

# IFSH High-Throughput Sequencing (HTS) Initiative

Behzad Imanian  
IFSH HTS Initiative, leader  
Research Assistant Professor, FSN  
Illinois Tech

September 4, 2024

# IFSH HTS Initiative

*Industry Advisory  
Committee*

33 members  
20 food companies

*Industry-Governmental  
Agencies Council*

*Proposal &  
Charter*

*Apr 2015*

*Funding from  
FDA*

*Sept 2015*

*HTS Initiative*

*Mar 2016*

FDA

CDC

FSIS

BioTech

Industry

Food companies → WGS of  
pathogens → GenomeTrakr

Promote HTS technology

Infrastructure

HTS Laboratory  
HTS Bioinformatics

Activities

HTS Services  
Collaborative Research  
Outreach

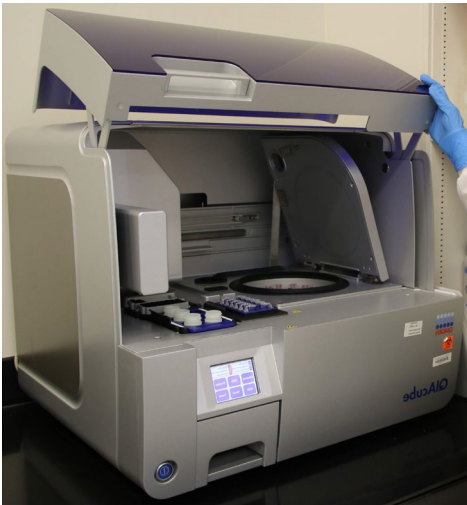
