In the present study, the effect of subinhibitory concentrations of tigecycline (TC) on biofilms formed by S. epidermidis strain RP62A was investigated using a quantitative global proteomic technique. As TC concentration increased, the number of viable cells in biofilms gradually decreased. Strain RP62A biofilms treated with 1/8 TC (T1) and 1/4 minimum inhibitory concentration (MIC) TC (T2) were much denser than the untreated biofilms. On the other hand, biofilms treated with 1/2 MIC TC (T3) were significantly dispersed. Overall, 413, 429, and 518 proteins were differentially expressed in T1, T2, and T3 biofilms, respectively. As the TC concentration increased, the number of induced proteins in each cluster of orthologous groups (COG) and Kyoto Encyclopedia of Genes and Genomes (KEGG) category increased. The TC concentration dependence of the proteome response highlights the diverse mechanisms of adaptive responses in strain RP62A biofilms. In both COG and KEGG functional analyses, most upregulated proteins belong to the cellular metabolism, suggesting that it may play an important role in the defense of strain RP62A biofilm cells against TC stress. Sub-MIC TC treatment against strain RP62A biofilms led to significant changes of protein expression related to biofilm formation, antimicrobial resistance, virulence, quorum sensing, ABC transporters, protein export, purine/pyrimidine biosynthesis, ribosome, and essential proteins. Interestingly, in addition to tetracycline resistance, proteins involved in resistance to various antibiotics, including aminoglycosides, antimicrobial peptides, β-lactams, erythromycin, fluoroquinolones, fusidic acid, glycopeptides, lipopeptides, mupirocin, rifampicin, and trimethoprim, were differentially expressed. Our study demonstrates that global protein expression profiling of biofilm cells to antibiotic pressure may improve our understanding of the mechanisms of antibiotic resistance in biofilms.

Novel aspect

Strain RP62A's proteome response depended on TC concentration, showcasing differential expression in functional protein groups (e.g., biofilm formation, antimicrobial resistance). Sub-MIC TC levels triggered active upregulation of metabolic and protective proteins, warranting future studies on biofilm protection against sublethal antibiotics.

06: Optimizing dia-PASEF isolation window schemes for proteomics measurements on a timsTOF ultra instrument

Markus Lubeck, Bruker Daltonics; ,Stephanie Kaspar-Schoenefeld, Bruker Daltonics; Christoph Krisp, Bruker Daltonics; Andreas Schmidt, Bruker Daltonics; Florian Busch, Bruker Daltonics; Eduardo Carrascosa, Bruker Daltonics; Oliver Raether; Gary Kruppa, Bruker Daltonics

Ryan.Marsico@bruker.com

Abstract

Data-Independent Acquisition (DIA) is widely used for proteomics as it usually outperforms Data-Dependent Acquisition (DDA) for protein identification and quantitation, due to its higher ion usage and reproducibility, resulting from a fixed scheme of rather broad isolation windows. This advantage can be further increased by combining it with trapped ion mobility separation (TIMS), as the additional separation dimension reduces complexity and the sequential elution of condensed ion packages from the TIMS device allows for even more efficient ion usage (dia-PASEF). The two-dimensional mass and mobility space enables method creation with extensively different window schemes.

Novel aspect

Highly sensitive proteomics acquisition for low sample applications such as single cell proteomics and immunopeptidomics.

07: Deeper plasma proteome coverage enables identification of novel biomarkers and classification of diseases

Shourjo Ghose, Bruker Daltonics; Andreas Schmidt, Bruker Daltonics; Zehan Hu, Bruker Daltonics; Claudia Martelli, Bruker Daltonics; Katharina Limm, Bruker Daltonics; Xaver Wurzenberger, Bruker Daltonics; Katrin Hartinger, Bruker Daltonics; Sebastian Mueller, Bruker Daltonics; Nils A. Kulak Bruker Daltonics
Shourjo.Ghose@bruker.com

Abstract

Blood plasma is one of the least invasive biopsies and a valuable specimen for clinical research and patient health monitoring. Since almost all tissues are sustained by the constant blood flow and proteins are constantly being actively secreted or leaked into the blood, plasma provides comprehensive information about health or disease state. However, access to proteome information is limited by the highly dynamic nature of protein abundance in plasma, which spans approximately 10 orders of magnitude and with only 22 proteins accounting for 99% of the whole protein mass. To address this challenge, we developed a novel workflow for LC-MS-based plasma proteomics that enriches low abundant proteins and enables an improved coverage of the plasma proteome.

Novel aspect

Deep plasma proteomics using novel sample preparation and trapped ion mobility mass spectrometry

09: Identifying metabolite markers for detection of flaviviruses in human induced Pluripotent Stem Cells (hiPSCs)

Tahira Fatima, FDA/CBER; Sandip De, FDA/CBER
Tahira.Fatima@fda.hhs.gov

Abstract

Background: Human cells, tissues and cellular and tissue-based products (HCT/Ps) are required to comply with the donor eligibility requirements as per 21 CFR part 1271 and applicable guidance documents. Communicable diseases screening now includes screening donors for risk of infection with Zika virus (ZIKV), and West Nile Virus. Moreover, rising global temperatures and expanding insect vector populations have led to a surge in vector-borne diseases, including flaviviruses. So, a rapid, sensitive, and precise pan-flaviviral detection tool may prove useful to detection of cells/tissues that may be contaminated and able to transmit disease to recipients.

Purpose: 1. Compare the sensitivity, speed and specificity of metabolomics to the NATs and immunogenic detection assays. 2. Develop metabolite markers for detecting flaviviruses in hiPSCs.

