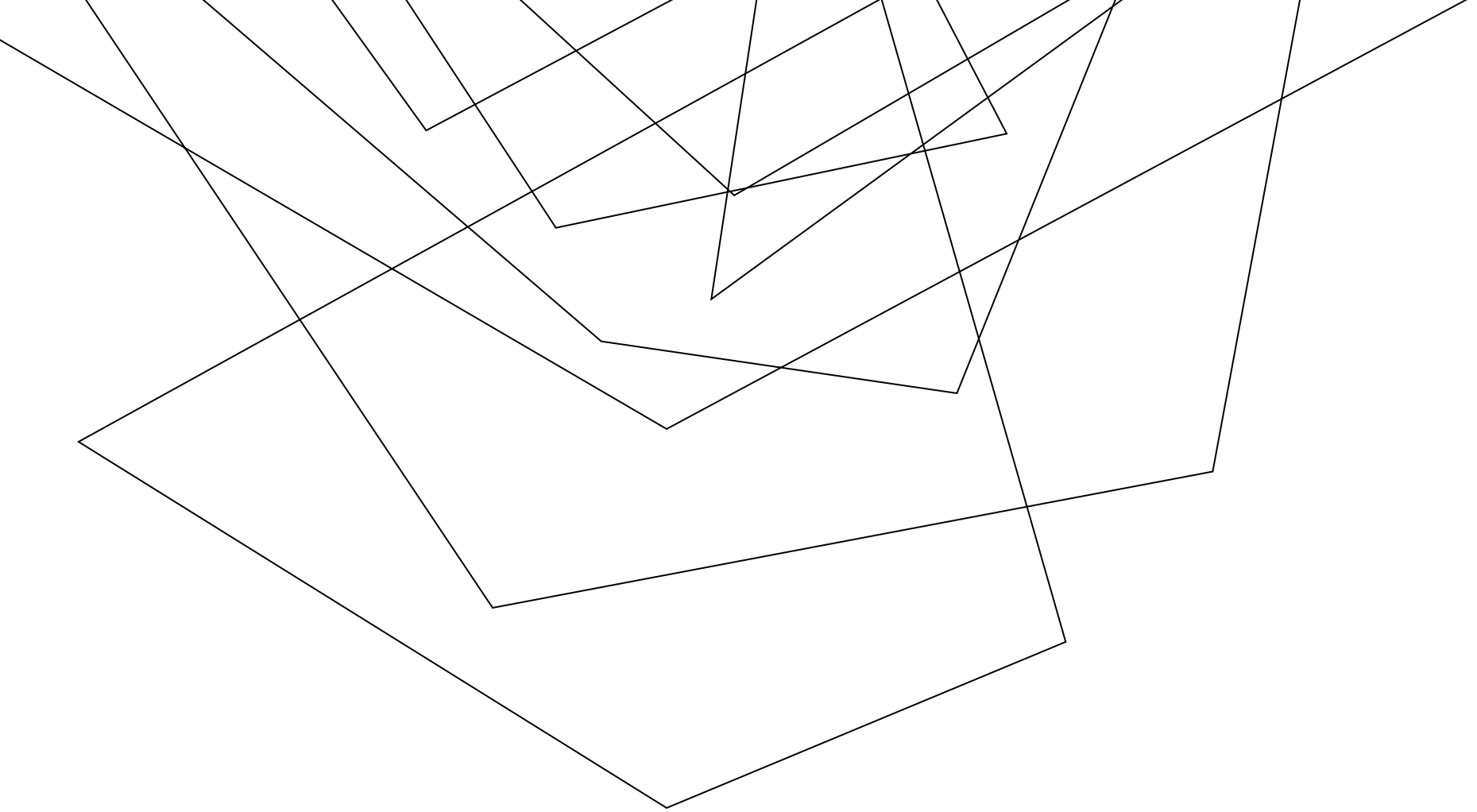


Royal Botanic Gardens Kew

DATA SCIENCE@WORK

Harnessing Botanical and Mycological Knowledge: Data-Driven Solutions for Global Challenges

Adam Richard-Bollans

DIGITISING PHYSICAL COLLECTIONS

The Herbarium and Fungarium at the Royal Botanic Gardens Kew houses approximately **8.25 million specimens**.

Kew is committed to Open Access Collections by digitising and building online systems to share images, datasets and metadata linked to specimens collected since the 18th century.

