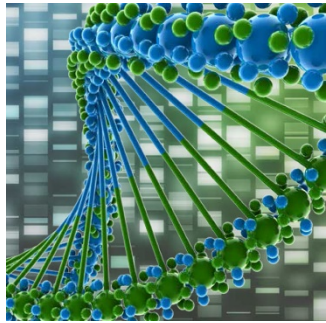
# Applications of NGS in Clinical Microbiology

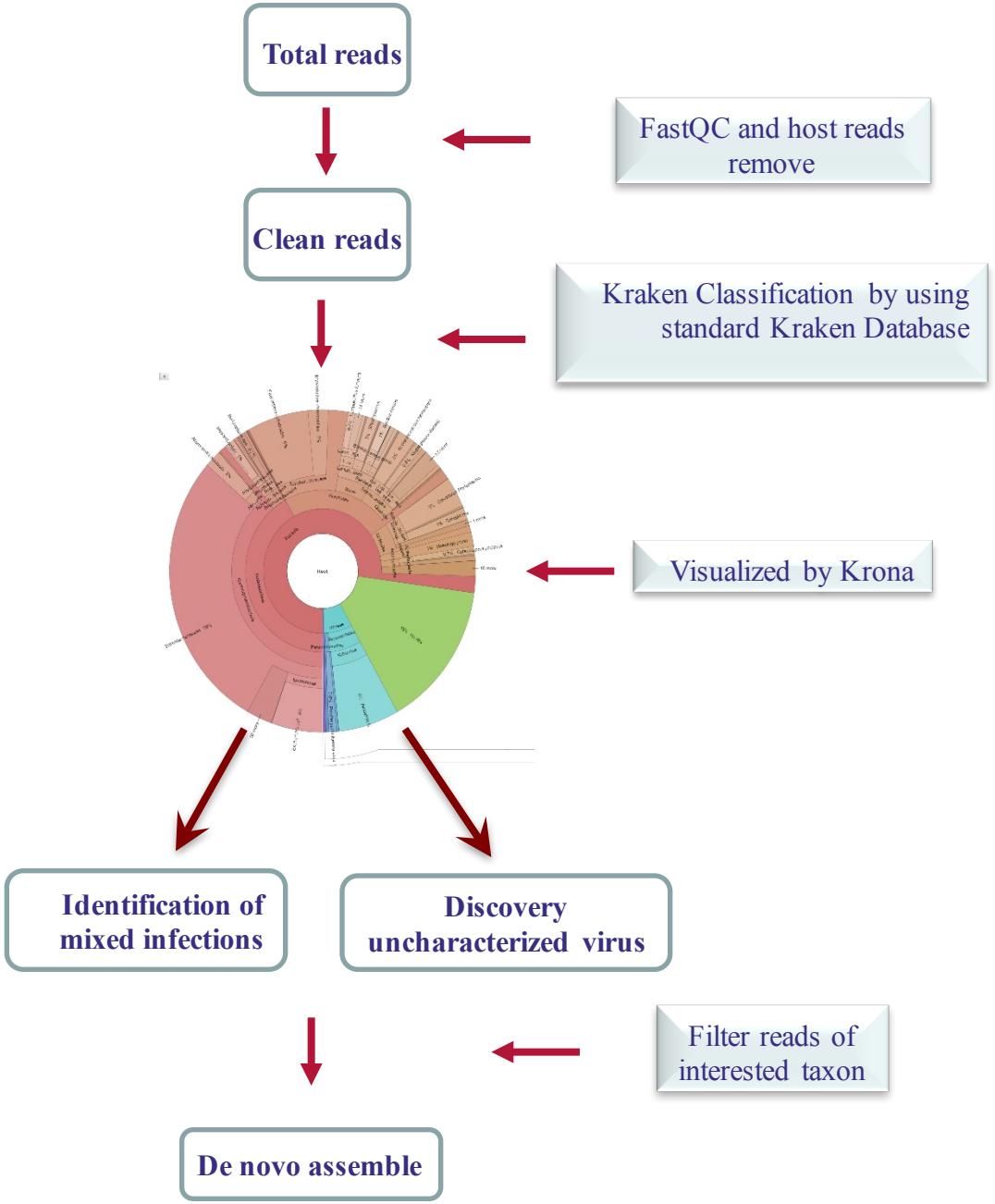
**Ganwu Li**

**Assistant Professor**

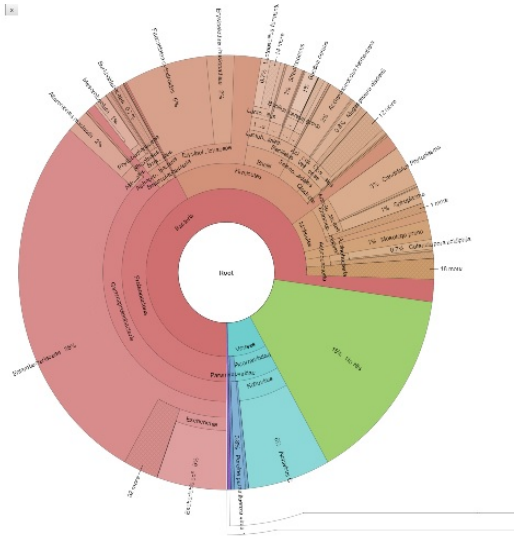
**Veterinary Diagnostic Lab at Iowa State University**