Methodology: hiPSCs were infected with DENV3 and ZIKV-MR766 strains. We assessed impact of viral infections on hiPSCs by examining cell viability, cytopathic effect (CPE), and viral load using immunofluorescence assay (IFA) and qRT-PCR. For metabolite extraction standardization, we analyzed 0.2, 0.8, 2, and 8 million control and infected cells using untargeted flow-injection mass spectrometry (FIA-MS). Ions were annotated by matching their inferred mass with compounds in the Human Metabolome Database. Currently, we are using untargeted LC-MS as a more sensitive and accurate method than FIA-MS to analyze cells infected by both viruses to identify all metabolites.

Results: We did not observe significant differences in cell viability and CPE in iPSCs infected with DENV3 and ZIKV vs control. However, IFA and qRT-PCR results indicated higher virulence of ZIKV compared to DENV3. FIA-MS detected same number of annotated compounds in all tested samples concluding that 0.2 million cells are sufficient for extracting maximum number of metabolites. Downstream analysis of the FIA-MS data identified 5 differentially expressed compounds in infected samples at 24 hpi. These findings also confirmed sensitivity and consistency of FIA-MS. Results (including LC-MS) will be presented in our poster.

Conclusion: We demonstrated that 'Metabolomics' is a reproducible, and sensitive method for detecting flaviviruses in infected cells. Our subsequent objective is to validate the ‘specificity’ of the identified metabolites and investigate the applicability of this technology to detect other infectious diseases.

Novel aspect

Creating sensitive and precise pathogen detection tools is critical for the success of cell and tissue therapy. Our study using DENV and ZIKV demonstrates that 'Metabolomics' is highly sensitive compared to standard NAT and immunogenic assays. We've discovered potential flavivirus detection markers in hiPSCs, currently validating their specificity.

10: A practical lock-mass calibrant introduction method for improved mass accuracy and reduced false positive identifications in non-targeted analysis

Christine Fisher (O'Donnell), FDA/CFSAN; Shannon Murphy, FDA/CFSAN; Ann Knolhoff, FDA/CFSAN
Christine.ODonnell@fda.hhs.gov

Abstract

In addition to isotopic distribution, accurate mass information obtained during non-targeted analysis (NTA) using liquid chromatography/high-resolution mass spectrometry (LC-HRMS) is vital for unknown identification in complex samples. Minimizing mass accuracy error can reduce false positive identifications. HRMS instruments can obtain lower mass accuracy errors using internal/lock-mass recalibration compared to external calibration. Background ions can be used as lock-masses; however, they require identification and can change over time. We present a reliable method for introducing lock-mass calibrant ions to a Thermo Q-Exactive using a calibrant reservoir kit (Bruker) placed ~2cm inside the source. Briefly, the kit consists of a stainless-steel reservoir that holds 2 sponges with absorbed calibrant and a perforated cap. The calibrants included methyl stearate ([M+H]+, m/z299), hexakis (2,2-difluoroethoxy)phosphazene ([M+H]+, m/z622; [M+formate]-, m/z666) and hexakis (1H,1H,3H-tetrafluoropropoxy)phosphazene ([M+H]+, m/z922; [M+formate]-, m/z966). Co-dosing the lower m/z calibrants on the same sponge under a 5-hole cap and the highest m/z calibrant on a separate sponge under a 12-hole cap generated stable signal for all calibrants for >36hrs. The lock-masses used in the method can be set to “best” (best calibrant used for recalibration) or “if all present” (“all”; all calibrants required for recalibration). Both settings corrected mass accuracy errors of 3-5 ppm to <1 ppm for >92% of the ~80 compounds in a standard mixture. In comparison, using the background ions diisooctyl phthalate ([M+H]+, m/z391) and sodium formate ([2M-2H+Na+]-, m/z112), reduced mass accuracy errors to <1ppm for 73% of the standard compounds. Using Compound Discoverer, top-ranked, correct molecular formula were obtained for 30/67 (“best”) and 35/67 (“all”) standard compounds compared to 23/67 without recalibration. MS/MS match scores improved for 43/52 compounds with lock-mass calibration. The “all” condition resulted in fewer recalibrated spectra throughout the chromatogram compared to the “best” condition. Therefore, the “best” condition is considered optimal for NTA. This method is amenable to other instruments with the HESI source.

Novel aspect

This simple method introduces lock-mass calibrant ions on the Q-Exactive series, improving mass accuracy errors (<1 ppm) over long queues (>36 hrs). This enables more accurate molecular formula assignment and MS/MS matches, which reduces false positives and improves confidence in annotations for NTA.

11: When is Enough Actually Enough- How Does the Number of Replicates Influence the Quality of Non-Targeted Analysis Results?

Karen Butler, FDA/CFSAN; Erica Bakota, FDA/ORA; Christine Fisher, FDA/CFSAN; Brian Ng, FDA/CFSAN; Ann Knolhoff, FDA/CFSAN
Karen.Butler@fda.hhs.gov

Abstract

Many routine analytical methods for food safety applications are targeted toward specific compounds or compound classes. In contrast, non-targeted analysis (NTA) using liquid chromatography coupled with high-resolution mass spectrometry (LC-HR/MS) can allow for a more global chemical analysis, facilitating the detection and identification of a wider range of compounds. The inherent chemical complexity of foods often results in the detection of thousands of compounds within each sample when assessed using an NTA approach, requiring careful data assessment to ensure compounds are reliably detected and extracted in each data file. The analysis of replicates can lend confidence to these assessments; however, to our knowledge, no studies have focused on investigating the impact of the number of extraction and injection replicates prepared and/or analyzed on the result output. This experiment was designed to determine the impact of extraction and injection replicates on molecular feature detection, result accuracy, and analysis time. Briefly, five replicates of a strawberry composite sample from the FDA Total Diet Study were prepared and analyzed in quintuplicate, with a single preparation injected for a total of thirty injections, using an NTA LC-HR/MS method. Different groups of sample preparation and injection replicates were processing together using Compound Discoverer 3.3. Each data analysis group was then assessed for the number of spiked quality control compounds and overall number of compounds extracted by the software, as well as their associated percent relative standard deviation (%RSD). Overall, more features were extracted by the software as increasing numbers of files were processed together, both for the total number of compounds extracted as well as quality control compounds. A comparable number of molecular features were extracted from the data for both triplicate injections of the same sample preparation and single injections of triplicate preparations. These efforts, with extension of the method to matrices in additional sectors of the AOAC food triangle, will facilitate practical guidance for the number/type of replicates considered when designing NTA experiments for analysis of foods.

