Management of Plant Pathogen Collections

2nd edition

ISBN: 978-0-646-88278-9

This document is available online at agriculture.gov.au/agriculture-land/plant/health or at appsnet.org/publications.

All enquiries should be directed to:

Australian Government Department of Agriculture, Fisheries and Forestry
GPO Box 858
Canberra ACT 2601
Australia
DAFF-TechnicalCapacityBuilding@agriculture.gov.au

Australian Government
Department of Agriculture, Fisheries and Forestry

APPS

The cover is from Goedsies van Wildebees. *Hemileia Hellride*, 2019.
Acrylic on canvas. Private collection, Queensland.

'Lists of records that cannot be verified are mere waste paper.'
R. W. G. Dennis, *British Ascomycetes* (1968)

'The affinities of all the beings of the same class have sometimes been represented by a great tree. I believe this simile largely speaks the truth.'
Charles Darwin, *On the Origin of Species* (1859)

'On the other hand, splitting fungi into more detailed taxonomic classifications provides useful evolutionary information, but also could provide useful clinical information by linking virulence factors or pathogenicity to a specific taxonomic classification.'
B. L. Wickes and N. P. Wiederhold in *Nature Communications 9* (2018)

Acknowledgments

This revision was funded by the Australian Government, Department of Agriculture, Fisheries and Forestry (DAFF), and the ASEAN-Australia-New Zealand Free Trade Agreement Economic Cooperation Work Programme. Special thanks are extended to Dr Ian Naumann and Ms Wendy Lee for making this possible.

Authors of this revision

Prof. Roger G. Shivas (University of Southern Queensland), Prof. Levente Kiss (University of Southern Queensland), Assoc. Prof. Andrew D. W. Geering (The University of Queensland) and Dr Alistair R. McTaggart (The University of Queensland)

How to cite

Shivas, R. G., Kiss, L., Geering, A. D. W., & McTaggart, A. R. (2023). *Management of Plant Pathogen Collections*, 2nd ed. Australian Government Department of Agriculture, Fisheries and Forestry, Canberra.

Preface

This handbook is a revised edition of *Management of Plant Pathogen Collections* first published in 2005. The original handbook was intended to assist plant pathologists collect, preserve and identify specimens of plant pathogens, as well as to provide a ready reference for the curators of reference collections.

Much has changed in the last 16 years. Foremost, a taxonomic revolution caused by the delimitation of species by molecular phylogenetic methods has accelerated the discovery of all types of organisms. Plant pathologists working in the biological disciplines must deal with many new published names for micro-organisms. Molecular methods are now routinely used in plant pathology laboratories worldwide to identify and classify microbial plant pathogens. Traditional diagnostic and taxonomic approaches based on phenotypic characteristics, e.g., morphology, host range and physiology, have been replaced by analyses of their genetic sequences.

One constant is that we need names to communicate, no matter whether taxonomy is based on morphology, phenotypes, single genes, multiple genes or genomes. The 2005 edition of this handbook did not cover molecular methods, which were only starting to gain traction in plant pathology laboratories at that time. In this revision, molecular methods take centre stage and are recognised as fundamental skills for curators and taxonomists in plant pathogen reference collections.

This revision was commissioned by the Australian Government, Department of Agriculture, Fisheries and Forestry. Its purpose is to assist curators and scientists manage reference collections of plant pathogens. Robust and credible pest-risk analyses are still fundamental in the negotiation of access to new markets for agricultural commodities.

Contents

1 Introduction

Plant pathogens have impacted food crops since our ancestors turned their hands to agriculture in the Fertile Crescent region about 10,000 years ago. Monocultures and clonal propagation of plants expanded ecological niches for plant pathogens. The introduction of plant species to new geographic regions allowed for chance encounters between plants and pathogens that were previously not physically possible. Despite advances in agriculture and plant biosecurity over time, new plant diseases continue to emerge and spread.

The global population will likely increase to about 10 billion by 2050. Food production must increase to meet demand, and this can at least be partially achieved by minimising yield losses from pests and pathogens. Diseases caused by plant pathogens can be controlled by deployment of resistance genes, by the provision of clean planting material (seed or vegetative propagules) and by use of agrichemicals. Biosecurity also helps prevent the spread of pathogens. The foundation of plant disease management is accurate diagnosis.

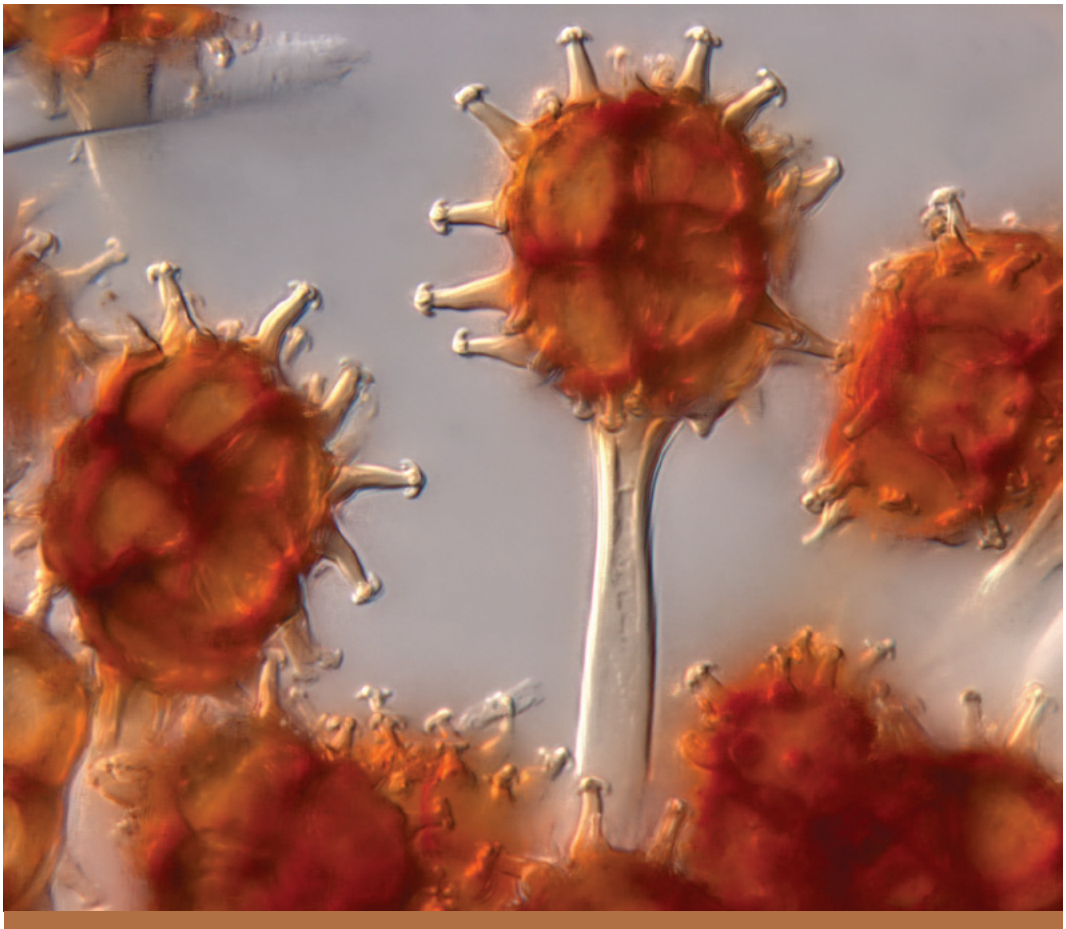
Specimens held in reference collections underpin stable taxonomies and are the foundation of universal names for organisms that cause plant disease. New roles for plant pathogen reference collections have manifested in recent years. Reference collections established to support agricultural research are now recognised for their biodiversity value. The remit of these reference collections has expanded to support research and innovation in biotechnology, bioprospecting, conservation, ecology, medical sciences and pharmacology.

1.1 INTERNATIONAL OBLIGATIONS

The International Plant Protection Convention (IPPC) and the Agreement on the Application of Sanitary and Phytosanitary Measures oblige exporting countries to provide a list of pests likely to be associated with the exported commodity. To meet these obligations, countries need to maintain reliable pest records. The International Standard for Phytosanitary Measures no. 8 (ISPM 8) states that all countries may use pest status information for:

- pest-risk analysis
- planning national, regional, or international pest management programs
- establishing national pest lists
- establishing and maintaining pest-free areas.

Published reports of the presence or absence of plant pathogens in certain regions or countries must be supported by reference specimens and collection records. Reference collections may contain multiple specimens of the same taxon from different populations on hosts from the same or different geographic and production areas. Reference specimens can be re-examined to determine or verify their identity. Unverifiable reports are those not supported by specimens. Unverifiable reports are a potential impediment to agricultural trade, as they are often difficult, time-consuming and expensive to disprove.



Teliospores of *Sphaerophragmium acaciae* on *Albizia* sp. in northern Queensland

In the last decade many well-established plant pathogens have been shown by molecular data to represent multiple closely related species that are morphologically indistinguishable. Specimens in reference collections provide a powerful tool to facilitate market access and justify measures that exclude potentially harmful plant pathogens. Living specimens, particularly cultures, may also have biodiversity or commercial value.

2 Specimens and collections

The primary purpose of a reference collection is to catalogue biodiversity. Reference collections provide a permanent repository of preserved specimens together with information about their identity and collection. Collections of living cultures safeguard against the loss of genetic diversity through habitat destruction and climate change. The specimens in a reference collection are used for taxonomic research, agricultural and environmental management, bioprospecting and biotechnology.

Reference specimens of plant pathogens:

- document microbial biodiversity
- preserve specimens to safeguard against the loss of genetic diversity
- support taxonomic research by holding type specimens
- help determine research priorities for agriculture and the environment
- facilitate agricultural trade by meeting international obligations
- provide specimens for bioprospecting and commercialisation
- give evidence for pathogen distribution and prevalence, especially the presence or absence of diseases
- facilitate the identification of exotic and emerging plant diseases
- allow the host range of plant pathogens to be determined
- evaluate the efficacy of control measures for plant diseases
- help identify pathogens with potential for biological control of weeds and insects
- verify and validate diagnostic methods.

2.1 SPECIMENS AND NAMES

Specimens with names are the foundations of taxonomy. Life on Earth is classified using a system of nomenclature proposed by Carolus Linnaeus in 1753. Linnaean nomenclature assigns genus and species names to all living organisms, including bacteria, fungi, nematodes and now, viruses. The naming of viruses using a binomial system is a recent concept, ratified in March 2021 and now required for all new taxonomic proposals. Linnaeus sought to establish a 'natural' system of classification, which today is considered to be one that reflects the evolutionary relationships between taxa in a universal tree of life. Linnaean names have become the backbone of biology, as they facilitate communication about all extant and extinct organisms.

Frequently asked question: **What is a type specimen?**

Answer: **A type specimen (or type) is the single specimen on which the name of a species was first proposed. Species of bacteria, fungi and nematodes have types. The description of a type may be based on phenotypic and/or genotypic characteristics, according to the nomenclatural codes. Types are usually physical specimens kept in reference collections. Type specimens of plants are placed in red packets or packets bordered or marked in red, to make them easily recognisable in a reference collection.**

The concept of a name-bearing type specimen or type species has not been adopted for viruses. Virologists recognise exemplar isolates of a species,

which serve many of the same functions of a type specimen by anchoring the taxonomy of the species. It is mandatory that entire genome sequences accompany exemplar isolates, although there is not a formal requirement to store a physical specimen in a recognised collection. While not compulsory, it is best to store a physical specimen of the exemplar isolate in case the genome sequence needs validation or there are questions about the biology of the isolate, e.g., whether satellite viruses are present.

Holotype – The one specimen designated by an author that typifies a new name.

Isotype – Any specimen that is a duplicate of the holotype.

Lectotype – A specimen designated from the original material as the type if no holotype was indicated at the time of publication, or if it is missing, or found to belong to more than one taxon.

Isolectotype – Any specimen that is a duplicate of a lectotype.

Syntype – Any specimen cited in the original place of publication when no holotype was designated, or any one of two or more specimens simultaneously designated as types.

Paratype – Any specimen cited in the original place of publication that is neither the holotype nor an isotype, nor one of the syntypes.

Neotype – A specimen selected to serve as a nomenclatural type if all the material on which the name of the taxon was based has been lost or destroyed.

Epitype – A specimen selected to serve as an interpretive type when the holotype, lectotype or previously designated neotype, or all original material associated with a validly published name, is demonstrably ambiguous and cannot be critically identified for the precise application of the name.

Topotype – Any specimen collected at the same location as the holotype.

Exemplar virus isolate – A virus isolate and accompanying genome sequence that is recognised by the International Committee on Taxonomy of Viruses as the pre-eminent example of the virus species. The genome sequence of the exemplar isolate is typically adopted as the reference sequence for the virus species by the National Center for Biotechnology Information (NCBI).

Frequently asked question: Is the description of a species based solely on the type specimen?

Answer: The description of a species must be based on the holotype and may include other specimens that have been examined to determine the range of variation in phenotypic characteristics. In the latter case, the other specimens examined should be verified by molecular methods.

Frequently asked question: What is a type specimen for a bacterium?

Answer: A type specimen for a bacterium is a living culture. It is also called a type strain.

Frequently asked question: What is a type specimen for a fungus?

Answer: A type specimen for a fungus must be either a permanently preserved, dried specimen or a culture preserved in a metabolically inactive state by freeze-drying or deep-freezing.

Frequently asked question: What is a type specimen for a nematode?

Answer: A type specimen for a nematode is a permanently preserved slide mount.

Frequently asked question: What is a type specimen for a virus?

Answer: Virus names are not based on type specimens but, instead, on exemplar isolates.

Frequently asked question: Can DNA serve as a type specimen?

Answer: No. According to the International Commission on Zoological Nomenclature (ICZN), DNA cannot be a type specimen as it is present in small amounts in cells and must be copied/amplified before sequencing. The copied DNA cannot be a type specimen, as it is not part of the original specimen. New species can be based on DNA sequences, which should form part of the description of the type specimen. The type specimen must be preserved and deposited in a reference collection.

Infrequently asked question: How can I locate the type specimens of
(i) *Ceratocystis manginecans*,
(ii) *Dickeya oryza*, and
(iii) *Laimaphelenchus australis*?

Answer:

(i) Use Species Fungorum and Index Herbariorum to find PREM 59612 at the Plant Protection Research Institute, National Collection of Fungi (PREM), Pretoria, South Africa.

(ii) Use LPSN – List of Prokaryotic names with Standing in Nomenclature to find type strains ACCC 61554 at the Agricultural Culture Collection of China (ACCC), Chinese Academy of Agricultural Sciences, Beijing, China, and JCM 33020 at the Japan Collection of Microorganisms (JCM), Saitama, Japan.

(iii) Search literature to find the original publication, which states that the type has been deposited in the Australian National Insect Collection (ANIC), Canberra, as slide no. 113.

Frequently asked question: Where are the names of new species of fungi registered?

Answer: There are three nomenclatural repositories recognised by the Nomenclature Committee for Fungi, i.e., Fungal Names, Index Fungorum and MycoBank.

Frequently asked question: What are evolutionary relationships?

Answer: Evolutionary relationships describe the common ancestors an organism or group of organisms have shared in their evolution. Organisms are considered closely related (or sister taxa) if they share a common ancestor. The evolutionary relationships of an organism are called its phylogeny.

Frequently asked question: What are the rules for naming organisms?

Answer: There are international committees and codes that provide rules for the naming of organisms. The respective nomenclatural codes for bacteria, fungi, nematodes and viruses are provided below.

The nomenclatural codes

International Code of Nomenclature of Prokaryotes (ICNP)

International Code of Nomenclature for algae, fungi and plants (ICN)

International Commission on Zoological Nomenclature (ICZN)

International Committee on Taxonomy of Viruses (ICTV)

Frequently asked question: What is the relationship between plant pathology and taxonomy?

Answer: Molecular data have shown that many names applied to plant pathogens represent complexes of morphologically indistinguishable species. Further, molecular data have also shown that many species thought to be closely related by morphology belong to different genera. This has accelerated species discovery and provided common and fertile ground for collaboration between plant pathologists and taxonomists. Occasionally some well-known plant pathogens have required name changes, which has caused tension between taxonomists and plant pathologists. Ultimately every species on our planet will find its place in the universal tree of life. The relationship between plant pathology and taxonomy over the last half century has been reflected in some paper titles (listed chronologically below; full citations listed alphabetically by author surname in Appendix 1):

'Mutual Responsibilities of Taxonomic Mycology and Plant Pathology.' Walker (1975)

'The Importance of Taxonomy in Plant Pathology.' van der Westhuizen (1991)

'Much more than Phylogenies: A Utilitarian View of the Taxonomy of Plant Pathogenic Fungi.' Glawe (2003)

'Plant Pathology is Lost without Taxonomy.' Crous (2005)

'Demystifying the Nomenclature of Bacterial Plant Pathogens.' Bull *et al.* (2008)

'Systematics of Plant Pathogenic Fungi: Why it Matters.' Rossman & Palm-Hernández (2008)

'A Case for Re-inventory of Australia's Plant Pathogens.' Hyde *et al.* (2010)

'Nematode Taxonomy: From Morphology to Metabarcoding.' Ahmed *et al.* (2015)

'Historical Virus Isolate Collections: An Invaluable Resource Connecting Plant Virology's Pre-sequencing and Post-sequencing Eras.' Jones *et al.* (2020)

Frequently asked question: What are some examples of pathogens that now have new names?

Answer: The fungus that causes wheat leaf blotch is now correctly called *Zymoseptoria tritici* (syn. *Septoria tritici*). The fungus that causes boil smut of maize is now correctly called *Mycosarcoma maydis* (syn. *Ustilago maydis*). The fungus that causes rice blast is now correctly called *Pyricularia oryzae* (syn. *Magnaporthe oryzae*). The bacterium that causes soft rot on several ornamental and horticultural host plants worldwide is now correctly called *Dickeya dadantii* (syn. *Erwinia chrysanthemi*).

Frequently asked question: What is the universal tree of life?

Answer: The universal tree of life is a hypothesis that shows evolutionary relationships among all organisms, both living and extinct. It was first described by Charles Darwin in his book *'On the Origin of Species'* (1859) as follows: 'The affinities of all the beings of the same class have sometimes been represented by a great tree. I believe this simile largely speaks the truth.'

Frequently asked question: Who is the authority for virus nomenclature?

Answer: The International Committee on Taxonomy of Viruses (ICTV) oversees the taxonomic classification and naming of viruses. The ICTV recognises several study groups comprising experts in various fields, who provide advice and present proposals each year for the recognition of new taxa. Taxonomic proposals can come from anyone, but the proposal must be vetted by the relevant study group and any feedback addressed. Finally, taxonomic proposals must be submitted on the official template, which can be download from the ICTV website.

Frequently asked question: Where can I find the most recent list of virus names?

Answer: Each year, the ICTV releases a new 'Master Species List: A Spreadsheet of Current Taxonomy'. This list can be found on the ICTV website. The ICTV also publishes a summary of all taxonomy, called 'Virus Taxonomy: The ICTV Report on Virus Classification and Taxon Nomenclature', which can also be found on their website. ICTV Virus Taxonomy Profiles are published by the Microbiology Society in the *Journal of General Virology*, e.g., ICTV Virus Taxonomy Profile: *Caulimoviridae*.

2.2 COLLECTIONS

Reference specimens of plant pathogens are found in national or regional culture collections, fungaria, herbaria or museums. Occasionally, significant reference collections of plant pathogens are found in private collections. The physical specimens of plant pathogens (bacteria, fungi, nematodes, and viruses) are mostly preserved as (i) dead, dried, pressed, or slide mounted material, together with (or often embedded in) parts of their plant hosts; (ii) living or cryopreserved cultures; and (iii) DNA extracts stored in ultralow freezers (-80°C). DNA forms part of most biological specimens (except for some RNA viruses), as it is present in all cells. The identity and collection metadata of a specimen should be databased and published (online) for publicly available specimens.

