Epigenetic Characterization and Observation (ECHO)

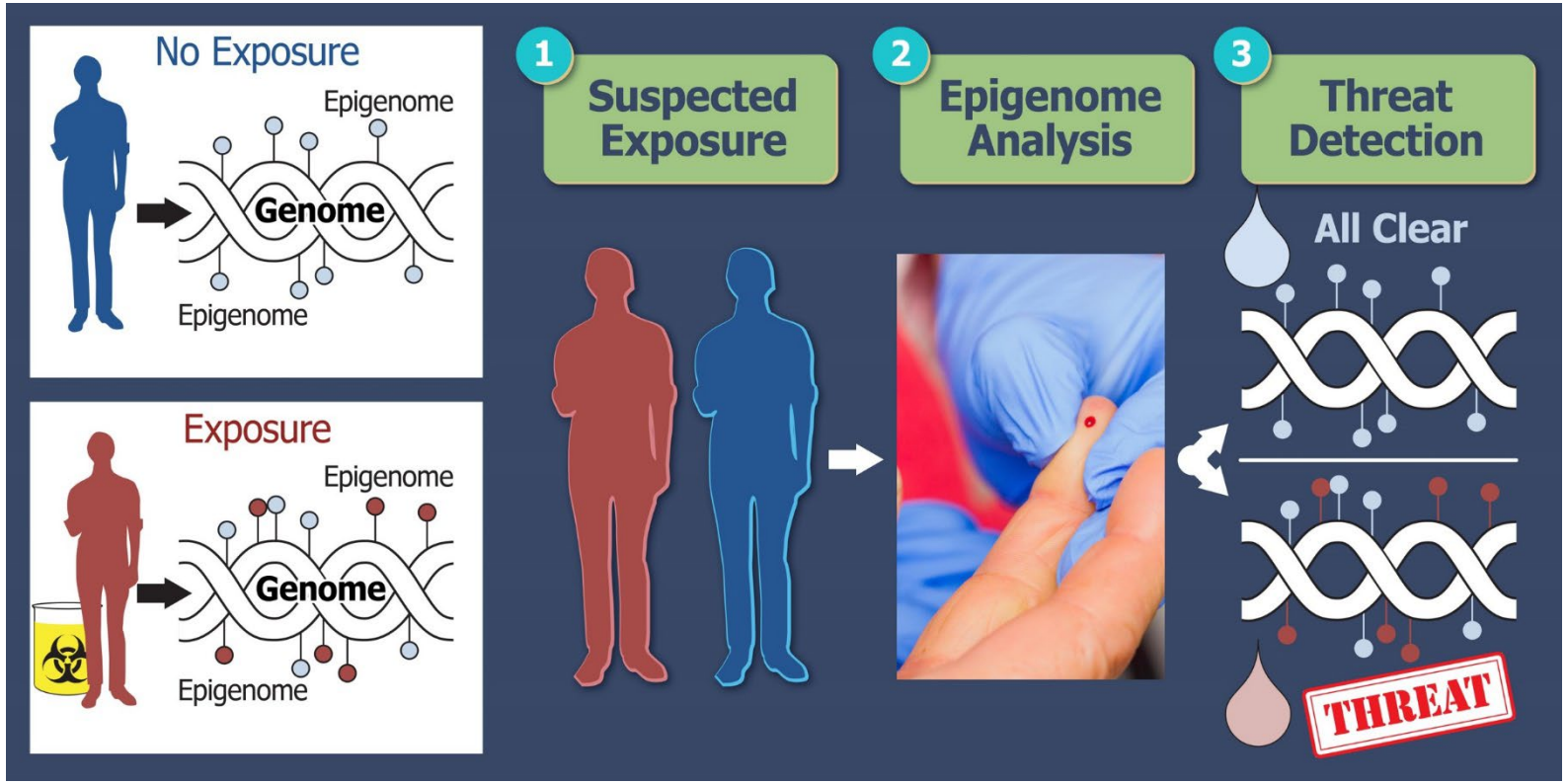
Jean-Paul Chretien, Ph.D.
Program Manager, BTO

September 7, 2023

Epigenetic Characterization and Observation (ECHO) Program

DoD Problem: Inability to rapidly and accurately identify exposure history of individuals exposed to CBRN threats is a big gap in our military's forensics and diagnostics ability impacting for national security

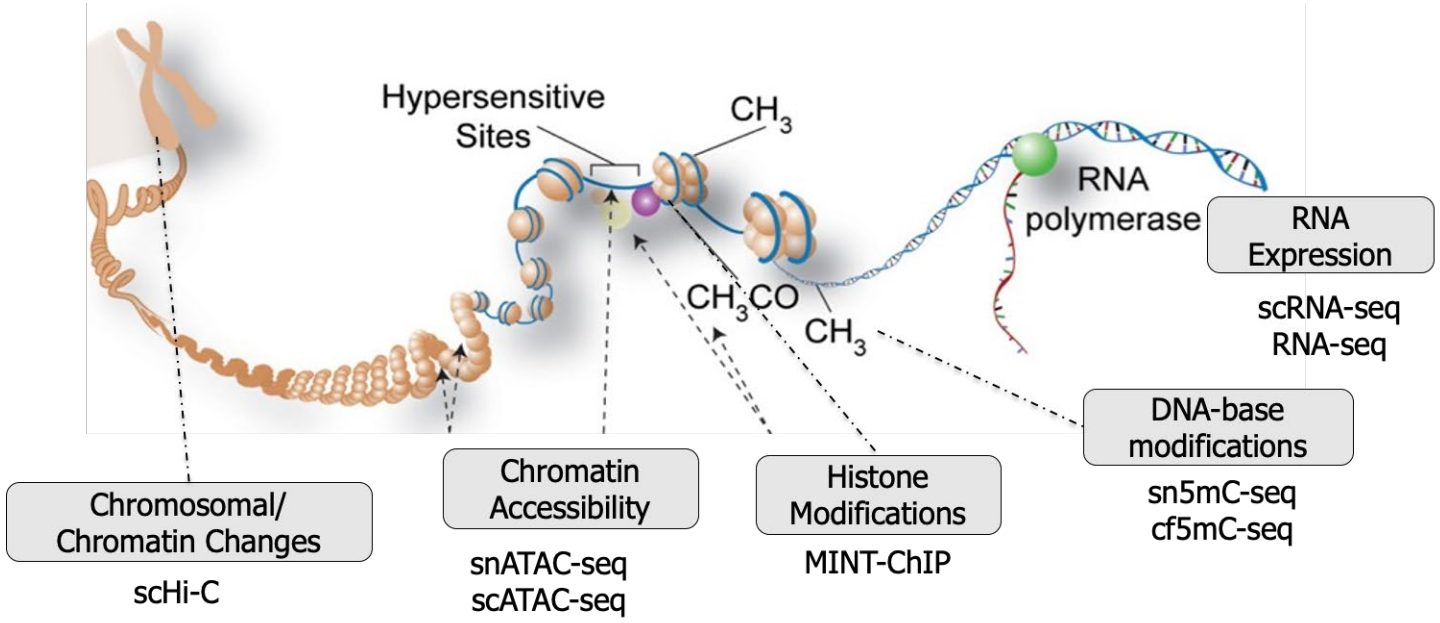
(U)



(U)

Vision: Attribution and diagnostics from a specific, temporal, human signature using the epigenome as the body's record keeper which can be obtained quickly with a field deployable platform in 30 minutes or less

ECHO: Combining Molecular Methods for Epigenetic Diagnosis



Exp #	Exposure	Source	Samples	Signatures	
1	HIV	GOV'T & DUKE	✓	✓	Months 1-6
2	MRSA	DUKE	✓	✓	
3	*B. anthracis* vaccination	BATTELLE	✓	✓	
4	Organophosphates (Chlorpyrifos)	DUKE	✓	✓	Months 6-12
5	SARS-CoV-2	ISMMS-DUKE-GOV'T	✓	✓	
6	Influenza H3N2	GOV'T	✓	✓	
	Fentanyl	ISMMS			
7	Fentanyl	GOV'T	✓		
8	Explosives (ANFO and PETN)	GOV'T	✓	✓ ✓	
9	Radiation (gamma)	ISMMS	✓		Phase 2
	EBV-convalescent	DUKE	✓		
10	EBV-vaccine	DUKE	✓		
11	*B. pseudomallei*	DUKE	✓		
12	*B. burgdorferi*	JHU	✓	✓	

ECHO <u>human</u> exposure samples

- Total number of samples committed to ECHO: **<u>6,724</u>**
- Total number of samples collected/ collecting/ on-hand: **<u>5,942</u>**
- **Longitudinal** collections have been given priority
- Matched geographically and ethnically **diverse control samples**
- Utilizing human samples from **other programs and research efforts** to compare against for **confounds**

ECHO sample groups

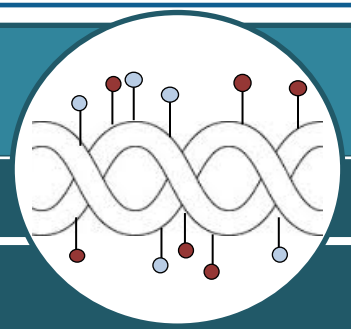
7

2

1

2

ECHO Technical Area 1: Metrics and Pressure Tests



TA1: Epigenetic Signature Identification

	Phase I (1-24 Mo)	Phase II (24-48 Mo)
Deliverables	1. Epigenetic data from molecularly analyzed samples 2. Signatures for WMD related exposures (**5 total**) 3. Algorithms for specific identification of exposure profiles and temporal resolution of last exposure event	1. Validated analytical algorithms 2. Signatures for WMD related exposures (**12 total**) 3. Finalized computational toolkit integrated into field forward system
Pressure Tests	**6 mo** – Molecular dataset release (**1** virus & **1** bacteria dataset per team) **12 mo -** Signature release (**5 total**; **2** viral**, 2** bacterial & **1** non-biological WMD exposure signatures per team) **42 mo -** Target pressure test (bacterial vs. viral differential distinction) **42 mo -** Multiple target pressure test (WMD, **+/- 1 year** temporal resolution of last exposure)	**36 mo** – Expansion of signatures, **7** additional signatures per team, algorithm and signature improvements to achieve **65% PPV** **48 mo** –Continued algorithm and signature matching improvements to achieve **85% PPV**