Novel aspect

This study investigates the impact of the number of replicates prepared and/or analyzed on the reliability of non-targeted analysis results in foods.

19: Metabolomics Evaluation of the Impact of Violet-Blue Light (405 nm) on Platelet Concentrate

Chintamani D. Atreya, CBER/FDA; Jinchun Sun, FDA/NCTR; Neetu Dahiya, FDACBER, Tom Schmitt, FDA/NCTR; Caitlin Stewart, U of Strath; John Anderson; U of Strath; Scott MacGregor, U of Strath; Michelle Maclean, U of Strath; Richard. D. Beger, FDA/NCTR
Chintamani.Atreya@fda.hhs.gov

Abstract

Introduction: Ex vivo stored human platelet concentrates (PCs) are an essential life-saving transfusion product. PCs are susceptible to bacterial contamination during the 5-7 day of storage at 22 ± 2°C in gas permeable bags. Current FDA approved pathogen reduction technologies (PRT) use external photosensitizers in combination with UV light irradiation. Since UV light is harmful to mammalian cells, we have been evaluating violet-blue light to determine whether it can serve as an alternate to PRT technology, without the need for additional exogenous photosensitizers. An LC/MS-based metabolomics analysis was conducted to evaluate the impact of violet-blue light on the platelets (PLTs).

Methods: Apheresis-collected human PCs were uniformly split into two transfer bags from each donor. One PC bag was used as control (no light treatment) and the other bag was used for the light treatment. PLTs in the bags were exposed to 405 nm light at an irradiance of approximately 54 J/cm2 for 1h, and 270 J/cm2 for 5h. The protocol was approved by the FDA Research Involving Human Subjects Committee. LC/MS-based metabolomics analysis was conducted to identify the metabolic changes in both PLTs and plasma.

Preliminary Data: After 5h treatment, the lactate level detected by LC/MS was 1.6-fold increased, which was consistent with the fold change (~1.7 fold) detected by the blood gas analyzer. No changes were observed either in platelet aggregation inhibitory factors, or platelet activation factors. No changes were observed in lysoPCs or PCs, which indicated that the integrity of PLTs was intact after the light exposure. After 5h treatment, the most distinctive changes were increases in hydroxy-fatty acids, OH-fatty acyl-carnitines, and aldehydes, indicative of lipid peroxidation induced by the light treatments. Lower levels of glutathione, vitamin A and uric acid were observed, which might represent a neutralization response to the reactive oxygen species (ROS) generated by the light treatments. Decreases in bilirubin compounds and vitamin A, increases in vitamin D derivatives, which are all endogenous photosensitizer compounds, indicated that no exogenous photosensitizers are needed for the violet-blue light treatment.

Novel aspect

The study identified that no exogenous photosensitizers are needed for the violet-blue light treatment of ex vivo human platelets stored for transfusion and the light treatment has potential as an alternative to current chemical/UV light based pathogen reduction technologies for platelet safety from infectious agents.

29: Application of Metabolomic Analysis Towards the Discovery of Biomarkers of Immunogenecity and Efficacy of Parasitic Vaccines

Nazli Azodi, FDA/CBER; Hannah Markle, FDA/CBER; Timur Oljuskin, USDA; Parna Bhattacharya, FDA/CDRH; Greta Volpedo, The Ohio State University; Abhay Satoskar, The Ohio State University; Sreenivas Gannavaram, FDA/CBER; Hira L. Nakhasi, FDA/CBER
Sreenivas.Gannavaram@fda.hhs.gov

Abstract

Background: Leishmaniasis is a neglected tropical disease caused by Leishmania parasites which is prevalent in approximately 88 countries, yet no licensed human vaccine exists against it and treatment options remain limited. Towards control of leishmaniasis, we have developed Leishmania major Centrin gene deletion mutant strains (LmCen-/-) as a live attenuated vaccine, which induces a strong Th1 response to provide IFN-γ-mediated protection to the host. However, the immune mechanisms of such protection remain to be understood

Purpose: Metabolic reprogramming of the host cells following Leishmania-infection has been shown to play a critical role in pathogenicity. Therefore, our goal was to study the metabolic changes associated with the LmCen-/- strain to identify the immune mechanism of protection and biomarkers of immunogenicity

Methodology: C57/BL6 mice were infected with wild type L. major (LmWT)and LmCen-/-. The infected ear tissues were collected 7 days post infection and analyzed by untargeted LC/MS mass spectrometry, and the data were analyzed with the Metaboanalyst 5.0 for pathway analysis and Metscape 3.1.1 for integrative network analysis. To verify the results from MS analysis, murine bone marrow-derived dendritic cells, were infected with LmWT and LmCen-/. BMDCs were cultured with inhibitors or agonists of tryptophan metabolism, and the expression levels of genes of interest were measured via qRT-PCR.

Results: Our results show that distinct metabolic reprogramming occurs in the host cells infected with virulent or live attenuated Leishmania parasites. We have identified that Tryptophan metabolism is differentially regulated between the LmWT infection and LmCen-/- immunization. The LmWT infection promotes the anti-inflammatory Kynurenine-AhR and FICZ-AhR signaling, while the LmCen-/- immunization uses tryptophan for the increased synthesis of the pro-inflammatory mediator, melatonin.