A reference collection should provide a safe and permanent place to store specimens. It should be insect-proof, fireproof and waterproof. A moderately large (up to 100,000 specimens) reference collection can be housed in a single room of about 9 m^2 if a compactus is used for storage. Metal storage shelving and cabinets are more insect-resistant than wood.

The reference collection should be kept at 20–23°C with humidity of 40–60%. This environment controls many insect pests, especially when combined with periodic freezing of specimens. Air-conditioners control temperature and dehumidifiers reduce humidity. Windows and doors should be kept shut to prevent insects from entering. Tinting on windows will reflect heat.

Virus specimens are normally stored in -20°C freezers as lyophilised leaf tissue in sealed and evacuated bottles/vials or in screw-cap plastic tubes that contain a layer of silica gel or anhydrous calcium chloride ($CaCl_2$) to keep the specimens dry. The sample should be separated from the desiccant by cotton wool or tissue paper. Fresh leaf tissue can also be stored frozen at -20°C or -80°C, although the specimens occupy more space in storage and are at risk of spoilage should the freezer break down.

Frequently asked question: **What are fungaria and herbaria?**

Answer: **Fungaria (sing. fungarium) and herbaria (sing. herbarium) are collections of fungal and plant specimens, respectively. Many herbaria also have collections of diseased plants.**

Frequently asked question: **Where is the largest fungarium in the world?**

Answer: **The Fungarium at the Royal Botanic Gardens, Kew, London, has over 1 million specimens of dried fungi and is the largest collection in the world with specimens dating back to the 1700s.**

Frequently asked question: **How do specimen collection records of plant pathogens support international trade?**

Answer: **The accidental movement of exotic pathogens between countries and regions is a risk to agriculture and the environment. Lists of plant pathogens in countries or regions help in the assessment and minimisation of this risk. These lists are essential when countries seek to export commodities to international markets. The only reliable lists are those that can be verified by specimens in reference collections.**

The reliability of information sourced for a particular specimen record will ultimately depend on:

- the availability of specimen collection records
- the method of pathogen identification
- the stability of the taxonomy on which the name of the pathogen is based.

Equipment (excludes consumables) in a reference collection

Pressed and dried specimens

- Dedicated insect-proof room
- Air-conditioner (20–23°C)
- Dehumidifier (40–60%)
- Shelves/compactus
- Fire alarm

Slide mounts

- Closed cabinet
- Heated water bath
- Microwave
- Glass microscope slides
- Coverslips

Cultures

- Laminar flow or clean bench
- Incubator
- Refrigerator (5°C)
- Freezer (-20°C)
- Ultralow freezer (-80°C)
- Freeze drier

Genomic DNA

- Ultralow freezer (-80°C)

Equipment (excludes consumables) in a reference collection

- Computer, printer, and scanner
- Database software
- Compound microscope
- Stereomicroscope
- Digital camera (microscope mounted)
- Packets, labels, annotation slips (acid-free)
- Access to scientific literature

Plant pathogen reference collections are dispersed around the world across multiple institutions. Many are in national and regional ministries of agriculture, forestry and the environment; museums of natural history; crop-specific research institutes; and universities. The cost of maintaining a reference collection has dictated that many are funded at a national level, rather than at a regional level. Many of the large reference collections of plant pathogens are held in government agencies that service agriculture and the environment, rather than in universities or the private sector. Reference collections benefit from national and international collaboration with research scientists and other collections.

Frequently asked question: How many culture collections are there in the ASEAN region?

Answer: The World Data Centre for Microorganisms (accessed 7 Sep 2023) recognised 112 microbial culture collections in ASEAN countries, with most in Thailand (65 collections), Indonesia (22) and Malaysia (13). Four ASEAN countries (Brunei, Cambodia, Laos, Myanmar) did not have registered microbial culture collections. This is inadequate for a region that includes four (Sundaland, Wallacea, the Philippines, Indo-Burma) of 25 biodiversity hotspots worldwide.

Frequently asked question: Where are the large culture collections in the ASEAN region?

Answer:

1 The National Center for Genetic Engineering and Biotechnology (BIOTEC), BIOTEC Culture Collection (BCC) Khlong Luang, Pathum Thani, Thailand, holds over 80,000 living specimens of actinomycetes, bacteria and fungi. Specimens in BCC are routinely tested by BIOTEC researchers for biodiscovery of secondary metabolites, enzymes and bioactive short peptides. BCC provides training, identification and liquid-drying preservation services to the public. Most microorganisms in the collection were isolated from insects (insect pathogenic fungi), with other specimens isolated from soil, seeds, decayed wood, plants, lichens, dung, fresh water and sea water. Specimens are either cryopreserved at -80°C, freeze-dried or stored in liquid nitrogen.

2 The Indonesian Culture Collection (InaCC), Microbiology Division, Research Center for Biology, Indonesian Institute of Sciences is in the Cibinong Science Center, West Java, Indonesia. InaCC holds about 3,000 living specimens of actinomycetes, archaea, bacteria, fungi, microalgae and bacteriophages. InaCC has modern laboratories with standard equipment for chemistry and molecular biology. Cultures are preserved (deep-freezer, L-dry ampoules, liquid nitrogen) and databased.

3 The Philippine National Collection of Microorganisms (PNCM), National Institute of Molecular Biology and Biotechnology (BIOTECH), University of the Philippines Los Baños, holds over 2,000 living specimens of algae, bacteria and fungi. PNCM provides services in the detection, enumeration and identification of microorganisms in food, feed, water, and agricultural and environmental samples.

2.3 MANAGEMENT CHALLENGES

Curators of reference collections will inevitably face challenges, including:

- keeping temperatures stable at 20–23°C and relative humidity at 50% or less
- preventing and controlling outbreaks of insect pests (cockroaches, mites, psocids, silverfish and tobacco beetles)
- managing small budgets and a lack of support for basic research.

The first detection of insect pests in a laboratory usually signifies a major outbreak. The most devastating pests in a mycological laboratory are mites that feed on fungi and cause contamination by transferring bacteria and fungi between cultures. Mites are often first detected by trails of contaminant fungi and bacteria they leave on plate cultures, or by trails left in droplets of condensation on lids of plates. Prevention of insect pests through sound hygiene is preferable to having to control an outbreak. Sound hygiene practices include:

- examination of all cultures that come into the laboratory for mites as soon as possible; if the culture is to be retained, take a subculture, and then discard the original plate by autoclaving
- separate incubators for clean cultures and primary isolation plates
- destruction of old cultures and surplus plant material as soon as possible by autoclaving
- regular cleaning of all surfaces with 70% alcohol
- seal culture plates with plastic film or cling wrap
- destroy infested and contaminated cultures immediately by autoclaving.

Frequently asked question: **How are mites controlled in culture collections?**

Answer: **Mites can become a chronic problem in fungal culture collections. These mites feed on fungal colonies and contaminate cultures by transferring bacteria and fungi among plates. Mites are introduced to laboratories on fresh plant material, shoes and clothing, the bodies of insects and cultures from other laboratories. They thrive in tropical conditions. If undetected, mites can devastate cultures in a laboratory in days, as they rapidly move from plate to plate in incubators or on benches.**

2.4 LOAN REQUESTS

Reference collections generally make specimens available via short-term loans for scientific research. Agreed loan conditions ensure the safety and security of specimens. Specimens should only be lent to other reference collections if safe transport and secure storage is certain.

Frequently asked question: **What conditions should reference collections attach to loans?**

Answer: **Most reference collections that lend specimens require a formal loan agreement (or material transfer agreement) that documents the conditions of the loan. The loan agreement may need to be signed by the head of the institution that maintains the reference collection. The loan agreement may specify the following conditions:**

1. **Purpose of the loan**
2. **Commercial use of specimens**
3. **Security of the loan**
4. **License fees**
5. **Transfer to third parties**
6. **Retention of material**
7. **Period of loan**
8. **Removal of material for destructive sampling, including for DNA extraction**
9. **Return of loan and packaging**
10. **Photography**
11. **Annotations**
12. **Intellectual property**
13. **Unpublished and manuscript names**
14. **Citing specimens**
15. **Acknowledgment.**

Specimens sent overseas are sometimes subject to biosecurity treatments on arrival. These treatments may include heat treatment, fumigation and gamma irradiation, which can damage specimens. If there is any doubt about how a specimen will be treated or whether a particular treatment is damaging, then the specimen should not be loaned.

3 Field collection

Many specimens in plant pathogen reference collections originate from material collected in either agricultural or natural environments by plant pathologists during targeted surveys. The objectives of the survey will determine the equipment, techniques, sites and skills required in the field. The degree of confidence needed to determine whether a plant disease is present or absent will dictate the scale of the survey. A rough estimate of sample size may be obtained from a pilot survey in which disease intensities are estimated for different sample sizes. Random sampling often yields good results because many plant diseases have a clustered distribution. The best sampling method must be chosen for any survey. Other considerations are the timing of the survey and the appropriateness of the sample of plant material as an indicator of disease.

Frequently asked question: **Who collects samples and specimens of plant diseases?**

Answer: **The primary collectors of specimens in plant pathogen reference collections are plant pathologists and researchers in agricultural and environmental science, most of whom work in government departments, research centres and universities. There are also private collectors, who may have a personal interest in the diversity of a group of organisms, e.g., larger fungi (mushrooms and toadstools).**

Frequently asked question: **Why submit samples and specimens of plant diseases to a reference collection?**

Answer: **Any specimen that represents a new record of a pathogen on a particular plant host or at a new location warrants deposition in a reference collection. Specimens in a reference collection provide a reliable way to verify or correct the identity of plant pathogens recorded in surveys and research studies. Most scientific journals will require evidence that specimens were deposited in an internationally recognised reference collection before considering manuscripts for publication, as this ensures that the reported experiment is repeatable. Most reference collections have taxonomic specialists, who can confirm the identity of specimens.**

3.1 COLLECTION PERMITS

Permission must be obtained from the owner of the land, whether publicly or privately owned, prior to collecting samples of diseased plants. In some places, official permits will be needed to collect and move samples for biosecurity reasons. The collection of biological samples may be restricted in ecologically sensitive areas. The movement of samples between countries will certainly require import and export permits, as well as biosecurity declarations.

Frequently asked question: **What is the Nagoya Protocol?**

Answer: **The Nagoya Protocol is a supplementary agreement to the 1992 Convention on Biological Diversity (CBD) that aims to ensure biodiversity**

conservation, sustainability and sharing of benefits that arise from research with biological specimens. The Nagoya Protocol came into force in 2014 as an international agreement. The key principles of the CBD are that countries own the biological resources within their borders and that access to their biological resources requires permission. Further, there should be acknowledgment and involvement of the holders of traditional knowledge. The Nagoya Protocol aims to ensure that the owners, including traditional owners, receive a fair share of any benefits that arise from research conducted with specimens collected from their land.

Frequently asked question: What are the consequences of the Nagoya Protocol for collectors and curators?

Answer: It means that collectors will need a permit stating what may be collected (and when and where) and how the collected material will be used. It may mean that curators can only lend specimens under material transfer agreements that stipulate specimens are for research only and may not be commercialised without a benefit-sharing agreement.

Frequently asked question: What biological specimens are covered by the Nagoya Protocol?

Answer: The Nagoya Protocol covers all plants, animals and microorganisms, whether living or dead material. It also covers DNA extracts.

Frequently asked question: How many countries have ratified the Nagoya Protocol?

Answer: As of July 2023, the Nagoya Protocol had been ratified by 139 countries.

3.2 FIELD SAMPLES

Samples collected in the field may be used to diagnose a plant disease or as a taxonomic resource. The best plant samples are those in early to middle stages of disease, when pathogens are still active. Severely diseased plant samples are often unusable because the pathogen may no longer be viable and saprobic organisms may have colonised necrotic tissues. This usually makes isolation of the pathogen impossible. Basic knowledge of the signs and symptoms of plant diseases is essential to ensure that the material is colonised by the pathogen. In some cases, symptoms may appear in one part of the plant, but the pathogen is found elsewhere. For example, wilt disease symptoms appear in the leaves, although the pathogen occurs in the vascular system of the roots and stems.

Frequently asked question: What is the difference between a sample and a specimen?

Answer: A sample is a portion of diseased plant material collected from the field for diagnosis. A specimen is a portion taken (or derived) from the sample that represents and enables the identification of the pathogen. The two terms are often interchangeable.

Frequently asked question: **What field equipment is needed to collect samples of diseased plants?**

Answer:

Field equipment for sample collection

Secateurs	Plant press	Newspaper	Labels
Hand lens	Trowel	Paper bags	Plastic bags
Machete	Ink markers	Envelopes	Pencils
GPS	Camera	Maps	Ice box

When collecting and handling diseased plants in the field:

- Determine the identity of the host plant. If the identity of the host plant is uncertain, collect healthy plant material, particularly flowers and fruits. Ensure that healthy material for host identification is selected from the same plant species as the diseased material. This may be difficult for pathogens that destroy the inflorescence, e.g., many smut fungi on grasses.
- Use paper bags for all samples except viruses (see virus-specific details below). Never wrap fresh plant material in plastic, as this causes the sample to sweat, which encourages saprobic organisms to colonise and decompose plant tissues. Plastic bags should only be used for short-term storage of damp specimens.
- Pack plant samples to avoid physical damage and condensation.
- Write labels in pencil, as ink may run if the paper becomes damp.

In some circumstances, it is prudent for the collector to wear disposable booties and gloves to prevent inadvertent spread of the pathogen, e.g., soil-borne fungal or oomycete pathogens in mud sticking to boots.

After field collection:

- Fresh material needs to be pressed and/or dried prior to freezing at -20°C to kill mites and insects. Specimens may be kept in the plant press during freezing. Plant viruses can be temporarily preserved in small desiccators.
- Bacteria and fungal pathogens that require isolation into pure culture should be kept fresh by refrigeration before the specimen is dried.
- Material that may deteriorate quickly, e.g., fleshy macrofungi, must be photographed and preserved.

Bacterial plant diseases

Samples of plants with suspected bacterial diseases may deteriorate rapidly. Place samples in paper bags and wrap in moistened newspaper to prevent drying. Keep the sample cool and out of direct sunlight. Pressed and dried specimens of leaf spots and blights can be used as reference specimens. Many bacterial pathogens survive for months, or even years, in dried material stored at room temperature.

Fungal plant diseases

Collect plant samples (leaves, shoots, roots, flowers, fruit) when surfaces are dry, or, if this is not possible, blot the leaves dry with absorbent paper. Spread and place the samples between layers of newspaper or other absorbent paper prior to pressing. If the samples are particularly fleshy, change the newspaper daily until dry.

Rust fungi have multiple life cycle stages, some of which are diagnostic for identification. Collect brownish-black pustules (telia), orange-yellow pustules (uredinia), as well as white or yellow cup-like structures (aecia). Smuts often destroy the inflorescence, or parts of the inflorescence,

of grasses and sedges. Fold diseased samples of rust and smut fungi in newspaper to contain spores. Stem pieces should have both healthy and diseased areas of tissue. Samples of fruit should be chosen in early stages of symptom development. Secondary rots and saprobes rapidly colonise fruit, which makes isolation of the causal pathogen difficult. Wrap fruit and stem pieces separately in dry newspaper.

Fleshy macrofungi, particularly *Agaricales*, deteriorate rapidly and are difficult to transport back to the laboratory. Do not break the stipe of macrofungi, whether in soil or on wood. Macrofungi should be wrapped individually in newspaper and placed in a container to avoid damage. Specimens of macrofungi need quick drying, which can be done in a fan-ventilated oven (45°C, overnight).

The pattern and colour of spore prints of gilled fungi aid their identification and provides a source of inoculum for cultures. A spore print is made by severing the pileus (cap) and placing it (gills downwards) on a piece of white card in still air for a few hours. A black card should be used for white-spored fungi.

Dry plant and fungal samples with silica gel if high-quality DNA is needed for downstream applications and cannot be extracted from fresh material. Silica gel helps remove water from the sample and preserves DNA before extraction. Alternatively, high-quality DNA may be extracted from fungal specimens preserved in absolute ethanol (>96%).

Nematodes and other soil-borne pathogens

Nematodes and other pathogens associated with roots and other below-ground plant organs are not easily detected. Do not pull plants and roots from the soil, as this may shear off diseased tissue or pathogens. Shake off excess soil from roots or wash the roots gently, unless the sample is to be tested for nematodes. Soil contains many microorganisms that colonise dead or dying tissue. These saprobes can interfere with the recovery of pathogens from diseased tissue. When removing soil from roots, do not scrub the roots, as this can lead to the loss of root tissue that may be useful in disease diagnosis. Wrap the roots in newspaper for transport to the laboratory.

Soil samples should be taken at least 5–10 cm below the surface in the root zone, where plant pathogens are most abundant. Individual soil samples should weigh approximately 250–300 g. Soil-borne pathogens are not evenly dispersed in the soil, tending to aggregate around infection points. If a plant shows patches of poor growth, separate soil samples should be taken from affected and healthy areas for comparison. Alternatively, the more random samples collected, the more accurate is the disease assessment. If sampling for nematodes, take care not to physically damage the nematodes through abrasion. Avoid sampling soil that is very wet or dry.

Roots should be either included in the sample or taken separately. For herbaceous plants, approximately 25–100 g of root tissue is sufficient. For woody plants, it may be necessary to excavate to a depth of up to 30 cm near the base of the tree or until roots are found that show the margin between healthy and diseased tissue.

Soil samples should be placed in strong plastic bags and labelled with pencil-written paper or plastic label placed inside the bag. Samples should be kept cool. Do not leave

soil samples in the sun or in a vehicle parked in the sun. Samples should be processed or despatched for analysis as soon as possible. Samples can be stored in a refrigerator at 4–8°C for several days without deterioration.

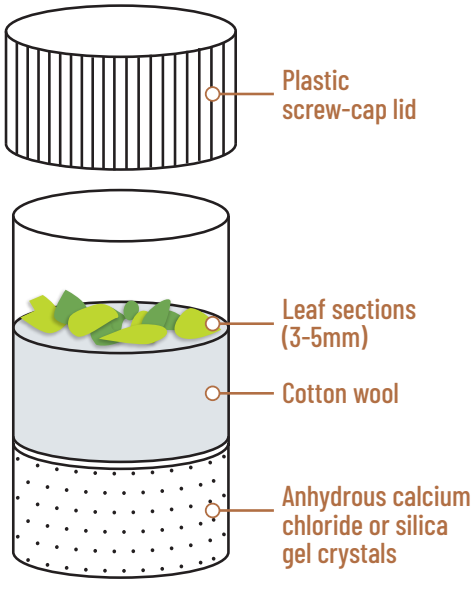
Viral plant diseases

Viral plant disease samples should be kept out of direct sunlight and at cool temperatures, e.g., in a cooler box or a refrigerator. It is critical that samples are kept fresh but not allowed to freeze, as this ruptures plant cells and organelle walls. Frozen tissue will rapidly deteriorate when thawed. To prevent samples from wilting, it is best to wrap them in slightly moistened paper towelling or newspaper and place in a zip-lock plastic bag. Viral plant disease samples should only be stored in this way for periods of days to a maximum of a week, depending on storage conditions and the nature of the plant material.

If plant material cannot be processed quickly, the tissue can be dried in small desiccators, e.g., small plastic containers with anhydrous calcium chloride ($CaCl_2$) crystals or silica gel up to one-third by volume. Swab leaves with water or alcohol if dusty or covered in sooty mould or scale insects. Use scissors or a safety scalpel blade to cut leaf tissue from near the centre of the lamina. Cut 3–5 mm squares and place 5–10 squares in each desiccator. Sterilise scissors or safety blades in alcohol or a 10% sodium hypochlorite (NaOCl) solution between samples to prevent cross-contamination. The sample should be separated from the desiccant by cotton wool or tissue paper. Store samples at 0–4°C.

