# MOLECULAR SURVEILLANCE OF HEPATITIS C

## STANDARDS FOR MOLECULAR EPIDEMIOLOGICAL ANALYSES

**Yury Khudyakov**

**Molecular Epidemiology & Bioinformatics Lab**
**Laboratory Branch**
**Division of Viral Hepatitis**
**National Center for HIV/AIDS, Viral Hepatitis, STD & TB Prevention**

# Molecular Surveillance Hepatitis C

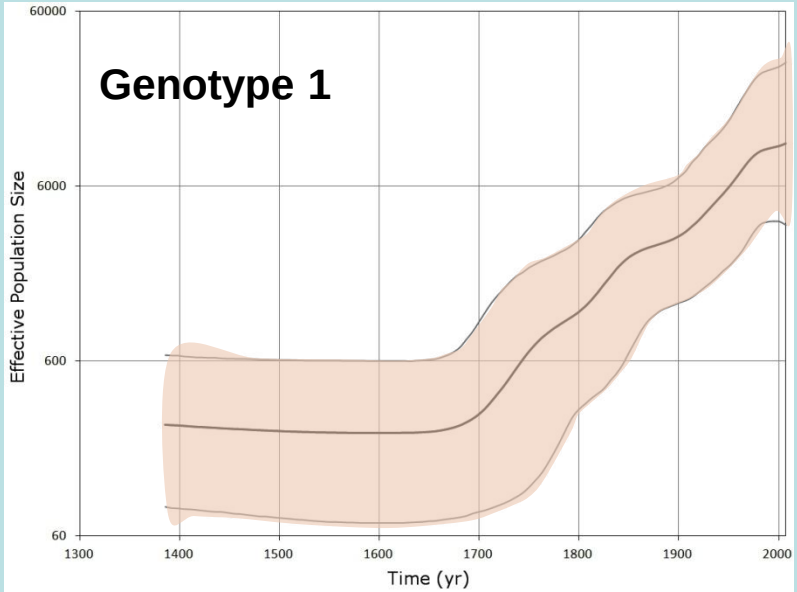
**Goals:**

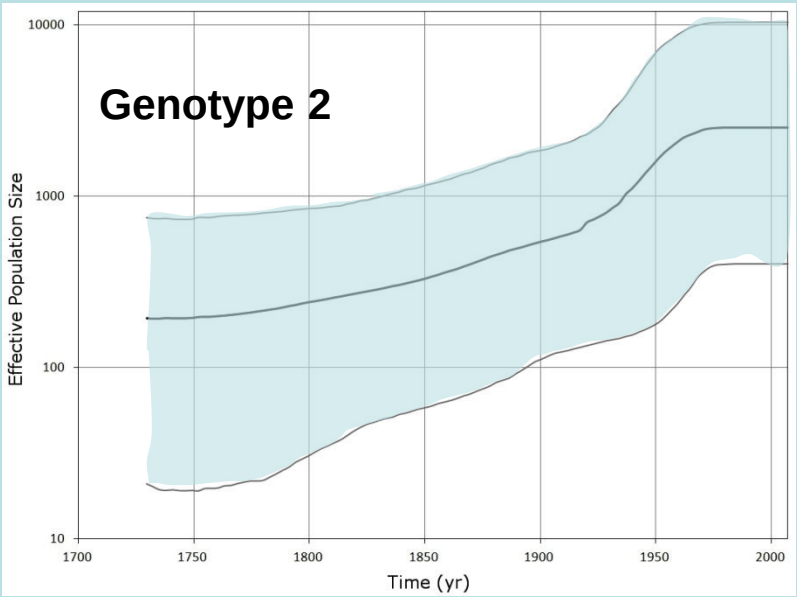
- Infection Control
- Effectiveness of PH Measures
- Decision making

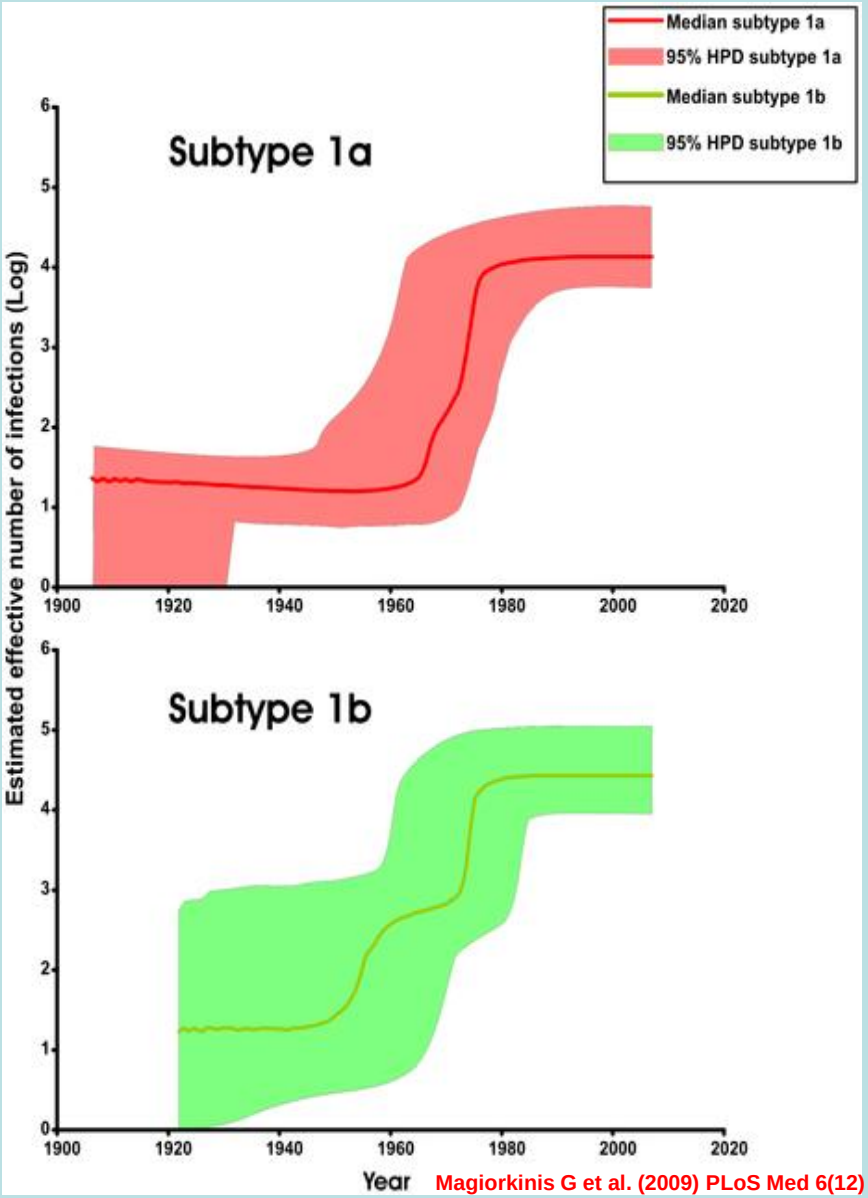
**Methods:**

- Strain Identification
- Strain Characterization

# HCV Epidemic History



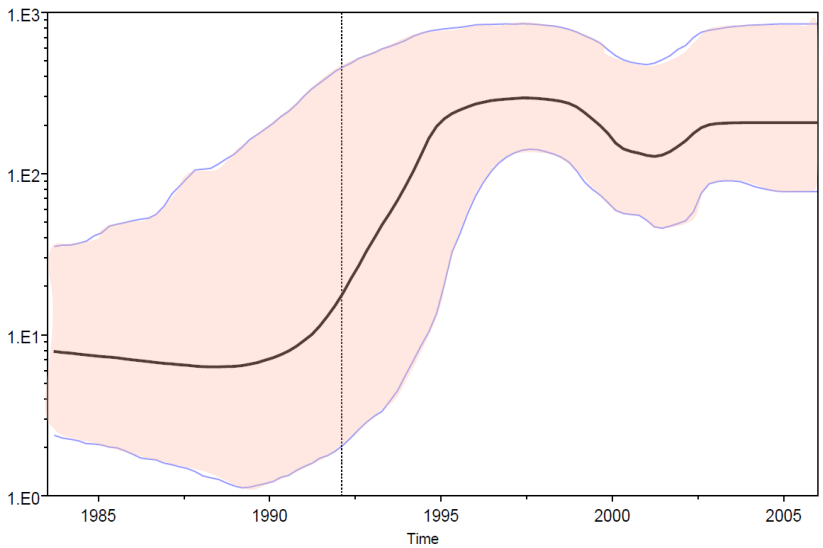




Magiorkinis G et al. (2009) PLoS Med 6(12)