Kew is:-

- Digitising 8.25M plant and fungal specimens

- Using a 114 strong team of data staff, curation staff and software developers

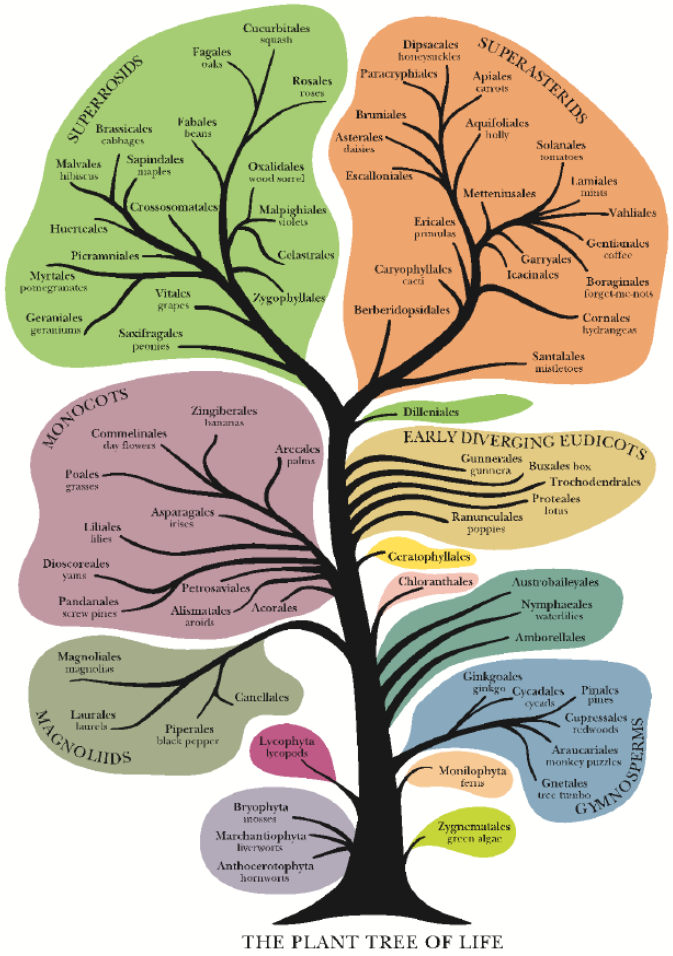
- ~70,000 specimens digitised each week

Herbarium specimens from Kew's Herbarium © RBG Kew

The Digitisation project in progress - a digitiser images a herbarium specimen © RBG Kew

