Finding the known and unknown, plant waterborne virus in Oklahoma water

Jon Daniels,

Suat Kaimak,
Francisco Ochoa Corona

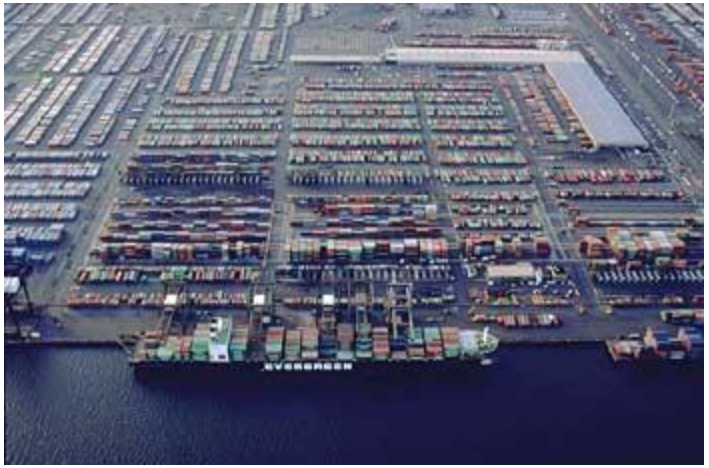
Oklahoma State University

National Institute for Microbial Forensics & Food and Agricultural Biosecurity

Dept. of Entomology & Plant Pathology

The predicament…

U.S. imports ~ 40 million tons of agricultural products



Only 2% inspected:
> 39 million tons uninspected

Plus:
Tourism, immigration, weather

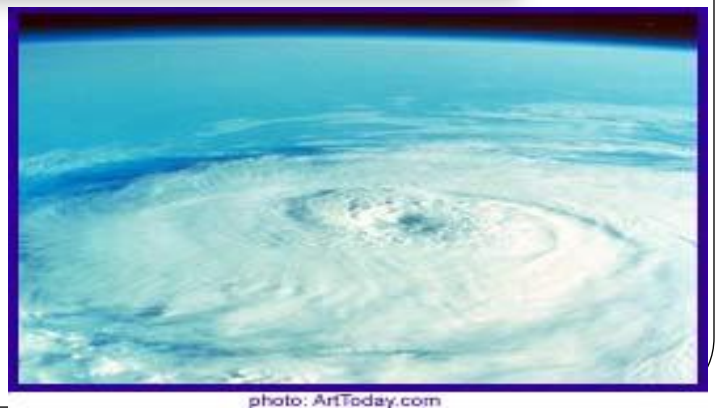


photo: ArtToday.com

The result

Citrus Greening

Plum Pox Virus

Citrus Leprosis

Soybean Rust

Canna Yellow Mosaic Virus

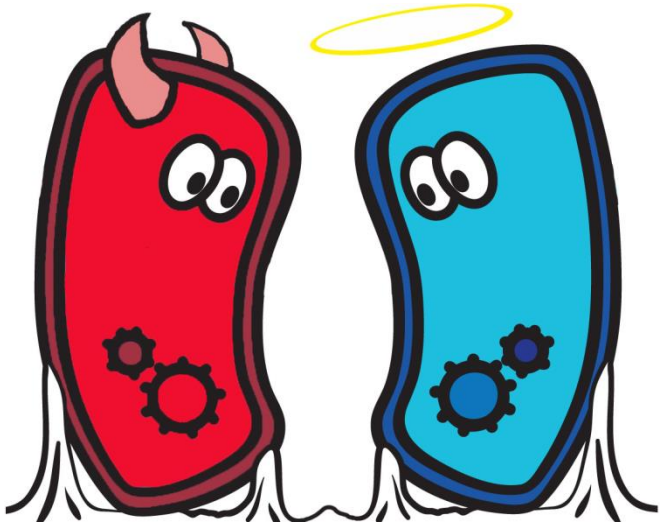
Brown spot

Lethal Yellowing of Palm

Southern Wilt

Challenges in plant pathogen diagnostics

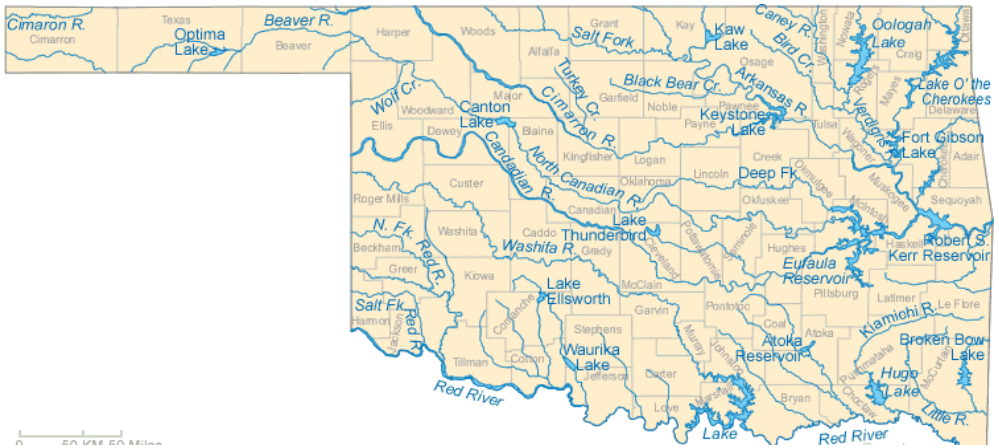
- **Multiple types of plant pathogens**
  - **Viral**
  - Bacterial
  - Fungal
  - Phytoplasma
  - Nematodes

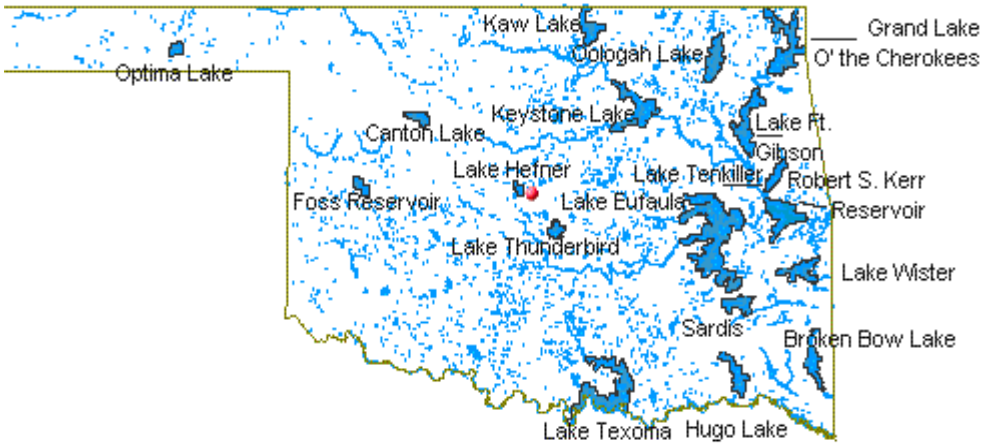


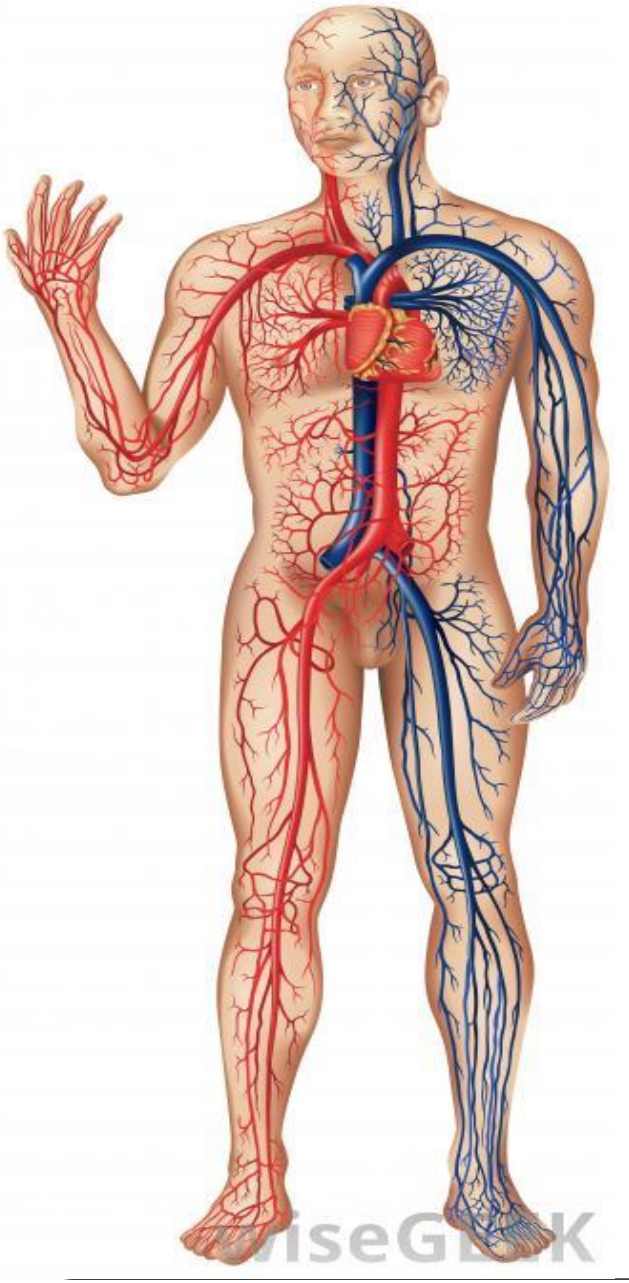
- **Non-pathogenic microbes living in the enviroment**
  - Who are the good guys and who are the bad guys?