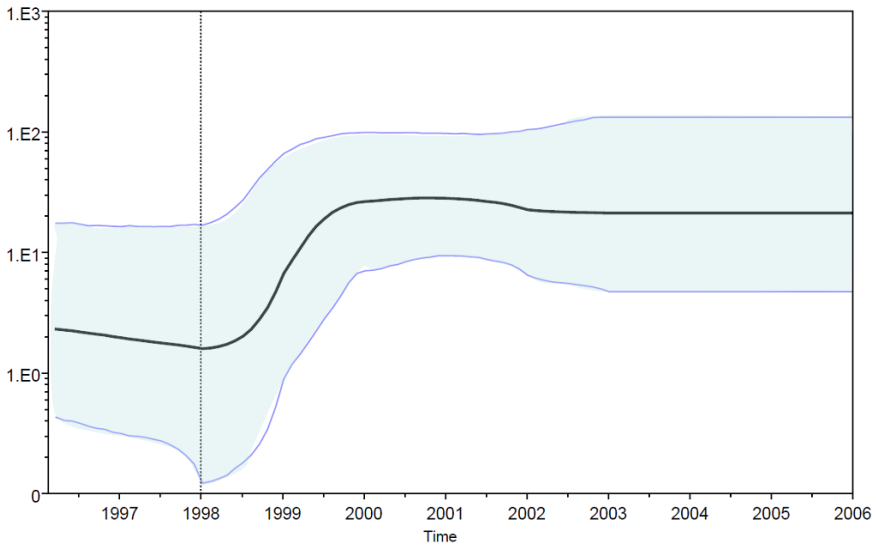
# HBV Epidemic History in the United States