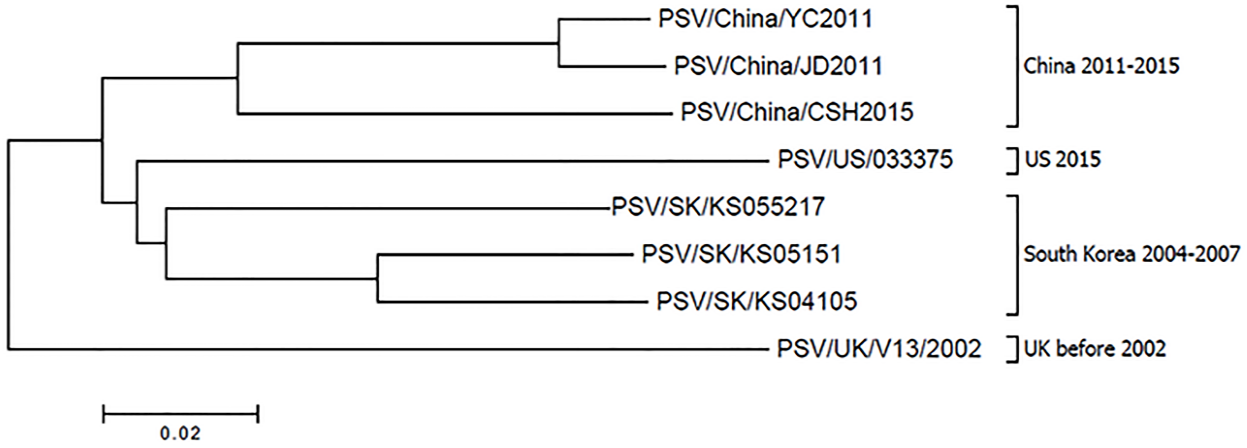


# Detect and Identify Pathogens

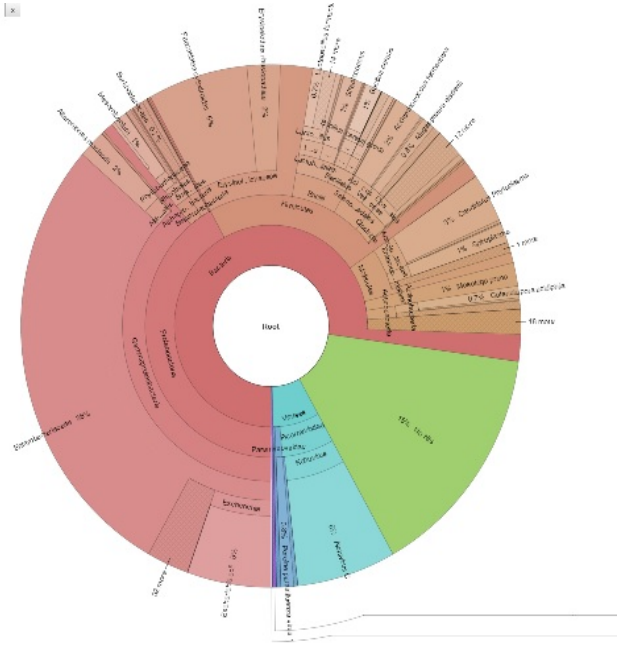

Total reads
FastQC and host reads remove
Clean reads
Kraken Classification by using standard Kraken Database
Visualized by Krona
Identification of mixed infections
Discovery uncharacterized virus
Filter reads of interested taxon
De novo assemble

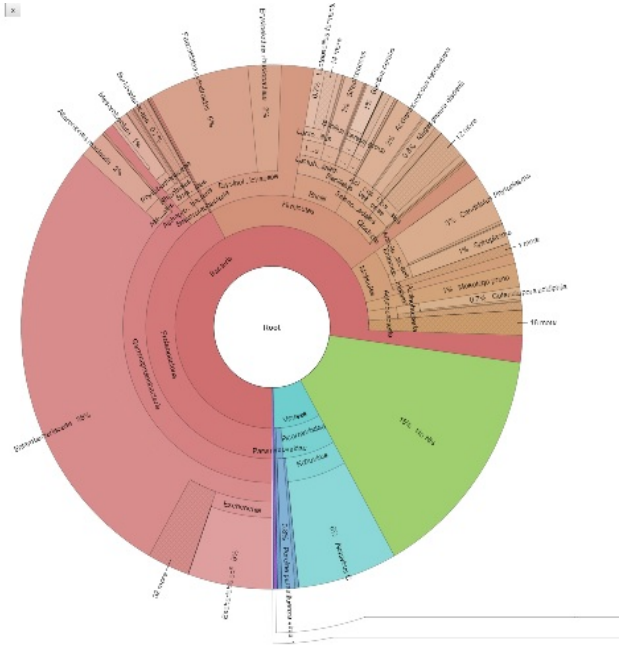
# Identification of Uncharacterized Virus from PEDV Positive Fecal Samples