Oklahoma water

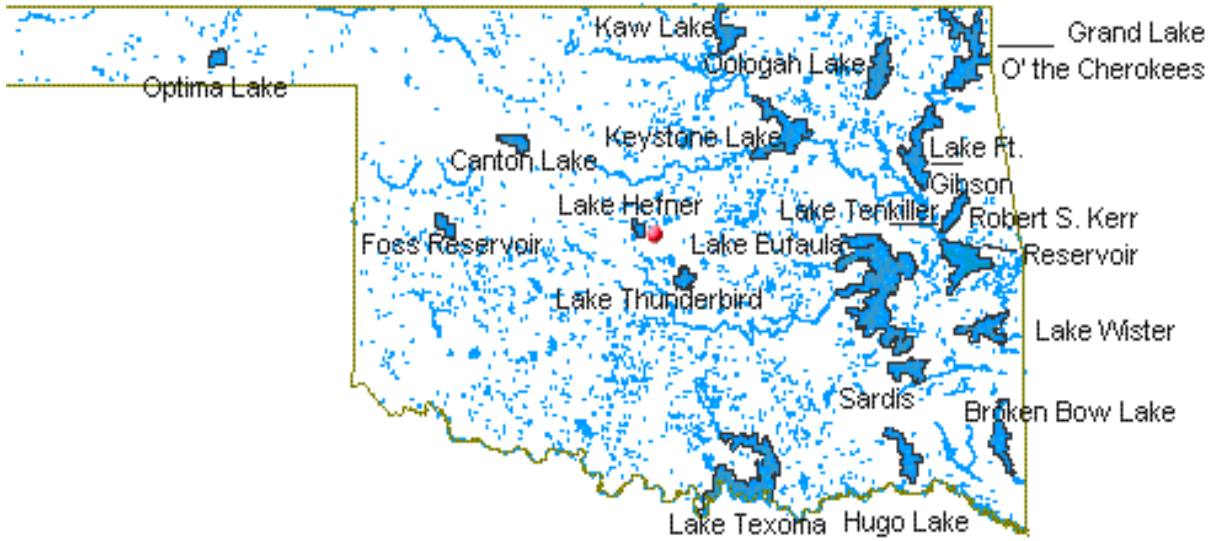
- Over 200 lakes
- Over 1 million surface acres of water
- Numerous rivers, streams & ponds
  - Boating & swimming








Cimaron R.
Cimarron
Texas
Optima Lake
Beaver R.
Beaver
Harper
Woods
Grant
Kay
Alfalfa
Salt Fork
Kaw Lake
Caney R.
Bird Cr.
Oologah Lake
Craig
Osage
Black Bear Cr.
Arkansas R.
Lake O' the Cherokees
Delaware
Wolf Cr.
Major
Garfield
Noble
Pawnee
Canton Lake
Turkey Cr.
Cimarron R.
Woodward
Ellis
Dewey
Blaine
Keystone Lake
Payne
Tulsa
Verdigris
Fort Gibson Lake
Wagoner
Adair
Kingfisher
Logan
Creek
Lincoln
Deep Fk
Canadian R.
North Canadian R.
Oklahoma
Okmulgee
Muskogee
Custer
Roger Mills
Okfuskee
McIntosh
Sequoyah
Canadian
Lake Thunderbird
Washita
Caddo
Grady
Washita R.
N. Fk. Red R.
Beckham
Greer
Kiowa
Pottawatomie
Cleveland
Hughes
Haskell
Robert S. Kerr Reservoir
Eufaula Reservoir
McClain
Pittsburg
Latimer
Le Flore
Lake Ellsworth
Salt Fk. Red R.
Harmon
Jackson
Comanche
Garvin
Pontotoc
Coal
Atoka
Klamichi R.
Broken Bow Lake
Stephens
Murray
Johnston
Atoka Reservoir
Waurika Lake
Tillman
Cotton
Carter
Jefferson
Pushmataha
Hugo Lake
McCurtain
Choctaw
Red River
Marshall
Love
Bryan
Little R.


Kaw Lake
Grand Lake O' the Cherokees
Oologah Lake
Optima Lake
Keystone Lake
Canton Lake
Lake Ft. Gibson
Lake Hefner
Lake Tenkiller
Robert S. Kerr Reservoir
Foss Reservoir
Lake Eufaula
Lake Thunderbird
Lake Wister
Sardis
Broken Bow Lake
Lake Texoma
Hugo Lake

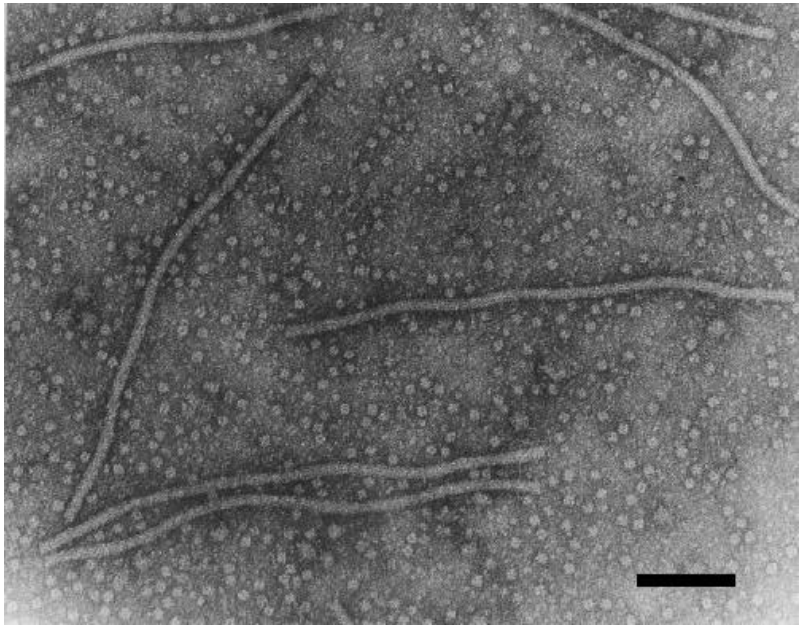
Waterborne viruses

- What do we mean by waterborne pathogens?
  - A ~~organism~~ virus capable of causing disease that can be transmitted via water
- ***Potexviruses***
- ***Tobamovirus***
- ***Tombusviridae***

Waterborne viruses

- What do we mean by waterborne pathogens?
  - A ~~organism~~ virus capable of causing disease that can be transmitted via water

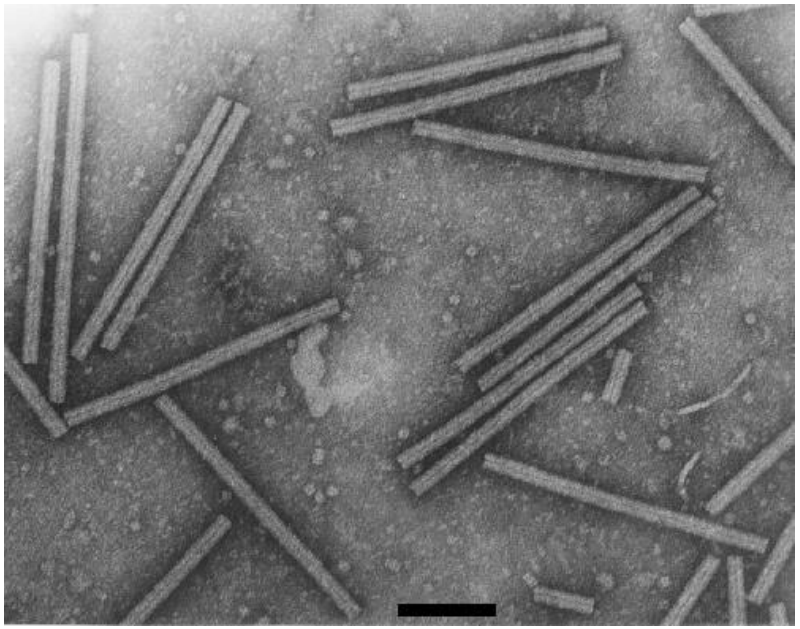
- ***Potexviruses***
- ***Tobamovirus***
- ***Tombusviridae***



Waterborne viruses

- What do we mean by waterborne pathogens?
  - A ~~organism~~ virus capable of causing disease that can be transmitted via water

- ***Potexviruses***
- ***Tobamovirus***

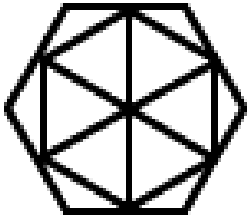


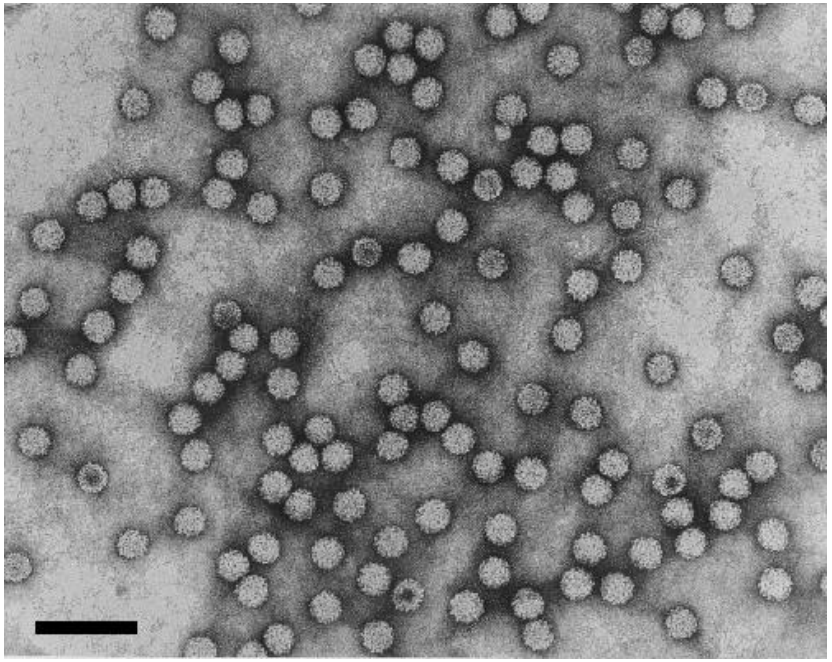
- ***Tombusviridae***

Waterborne viruses

- What do we mean by waterborne pathogens?
  - A ~~organism~~ virus capable of causing disease that can be transmitted via water

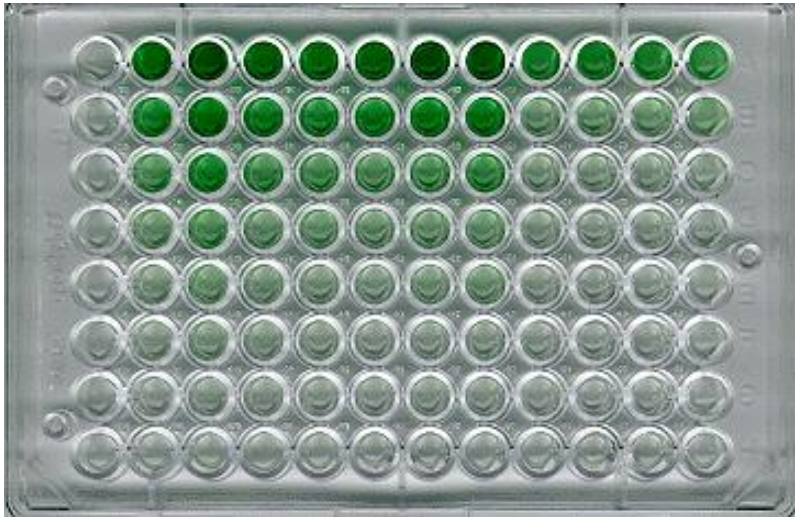
- ***Potexviruses***
- ***Tobamovirus***
- ***Tombusviridae***