GENERATING, CURATING AND PRESERVING KNOWLEDGE

Medicinal Plant Names Services

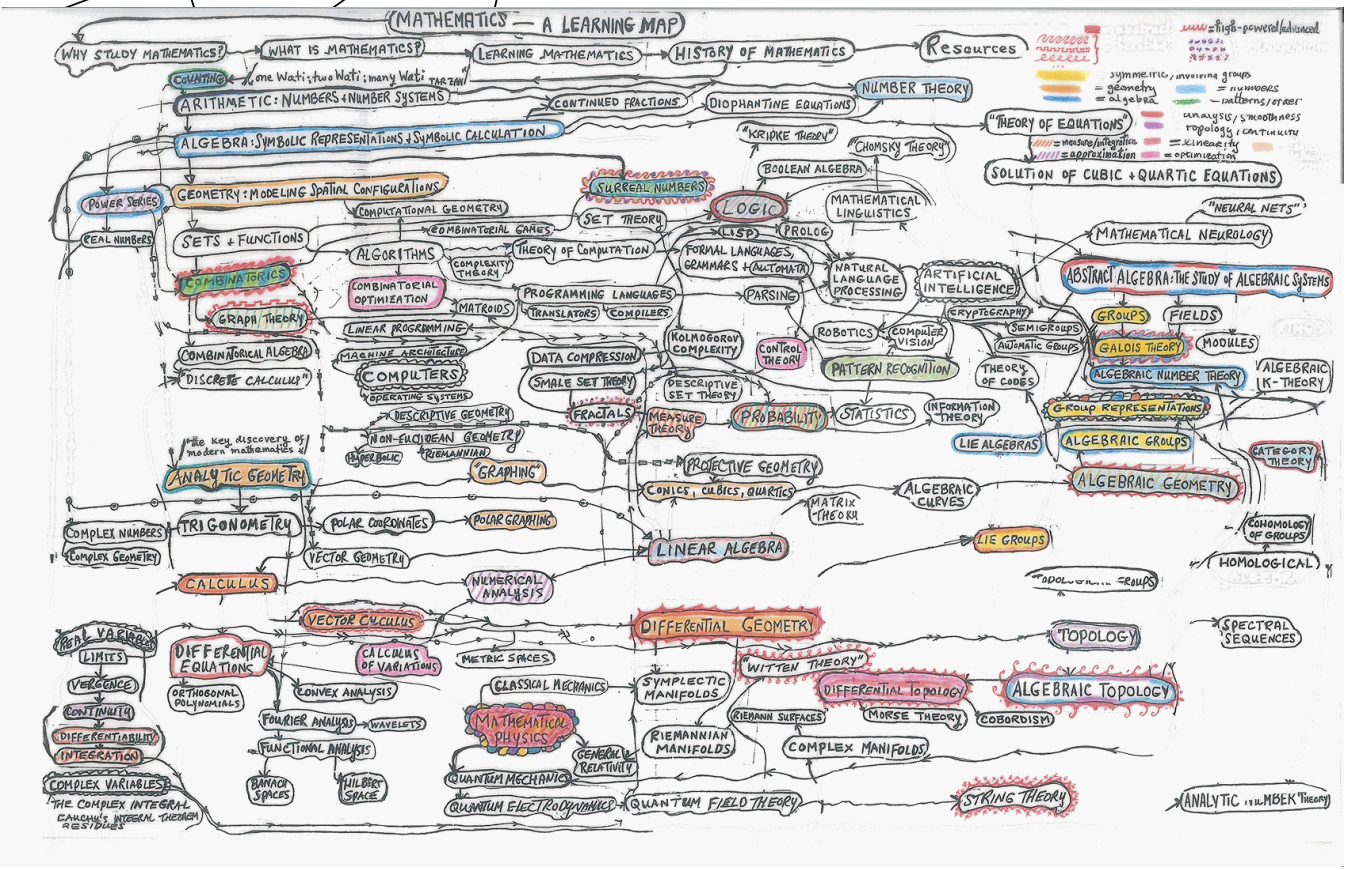


THE PLANT TREE OF LIFE

Royal Botanic Gardens Kew | Plants of the World Online

I · P · N · I

BACKGROUND



BSc, MSc Mathematics 2012/2016
University of Manchester/Leeds