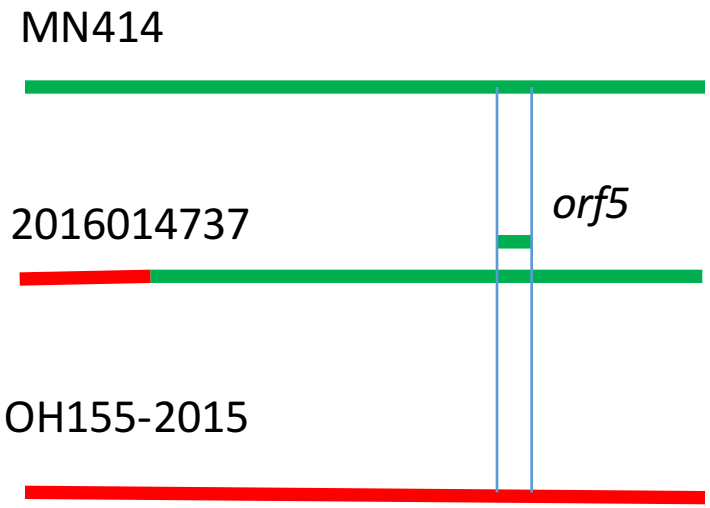


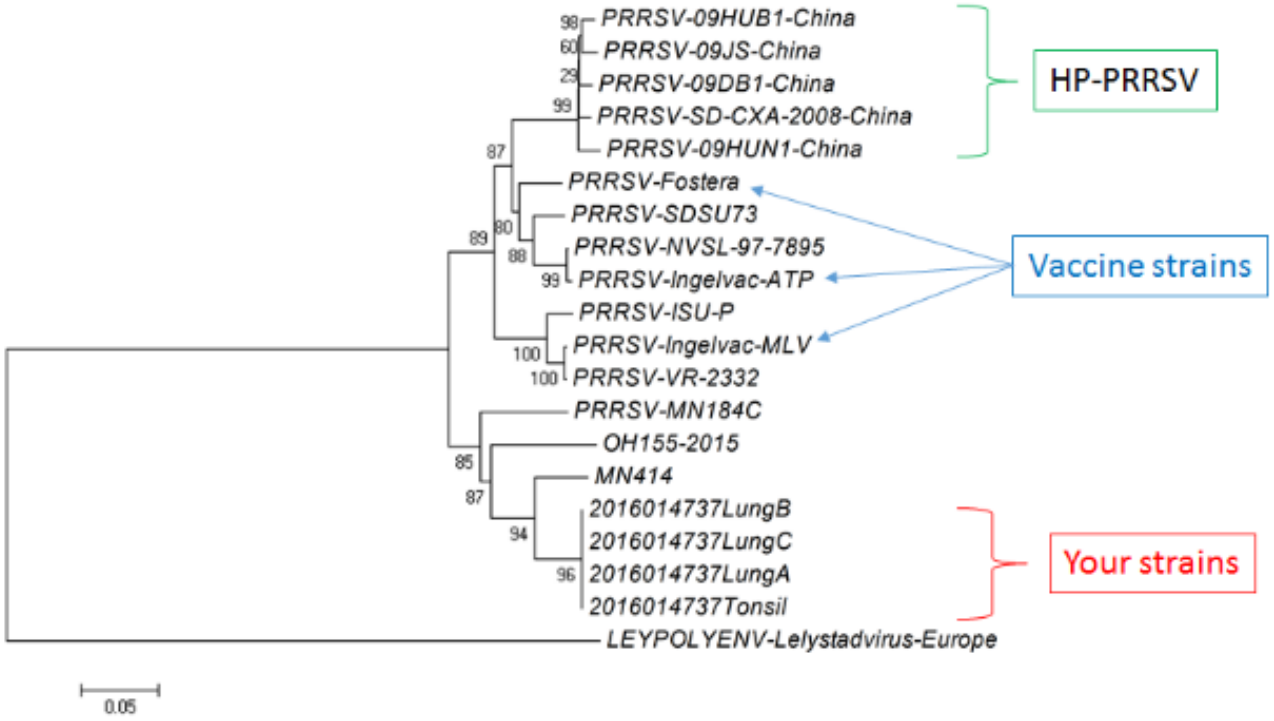
# Identification of Mixed-Infections from PEDV Positive Fecal Samples