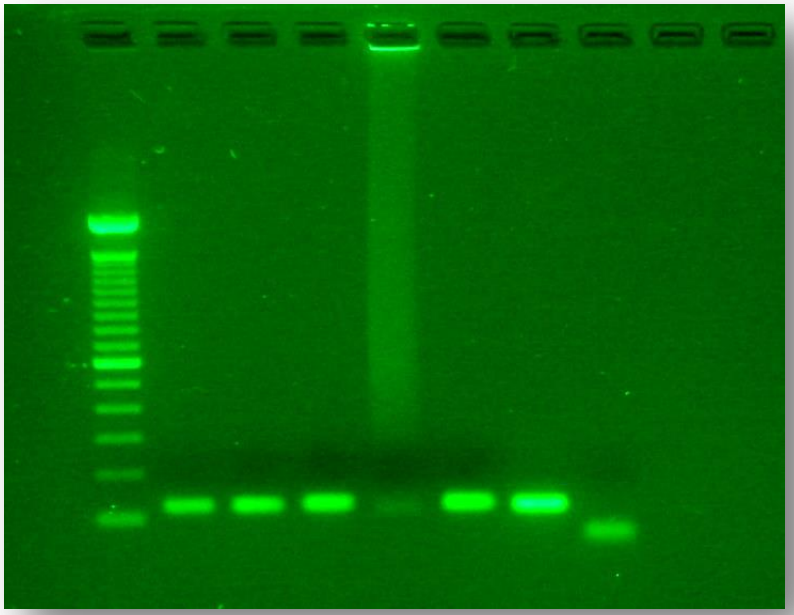
Current plant pathogen detection technologies

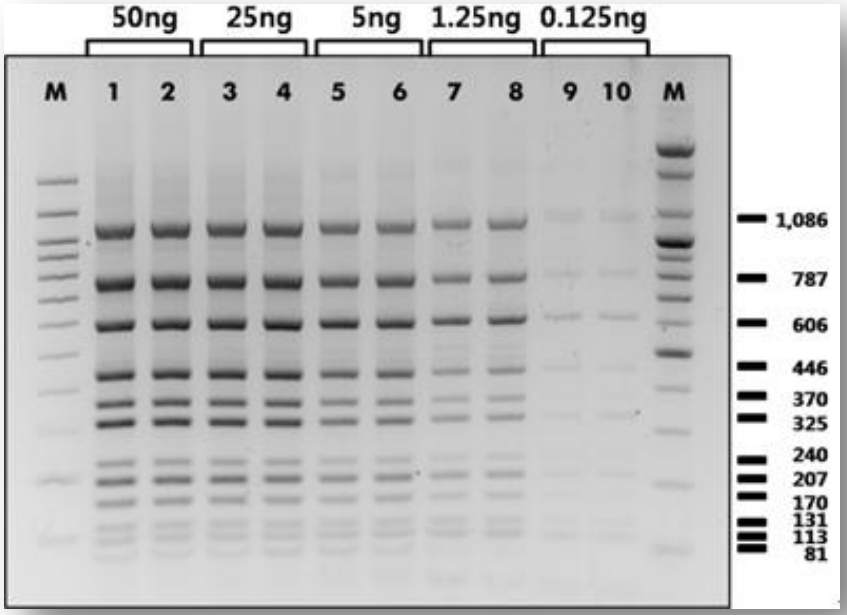
- **Protein based immunological assays**
  - Immunostrip test
  - ELISA



Current plant pathogen detection technologies

- **Nucleic acid based assays**
  - Polymerase chain reaction (PCR)
  - qPCR
  - Multiplex PCR



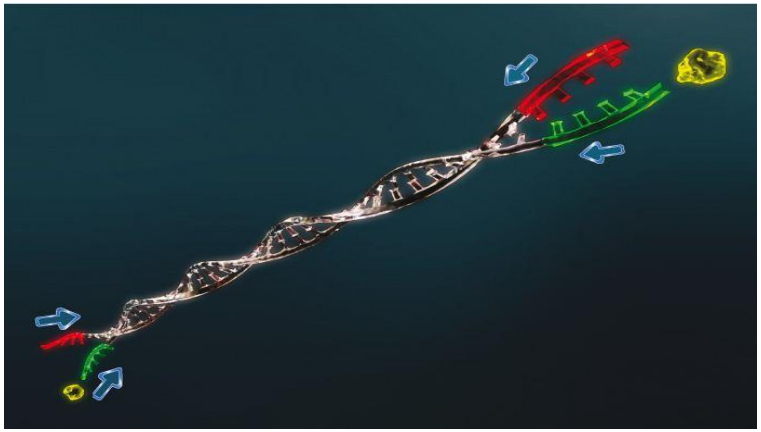


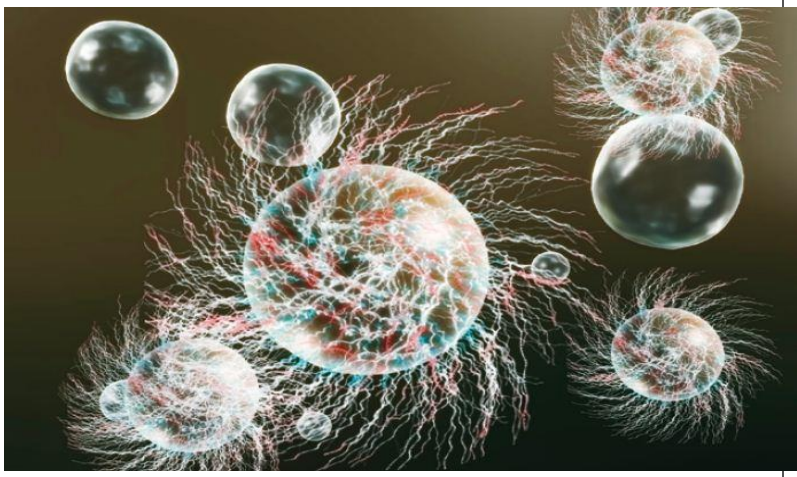
Current plant pathogen detection technologies

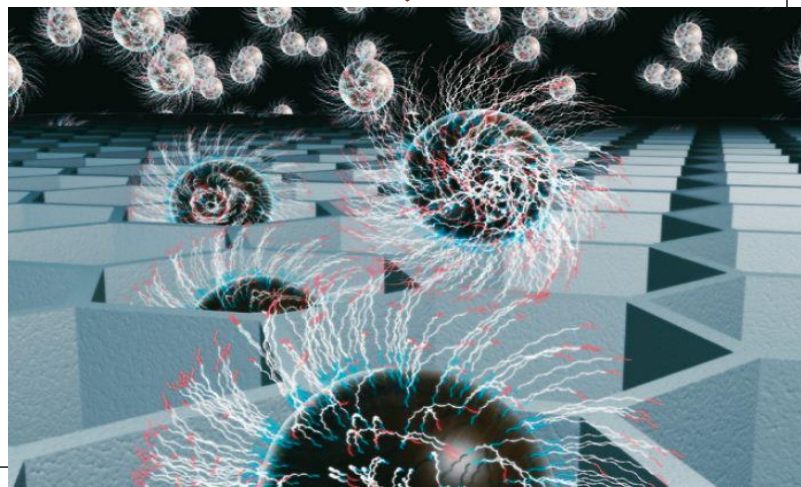
- **Protein based immunological assays**
  - Immunostrip test
  - ELISA
- **Nucleic acid based assays**
  - Polymerase chain reaction (PCR)
  - qPCR
  - Multiplex PCR
- **Can only detect a few pathogens at a time**
- **Prior characterization of the pathogen is required for all**

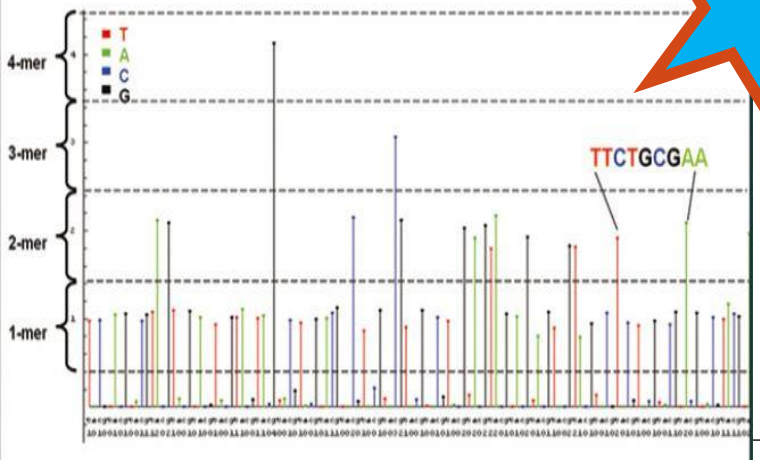
Next generation sequencing technology

- 454 Pyrosequencing









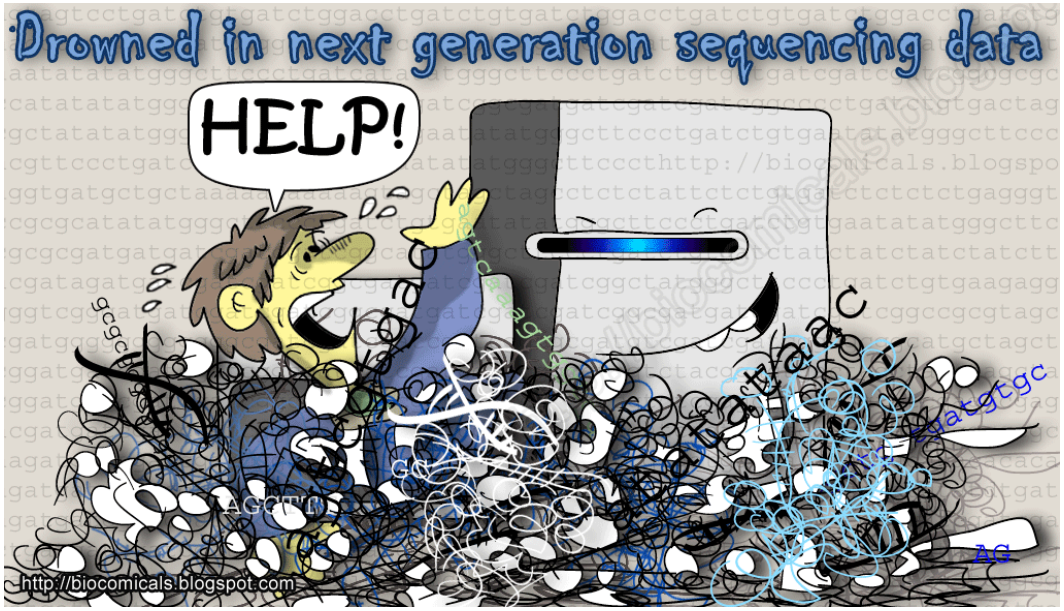
Next generation sequencing technology

- 454 Junior Pyrosequencing Power
  - Generates 35 million bases (Roche)
    - OSU Core facility over 61million bases
    - Read length ~400 bp
  - 100,000 shotgun reads
- Roche 454
- SOLiD
- Illumina
- Ion Torrent

http://www.sciencedirect.com/science/article/pii/S0002929709002833

Why use next generation sequencing?