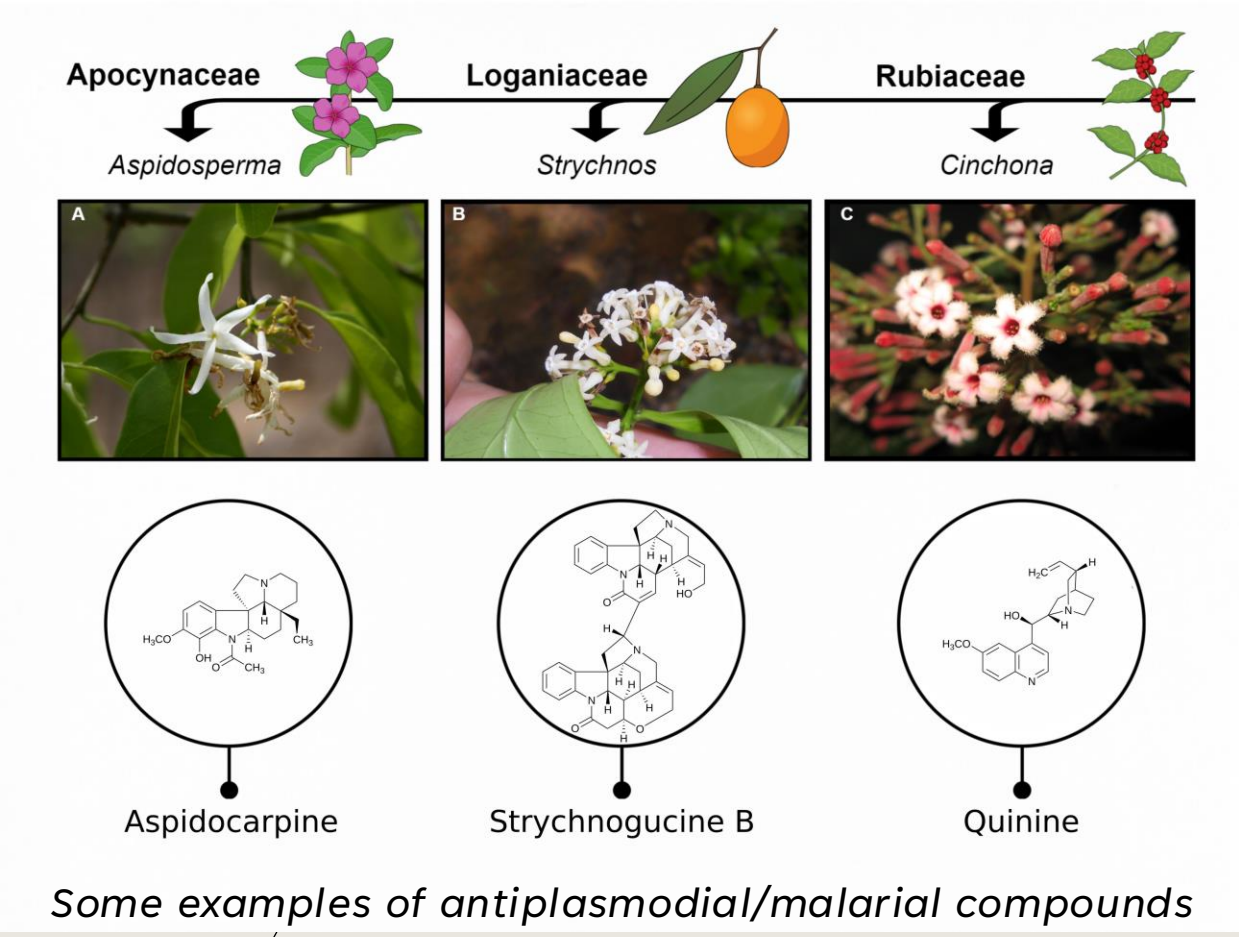
PhD Computing 2021
University of Leeds

Artificial Intelligence Researcher, 2021-22
RBG, Kew

Future Leader Fellow, 2022 - Present
RBG, Kew

FINDING NEW MEDICINAL COMPOUNDS

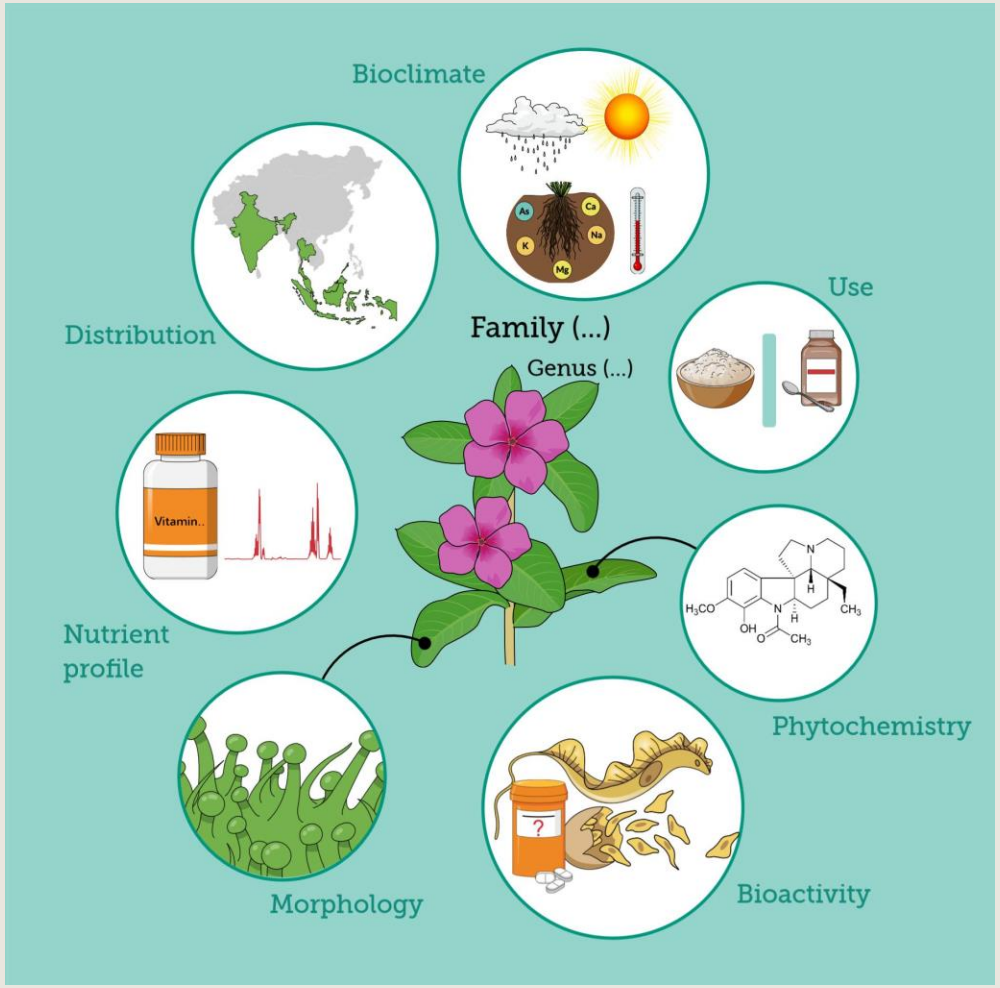
- Plants have provided or inspired the development of numerous pharmaceutical drugs