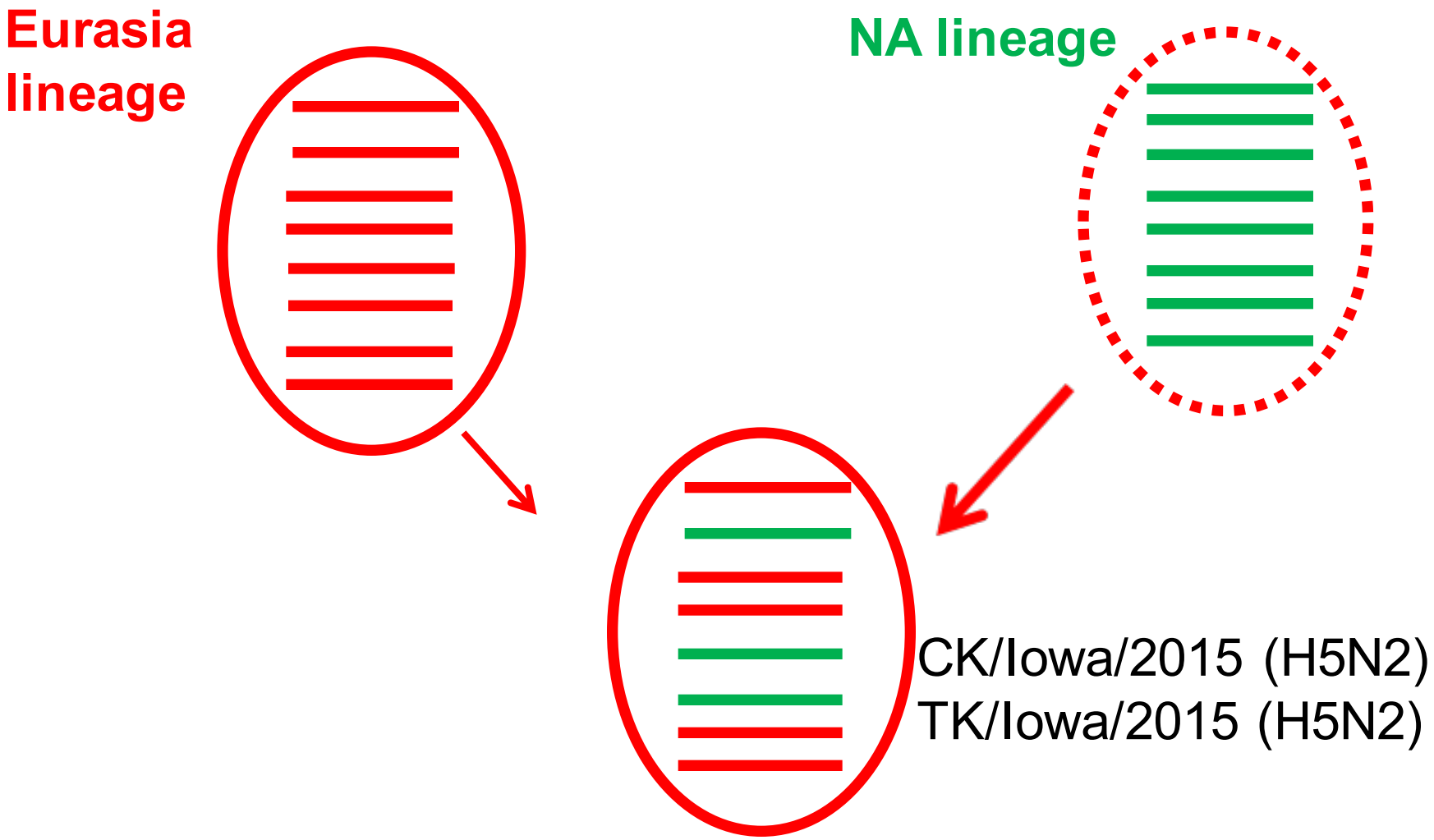
# NGS Applications in Viral Comparison (1)



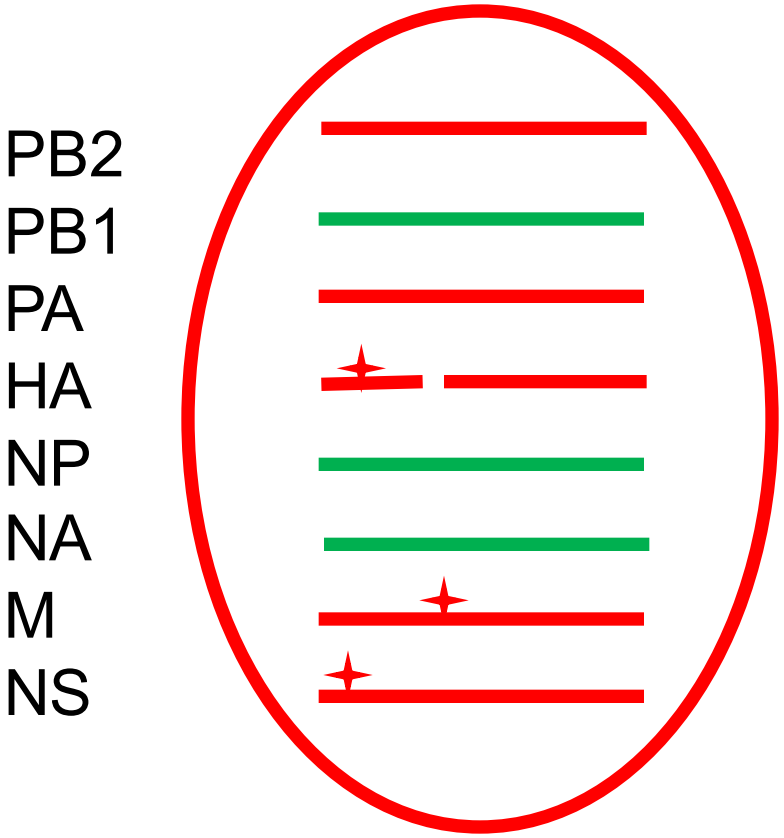



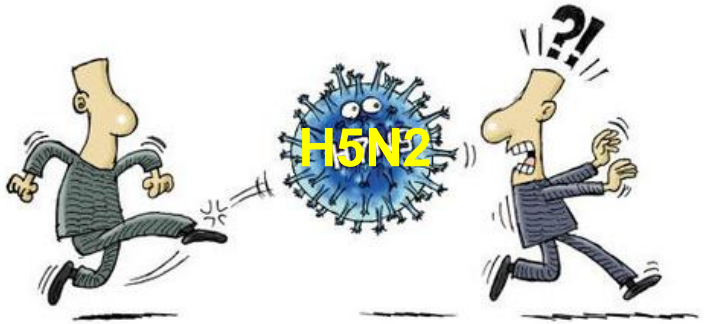

Can the sequence of single gene provide the breadth of evidence needed for differentiating the whole virus? No!