- We get an enormous amount of data
- What to do with all that sequence data?
- We do not need all of the data
- Eliminate all irrelevant sequence data and reduce the bioinformatic load

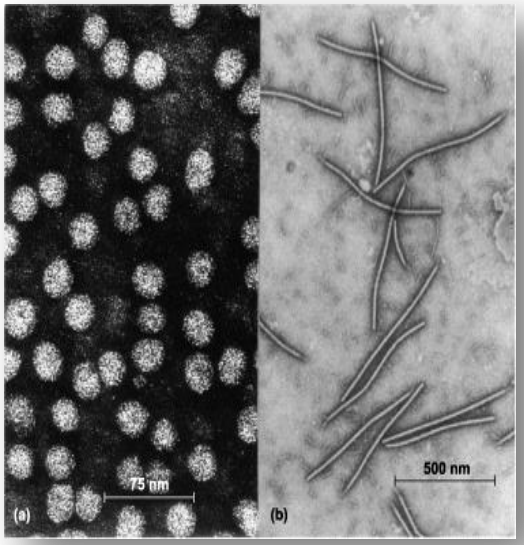


Justifications

- One-sixth of the GDP = $1 trillion annually
- 1 in every 8 individuals is employed by the **total agricultural industry**
- $50 billion in exports annually
- Single largest positive contributor to the U.S. economy
- In Oklahoma, ~$1 billion annually (2005)
- NGS allows the possibility of simultaneously **detecting** and **identifying** unknown waterborne plant pathogens

Purpose

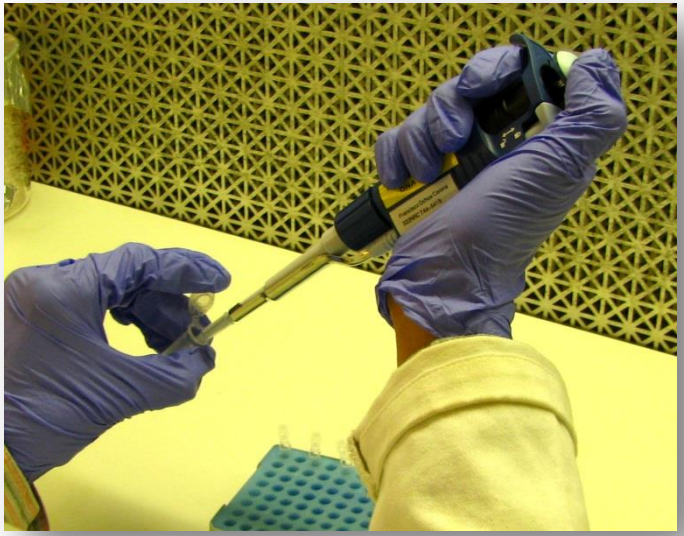
This study addresses the current need for detecting water borne viruses, in a single assay by combining metagenomics, NGS and bioinformatics. Our goal is to develop simple methods to facilitate the role of first detectors and managers monitoring water resources.



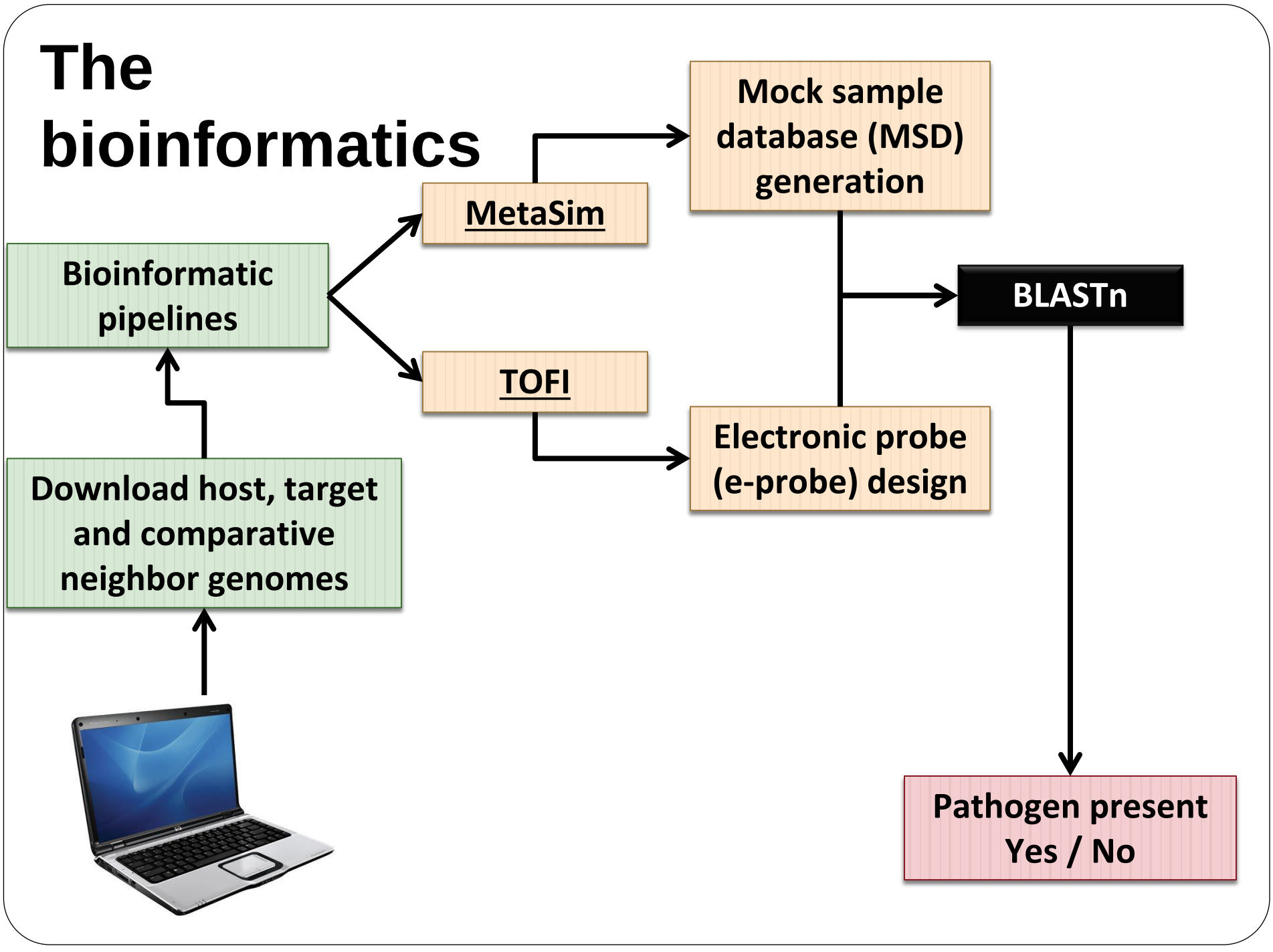
Virus

How can we detect waterborne viruses?