Conclusion: Application of metabolomic analysis to vaccine studies identified immune mechanisms of protection or pathogenicity and may help identify novel biomarkers of vaccine efficacy of a live-attenuated vaccine candidate for Cutaneous Leishmaniasis.

Novel aspect

We applied the emerging field of metabolomics to study the metabolic biomarkers of immunogenicity in a parasitic vaccine. The results showed that infection and immunization environments induce distinct metabolic profiles, and that melatonin enrichment could be used as a metabolic biomarker of immunogenicity to evaluate the efficacy of the vaccine.

30: Evaluation of plasma proteome and miRNA changes related to COVID-19 patient severity response

Richard Beger, FDA/NCTR; Li-Rong Yu, FDA/NCTR; Han Tao, FDA/NCTR; Vikrant Vijay, FDA/NCTR; Lisa Pence, FDA/NCRT; Jinchun Sun, FDA/NCTR; Mallikarjun Bidarimath, FDA/NCTR; Thomas Schmitt, FDA/NCTR; Kelly Mercer, FDA/NCTR; Elysia Masters, FDA/NCTR; Jessica Hawes, FDA/NCTR; Heather Smallwood, Univ. Tennessee Health Science Center; John Arthur, Univ. Arkansas Medical Sciences; Sudeepa Bhattacharya, ASU; Keith Burkhart, FDA/CDER
Richard.Beger@fda.hhs.gov

Abstract

Background and Purpose: COVID-19 remains a worldwide pandemic where underlying health conditions like diabetes, cancer, obesity, high blood pressure, asthma, and smoking are all risk factors related to the severity of outcome from a SARS-CoV-2 infection. Therefore, gaining mechanistic insights at the molecular level to understand the differences in severity of the infection and discovering early biomarkers that enable prediction of outcomes among COVID-19 patients remains urgently needed. These insights could help ease burden of care and aid in evaluation of treatments or development of cures.

Methodology: COVID-19 positive patients were categorized into 3 symptom response categories: mild, out-patient; moderate, hospitalization without intensive care; and severe, hospitalization with intensive care. Blood samples from 93 COVID-19 patients collected at the time of initial diagnoses were processed to plasma and deidentified for proteomic and miRNA analysis.

Results: A total of 2939 proteins and 2097 miRNAs were analyzed in the plasma samples. Student's t-test was used for statistical significance with a p-value < 0.05 and fold change (FC) > 2.0 as the criteria. The number of significantly changed proteins was 369 (13%) for severe vs. mild%), 135 (5%) for severe vs. moderate, and23 (1%) for moderate vs. mild. The number of significantly changed miRNAs was 578 (28%) for severe vs. mild, 386 (18%) for severe vs. moderate, and 122 (6%) for moderate vs. mild. Many of the most significant protein differences in patients with severe COVID-19 (p-value < 10-5, FC > 5) were involved in inflammation and cardiac injury, which confirms earlier reports of "cytokine storm" and cardiovascular events.

Conclusions: Further investigations will be conducted, including analyzing more samples, evaluating metabolites and lipids to discover other outcome biomarkers, and, most importantly, aligning omics data with demographics, and clinical endpoints. The combination of omics and clinical data will be further evaluated for pathways analysis.

Novel aspect

Gaining mechanistic insights at the molecular level to understand the differences in severity of the COVID-19 infection should enable improve patient care. Analysis of the protein and miRNA data showed that immune function, inflammation, wound healing, and cell survival pathways were altered in the severe versus mild cases.

33: Applying MHC-II Associated Peptide Proteomics (MAPPs) assay to study the Immunogenicity of Staphylococcus aureus Cas9

Louis Hopkins, FDA/CBER; Joseph R. McGill, FDA/CBER; Swati Mukherjee, Editas; Kate Zhang, Editas; Zuben E. Sauna, FDA/CBER
Vijaya.Simhadri@fda.hhs.gov

Abstract

Background: The Major histocompatibility complex (MHC)- Associated Peptide Proteomics (MAPPs) is a mass spectrometry- based approach. MAPPs is a powerful tool to identify peptides presented by MHC proteins (an early step in the immune response). MAPPs was used in this study to identify Staphylococcus aureus (Sa) Cas9 derived peptides presented by the MHC Class II (MHC-II) proteins (i.e., putative T cell epitopes).

Purpose: MAPPs assay identifies the naturally processed peptides that are presented by MHC-II proteins. The MHC-II-peptide complex could potentially activate CD4+ T cells and elicit an immune response against protein therapeutics. Used in the preclinical development phase, the MAPPs assay can be used to identify immunogenic regions of a therapeutic protein. The information can be exploited to modulate immunogenicity.

Methods: The MAPPs assay was used to identify peptides derived from the SaCas9 protein on MHC-II molecules on antigen presenting cells. A panel of 18 MHC typed peripheral blood mononuclear cell (PBMC) samples from healthy human donors were used to generate the immature monocyte derived dendritic cells in vitro and matured in the presence of full length SaCas9 protein. Dendritic cells were then harvested and lysed. The MHC-II molecules were isolated by immunoaffinity chromatography. Peptides were then recovered from the MHC-II complexes and processed for further analysis by sequencing mass spectrometry.

Results: We have identified 860 peptides from a total number of 18 donors with a minimum length of 7 amino acids (AA) and maximum length of 24 AA. The average peptides per donor is 37.8 and we found 26 unique peptides.

Conclusion: SaCas9 is a large protein of bacterial origin and thus completely foreign to human immune systems. Of the 1,053 amino acids in the SaCas9 protein, we identified 26 peptides that were presented by MHC-II variants on dendritic cells from 18 individual donors. As SaCas9 is derived from bacteria, it is possible that all the peptides presented by antigen-presenting cells are likely to be recognized by T-cell receptors. By identifying specific regions of the protein that are most immunogenic we have targets for the potential engineering of SaCas9 for modulating its immunogenicity.