# NGS Applications in Viral Comparison (2)

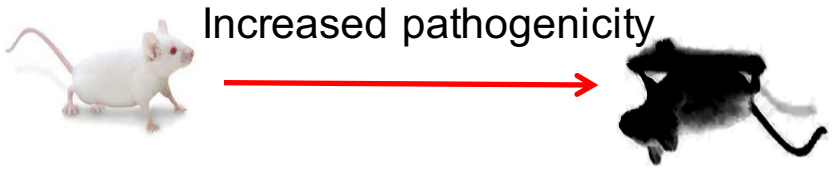


# NGS Application in Public Health Risk Assessment



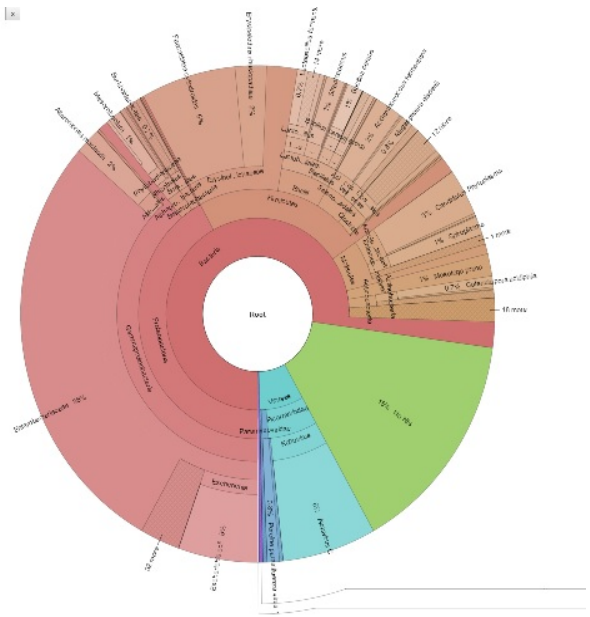

HA1:S138A

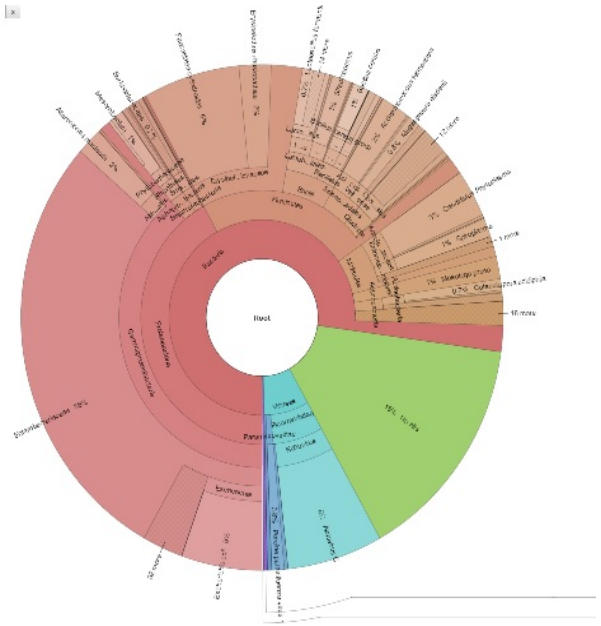



M1:T215A

NS1: P42S

# NGS Application in Detection of Cell Line Contamination





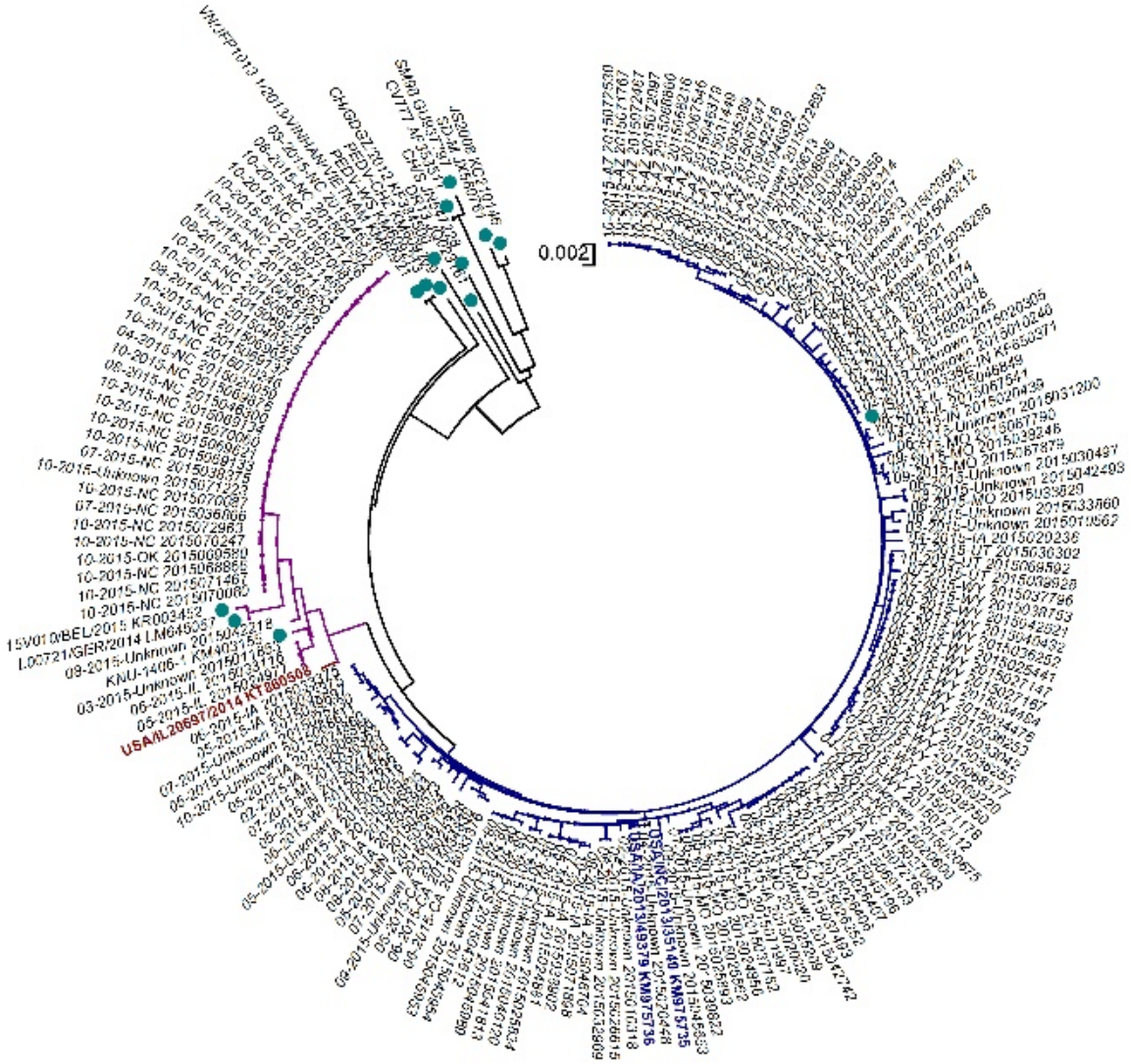
# PRRSV – Clinical Serum Samples

| Sample ID | Ct Value | # Total reads | # Reads mapped | % reads mapped | Full length genome | Total bases assembled | Longest contig | Total gaps | Gap <150 | Gap <750 | Gap <1500 | Gap >=1500 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 14994-20VD_S12 | 15.1 | 1087018 | 569911 | 52.43% | Yes | 14997 | 14997 | 0 | 0 | 0 | 0 | 0 |
| 26109_S8 | 17.1 | 308880 | 17838 | 5.78% | Yes | 14999 | 14999 | 0 | 0 | 0 | 0 | 0 |
| 1295_S9 | 19.1 | 411092 | 79321 | 19.30% | Yes | 15510 | 15510 | 0 | 0 | 0 | 0 | 0 |
| S4099_S3 | 19.7 | 1180906 | 24489 | 2.07% | Yes | 15025 | 15025 | 0 | 0 | 0 | 0 | 0 |