- Use tools/methods that combine **bioinformatics with plant pathology**



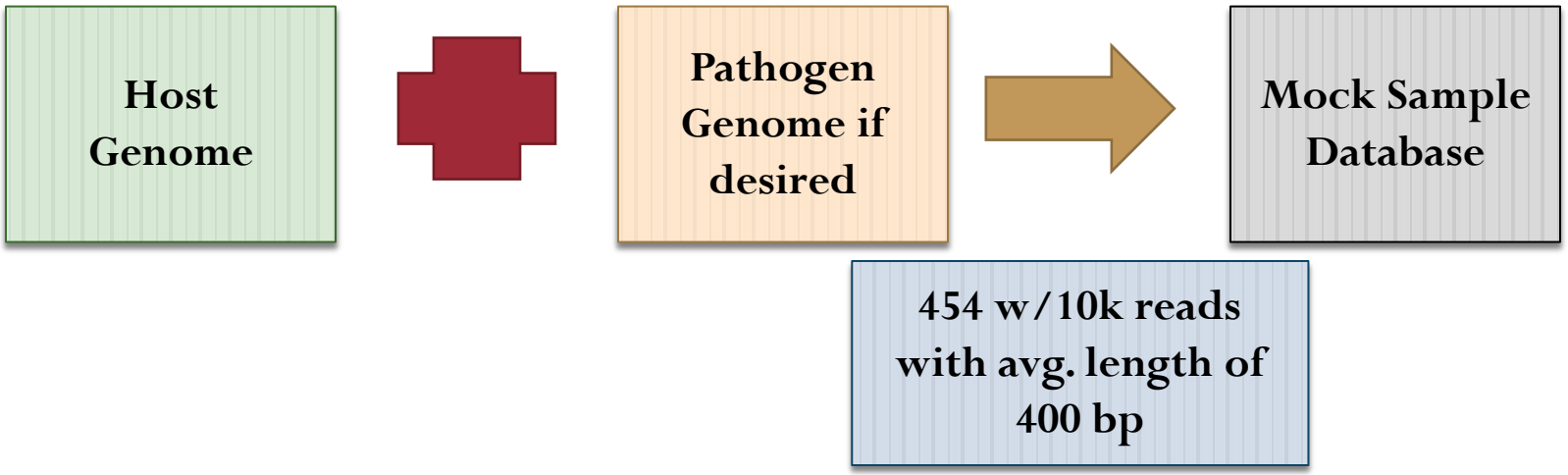
The bioinformatics



Mock Sample Database Generation

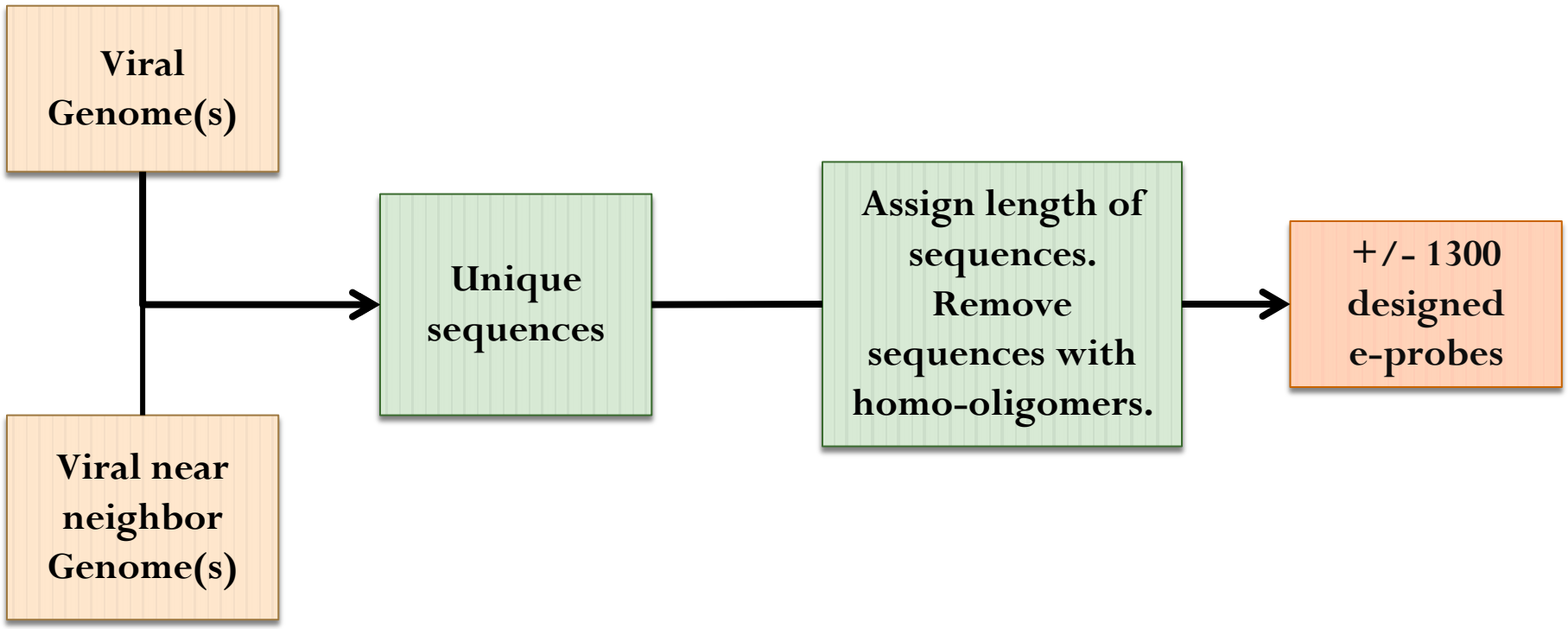
- Develop a Mock Sample Database for BLAST

MetaSim
Decreased control over the ratio, but offers a more realistic 454 sampling
Sequence errors are included
Run time ≤ 5 minutes



Electronic probe (e-probe) design

- The pipeline for **e-probe** generation by using parts of **T**ools for **O**ligo **F**ingerprint **I**dentification (**TOFI**).



E-probe design

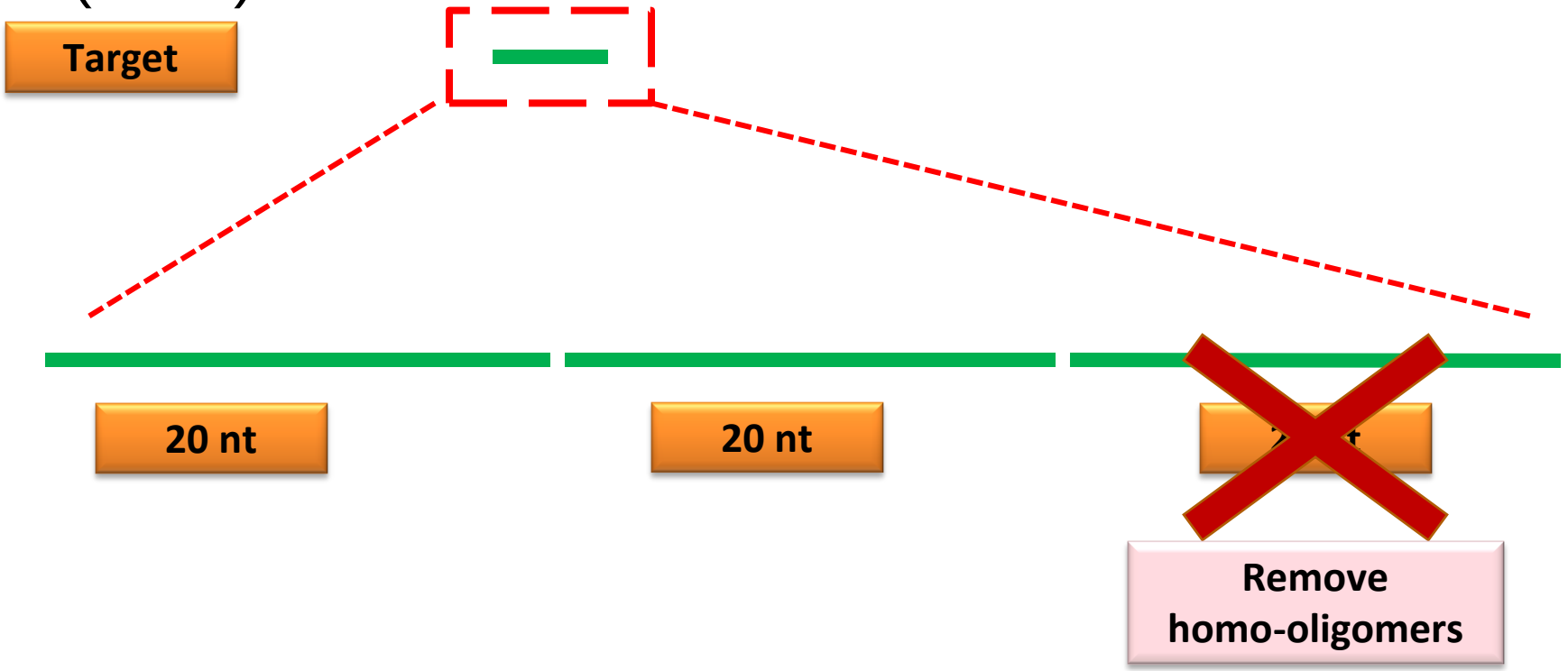
- Developed a pipeline for query generation by using parts of Tools for Oligo Fingerprint Identification (**TOFI**).

E-probe design

- Developed a pipeline for query generation by using parts of Tools for Oligo Fingerprint Identification (**TOFI**).

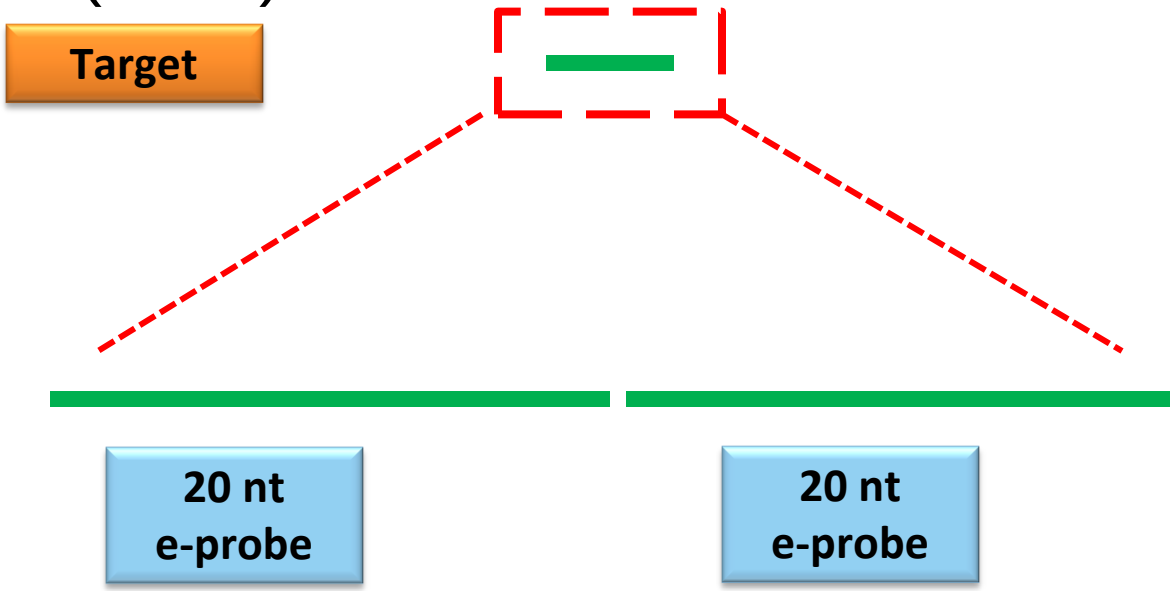
E-probe design

- Developed a pipeline for query generation by using parts of Tools for Oligo Fingerprint Identification (**TOFI**).



E-probe design

- Developed a pipeline for query generation by using parts of Tools for Oligo Fingerprint Identification (**TOFI**).