Leaf samples of some plant species can be stored in tubes of 90% ethanol and desiccated later in the laboratory. This technique does not work for all plant species, and it is necessary to test whether nucleic acid extracts can be amplified before using this method on a survey.

Finally, there are commercial products that preserve plant sap extracts in a state that will allow downstream use in molecular diagnostic assays, such as PCR (polymerase chain reaction). One option is to use FTA® cards, which are pieces of filter paper that have been impregnated with chemicals that lyse cells, denature proteins and protect nucleic acids from nucleases, oxidation, and UV damage.



Desiccator for virus-infected leaf samples

3.3 LABELLING FIELD SAMPLES

A mobile phone, digital app or notebook may be used to record information about samples collected in the field. All samples submitted to a diagnostic laboratory or reference collection should be preceded by their collection details in a database (spreadsheet) or accompanied with completed field sample labels. Samples (and specimens) that do not have collection details have little value. The name of the collector (and/or person submitting the specimen) and contact details are essential information. The reason for sample submission must be clearly stated, for example, for diagnosis or deposition in a reference collection.

Frequently asked question: What details are recorded when a specimen is collected in the field?

Answer: An example of a field collection slip is shown below.

Frequently asked question: Are there standards for recording specimen collection details?

Answer: Darwin Core is an international standard for recording specimen collection details.

BRIP 46795

Genus Ustilago ~~triodiae~~ lituana R.G.Shivas & K.Vánky

Species

Host genus Triodia Host species ~~sp.~~ epatica

Host variety/cultivar Common name

Symptom(s) smut

Collector(s) Ryley, M.J Marney, T.S. Shivas, R.G

Date collected 12/8/2005 Collector No. Grower

Precise location Foreshore (Cooke Point) off Dempster St.

Town Port Hedland State W.A. Lat. Long.

Additional information alt. ca. 3m

same host as BRIP 46796

PERTH 07094574 (several images of germinating spores)

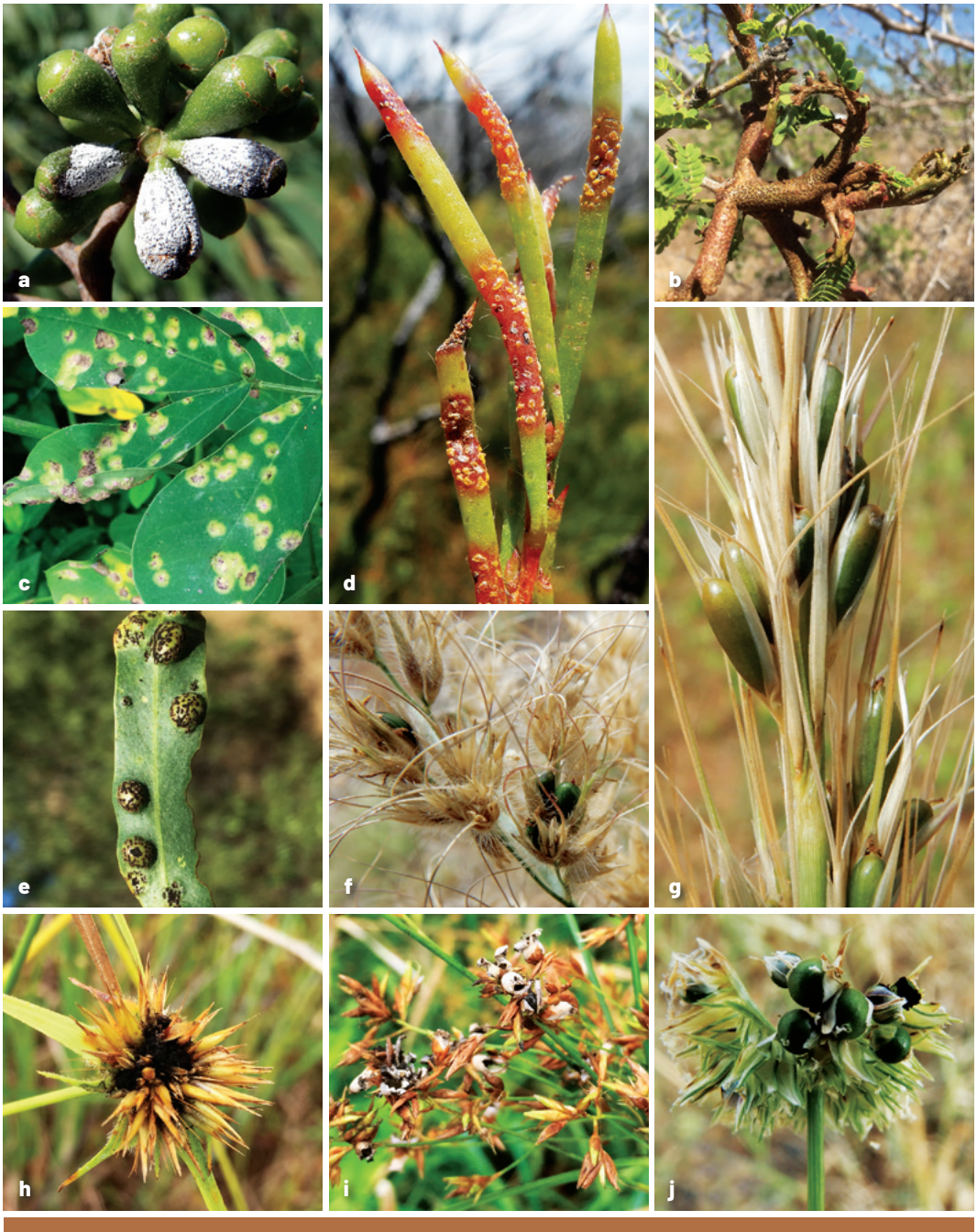
3.4 SIGNS AND SYMPTOMS OF PLANT DISEASE

Diseased plant samples are recognised by signs and symptoms. A sign of a plant disease is the visible presence of the pathogen, e.g., a fungal structure or discharge clearly associated with the diseased plant tissues. Some common signs of plant diseases are:

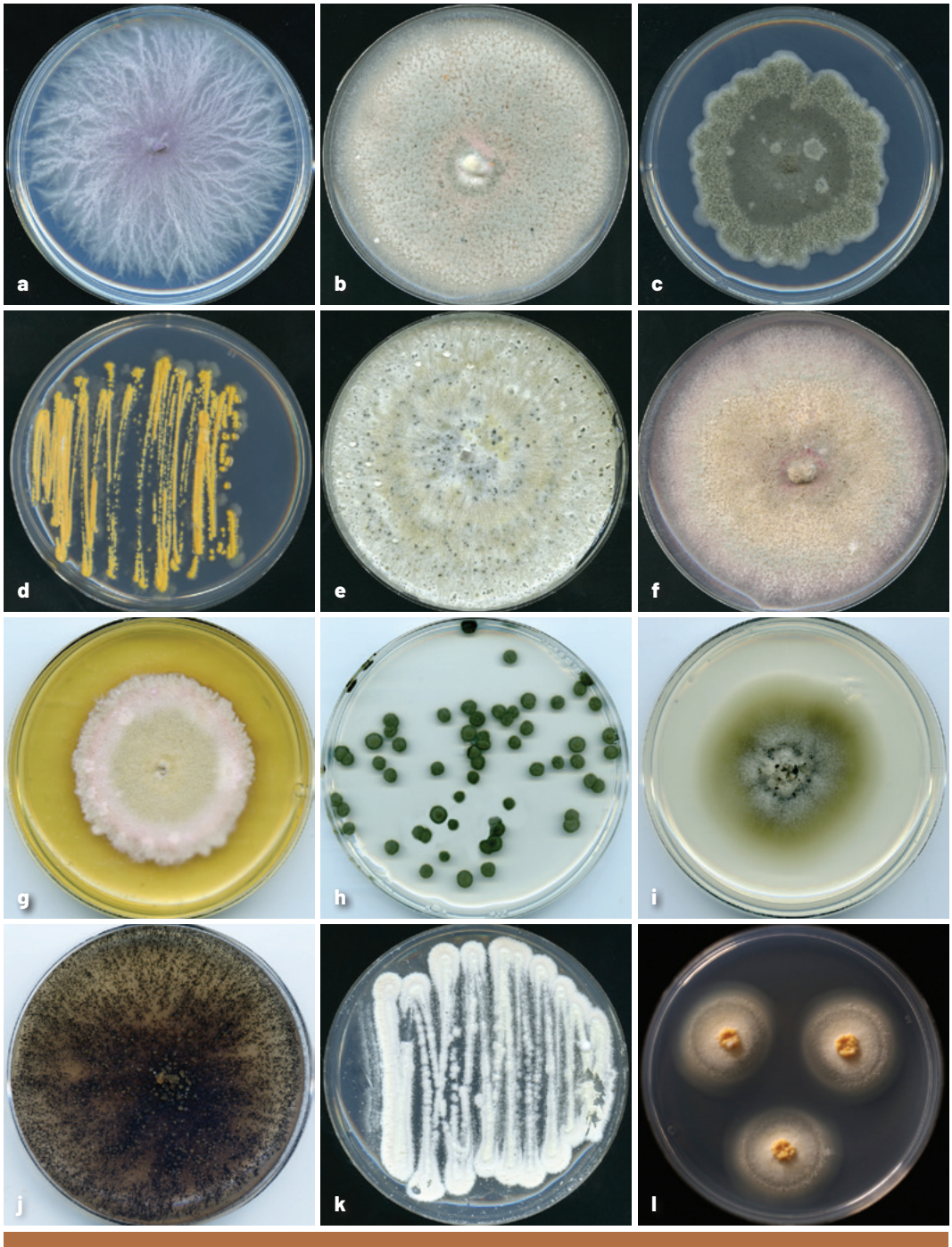
- ascomata, conidiomata, conidiophores (fungal fruiting structures that produce conidia)
- basidiocarps (the fruiting body of a polypore or agaric)
- mycelia (the mass of hyphae)
- ooze (sticky fluid exuded from a wound or opening)
- rhizomorphs (string-like strands of fungal hyphae, often dark)
- foliar mosaic, mottling, stippling or ringspot patterns
- abnormal growth patterns such as stunting, epinasty, distorted leaves, phyllody, fasciation and shoot proliferation.

A symptom is the visible change in the appearance of a plant or its parts arising from a disease.

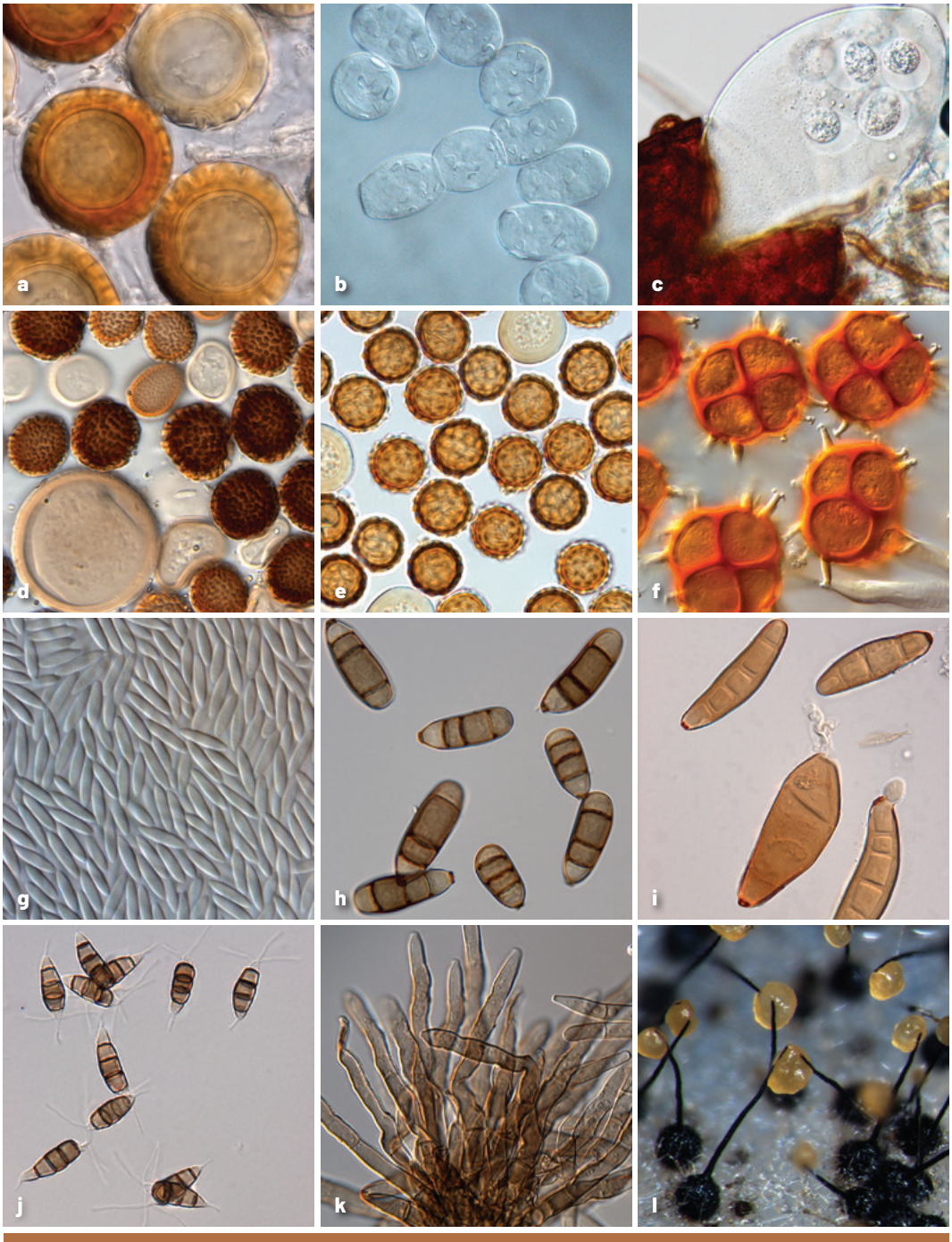
Plant disease symptoms caused by oomycetes and ascomycetes. a. ***Plasmopara sphagneticolae*** **on** ***Sphagneticola trilobata*****; b.** ***Nigrocornus scleroticus*** **on** ***Triodia*** **sp.; c. downy mildew on** ***Triodia*** **sp.; d.** ***Ustilaginoidea*** **sp. on** ***Scleria levis*****; e.** ***Podosphaera xanthii*** **on** ***Elatostema reticulatum*****; f. zonate leaf spot on** ***Ficus*** **sp.; g. cercosporoid fungus on** ***Ziziphus mauritiana*****; h.** ***Pseudocercospora fijiensis*** **on** ***Musa*** **sp.; i.** ***Pseudocercospora robertsiorum*** **on** ***Senna tora*****.**



Plant disease symptoms caused by basidiomycetes. **a.** *Quambalaria pitereka* on *Corymbia tessellaris*; **b.** *Cephalotelium neocaledoniense* on *Vachellia farnesiana*; **c.** *Puccinia arachidis* on *Arachis hypogaea*; **d.** *Austropuccinia psidii* on *Melaleuca nodosa*; **e.** *Endoraecium* sp. on *Acacia* sp.; **f.** *Macalpinomyces eriachnes* on *Eriachne cliliata*; **g.** *Langdonia aristidicola* on *Aristida* sp.; **h.** *Cintractia amazonica* on *Rhynchospora heterochaeta*; **i.** *Trichocintractia utricolicola* on *Rhynchospora corymbosa*; **j.** *Ustilago sparsa* on *Dactyloctenium radulans*.



Fungal colonies. **a.** *Fusarium ramsdenii*; **b.** *Colletotrichum acutatum*; **c.** *Leptosphaerulina queenslandica*; **d.** *Aschersonia* sp.; **e.** *Diaporthe searlei*; **f.** *Colletotrichum fioriniae*; **g.** *Stagonospora tautonensis*; **h.** *Pseudocercospora* sp.; **i.** *Periconia cyperacearum*; **j.** *Nigrospora cooperae*; **k.** *Beauveria malawensis*; **l.** *Moelleriella* sp.



Spores. **a.** *Peronosclerospora aristidiae*; **b.** *Podosphaera xanthii*; **c.** *Podosphaera plantaginis*; **d.** *Tilletia marjaniae*; **e.** *Tilletia palpera*; **f.** *Sphaerophragmium quadricellulare*; **g.** *Colletotrichum fioriniae*; **h.** *Curvularia tanzanica*; **i.** *Bipolaris* sp.; **j.** *Neopestalotiopsis macadamiae*; **k.** *Pseudocercospora proiphydis*; **l.** *Ceratocystis corymbiicola*.

Cattleya orchid infected with a sadwavirus; garlic infected by a mixture of poty-, carla- and allexiviruses; and tomato infected by tomato yellow leaf curl virus (left to right).
Images provided by John E Thomas and Nga Tran

Some common symptoms and descriptions of plant disease follow.

Symptom	Description
anthracnose	black, necrotic lesions often caused by *Colletotrichum*
black mildew	dense, black colonies of parasitic fungi *(Meliolales)*, often on leaf surfaces
blight	widespread rapid death of plant tissue
canker	sunken necrotic lesion, often on stems, branches or roots
damping-off	rot of seedlings near soil level, often before emergence, caused by some fungi *(Rhizoctonia)* and oomycetes *(Pythium)*
dieback	partial defoliation or branch death
downy mildew	whitish bloom on leaves and stems caused by some oomycetes *(Peronosporales)*
enation	small abnormal outgrowth of host tissue, flat extensions from veins on leaves and flowers
fasciation	shoot proliferation appearing as thin, flattened, curled shoots
gall	an abnormal swelling
gummosis	leaking of gum from host tissues
lesion	localised diseased tissue (a wound)
mosaic	patchy light and dark green colour in leaves, symptomatic of viral diseases

Symptom	Description
phyllody	flowers transformed into leaf-like tissue
powdery mildew	white, powdery bloom on leaves and stems caused by *Erysiphaceae*
pustule	blister from which fungal spores erupt
root knot	a swelling or gall on a root caused by some nematodes *(Meloidogyne)*
rot	softening and disintegration of plant tissues
rust	pustules formed by some fungi *(Pucciniomycotina)*
scab	crust-like, superficial, roughened plant tissue
scald	dried, seemingly burnt, plant tissue
shothole	leaf lesions in which the central necrotic tissue falls out leaving holes
smut	black fungal spore masses (*Microbotryomycetes* and *Ustilaginomycotina*) in the inflorescences of many grasses and some dicots
sooty mould	saprobic mats of dark superficial fungi living on exudates from insects (often aphids and scale) on leaves and stems
virescence	greening of flowers
wilt	drooping of leaves and stems
witches' broom	excessive proliferation of buds and shoots

4 Specimen preservation and collection records

The different types of physical specimens found in plant pathogen reference collections include pressed and dried specimens, living cultures, glass-mounted slides and DNA. Information, or metadata, associated with specimens is also managed by the collection curator. Such information includes collection details, images, publications and DNA sequences. The curators of these collections are often plant pathologists, who have the skills to work with different types of specimens and maintain collection records.

Specialised collection management software systems are available that store and share specimen collection records. Often the software needed is inexpensive, although the time and effort of adding records to a database is costly. Data entry remains a persistent and widespread problem for reference collections that have thousands, sometimes hundreds of thousands, of uncaptured specimen records. Another consideration in the development of a database is the quality and quantity of the underlying information. In many institutions, a significant amount of taxonomic work is needed, especially the molecular validation and identification of specimens.