ECHO Technical Area 2: Metrics and Pressure Tests

TA2: Deployable Platform Development

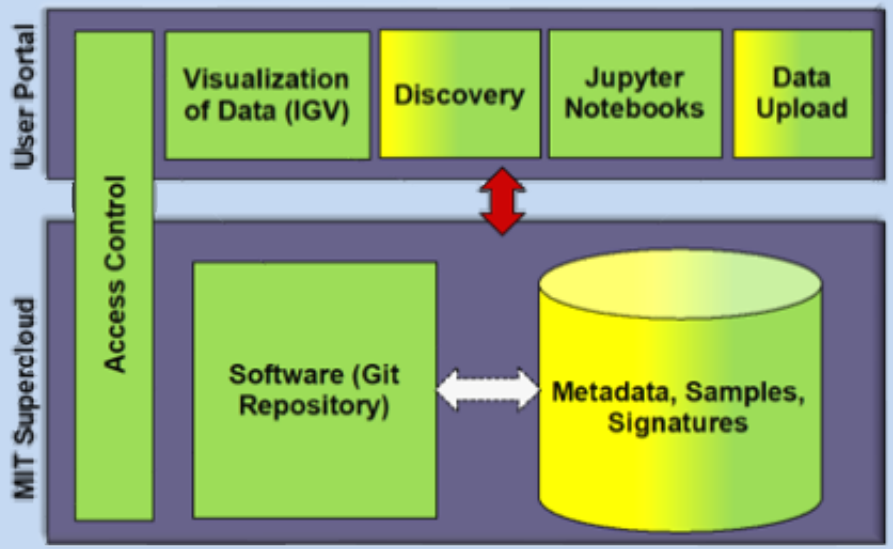
	Phase I (1-24 Mo)	Phase II (24-48 Mo)
Deliverables	1. Sample collection and preparation system, air gap permitted 2. Onboard computational system that implements spectral matching algorithms 3. Deliver and functional testing of system module prototypes, with preliminary design review of final system	1. Sample-answer device with **zero air-gaps** and onboard computational capability 2. Size (**1 ft^3**), weight (**< 10 lbs**), and power (**< 20 W**) footprint equivalent to today's POC diagnostics systems 3. Sample to answer time **< 30 min**
Pressure Tests	**18 mo –** Molecular reaction development at large-scale, incorporating the pre-sequencing preparation for the 12+ emerging epigenetics analysis methods; QoS – **100%** **24 mo –** Assembly and separate functional testing of small-scale system modules: 1) nucleic acid extraction module, 2) pre-sequencing preparation, 3) sequencing; QoS – **100%** with intermittent coverage	**30 mo –** Transition of all molecular analysis steps to the small scale system, air-gapped demonstration of epigenetic tests across all system modules; QoS – **50%** with intermittent coverage **36 mo –** Epigenetic test of all modules with **<50,000 cells**, **65% PPV;** QoS – **25%** with intermittent coverage **48 mo –** System Demonstration, **<50,000 cells**, **under 30 min**, **85% PPV**, QoS – **10%** with intermittent coverage with minimally-trained operators

ECHO Performer Teams

Performer	Focus of work
Icahn School of Medicine at Mount Sinai (ISMMS)	• TA1 and TA2 Performer • Sourced samples from human exposures • Performed epigenetic sequencing • Develop epigenetic signatures by using ML algorithms • Develop device for deployment in austere conditions
Duke University	• TA1 Performer • Sourced samples from human exposures • Performed epigenetic sequencing • Develop epigenetic signatures by using ML algorithms
Battelle Memorial Institute	• TA1 Performer • Sourced samples from human exposures • Develop epigenetic signatures by using ML algorithms

ECHO Independent Validation & Verification

MITLL-ECHO Knowledge Platform Database



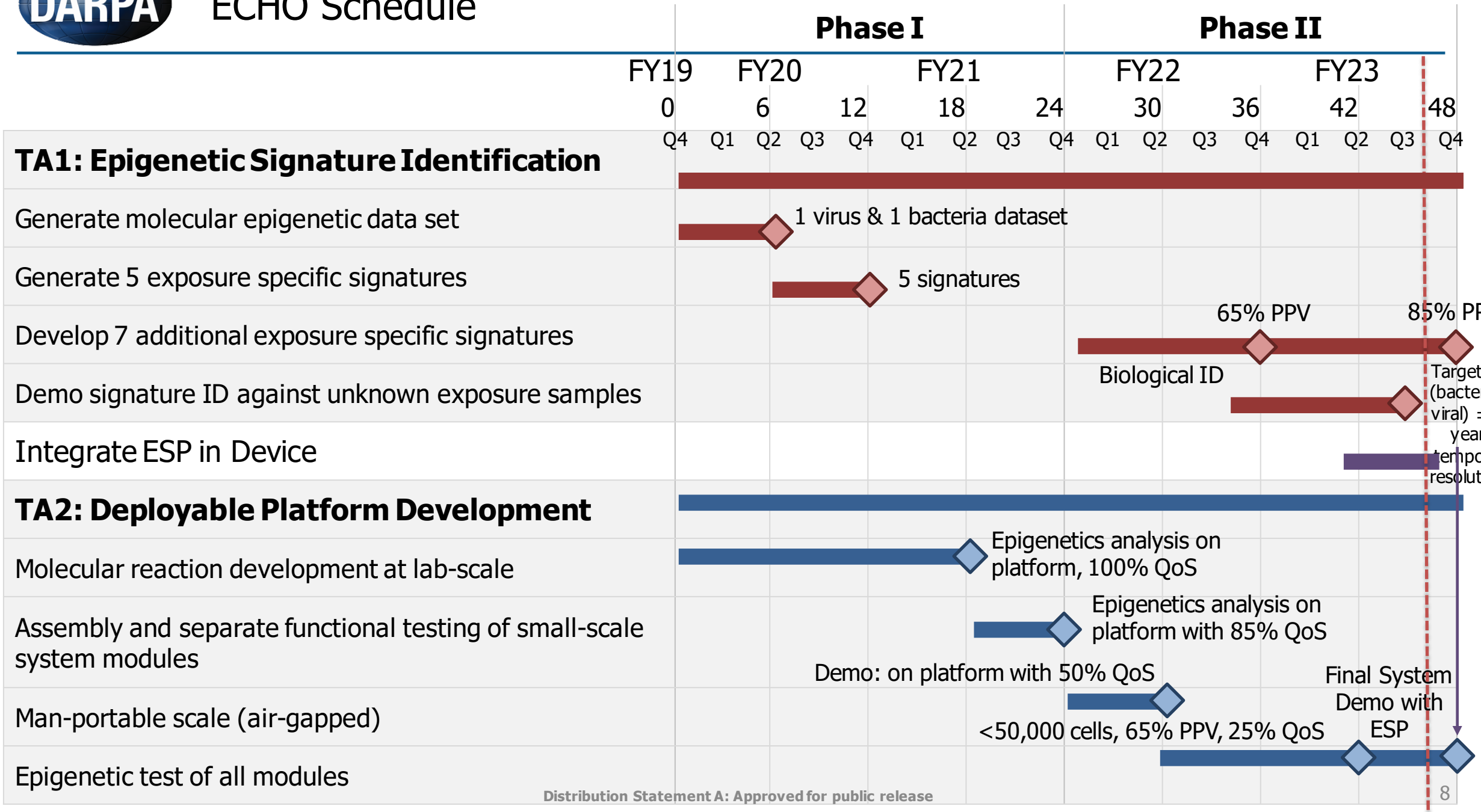
- Over 1.5 PetaByte of epigenetic datasets (1,467 datasets) across 12 different epigenetic protocols (bulk and single cell)
- Method-Specific QA/QC for all 12 epigenetic approaches under ECHO, improving on SoA (ENCODE) developed using Jupyter

Walter Reid Army Institute of Research

- MHRP (U.S. Military HIV Research Program) sourced human exposure samples to help start the ECHO program (POC COL Julie Ake)
- DCB (Diagnostics and Countermeasure Branch) worked together with the DARPA team for the ECHO blinded test design and execution (POC Sheila Peel)
- Blinded test samples to be sent to performer teams in August; Teams will be asked to discriminate infection from non infection, pathogen that caused infection and time since exposure

The ECHO program has developed an extensive IV&V network with the intention of secure sharing of high-quality ECHO-developed data through rigorous QA/QC across the performer teams and stakeholders (DTRA, FDA)

