FACT SHEET

National Clean Plant Network

Start clean, stay clean.

High Throughput Sequencing (HTS)

A variety of test methods - biological, serological, biochemical and molecular - are used currently in plant diagnostics and each has advantages and disadvantages. Classic diagnostic methods require prior knowledge of the pathogen in question and in some cases, require 2-4 years to complete. High throughput sequencing (HTS), also known as next generation sequencing (NGS), is a new, efficient method for obtaining nucleic acid sequences.

What is high throughput sequencing?

HTS provides the nucleic acid sequences in a sample which are compared with sequences known to be common to pathogens. For example, a **motif** or signature sequence in the family Geminiviridae is a unifying feature that, along with other identifying features, can place other viruses within the same family. Further analysis will determine if this virus is a known or a novel virus. A huge advantage of HTS is that it can identify a potential pathogen without having prior knowledge of that pathogen by using this principle of **conserved sequences** between related pathogens.

A huge advantage of HTS is that it can identify a potential pathogen without having prior knowledge of that pathogen.

Definitions

Motif - a nucleotide or amino-acid sequence pattern that is a defining attribute

Conserved sequence - similar or identical sequences that are indicative of how closely organisms are related

Read sequence ('read') - a nucleotide sequence that may have come from anywhere in the sample

Nucleotides - the four bases guanine (G), cytosine (C), adenine (A) and thymine (T) that make up a DNA strand

DNA library - a collection of labeled DNA fragments to be sequenced

Contig (from contiguous) - a set of overlapping DNA segments that represent a consensus sequence of DNA

Koch's postulates - the criteria to establish a causative relationship between a microbe and a disease

How is it done?

There are many variations of HTS sequencing platforms and different nucleic acid templates, such as small RNA, double-stranded RNA, or total nucleic acid can be used. Amplifying and sequencing purified DNA occurs when DNA molecules are attached to a chip and the sequencer reads millions of sequences per run. '**Reads**' are built one **nucleotide** base at a time during the sequencing operation. One HTS run takes about one day and can produce hundreds of millions of reads. This large-scale simultaneous synthesis of reads is what makes the process high throughput.

Simply, the steps are:

1. Collect the plant samples; extract and purify the total nucleic acid (TNA).

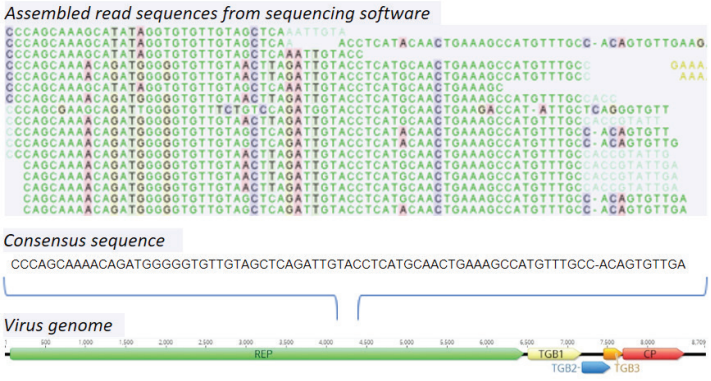
2. Evaluate the quality and quantity of the TNA.

3. Prepare **DNA libraries**; add specific adapters (labels) for each sample.

4. Quantify the libraries, combine, and load into a sequencer.

Preparing the library is an important step requiring 1-3 days.