Novel aspect

Studies with Cas9 immunogenicity are limited but there is considerable experience with proteins that are used in clinical applications. Our study provides important tools for assessing the immunogenicity risk for direct delivery of CRISPR/Cas9 system as a ribonucleoprotein (RNP) complex (Cas9 protein and guide RNA).

35*: Lipidomics Evaluation of the Impact of Fentanyl Treatments on Neural Stem Cells

Richard Beger, FDA/NCTR; Jinchun Sun, FDA/NCTR; Rohini Donakonda, FDA/NCTR; Shuliang Liu, FDA/NCTR; Fang Liu, FDA/NCTR; Cheng Wang, FDA/NCTR
Richard.Beger@fda.hhs.gov

Abstract

Fentanyl is a potent and short-acting opioid medication that is often given to pediatric patients during surgery to relieve pain and as an adjunct to anesthesia. Because it is difficult to assess the adverse effects on human infants and children, the utilization of human-derived neural stem cell models, might be a good tool to evaluate the vulnerability of the developing nervous system to fentanyl exposure. Since neural cells contain a wide variety of lipid classes and lipid species, lipidomics analysis using ultra-high-performance liquid chromatography (UHPLC) coupled with high-resolution mass spectrometry (HRMS) was conducted to investigate the impacts of different doses of fentanyl on neural stem cells (NSCs) and neural cells differentiated/derived from NSCs. The neural cells were treated with vehicle (control), or fentanyl at 1, 10, or 100 µM for 24 hours. Although 24-hour fentanyl exposure of NSCs resulted in a dose-related increase (not significant) in the release of lactate dehydrogenase into the cell culture medium (indicator of cell death/damage), no significant reduction in the mitochondrial health marker [3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetra-zolium bromide (MTT) was observed vs control. Lipidomics analysis detected 1830 lipid species from 20 lipid classes. Consistent with MTT data, palmitoylcarnitine (indicating mitochondrial functioning) did not significantly accumulate after fentanyl exposure. Among the 20 lipid classes detected, the total abundance of cholesterol ester and sphingosine classes significantly decreased while ceramide and hexosylceramide classes significantly increased (>2 folds increases) in the high-dose group vs the control. This preliminary data indicated that the ceramide pathway might be disturbed by fentanyl treatments, which might provide the underlying mechanisms of fentanyl-induced neurotoxicity on developing neural cells.

Novel aspect

Lipidomics was conducted to investigate the adverse effects of fentanyl exposure to human-derived neural stem cells.

37*: MSlineaR: a new tool to assess linearity, improving statistical robustness and quality assurance in untargeted metabolomics

Janine Wiebach, Berlin Institute of Health at Charité – Universitätsmedizin Berlin; Jennifer Kirwan, Berlin Institute of Health at Charité – Universitätsmedizin Berlin; Álvaro Fernández Ochoa, University of Granada
janine.wiebach@bih-charite.de

Abstract

An ongoing challenge in untargeted mass spectrometry based metabolomics is cleaning up datasets for later statistical analysis. The use of serial diluted pooled QC samples can be used for this purpose to restrict the final dataset to only features which show a linear response. The resulting dataset is more statistically robust and comprises a greater proportion of biologically relevant peaks than the original. Existing tools (e.g. mzMatch, Nypc-Toolbox, MRC.R), are all currently based on using coefficient of determination (R2 values) or metrics of linear regression to identify these features.

To considerably improve this approach, an R-based tool, MSlineaR, was developed. It initially uses three curve fitting models to determine which signals are outliers and discount them which results in a considerably improved final linear model. Additionally, it identifies the non linear portions at the beginning and end of the concentration ranges, enabling the exact linear range to be determined per feature. A graphical function enables the user to inspect the distribution of study samples along the linear and non linear parts of the dilution curve. Finally, it flags features which show no, or poor linear behavior and can exclude them from the final dataset.

Using this tool, it was possible to reduce an untargeted dataset with ~ 60,000 features by 83 % in 20 minutes, based on their linear response behavior. To benchmark MSLineaR, 1410 heavy labeled C13 metabolites from the IROA internal standard were used as negative controls. MSlineaR correctly identified all 1410 features as non linear. As positive controls, a targeted data set was used with 68 standard calibrations which were assessed by eye and 63 of them show a linear response. MSlineaR correctly assigned all compounds according to their linear behavior. Early benchmarking against existing tools appears favourable. Use of this tool promises to be quick and will improve statistical robustness by filtering poor quality data.

Novel aspect

MSlinear highly improves the quality assurance in untargeted metabolomics by cleaning up data sets in an automatized and fast way by removing noise and low quality peaks. It is the only accessible tool which includes an outlier detection and determines not only the linearity but also the linear range.

41: Identification of Circulating Pharmacodynamic Biomarkers of IL-5 Inhibitors using proteomics approach.

Deepti Samarth, FDA/CDER; Lakshmi Manasa Chekka, FDA/CDER; Esraa Mohammed, FDA/CDER; Erica Decker, FDA/CDER; Yan Guo, FDA/CDER; Will Wheeler, IMS Inc.; Kristina Howard, FDA/CDER; Sarah J. Schrieber; Jeffey Florian, FDA/CDER; Yow-Ming Wang, FDA/CDER; David G. Strauss, FDA/CDER; Paula L. Hyland, FDA/CDER
Deepti.Samarth@fda.hhs.gov

Abstract

Background: Proteomics can identify pharmacodynamic (PD) biomarkers to support clinical pharmacology studies for biosimilar drugs development and approval. Mepolizumab and reslizumab are two interleukin 5 (IL-5) inhibitors approved for the treatment of severe asthma with an eosinophilic phenotype. Peripheral blood eosinophil count is used for dose selection in the development programs and has been discussed as a PD biomarker for biosimilar development, though variability may limit its utility.

Purpose: The aim of the study is to assess the utility of plasma proteomics for the identification of additional circulating PD biomarkers of IL-5 inhibitors.