## estimated from whole-genome sequences

### HBV A2

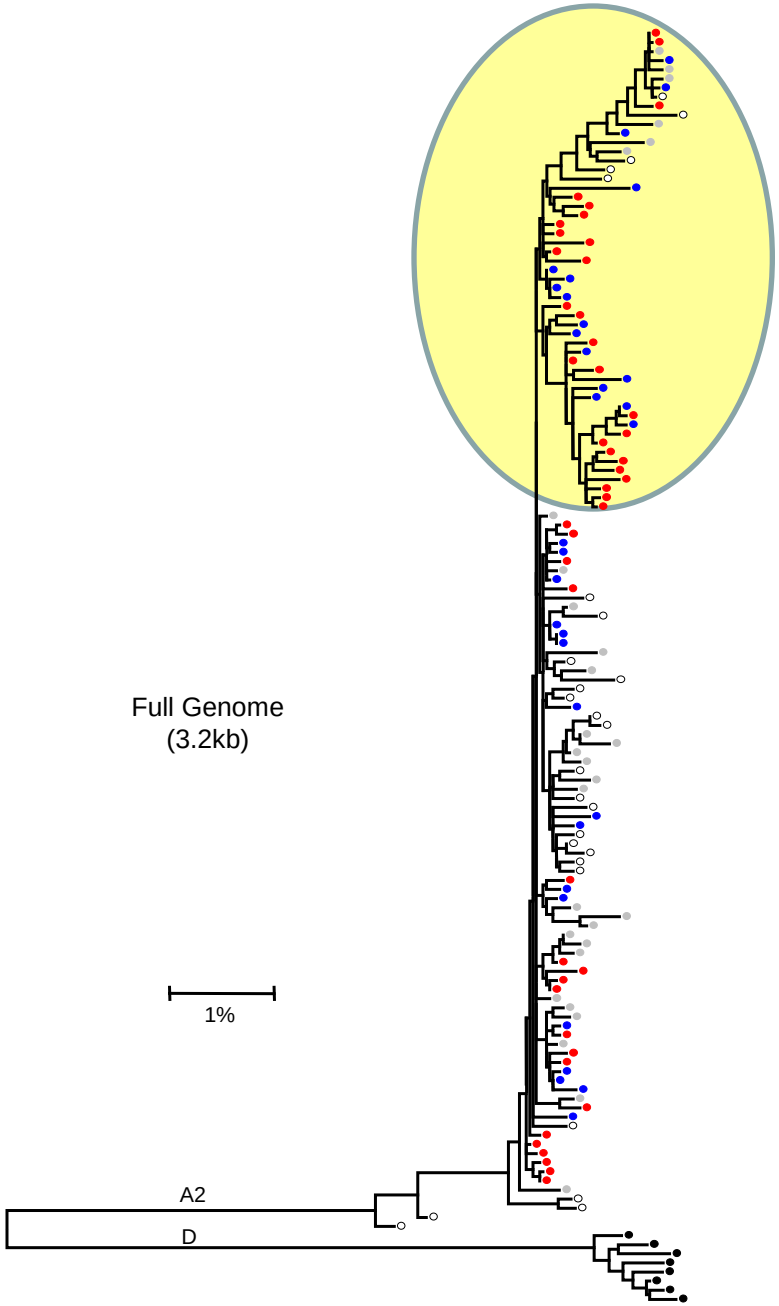


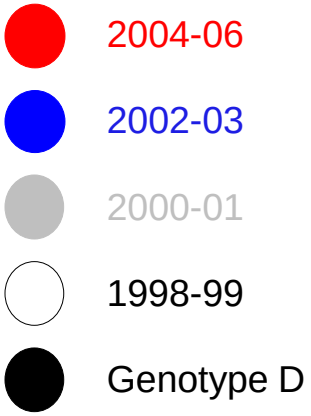

### HBV D3

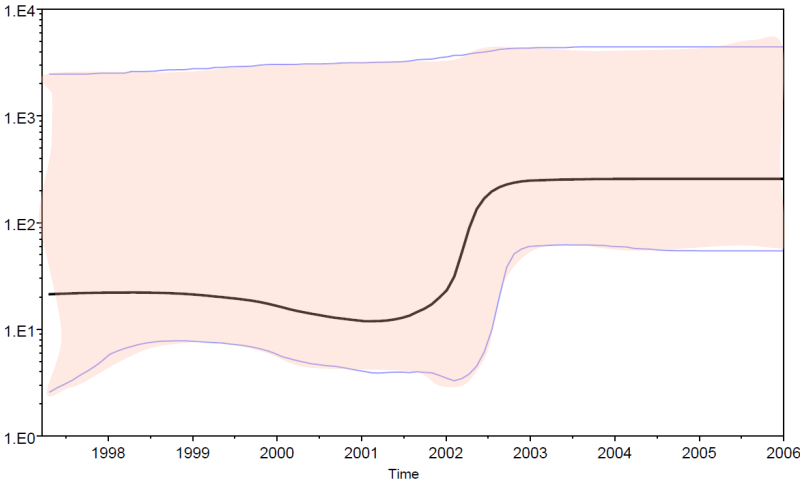



**Sentinel Counties**

# Epidemic History Estimated for a Single HBV A2 Cluster



# Molecular Surveillance Hepatitis C

-**Transmission**

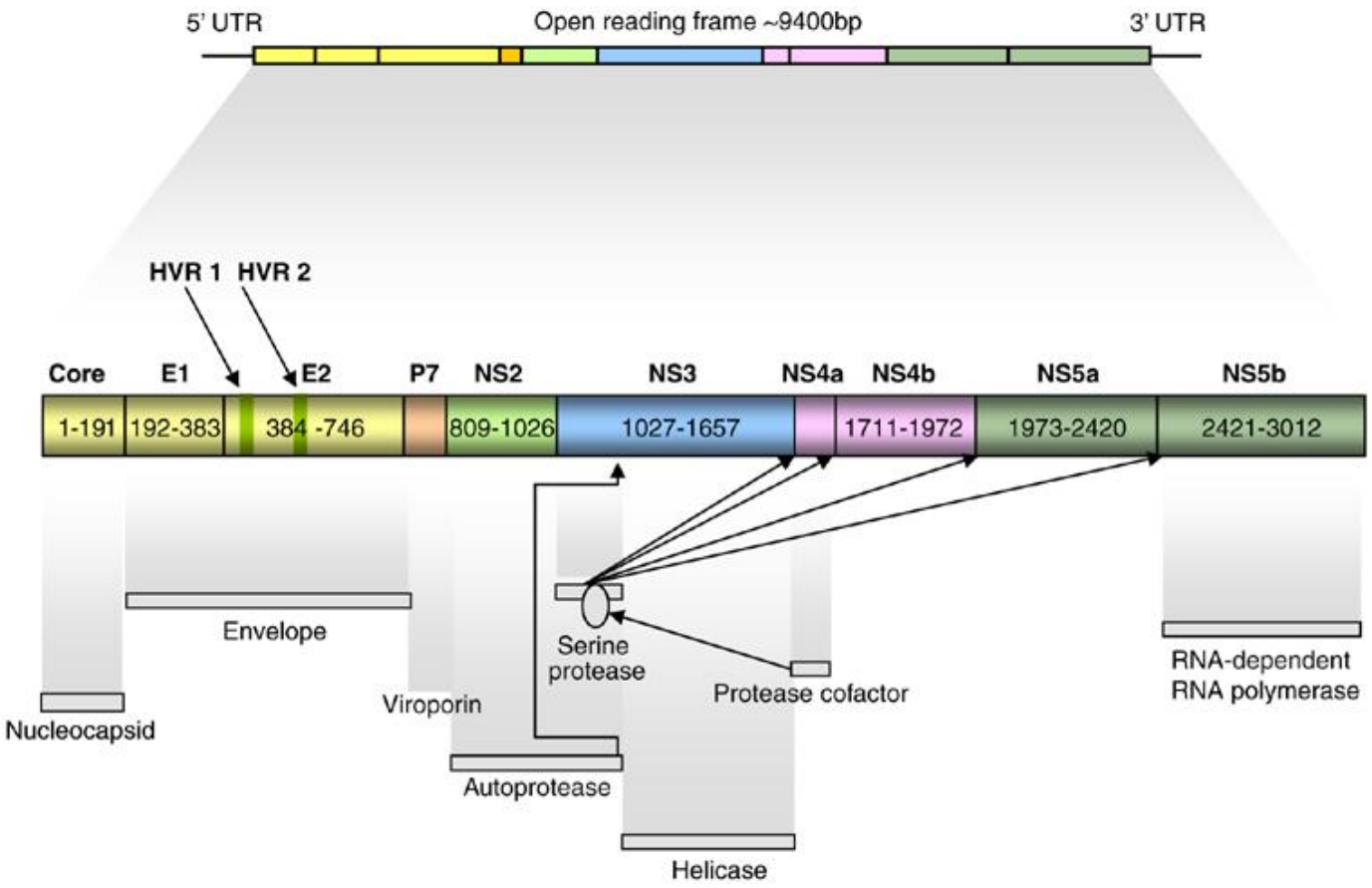
- **Virulence**

# Transmission

**Strain Identity**

**Outbreak Investigation**

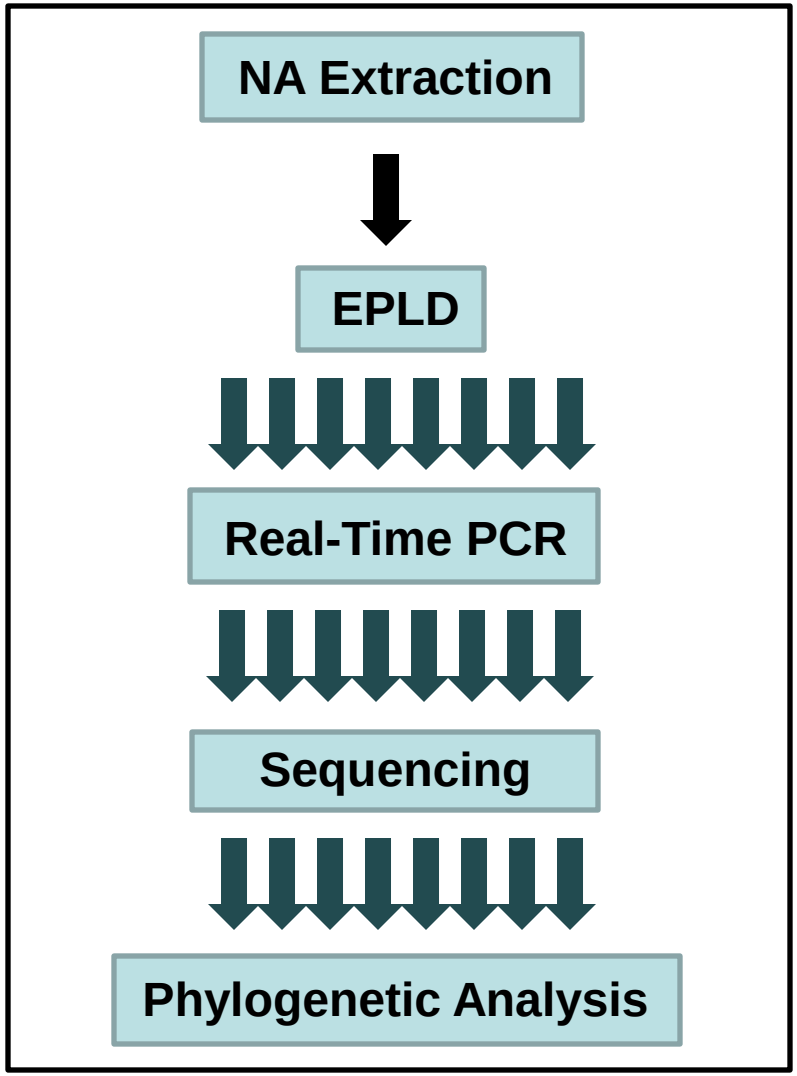
