Rapid, Simultaneous Viral Pathogen Detection, Subtyping and Strain Differentiation Through Unbiased Metagenomic Sequencing with Delve Detect CSF

Sirisha Thippabhotla, Brian O'Donovan, Emily Crawford, Amy Wong, Coren Milbury, Steve Miller

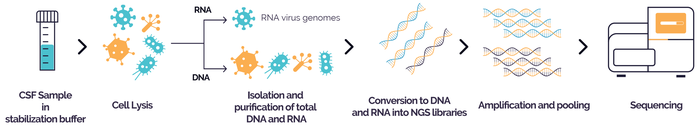
Delve Bio, Inc, Boston, MA

Background and Introduction

Delve Detect CSF: Metagenomic Next-Gen Sequencing (mNGS) for the Rapid, Unbiased Detection of DNA and RNA pathogens in Cerebrospinal Fluid (CSF)

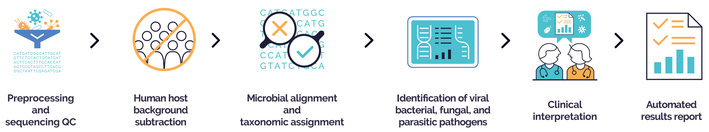
Wet Lab Processing

Automated Protocol Focused on Precision, Sterility, Speed, and Throughput



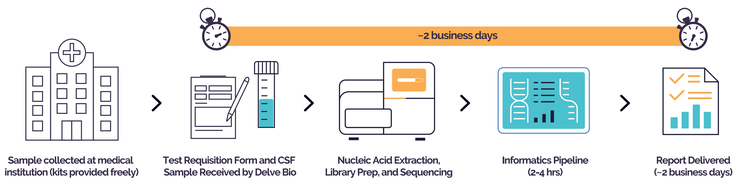
Dry Lab - Delve Decide

Designed for Speed, Accuracy, Scalability, and Auditability



Optimized for Accuracy and Turnaround Time

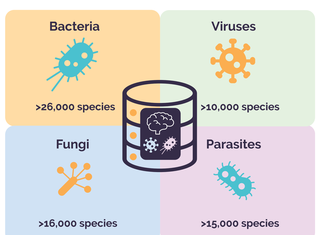
Fully Automated and Integrated Workflow Delivers Results in ~2 Business Days



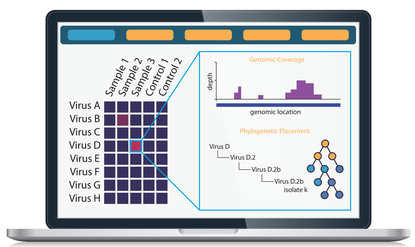
Pipeline Delivers Species-Level Taxonomic Assignments

Higher Resolution Viral Subtyping Is Possible Where Clinically Relevant

Curated Database of Known & Potential CNS Pathogens



Delve Clinical Interpretation and Decision Engine (D-CIDE)



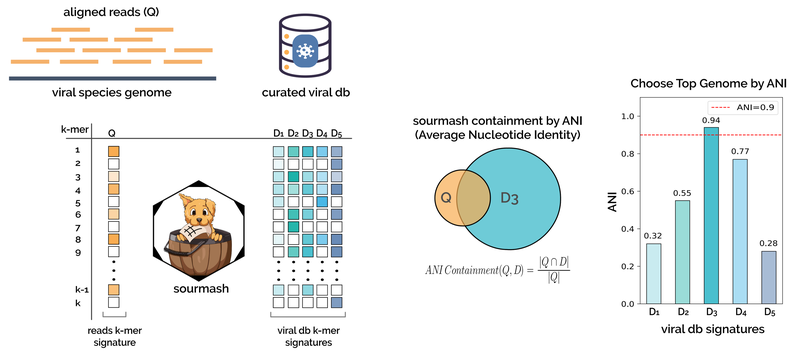
Objective

Delve Detect CSF was developed and validated to serve as a mNGS platform for hypothesis-free, unbiased detection of DNA and RNA pathogens in cerebrospinal fluid (CSF) from patients with suspected infectious meningitis/encephalitis diagnosis. The platform provides agnostic and comprehensive detection of viral, bacterial, fungal, and parasitic species in a single test.

This study investigated whether mNGS sequencing data generated during routine clinical testing could be further leveraged to achieve higher-resolution taxonomic assignments for viral strain and subtyping, potentially expanding the clinical and epidemiological applications of the Delve Bio platform.

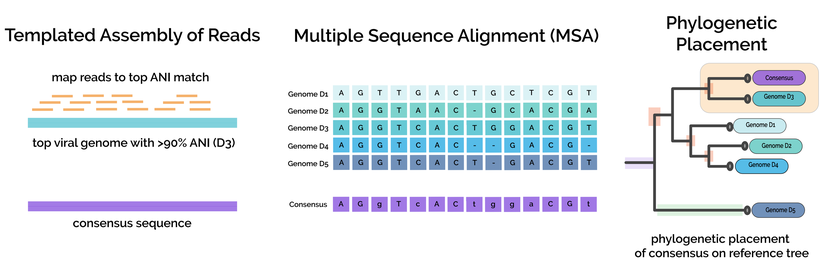
Methods

Computational Viral Subtyping



Viral reads aligning to a species of interest identified by Delve Bio's proprietary bioinformatic pipeline, Delve Decide, are used to generate a kmer-based hash vector - or 'signature' - using the sourmash package [1]. This kmer signature is compared to a curated database of signatures from genomic sequences assigned to that species in the NCBI NT database. A similarity measure based on kmer containment - average nucleotide identity (ANI) - is used to select the representative genome used for templated assembly of the reads to generate a consensus sequence/contig. This sequence is then taxonomically 'placed' within phylogenetic context using a multiple sequence alignment (MSA) based approach (IQ-TREE) [2].