- An estimated 343,000 known vascular plant species that remain largely unexplored scientifically
- *Can machine learning models trained on plant trait data reliably predict the bioactivity of plants?*



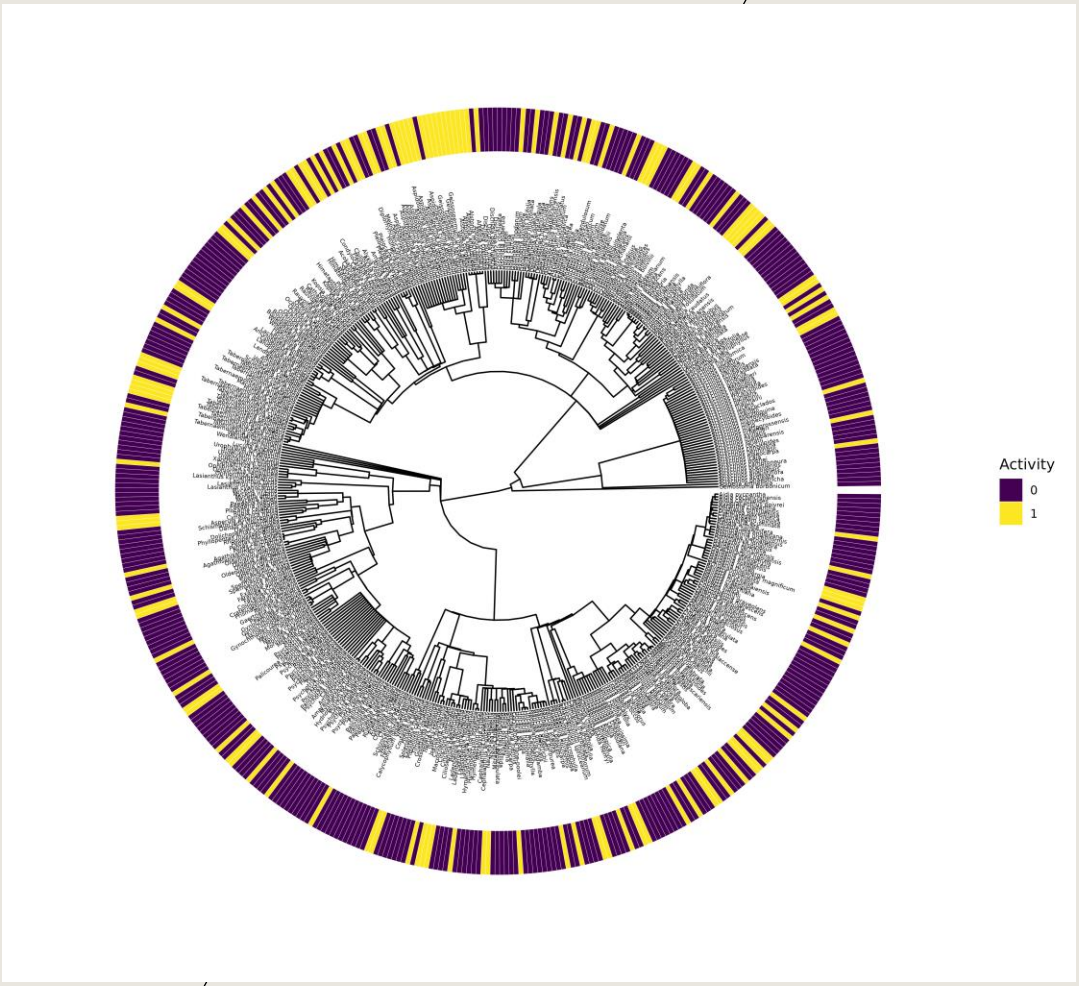
Some examples of antiplasmodial/malarial compounds