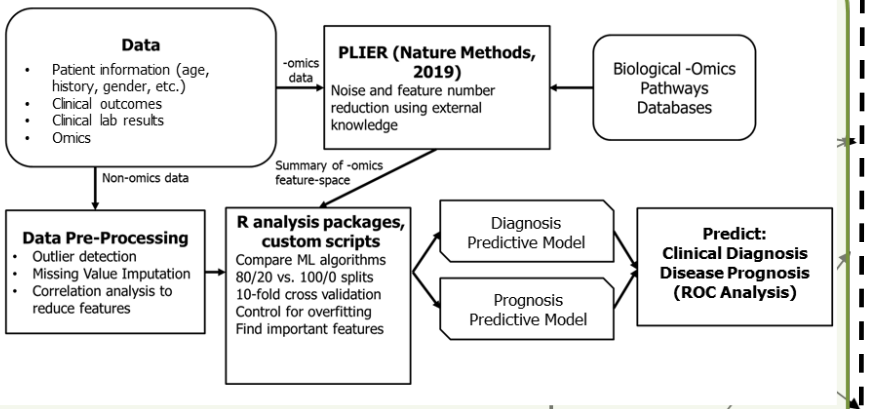
ECHO Schedule

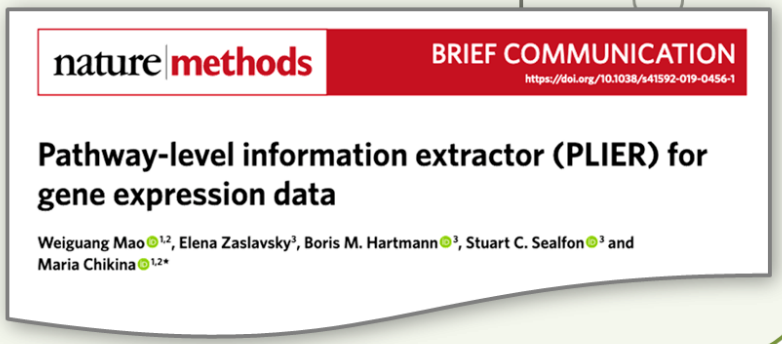


S. aureus ECHO exposure

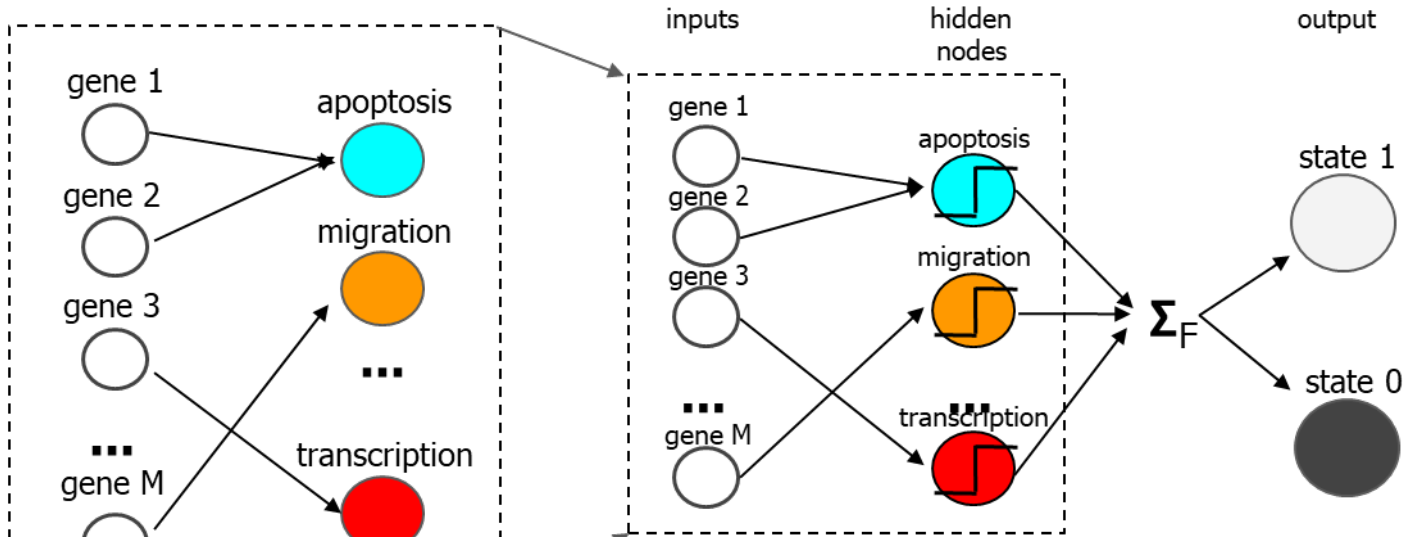
BioNNET/PLIER – AI Neural Network developed for ECHO

Machine Learning Pipeline





ECHO-Developed BioNNET



-**complexity reduction**: less inputs, less weights

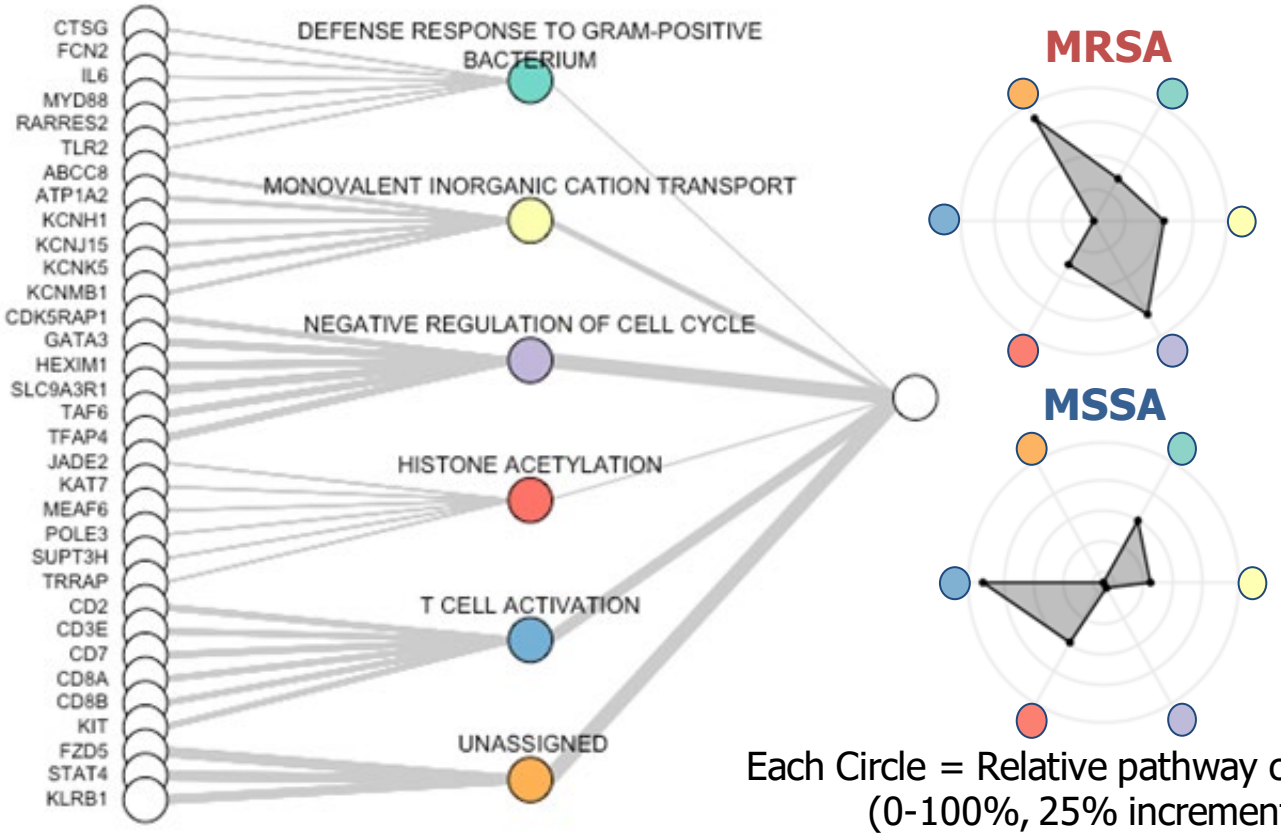
-**interpretation:** nodes have clear biological roles

-**modularity**: builds from established –omics tools

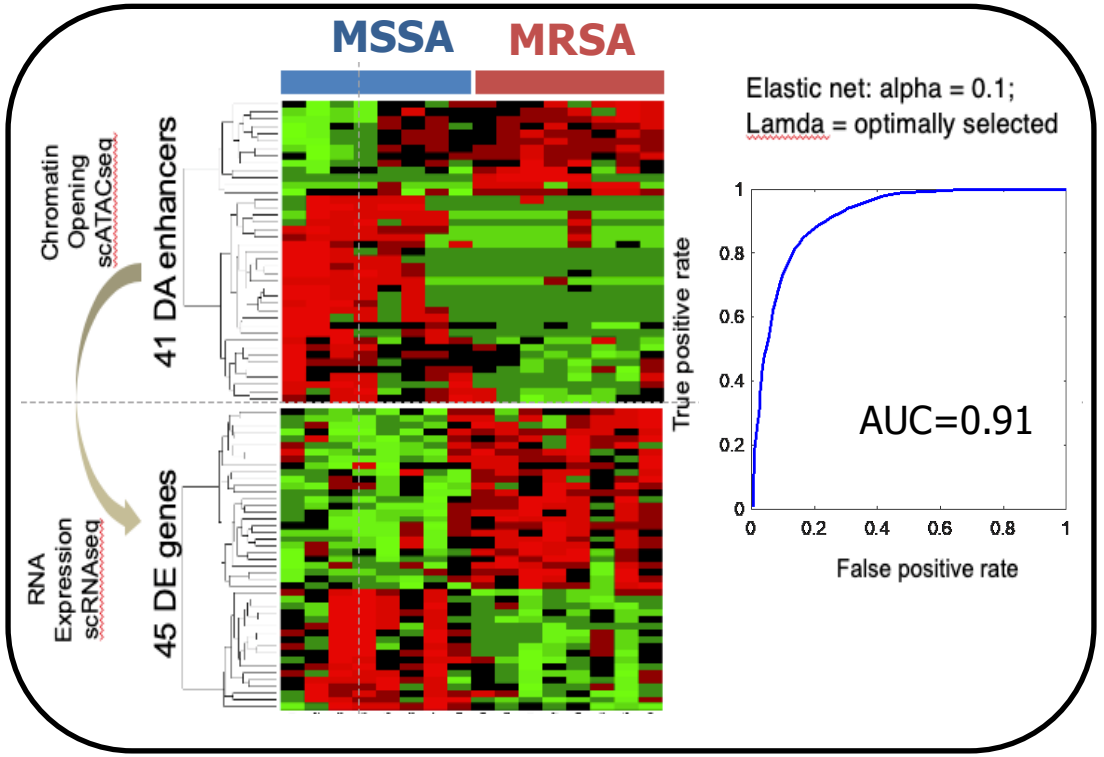
BIONNET/PLIER is a modular approach which builds on established tools for -*omics* data analysis to identify networks and genes in human exposure samples