Frequently asked question: **What is the specimen accession number?**

Answer: **Every specimen in a reference collection is assigned a unique identifier (called the specimen accession number) that distinguishes it from all other specimens in the world. This unique identifier accompanies all physical specimens as well as databased specimen collection details. This specimen accession number comprises the reference collection code followed by a number (often sequential). If the specimen accession number of a particular specimen is known, then the physical specimen (and sometimes also online specimen collection records) can be found with the aid of online directories.**

If a specimen contains more than one taxon, a suffix (a, b, c, etc.) is placed after the specimen accession number to represent each taxon. Living cultures derived from specimens are best given a different specimen accession number. If the specimen is divided into duplicate specimens, these should have the same specimen accession number as the specimen from which they were divided. Duplicate packets may be loaned, gifted or exchanged for specimens in other reference collections.

Frequently asked question: **What is the reference collection code?**

Answer: **The reference collection code is an abbreviation or initialism, which is often obtained from the name of the institution that houses the collection. Many fungaria and herbaria are registered in Index Herbariorum, which provides a searchable online global directory.**

Frequently asked question: **What information about the specimen collection should be kept?**

Answer:

Specimen collection details

1. Pathogen scientific name (genus, species, infraspecies)
2. Collection accession number
3. Type of specimen (dried, culture, slide mount, DNA)
4. Identification method (morphology, molecular barcode)
5. Collector(s), collector's number (if applicable), date of collection
6. Collection location
   - a. Precise place, town, district, province, country
   - b. Latitude and longitude (GPS coordinates)
7. Host scientific name (genus, species, cultivar)
8. Host symptoms (part of plant affected, severity)
9. Field notes
10. Field images

4.1 DRIED SPECIMENS

Dried specimens of foliar fungi and plant material are prepared in a plant press with straps to apply pressure until the material is completely dried. The paper in the press absorbs moisture and should be changed regularly over a period of 5–10 days. The press can be placed in the sun or in an air-conditioned room to accelerate drying. Removing plants from the press too early may result in wrinkling or the growth of saprobes on the specimen. Before incorporation into the reference collection, specimens should be frozen for 14 days at -20°C to kill mites and insects. Macrofungi are best preserved by rapid drying in a fan-forced oven. Small plastic boxes with lids or zip-lock plastic bags with silica gel are suitable for long-term storage.

Frequently asked question: **What insects can infest and damage a collection of dried specimens in a reference collection?**

Answer: **Many insects, including psocids, tobacco and herbarium beetles, silverfish and cockroaches can damage stored plant material in reference collections. Conditions in tropical regions with high temperatures and high humidity suit the rapid development insect pests that can destroy collections.**

Frequently asked question: **How are insect pests controlled in a reference collection?**

Answer: **Deep-freezing dried specimens at -20°C or below for 14 days is the best way to kill potentially damaging insects and their eggs. All new specimens should be treated this way prior to deposition into a reference collection. Existing specimens should be rotated through the freezer periodically.**

Specimens must be wrapped in plastic bags or placed in polystyrene boxes with tight lids

to prevent condensation of moisture during deep-freezing. After freezing, the specimens should be allowed to warm gradually by placing in an air-conditioned reference collection until they reach room temperature.

Fresh plant specimens should not be stored, nor examined, near the reference collection. Reference specimens removed for examination or loan from the environmentally controlled conditions should be returned to their position only after having been frozen at -20°C or below for 14 days.

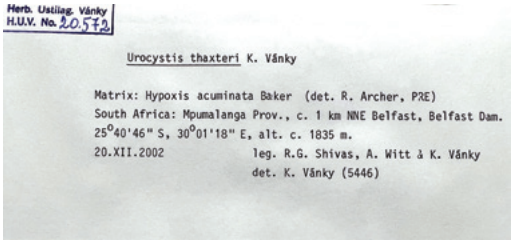
Dried specimens should be stored in archival (acid-free) paper or card packets. Specimens may be accompanied with annotation slips, correspondence and notes, dried cultures, images and illustrations. Each item inside the specimen packet should be clearly labelled with the accession number of the specimen. Microscope slides are best stored separately.

Frequently asked question: **What is a voucher specimen?**

Answer: **In a strict sense, a voucher specimen is a pressed, dried specimen of a plant that has been attached to a sheet of paper. In a general sense, a voucher specimen is a part or whole of any preserved organism that has been collected and deposited in a reference collection. Living specimens should not be used for voucher specimens.**

Frequently asked question: **How are specimen packets labelled in a reference collection?**

Answer: **An example of a specimen packet kept at the Queensland Plant Pathology Herbarium (BRIP) follows.**


Herb. Ustilag. Vánky
H.U.V. No. 20.572

Urocystis thaxteri K. Vánky

Matrix: Hypoxis acuminata Baker (det. R. Archer, PRE)
South Africa: Mpumalanga Prov., c. 1 km NNE Belfast, Belfast Dam.
25°40'46" S, 30°01'18" E, alt. c. 1835 m.
20.XII.2002 leg. R.G. Shivas, A. Witt & K. Vánky
det. K. Vánky (5446)

The physical arrangement of specimens in a reference collection is often at one of the higher taxonomic ranks to reflect their phylogeny. For example, fungal specimens may be arranged as either *Ascomycota* or *Basidiomycota*, with further arrangement within these groups alphabetically under genus name. Some large groups of plant pathogens may be kept separately. For example, rusts (*Pucciniomycotina*), smuts (*Ustilaginomycotina* and *Microbotryomycetes*) and powdery mildews (*Erysiphaceae*), with further arrangement within these groups alphabetical by host genus (because many species of rust, smut and powdery mildew are host-specific).

Dried specimens are often exchanged between collections as duplicates or exsiccata (a set of dried specimens with labels). Gifts and purchases are another means of adding to a collection. Permanent loans, where an entire collection is lent indefinitely from one institution to another, are a useful way to preserve valuable collections that are either inaccessible or unable to be adequately housed and maintained.

Viral disease samples are best stored as lyophilised tissue under vacuum and at -20°C. Commercially available, refrigerated, vacuum freeze-drying machines are available to prepare the samples for storage. Some viruses such as orthtospoviruses only retain infectivity when stored in ultracold conditions, such as in liquid nitrogen dewars.

4.2 CULTURES

The primary objective of keeping a bacterial or fungal culture is to maintain it in a viable state without morphological, physiological or genetic change for as long as required. Most bacterial and fungal cultures may be kept viable indefinitely by cryopreservation in an ultralow freezer (-80°C) in 15% and 30% glycerol, respectively.

Frequently asked question: **What is the procedure for the cryopreservation of bacteria at -80°C in 15% glycerol?**

Answer: **Prepare a stock solution of 30% glycerol (v/v) by mixing 30 mL of glycerol with 70 mL of water. Transfer the solution to a screw-capped glass bottle and sterilise by autoclaving at 121°C for 15 min. Loosen the cap during autoclaving. Transfer 500 µL of sterile 30% glycerol solution into sterile 2 mL microfuge tubes. Add 500 µL of bacterial culture to the tube (which dilutes the glycerol to 15%) and mix with the glycerol by shaking vigorously or using a vortex mixer. Label the tube and place in the -80°C freezer and record its location. The addition of sterilised porous ceramic beads (cryobeads) or glass beads to the glycerol in the microfuge tubes offers a surface for adsorption of bacterial and fungal cells. To recover the culture, an individual bead is removed and used to inoculate a liquid medium or streak onto the surface of an agar plate.**

Frequently asked question: **What is the role of glycerol in cryopreservation?**

Answer: **Glycerol prevents the formation of ice crystals in cells during freezing by lowering the temperature at which the glycerol/water solution freezes. Glycerol has the capacity to permeate cells walls with little to no toxicity. Without glycerol, ice crystals that form in cells can puncture cell membranes and damage cells.**

Some other methods that may be used to preserve and store bacterial and fungal cultures follow.

On agar slopes

An agar slant (or agar slope) is a small screw-capped bottle or cotton-wool-plugged tube, partly filled with a nutrient agar that has been allowed to cool with the tube laying at an angle, which increases the surface area for colony growth. The inoculated slants are incubated for about 3–5 days at room temperature until good colony growth appears. An option at this stage is to cover the slants with sterile mineral oil to a depth of 1 cm above the tip of slant surface. The oil must be sterilised by autoclaving twice at 121°C for 15 min. The slants are then stored in a refrigerator or a cold room at 5–8°C. The lids or mouths are sealed with plastic film or aluminium foil before incubation and during storage. The slants need regular inspection to ensure that agar has not dried or become contaminated. The cultures may need to be subcultured to fresh media every six months. Frequent subculturing takes time and can accelerate morphological change, loss of pathogenicity and decreased sporulation of cultures.

Water storage

Agar blocks are excised from the edges of young (one week) fungal cultures and submerged in sterile distilled water in 5–20 mL vials. The caps are screwed down and sealed with film or cling wrap. The vials are then stored at room temperature. This storage method is suitable for *Pythium* and *Phytophthora* as well as bacteria.

Freeze-drying

Freeze-drying removes water under reduced pressure from cultures in a frozen state by sublimation, which occurs when a frozen liquid goes directly to the gaseous state without passing through the liquid phase. Freeze-dried cultures are sealed and stored in glass ampoules or vials. Disadvantages of this method are that the freeze-drying apparatus is expensive to buy and maintain; experienced staff are required to operate the system; and viability should be checked after each run, as not all species survive the process. One of the advantages of this method is that the freeze-dried cultures may be stored at room temperature. Freeze-dried bacteria and yeasts need to rehydrate, either in a broth or distilled water before regeneration.

Liquid nitrogen

The storage of microorganisms at very low temperatures (down to -196°C) in liquid nitrogen is the best preservation method for most fungi and bacteria. Cultures, tissue or spore suspensions are treated with a cryoprotectant (10% glycerol) before aseptic transfer into sterile ampoules. Disadvantages of this method are that slow cooling rates are critical; equipment is expensive; a regular and reliable source of liquid nitrogen is essential; and experienced staff are needed.

Slide mounts

Permanent slide mounts are essential for nematodes as they can serve as type material. Permanent slide mounts in lactic acid are not recommended for fungi, as slides sealed with nail varnish dry out after a few years. Microscope slides should be kept in special-purpose microscope slide drawers. Each slide should be cross-referenced with its specimen accession number.

4.3 ON LIVING PLANTS

Obligate biotrophic plant pathogens (powdery mildews, rusts, downy mildews, and viruses) can be maintained on living host plant tissues. Some institutions maintain monosporic isolates of these microorganisms, on either potted plants isolated in finely meshed cages in an experimental glasshouse, or on detached surface-sterilised leaves kept on mannitol sucrose agar (mannitol, 20 g/L; sucrose, 10 g/L; agar, 7 g/L; tetracycline hydrochloride, 25 mg/L) plates or tubes. Continuous cultivation of host plants ensures that fresh leaves or young plants are always available. Surface-sterilised leaves kept on plates or in tubes are inoculated with the spores of the plant pathogen and kept in climate chambers at 16–20°C with 16 hour daily illumination.

The maintenance of living isolates of obligate biotrophic plant pathogens is laborious and needs preliminary experimentation to develop the best method for each host-pathogen system.

Virus isolates that are maintained on living plants will evolve over relatively short periods and may lose important biological traits, such as aphid transmissibility, if they are

continually propagated using methods such as mechanical inoculation.

4.4 DIGITAL IMAGES

If possible, take digital images of the plant disease specimen in the field at the time of collection. Digital images taken in the field of the specimen can be stored on a database together with the specimen collection records. The metadata linked to the digital image will include information that is useful in the compilation of specimen records. Digital image technology has provided high-resolution displays, as have smartphones with high-quality cameras. This technology combined with artificial intelligence (AI) heralds a new era for plant disease identification.

Frequently asked question: **What is a digital image?**

Answer: **A digital image is the representation of a real image as a set of numbers that can be stored on a computer.**

Frequently asked question: **What is a pixel?**

Answer: **A pixel (derived from the words 'picture element') is the smallest unit of a digital image that can be displayed on a display screen (computer, telephone, television). Pixels are the building blocks of digital images.**

Frequently asked question: **How is the digital information in a digital image stored?**

Answer: **Digital images can be stored in several file formats, including JPG, TIF, GIF and PSD.**

Frequently asked question: **What is metadata?**

Answer: **When an image is saved, the file also contains data about the image, known as metadata. Digital cameras automatically give image files unique names, e.g., IMG_001.jpg. The metadata includes the camera brand and model, image settings such as aperture, shutter speed, ISO number, focal depth and dots per inch. Other metadata includes the date and time when the image was created and the location (GPS) of the image.**

4.5 DATABASES

Collection management software systems store plant disease records in a database that allows rapid access to information, without the need to examine numerous physical specimens and labels. Data are stored in an organised manner that can easily be searched, retrieved, analysed and updated. The information in the database can be used to map the distribution of plant pathogens. A database may be as simple as a table set up in a spreadsheet program, such as Microsoft Excel, or it may be part of a sophisticated collection management system.

Collection management software systems allow the attachment of many types of multimedia, including digital images, word documents, PDFs, html, videos, DNA sequence data and chromatographs. Built-in reporting tools enable the rapid analysis of information. Contact information can be kept about collectors, determiners and clients. Task templates provide alerts about overdue loans, culturing needs and other events. Regardless of the complexity of the database being used to house information about the specimens, the data stored in the database must be accurate.

On a national scale, the databases of dispersed collections held by different agencies can be linked by a web-based system to create a virtual collection. On an international scale, searchable databases have linked primary data on biodiversity worldwide and made these data freely and universally available.

Frequently asked question: **Is a single large, centralised national reference collection better than having several small, widely distributed regional collections?**

Answer: **There are advantages and disadvantages for both centralised and distributed reference collections. Advantages of the centralised model are (i) more reliable flow of information, (ii) better quality of data, and (iii) it is cheaper to run due to economies of scale. In the distributed model, the flow of information is often impeded, as there are more computer systems that may fail. Further, in a distributed model the quality of data is often highly variable between different collections. The main disadvantage of the centralised model is that it is at greater risk of calamity, such as a natural disaster, by having everything in one place.**

5 Phenotypic identification of plant pathogens

The identification of plant pathogens starts in the field with the naked eye or a hand lens. Specimen identification in the laboratory begins with the observation of phenotypic characteristics under a microscope (stereomicroscope and compound microscope). Microscopic examination looks for clues that indicate the type of pathogen/s (bacterium, fungus, nematode, virus). Notes and images may be taken during the initial field and laboratory examinations. Later, these may be kept with the preserved specimen in the reference collection.

Frequently asked question: **What are phenotypic characteristics?**

Answer: **The phenotype is all the observable characteristics of an organism. Phenotypic characteristics include organism morphology, colony appearance in culture and host range.**

Frequently asked question: **What influences phenotypic characteristics?**

Answer: **Phenotypic characteristics are the outcome of gene expression (genotype), as influenced by the environment.**

Frequently asked question: **Can I confidently identify a plant pathogen using phenotypic identification?**

Answer: **Not at species rank, as there are many closely related cryptic species that have near identical or convergent phenotypic characteristics. At higher taxonomic ranks, genera or above, phenotypic characteristics are more useful, but, even then, not always reliable.**

5.1 LIGHT MICROSCOPY

A stereomicroscope is useful for initial laboratory examination of specimens. A compound microscope with magnification up to 1000 x is used to examine the morphology of bacteria, fungi and nematodes.

For fungi, a minute portion of sporulating tissue is removed under a stereomicroscope with the aid of a mounted needle or pointed scalpel and placed in a small drop of lactoglycerol. If the plant material is bulky, light pressure on the cover slip will help to disperse structures into a thin optical plane. Air bubbles can be removed from the mount by gently passing the microscope slide through a flame, making sure not to deposit carbon on the lower surface of the glass slide. If the slide is heated too vigorously the cover slip will pop off.

The adhesive tape technique is useful for showing the orientation of dry spores (as well as spore chains and the shape of conidiophores). A small piece of tape is held with forceps against a fungal colony (or leaf surface). It is then placed sticky-side up on a small drop of lactoglycerol on a clean microscope slide. A cover slip is placed over the drop for examination under a microscope.

Slide cultures can be used to examine fungi that have fragile chains of spores. Slide cultures are made using a Petri dish as a moist chamber and placing a bent glass tube inside the Petri dish, resting on

moistened filter paper. A sterile block of agar (approximately 1 cm^2) is placed on a flame-sterilised microscope slide resting on the bent glass tube. The fungus is inoculated at the four edges of the sterile agar block and a sterile cover slip is placed over it. After a few days, the slide can be mounted on a microscope and fungal structures viewed as they grow.

Tissue sections viewed under a microscope can show where fungal fruiting bodies occur in host tissues. Tissue sections are best cut by hand with a scalpel under a stereomicroscope.

Maintaining the optical quality of a microscope requires regular maintenance. Each laboratory microscope should be serviced once a year. In humid tropical environments, fungi can grow on the lenses and optical surfaces, affecting the optical qualities of the microscope. The lenses may become permanently damaged if left unchecked. Ideally, microscopes should be kept in an air-conditioned and dehumidified room. An alternative to using a commercial dehumidifier is to store the microscope in a dehumidifying cabinet or 'hot box' when not in use. A hot box is a reasonably airtight wooden or plastic box with a 25 W light bulb. The light bulb acts as a dehumidifier by providing enough heat to vaporise the humidity in the box, keeping the contents dry. In the tropics, camera equipment should also be stored in a hot box.

5.2 BACTERIA

Bacteria are single-celled, prokaryotic microorganisms that lack organelles and chlorophyll. Bacteria are ubiquitous, physiologically diverse, occupy a wide range of ecological niches and reproduce rapidly. Most plant pathogenic bacteria infect plants through wounds or natural openings, e.g., hydathodes, lenticels, and stomata, and colonise the intercellular spaces, rarely penetrating host cells. Plant pathogenic bacteria are spread by infected seeds, infected plant propagating material, water splash, insects and machinery. These bacteria colonise a host plant as epiphytes before the pathogenic phase. Some specialised bacteria invade vascular systems of plants and are spread by sap-sucking insects.

The bacteria that colonise the vascular systems of plants comprise the wall-less mollicutes (phytoplasmas and spiroplasmas), walled phloem-inhabiting bacteria and walled xylem-limited bacteria. Plant pathogenic bacteria have evolved an arsenal of virulence factors, including plant cell wall degrading enzymes, phytotoxins, extracellular polysaccharides and phytohormones.

The presence of ooze and the microscopic observation of bacterial streaming from cut surfaces under water are key indicators of bacterial infection. Dilution methods to isolate bacteria are frequently used. It is essential to obtain a single colony that represents a pure culture of the bacterium. Working with mixed cultures is wasted effort for meaningless results.

Frequently asked question: **What are the phenotypic tests for bacteria?**

Answer: **Colony appearance, cell morphology, Gram stain, biochemical tests (catalase and oxidase), substrate utilisation and physiological growth requirements. Phenotypic tests for bacterial identification have been replaced in most laboratories by molecular methods.**

Frequently asked question: **What is the Gram stain?**

Answer: **The Gram stain is a differential method of staining that assigns bacteria to one of two groups (Gram-positive or Gram-negative) based on the properties of their cell walls. Bacteria are observed microscopically on a glass slide using the Gram stain. Gram-positive bacteria retain a complex of crystal violet and iodine against elution with ethanol, whereas Gram-negative bacteria do not. Observation of unstained Gram-negative bacteria is enhanced by safranin counterstain. The Gram stain is useful as most plant pathogenic bacteria are Gram-negative. The procedure and reagents for the Gram stain are given in Appendix 2.**

Frequently asked question: **Can the potassium hydroxide (KOH) solubility test be used instead of the Gram stain?**

Answer: **The KOH solubility test (described in the next paragraph) is useful as a complement to the Gram stain. A negative KOH solubility test does not necessarily show that a bacterium is Gram-positive. Cultures more than 48 hours old may show a positive reaction after 30 sec of mixing the bacteria in the KOH solution. False positive results in the KOH test can occur if too much inoculum is used or if bacterial colonies are mucoid. False negative results in the KOH test can occur if too little inoculum is used or too much KOH.**

Take a loopful of bacteria from a young agar culture and mix in a drop of 3% aqueous KOH on a clean glass microscope slide until an even suspension is obtained. Lift the loop a few centimetres from the slide. If a string of slime is lifted with the loop (about 5–20 mm in length), the bacterium is Gram-negative. If a watery suspension is produced and no string of slime observed after repeated strokes of the loop, the culture is Gram-positive. The destruction of the cell wall of Gram-negative organisms and subsequent liberation of DNA, which is viscid in water, produces the string of slime. The Gram-positive wall is more resistant to KOH and remains intact, thus no DNA is released. This test is useful in cases of doubtful stain results.

