Challenges, opportunities and strategies for building national computational platforms for bioinformatics

Steven Manos, Rhys Francis, Ziad Al Bkhetan, Ove Johan Ragnar Gustafsson, Uwe Winter, Andrew Lonie

ALCS 23

AUSTRALASIAN LEADERSHIP COMPUTING SYMPOSIUM

Thursday 15 June, 2023

Australian BioCommons

NCRIS National Research Infrastructure for Australia
An Australian Government Initiative

BIOPLATFORMS AUSTRALIA

The Australian BioCommons

Mission:

- To **support (Australian) life science research communities** with **community scale digital infrastructure**

- To enable **access to computational services** that:

    - Provide sophisticated analysis capabilities (including software and hardware platforms)
    - Support digital asset stewardship and management, retention, integration and publication

- To provide **training and support** solutions that enable the rapid and broad based adoption of the above

BIOPLATFORMS AUSTRALIA

NCRIS National Research Infrastructure for Australia
An Australian Government Initiative

Australian BioCommons

We engage communities through broad topic areas

genome assembly

genome annotation

microbiome analysis

metabolomics

proteomics

comparative genomics

human genomics

single cell omics

cmd line users

Australian BioCommons

… and facilitate interactions with computational providers to build a more fit-for-purpose, flexible and cohesive data analysis ecosystem

genome annotation

genome assembly

proteomics

metabolomics

metagenomics

human genomics

Australian BioCommons

ABLeS

The Australian BioCommons Leadership share

Australian BioCommons

Zero Childhood Cancer Consortium

Australia's world-leading precision medicine program for children with cancer.

Genomics for Australian Plants

Producing sequence data for 20 key plant species.

Plant Pathogen 'Omics Initiative

Generate important data resources for priority plant pathogen species.

ZERO Transcriptomics

Analyse the splicing profiles and structural rearrangement events within the RNA samples

Australian Amphibians and Reptiles Genomics

Understanding of evolution and conservation of Australia's unique native Amphibians and Reptiles