Single cell epigenetics and unique AI analytics identify distinct MSSA and MRSA Host Response Pathways

Staphylococcus aureus strains MRSA and MSSA are virtually indistinguishable requiring time consuming bacterial blood culture to discern with major implications to treatment plans in our military



Each Circle = Relative pathway correlation (0-100%, 25% increments)



N=8 per group, 8192 cells per subject

Chen X et al, Nat Comp Science 2023

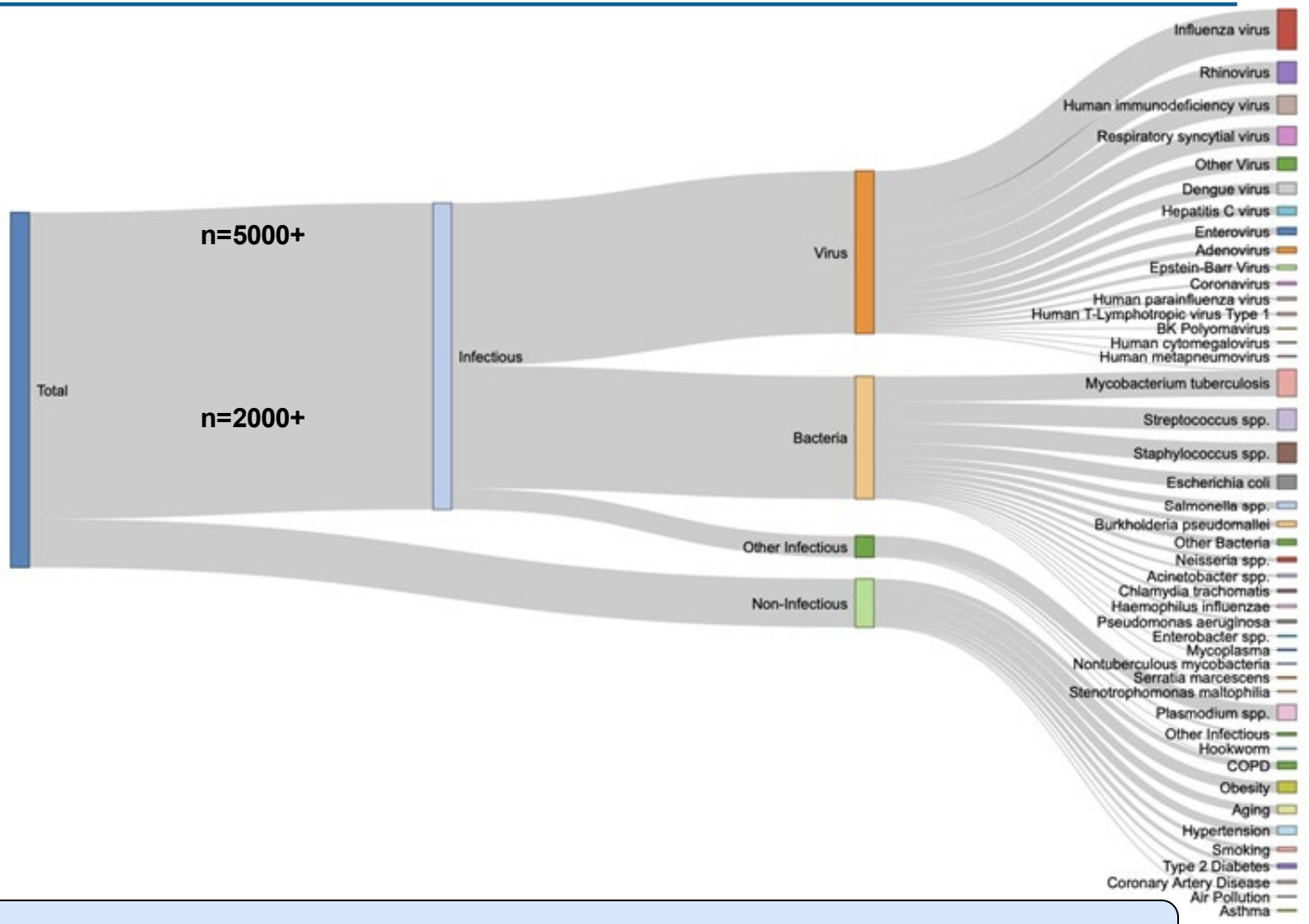
Single cell epigenetic approaches identify host networks and genes that discriminates MRSA from MSSA even in small sample sizes

Curated Compendium to assess specificity of ECHO developed signatures

- curated a compendium of blood gene expression data on viral, bacterial, parasitic, and non-infectious conditions (e.g., COVID-19 risk factors) from human subjects *in vivo*

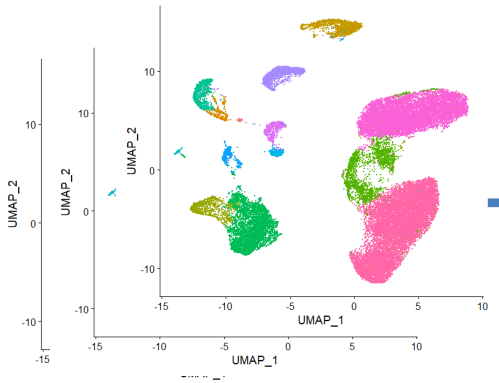
- retrieved from Gene Expression Omnibus (GEO) and pre-processed with a consistent pipeline for most Illumina & Affymetrix arrays

- the compendium enables to validate signature's robustness and pathogen-specificity
  - **16,677 samples**
  - **170 studies**
  - **15+ viruses, 17+ bacteria, 3 parasites, 9 non-infectious conditions**

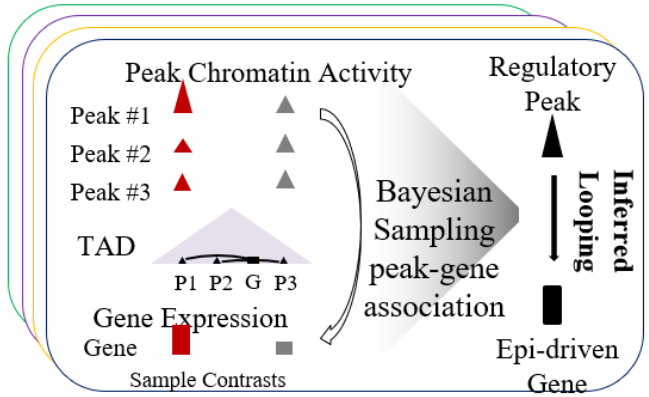


Compendium Emerged as a critical asset for development and validation of ECHO signatures