# HCV GENOME




Lloyd et al. 2007. Host and viral factors in the immunopathogenesis of primary hepatitis C virus infection. Immunology and Cell Biology. 85, 24–32.

# QUASISPECIES ANALYSIS

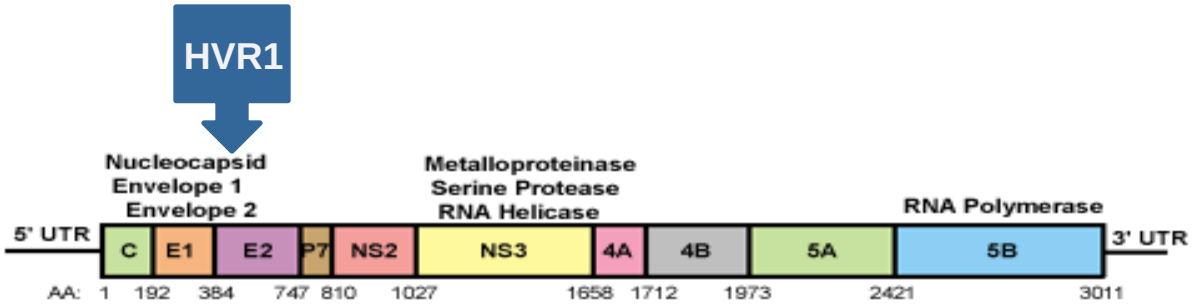
## END-POINT LIMITING-DILUTION REAL-TIME PCR ASSAY

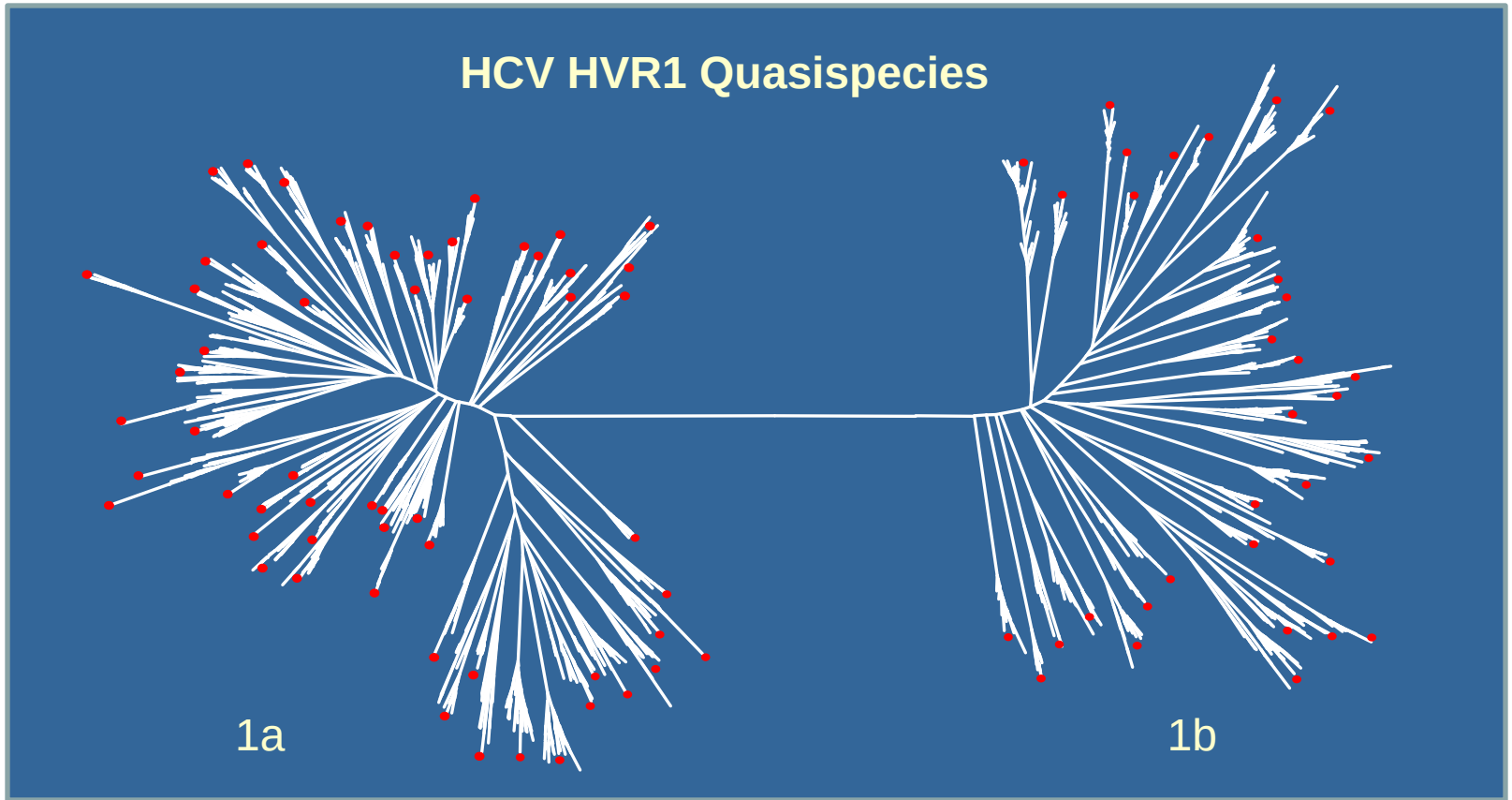



Ramachandran, S. et al. End-point limiting-dilution real-time PCR assay for evaluation of hepatitis C virus quasispecies in serum: Performance under optimal and suboptimal conditions. J. Virol. Methods (2008) 151, 217-224.