**ILLINOIS TECH**

# Laboratory

TapeStation 4200



QIAcube



MinION



MiSeq



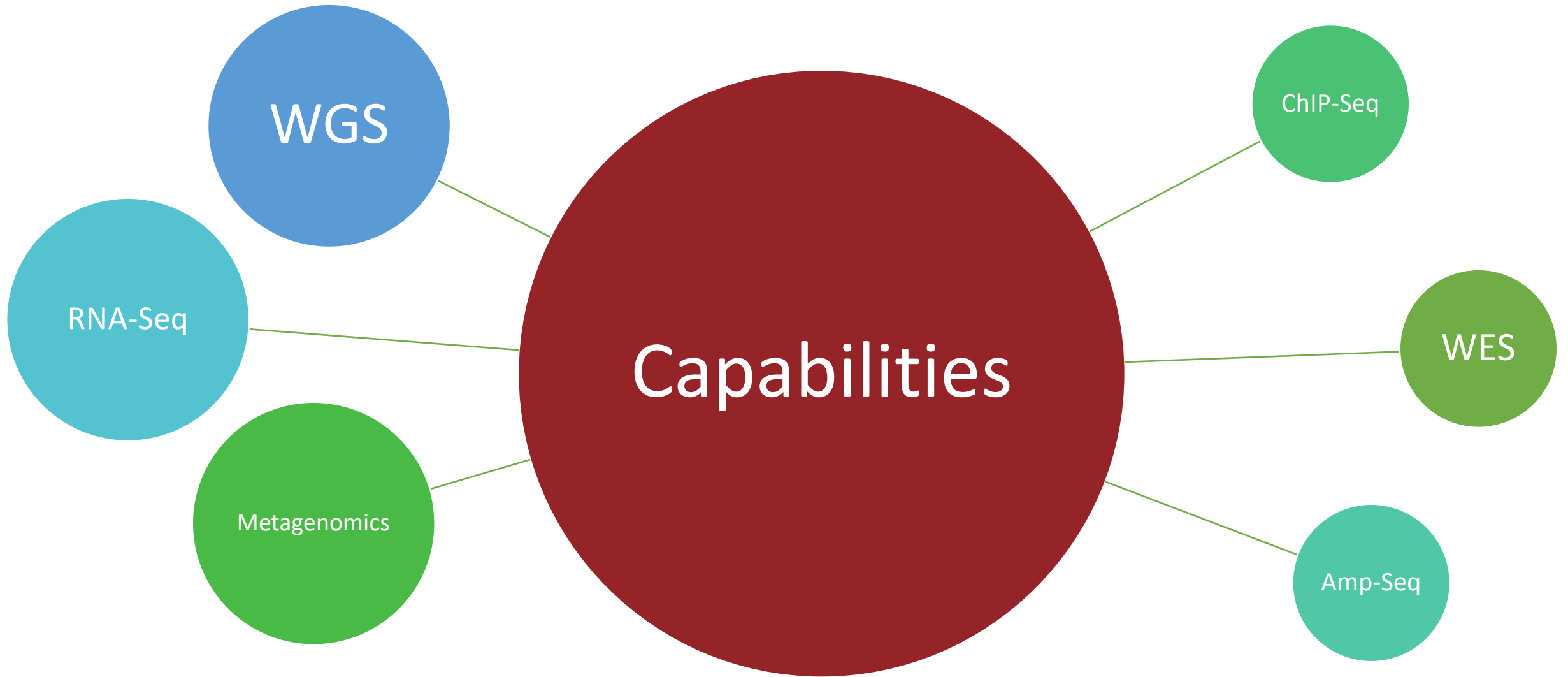
15 GB      50 x 10<sup>6</sup> reads

NextSeq 2000



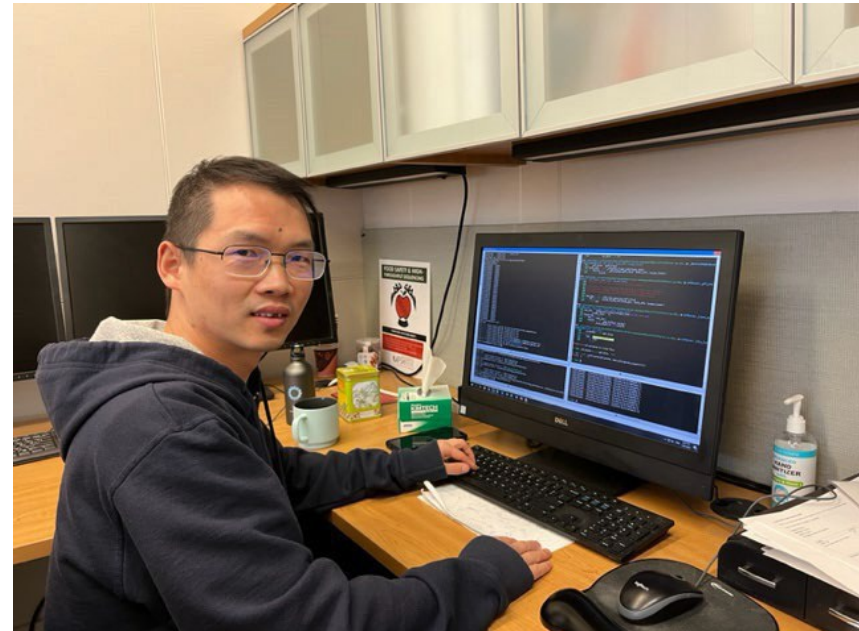
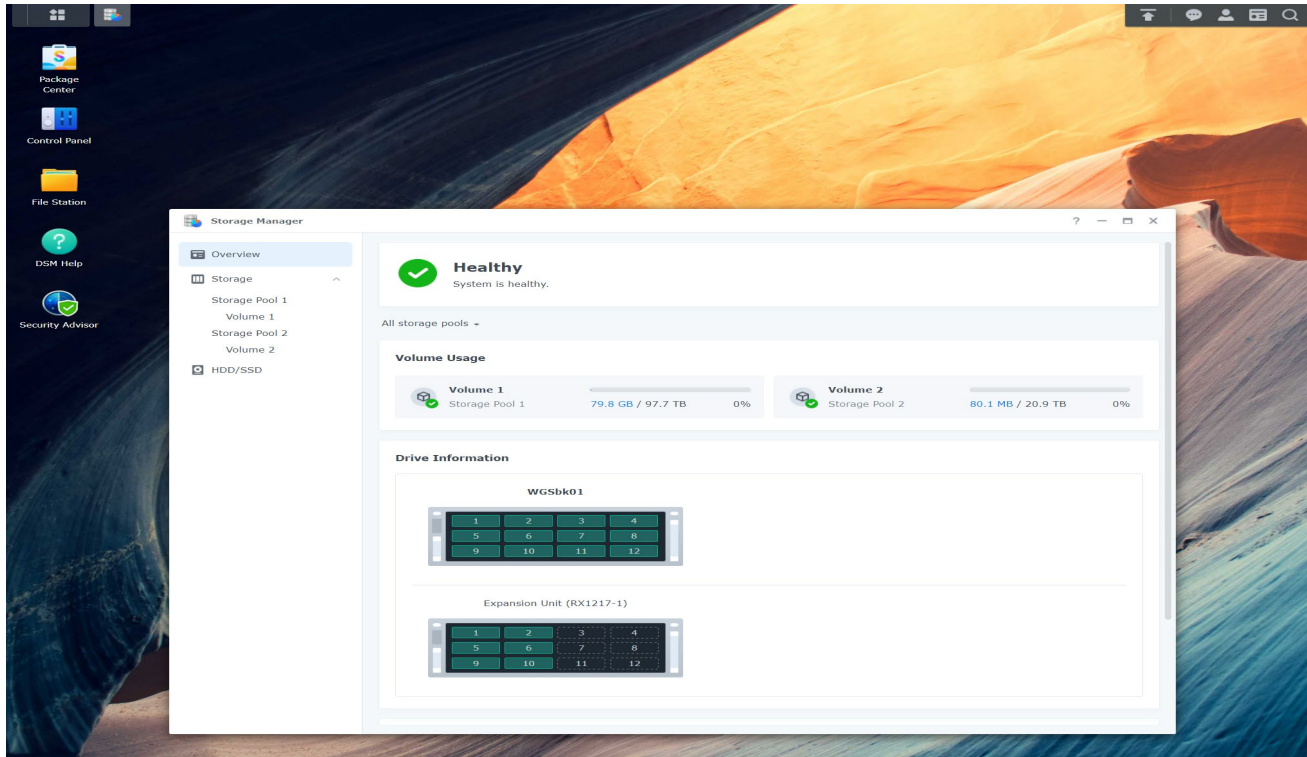
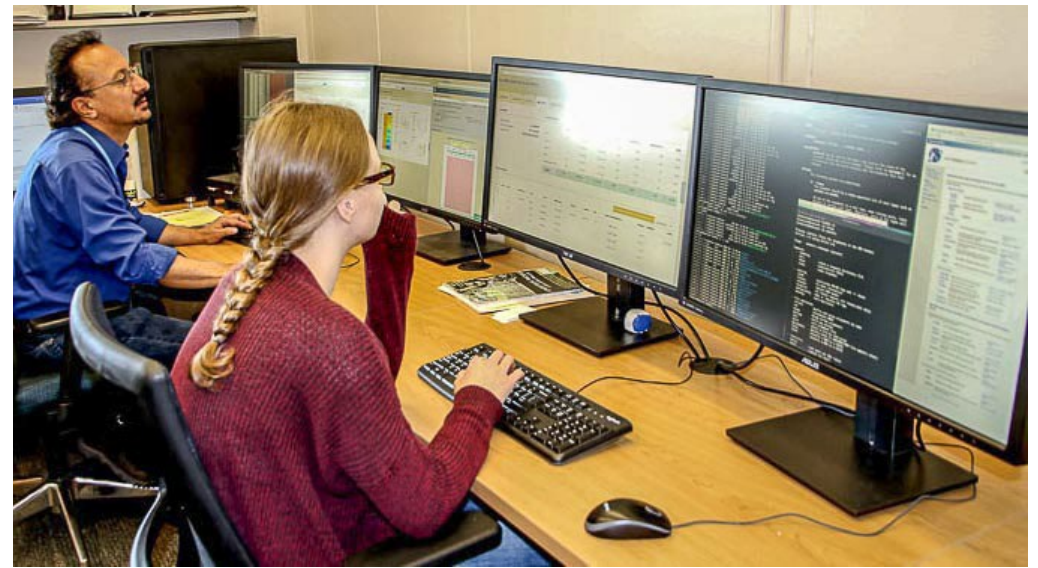
540 GB      3.6 x 10<sup>9</sup> reads