Specificity of ECHO developed signatures



Single cell RNA/ATAC MRSA/MSSA Data

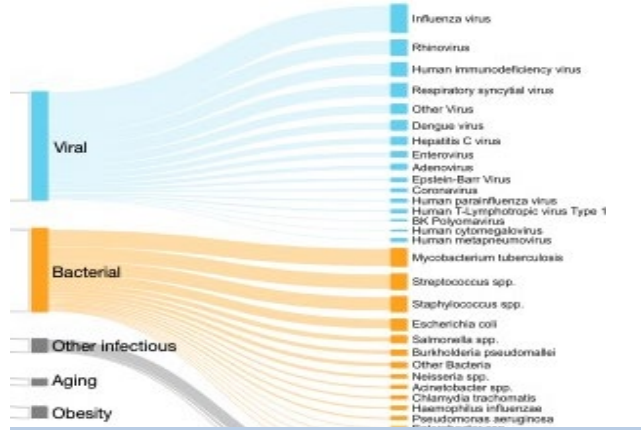


ECHO developed AI pipeline down-sampling

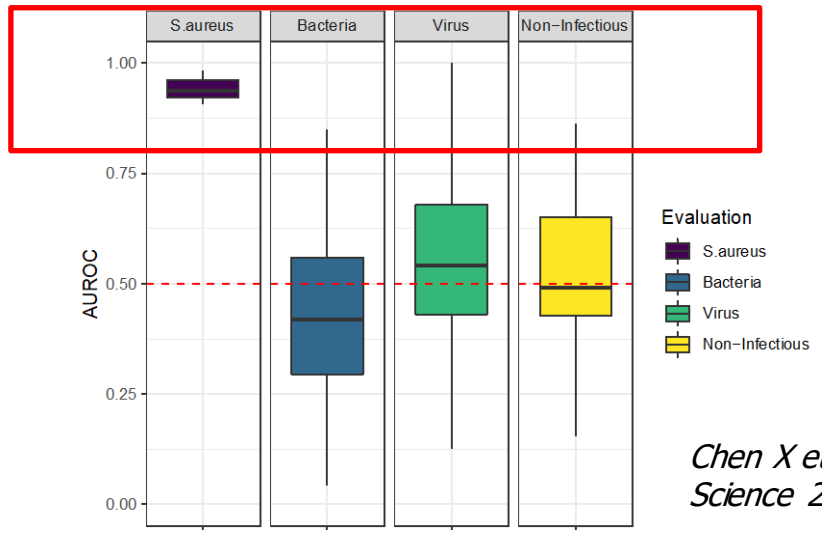
Chromatin site-RNA Signature

Robustness of Signature on data

Signature Specificity Module



Specificity test on massive ECHO data compendium

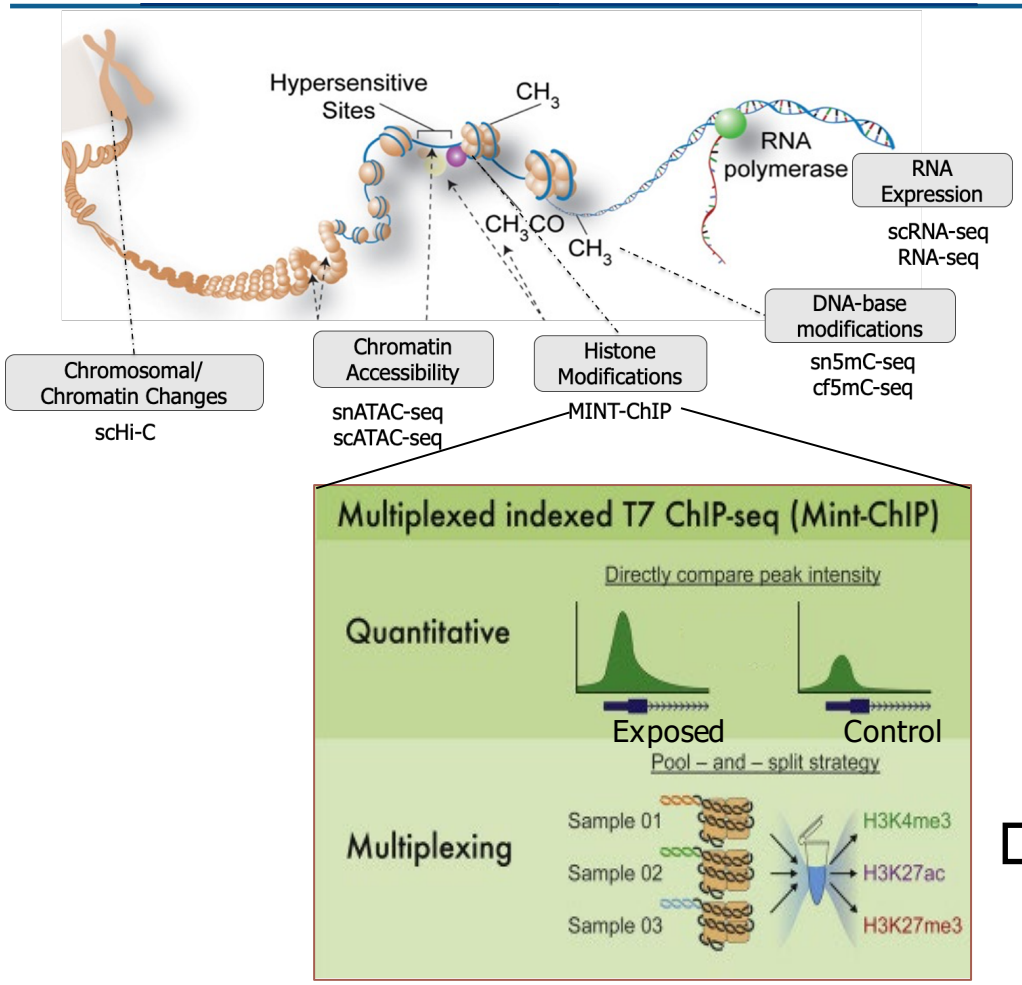


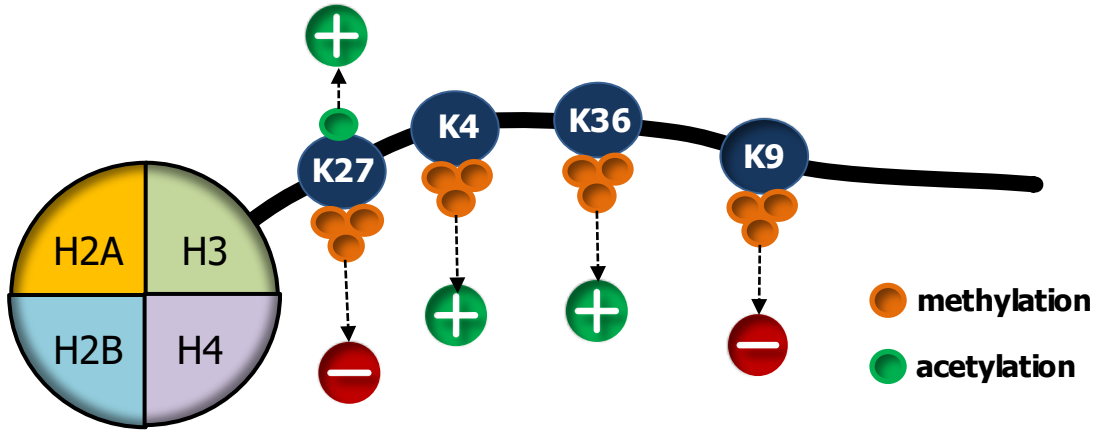
Chen X et al, Nat Comp Science 2023

S. aureus ECHO Signature accurately and specifically classifies 12,202 samples from 138 studies

Organophosphates (Chlorpyrifos) ECHO exposure

Profiling chromatin modifications across human exposures: MINT-ChIP



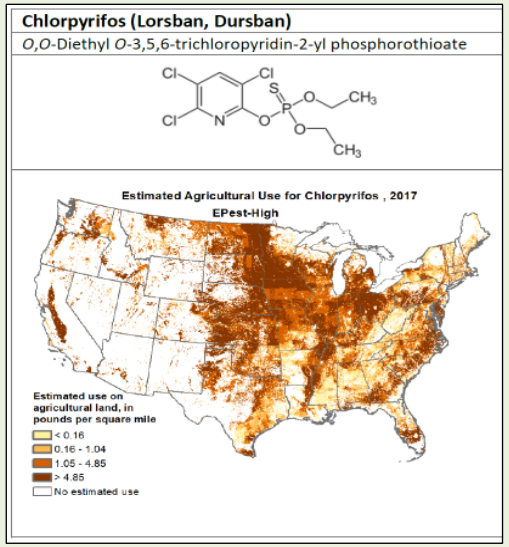


Histones	Transcription	Function
H3K27ac	+	Active transcription
H3K4me1	+	Active poised enhancer
H3K4me3	+	Active promoter
H3K27me3	−	Hereditably repressed genes
H3K36me3	+	Hereditably active genes
H3K9me3	−	Hereditably repressed genes

ECHO performer Duke deploys the next generation of chromatin IP methods requiring 1000x fewer cells to assess histone marks in human samples

OP exposure induces stable histone modification changes in chromatin

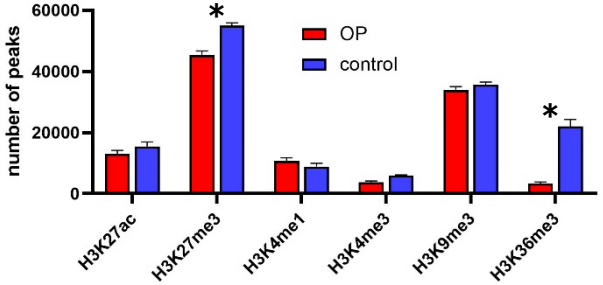
Human OP exposure samples (Chlorpyriphos - TCPY)

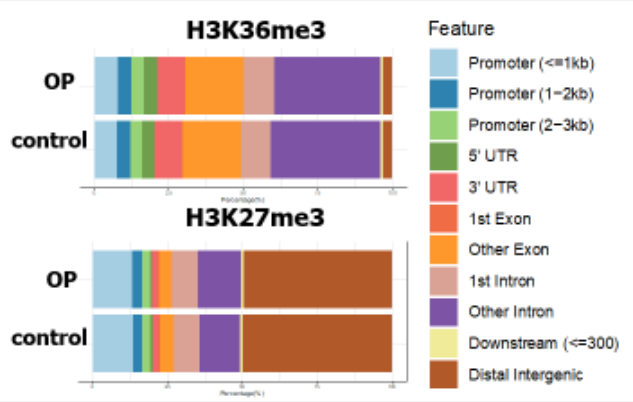


- TCPY exposure in farm workers followed for 2 months (**N=30**) and profiled by MINT-ChIP
- OP insecticides are related to nerve agents and share same mode of action with V-series nerve agents
- Most widely used OP in agriculture; Highly toxic oxon formed *in vivo*

MINT-ChIP identifies histone modifications enriched in TCPY exposure

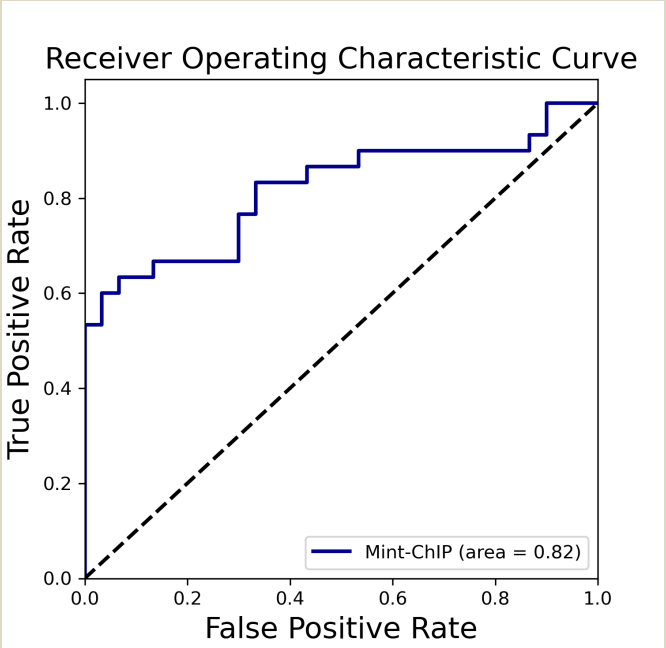
Significantly enriched peaks show global **H3K36me3** and **H3K27me3** reductions specific to OP exposure





Enriched peaks were mapped to genetic elements, confirming predicted biological function: **H3K36me3** marks elongation of exons, and **H3K27me3** marks closed promoter and enhancer regions