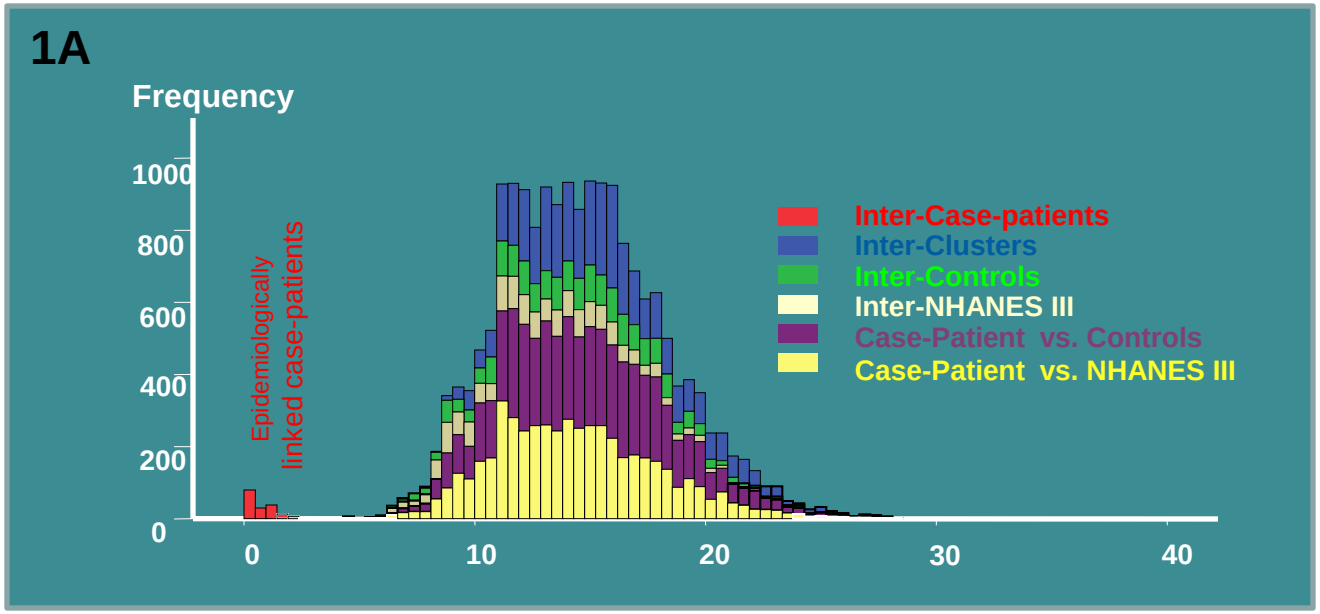
# QUASISPECIES ANALYSIS

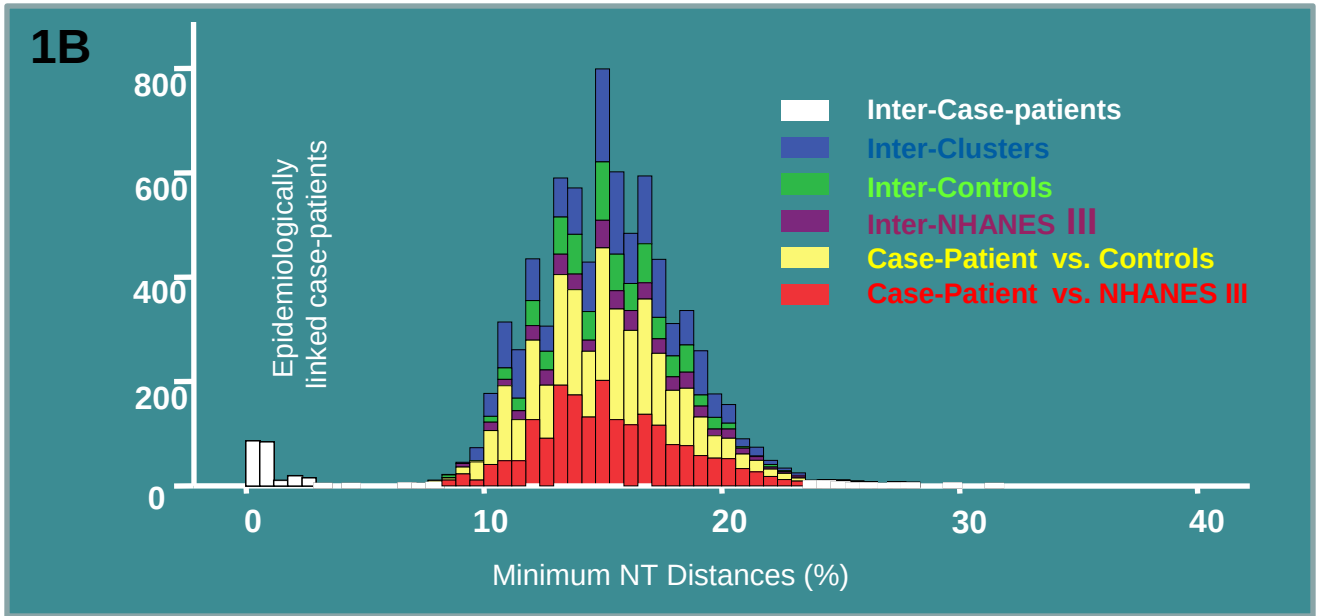
## OUTBREAK INVESTIGATION



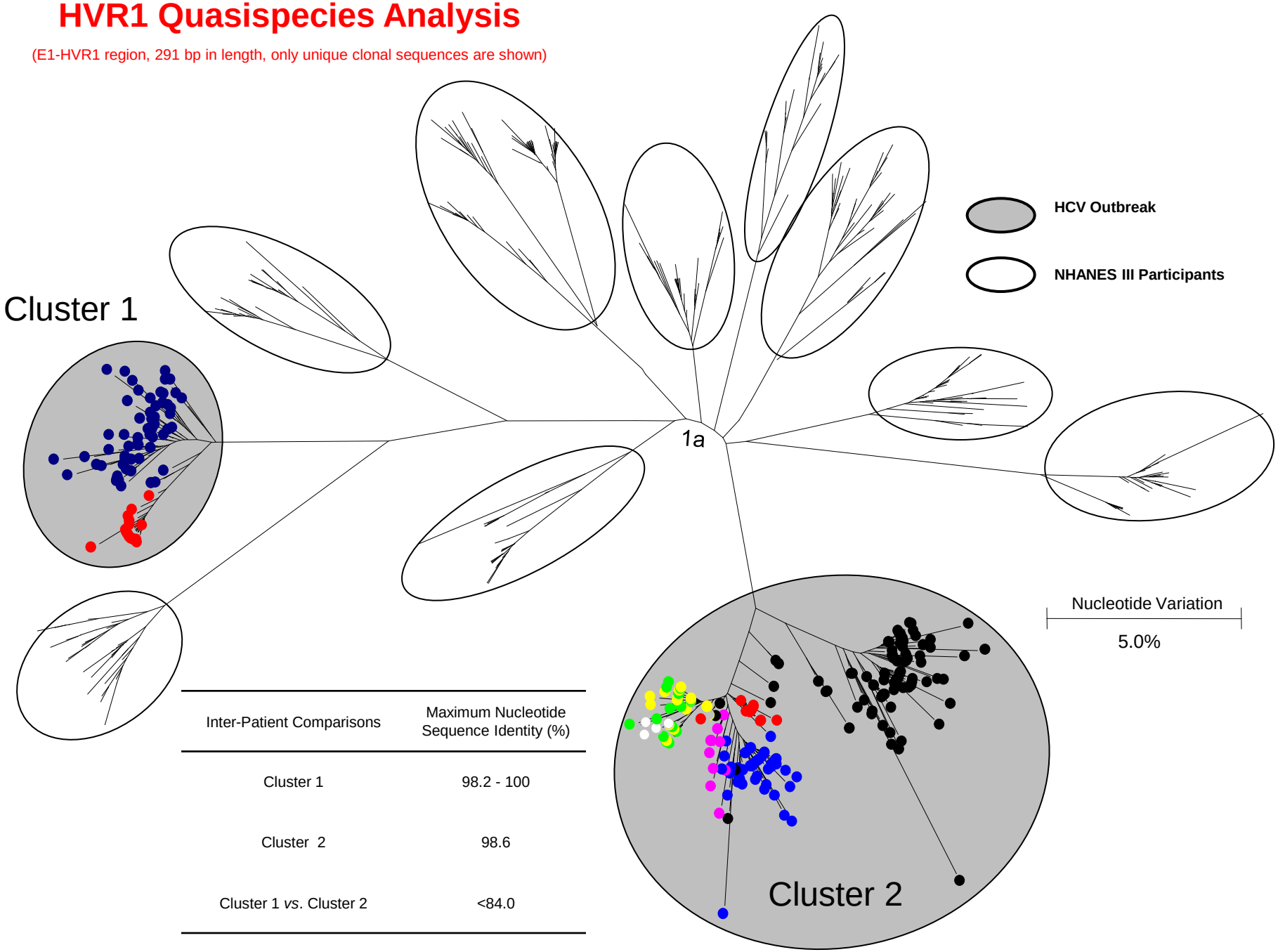
# Distribution of Minimum NT Distances



# HVR1 Quasispecies Analysis

(E1-HVR1 region, 291 bp in length, only unique clonal sequences are shown)




| Inter-Patient Comparisons | Maximum Nucleotide Sequence Identity (%) |
|---|---|
| Cluster 1 | 98.2 - 100 |
| Cluster 2 | 98.6 |
| Cluster 1 *vs*. Cluster 2 | <84.0 |