# HTS Laboratory



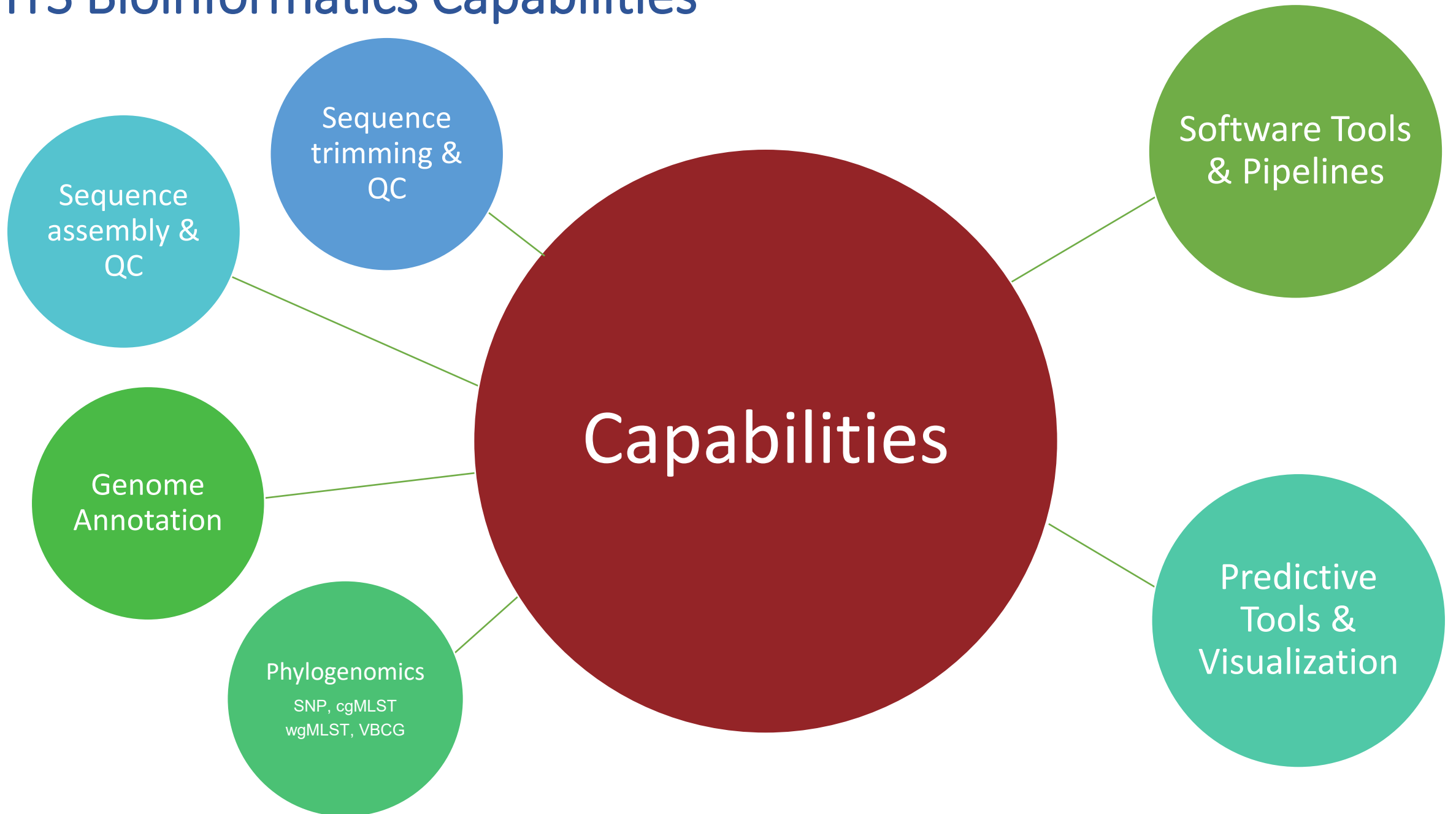
# Bioinformatics Facility

- Four HPCs (512 GB - 1 TB RAM; up to 128 CPUs)
- A back-up system (> 140 TB current capacity – expandable)
- Learning and training (e.g. ML and AI)