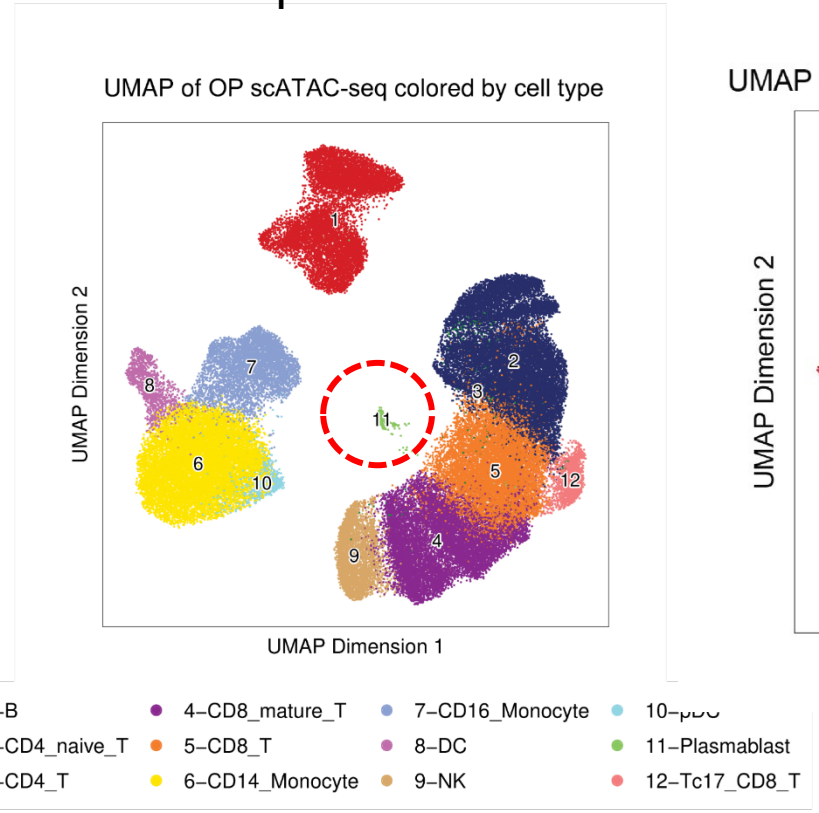
Human OP exposure signature

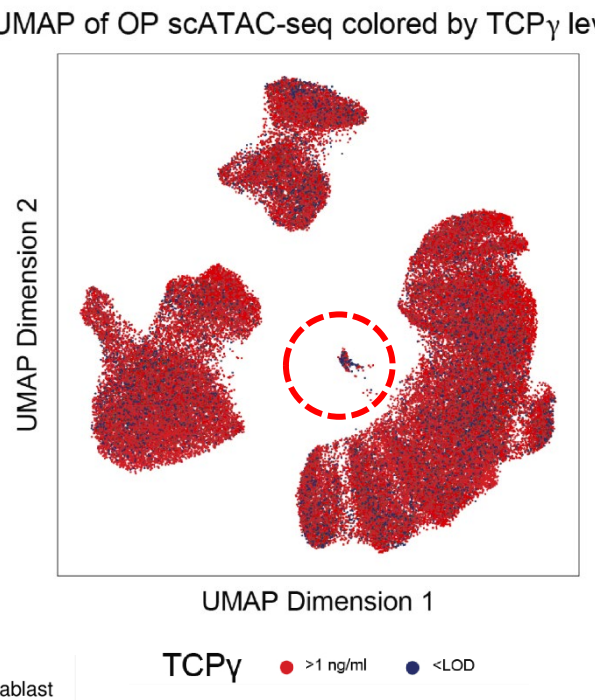


OP exposure elicits strong epigenetic mark changes at the chromatin histone level which identifies exposure with 82% performance compared to non-exposed control populations

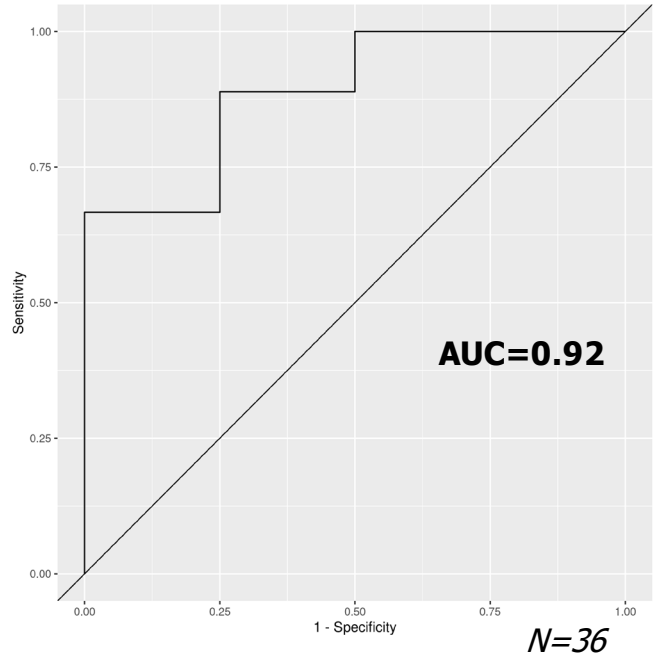
Single-cell chromatin accessibility of OP exposure identifies unique cell types correlating with OP exposure

- scATAC-seq data from 36 samples
- 12 distinct cell types found, plasmablasts only cell type not enriched with OP
- 1622 peaks associated with OP exposure; 2646 peaks associated with controls





scATAC-seq OP signature



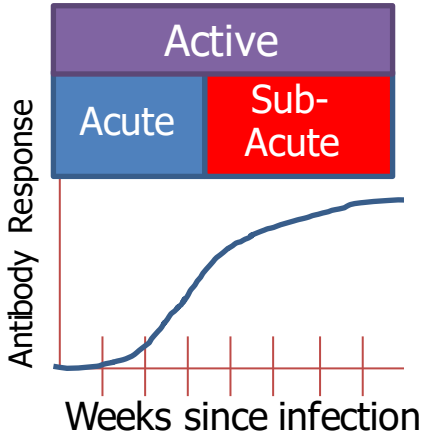
N=36

OP exposure is identified epigenetically by transposase accessible chromatin and identifies plasmablasts inversely corelated with exposure

B. burgdorferi (Lyme Disease) ECHO exposure

Program Accomplishments: ECHO delivers first in class Lyme disease assay

- B. Burgdorferi causes Lyme disease through tick bite
- Debilitating neurological condition affecting many organs; early diagnosis critical
- SoA test uses technology from 1980s (ELISA + Western Blot)
- Over 470,000 case per year (CONUS); active Lyme results in 100% disability rating (DoD Diagnostic Code 6319
- 75% of military installations (CONUS) in Lyme Disease high prevalence areas

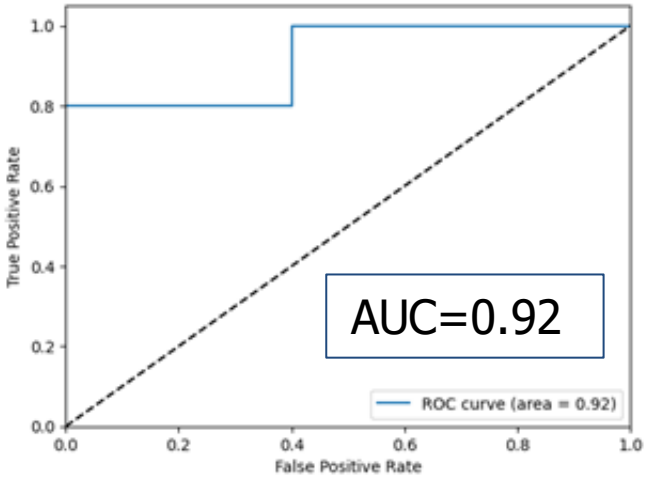


Lyme disease categories

Acute: Active Infection, within weeks of initial infection (seroconvert at second visit only)
Subacute: Active Infection, weeks after initial infection (PCR+, seropositive at first visit)
Active: Acute + subacute

ECHO simultaneous single cell RNA-seq and ATAC-seq multiome assay

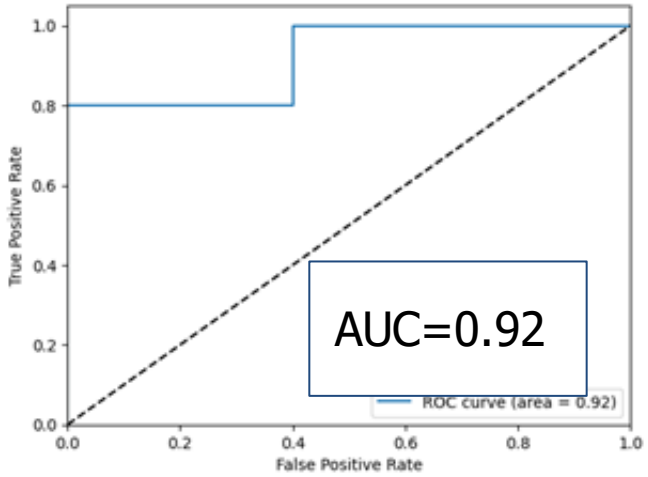
Comparison	Cell Types (Marker #)
Acute vs. Uninfected	CD4 naive (73); CD4 mem (56); CD14 mono(52)



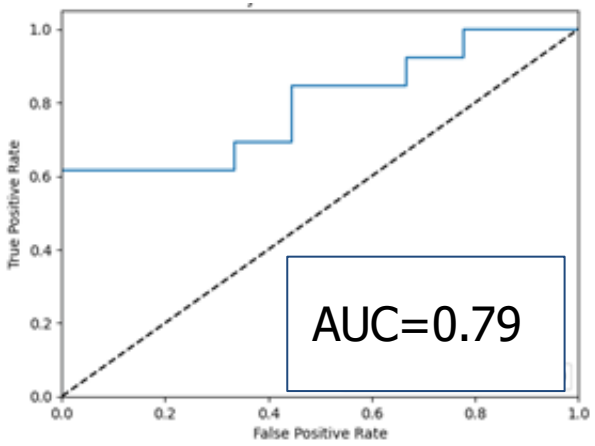
ECHO developed single cell multiome assay identifies acute Lyme from uninfected individuals and determines cell types responsible for acute Lyme disease