# Transmission

## Throughput Increase

- Next Generation Sequencing
- Mass Spectrometry

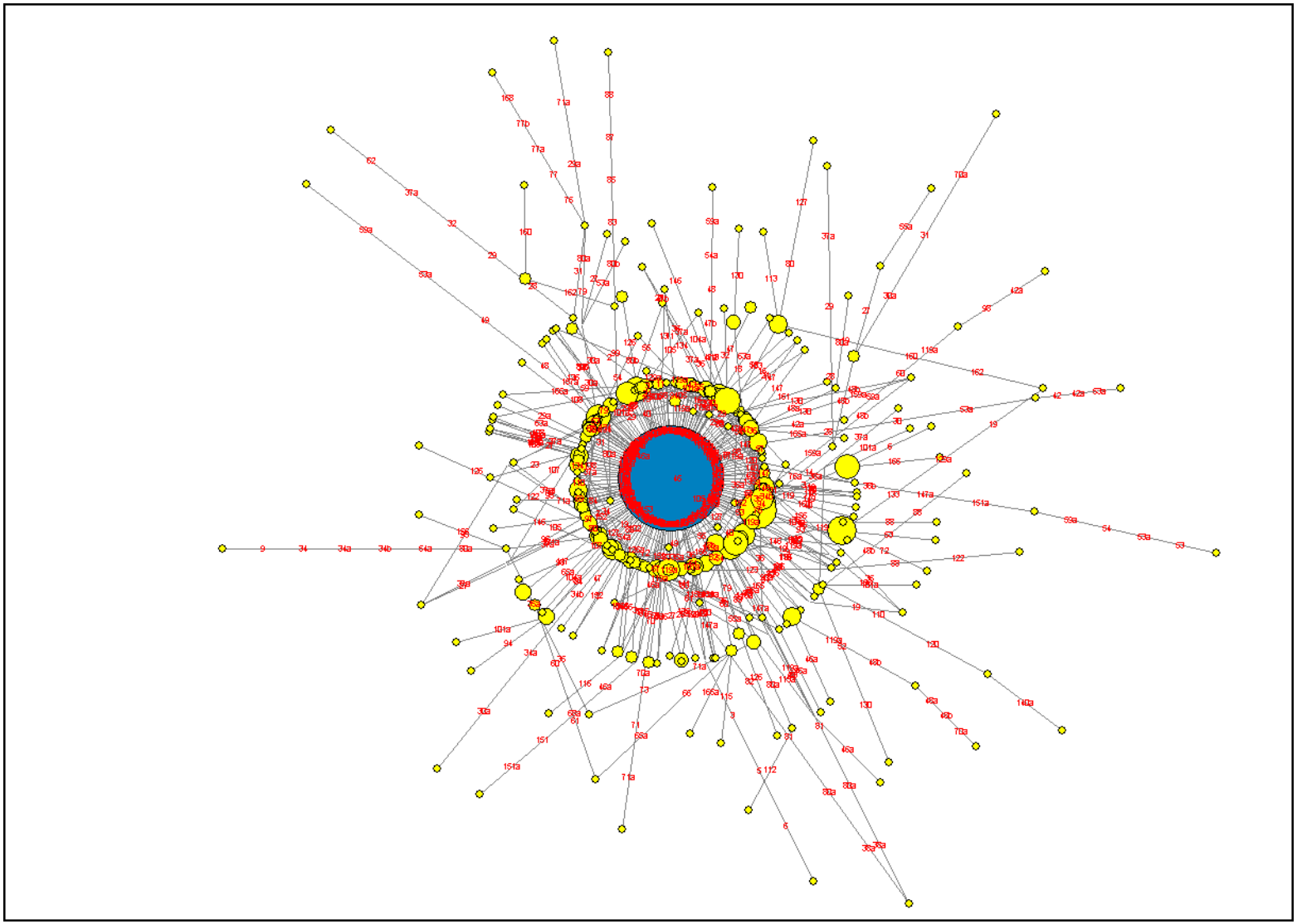
# Next Generation Sequencing

- Read length
- Errors
- Data processing
- Throughput

# Single-Clone Sequencing

## Median Joining Network

unique sequences

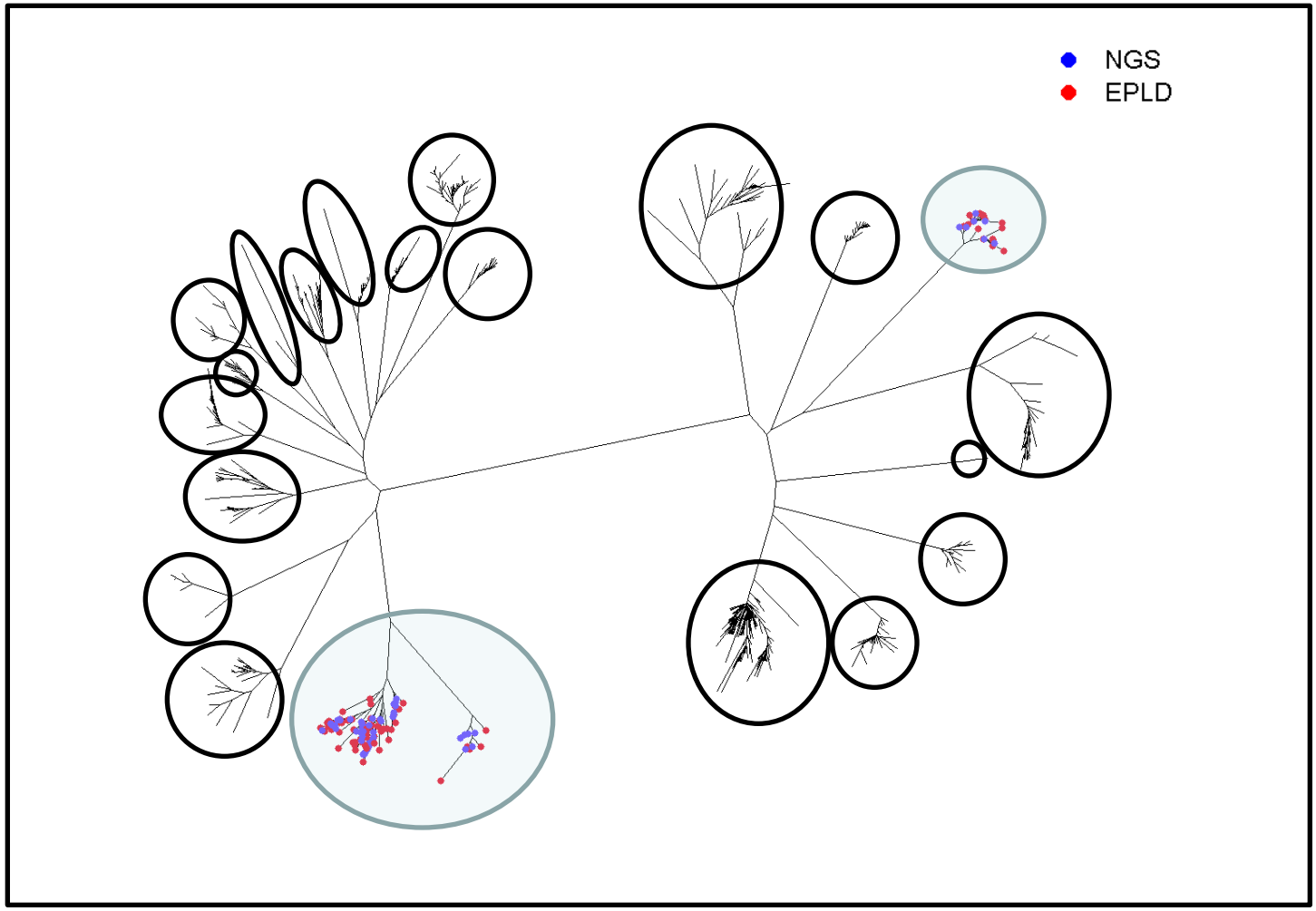


# Next Generation Sequencing

- Read length
- Errors
- Data processing
- Throughput

# HVR1 Quasispecies Analysis

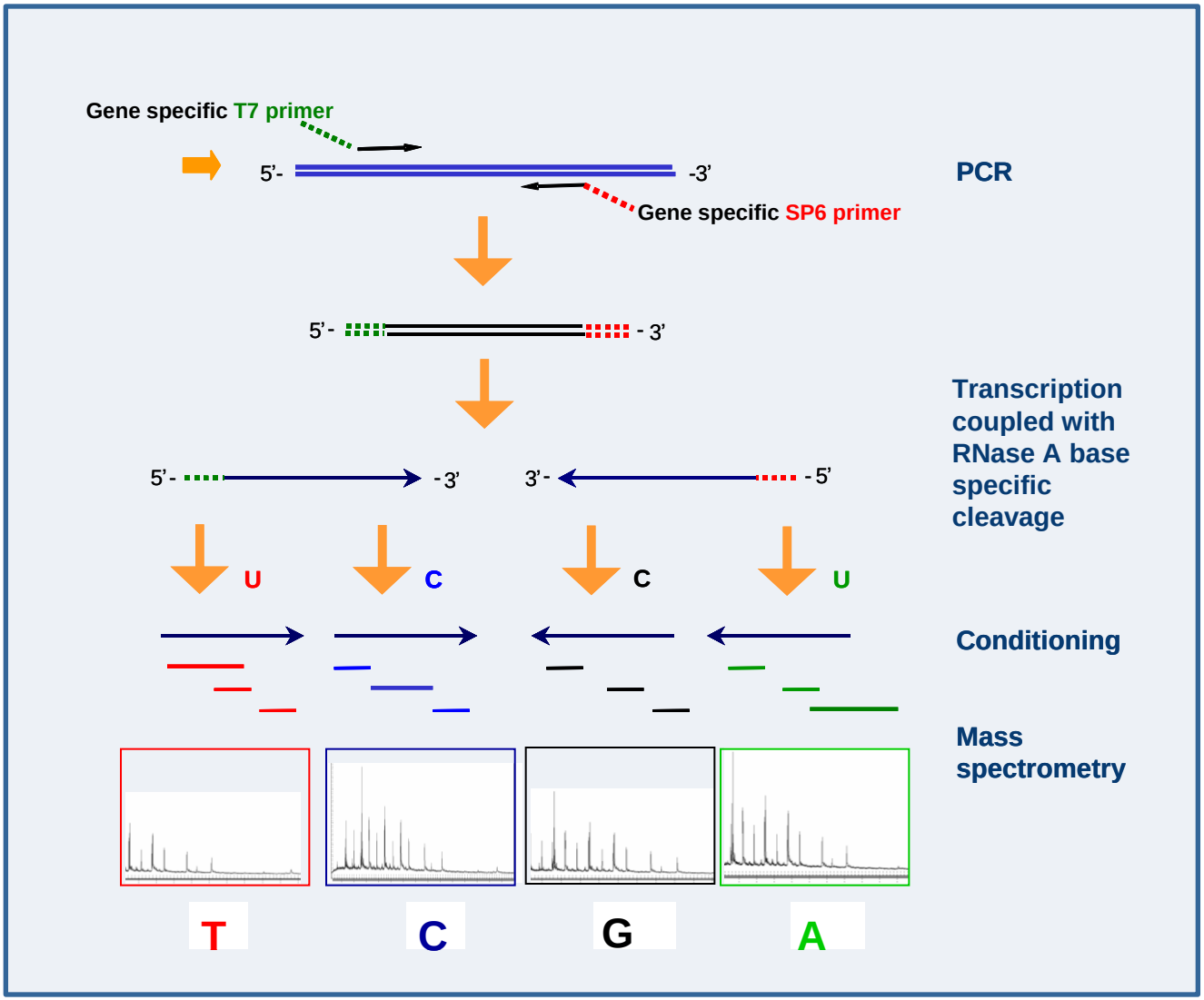
## Next-Generation Sequencing *vs.* EPLD-Sequencing



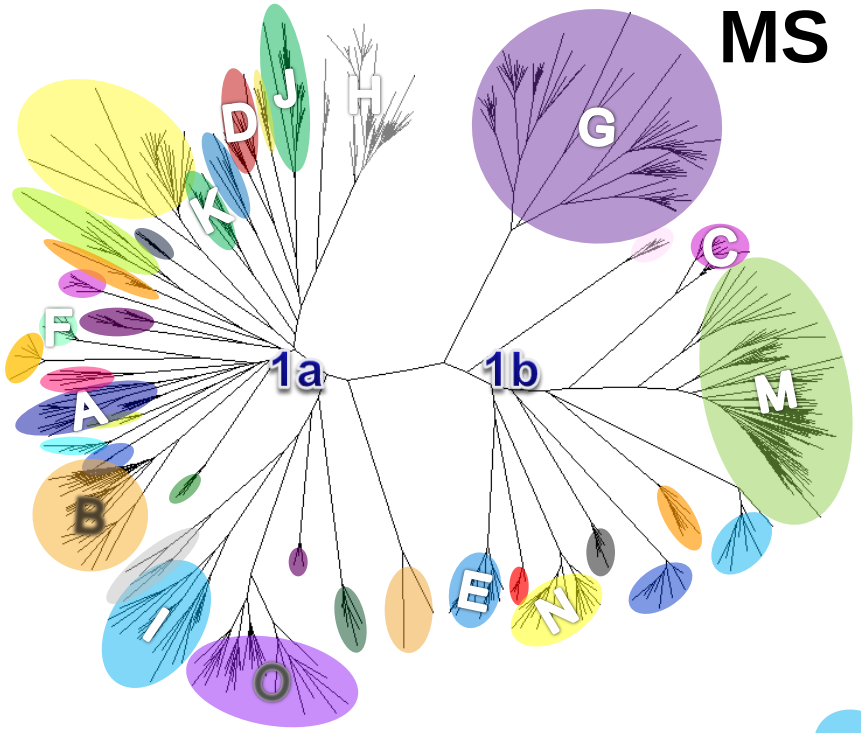
# Mass Spectrometry

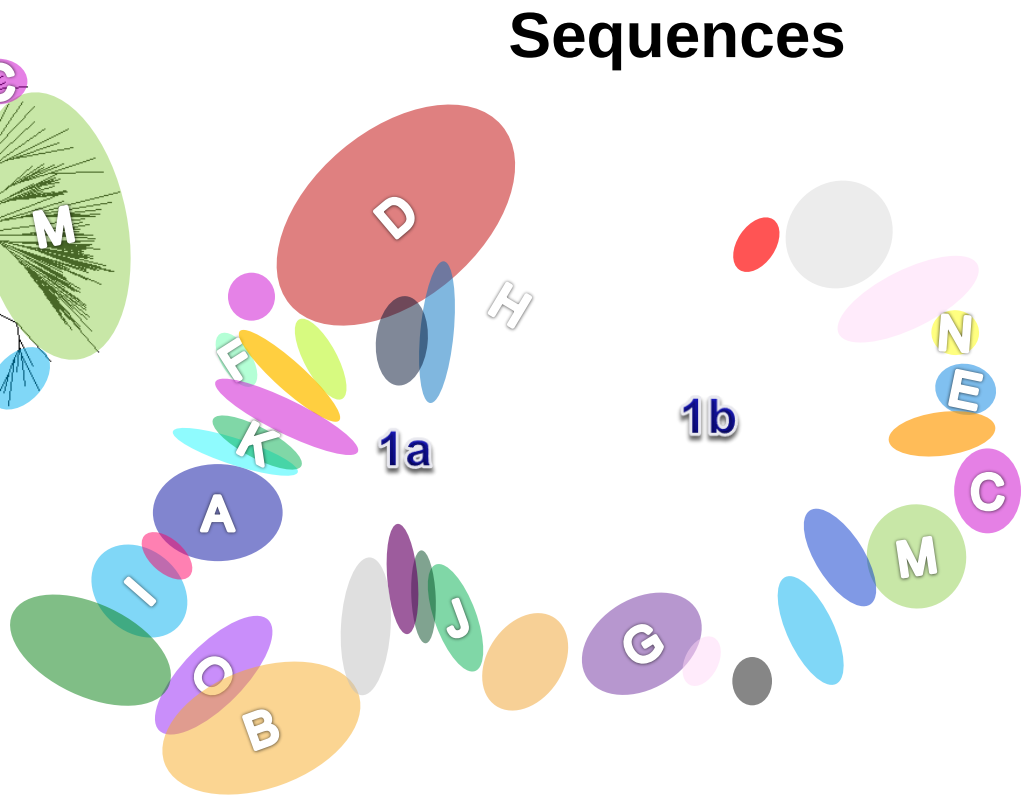
- ~~Read length~~
- ~~Errors~~
- Data processing
- Scale flexibility
- Throughput

