

**Genomics for Australian Plants (GAP):**

**Call for species suggestions 2020**

**GAP Background**

Australia has a unique flora of over 21,000 (flowering plants) species. The [Genomics for Australian Plants Initiative (GAP)](https://www.genomicsforaustralianplants.com/) was catalysed by Bioplatforms Australia in partnership with researchers from the Australian State and National Herbaria and Botanic Gardens. This nationally-inclusive and collaborative initiativeis supported by funding from Bioplatforms Australia through the Australian Government National Collaborative Research Infrastructure Strategy (NCRIS), the Ian Potter Foundation, Royal Botanic Gardens Victoria, the Royal Botanic Gardens and Domain Trust, CSIRO, Centre for Australian National Biodiversity Research and the Department of Biodiversity, Conservation and Attractions, Western Australia.

The central resource for this initiative will be derived from herbaria and botanic gardens (living collections) around the country. The addition of genome sequencing data will add significant value to the collections and contribute to the development of new methods and capabilities. This project is aligned with and will help deliver on strategic actions identified in the recently released [“Discovering biodiversity: a decadal plan for taxonomy and biosystematics in Australia and New Zealand 2018–2027”](http://www.science.org.au/support/analysis/decadal-plans-science/discovering-biodiversity-decadal-plan-taxonomy).

The GAP initiative is being will be driven by the plant research collections community and brings together researchers, data specialists, State and Commonwealth government agencies, and plant conservation and research agencies with the aim of using genomics approaches to enable the community to better understand, utilise and conserve Australia’s unique plant diversity.

Bioplatforms’ investment in genomics data generation is augmented with in-kind contributions from the wider plant community in sourcing plant material, preparation of material for genomic sequencing, development of bioinformatics pipelines for data analysis and data sharing. This includes. contributing to and benefiting from collaborations with other international initiatives.

The three main aims in the initiatives are in the generation, analysis and dissemination of genomics data in the following areas:

* Reference genomes
* Phylogenomics
* Conservation genomics

**Lessons learnt from the Reference Genomes pilot**

The GAP Initiative was launched at the Australasian Systematic Botany Society Inc. (ASBS) conference in Brisbane on Dec 6, 2018. Following a call for proposals from the plant community, three species were chosen for the reference genome pilot: *Acacia pycnantha*, Telopea speciosissima and Areocleome oxalidea. For more details, please click [here](https://www.genomicsforaustralianplants.com/2018/12/18/dec-2018-plants-launch/). The aim is to assemble DNA sequence data into annotated genome sequences, that can act as a platform for ongoing genomic-level analysis. These genome sequences will sit at some compromise between cost, completeness, contiguity, and correctness – depending on the decisions of the group as to what best provides a suitable basis for answering – or sparking new - research questions. Genomes are called “reference genomes” as they provide a research base, rather than meet particular pre-defined specifications (e.g. N50 length, Busco scores). To best facilitate continuous improvements in genome sequence quality (under any of these metrics), assembled genomes will have version numbers.

The pilot projects were chosen for GAP to identify the potential challenges and pitfalls working with Australian native plants. Unsurprisingly, the pilot projects presented a significant challenge to the teams. After 15 months, the Areocleome project was unable to continue due to difficulties in obtaining plant material. Both *Acacia* and Telopea have generated long and linked read data and are currently working towards generating additional data (e.g. RNA, Hi-C genome assembly data). Unfortunately, the Telopea plant chosen for the pilot perished in the recent bushfires, highlighting the need to prepare clonal material and ensure that strategies are in place to preserve nominated specimen plants for future sampling.

Another major hurdle that was not anticipated was the difficulty in obtaining plant material (heat stress due to extreme weather conditions, timing of plant growth) and challenges in extracting DNA suitable for long read sequencing. The DNA extraction methods successfully used by the Acacia and Telopea group can be found [here](https://www.genomicsforaustralianplants.com/protocols/).

We have also learnt from challenges in the data generation and bioinformatics space. These include the type and combination of sequencing technologies to be used, compute challenges (capacity, access and configuration), and rapidly evolving tools available to assemble the genomes.

For more details on the lessons learnt, please click [here.](https://docs.google.com/document/d/1RNghfr5Ivs2Bxo-_CTCCfAWWq67g7oRs0ROYM6dCPyU/edit) It is recommended that all EOI submitters familiarise themselves with the potential challenges and address these in their EOI submissions.

**Recommendation for a successful outcome**

Based on the pilot studies, the GAP Steering Committee has prepared a set of recommendations to ensure a successful outcome in generating more routine sequence data and genome assemblies.