Templated Assembly, MSA and Phylogenetic Placement



Results

Sub-Species Calls in Clinical CSF Samples

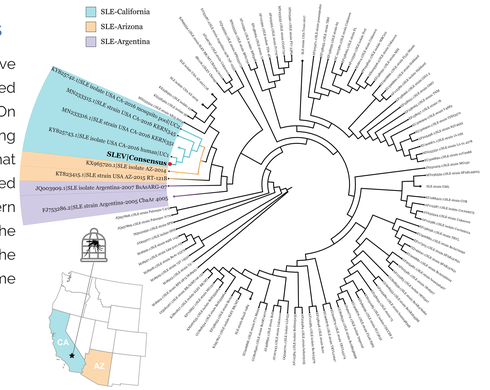
Reported Species	Aligned Reads	Unique Reads	Covered Bases	Subtype (Taxid)	Accession (ANI)
Varicellovirus humanalpha3 (VZV)	10328	5394	>117000	Human herpesvirus 3 isolate DE10-2660, partial genome (10335)	KP771893.1 (0.99)
Roseolovirus humanbeta 6b (HHV-6b)	675364	119447	>160000	Human betaherpesvirus 6B strain GTEX-13FLV, partial genome (32604)	MH698399.1 (0.99)
Lymphocryptovirus humangamma4 (EBV)	8581	3666	>73000	Human gammaherpesvirus 4 isolate IMS_Saliva_90_BWA, complete genome (10376)	MG298924.1 (0.99)
Betapolyomavirus secuhominis (JCPv)	45	20	1244	JC polyomavirus DNA, complete genome, isolate: FL-8 (10632)	AB127350.2 (1.0)
Human Immunodeficiency Virus 1 (HIV-1)	32973	2508	3598	HIV-1 isolate 99KE_KISII5003 from Kenya, partial genome (11676)	AF457059.1 (0.99)
West Nile Virus (WNV)	9380	739	8226	WNV strain USA Sacramento CA-2012 W0685, partial genome (11082)	OQ721247.1 (0.99)
St. Louis Encephalitis Virus (SLEV)	19736	1748	10859	SLEV isolate USA CA-2016 human UC-1, complete cds (11080)	KY825743.1 (0.99)
Enterovirus A	1577	121	3810	Enterovirus A71 strain EV-A71 2019 FR-0622, complete genome (138948)	MW030642.1 (0.99)
Enterovirus B	6455	513	4051	Coxsackievirus B4 isolate B4T051217 polyprotein gene, complete cds (138949)	MG845889.1 (0.94)

Table of 9 selected results from clinical CSF samples that have >1000 bases of coverage. The highlighted rows indicate the species with phylogengy results presented below.

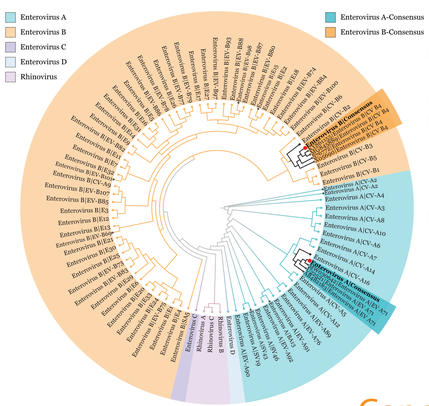
Phylogenetic Placement of Specific Results

St. Louis Encephalitis Virus

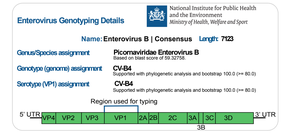
A patient sample previously reported as positive for St. Louis Encephalitis Virus (SLEV) yielded 19,736 reads aligning to the SLEV genome. On running this sample through our subtyping workflow, the phylogenetic analysis showed that the sample was most similar to a strain sequenced from a mosquito pool collected in 2016 from Kern county. These results are concordant with the published case from UCSF [3] that reported the patient as being an oil field worker from the same geographic region.



Enterovirus



Two Enterovirus detections from the Delve Detect pipeline were further strain-typed using our subtyping approach. The Enterovirus A sample most closely matched to an A71 strain (highlighted in blue), while the Enterovirus B sample showed the highest similarity to a Coxsackievirus B4 strain (highlighted in orange). To validate these results, we also analyzed the samples using the EV-genotyping tool developed by RIVM [4], and the findings were concordant.



Conclusion

This study highlights the capability of mNGS to provide sub-species taxonomic resolution for viral pathogens in a rapid manner. The ability to perform strain-level classification and phylogenetic analysis within a standard diagnostic workflow has important implications for viral outbreak surveillance, lineage tracing, and public health applications. These results suggest that metagenomic sequencing can contribute actionable epidemiological insights for infectious disease monitoring and response, in addition to informing treatment in routine clinical settings.

[1] Pierce NT, Irber L, Reiter T et al. Large-scale sequence comparisons with sourmash. F1000Research 2019

[2] Lam-Tung Nguyen, Heiko A. Schmidt, Arndt von Haeseler, Bui Quang Minh, IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies, Molecular Biology and Evolution, January 2015

[3] Chiu, Charles Y et al. "Diagnosis of Fatal Human Case of St. Louis Encephalitis Virus Infection by Metagenomic Sequencing, California, 2016." Emerging infectious diseases, Oct 2017

[4] Kroneman A, Vennema H, Deforche K, et al. An automated genotyping tool for enteroviruses and noroviruses. J Clin Virol. 2011