FINDING NEW MEDICINAL COMPOUNDS

Plant Traits

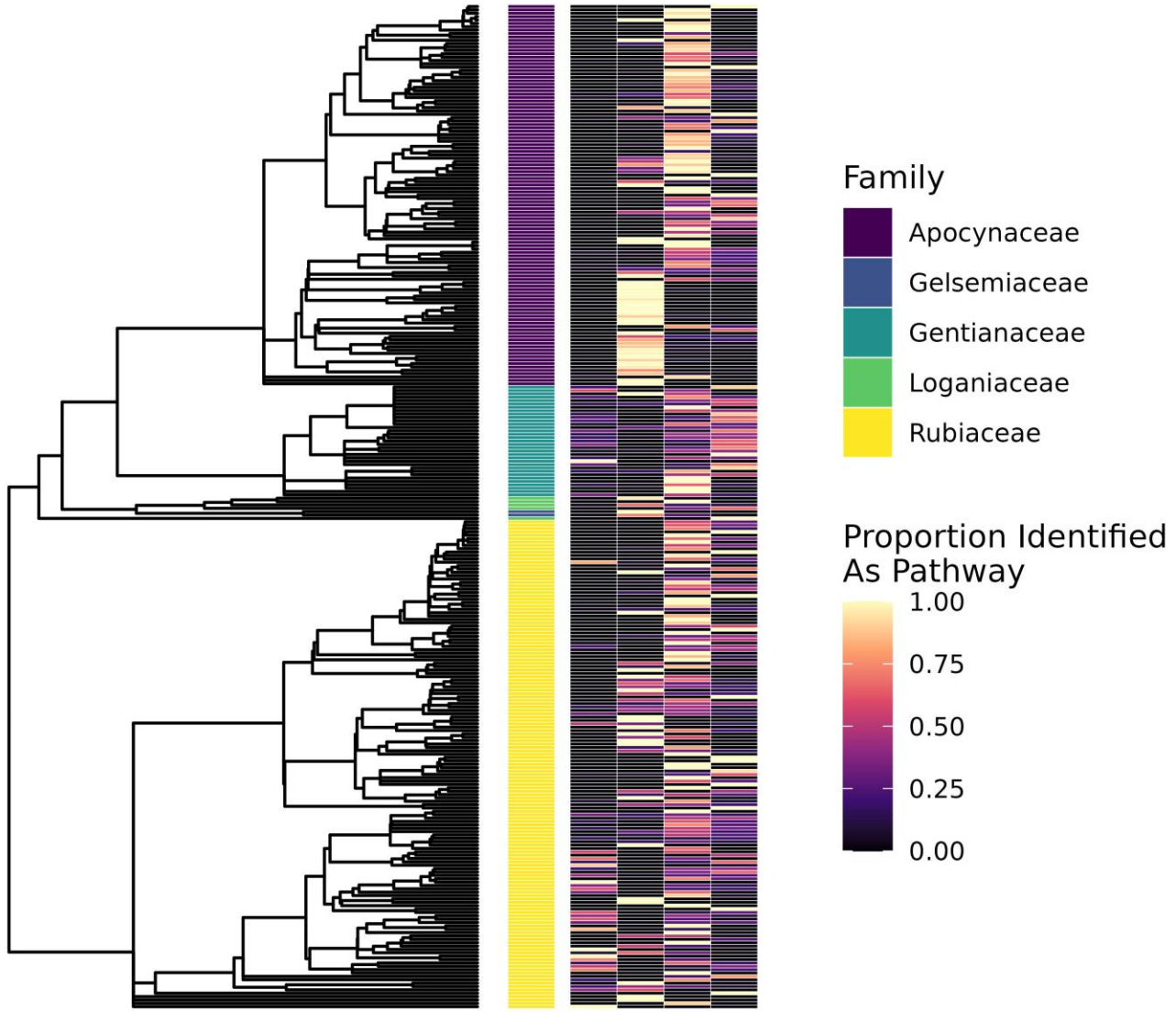


Phylogenetic Patterns



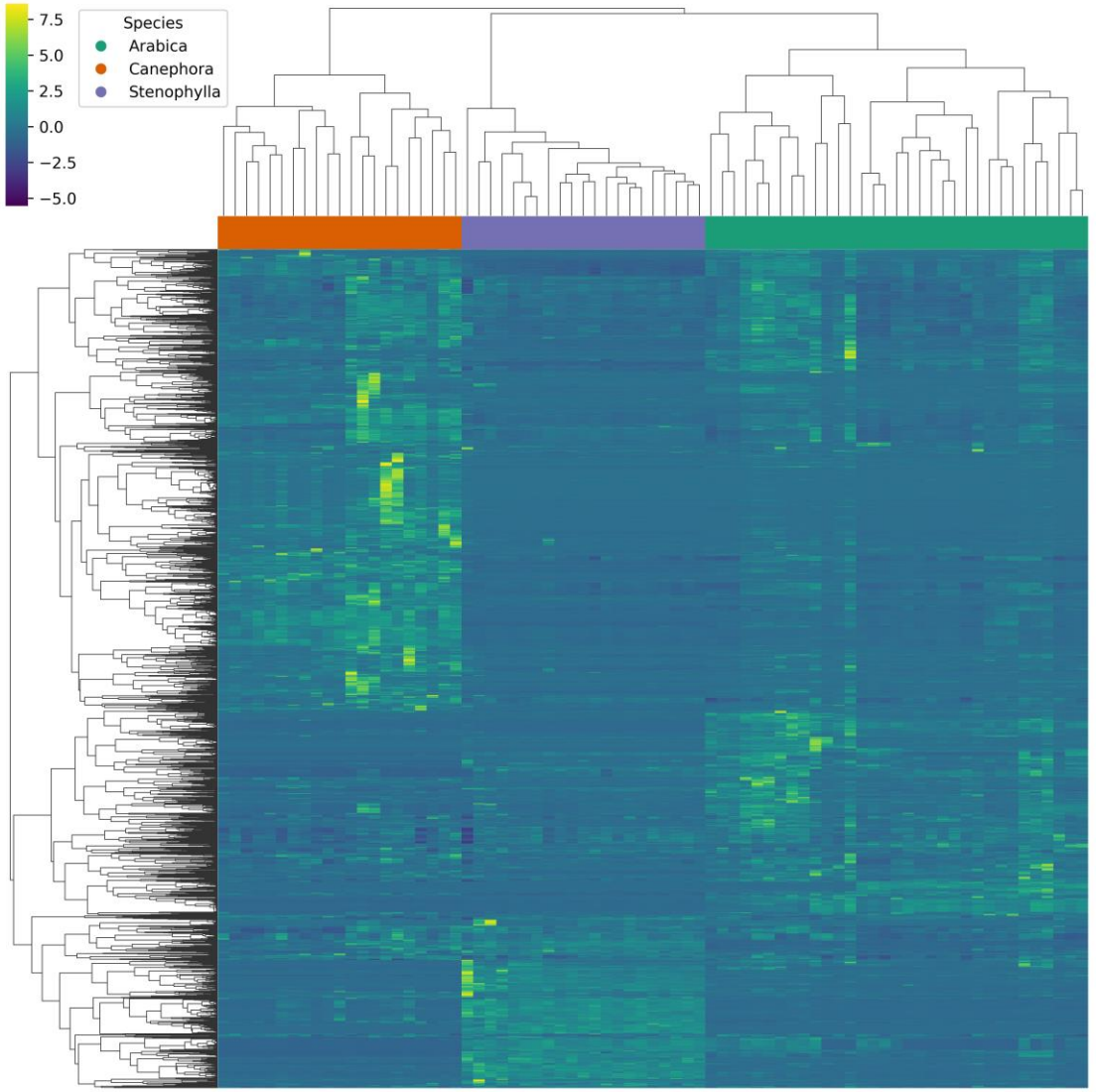
ANALYSING PHYTOCHEMICAL DIVERSITY

- *What are the evolutionary drivers of phytochemical diversity and how are different compound classes distributed across the tree of life?*
- Using large-scale datasets may bring better understanding to chemotaxonomy as well as helping to inform conservation efforts