**Plant**

1. Genome size and complexity – e.g. is your genome a polyploid? Is there any cytological evidence of clustered or dispersed distribution of heterochromatin in chromosomes? Does the group have the plan and ability to meet the challenges of potentially large/polyploidy genome including personnel resources (e.g. bioinformatician time), compute resources and additional planning requires to research tools required for resolving polyploid chromosomes etc? Can the group provide evidence of genome size and ploidy level as part of the EOI?
2. Availability of plant material – Is your plant material vouchered? (if not, what steps will you take to ensure voucher specimens are secured before or when DNA is sampled?) Do you have clones of the chosen plant in one or more location? What steps will be taken to protect the plant/clonal lineage?
3. Experience extracting DNA – Do you have experience extracting long read DNA from your species? Do you have collaborators who can assist with DNA extraction\*? Are you willing to spend time talking to other researchers and commit resources to extracting high quality, long read DNA? \*If not, then the GAP Steering Committee can assist in putting you in contact with relevant labs.

**Personnel**

1. Dedicated team leader – Can the nominated team leader commit time to the project? If you cannot dedicate adequate time to manage the project, the chances of the project succeeding are very low.

Team leader expectations:

1. Point of contact with the GAP Project Manager, GAP Steering Committee, wet lab and dry lab teams
2. Actively manage both the wet lab and dry lab teams, communicating progress and issues with the GAP Project Manager
3. Meet once a month with other GAP reference genome teams and bioinformaticians to discuss project progress (wet lab and dry lab) via video conference.
4. Dedicated personnel to extract DNA – Do you have dedicated personnel to extract DNA? Long read DNA extraction poses a significant challenge, even for teams who routinely extract DNA from the same species. Successful projects have dedicated personnel who have spent months continuously trying new protocols and speaking to other groups who have experience in extracting DNA and discussing the minimum QC requirement for the samples to be run.
5. Dedicated bioinformaticians – Do you have dedicated bioinformaticians who can work on the project? Can they commit the time to perform the bioinformatics analyses? \*It is expected that the reference genome bioinformaticians meet once a month with other GAP reference genome teams and bioinformaticians to discuss project progress (wet lab and dry lab) via video conference. \*If not, then the GAP Steering Committee can assist in putting you in contact with relevant institutions.

**Expected time frames**

1. DNA extraction – 4 months from project commencement
2. Draft genome – 1 year from project commencement

**Call for candidate species suggestion 2020**

In the first stream of activity, the Genomics for Australian Plants Initiative aims to sequence and assemble genomes of representative Australian plant taxa, these reference genomes will be important for identifying universal markers that can be used in phylogenomic approaches to assemble a tree of life for Australian vascular plant genera and develop markers for use in conservation genetics and other applications. Based on the recommendations from the two working groups and lessons learnt from the pilot projects, a suite of criteria has been developed to enable selection of taxa for reference genome assembly. **We are now calling for suggestions from the community for candidate species for the GAP Reference Genome project. Please complete the questionnaire below on pages 7 - 11 and return to Mabel Lum** **mlum@bioplatforms.com** **by** **COB Monday, 6th July 2020.**

We request a response to the questions posed below to enable the Steering Committee to select the most appropriate candidate species.

It should be noted that this is NOT a grant opportunity.

The sequencing will be undertaken by Bioplatforms Australia genomics facilities through their Framework Initiative strategic investments (<http://www.bioplatforms.com/projects/>), while the data assembly is to be supported by research community co-investment.

EOI submitters and collaborators will be required to undertake the following activities using their own resources:

1. Project design
2. Sample preparation and supply of metadata
3. DNA extraction from samples
4. Data analysis
5. Publication
6. Raw sequence data and assembled genome submission to international repository

Bioplatforms will provide the following resources:

1. Sequencing at Bioplatforms facilities
2. Project management and data management (raw data storage and access, with associated metadata)

**Genomics for Australian Plants FAQ**

1. **How can I get involved?**

All Consortium members are required to agree and sign the GAP Collaborative Agreement.

GAP Agreements and Policies can be found here:

<https://www.genomicsforaustralianplants.com/agreements-policies/>

GAP webpage:

<https://www.genomicsforaustralianplants.com/>

GAP Data Portal:

<https://data.bioplatforms.com/organization/about/bpa-plants>

Note: Access to GAP data is restricted to Consortium members for period of 12 months.

Please see the Data Policy for more details:

<https://www.genomicsforaustralianplants.com/agreements-policies/>

1. **How is the GAP project funded? How long is the GAP project funded?**

GAP is funded by Bioplatforms through the Australian Government National Collaborative Research Infrastructure Strategy (NCRIS), with additional contributions from the Botanic Gardens and Herbaria.

Bioplatforms funding for the project is for 3 years, from December 2018.

The Ian Potter Foundation has also provided substantial funds for 5 years (bioinformatics position, bioinformatics training and some funds for wet lab).