E-probes designed:

- ***Potexvirus***
- ***Tobamovirus***
- ***Tombusvirus***

In silico results from a 100-year agricultural field metagenomic database

NGS Simulator MetaSim	Total # of viral genomes	Avg. individual virus load	Total viral load	Total matches	Pathogenic virus present yes/no
NGS simulator (non-spiked)	unknown	unknown	unknown	1	Yes
NGS simulator (spiked)	27	0.14%	3.76%	343	Yes

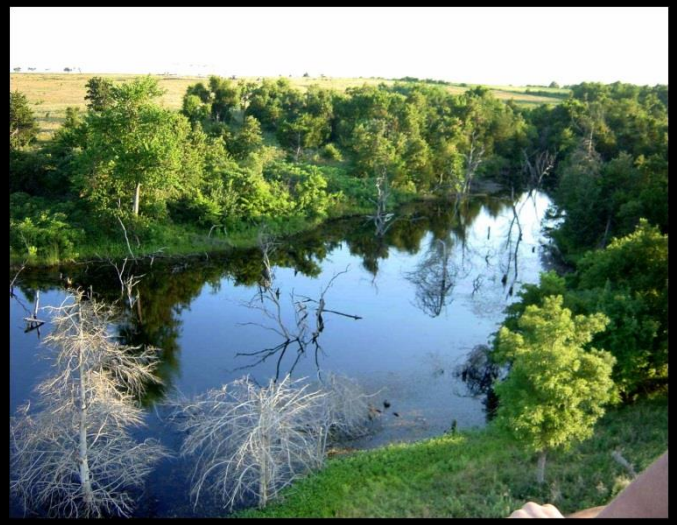
E-probes used:

- *Potexviruses*
- ***Tobamovirus***
- *Tombusviridae*

***Odontoglossum ringspot virus* (ORSV)**

- **ssRNA**
- **Plant pathogenic virus affecting orchids**

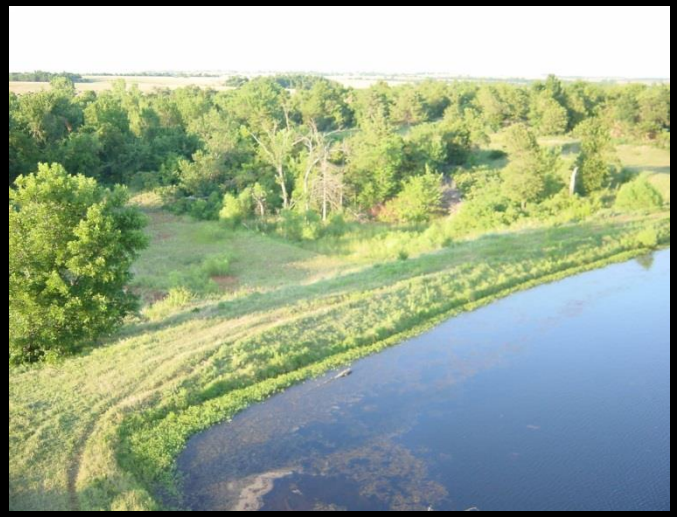
Experiment
Collection sites

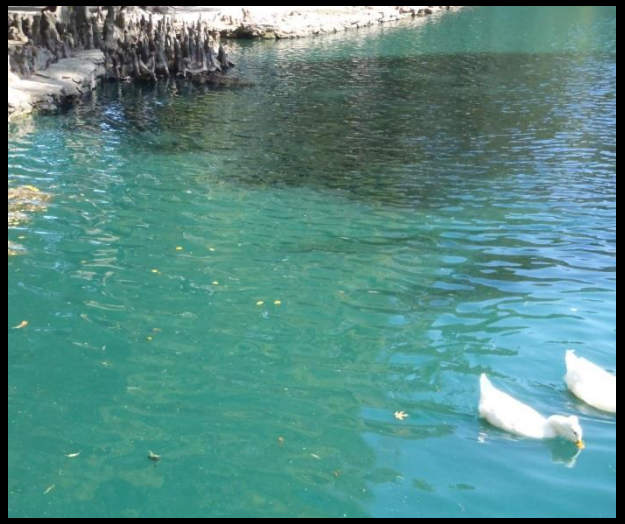


1

2

3

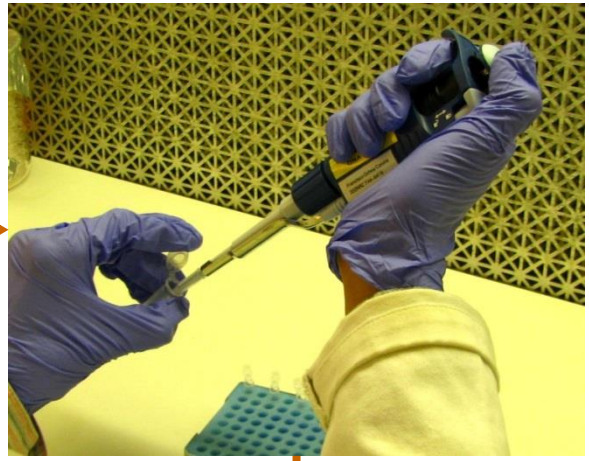




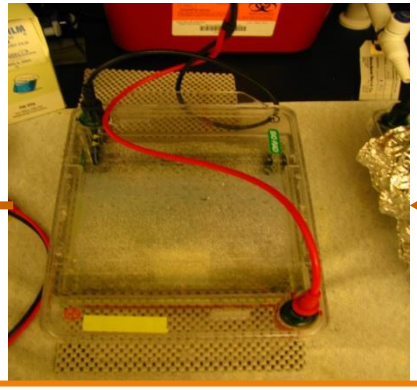
Experiment:
Basic procedures

Collection

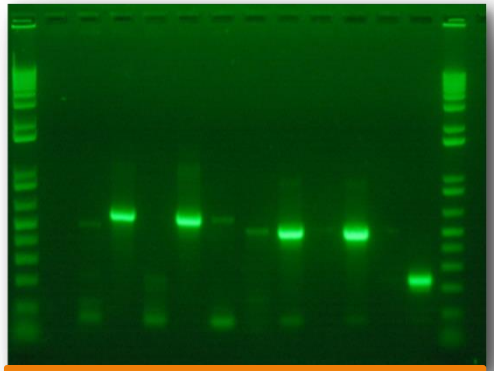
Extraction



PCR amplification

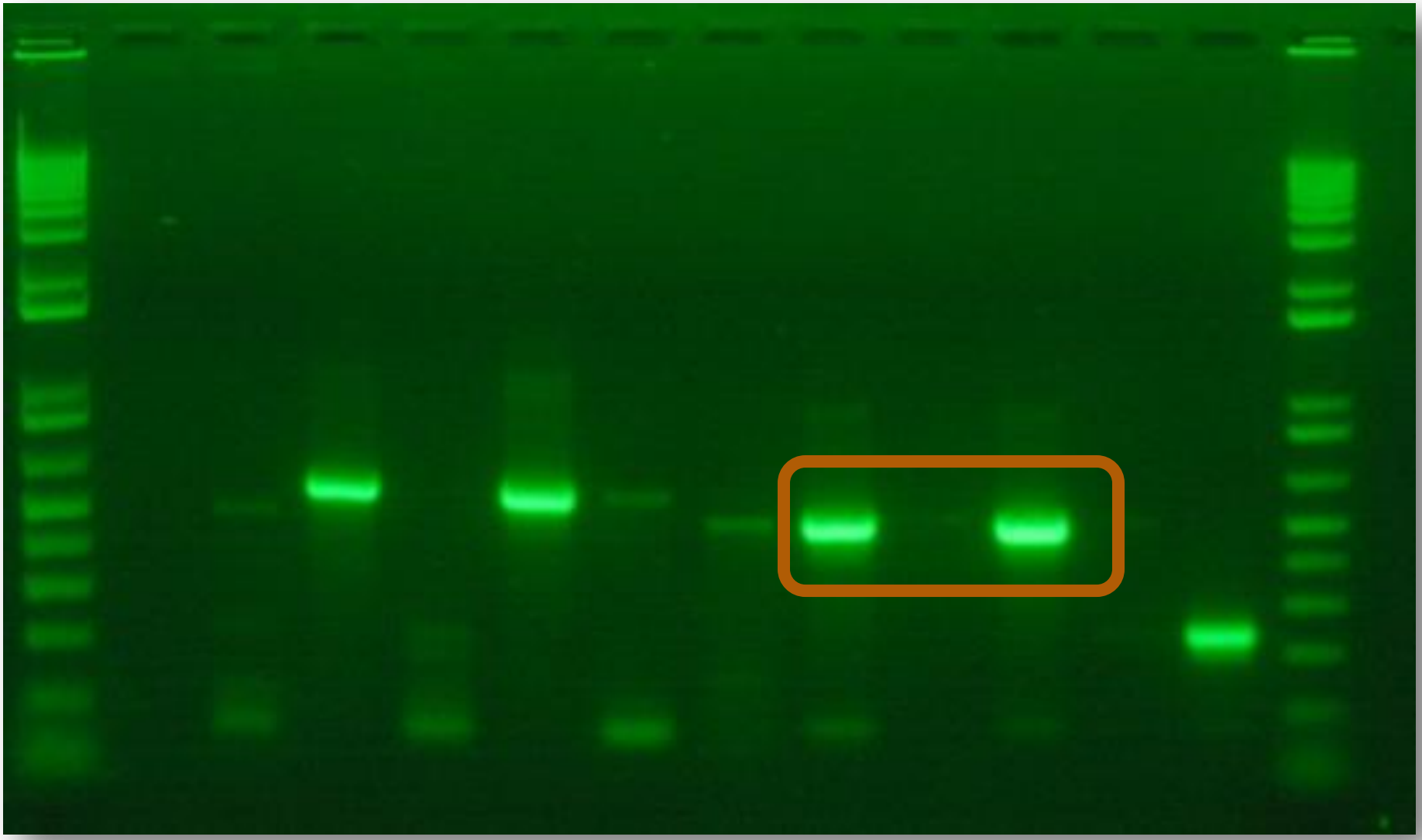


Gel electrophoresis

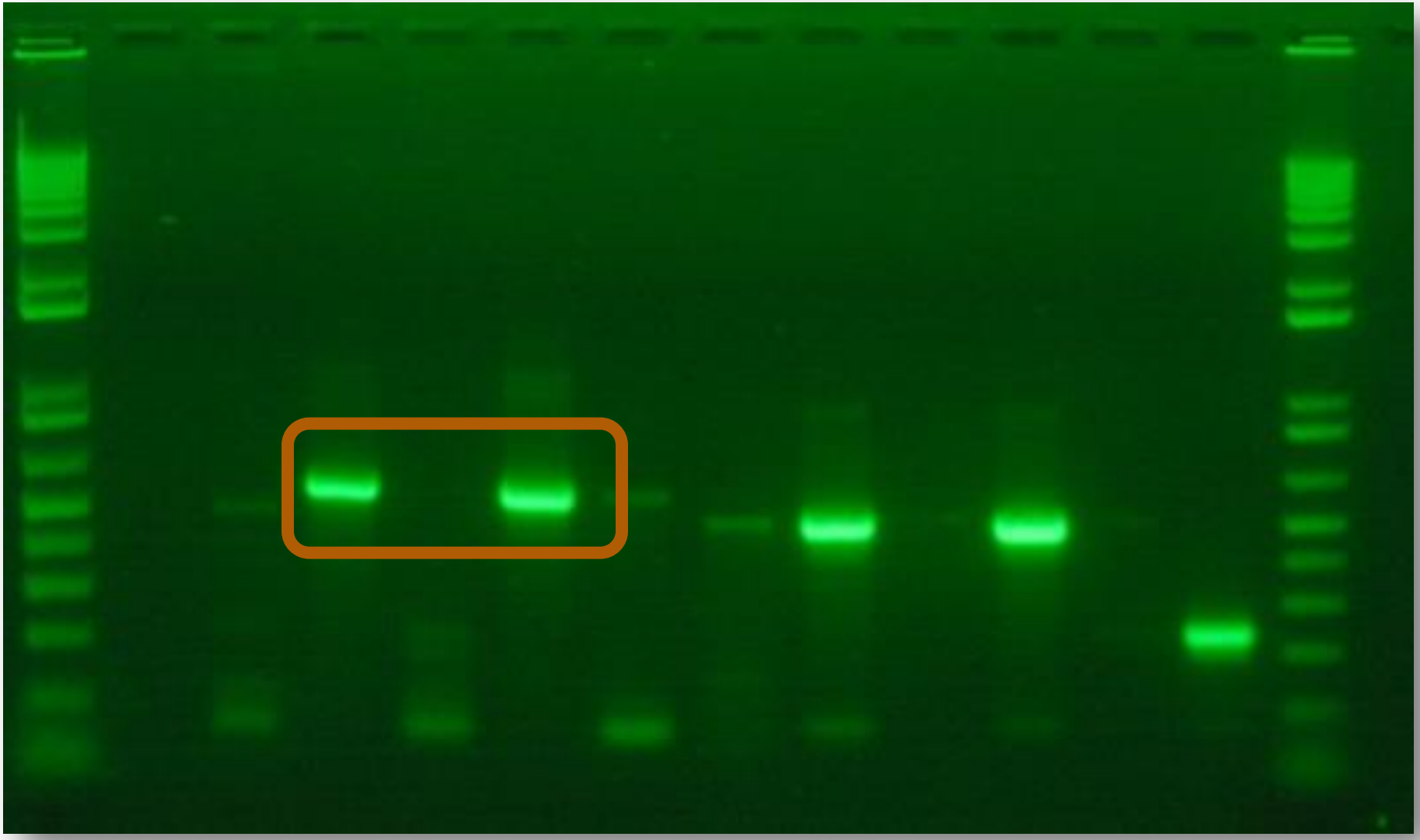


Products

Experiment
Gel electrophoresis: *Tombusvirus*?



Experiment
Gel electrophoresis: *Potexvirus*?



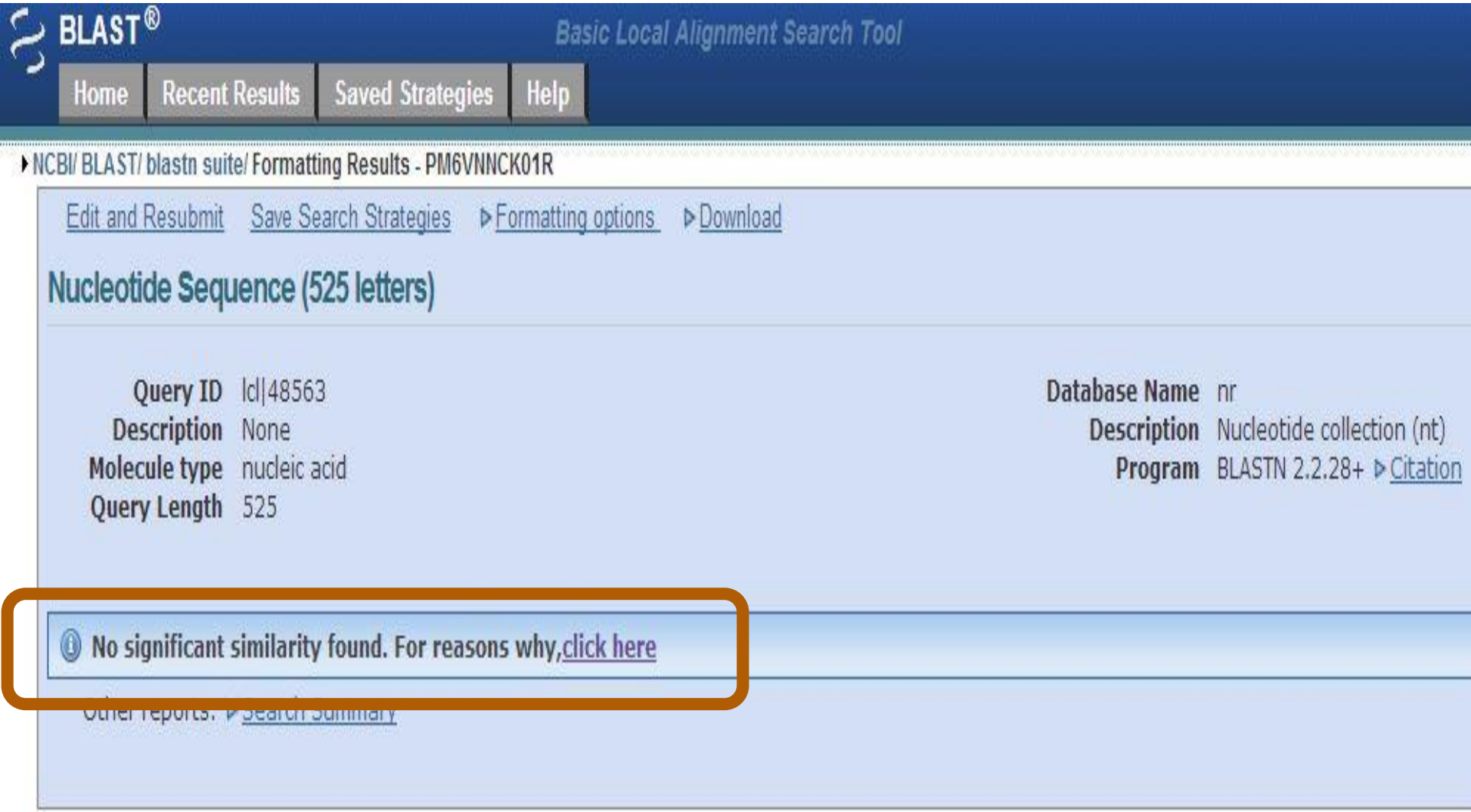
Sequencing the bands

- ***>Tombusvirus***
  - AAGGGTAAAAGGATGGTGAGGTATTCCGTTTTGACCATGTCGCCTGCGATCCAGGCATTGAGGGATA
    GGCATTGAGGGATCGCTTGTCCAACAAGGATTGGCGGATGCGGAACCTTTACCTCATCCTCGACGA
    GGAGGGTAGGACCGTTCCGTTTGTTATGCGGGGCGAGCAGGAGCAGTTCCTGCGGGAGAGGCACA
    ACCGGAATTTCATCCCCAAAGCTCGGAAGCTTGGGATGTCTACTGCCATCGTGCTAGCCAACCTCGA
    TGACTGCATTTTCAACGGCAATCTCGCGGCAGGGATCATCGACCTCACCAAGGACGATGCCTTCGC
    GAAACTGTCGATGGCGAGGTTCGCATGGGCATCTGGGCCGCTCCATCCAGATCCCGCTATCGGCGC
    ACTTTGGAAGTCACTGCACAAGGCCAATCCGCTTGAGAAGGATGCCGGGGGCGAGATGGTCTGGC
    GCAACGGATCAAGGATCACCGCAGGTGTAGCCTTCACGGGCCGCACTCCACACCCCTACCAAA

- ***>Potexvirus***
  - GGGGCACCCGACTATCGAGATCGGATGACTTGTCGACCCACCGAGTCGACACGGCTGACTCGCCC
    GGCAGCCGGAAAGCCAACGGTGGCTTGCTGCCACCTGCCAGCTGATGCAGCTGGGTTTCTAGCCC
    ACGAGTCGTTTTATCGATCGAACCAAGCGCCGGCCGAACCAGTGCCTTCCAATCCGACTCGACCTT
    CTCCTTCTGAGCTCGCAGAGATGCGATTTCCGTCCGCTTGGGGTCCTTTTCCAGCGAAGCAATTCGC
    TTTCGGGAGTTCTGATAATCCTGGATGTCCTTCGACCACTGAGCGTAGATCGAATCGATTTGGTCGAT
    CGCCTTATTGGCCAGCTTGTCGGCTTCCTGCTTGTTCTCGAATCCATAGTAGGTCGAGGCCTGCGCG
    ACATAGTCGCCGATGACCTGAGCCATCTTGTCGCGGTCTAAGCGGATCGTCCCATCCTGGTCCGGG
    ATCAACGCCTGGAACTGCCCCGCCAGCGGCCCTTTGGCAGCCTGCAGAAAGCCAATCGAGGAAAA
    CTTGCCGTCCCGAACCTTCGACCCAA