# Nucleic Acid MassCLEAVE ™ and MassARRAY™

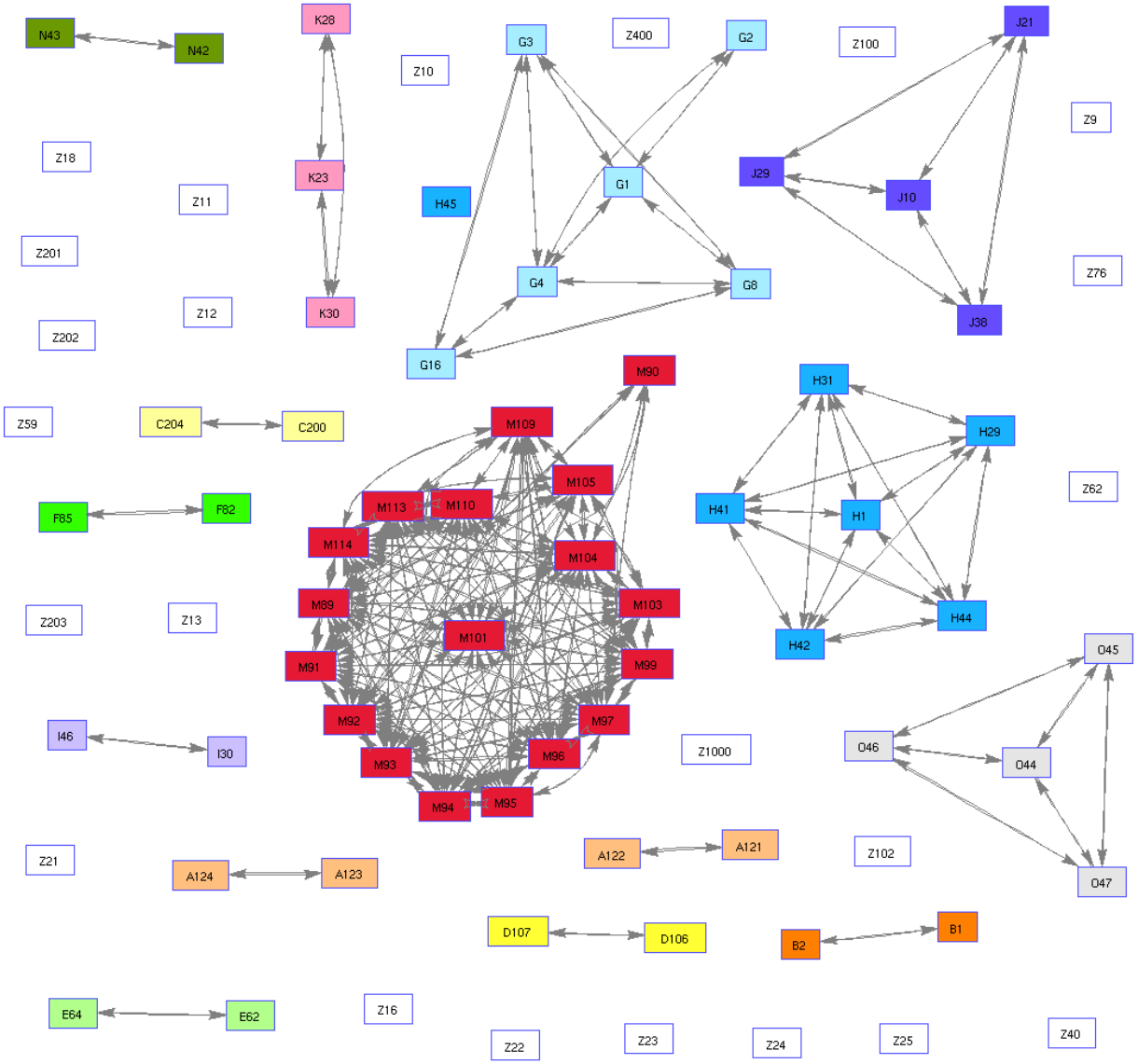


# NEIGHBOR JOINING TREES



# Linkage Networks of HCV Strains Involved in Outbreaks



# Virulence

## HCV Interferon Resistance

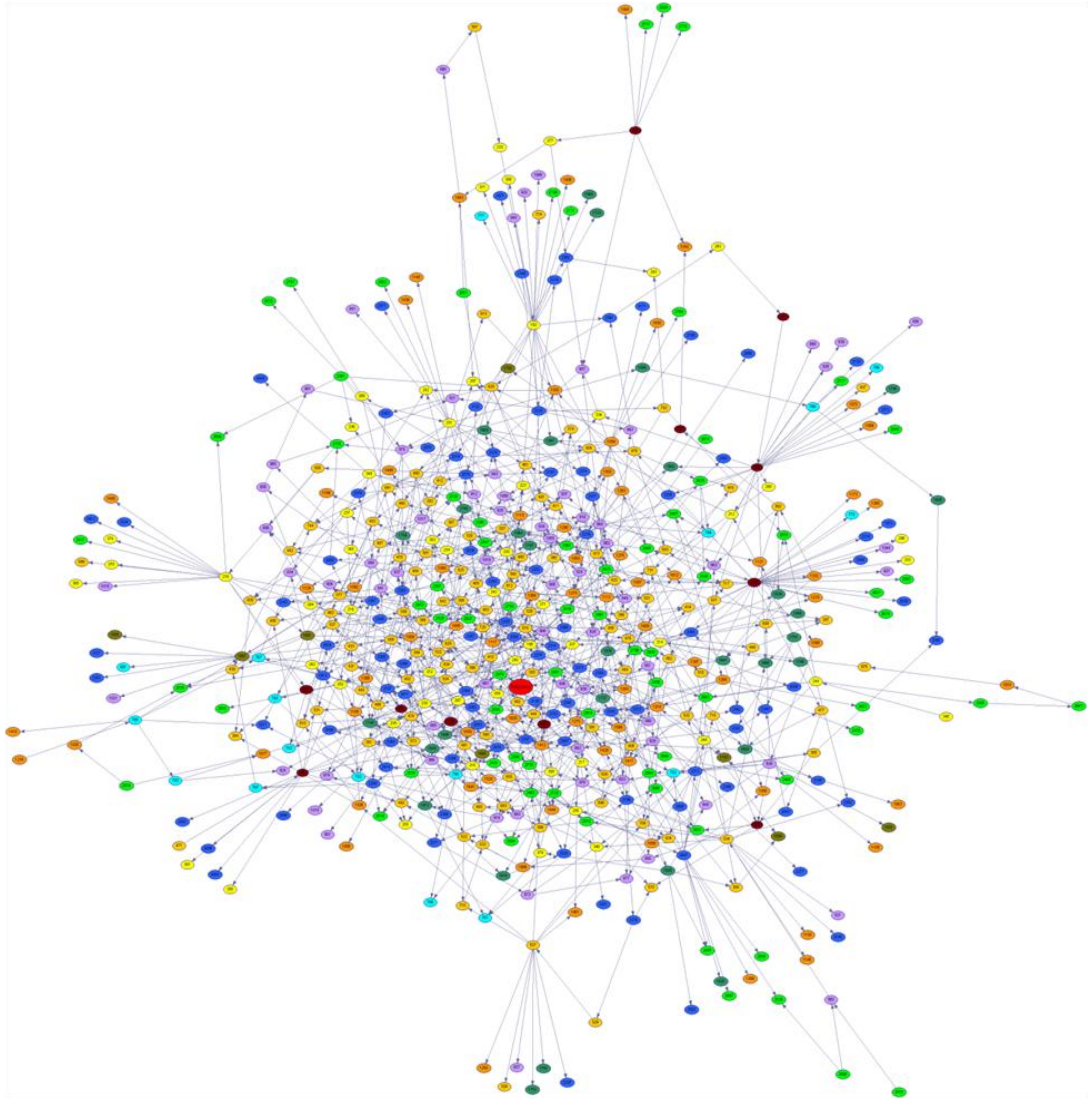
# IDENTIFICATION OF INTERFERON RESISTANCE