Frequently asked question: **What is a false positive? What is a false negative?**

Answer: **A false positive is a result that appears positive when it should be negative. A false negative is a result that appears negative when it should be positive.**

Frequently asked question: **Are plant pathogenic bacteria Gram-positive or Gram-negative?**

Answer: **Most species of plant pathogenic bacteria are Gram-negative (stain reddish pink).**

Frequently asked question: **What are phytoplasmas and what disease symptoms do they cause?**

Answer: **Phytoplasmas are insect-transmitted plant pathogenic bacteria that belong to the genus *Phytoplasma*. These bacteria are found in the sieve tube cells of plant phloem tissue, and most are transmitted by phloem-feeding leafhoppers and planthoppers. Phytoplasmas cause a wide range of plant diseases, producing symptoms that include leaf discolouration, stunting, dieback, reduced leaf size (little leaf),**

witches' broom, phyllody, virescence and floral gigantism (big bud).

Frequently asked question: **How are phytoplasmas detected?**

Answer: **Phytoplasmas are detected using a 16S rDNA gene PCR using universal primers for all known phytoplasmas.**

Frequently asked question: **How many species of plant pathogenic bacteria are known?**

Answer: **There are about 150 species of known plant pathogenic bacteria. Most plant pathogenic bacteria belong to the genera *Acidovorax, Agrobacterium, Brenneria, Burkholderia, Clavibacter, Dickeya, Erwinia, Liberibacter, Lonsdalea, Pantoea, Pectobacterium, Phytoplasma, Pseudomonas, Ralstonia, Spiroplasma, Streptomyces, Xanthomonas, Xylella* and *Xylophilus*.**

5.3 FUNGI

Fungi are eukaryotic organisms that digest organic matter externally and then absorb it directly through their cell walls. Fungi differ from other eukaryotes by having cell walls composed mainly of glucans, chitin, and glycoproteins, and cell membranes that contain ergosterol.

The vegetative body of a fungus is either unicellular or multicellular. Unicellular fungi are called yeasts. Multicellular fungi are filamentous, with individual filaments (hyphae) collectively called mycelium. Growth of the mycelium occurs at the tips of the hyphae. Fungi reproduce by spores, which are specialised reproductive bodies consisting of one or a few cells. Spores may be formed asexually or as the result of a sexual process. Sexual reproduction occurs in most groups of fungi. In some, two cells (gametes) unite to produce a zygote called a zygospore. In the ascomycetes, the sexual spores are produced within the zygote cell, the ascus, and the spores are called ascospores. In the basidiomycetes, the zygote cell is called the basidium and the spores are called basidiospores.

Fungal morphology is commonly observed microscopically on a glass slide in a droplet of water under a cover slip. Lactic acid or lactoglycerol is the most used mounting medium for fungi. Staining fungal structures with dyes should be avoided, as some of the dyes and solvents are harmful to human health.

Lactoglycerol

Lactic acid	25 mL
Glycerol	50 mL
Distilled water	25 mL

Frequently asked question: **Are phenotypic methods used to identify fungi?**

Answer: **Molecular methods have replaced phenotypic characteristics (morphology, host, pathogenicity) as a reliable means of fungal identification.**

Frequently asked question: **What is a kingdom and how many kingdoms of life are there?**

Answer: **A kingdom is a taxonomic rank comprised of phyla. There are seven kingdoms of cellular life: *Animalia, Plantae, Fungi, Protozoa, Chromista, Archaea* and *Bacteria*. The taxonomy of viruses is independent of all other life**

forms, and there are six realms and ten kingdoms of viruses.

Frequently asked question: **Which kingdoms contain plant pathogens?**

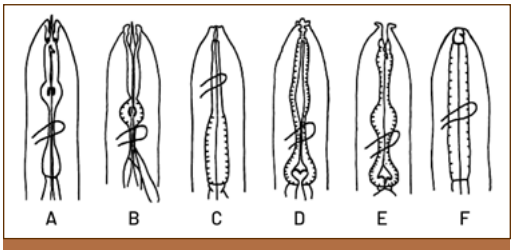
Answer: ***Animalia*** **(e.g., nematodes),** ***Bacteria, Chromista*** **(e.g., downy mildews, as well as** ***Pythium*** **and** ***Phytophthora*****),** ***Fungi, Protozoa*** **(e.g.,** ***Phytomonas*****).**

Frequently asked question: **What are the main groups of the kingdom** ***Fungi*****?**

Answer: **A recent classification of the kingdom Fungi divided it into seven phyla, namely,** ***Chytridiomycota, Blastocladiomycota, Neocallimastigomycota, Microsporidia, Glomeromycota, Ascomycota*** **and** ***Basidiomycota*** **(the latter two being combined in the subkingdom** ***Dikarya*****).**

5.4 NEMATODES

The identification of plant parasitic nematodes by phenotypic characteristics, specifically, morphology and host, is difficult. Plant-feeding nematodes have stylets (protrusible hollow feeding structures) in their oral cavity, which distinguish them from species that feed on other substrates. Some nematode species with stylets feed on fungi, algae and lichens, and others are predatory on soil microfauna. The classification of live specimens into presumptive trophic groups is possible with a good-quality dissecting microscope. For survey purposes, it may be sufficient for diagnostic nematologists to make tentative identifications, with molecular identification sought later, especially for new records or unidentified species. New records should be lodged in a reference collection.



Comparative anterior morphology of some groups of nematodes

Frequently asked question: **Are phenotypic methods used to identify plant nematodes?**

Answer: **Molecular methods have mostly replaced morphology and host range as a way to identify nematodes.**

For morphological examination nematodes are mounted on glass slides. The most widely used technique is the Seinhorst method, which involves hot fixation and stepwise transfer of nematodes to glycerol. There are many modifications of this procedure, which requires delicate and skilled labour, as individual nematodes must be transferred in a series of solutions.

Seinhorst solution I

Ethanol (95%)	20 mL
Glycerol	1 mL
Distilled water	79 mL

Seinhorst solution II

Ethanol (95%)	95 mL
Glycerol	5 mL

5.5 VIRUSES

Viruses are composed of a protein coat or shell called a capsid, which surrounds one or more molecules of nucleic acid. Plant viruses vary as to the type of nucleic acid, which may be single-stranded RNA (ssRNA), double-stranded RNA (dsRNA), single-stranded DNA (ssDNA), or double-stranded DNA (dsDNA). Viruses multiply in living cells by taking command of the genetic processes of an infected plant. The energy resources of the plant are thereby redirected to reproduce viruses. Viral infection impairs normal functions of plants, such as photosynthesis and growth. Sap-sucking insects, such as aphids and leaf hoppers, often spread viruses. Infected vegetative propagation material is a major concern in the movement of viruses. Viruses often survive in alternative weedy hosts.

Viroids are small circular ssRNA molecules that lack a protein coat. Viroids infect plant cells, replicate and cause plant diseases. Viroids are transmitted mechanically during pruning, by seed transmission and by vegetative propagation such as grafting.

Some of the phenotypic characteristics of viruses are:

- symptoms on naturally infected plants at different stages of infection
- modes of transmission, e.g., contact, seed, pollen, insect vectors
- the range of species that are susceptible to natural and experimental infection
- the shape of the virus particles as seen under a transmission electron microscope
- serology.

Frequently asked question: **Why are mixed infections of viruses common in individual plants?**

Answer: **Mixed infections of different virus species in individual plants are common because of repeated inoculations. Infection by one type of virus does not prevent infection from a different type of virus.**

Frequently asked question: **Are phenotypic methods used to identify plant viruses?**

Answer: **Specific serological (enzyme-linked immunosorbent assay, ELISA) methods are used to detect and identify some plant viruses.**

Frequently asked question: **How do viruses enter plant cells?**

Answer: **Viruses enter plant cells through wounds (broken epidermal hairs or small abrasions) or are injected into the leaf by feeding insects, mites or even nematodes.**

Frequently asked question: **What are some of the symptoms of a virus infection in plants?**

Answer: **Hypersensitive reaction, mosaic, leaf streak, mottle, chlorotic flecking, spotting, blotching, yellowing, ringspotting, cell necrosis, stunting, dwarfing, abnormal growth (leaf and stem distortion), hyperplasia (an increase in the number of cells), hypertrophy (an increase in the size of cells), enations, galls, colour break in petals, misshapen fruit, inclusion bodies in host cells, lignification of xylem elements and degeneration of phloem cells.**

Frequently asked question: How does one tell the difference between a virus infection, nutrient disorder and herbicide damage?

Answer: This can be difficult as the symptoms are similar. The distribution of affected plants provides circumstantial evidence. In the case of herbicide damage, the distribution of damaged plants may be consistent with spray drift, e.g., along the edge of paths or roads. In the case of nutrient disorders, the affected plants may reflect soil type or the application of chemicals. Viruses that are spread by insect vectors have clumped distributions or form a gradient of incidence from the source of infection, e.g., weeds or seed-borne infected seedlings. Other considerations are when the symptoms first appeared and whether more than one plant species is affected.

Frequently asked question: What famous role did tulip breaking virus (TBV) play in 17th-century Dutch history?

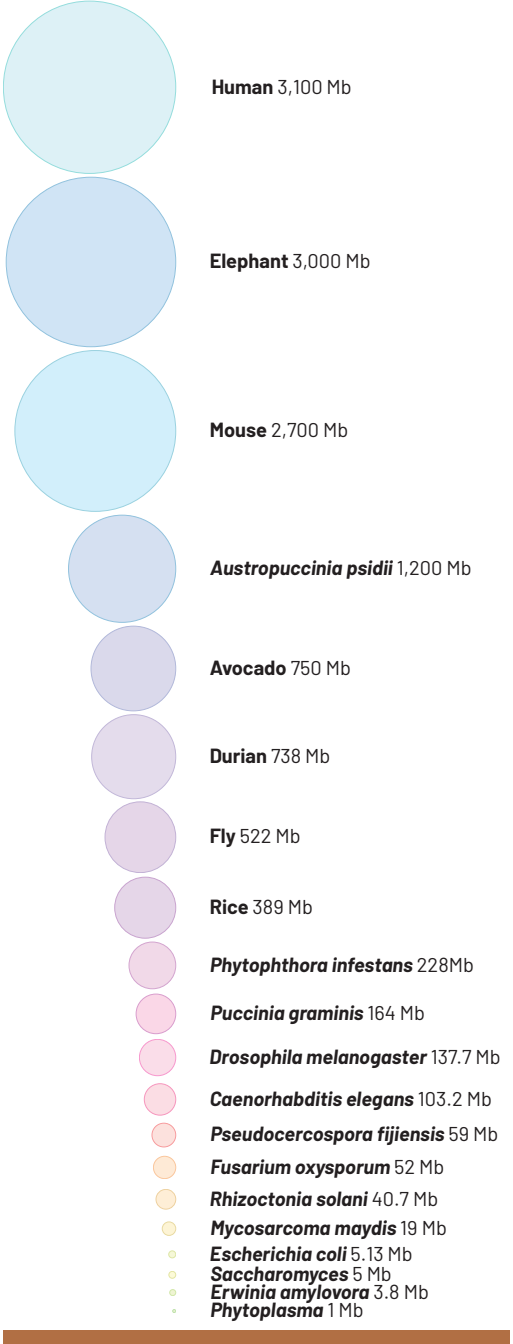
Answer: In the 1590s tulips were introduced from Turkey to Holland, where they became fashionable in Dutch high society. Tulip flowers with colour break in the petals became particularly desirable in early 17th-century Holland. This created a speculative bulb market that caused a bubble of high prices that inevitably crashed in 1636–37. This period of Dutch history is now known as 'Tulip mania' and represents one of the most famous market bubbles and crashes in history. Tulip breaking virus caused the colour break in the petals of the tulips in Holland. Of course, at that time the viruses had yet to be discovered.

6 Molecular identification of plant pathogens

Molecular methods have replaced phenotypic (morphological and biological) approaches to classify and identify life on earth. This change was needed because organisms that look the same (have similar morphological traits) or behave the same (have similar biological traits), may not be the same or even closely related. For example, birds, bats and insects have all independently evolved wings for flight from unrelated ancestors. This phenomenon is known as convergent evolution.

6.1 DNA IS A MOLECULE

Many individual species of microbial plant pathogens (bacteria, fungi, nematodes and phytoplasmas) that were delineated solely by morphology have since been shown by molecular (DNA) studies to represent many cryptic, phylogenetic species. These phylogenetic species, often geographically separated, challenge many aspects of plant health management. The DNA molecules of all organisms have genetic differences that can distinguish them through all taxonomic ranks (and as individuals). Molecular methods are now routinely used to identify and classify microorganisms. This is particularly useful for microorganisms, which have few reliable morphological or biological traits that distinguish species. An imperative for microbiologists is the ability to distinguish microbial pathogens from closely related non-pathogens that are morphologically identical.



Comparative genome sizes

Frequently asked question: What is a DNA molecule?

Answer: Deoxyribonucleic acid (DNA) is the chemical name for the molecule that carries genetic instructions needed for all living organisms (animals, bacteria, fungi, plants and some viruses) to develop, survive and reproduce. The DNA molecule consists of two strands that wind around one another to form a shape that looks like a twisted ladder, known as a double helix. The sides of the DNA molecule are complementary strands of DNA made of alternating sugar (deoxyribose) and phosphate molecules. Each sugar is attached to one of four bases: adenine (A), cytosine (C), guanine (G) and thymine (T). A base attached to a sugar and a phosphate molecule is called a nucleotide. The two strands in DNA are held together by bonds between the base pairs that form the rungs of the ladder, such that adenine only bonds with thymine, and cytosine only bonds with guanine.

Frequently asked question: How is the size of the DNA molecule measured?

Answer: Base pairs (bp) are used to measure the size of the whole or part of the DNA molecule. The total number of base pairs is equal to the number of nucleotides in one of the strands.

Frequently asked question: Who discovered the structure of the DNA molecule?

Answer: In 1951, Rosalind Franklin (1920–1958) produced data, including an X-ray diffraction photograph, which showed the DNA molecule was a helix. This laid the foundation for James Watson (1928–present) and Harry Crick (1916–2004) to show that DNA was structured as a double helix. Watson, Crick and Maurice Wilkins (1916–2004) were awarded the Noble Prize in Medicine in 1962 for this discovery.

Frequently made comment: It seems unfair that Rosalind Franklin was not also awarded the Nobel Prize.

Answer: The Nobel committee does not consider posthumous candidates.

Frequently asked question: How does a DNA molecule carry genetic information?

Answer: The sequence (order) of the bases along the strands of the DNA molecule carries the instructions (genetic code) for making RNA molecules and proteins. The significance of the discovery was noted in 1953 by Rosalind Franklin and Raymond Gosling (1926–2015) who wrote, 'an infinite variety of nucleotide sequences would be possible to explain the biological specificity of DNA'.

Frequently asked question: How does DNA make RNA molecules and proteins?

Answer: DNA is organised into genes, which are sections of DNA that lead to observable traits. The central dogma of molecular biology is that genes, which are made of DNA, are transcribed into RNA, which are then translated into proteins.

Frequently asked question: What is the difference between transcription and translation?

Answer: Transcription is the conversion of DNA in the cell nucleus into small, single-stranded messenger RNA (mRNA) molecules. Translation is the process by which ribosomes make proteins from the sequences carried by mRNA molecules.

Frequently asked question: Where are ribosomes found?

Answer: Ribosomes occur freely in the cell cytoplasm and are also attached to the endoplasmic reticulum in eukaryotic cells. Ribosomes are formed from a complex of RNA (transcribed from ribosomal DNA) and proteins.

Frequently asked question: Where is DNA found?

Answer: DNA is only found in the cells of living organisms (animals, bacteria, fungi, plants) and some viruses. In eukaryotes, DNA is found in cells in a membranebound structure called the cell nucleus (where it is called nuclear DNA or nDNA), with a small amount of DNA also found in cells in double-membraned structures called mitochondria (where it is called mitochondrial DNA or mtDNA). DNA is also found in the chloroplast (cpDNA) of plants and photosynthetic algae.

In eukaryotes (animals, fungi and plants), different linear dsDNA molecules wrap around histone proteins to form chromosomes, whereas in prokaryotes (bacteria and archaea), there is just a single chromosome composed of a circular dsDNA molecule and bacteria do not have histone proteins. The complete set of nDNA (all the chromosomes) is called a genome.

Frequently asked question: What are mitochondria (sing. mitochondrion) and what do they do?

Answer: Mitochondria are structures (organelles) in cells that likely originated billions of years ago when a bacterial cell was engulfed by another cell. The bacterial cell was not digested and formed a symbiotic relationship with its host. Mitochondria divide using their own circular strand of DNA (mtDNA) and there are many mitochondria in one cell.

Mitochondria convert oxygen and nutrients within the cell to supply energy for host cells. The process is called oxidative phosphorylation. Large numbers of mitochondria are found in cells where there is a high energy demand. In fungi, mitochondrial activity has been implicated in multiple aspects of fungal cell biology, including antifungal drug resistance, pathogenicity, senescence and virulence.

Frequently asked question: Is mtDNA useful for DNA barcoding?

Answer: mtDNA is widely used in phylogenetics, especially for animals and plants, as it is independent from nDNA and has maternal inheritance, a high copy number, and longer sequence length than the rRNA gene. Cytochrome c oxidase subunit I (COI or COX1) is a mitochondrially encoded gene, which is an extremely useful DNA barcode capable of species identification in a broad range of eukaryotes, including oomycetes and nematodes.

6.2 DNA BARCODES

Species identification for most eukaryotes (algae, animals, fungi, plants) relies on comparison of sequences of standard DNA barcodes. A molecular barcode is a sequence of a short region (250–1,000 base pairs) of DNA that uniquely characterises an organism at different taxonomic ranks. A DNA barcode may be a housekeeping gene or a unique and conserved genomic region.

The premise of DNA barcoding is that, by comparison with a reference library of DNA sequences, an unknown species can be identified. Molecular barcodes should be variable between species and conserved within species. DNA barcodes are also used to catalogue taxa and to determine species boundaries. DNA metabarcoding with high-throughput sequencing is used to identify organisms from a sample containing DNA from multiple organisms.

The internal transcribed spacer (ITS) region of the nuclear ribosomal DNA (nrDNA) operon is the most widely used DNA barcode for fungi, nematodes and plants. The ITS region is part of the nuclear DNA coding region transcribed into ribosomal RNA, which then binds with proteins outside of the nucleus to form ribosomes. The nrDNA operon is repeated in tandem many times in the genome and includes ITS spacer regions that do not form structural parts of the ribosome. The ITS region is flanked by the 18S small subunit ribosomal RNA gene (SSU) and the 28S large subunit ribosomal RNA gene (LSU), which form functional parts of the ribosome.

The most widely used barcodes for plant pathogens are 16S rRNA for bacteria; ITS for fungi; ITS, LSU, SSU and COI for nematodes; and the cytochrome C oxidase subunit 2 gene (COX2) for oomycetes. These gene regions vary among closely related species and are flanked by conserved regions that can be targeted for universal primer design. Genomic approaches have become standard to identify and classify plant viruses, which lack a universally conserved barcode.