# HTS Bioinformatics Capabilities

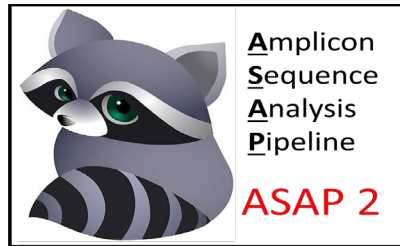


# New Software & Tools

## Phylogenomic Analysis - VBCG



## Amplicon Sequence Analysis Pipeline - ASAP2



## PlasmidHunter



ILLINOIS TECH

**Galaxy** Workflow Visualize Shared Data Help Log in or Register

Tools  
search tools  
Upload Data

### HTSI Initiative

**01 Genomic Data Analysis - GAAC**  
Automatically analyze prokaryotic genomic data including raw FASTQ reads and assembled genomes.

**02 KEGG Pathway Annotation and Comparison - KEGG\_Anno**  
Annotate protein sequences and compare in KEGG pathways at multiple levels.

**03 Illumina Data Preprocessing - seqQC**  
Evaluate sequence quality profile, remove artificial duplicates, trim low-quality regions and adaptors, filter low-quality reads.

**04 Phylogenomic Analysis - VBCG**  
Build a phylogenomic tree based on 29 validated bacterial core genes (VBCG) with input of whole genome sequence FASTA files.

**05 Genomic Contig Identification - PlasmidHunter**  
Predict the location (chromosome or plasmid) of contig sequences using machine learning

**06 Amplicon Data Analysis - ASAP2**  
Automatically analyze prokaryotic 16S rRNA gene amplicon sequencing data using QIIME 2 and other tools.

**07 Transcriptomic Data Analysis - rnaSeq**  
RNA-seq raw data quality control, read mapping, gene expression quantification.

Unzip Unzip a file

Statistics

**WORKFLOWS**  
All workflows

## High Throughput Sequencing Initiative (HTSI)

Institute for Food Safety and Health, Illinois Institute of Technology

### About Our Lab

Our lab mainly focuses on microbiological research and bioinformatics tool development. We are interested in applying AI and machine learning in solving biological problems using big data.

### Lab Members

- Behzad Imanian, PhD, Research Assistant Professor
- Renmao "Tim" Tian, PhD, Research Scientist

### Recent Projects

#### Microbiology

- A novel technology to isolate bacteria from soil environments using machine learning
- Genomics and evolution of Clostridium botulinum
- Ecology and evolution of Pathogenic and non-pathogenic Escherichia coli

#### Bioinformatics Tools

Galaxy Server

# Recent Publications



Briefings in Bioinformatics, 2024, 25(4), bbae322  
<https://doi.org/10.1093/bib/bbae322>  
Problem Solving Protocol

## PlasmidHunter: accurate and fast prediction of plasmid sequences using gene content profile and machine learning

Renmao Tian<sup>1</sup>, Jizhong Zhou<sup>2</sup>, Behzad Imanian<sup>1,3,\*</sup>

<sup>1</sup>Institute for Food Safety and Health, Illinois Institute of Technology, 6502 S Archer Rd, Bedford Park, IL 60501, United States

Tian and Imanian *Microbiome* (2023) 11:247  
<https://doi.org/10.1186/s40168-023-01705-9>

Microbiome

SOFTWARE

Open Access

## VBCG: 20 validated bacterial core genes for phylogenomic analysis with high fidelity and resolution

Renmao Tian<sup>1</sup> and Behzad Imanian<sup>1,2\*</sup>

Tian and Imanian *BMC Bioinformatics* (2022) 23:27  
<https://doi.org/10.1186/s12859-021-04555-0>

BMC Bioinformatics

SOFTWARE

Open Access

## ASAP 2: a pipeline and web server to analyze marker gene amplicon sequencing data automatically and consistently

Renmao Tian<sup>1</sup> and Behzad Imanian<sup>1,2\*</sup>

npj | science of food

[www.nature.com/npjfood](http://www.nature.com/npjfood)