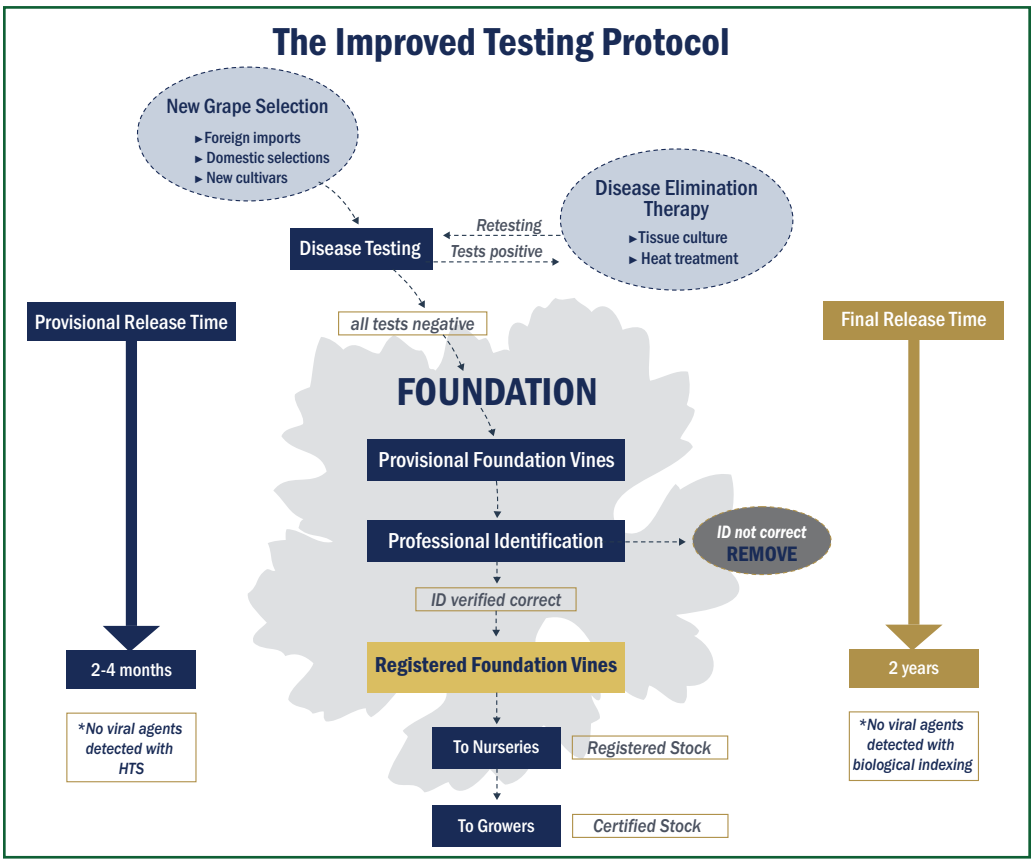
5. For each sample, use bioinformatics software to analyze reads and assemble into larger **contigs**.



Analysis process from read sequences to virus identification.

What are the benefits of HTS?

HTS is rapid, accurate, and efficient in the detection and identification of viral pathogens. It provides a comprehensive picture of the entire microbial profile of a sample. In studies comparing the efficacy of HTS and conventional assays used to detect grapevine and fruit tree viruses, researchers found that HTS was superior in its ability to detect viruses of economic significance (including low titer viruses), comprehensiveness, speed of analysis, and ability to discover novel, uncharacterized viruses. USDA APHIS has recently approved the provisional release of propagative plant material that has been HTS screened for pathogens. This has greatly reduced the wait time for clean plant material as selections under provisional release may be available in only 2 to 4 months.



For example, HTS testing has greatly reduced the time required to test grape selections. If a selection tests clean by HTS, it can be made available in as little as 2 to 4 months vs. 2 years or longer for bioassay results.

What are the limitations of HTS?

While HTS remains a powerful new technology with significant benefits, there are challenges associated with the technology. Due to its sensitivity, microbes of unknown pathogenicity are detected. Detection of a given microbe does not mean that it is responsible for disease. It is essential to establish biological significance to determine if the microbes sequenced are indeed biologically important. Biological significance is assessed by performing graft transmission, fulfilling **Koch's postulates**, analyzing spread and distribution, and assessing economic significance of symptoms.

No detection methodology is perfect. HTS and bioinformatics tools are only as good as the reference databases used. For example, novel virus sequences may be quite different from those deposited in databases. However, new methods that do not depend on sequence similarity are being explored and as knowledge expands so will the ability to identify novel species. In addition, efforts are underway to standardize HTS methodology across laboratories.

References

Al Rwahnih, M., Daubert, S., Golino, D., Islas, C. and Rowhani, A. 2015. Comparison of next-generation sequencing versus biological indexing for the optimal detection of viral pathogens in grapevine. Phytopathology, 105: 758-763.

Rott, M., Xiang, Y., Boyes, I., Belton, M., Saeed, H., Kesanakurti, P., Hayes, S., Lawrence, T., Birch, C., Bhagwat, B. and Rast, H. 2017. Application of Next Generation Sequencing for Diagnostic Testing of Tree Fruit Viruses and Viroids. Plant Disease, 101: 1489-1499.

About NCPN

Established in 2008 and supported by the US Department of Agriculture, the NCPN is a national network of clean plant centers, scientists, educators, regulators and industry representatives who are concerned with the health of vegetatively propagated specialty crops.

nationalcleanplantnetwork.org

Start clean, stay clean
NCPN
National Clean Plant Network