| 30820_S7 | 19.8 | 333428 | 52996 | 15.89% | Yes | 14959 | 14959 | 0 | 0 | 0 | 0 | 0 |
| S5342_S10 | 20.5 | 1554098 | 24895 | 1.60% | Yes | 15002 | 15002 | 0 | 0 | 0 | 0 | 0 |
| PRRpig1S4 | 22 | 1029532 | 9086 | 0.88% | Yes | 15340 | 15340 | 0 | 0 | 0 | 0 | 0 |
| S0652A_S5 | 22.2 | 1084508 | 8503 | 0.78% | Yes | 15000 | 15000 | 0 | 0 | 0 | 0 | 0 |
| PRRpig11_S4 | 23 | 970396 | 112907 | 11.64% | Yes | 15398 | 15398 | 0 | 0 | 0 | 0 | 0 |
| S5432_S7 | 23.6 | 2223592 | 19699 | 0.89% | Yes | 15396 | 15396 | 0 | 0 | 0 | 0 | 0 |
| S4969_S6 | 24.3 | 1650972 | 2305 | 0.14% | No | 14614 | 4296 | 9 | 8 | 1 | 0 | 0 |
| S4825_S12 | 25 | 1681672 | 422 | 0.03% | No | 2155 | 270 | 11 | 1 | 5 | 3 | 2 |
| S2154_S13 | 25.2 | 1457568 | 582 | 0.04% | No | 4354 | 580 | 24 | 8 | 10 | 4 | 2 |
| S13059316_S9 | 26.2 | 1416986 | 2302 | 0.16% | No | 10546 | 1485 | 19 | 11 | 7 | 1 | 0 |
| S0652B_S11 | 26.8 | 1327504 | 529 | 0.04% | No | 4491 | 649 | 19 | 3 | 12 | 2 | 2 |