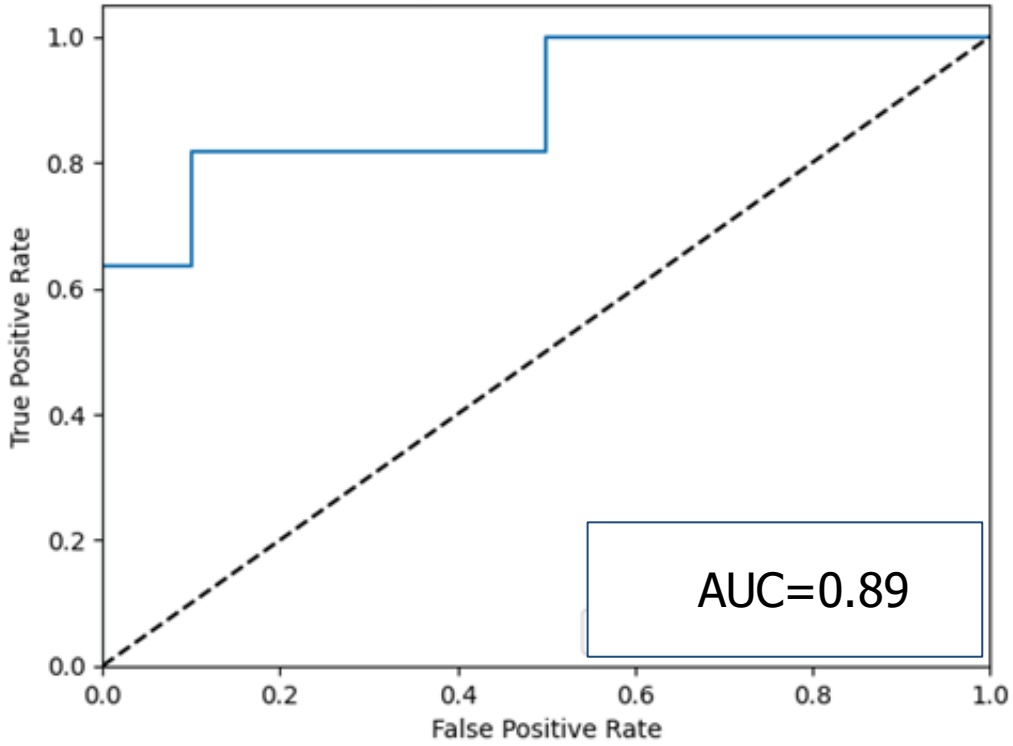
Diagnosis of acute Lyme and acute vs. subacute active infection

Held out test set validation



Apply to public data mixed 8 acute/ 14 subacute cohort





2 fold cross validation

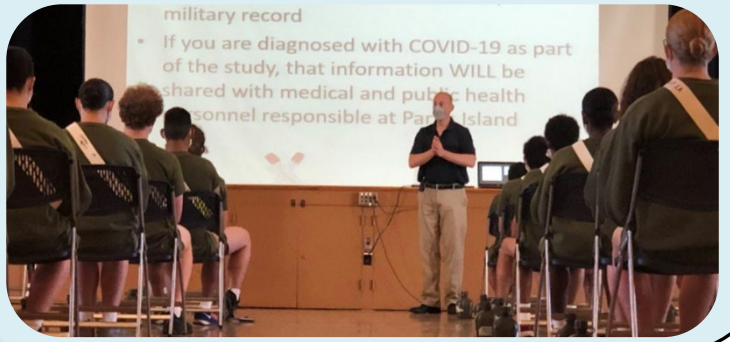
infection before positive serology and stage of infection can be identified with clinical grade sensitivity

SARS-CoV-2 ECHO exposure

ECHO COVID-19 Response: Major updates providing tangible results

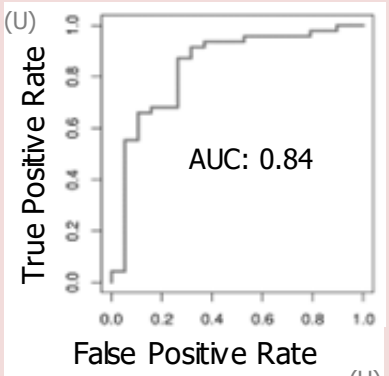
USMC COVID Health Action Research for Marines (CHARM)

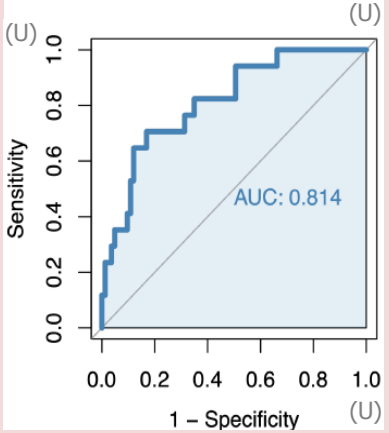
- Ensured force readiness for Parris Island Marines (> 54000 tests)
- New quarantine guidelines (from 14 to 10 days) to control COVID-19 for USMC, military, and public (CDC) (*Letizia AG et al, NEJM 2020*)
- Identified 10% re-infection attack rate and inverse correlation with protective antibody concentrations (*Letizia AG et al, Lancet Resp Med 2021*)



New Class of Host Based COVID-19 Diagnostics and Prognostics

- (U) Identified novel host-based epigenetic signatures leveraging unique biological traits of COVID-19 (RNA alternative splicing)
- (U) First host-based prognostic tool to predict COVID-19 disease outcome
  Wilk AJ et al, J Exp Med; 2021
- (U) Identified correlation between COVID-19 disease severity and single cell epigenetics
  Submitted: Immunity; posted: https://www.biorxiv.org/content/10.1101/2020.12.04.412155v1





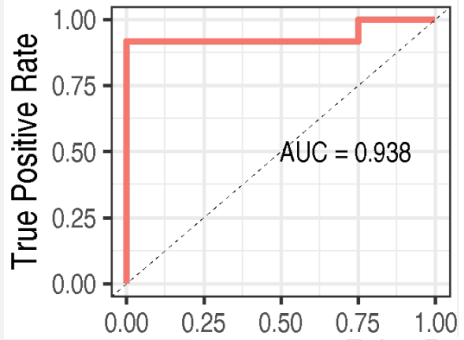
New Diagnostic Tools

- Two COVID-19 diagnostic Emergency Use Authorization (EUA) tests
- >5 million tests performed
- One EUA test in preparation (host-based test)
- New Point-of-Care, high-throughput microfluidics testing platform

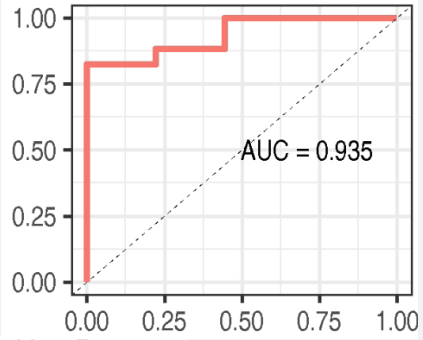
DNA methylation tracks with COVID-19 infection status, imprints genetic locations important for immunity and predicts time since infection

DNA methylation strongly correlates with COVID-19 infection status

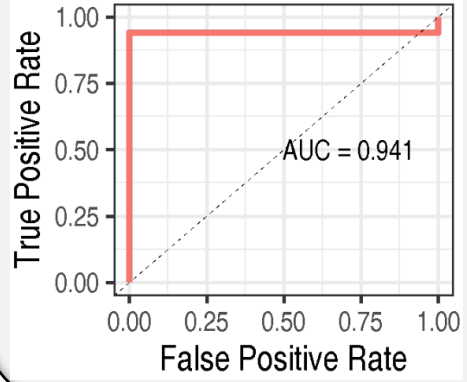
- **Infected *vs* Uninfected**



- **Infected *vs* Post infection**



- **Uninfected *vs* Post infection**



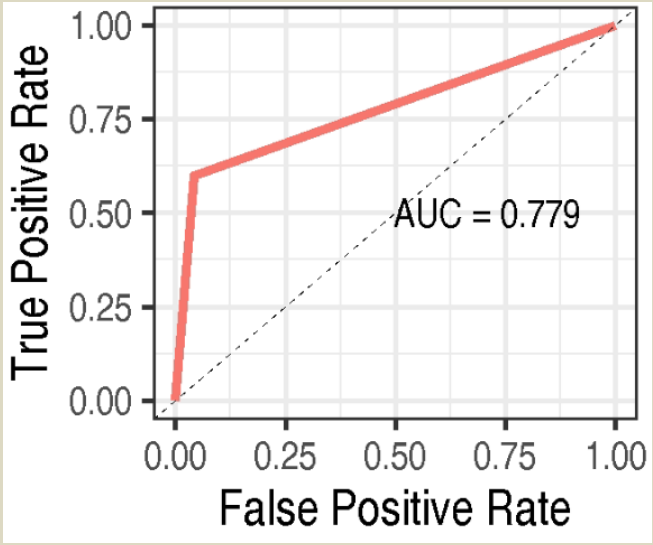
With DNA methylation, it is possible to distinguish between **all infection status scenarios** (direct-viral PCR does not cover this)

DNA methylation data: 31 Uninfected, 59 Infected, 67 Post infection

Mao W et al, Mol Systems Biol 2023

DNA methylation predicts time since infection with SARS-CoV-2

- Time since **first diagnosis ≥ 28 days**
- Pathways reflected by epigenetic response are unique to different time periods-post-infection
- Based on 31 uninfected, 59 infected