REVIEW ARTICLE OPEN



## The power, potential, benefits, and challenges of implementing high-throughput sequencing in food safety systems

Behzad Imanian<sup>1,2\*</sup>, John Donaghy<sup>2</sup>, Tim Jackson<sup>1,14</sup>, Sanjay Gumma<sup>14</sup>, Balasubramanian Ganesan<sup>15</sup>, Robert C. Baker<sup>2</sup>, Matthew Henderson<sup>6</sup>, Emily K. Butler<sup>7</sup>, Yingying Hong<sup>7</sup>, Brendan Ring<sup>8</sup>, Clare Thoop<sup>8,15</sup>, Ramin Khaksar<sup>9</sup>, Mansour Samadpour<sup>10</sup>, Kahlil A. Lawless<sup>11</sup>, Iain MacLaren-Lee<sup>12</sup>, Heather A. Carleton<sup>13</sup>, Renmao Tian<sup>1</sup>, Wei Zhang<sup>1</sup> and Jason Wan<sup>1</sup>

genes



Article

## The Light Chain Domain and Especially the C-Terminus of Receptor-Binding Domain of the Botulinum Neurotoxin (BoNT) Are the Hotspots for Amino Acid Variability and Toxin Type Diversity

Renmao Tian<sup>1</sup>, Melissa Widel<sup>1</sup> and Behzad Imanian<sup>1,2,\*</sup>

<sup>1</sup> Institute for Food Safety and Health, Illinois Institute of Technology, Bedford Park, IL 60501, USA

toxins



Article

## Integration of Complete Plasmids Containing *Bont* Genes into Chromosomes of *Clostridium parvotulinum*, *Clostridium sporogenes* and *Clostridium argentinense*

Theresa J. Smith<sup>14</sup>, Renmao Tian<sup>24</sup>, Behzad Imanian<sup>23</sup>, Charles H. D. Williamson<sup>1</sup>, Shannon L. Johnson<sup>4</sup>, Hajnalka E. Daligault<sup>4</sup> and Kristin M. Schill<sup>5,6,4</sup>

<sup>1</sup> Pathogen and Microbiome Institute, Northern Arizona University, Flagstaff, AZ 86011, USA; terrys2much@comcast.net (T.J.S.); chase.williamson@nau.edu (C.H.D.W.)

ILLINOIS TECH

IFSH | INSTITUTE FOR FOOD SAFETY AND HEALTH ILLINOIS TECH

Innovation Through Collaboration



# Symposiums

2016



2017



# Workshops

2018

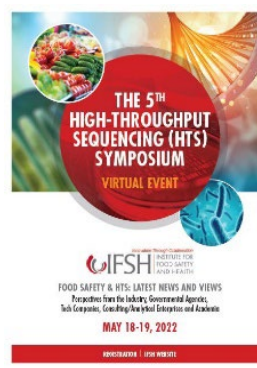


2019



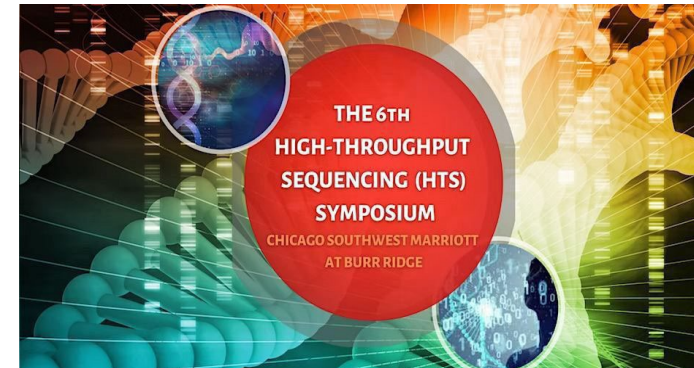
# Webinars

2022



# Roundtables

2023



# ILLINOIS TECH

# IFSH

INSTITUTE FOR  
FOOD SAFETY  
AND HEALTH  
ILLINOIS TECH

Innovation Through Collaboration

# The 7<sup>th</sup> IFSH HTS Symposium

## May 2025 - Chicago

# Ongoing and Future Projects

## Developing Analytical and Predictive Tools

Utilizing ML and AI

## Evolution of Pathogenicity

Genomic distinctions between pathogens and non-pathogens (genes, genome content, kmers, ...)  
LLM (genome and protein language models)

## Microbiomes Characterization

Characterization of microbial communities (composition, succession and links to events/traits in a longitudinal study in the State of Illinois)

## Microbiomes Manipulation

Study the effects of manipulation of microbiomes in the environment by addition & removal of microbes/chemicals

# Questions/Comments?