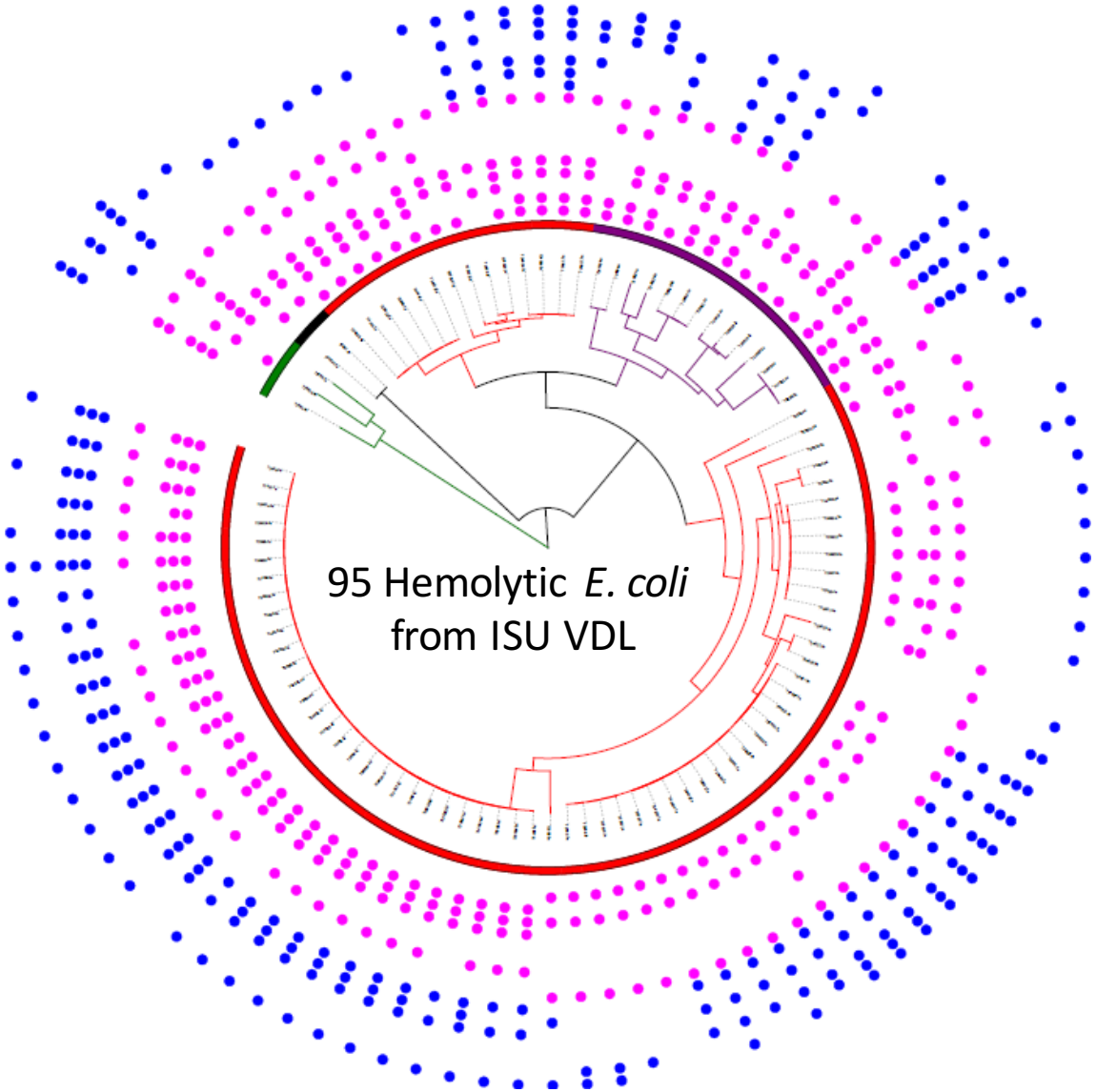
# PRRSV – Clinical Lung Samples

| Sample ID | Ct Value | # Total reads | # Reads mapped | % reads mapped | Full length genome | Total bases assembled | Longest contig | Total gaps | Gap <150 | Gap <750 | Gap <1500 | Gap >=1500 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 30324_S4 | 16.4 | 3367976 | 136882 | 4.06% | Yes | 15068 | 15068 | 0 | 0 | 0 | 0 | 0 |
| 31034_S3 | 16.6 | 5915908 | 100744 | 1.70% | Yes | 14981 | 14981 | 0 | 0 | 0 | 0 | 0 |
| 30131_S6 | 17.3 | 3418010 | 39526 | 1.16% | Yes | 15011 | 15011 | 0 | 0 | 0 | 0 | 0 |
| 31861_S2 | 19.2 | 6332688 | 60408 | 0.95% | Yes | 15000 | 15000 | 0 | 0 | 0 | 0 | 0 |
| 14-1179A_S6 | 21.0 | 963362 | 18666 | 1.94% | Yes | 15377 | 15377 | 0 | 0 | 0 | 0 | 0 |
| 4302_S16 | 21.3 | 910914 | 12116 | 1.33% | No | 10499 | 2043 | 23 | 11 | 11 | 1 | 0 |
| 4945_S12 | 22.2 | 845346 | 1388 | 0.16% | No | 9897 | 1713 | 21 | 15 | 5 | 0 | 1 |
| 4902_S7 | 22.9 | 1114286 | 6176 | 0.55% | No | 6545 | 698 | 26 | 9 | 14 | 3 | 0 |
| 4183_S9 | 23.1 | 1061682 | 6414 | 0.60% | No | 1222 | 360 | 6 | 0 | 2 | 1 | 3 |
| 4210_S6 | 23.3 | 1243518 | 16497 | 1.33% | No | 1433 | 304 | 8 | 1 | 3 | 0 | 4 |
| 14-8412_S5 | 26.3 | 912364 | 1239 | 0.14% | No | 712 | 150 | 4 | 0 | 0 | 2 | 2 |
| 7033A_S5 | 31.3 | 916084 | 398 | 0.04% | No | | | | | | | |
| 71759A_S6 | 32.0 | 1354144 | 576 | 0.04% | No | | | | | | | |
| 70451A_S7 | 35.6 | 1498278 | 740 | 0.05% | No | | | | | | | |