1. **What are Bioplatforms’ investments in GAP?**

Bioplatforms investment is in genomics data generation, raw data storage and project management.

1. **What in-kind contributions are expected of the researchers?**

In-kind contributions from the wider plant community include sourcing plant material, preparation of material for genomic sequencing (e.g. DNA extractions), development of bioinformatics pipelines for data analysis and data sharing (e.g. contributing to and benefiting from collaborations with other international initiatives).

1. **Can I apply for funding from GAP for my own research?**

GAP is a nationally-inclusive and collaborative initiative. We do not fund individual research projects. The aim of the initiative is to build resources which will benefit the plant community.

1. **What is the GAP Data Policy? Who can access data from my samples if I am part of GAP?**

Data access falls broadly into two phases: a “mediated-access” phase, where access to the data will be limited to members of the Consortium and other authorised parties; and an “open-access” phase where the data will be made openly available from resources including International Data Repositories.

Consortium members who have agreed and signed the GAP Collaborative Agreement will be given access upon registration at the Data Portal.

The “open-access” phase is set at 12 months from deposition of data into Bioplatforms Data Portal to allow the members of the Consortium and other authorised parties to progress analysis and publications, but not hold up the release of the data to be made openly available from resources including International Data Repositories.

GAP Data Portal:

<https://data.bioplatforms.com/organization/about/bpa-plants>

GAP Agreements and Policies can be found here:

<https://www.genomicsforaustralianplants.com/agreements-policies/>

1. **Will I be acknowledged if other Consortium members use data from my samples in their publication?**

All communications (scientific or general publications and presentations) that arise from the Consortium’s work will appropriately acknowledge the input of all relevant contributions.

The publications should specify the collaborative nature of the project, and authorship is expected to include all those contributing significantly to the work.

Authorship for publications will be determined on a case-by-case basis.

For more details, please see the GAP Agreements and Policies:

<https://www.genomicsforaustralianplants.com/agreements-policies/>



**Genomics for Australian Plants:**

**Call for species suggestions**

**Instructions for filling in this document:**

1. Complete one document for each species.
2. Address all the criteria listed.
3. Provide a short 1 paragraph answer for each question.
4. Answers to all the criteria should not exceed 4 pages.

**Dateline:** COB Monday, 6th July 2020

Please email your completed documents to Mabel Lum mlum@bioplatforms.com

**Contact information:**

David Cantrill | Project information | T: +61 3 9252 2301 | [David.Cantrill@rbg.vic.gov.au](http://www.bioplatforms.com/australian-plants/David.Cantrill%40rbg.vic.gov.au)
Mabel Lum | Project Manager | T: +61 2 9850 1174 | mlum@bioplatforms.com

**Please fill in your details below:**

|  |  |
| --- | --- |
| Name: |  |
| Affiliation: |  |
| Contact details: | Email:Phone: |
| Species: |  |
| Genome size: |  |
| Interest in the project: |  |
| Expertise: |  |
| Conflict of interest to declare: |  |

**List of criteria to be addressed:**

4. Does the species resonate with the public or have political value in additional to scientific value?

3. Does the genus lack a reference genome such that a genome of this species will contribute to furthering plant research in Australia? What is the phylogenetically closest genome currently available or being developed?

2. How much additional research can this genome leverage e.g. opening up new areas or research or leverage new funding (local and international)? Quantify, if possible and provide evidence.

1. Is the species currently being worked on so that the initiative can add value to it?

**List of criteria to be addressed:**

8. Is there access to a sample of required quality and quantity for 'reference' genome analysis and that can be vouchered?

7. Is the species of a genome ploidy and size that makes a 'reference' genome feasible?

6. Does the species have commercial or conservation value?

5. Is the species phylogenetically representative of the genus?

**List of criteria to be addressed:**

11. Are there groups that are willing to contribute to genome assembly?

Is the group willing and able to commit time required for genome assembly and attending monthly teleconference meetings?

10. Is there a group that is willing to undertake the sample preparation of this species (high quality DNA and RNA)?

Is the group willing and able to dedicate the time required for DNA and RNA extractions?

9. Are these samples available from an accessioned collection and an institution that has a

commitment to maintain and clonally propagate if necessary?

Is the group willing and able to take steps to protect the plant? Please outline below.

12. Are there any other factors to be considered for this species?

**List of criteria to be addressed:**

14. Is the team leader able and willing to commit the time required for the project?

Please see page 2 for an outline for the team leader expectations from GAP

15. Are you able to commit to the expected time frames below?

1. DNA extraction – 4 months from project commencement
2. Draft genome – 1 year from project commencement

13. Have you and all your named collaborators read and understood the GAP Agreements and Policies? Please contact Mabel Lum mlum@bioplatforms.com if you have any questions.

Link: <https://www.genomicsforaustralianplants.com/agreements-policies/>