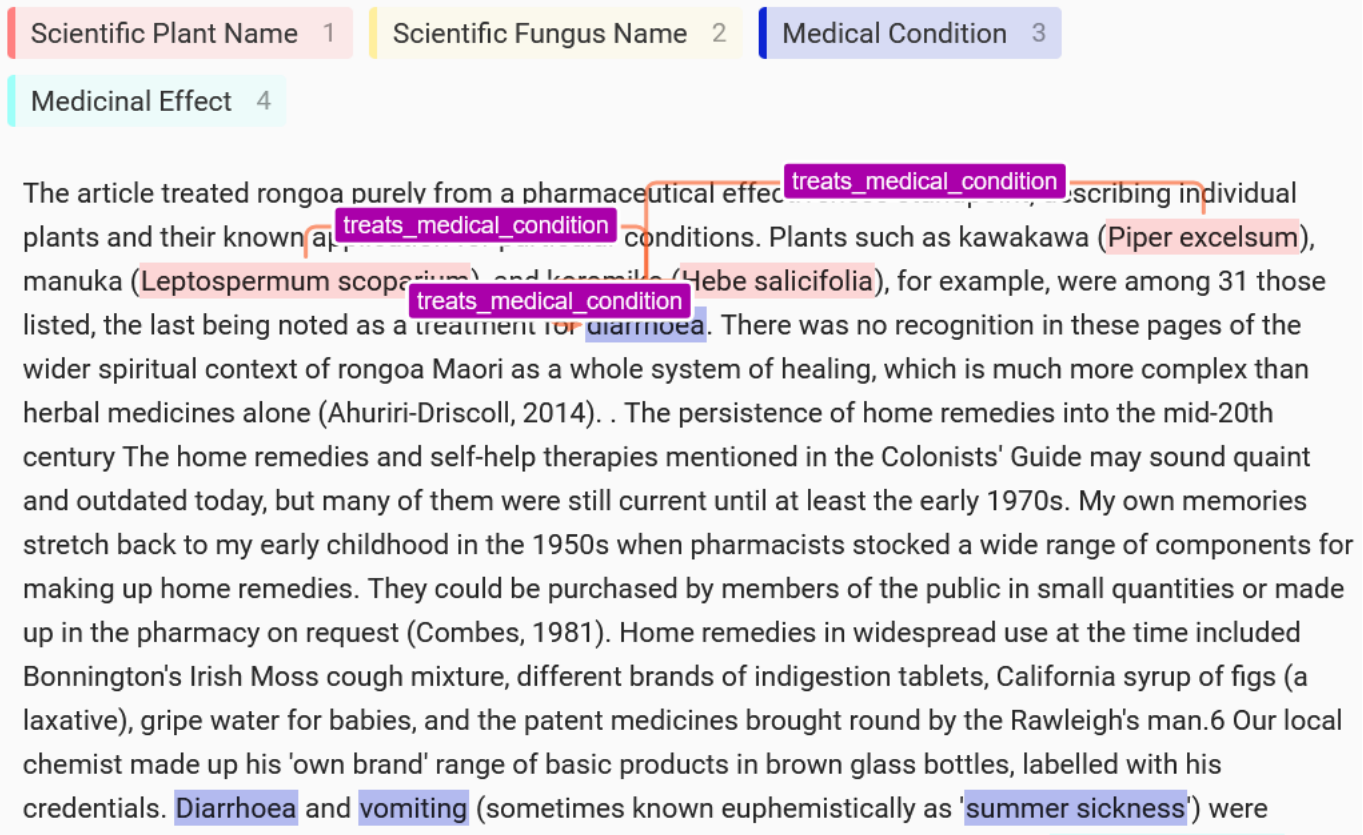
COFFEE METABOLOMICS

- Exciting ‘re-discovered’ coffee species and its chemical similarities and dissimilarities with arabica and robusta coffees
- One of many advantages of working at Kew!



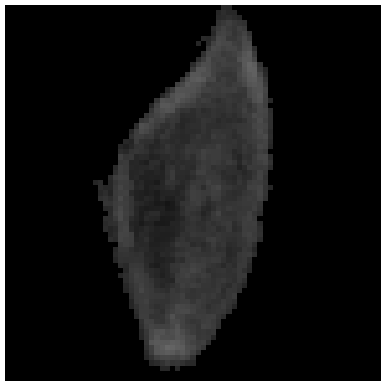
TEXT MINING

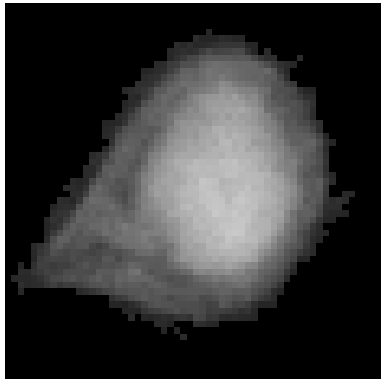
- Kew provides some key global datasets for plant and fungal knowledge
- Enormous amounts of new knowledge and information are published globally
- Working on methods to integrate text mining tools for a new Plants for Health dataset and web portal