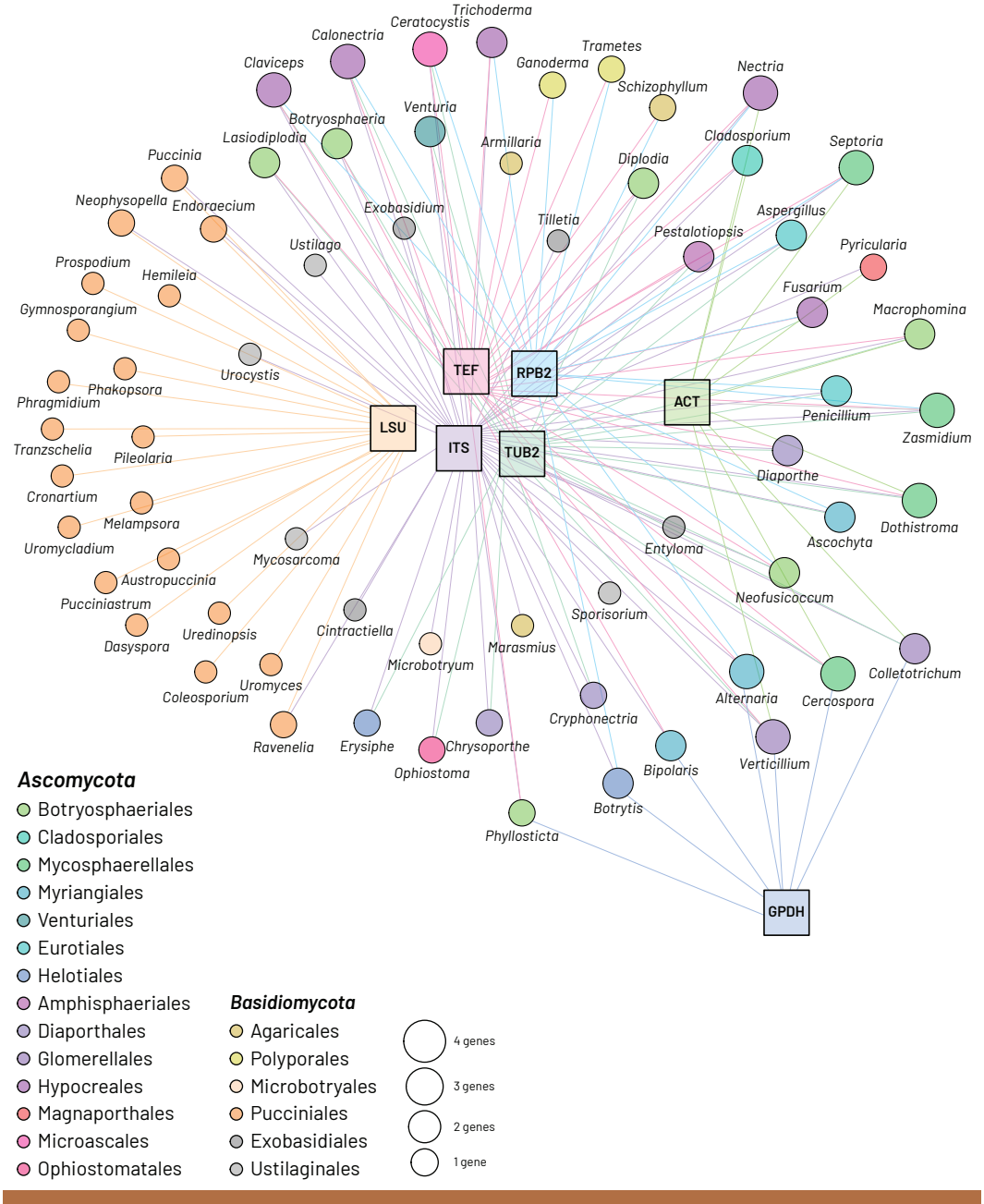
An advantage of DNA-based methods is that extremely small (picogram) quantities of DNA can be detected. These methods also have the advantage of speed and sensitivity. Disadvantages of DNA-based methods are their expense (equipment, reagents, facilities, staff training) and sample contamination caused by enhanced sensitivity.

Frequently asked question: **What DNA barcode region should I use to identify species of bacteria?**

Answer: **Start with 16S SSU rRNA and check recent literature to determine which barcode regions are best used to identify species. Bacteria have small, circular genomes, up to 5 million base (Mb) pairs, and their identification based on whole genome sequencing is gaining popularity.**

Frequently asked question: **What DNA barcode region should I use to identify species of fungi?**

Answer: **Start with ITS region to identify family, genus or species ranks of unknown fungi. It is the most widely sequenced locus, with over 1 million sequences available for comparison on the UNITE database. There are some genera of plant pathogens that lack variability in the ITS region. For example, species of *Aspergillus* are best differentiated by TUB2; species of *Diaporthe* by combined**



Network of barcode genes recommended for species-rank identification in common genera of plant pathogenic fungi

ITS, TUB2, CAL and TEF1-α; and species of *Fusarium* by TEF.

Frequently asked question: **What DNA barcode regions should I use to identify species of nematodes?**

Answer: **Start with 18S SSU rRNA for a preliminary identification to genus rank and check the recent literature to determine which barcode regions are best used to identify species.**

Frequently asked question: **What barcode regions should I use to identify plant viruses?**

Answer: **Universal PCR primers have been designed for many virus genera (e.g., potyvirus, orthotospovirus, tobamovirus, ampelovirus, badnavirus, begomovirus) but rarely for higher taxonomic ranks. DNA barcoding regions across different viral genera are not comparable and cannot be used to assemble a universal classification system for all viruses. Virus genomes can be assembled from high-throughput sequencing data, which is becoming standard practice within the virology community.**

Frequently asked question: **Why do plant viruses often have high genetic diversity?**

Answer: **Plant viruses replicate rapidly and generate large populations, which increases the likelihood of mutations. Many plant viruses have RNA genomes, which have a high mutation rate as RNA polymerases lack proofreading activity. The high mutation rate can give rise to populations of mutant sequences (haplotypes or variants) grouped around a master or modal sequence. The population of sequence variants of a virus within a plant is termed a 'quasispecies'.**

6.3 EXTRACTION OF GENOMIC DNA

DNA is extracted and purified in four basic steps.

1. Cells are lysed, whether mechanically (by freezing, grinding, pulverising with beads or heating) or enzymatically.
2. Proteins are removed, usually by changing their solubility with salts, a column or phenol.
3. DNA is precipitated (whether by a column, magnetic beads, or 100% ethanol or isopropanol) and washed.
4. DNA is made soluble in an elution buffer that prevents degradation, then stored at -20°C.

Genomic DNA is most conveniently extracted from microbial plant pathogens (bacteria, fungi, nematodes, viruses) with commercial kits. Kits allow the process to be precisely replicated. Most kits contain complete sets of nontoxic reagents, consumable equipment and instructions. The yield and purity of DNA may differ between extraction methods, e.g., the extraction method for a diagnostic PCR may differ from one that extracts high-molecular weight DNA for genome sequencing.

Frequently asked question: **What commercial kits are available for genomic extraction of plant pathogens?**

Answer: **There are many. Choosing the correct extraction kit is critical and will depend on the target organism, starting material and end uses. Ask colleagues, contact local suppliers and check the literature.**

Frequently asked question: **How long does it take to extract DNA using a kit?**

Answer: **The extraction process usually takes about two to three hours. High-throughput extractions will be limited by how many samples can be handled at once.**

Frequently asked question: **What is the starting material for genomic extraction of plant pathogens?**

Answer: **Culturable fungi should be grown from single spores or hyphal tips on PDA for seven days at room temperature (approx. 25°C) to build enough material that can be added to an Eppendorf tube for extraction. Non-culturable fungi (*Phyllachora*, powdery mildews, rusts, smuts) and oomycetes (downy mildews) should be extracted from small amounts (tip of a scalpel) of fungal material directly from the fungal structures (ascomata, pustules, sori, epiphytic mycelium and conidiophores) on the host plant. Smaller amounts should be used to avoid cross-contamination or genetic differences that may occur between pustules/structures. Avoid harvesting host tissue as far as possible, as this can inhibit PCRs or provide a host DNA template that is amplified preferentially to fungi. Elute DNA from small amounts of material in smaller volumes, about 15 µL, depending on the extraction protocol and downstream uses.**

6.4 PCR AMPLIFICATION OF DNA

Polymerase chain reaction (PCR) is a laboratory technique used to amplify specific regions of DNA. PCR uses two short DNA sequences called primers that target a portion of the genome for amplification.

Forward and reverse primers are designed to attach on opposite strands of the DNA target region. The technique can produce a billion copies of the target DNA sequence in a few hours. PCR produces a double-stranded product called an amplicon.

A typical PCR cycle has three steps: denaturation, annealing and extension. After each cycle, the amount of DNA is doubled. After 30 cycles under ideal conditions from a single target there will be 2^{10} copies (1,073,741,824).

Frequently asked question: **What equipment is needed for DNA amplification?**

Answer: **A thermocycler to control different temperatures in the PCR, pipettes and PCR tubes.**

Frequently asked question: **What reagents are used in a PCR?**

Answer: **The key reagents in a PCR are primers, nucleotides, a thermostable DNA polymerase, and the template DNA in PCR buffer. In PCR, magnesium ions (in the form of $MgCl_2$) are an essential cofactor that enhances the enzymatic activity of DNA polymerase, thereby boosting DNA amplification. Other chemicals can be added to aid PCR, e.g. dimethyl sulfoxide (DMSO) helps denature GC-rich templates, and bovine serum albumin (BSA) may help reactions overcome PCR inhibitors.**

Frequently asked question: What happens at each step in the PCR?

Answer:

1. **Denaturation:** A mixture of template DNA, PCR primers, nucleotides, PCR buffer and DNA polymerase in a buffer is heated to more than 90°C (usually 95–98°C, depending on the polymerase enzyme). At this temperature, the strands of the DNA double helix separate into two single strands (called denaturing).
2. **Annealing:** This mixture is cooled to 50–60°C and primers anneal to separated strands of DNA. Lower annealing temperatures will decrease the specificity of primers, and they may bind to a template despite mismatches with their sequence. Higher annealing temperatures (≥60°C) increase specificity of annealing, and primers will only bind if they are a direct match to the target.
3. **Extension:** This mixture is heated to 72°C for the DNA polymerase to synthesise new complementary strands of DNA starting at the bound primers.

Frequently asked question: What is *Taq* DNA polymerase and what role does it play in PCR?

Answer: Taq DNA polymerase is a heat-stable DNA polymerase named after the thermophilic bacterium *Thermus aquaticus*, from which it was first isolated in 1976. *Taq* DNA polymerase plays a critical role in PCR as it is chemically reactive at 72°C and synthesises new strands of DNA by extension of the primers. Alternative DNA polymerases to *Taq* include proofreading enzymes that have low error rates of DNA replication, although these are generally more expensive per reaction.

Frequently asked question: What is reverse transcription PCR (RT-PCR)?

Answer: RT-PCR is a modification of the standard PCR method, which is used when the starting nucleic acid template for the reaction is either ssRNA or dsRNA. The RNA molecule must first be converted to DNA before it can be processed by *Taq* DNA polymerase. This conversion step is catalysed by an enzyme called a reverse transcriptase. Reverse transcription occurs in a single incubation step before the PCR commences, usually at 42–50°C. Reverse transcription and PCR can be done in separate reactions or commercial, one-step RT-PCR kits are now available, which contain all the reagents needed for reverse transcription and PCR in the single tube.

Frequently asked question: How do I know if my PCR has worked?

Answer: A PCR product can be visualised using gel electrophoresis. If the sequence of interest has been copied, it will appear as a single band in the gel (as all the copied sequences will be the same size and run the same distance through the gel).

Frequently asked question: Why is there more than one band on my gel after a PCR?

Answer: Non-specific primer annealing is one possible cause that can be resolved by (i) increased annealing temperatures, (ii) increased concentration of $MgCl_2$, or (iii) decreased concentrations of primers in the final reaction volume. Too many

(>30) PCR cycles may also cause multiple bands due to an increased chance of error with each cycle. Multiple bands can occur when DNA is contaminated with more than one species.

Frequently asked question: What are primer dimers?

Answer: Primer dimers are the unwanted hybridisation of two primer molecules in a PCR caused by strings of complementary bases. As a result, PCR amplifies the dimer, which leads to competition for PCR reagents and inhibits amplification of the target DNA.

6.5 DNA SEQUENCING USING THE SANGER METHOD

Sanger sequencing is a method that determines the nucleotide sequence of DNA. The method was developed by Frederick Sanger and his colleagues in 1977. In automated Sanger sequencing, the four nucleotides in DNA (A, T, G, C) are uniquely labelled with a fluorescent dye and mixed in a single reaction (big-dye terminator reaction), similar to PCR. The product is run through a capillary and a computer reads fluorescent terminal bases to call the identity and position of each nucleotide. A laser excites the fluorescent tags in each band, and a computer detects the emitted light. The output is called a chromatogram, which shows the fluorescent peak of each nucleotide along the length of the template DNA.

Frequently asked question: Can a PCR mixture be used for Sanger sequencing?

Answer: Prior to sequencing PCR amplicons, a clean-up step is necessary, as the PCR mixture will also contain unincorporated primers, nucleotides, enzymes, short-failed PCR products and salts from reactions that may interfere with Sanger sequencing. There are several methods for PCR clean-up, including ethanol precipitation, bead or column-based purification and enzymatic approaches. Commercial kits are available for clean-up from either the reaction mixture or from single bands on an agarose gel. Many commercial sequencing providers also offer a PCR amplicon clean-up service – ask!

Frequently asked question: Should I use the same primers in PCR for sequencing?

Answer: Sanger sequencing and PCR can use the same primers, but each has a different purpose. PCR is used to amplify a target region of DNA in its entirety and produce an amplicon. The goal of Sanger sequencing is to generate every possible length of DNA up to the full length of the target DNA with a labelled base. Two PCR primers are needed in a PCR (forward primer and reverse primer), whereas one primer is needed for a Sanger reaction so that the product has only one fluorescent base when run in a capillary.

Frequently asked question: How should I get my PCR products sequenced?

Answer: Many commercial laboratories provide Sanger sequencing services. Ask colleagues, contact local suppliers and read the recent literature.

Frequently asked question: For what research did Sanger win Nobel Prizes?

Answer: Frederick Sanger (1918–2013) won the Nobel Prize in chemistry twice,

the first time in 1958 and again in 1980. The 1958 award was given for his work on the structure of the insulin molecule and the 1980 award for determining the base sequence of nucleic acids.

6.6 TESTING THE QUALITY OF SEQUENCES

Chromatograms from Sanger sequencing can be visualised in software to ensure they have clean, single peaks for each base call. Freely available software to visualise Sanger sequencing outputs (*.ab1 files) includes MEGAX, BioEdit and UGene, to name a few. If both forward and reverse directions of PCR have been sequenced, these should be assembled to produce a contiguous sequence (contig) of DNA. A contig has two advantages; (i) it confirms the order of nucleotides bases, and (ii) it has a sequence length closer to that of the sequenced product.

Sequences should be confirmed they match the putative target organism, rather than a contaminant or something unexpected. Sequence contigs, or one of the sequence products from the forward or reverse primer, can be identified with a nucleotide BLAST search against a sequence database, e.g., GenBank. A BLAST match will provide a tentative identification and help determine whether to proceed with analysis of the sequence data.

6.7 HIGH-THROUGHPUT SEQUENCING

High-throughput sequencing is a collective term for a number of new sequencing technologies such as Illumina, PacBio, Ion Torrent and NanoPore. Each method has its advantages and disadvantages, and factors such as price, read length, sequencing accuracy, and the quality and quantity of nucleic acid needed for sequencing must be considered. Most people outsource high-throughput sequencing to commercial sequencing service providers, although the MinION (NanoPore sequencing) is affordable enough to be done in-house. High-throughput sequencing delivers very large sequence datasets, which must be edited, assembled and interpreted. Advanced bioinformatic skillsets are required, and often high-performance computers are needed to handle data. Many commercial sequencing providers offer a bioinformatics service.

Frequently asked question: **Which high-throughput sequence technology should I use?**

Answer: **Each technique has its advantages and disadvantages, and the choice of technology depends on individual sequencing needs. Ask colleagues, contact local suppliers and read the recent literature.**

Frequently asked question: **Can bioinformatics be done on a home computer?**

Answer: **Smaller datasets less than 2 gigabytes of sequencing data can be analysed on laptop or desktop computers. Bacterial genomes and smaller fungal genomes less than 70 million base pairs can be assembled on personal computers, which may also be adequate for downstream analyses, such as annotating genomes, BLASTing genomes and comparative genomics. Larger datasets more than 2 gigabytes of data are best analysed using high performance computing.**

6.8 PHYLOGENETIC METHODS

Taxonomy should reflect phylogeny, whereby organisms are grouped according to their evolutionary history (phylogeny).

The connection between taxonomy and phylogeny explains why Linnaean names are shuffled and changed as our understanding of evolutionary history advances with new discoveries and new sequence data.

One of the most iconic images in biology is a bifurcating diagram that depicts how a group of taxa are related. Evolutionary relationships have since been visualised by phylogenetic trees, which are diagrams that represent a hypothesis of the phylogeny.

Phylogenetic species recognition methods are widely used to delimit species of microbial plant pathogens, particularly fungi and nematodes, by their relationships to closely related taxa/sequences in a phylogenetic tree. Phylogenetic species are recognised as monophyletic groups that are either (i) concordantly supported by most genes/loci (genealogical concordance), (ii) supported by at least one gene/region and not contradicted by any other genes/regions, or (iii) recovered by concatenation of two or more genes. Whether genes should be compared separately or concatenated is an ongoing debate in the phylogenetic community.

Phylogenetic trees can be used to identify a taxon, determine whether taxa are monophyletic, show common ancestry of taxa and determine how long ago taxa diverged. The root of a phylogenetic tree represents the ancestral lineage, and the tips of the branches represent the descendants of that ancestor. As one moves from the root of the tree to the tips (or crown), one is essentially moving forward in geological time. Phylogenetic trees may be constructed from comparison of any characteristics of the species or groups studied; however, DNA sequence data are most used. The steps to build a phylogenetic tree are (i) taxon selection, (ii) sequence alignment, (iii) phylogenetic analysis of the sequence data, and (iv) visualisation of the phylogenetic tree.

Frequently asked question: **Who first came up with the idea of a species tree?**

Answer: **Charles Darwin sketched the first species tree in 1837, and later included a species tree as the only figure in his book *'On the Origin of Species'* (1859).**

Taxon selection

The selection of sequences should reflect the aim of the phylogeny. For example, if the phylogeny is to reflect the diversity within a taxon, then it must be based on more than a single specimen of that taxon. If the aim of the phylogeny is to show the identity at genus rank, it should include related genera and sampling from within the identified genus. Or, if the aim is to show the species identity, the sampling should include, ideally, a sequence from the type and sequences from closely related taxa.

Selected sequences must be homologous, i.e., they must represent the same gene/locus/region in the genome. A starting point is to find near-identical sequences in GenBank using BLAST. The input for BLAST searches is either a FASTA file or a GenBank accession number. A FASTA file has a text-based format that represents nucleotide sequences using single-letter codes for base pairs. The FASTA format begins with a description line for sequence names and comments immediately after a > sign, with the sequence following on the next line. FASTA files can be saved in a text editor such as 'Notepad++' on PC or BBEdit on Mac.