Mao W et al, Mol Systems Biol 2023

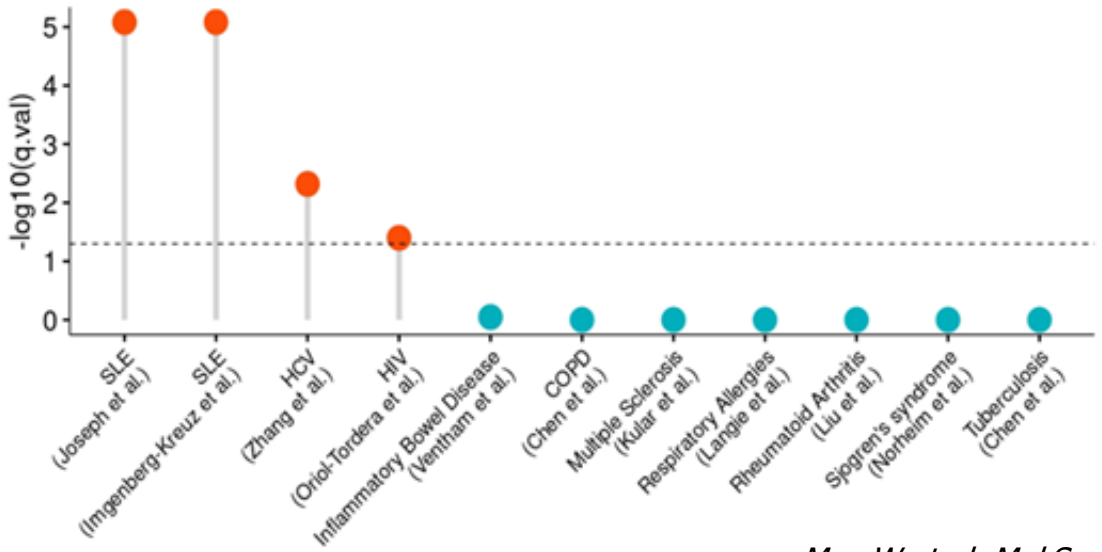
Epigenome (methylation) indicates that the host immune response (Interferon) continues long after acute infection

If only RNA-seq data was used, it would lose information (Interferon response gene returns to normal). **Only with DNA methylation is there a clear differentiator of interferon response**

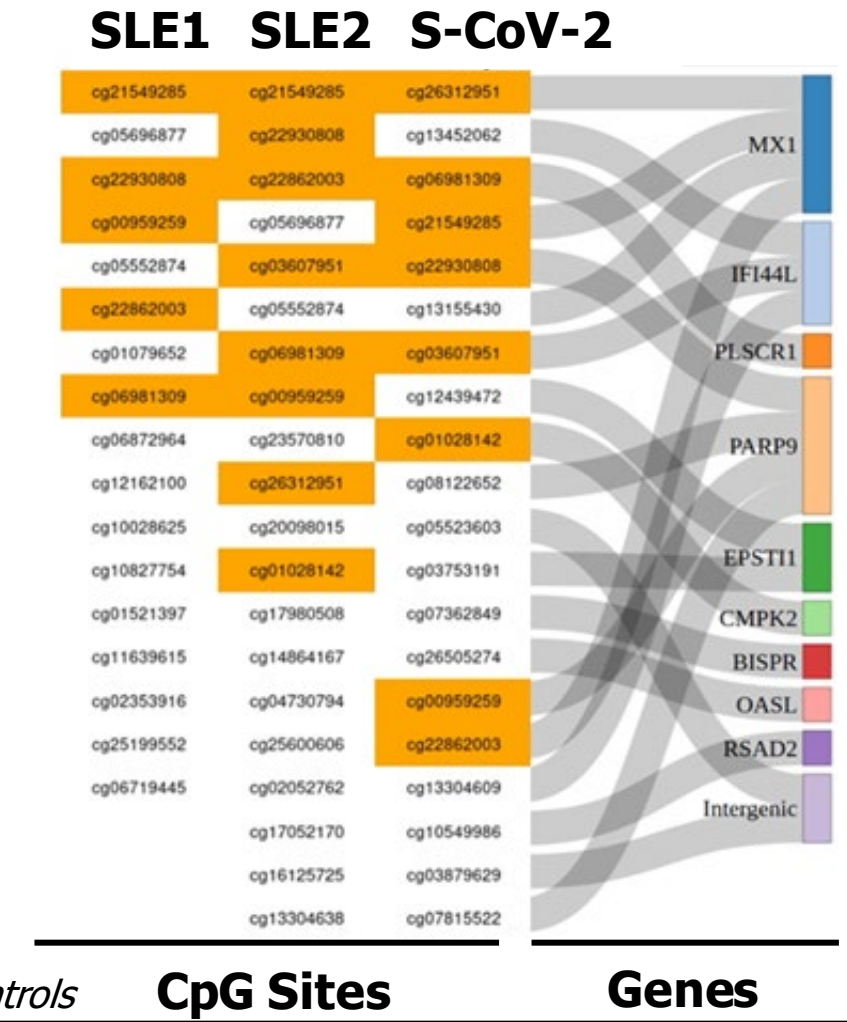
Mao W et al, Mol Systems Biol 2023

Big win: DNA methylation identifies COVID-19 infection status with over 93% sensitivity, predicts time since infection, and provides biological clues to long COVID-19 syndrome as a result of active interferon host response

COVID-19 Methylation patterns resemble autoimmune Systemic Lupus Erythematosus (SLE)



Mao W et al, Mol Systems Biol 2023



N=181 vs 63 controls

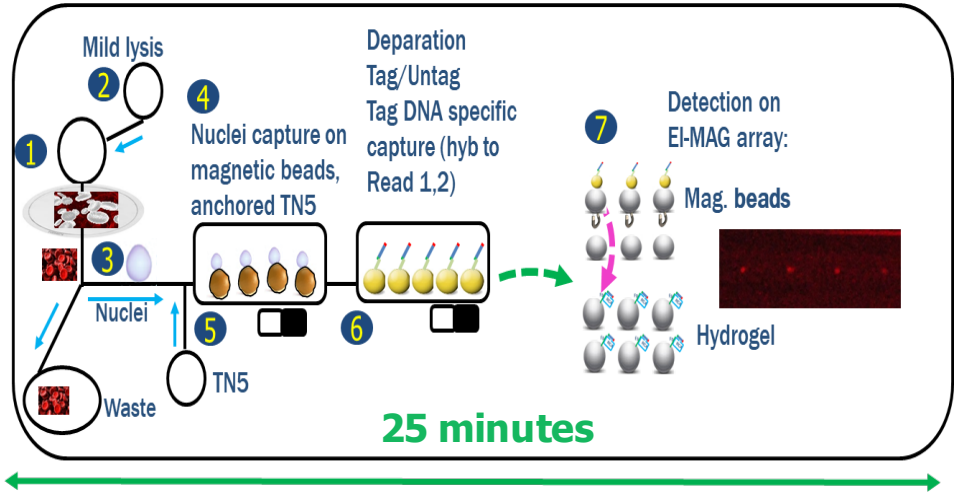
CpG Sites **Genes**

SARS-CoV-2 infection rewires the host immune epigenome, inducing hypomethylation which resembles autoimmune SLE hypomethylation

ISMMS ECHO POC TA2 epigenetic device

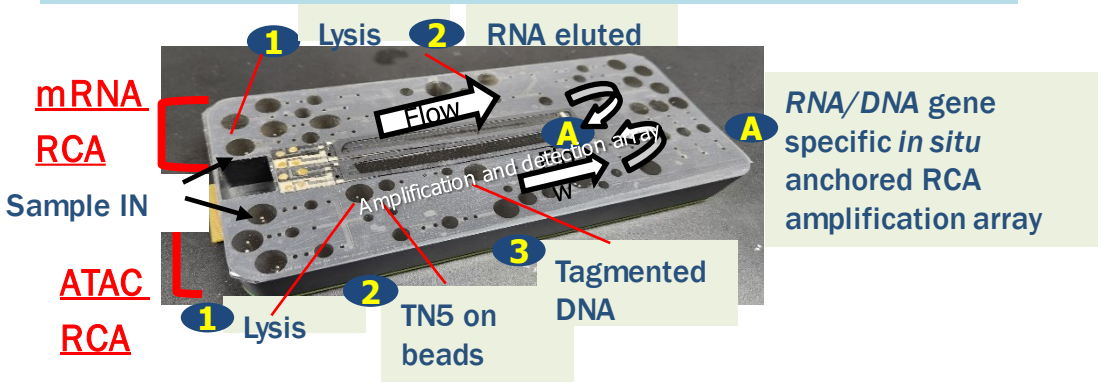
MC² DUAL ASSAY INSTRUMENT

<25 min <0.7 ft³ <25 W

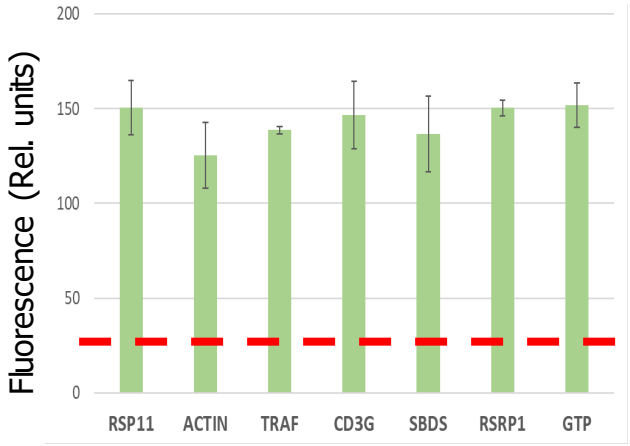


Dual assay cartridge

- 6 pumps for sample processing
- 36 microfluidics chambers (sample processing and washing)
- Rolling Circle Amplification (RCA) Array for isothermal detection



RNA/DNA gene specific *in situ* anchored RCA amplification array



N = 4, STDEV error = 2.7 % – 13.9 %

ATAC RCA Assay
Whole blood on el-mag cartridge predicting S. aureus in human exposure samples

Nexogen host ATAC-RCA demonstrates feasibility of field forward epigenetic device with sample-to-answer in 25 min

Summary

- Single cell epigenetics and in particular multiomic assays (simultaneous RNA and ATAC seq) substantially increase sensitivity and specificity of developed signatures
- ECHO developed compendium of published signatures dramatically increases specificity of the developed signatures
- Chemical exposures (OP Chlorpyriphos) imprint the epigenome of the immune system enabling the detection of exposure
- ECHO developed signatures have substantial prognostic capabilities due to the detection of immune dysregulation at the single cell level
- ECHO technologies can potentially develop tests even for exposures that are either difficult to assess due to heterogeneity (Lyme Disease)
- ECHO signatures can be translated to a point of care device that can be deployed under austere conditions with minimally trained personnel

www.darpa.mil