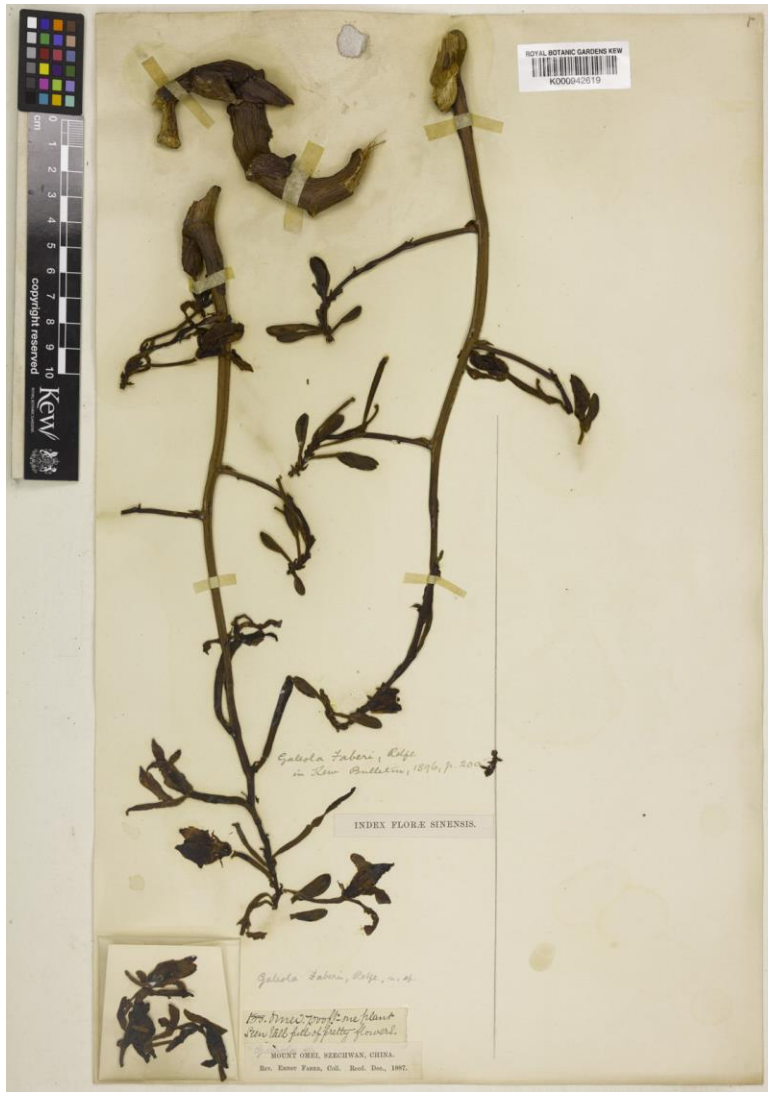
SEED GERMINATION

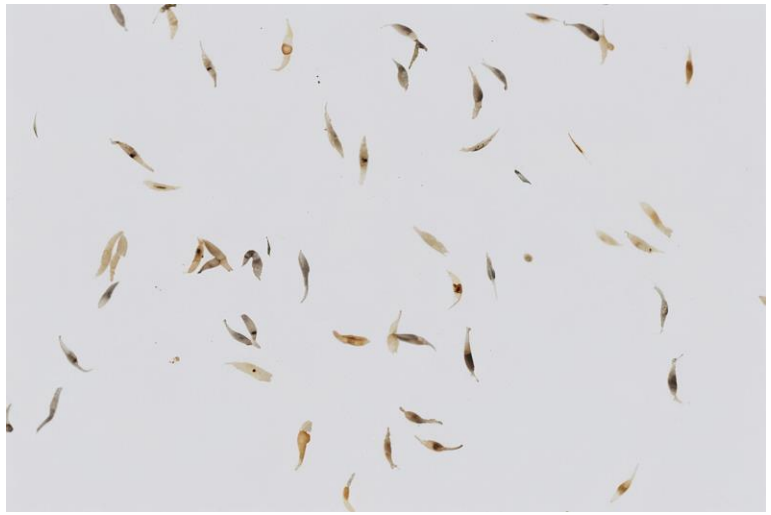
- Kew's Millenium Seed Bank holds over 2.4 billion seeds from around the world
- Understanding viability of seeds added to the collection or given out is vitally important
    - Basic features from xrays of pine seeds can provide very reliable predictions of germination rates
    - *Can CNNs provide further insights and improve on existing methods?*

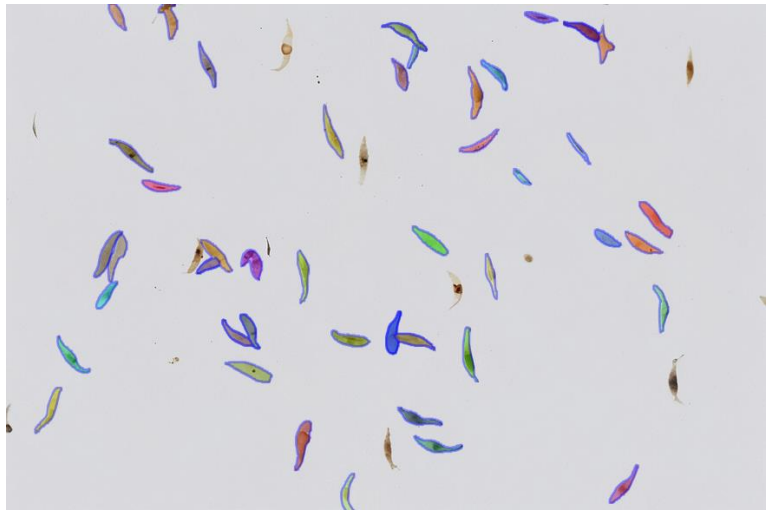




SEED GERMINATION







- Stained orchid seeds can be reliably classified as viable or non-viable by experts
- This work is labour intensive and orchid seeds can be tiny! (c. 0.05 to 6.0 mm in length)
- *Can we reliably automate the counting and classifying of these seeds?*

THANK YOU!

Other advantages of working for Kew...