Example of a FASTA file

>NR_138392.1 Uromycladium falcatarium BRIP 57477 ITS region; from TYPE material
TGTGTGTGGCTTTACTTTTTTCATAAAGTTTTTCTTTTCTTCTTTCACCCCCCACCTTAATTTTGGC
TCCTTGACGTGAGACATCCCTCACAGTCATTGAGCAGAAGCTCTGGAGAAGAGAAATGTGCCTACA
CAATCTCTTTTTGTGTAGGGTGATATCTCTCCCAGAGCATTAACAACCCCCCTAAAACCATGTTTTT
TTATATGTTTGTTTTTTTGAATATCATTACAATTGTGGAAAATAACTTTTAACAATGGATCTCTTGG
CTCTCACATCGATGAAGAACACAGTGAAATGTGATAAGTAATGTGAATTGCAGAATTCAGTGAATCA
TTGAATCTTTGAACGCACCTTGTGCCTCTTGGTATCCCAAGAGGCACACCTGTTTGAATGTCATGA
AAAAAATCTCTCAAACACATTTTTTTTTATATATTTACCATCTTATGTAAGATGGAAAAAATTTAAA
AAACAAAATGTGTTTGGATGATGAGTGCCTGCTGTGATGTTCAGCTCACTTTGAATAAATAAGCTC
TTCAGTGTTATAGCTGAGTTTGGAGTACAAAGGACTTGGTGTGATATAACAAAAGAAATGTCTTCAT
CAAGGAGTGCATGACCTTATTTAACTATGAGGAATGCTGCCAAGCTGTTTGAAAGTTGTACACATAT
GAGGGCTTTCTAAACTTTGTATATATATATATATATATCCC

Homologous sequences from type specimens (these usually begin with NR_ on GenBank) should be included where possible in taxon selection. Sequences published in taxonomic papers that report on the group under investigation should be included as these are often verified by taxonomists. Sequences should be avoided if from environmental and metagenomic studies or if specimens have not been lodged in a reference collection, as these cannot be verified.

Frequently asked question: **What is the meaning of the BLAST acronym?**

Answer: **BLAST stands for Basic Local Alignment Search Tool (BLAST). There are three BLAST algorithms that can compare nucleotides (megablast, discontiguous megablast and blastn) freely available at the National Center for Biotechnology Information. Megablast is best for comparing closely related sequences if the target per cent identity is 95% or more.**

Frequently asked question: **For what does FASTA stand?**

Answer: **FASTA stands for 'fast all' and is pronounced 'fast A'.**

Sequence alignment

Sequences must be aligned in a multiple sequence alignment (MSA), as a fundamental assumption is that characters are homologous in a phylogenetic analysis. Homologous positions are aligned by inserting gaps of varying length within sequences when compared to each other. In an evolutionary context, these gaps represent insertions and deletions (indels) within the genome that may have occurred during evolution from a common ancestor. An alignment will reveal point mutations or single-nucleotide polymorphisms (SNPs), which are the substitution of a single nucleotide at a specific position in the genome. SNPs, as well as indels, can be phylogenetically informative.

The manual curation of sequences is applicable for small datasets but is

impractical and less reproducible for larger datasets, i.e., those that have many long sequences. Increasingly, computational methods are used to generate MSAs. There are several computer programs that will generate MSAs from FASTA formatted sequence data.

Alignments of different genes/regions may be concatenated (joined end to end). An advantage of concatenation is that more data are used in the phylogenetic analysis. A potential disadvantage is that concatenated genes/regions with different histories of evolution may produce less accurate phylogenetic trees. There are several methods for reconstructing phylogenetic trees that can be used to infer evolutionary relationships.

Frequently asked question: **What alignment programs will generate MSAs from FASTA files?**

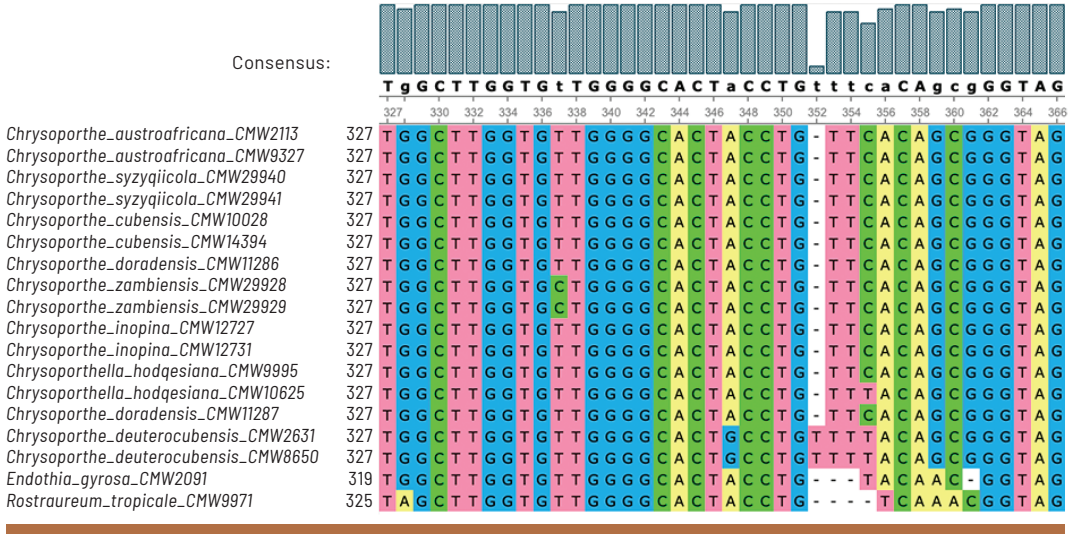
Answer: **ClustalW, Kalign, MAFFT, MUSCLE, T-Coffee, PROBCONS.**

Frequently asked question: **How can ambiguously aligned regions and indels be removed from MSAs?**

Answer: **By visual inspection or with a computer program, e.g. Gblocks.**

Frequently asked question: **Why may gene trees and species trees differ for a group of organisms?**

Answer: **It is common that different genes/loci in a group of organisms will have different gene tree topologies. One reason for this is that different genes may have different rates of evolution. Other reasons are incomplete lineage sorting (in which there may be different selection pressures on a gene in different populations), hybridisation, gene duplication and horizontal gene transfer.**



Nucleotide alignment of the ITS region of rDNA visualised with UGene software

Frequently asked question: **What methods are used for species delimitation in bacteria?**

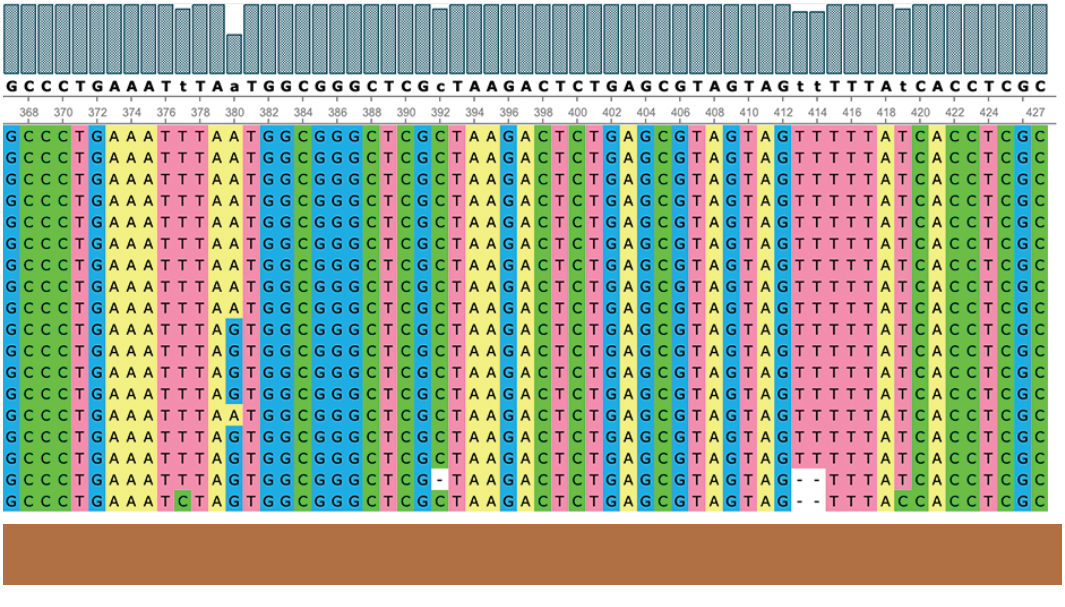
Answer: **Average nucleotide identity analysis of genomes is used to verify taxonomic identities in bacteria with a value of about 96% used to define species boundaries. A phylogenetic approach to recognise species of bacteria has gained traction as more bacterial genomic data are made public.**

Phylogenetic analyses

The set of all possible tree topologies is referred to as tree-space. Phylogenetic tree-space describes all the possible evolutionary relationships among taxa. The number of phylogenetic tree topologies can increase exponentially as more tips (sequences) are added. More precisely, the number of tree topologies will increase exponentially as more nodes (common ancestors) are generated. A computational challenge is to find trees that best match an alignment. The complexity increases when models of evolution are applied to account for different rates of change/substitutions in nucleotides or amino acids.

Distance-based methods are not computationally expensive because they compute pairwise distances among sequences and put tips in relationships based on how much they differ. Distance-based methods, such as neighbour joining and UPGMA (unweighted pair group method with arithmetic mean), are useful to show relationships among populations and closely related taxa. They are less well-regarded for hypotheses of evolutionary relationships above species rank because they do not use a model of evolution.

Other phylogenetic methods search tree-space to find trees that best match an alignment. These methods include parsimony, maximum likelihood and Bayesian



inference. Parsimony treats all data, whether from DNA, presence/absence data, or morphology in the same way, and does not use a model of evolution. Statistical methods that analyse tree-space are based on assumptions about DNA sequence evolution and a model of evolution.

The best nucleotide substitution models for a particular MSA can be chosen from comparison via likelihood ratio tests (LRTs) or information criteria, e.g., the Akaike information criterion, the Bayesian information criterion. This step of selecting a model of evolution is often unnecessary as most models will produce trees with similar topologies. The generalised time-reversible model (GTR +I +G) accounts for all possible changes among nucleotides at different rates. It usually fits real data better than other models and has been implemented in most software designed for phylogenetic reconstructions.

Relationships recovered in phylogenetic trees can be assessed by statistical tests of support. NJ (neighbour joining), parsimony, and ML (maximum likelihood) use bootstrap tests that re-sample the alignment to make artificial datasets based on the original. A tree is searched in each of these re-sampled alignments and a bootstrap value reflects the number of times a node was recovered in the artificial datasets. Relationships are supported in BI (Bayesian inference) by the probability that a node/branch was sampled in a search of tree-space. Other measures of support include approximate likelihood ratio tests, ultrafast bootstraps and genealogical concordance factors.

(i) Neighbour joining

Neighbour joining (NJ) is a distance-based method that assumes the tree with the smallest sum of branch length estimates is most likely to be the true one. It uses a distance matrix obtained from the MSA that records the number of differences in bases at homologous positions for two equal-length sequences from different specimens. NJ is fast, which makes it useful for analysing large datasets. NJ is useful for checking the initial MSA, as long branch lengths may reveal if alignments have shifted. Disadvantages of NJ are that (i) it only produces one tree and neglects other possible trees in tree-space, and (ii) errors in distance estimates are exponentially large for long distance, which will yield a biased tree. NJ has been superseded by ML and BI methods that offer superior accuracy to test hypotheses of evolution from molecular data.

(ii) Maximum likelihood

Maximum likelihood (ML) returns the most likely phylogenetic tree based on probabilities assigned to nucleotide substitutions (point mutations) in MSAs (rather than just counting them as in NJ). ML assigns probabilities for nucleotide substitutions based on assumptions, e.g., that substitutions between bases with similar chemistry (transitions) occur more often than substitutions between bases with different chemistry (transversions), or that the rate of evolution differs across sites in the MSA. In ML analysis it is not necessary to know the actual values of these probabilities, as they are varied in the tree evaluation process. ML compares all possible phylogenetic trees that predict the observed data to find the one that has the highest

probability of producing the observed sequences.

ML reconstructs ancestors at all nodes of each considered tree and assigns branch lengths based on the probabilities of mutations. ML is recognised as a consistent and powerful basis for phylogenetic, and thereby taxonomic, inference. A disadvantage of ML is that large datasets (>100 sequences) require considerable computational time to search trees. Maximum likelihood estimation may be subject to systematic errors if the model of evolution used to evaluate the likelihood of given trees does not reflect the actual evolutionary processes.

Frequently asked question: **What computer programs should I use to generate ML phylogenetic trees?**

Answer: **IQTree, PhyML, RAxML and FastTree (for datasets >1000 sequences).**

Frequently asked question: **What is the bootstrap method and what are bootstrap values in ML phylogenetic analysis?**

Answer: **Given an MSA, bootstraps replicate MSAs (also called bootstrap pseudoreplicates) by sampling columns uniformly at random with replacement. The tree is rebuilt from each replicate MSA and tested to see if the same branches were recovered. The bootstrap value indicates the percentage of times (usually out of a 100 or 1,000) that the same branch was recovered.**

Frequently asked question: **What does a bootstrap value of 100% mean?**

Answer: **A bootstrap value of 100% means that a branch showed up in all trees generated from the bootstrap replicates. Generally a bootstrap value ≥70% is considered a well-supported relationship. Ultrafast bootstraps are suitable for large datasets, with shortcuts that reduce time in the analysis. Ultrafast bootstraps ≥95% are considered well-supported.**

(iii) Bayesian inference

Bayesian inference (BI) examines phylogenetic relationships among species from MSAs. BI uses probability distributions that describe the uncertainties around trees and MSAs. BI is a powerful technique, as the calculated posterior probability distribution is used as the prior probability distribution in another Bayesian calculation. Critical to the success of BI is the use of Markov chain Monte Carlo (MCMC) algorithms to estimate posterior probability distributions. A disadvantage of BI is that users tend to use default settings in their computer analyses because of the inherent complexity of the method.

Frequently asked question: **What is Bayes' theorem?**

Answer: $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A) \times P(B|A)}{P(B)}$

where P(A) is the probability of event A; P(B) is the probability of event B; P(A|B) is the probability of event A given B has occurred; P(B|A) is the probability of event B given A has occurred; and P(A∩B) is the probability of both A and B occurring.

Frequently asked question: How is Bayes' theorem used in BI for phylogenies?

Answer:

$$P(\text{tree}|\text{data}) = \frac{P(\text{tree}) \times P(\text{data}|\text{tree})}{P(\text{data})}$$

BI uses Bayes' theorem to estimate the posterior probability of trees P(tree|data), which is the probability that the tree is correct given the data. Inputs are the prior probability of a tree P(tree) based on knowledge before the analysis of the data, and the likelihood of the data P(data|tree). P(data) is a normalising constant, sometimes called the evidence. Bayes' theorem may be expressed as:

$$\text{posterior} = \frac{\text{prior} \times \text{likelihood}}{\text{evidence}}$$

Frequently asked question: What software should I use to generate BI phylogenetic trees?

Answer: MrBayes, BEAST2.

Frequently asked question: What is the role of Markov chain Monte Carlo (MCMC) in BI?

Answer: MCMC refers to algorithms that sample from a known probability distribution. MCMC estimates the posterior probability distribution from the prior probability distribution and likelihood. MCMC algorithms are sensitive to their starting point, and often require a warm-up phase or burn-in phase, after which prior samples can be discarded and useful samples can be collected.

Frequently asked question: What is the origin of the name Markov chain Monte Carlo?

Answer: Monte Carlo is named for the casino in Monaco. Monte Carlo methods estimate a distribution by random sampling. A Markov chain takes the name from Andrey Markov (1856–1922), who was a Russian mathematician. In a Markov chain, the probability of the next computed estimated outcome depends only on the current estimate and not on prior estimates.

Frequently asked question: Which method is better, ML or BI?

Answer: Different phylogenetic criteria are better suited to different datasets, research questions or computational resources. ML and BI analyses are both suitable to recover hypotheses of evolution from molecular data, and, with the same models of evolution, will mostly recover the same topology. Both BI and ML analyses are computationally expensive and require time to create reliable trees.

(iv) Coalescent/Supertree methods

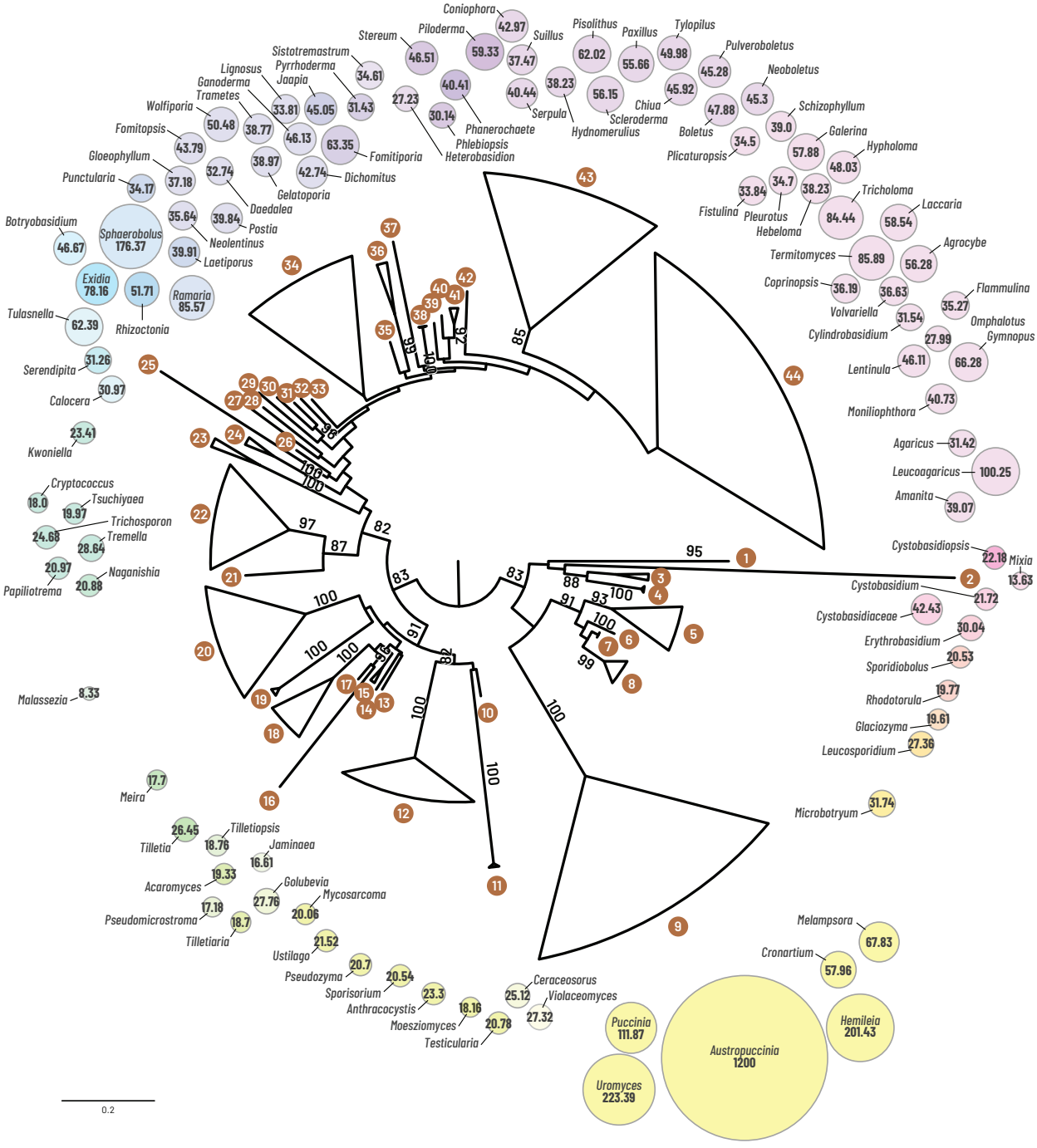
Coalescent-based species delimitation (CBD) methods account for different evolutionary histories in different genes/loci as a result of deep coalescence from random genetic drift in branches of the species tree. CBD accounts for differences in gene trees that may result from biological processes, including incomplete lineage sorting, horizontal gene transfer, hybridisation, recombination and gene duplication or loss. CBD is different from sequence-based methods, which assume all genes have the same evolutionary history across concatenated genes/regions. CBD infers a species tree from differing topologies of gene trees. Computer programs for CBD phylogenetic analyses, such as ASTRAL, summarise gene trees from any phylogenetic criteria to produce a consensus species tree.