Methodology: A discovery pilot was conducted in 266 plasma samples from 32 healthy subjects from a placebo-controlled randomized single dose clinical study with IL-5 inhibitors by the FDA. Using the SOMAscan® assay (SomaLogic, v4.1), 7288 analytes were measured at 11 timepoints over 123 days in mepolizumab (n=8 [24 mg]), reslizumab (n=8 [0.8 mg/kg]), and placebo groups (n=8). ANOVA was conducted on linear-mixed effect models regressing protein level changes with treatment, time and their interaction. Analytes with p-values< 6.82E-06 (Bonferroni-adjusted alpha) for the interaction term were considered differentially expressed. Proteins were further prioritized based on biological relevance, peak change, and area under the effect curve (AUEC) for both products.

Results: Three candidate proteins, pappalysin (PAPPA) for mepolizumab, and proteoglycan-3 (PRG-3) and follicular dendritic cell secreted peptide (FDCSP) for reslizumab were identified as differentially expressed upon treatment. PAPPA was also associated with response to reslizumab (p= 7.16E-05) and PRG-3 with mepolizumab (p= 7.41E-06), but at a lower significance threshold. Further analysis of FDCSP response to reslizumab showed that the original association was driven by variance in the placebo group over the study time. A significant difference in AUEC of PAPPA compared to placebo was observed for mepolizumab (t-test p=1.98E-02) and reslizumab (p=1.39E-04) as well as AUEC of PRG-3 for reslizumab (p=8.6E-04), but not mepolizumab (t-test p=0.19) compared to placebo.

Conclusion: Using proteomics and a discovery cohort, we identified PAPPA and PRG-3 as potential PD biomarkers of IL-5 inhibitors for future investigation.

Novel aspect

The study explores use of plasma proteomics to uncover additional IL-5 inhibitors pharmacodynamic biomarkers viz. Mepolizumab and Reslizumab. Traditional biomarkers like eosinophil count have limitations. We identifies three promising candidate proteins— PAPPA, PRG-3, and FDCSP from the proteomic analysis. After future investigation, these biomarkers may also support biosimilar drug development.

43: Lysophosphatidylcholine mediates neutrophil activity through early metabolic modulation following immunization with a live-attenuated Leishmania vaccine

Hannah L Markle, FDA/CBER; Thalia Pacheco-Fernandez, FDA/CBER; Parna Bhattacharya, FDA/CDRH; Jinchun Sun, FDA/NCTR; Nazli Azodi, FDA/CBER; Grace Kitthanawong, FDA/CBER; Caroline Hobson, FDA/CBER; Richard Beger, FDA/NCTR; Sreenivas Gannavaram, FDA/CBER; Hira Nakhasi, FDA/CBER
Hannah.Markle@fda.hhs.gov

Abstract

Leishmaniasis, caused by Leishmania parasites, spreads via sandfly vectors and blood transfusions. We are evaluating a centrin-deleted Leishmania major (LmCen-/-) parasite as a live attenuated vaccine that has shown safety and efficacy against challenge with wild type L. major (LmWT) and L. donovani in animal models. Differences in immunogenicity between LmCen-/- and LmWT infections exist; notably, neutrophil-mediated pathogenicity reported in virulent infections is absent in LmCen-/- infections. Metabolic regulation that may underlie neutrophil recruitment and functionality remains unexplored. Therefore, we analyzed metabolic reprogramming in neutrophils and their role in LmCen-/- immune protection.

C57Bl/6 mice were intradermally infected with LmWTmCherry or LmCen-/-mCherry. 48 hours post-infection, 2.5-3x106 parasitized and non-parasitized neutrophils were sort-selected from draining lymph nodes (dLNs). Untargeted metabolomic analyses were performed on neutrophils by mass spectrometry. Neutrophil migration to dLNs was measured via flow cytometry at 24, 48, and 72 hours post-intradermal inoculation with LmWT and LmCen-/-. Transwell assays were performed on parasitized neutrophils to measure migration in the presence and absence of exogenous LysoPC. Simultaneously, in vitro phagocytosis and migration were measured by flow cytometry. Neutrophil activation with and without LysoPC supplementation was measured in the ear and dLNs via flow cytometry.

Mass spectrometry revealed significant enrichment of the bioactive lipid, lysophosphatidylcholine (LysoPC) in neutrophils isolated from LmCen-/--infected mice compared to naïve or LmWT-infected mice. Peak neutrophil influx to dLNs occured at 48 hours post-infection. Transwell assays revealed enhanced neutrophil migration in presence of exogenous lysoPC in both LmWT and LmCen-/- infections. In vitro infection of neutrophils revealed increased phagocytic cells and migrated neutrophils measured by LFA-1 expression in the LmCen-/- group compared to LmWT. LFA-1 expression was enhanced in the LmCen-/- group supplemented with LysoPC. In addition, neutrophils in the LmCen-/- group expressed greater CXCR2 in the ear, and greater CXCR2 and CXCR4 in the dLNs, compared to LmWT when supplemented with LysoPC.

LysoPC enrichment in neutrophils from LmCen-/- parasites may promote immune protection following immunization. Further investigating the functional role of neutrophils and the activity of LysoPC in modulating immu

Novel aspect

LysoPC enrichment in neutrophils from LmCen-/- parasites may promote immune protection following immunization. Further investigating the functional role of neutrophils and the activity of LysoPC in modulating immune responses could aid in the discovery of novel metabolic immune mechanisms and biomarkers of vaccine-induced immunity.