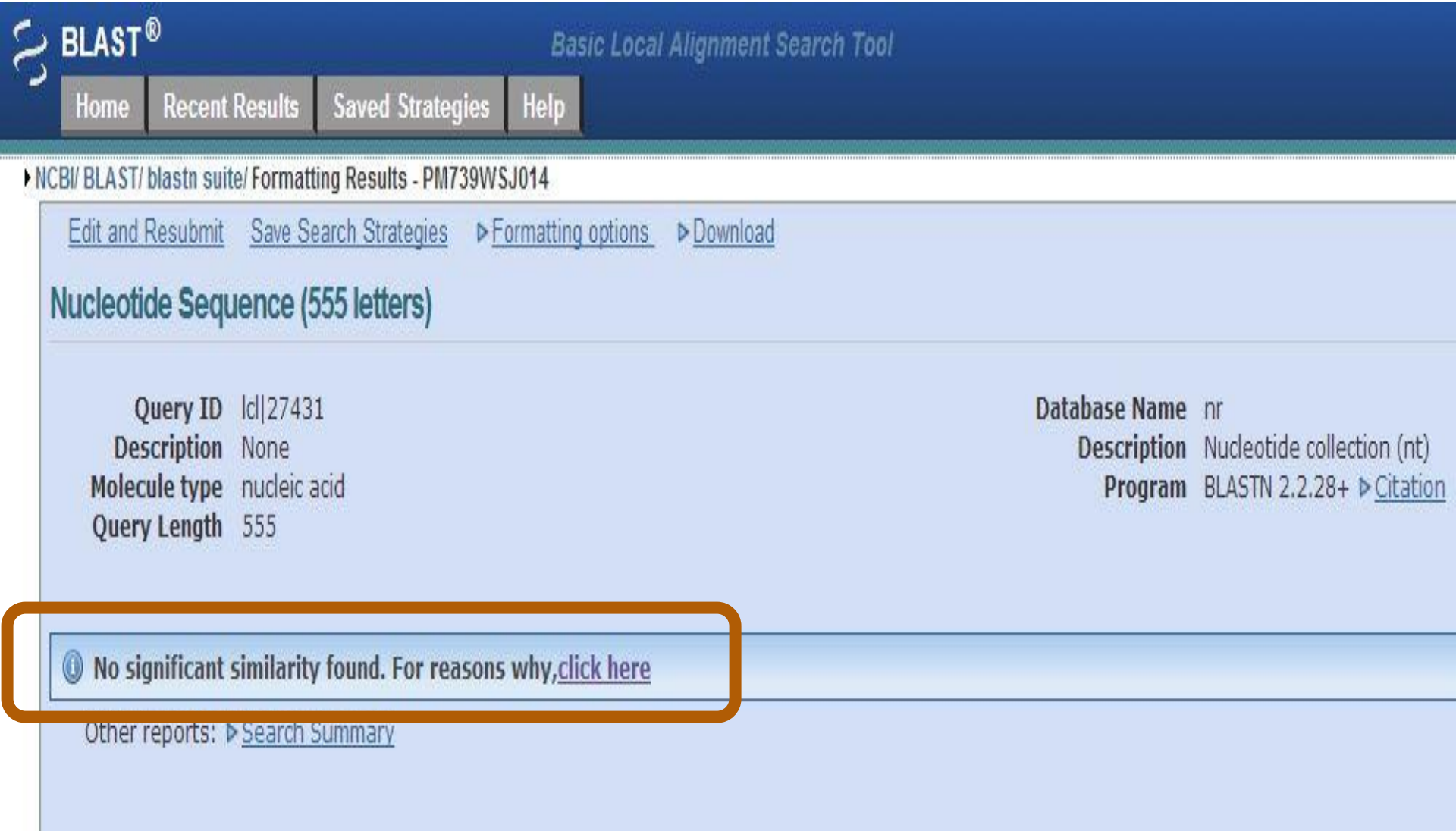
BLAST the sequences

- ***>Tombusvirus***

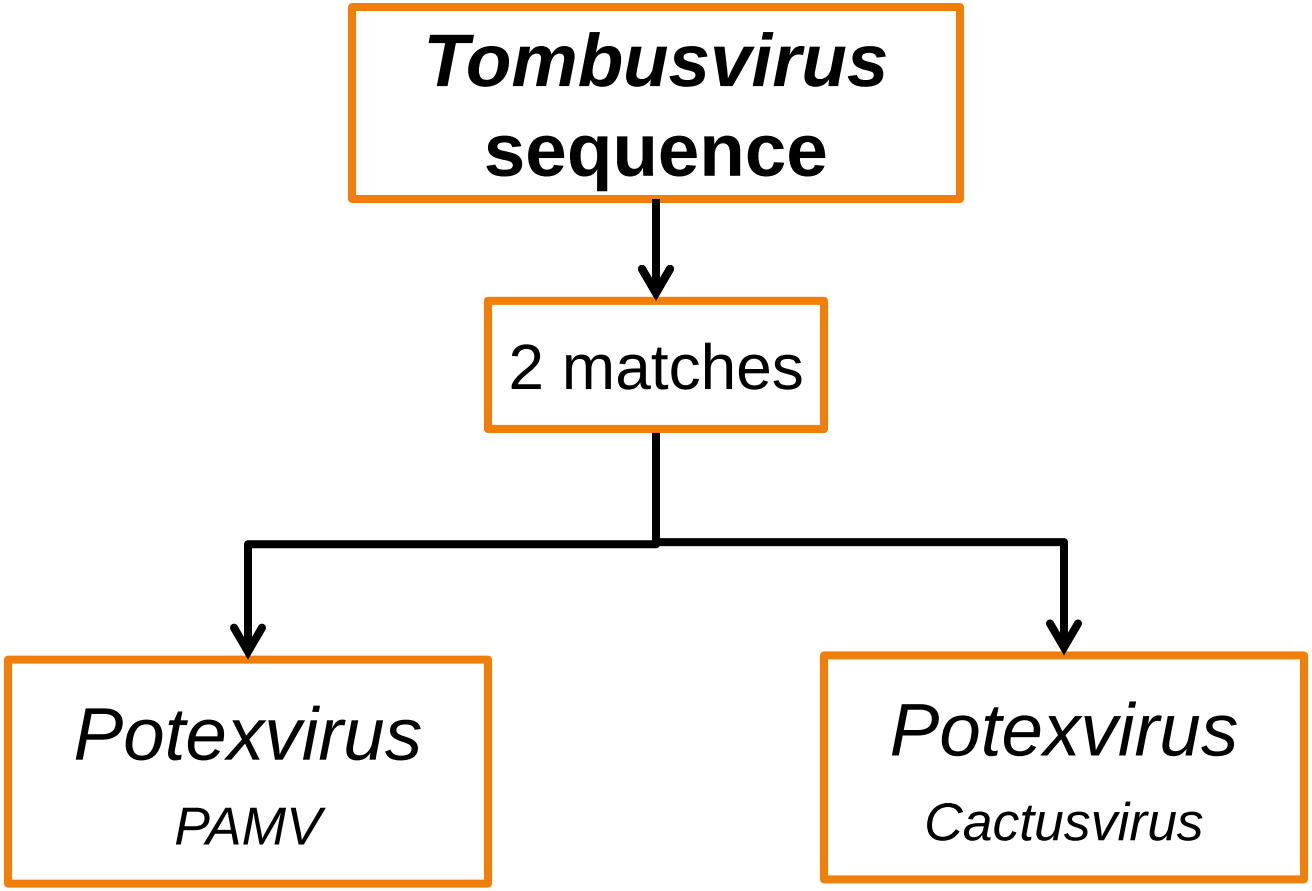


BLAST the sequences

- ***>Potexvirus***

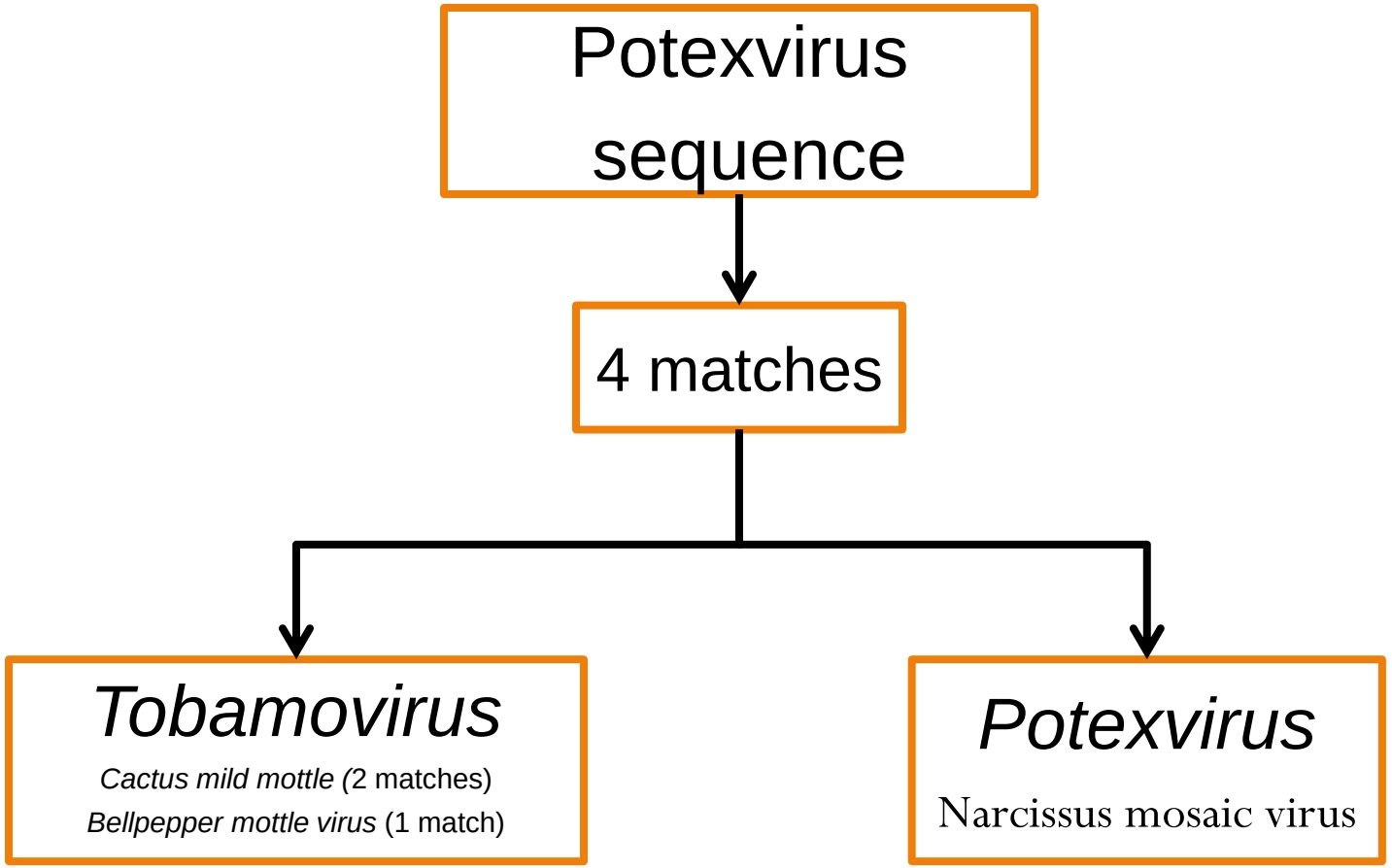


Using e-probes to BLAST the sequences



Both are plant pathogens!

Using e-probes to BLAST the sequences



All are plant pathogens!

Validation

Validation

Conclusion

- Using bioinformatic tools in combination with molecular tools enables to detect waterborne viruses.
- Positive detection is possible from 100 ml samples.
- Because of the volume, dilution and putative water dynamics of water as virus pathway,

  finding known and unknown plant pathogenic viruses is a very difficult specially for unknowns.
- The proposed database of waterborne viruses is dynamic and can be updated periodically.

Acknowledgements

AFRI Plant Biosecurity Program 2009-2013

Massively Parallel Sequencing (MPS) as a diagnostic and forensic analysis tool for plant pathogens

Division of Agricultural Sciences and Natural Resources

Vice President for Research and Technology Transfer

Please feel free to contact anytime!

Webpage:

http://entoplp.okstate.edu/profiles/ochoa-corona.html

Dr. Francisco Ochoa Corona: ochoaco@okstate.edu

Jon Daniels: jon.daniels@okstate.edu