(v) Phylogenies as trees or networks

Phylogenies are often represented as stick diagrams that may be trees or networks. Both trees and networks have branches (called edges in networks) and nodes. Branches represent evolutionary change over time. Branches may be drawn as straight diagonal lines, kinked to make a rectangular tree, or curved to make a circular tree. The branch length is measured in nucleotide changes per sequence site. The longer a branch, the greater the genetic change.

A node in a tree or network is where branches meet. A node represents a common ancestor of descendants. The most recent common ancestor is the youngest common ancestor shared by a group of descendants. Internal branches connect two nodes, whereas external branches connect a tip and a node. The tips of branches usually represent extant species.

Trees are networks that have nodes with only one most recent common ancestor. Networks have nodes that may have two (or more) common ancestors. Consequently, networks may better express evolutionary history when hybridisation, horizontal gene transfer, recombination, or gene duplication and loss are involved. If there are incongruences in the topological variation in gene trees, then a network may better represent the phylogeny.



Phylogram of taxa with sequenced genomes in *Basidiomycota* based on the LSU region rDNA. Average genome size for genera of *Basidiomycota* is visualised by bubbles at tip ends.

1. *Cystobasidiopsis lactophilus*
2. *Mixia osmundae*
3. *Cystobasidium pallidum*
   Cystobasidiaceae sp.
4. *Erythrobasidium hasegawianum*
   Erythrobasidium yunnanense
   Sporidiobolus salmonicolor
5. *Rhodotorula kratochvilovae*
   Rhodotorula graminis
   Rhodotorula toruloides
   Rhodotorula taiwanensis
   Rhodotorula sp.
   Rhodotorula mucilaginosa
6. *Glaciozyma antarctica*
7. *Leucosporidium creatinivorum*
   Leucosporidium scottii
8. *Microbotryum intermedium*
   Microbotryum lychnidis-dioicae
   Microbotryum silenes-dioicae
   Microbotryum saponariae
   Microbotryum violaceum
9. *Melampsora pinitorqua*
   Melampsora abietis-canadensis
   Melampsora occidentalis
   Melampsora medusae
   Melampsora laricis-populina
   Melampsora aecidioides
   Melampsora allii-populina
   Cronartium ribicola
   Cronartium harknessii
   Cronartium quercuum
   Hemileia vastatrix
   Austropuccinia psidii
   Puccinia graminis
   Uromyces transversalis
   Puccinia coronata
   Puccinia triticina
   Uromyces viciae-fabae
   Puccinia striiformis
   Puccinia sorghi
   Puccinia horiana
10. *Violaceomyces palustris*
11. *Ceraceosorus guamensis*
    Ceraceosorus bombacis
12. *Testicularia cyperi*
    Ustilago esculenta
    Pseudozyma aphidis
    Pseudozyma tsukubaensis
    Meosziomyces antarcticus
    Anthracocystis flocculosa
    Kalmanozyma brasiliensis
    Sporisorium reilianum
    Sporisorium scitamineum
    Ustilago vetiveriae
    Pseudozyma hubeiensis
    Ustilago trichophora
    Ustilago xerochloae
    Ustilago cynodontis
    Ustilago tritici
    Mycosarcoma maydis
    Ustilago hordei
    Ustilago bromivora
13. *Golubevia pallescens*
14. *Tilletiaria anomala*
15. *Pseudomicrostroma glucosiphilum*
    Jaminaea rosea
16. *Acaromyces ingoldii*
17. *Tilletiopsis washingtonensis*
18. *Tilletia buchloeana*
    Tilletia horrida
    Tilletia walkeri
    Tilletia indica
    Tilletia caries
    Tilletia controversa
19. *Meira miltonrushii*
    Meira nashicola
20. *Malassezia cuniculi*
    Malassezia obtusa
    Malassezia furfur
    Malassezia japonica
    Malassezia yamatoensis
    Malassezia vespertilionis
    Malassezia slooffiae
    Malassezia pachydermatis
    Malassezia restricta
    Malassezia globosa
    Malassezia nana
    Malassezia sympodialis
    Malassezia caprae
    Malassezia equina
    Malassezia dermatis
21. *Naganishia albida*
22. *Papiliotrema flavescens*
    Papiliotrema laurentii
    Tremella mesenterica
    Trichosporon asahii
    Cryptococcus amylolentus
    Tsuchiyaea wingfieldii
    Cryptococcus depauperatus
    Cryptococcus neoformans var. grubii
    Cryptococcus gattii
    Kwoniella heveanensis
    Kwoniella dejecticola
    Kwoniella pini
    Kwoniella bestiolae
    Kwoniella mangrovensis
23. *Calocera viscosa*
    Calocera cornea
24. *Sebacina vermifera*
    Serendipita indica
25. *Tulasnella calospora*
26. *Exidia glandulosa*
27. *Botryobasidium botryosum*
28. *Rhizoctonia solani*
29. *Sphaerobolus stellatus*
30. *Ramaria rubripermanens*
31. *Turbinellus floccosus*
32. *Laetiporus sulphureus*
33. *Punctularia strigosozonata*
34. *Neolentinus lepideus*
    Gloeophyllum trabeum
    Postia placenta
    Daedalea quercina
    Fomitopsis pinicola
    Wolfiporia cocos
    Gelatoporia subvermispora
    Trametes versicolor
    Lignosus rhinocerotis
    Dichomitus squalens
    Ganoderma lucidum
    Ganoderma sinense
35. *Jaapia argillace*
36. *Fomitiporia mediterranea*
    Pyrrhoderma noxium
37. *Sistotremastrum niveocremeum*
38. *Heterobasidion irregulare*
    Heterobasidion annosum
39. *Stereum hirsutum*
40. *Phlebiopsis gigantea*
41. *Phanerochaete chrysosporium*
    Phanerochaete carnosa
42. *Piloderma croceum*
43. *Coniophora puteana*
    Serpula lacrymans
    Suillus alpinus
    Suillus luteus
    Hydnomerulius pinastri
    Pisolithus tinctorius
    Scleroderma citrinum
    Paxillus rubicundulus
    Paxillus involutus
    Tylopilus plumbeoviolaceoides
    Chiua virens
    Pulveroboletus ravenelii
    Boletus magnificus
    Boletus bicolor
    Neoboletus brunneissimus
44. *Plicaturopsis crispa*
    Schizophyllum commune
    Fistulina hepatica
    Pleurotus ostreatus
    Galerina marginata
    Hypholoma sublateritium
    Hebeloma cylindrosporum
    Tricholoma terreum
    Termitomyces eurrhizus
    Laccaria bicolor
    Laccaria amethystina
    Coprinopsis cinerea
    Agrocybe cylindracea
    Volvariella volvacea
    Cylindrobasidium torrendii
    Flammulina velutipes
    Omphalotus olearius
    Lentinula edodes
    Gymnopus luxurians
    Moniliophthora perniciosa
    Moniliophthora roreri
    Amanita inopinata
    Amanita thiersii
    Agaricus bisporus
    Leucoagaricus gongylophorus
    Amanita jacksonii
    Amanita muscaria
    Amanita brunnescens
    Amanita polypyramis
    Amanita pseudoporphyria

Frequently asked question: What are posterior probability values in BI phylogenetic analysis?

Answer: Phylogenies from BI provide statistical support for any branch measured by a Bayesian posterior probability. Generally a posterior probability ≥0.95 is significant for a well-supported relationship.

Frequently asked question: What is a clade?

Answer: A clade is a group of organisms that includes a common ancestor and all its descendants.

Frequently asked question: How does a phylogram differ from a cladogram?

Answer: The lengths of branches in a phylogram represent evolutionary change over time, while the branches in cladograms are uniform in length and have no meaning.

Frequently asked question: What is the difference between a monophyletic, paraphyletic and polyphyletic taxon?

Answer: A monophyletic taxon includes the most recent common ancestor of a group of organisms, and all its descendants. A monophyletic taxon is represented as a clade in a phylogenetic tree. A single branch can be a clade. A paraphyletic taxon includes the most recent common ancestor, but not all its descendants. A polyphyletic taxon does not share a common ancestor. Polyphyletic taxa are often the result of convergent evolution, and have multiple ancestors.

Frequently asked question: What are sister groups (or species)?

Answer: Groups of descendants (including species) are considered sister groups if they shared a most recent common ancestor.

Frequently asked question: What is the difference between a split network and a reticulate network?

Answer: A split network represents all splits in two (or more) trees, with parallel edges each corresponding to a branch contained in one of the input trees. The nodes do not necessarily correspond to hypothetical ancestors in a split network. A reticulate network represents two (or more) trees by postulating reticulations that describe an evolutionary history, with internal nodes corresponding to ancestral taxa, and the edges representing evolutionary pathways.

Frequently asked question: What computer programs should I use to visualise phylogenetic trees and networks?

Answer: For trees try FigTree, Treeview and MEGAX, and for networks use SplitsTree or PopArt.

Frequently asked question: What are the steps in tree reconstruction?

Answer:

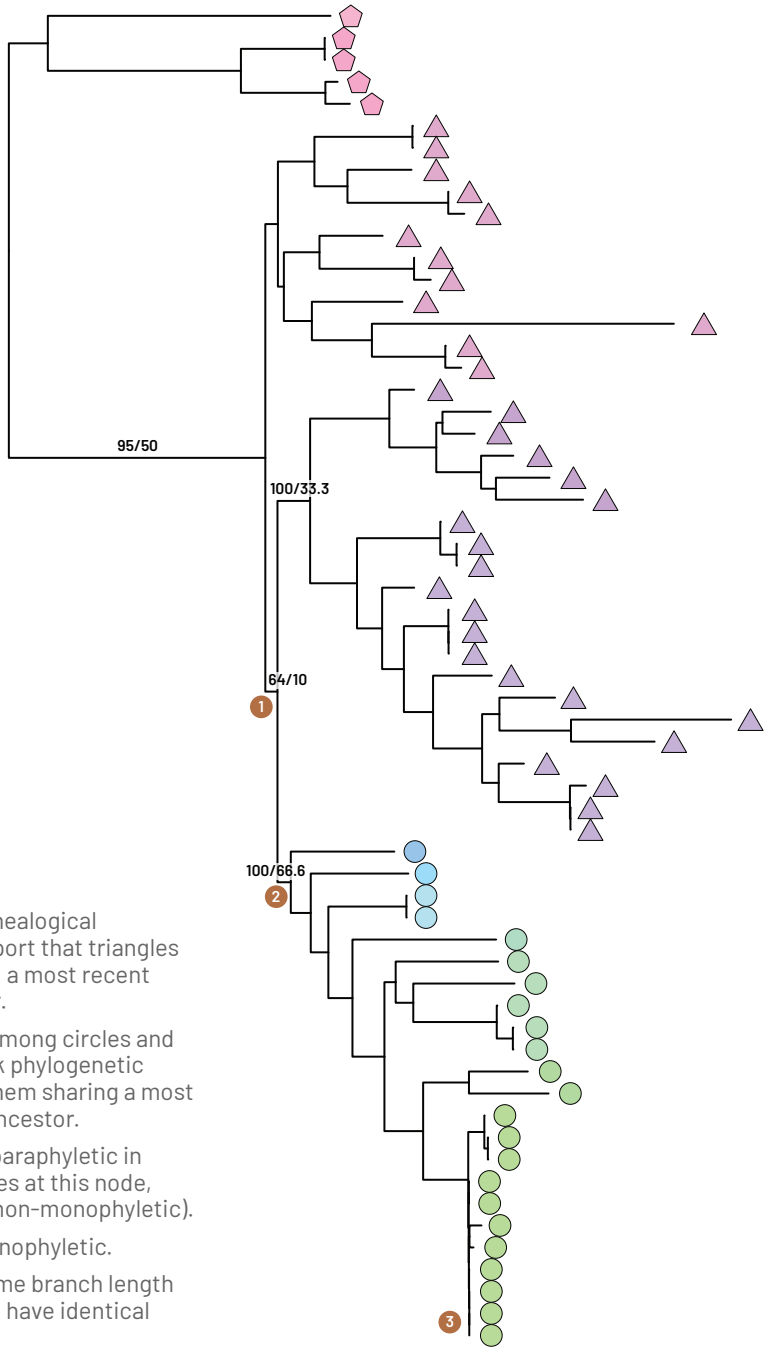
1. Taxon selection
2. Alignment
3. Concatenate and partition genes
4. Phylogenetic analysis
5. Visualisation

Frequently asked question: What is long branch attraction?

Answer: Long branch attraction in phylogenetic analyses occurs when two or more lineages are inferred to be closely related, regardless of their true evolutionary relationships. The problem arises when lineages evolve rapidly.

Frequently asked question: What do apomorphy and synapomorphy mean?

Answer: An apomorphy is a trait shared exclusively by a group or species and not present in an ancestor. A synapomorphy is a trait shared exclusively by a clade (which includes the most recent common ancestor). A synapomorphy distinguishes a clade from other organisms and provides evidence for common ancestry.



Bootstrap and genealogical concordance support that triangles and circles shared a most recent common ancestor.

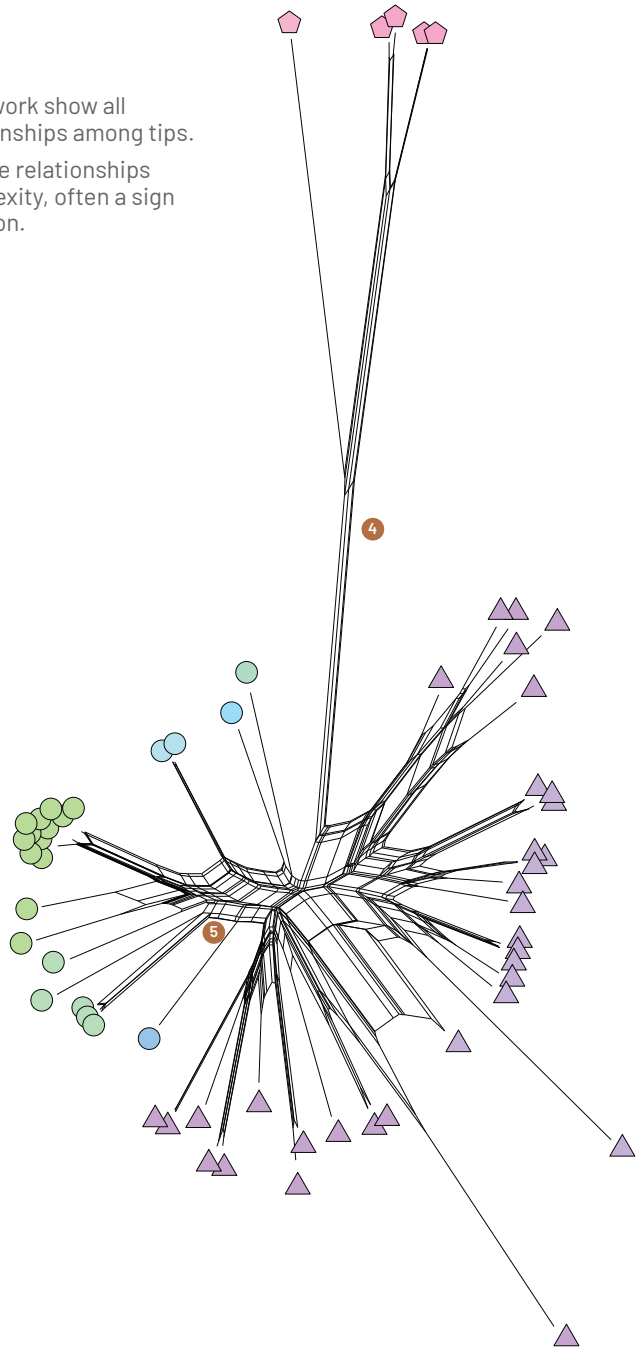
The relationship among circles and triangles has weak phylogenetic support despite them sharing a most recent common ancestor.

1. Triangles are paraphyletic in regard to circles at this node, (triangles are non-monophyletic).
2. Circles are monophyletic.
3. Tips on the same branch length in a phylogram have identical sequences.

Phylogram to illustrate monophyletic, paraphyletic, and polyphyletic relationships

4 Edges of a network show all possible relationships among tips.

5 Highly reticulate relationships indicate complexity, often a sign of recombination.



SplitsTree network based on the same dataset as used on page 66 to illustrate non-dichotomous evolutionary relationships

Appendix 1

REFERENCES

Ahmed, M., Sapp, M., Prior, T., Karssen, G., & Back, M. (2015). Nematode taxonomy: From morphology to metabarcoding. *SOIL Discussions* 2: 1175–1220.

Bull, C. T., De Boer, S. H., Denny, T. P., Firrao, G., Fischer-Le Saux, M., Saddler, G. S., Scortichini, M., Stead, D. E., & Takikawa, Y. (2008). Demystifying the nomenclature of bacterial plant pathogens. *Journal of Plant Pathology* 90(3): 403–417.

Crous, P. W. (2005). Plant pathology is lost without taxonomy. *Outlooks on Pest Management* 16(3): 119–123.

Glawe, D. A. (2003). Much more than phylogenies: A utilitarian view of the taxonomy of plant pathogenic fungi. *APSnet Features*. Online.

Hyde, K. D., Chomnunti, P., Crous, P. W., Groenewald, J. Z., Damm, U., Ko Ko, T. W., Shivas, R. G., Summerell, B. A., & Tan, Y.P. (2010). A case for re-inventory of Australia's plant pathogens. *Persoonia* 25: 50–60.

Jones, R. A. C., Boonham, N., Adams, I. P., & Fox, A. (2020). Historical virus isolate collections: An invaluable resource connecting plant virology's pre-sequencing and post-sequencing eras. *Plant Pathology* 70(2): 235–248.

Rossman, A. Y., & Palm-Hernández, M. E. (2008). Systematics of plant pathogenic fungi: Why it matters. *Plant Disease* 92(10): 1376–1386.

van der Westhuizen, G. C. A. (1991). The importance of taxonomy in plant pathology. *South African Journal of Science* 87(1): 8–11.

Walker, J. (1975). Mutual responsibilities of taxonomic mycology and plant pathology. *Annual Review of Phytopathology* 13: 335–355.

Appendix 2

THE GRAM STAIN PROCEDURE

Crystal violet solution	
Crystal violet	2 g
Ethanol (95%)	20 mL
Ammonium oxalate	0.8 g
Distilled water	80 mL

Dissolve 2 g crystal violet in 20 mL 95% ethanol. Dissolve 0.8 g ammonium oxalate in 80 mL distilled water. Mix the two solutions.

Lugol's iodine solution	
Iodine	1 g
Potassium iodide (KI)	2 g
Distilled water	300 mL

Grind together the iodine and potassium iodide, and then dissolve in the water and stir in a closed container for several hours to complete dissolution.

Safranin counterstain	
Safranin O	2.5 g
Ethanol (95%)	20 mL

Prepare a stock solution of safranin counterstain by dissolving 2.5 g safranin O in 100 mL 95% ethanol. Dilute 1:10 in distilled water for use.

Make a turbid suspension of bacterial cells from an actively growing culture (usually 24–48 hours old) in sterile water. Smear a loopful of this suspension on a clean glass microscope slide. The smear should be about 1 cm^2 and just visible. Air dry the slide, then fix by passing the slide, bacterium up, several times through the flame of a Bunsen burner. Do not overheat the slide. Fixing ensures the bacteria will adhere to the slide surface during the staining procedure. Flood the slide with crystal violet for 1 min. Pour off excess fluid and wash in a gentle stream of tap water until no more stain is removed from the smear. Flood the slide with Lugol's iodine for 1 min. Wash in a gentle stream of tap water and blot dry. Wash in a gentle stream of 95% ethanol for not more than 30 sec to remove any stain and blot dry. Counterstain by flooding with safranin for 20 sec. Wash again with water and blot dry. Examine the slide under x1000 magnification with oil immersion. Gram-positive bacteria stain blueish violet, while Gram-negative stain reddish pink.