44: A New Generation of Reference Materials to Promote High Quality Data in Untargeted Metabolomics

Amanda Bayless, NIST/CSD; Clay Davis, NIST/CSD; Fabio Casu, NIST/CSD; Tracey Schock, NIST/CSD
amanda.bayless@nist.gov

Abstract

Quality assurance (QA) and quality control (QC) practices lend confidence to the accuracy of acquired analytical data, which is crucial for clinical applications. As QC materials, reference materials (RM) help distinguish experimental variance from systematic error occurring during sample preparation, instrument analysis, and data processing. Reference materials enable measurement harmonization and comparability across instrumentation, protocols, and laboratories since they are well-defined, stable, and homogeneous. Untargeted metabolomics approaches aim for global detection and identification of metabolites, presenting unique challenges for reference material development and measurement traceability. The complexity of the measurement hinders the feasibility of developing a reference material with a wide coverage of concentration values for these chemical constituents. Therefore, the National Institute of Standards and Technology (NIST) is developing a new generation of RMs for multi-platform (LC-MS and NMR) metabolomic analyses which aim to provide the metabolomics community with more economical reference materials with qualitatively characterized metabolite profiles for comparison, benchmarking, and harmonization of chemical identification. This process has the benefit of accelerating the RM development and certification procedures in delivering a fit-for-purpose material for the community at a reasonable price. Metabolomics RMs currently under development include human biofluid-based materials (urine and plasma), which are commonly used in clinical diagnostics with benefits of being non- or minimally invasive and contain a wide coverage of the human metabolome. Metabolomics RM development also includes tissue (liver) and an alternative biological matrix containing a complex microbial system (feces). Each material will be developed as a suite to include multiple phenotypes with distinct metabolic profiles to facilitate differential analysis, which can provide relative quantitative values, and will include reference datasets providing highly confident identifications of molecular components to validate analytical workflows and data processing software. The development of unified multi-platform QA/QC tools that include associated reference data was prompted by an urgent need in the community to increase measurement reproducibility while also improving transparency in scientific data reporting. The future goal is to reduce uncertainty within experimental workflows and enhance confidence in the results obtained from untargeted metabolomic studies.

Novel aspect

This new generation of RMs will contain associated lists of confident annotations which have been measured using multiple platforms, multiple instruments of the same platform, multiple annotation algorithms, and multiple libraries. The convergence of metabolite and lipid annotations across methods provides confident and novel benchmarks for untargeted workflows.

45*: High resolution MALDI imaging mass spectrometry to assess spatial lipidomics of mouse fetal neural tube defects after maternal opioid exposure

Dustyn Barnette, FDA/NCTR; E. Ellen Jones, FDA/NCTR; Pravin Kaldhone, FDA/CBER; Grace Lee, FDA/CDER; Kelly Davis, FDA/NCTR; Sumit Sarkar, FDA/NCTR; Pritpal Malhi, FDA/CDER; J. Edward Fisher, FDA/CDER; Joseph P. Hanig, FDA/CDER; Richard D. Beger, FDA/NCTR; Richard D. Mellon, FDA/CDER; Amy L. Inselman, FDA/NCTR
Dustyn.Barnette@fda.hhs.gov

Abstract

In 2015, FDA released a Drug Safety Communication regarding a possible link between opioid exposure during early pregnancy and an increased risk of neural tube defects (NTDs) based on previous reports. At the time, FDA did not make new recommendations for opioid use during pregnancy due to incomplete maternal toxicity data and limitations in human and animal studies. Since then, FDA scientists have conducted multiple comprehensive studies designed to determine whether opioid-induced maternal toxicity is associated with fetal NTD development. In this study, lipid changes in mouse fetuses following exposure to morphine (100 or 400 mg/kg BW), the positive control valproic acid (VPA) (300 or 500 mg/kg BW), or the vehicle negative control were evaluated using matrix assisted laser desorption ionization imaging mass spectrometry (MALDI IMS). MALDI IMS is a mass spectrometry-based approach that provides the distribution and localization of an analyte(s) of interest across an organ or whole-body tissue section. Following maternal exposure to the drugs on gestational day (GD) 8, whole-body mouse fetal sagittal sections with and without treatment-related NTDs were analyzed using MALDI IMS on GD 18. Differential lipid distributions related to dose and exposure were identified for several phosphatidylcholine (PC) classes in the fetal brains, including PCs 34:1, 34:2, and 36:2, all of which have been previously associated with hypoxia. Follow-up high resolution imaging of horizontal sections revealed regional increases and decreases in PC levels in the cerebral cortex, thalamus, hypothalamus, and hippocampus. Additionally, an increase in the distribution of lyso PC 16:0 was observed across the brain with drug exposure. Lipid identities were confirmed with collision-induced dissociation (CID) for analyte fragmentation. MALDI images were also aligned to hematoxylin and eosin (H & E), Cresyl violet, and Fluoro-Jade C (FJC) staining of serial sections to map these distributions to histopathology. These findings represent the first MALDI IMS study of whole-body fetuses with opioid exposure during pregnancy. The observed changes in lipid distribution within fetal neural tissues and drug related NTDs associated with maternal opioid exposure provide leads to investigating possible mechanisms underlying opioid-induced NTDs and further support hypoxia as a likely part of the molecular mechanism.

Novel aspect

These findings represent the first MALDI IMS study of whole-body mouse fetuses with opioid exposure during pregnancy.