Threatened Species Initiative

Improve conservation practices using cutting-edge genomics technology

Australian BioCommons

GAP Phylogenomics – Australian Angiosperm Tree of Life

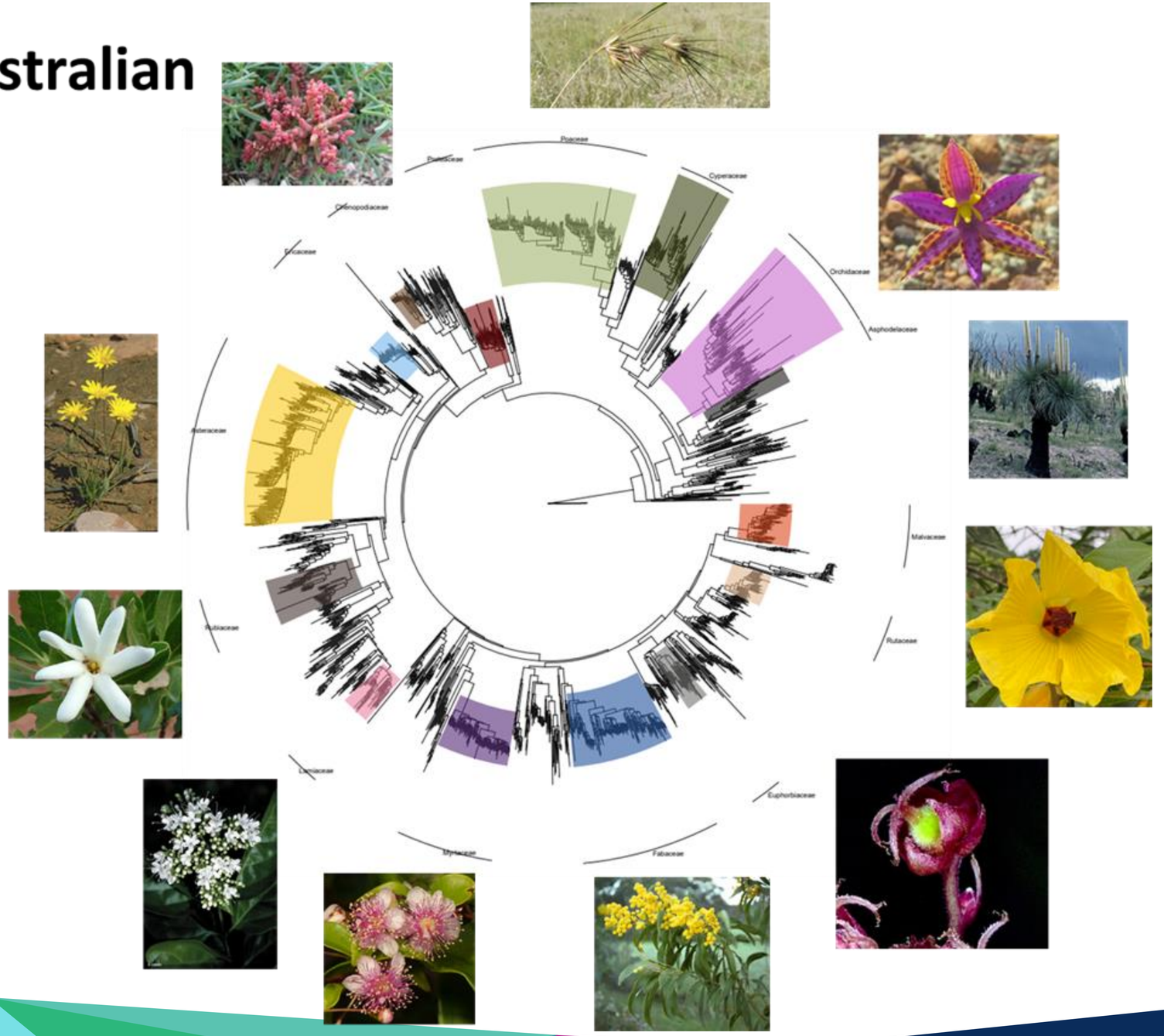
Final tree now complete

95% of all Australian angiosperm genera

2232 species' DNA sequenced

353 genes

438 Mbp of aligned sequence



Genomics for Australian Plants

ROYAL BOTANIC GARDENS VICTORIA

Australian BioCommons

Computing features of life sciences

Time

Bioinformatics is an inherently sample based discipline and therefore any downstream compute is episodic in nature.

Communities form around disciplines and in turn around methods and data over multiple years.

Skills

Augment scientists with analysis capabilities.

Provide human usable interfaces to otherwise complex CLI environs and ecosystems.

Adapt to the community nature of dispersed and diverse skills.

Community

The Structure and Timing of communities needs to be reflected in new access policies.

Data

The nature of data is different here.

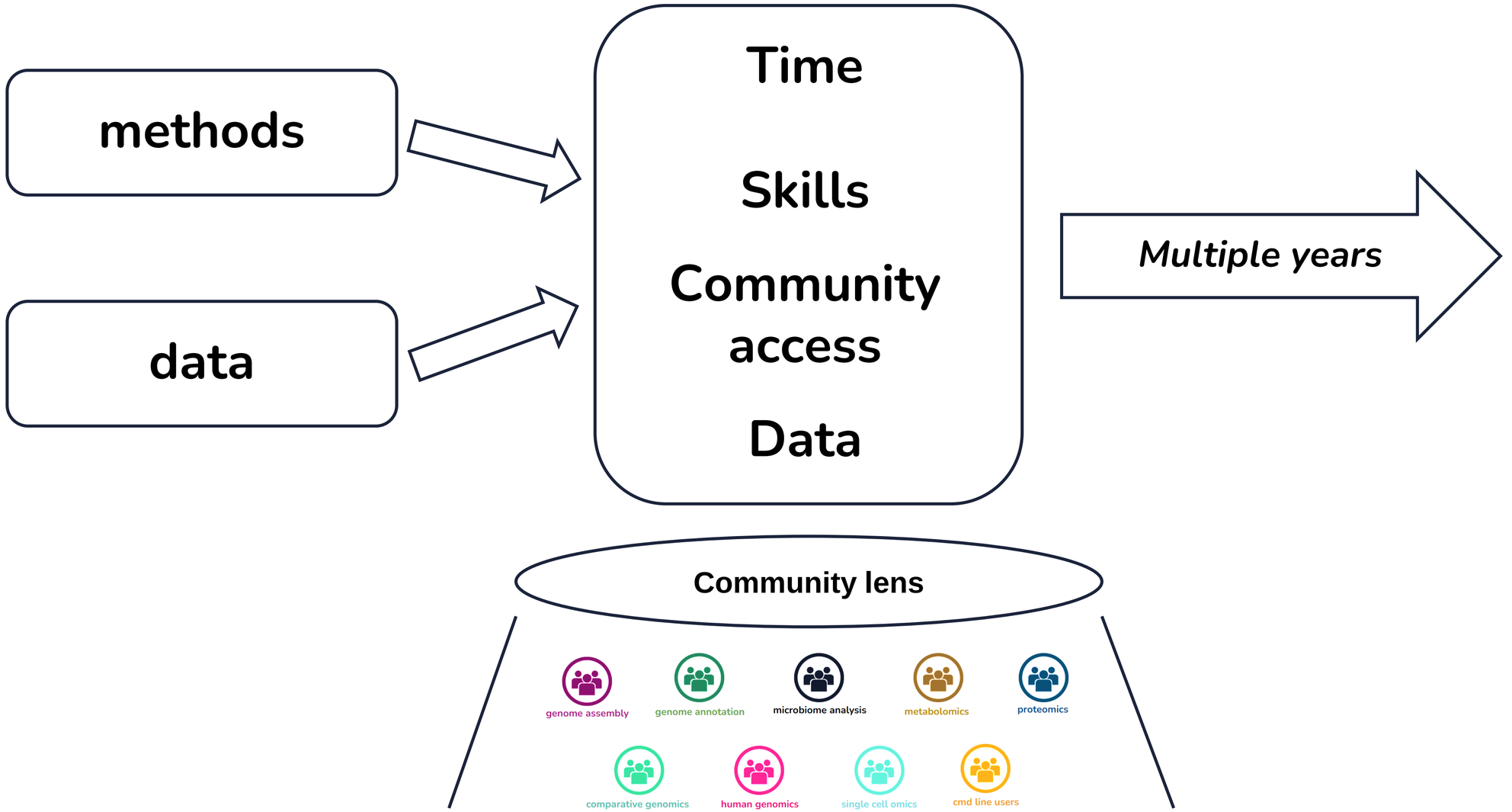
It’s dispersed.

It’s actually core and key, both an input and an output.

The outcome is often data in a repo elsewhere.

Compute is just an intermediary.

Human-actionable ⟷ Machine actionable



Australian BioCommons

Thanks!

Any questions?

steven@biocommons.org.au

Australian BioCommons