ICMP culture collection: Informatics and genomics

Dr Bevan Weir, ICMP Curator & Director
Host Institute of ICMP

- Government Research Institute
  - Manaaki Whenua
- Environmental Research
  - Greenhouse gasses
  - Pest control
  - Soils research
  - Sustainable land use
  - Biological Systematics
- Holds National Biological Collections
- 40% Government funded
The ICMP culture collection

- 21,000 cultures
- 50:50 Fungi : Bacteria, since 1952
- Broad taxonomic diversity
- Plant pathogens, ‘mushrooms’, beneficials e.g. rhizobia
- Many cultures found only in NZ
- Almost all bacterial plant pathogens
- Database: Taxonomy, history, strain properties, GenBank links all online

http://www.landcareresearch.co.nz/resources/collections/icmp
Biodiversity informatics
Biodiversity informatics

• A single custom built software package for all collections
  - Expensive and slow to implement, but flexible and responsive to needs

• Open data is a New Zealand government mandate
  - (almost) all data online and CC-BY 4.0 INT
  - Federated into GBIF and other platforms
  - All sequences in GenBank

• GBIF
  - Publish to an IPT server weekly
  - Very useful for data quality metrics: geolocation and date errors
  - Download stats
  - Citable DOI datasets
  - Enables biodiversity science
Genomics
**Pseudomonas syringae genomes**

- First in-house genome November 2010 on 454 Jr
  - Psa disease of kiwifruit in New Zealand, effector genes
- Recently sequenced with D. Guttman 391 ICMP *P. syringae* genomes (in press)
  - *P. syringae* species complex is subdivided into primary (agricultural) and secondary (environmental)
  - inter-phylogroup recombination is rare
  - This level of genetic cohesion and the shared plant-associated niche argues for considering the primary phylogroups as a true biological species.
Leotiomycetes phylogenetic classification

- **Leotiomycetes** are ecologically diverse fungi, including mycorrhizas, endophytes, pathogens, and saprobes.
- Higher level classification based on morphologically defined taxa.
- Several known useful genes but hard to design working primers, no need with full genomes.
- Tree based on 3,156 genes from 51 selected genomes, 10 new ICMP genomes (in press).
- A framework for enabling future taxonomically targeted studies using deliberate specimen selection.
**Epicoccum secondary metabolites**

- Secondary metabolites do interesting things
- *Epicoccum* red pigment has anti fungal activity
- Custom genome annotation pipelines for these biosynthetic gene clusters
- Supports recently described *Epicoccum* species
- Informs biodiscovery of new biological compounds
Our Land, Our Future
Tō tātou whenua, mō āpōpō