48*: Implementing a high-throughput nanoflow proteomics workflow using a dual-trap single column approach with in-plate resuspension of peptides

Benjamin Neely, NIST; W. Clay Davis, NIST
benjamin.neely@nist.gov

Abstract

Recent years have seen an explosion in sample sizes, not just in large consortia studies like the The SCALLOP consortium (Systematic and Combined AnaLysis of Olink Proteins) or UK Biobank Studies that use 10 000s of patient samples, but in fields like single-cell proteomics that require analyzing many 100s of samples per treatment. For this reason, there is a continuing effort to scale analyses to the 100s and 1000s (and beyond) sample scale. In mass spectrometry-based proteomics the primary limitation is how to operate the liquid chromatography system over time to avoid replacing columns, while also running fast enough to avoid re-calibration mid-run. One recent solution is the dual-trap single column approach that essentially operates LC steps in parallel such that one sample is loaded while the other sample is eluted onto the mass spectrometer. We have implemented this system at NIST using nanoflow to preserve high-sensitivity for single-cell applications. Likewise, we have demonstrated the autosampler can successfully resuspend dried peptides in wells immediately before applications. This allows our lab to receive pre-digested and cleaned samples from remote collaborators and run them with minimal effort. Though we detect approximately 50 % fewer proteins than our typical 10 samples per day method on our system (i.e., 2000 instead of 4000 proteins from a HeLa digest, and 200 instead of 400 proteins from undepleted plasma), we are now able to run nearly 60 samples per day in a robust manner. Overall, this nanoflow dual trap single column setup allows for ongoing and future studies benchmarking single-cell proteomics as well as embarking on large-scale plasma proteomics studies

Novel aspect

Benchmarking a method requisite for large-scale proteomics analysis

Data Integration & Data Management

17: Fostering Public Health Bioinformatics and Collaboration with GalaxyTrakr

Jayanthi Gangiredla, FDA/CFSAN; Hugh Rand, FDA/CFSAN; Daniel Benisatto, DRT Strategies; Justin Payne, FDA/CFSAN; Charles Strittmatter, FDA/CFSAN; Jimmy Sanders, SDS Solutions, Inc.; William J. Wolfgang, New York State Department of Health/Wadsworth Center; Kevin Libuit, Division of Consolidated Laboratory Services/Department of General Services; James Herrick, James Madison University/Biology Department; Melanie Prarat, Ohio Department of Agriculture/Animal Disease Diagnostic Laboratory; Magaly Toro, Universidad de Chile/Laboratorio de Microbiología y Probióticos; Thomas Farrell, FDA/CFSAN; James Pettengill, FDA/CFSAN; Errol Strain, FDA/CFSAN
Jayanthi.Gangiredla@fda.hhs.gov

Abstract

Background:In the United States, surveillance activity of infectious disease - foodborne, hospital-acquired, zoonotic, or otherwise – is addressed by a federated system of county, state, and national agencies managing different streams of data relatively independently. This poses a challenge to the dissemination of techniques, tools, resources, data, and analysis among these disparate groups of public health scientists, despite their aligned aims.

Purpose: The objective is to create a cloud based user-friendly Bioinformatics platform that enables scientists from public health and food safety research labs without any bioinformatics knowledge, to run queries and obtain reliable, comparable and consistent results. This helps in harmonized interpretations of WGS results across laboratories by providing tools optimized for food pathogen surveillance.

Methodology: The US Food and Drug Administration's Center for Food Safety and Applied Nutrition (FDA-CFSAN) addressed this challenge by creating GalaxyTrakr, a cloud-hosted Galaxy environment with curated tools for pathogen biosurveillance of sequencing data generated by GenomeTrakr and from other sources. A cost-effective scaling architecture in Amazon Web Services now addresses the needs of an increasing number of users executing an increasing number of jobs, exploring an increasing number of bioinformatics tools, collaborating on an increasing number of shared data sets, and developing an increasing number of formal analysis protocols based on the GalaxyTrakr platform.

Results: A number of laboratory sites, based in the US and internationally, now use GalaxyTrakr to overcome key infrastructure and resource challenges. Finally, GalaxyTrakr serves as a platform to distribute and host bioinformatics tools developed by researchers at FDA-CFSAN to users without the local capacity or ability to install and run them, and we now target GalaxyTrakr specifically as a release platform for new bioinformatics tools. These tools support coordinated analytic methods and consistent interpretation of results across laboratories.

Novel aspect

GalaxyTrakr advances food safety by providing reliable and harmonized WGS analyses for public health laboratories and promoting collaboration across laboratories with differing resources. Anticipated enhancements to this resource will include workflows for additional foodborne pathogens, viruses, and parasites, as well as new tools and services.

36: Impact of FDA's Low- or No-Cost Tech-Enabled Traceability Challenge on Strengthening Traceback Investigations

Tanya Gupta, FDA/CFSAN; Jorge Hurtado Maturana, FDA/CFSAN; Sharon Seelman, FDA/CFSAN; Adam Friedlander, FDA/CFSAN; CAPT Kari Irvin, FDA/CFSAN
Tanya.Gupta@fda.hhs.gov

Abstract

Since 2020, the FDA New Era of Smarter Food Safety initiative has highlighted the benefits of operationalizing tech-enabled traceability among food supply chain stakeholders. Throughout the summer of 2021, the New Era Technology team and precisionFDA hosted the Low- or No-Cost Tech-Enabled Food Traceability Challenge. A diverse array of industry stakeholders across many disciplines were encouraged to share traceability solutions addressing how data associated with food shipments could be more easily tracked throughout the supply chain at little or no cost to the end user. The Challenge garnered 90 submissions internationally, and 12 teams were proclaimed as winners later that year. Since the Challenge concluded, the FDA has collaborated with organizations to encourage adoption and innovation of these software, hardware, and data analytics technologies. For example, FDA contracted with a non-profit organization to produce an independent report analyzing the outcomes and themes from the Challenge and is currently producing a video series with participants from the Challenge. For this poster presentation, the FDA will highlight the methodologies (such as defining the scope of the Challenge and scoring processes), goals, outcomes, and benefits from the hosting this event. This poster will conclude with a discussion about how industry's innovative traceability technology solutions will improve data harmonization and data interoperatiility and can strengthen FDA's foodborne outbreak investigation processes.

Novel aspect

This poster summarizes the benefits of FDA hosting the 2021 FDA Tech-Enabled Traceability Challenge, the various resources that have been created to help industry stakeholders adapt tracing technologies since the Challenge concluded, and the impact the Challenge has on improving traceback investigations.