## Molecular Markers

**-Host genetic markers**

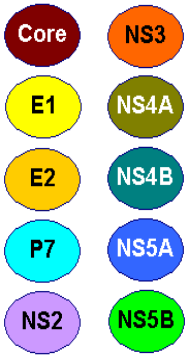
**-HCV genetic markers**

# NETWORK
# OF INTERACTIONS AMONG HCV POLYPROTEIN SITES AND THERAPY OUTCOME



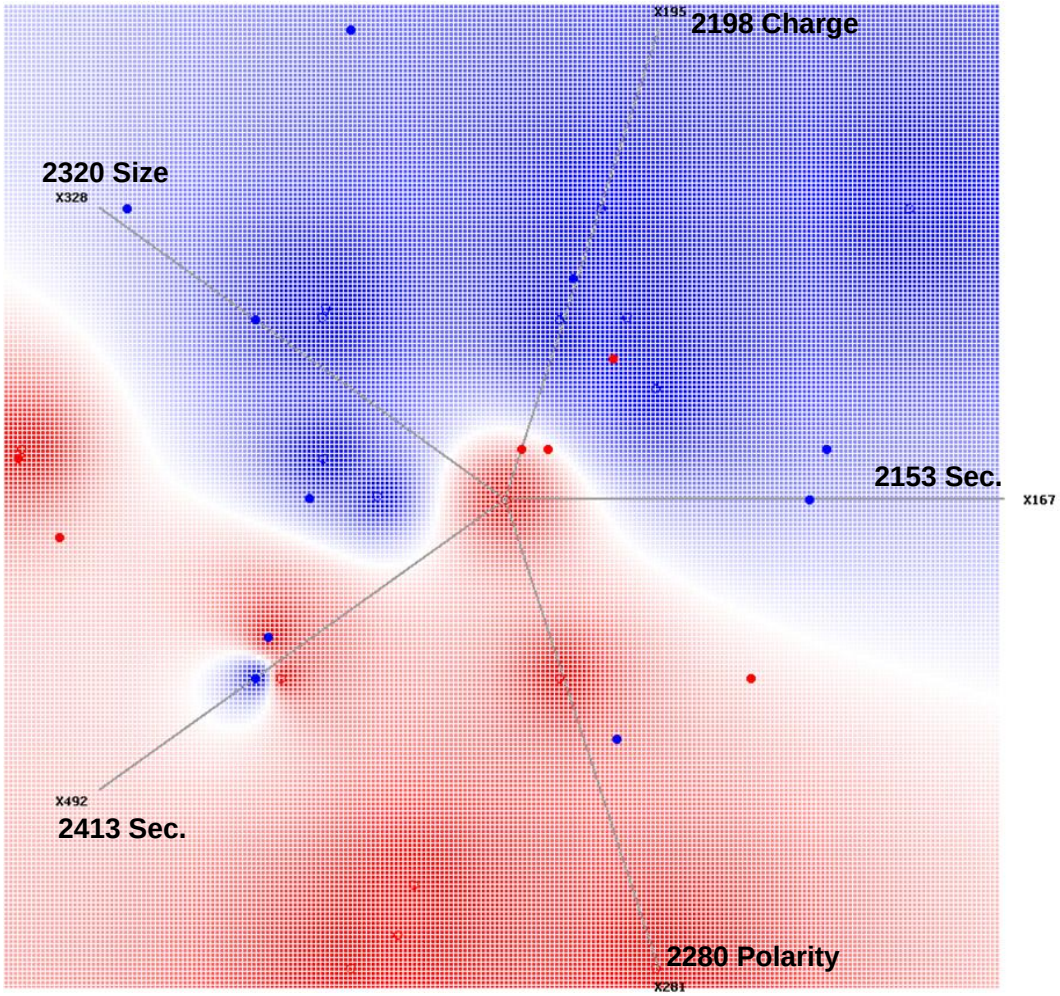
**PROPERTIES OF THE HCV POLYPROTEIN-BN**

| Region | No. Sites | Outdegree | Indegree | Edges to response |
|---|---|---|---|---|
| Core | 11 | 136 | 11 | 0 |
| E1 | 55 | 248 | 113 | 0 |
| E2 | 118 | 239 | 419 | 0 |
| P7 | 17 | 45 | 39 | 0 |
| NS2 | 67 | 151 | 154 | 2 |
| NS3 | 73 | 135 | 144 | 0 |
| NS4A | 8 | 19 | 17 | 0 |
| NS4B | 30 | 71 | 57 | 1 |
| NS5A | 100 | 165 | 217 | 2 |
| NS5B | 72 | 107 | 143 | 1 |
| Polyprotein | 551 | 1318* | -- | 6 |




- **Extensive coordination**
- **Topological contribution**
  - **Sites**
  - **Proteins**
- **Direct links to outcome**
- **Significant coevolution**

# GENETIC MARKERS FROM HCV NS5A
## FOR PREDICTION OF INTERFERON RESISTANCE




| Classifiers | Evaluation | Validation (% accuracy) | | |
|---|---|---|---|---|
| | | Overall | NR class | BT |
| DTNB | 97.5† | 72.2 | 75.0 | 66.7 |
| Linear Projection | 95.2* | **83.3** | **83.3** | **83.3** |

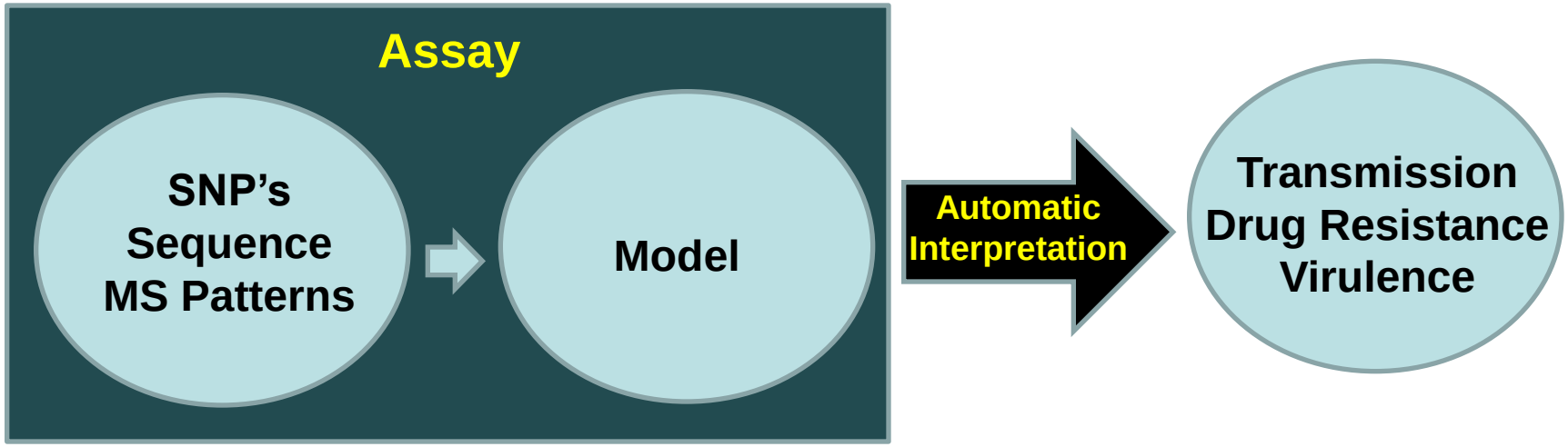
**NS5A features (n=5)**

**Site 2153 – Secondary Structure**
**Site 2198 – Charge**
**Site 2280 – Polarity**
**Site 2320 – Size**
**Site 2413 – Secondary Structure**

**NS5A from HALT-C: 12 NR and 6BT**

# Molecular and Computational Tools for Molecular Surveillance

**“*New*” standards for molecular epidemiological analyses**




- Accuracy
  - Sensitivity
  - Specificity
- Efficacy
  - Throughput
  - Time
  - Cost
  - Automation
  - Interpretation
    - Operator independent
    - Uniform

# Molecular Surveillance

- Detection and characterization of HCV strains
- Transmission and Virulence
- Novel Assays
  - Accurate
  - High throughput
  - Cost-effective
- Discovery of novel markers
  - Molecular data
  - Automated interpretation

**David Campo**
**Zoya Dimitrova**
**Joseph Forbi**
**Lilia Ganova-Raeva**
**James Lara**
**Yulin Lin**
**Michael Purdy**
**Sumathi Ramachandran**
**Pavel Skums**
**Hong Thai**
**Gilberto Vaughan**
**Guo-liang Xia**

***Molecular Epidemiology and Bioinformatics Lab***

**Division of Viral Hepatitis, NCHHSTP, CDC, Atlanta, GA**