# PEDV 04/2015-02/2016



# Bacterial Genomic Epidemiology



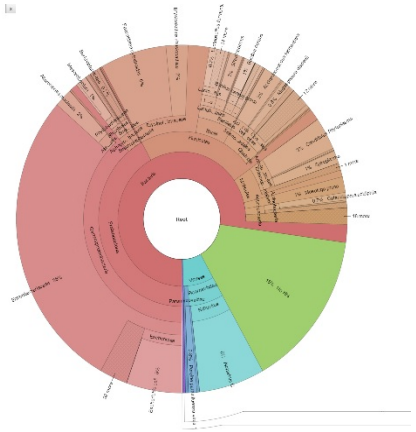

LT: 58.9% (56/96)
STa: 31.3% (30/96)
STb: 54.2% (52/95)
EASTA: 36.5% (35/95) (64/96)(67.4%)
Stx1 :3.1% (3/95) (17/95) (17.9%)
STX2=STX2e: 21.1% (20/95)

987p:0%
F41: 0%
K99: 0%
K88: 40%
F18: 43.2%
Paa: 36.8%
Novel adhesin: 33.7%

# Bacterial Resistome Analysis

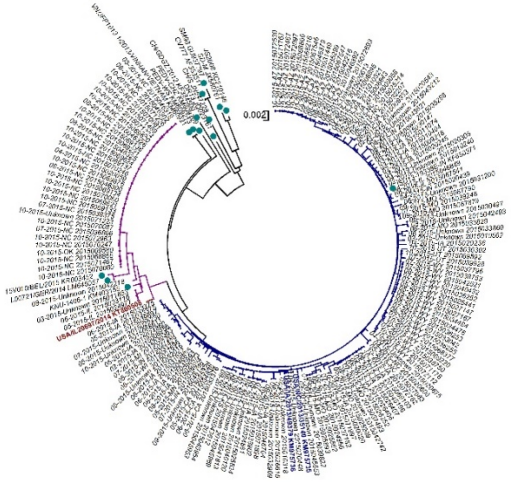
| Sample | AMPH_Ecoli_Bla | Beta-Lactose | | | | | | | | | | Agly | | | | | | | | | |

# Microbiota Analysis



# Next-Generation Sequencing Center at the Iowa State University Veterinary Diagnostic Laboratory

The ISU VDL Next-Generation Sequencing Center offers a range of services based on genomic epidemiological, metagenomics, and transcriptomic analyses using Next-Generation Sequencing Technology combined with bioinformatics analyses. This novel and powerful technology provides a potentially game-changing opportunity in veterinary diagnostic and clinical medicine



**Services provided:**

- Novel/uncharacterized pathogen detection — $400.00/library/sample
- Mixed infection detection — $400.00/library/sample
- Viral whole genome sequencing — $300.00/cell culture isolate sample; $300.00/infected or clinical tissue sample*
- 16s metagenomics analysis — Inquiry
- Shotgun metagenomics sequencing — $300.00/sample ($150.00/sample if sample number is more than 96)
- Bacterial genomic epidemiological analysis — $200.00/strain ($100.00/strain if strain number is more than 96)
  - Phylogenetic analysis based on maximum-common genome sequence
  - Virulome (a set of virulence genes) analysis
  - Resistome (a set of antibiotic resistance genes) analysis

**Turnaround time: 1-2 weeks (bacterial genomic analysis: 1-2 months). The service fee includes costs for NGS library preparation, library sequencing, and bioinformatics analysis, but not the submission fee. *Please inquire if the Ct value of the clinical samples is more than 24.**

**Advantages:**

- The hypothesis-free metagenomics strategy enables NGS to simultaneously detect mixed infections of different microorganisms and identify novel and/or uncharacterized pathogens directly from clinical samples requiring no prior knowledge of the pathogen(s).
- Compared to traditional epidemiology, genomic epidemiology based on whole genome sequences is the ultimate source of information and may never be superseded.


Contact Information

*Next-Generation Sequencing Center, Veterinary Diagnostic Laboratory*

College of Veterinary Medicine
Iowa State University
1802 University Blvd., VMRI#1
Ames, IA 50011

Phone: 515-294-3358
Email: liganwu@iastate.edu