

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

**Australian Angiosperm Tree of Life (AAToL)**

**Stage 1**

Expression of Interest

**GAP Background**

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 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.

Bioplatform’s 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 (e.g. 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

**GAP Phylogenomics**

GAP Phylogenomics has the ambitious long-term goal of resolving the Australian Angiosperm Tree of Life (AAToL) to species level using a common set of hundreds of (predominantly nuclear) molecular markers. This is to be achieved by supporting collaborators to accomplish their research goals that align with GAP Phylogenomics goals, and coordinating these efforts across the botanical research community. Work will proceed in three stages (as follows). Current resources limit GAP to completion

of the first two stages. It is anticipated that the genomic and bioinformatic resources and research capabilities built within the Australian plant research community will help leverage funding for the third stage.

**Stage 1** Resolve the AAToL to genus level using one exemplar species of at least 95% of accepted ([APC](https://biodiversity.org.au/nsl/services/APC)) Australian angiosperm genera.

**Stage 2** Generate phylogenomic datasets with denser sampling within genera to address questions of monophyly, evolution, and biogeography.

**Stage 3** Resolve the AAToL to species level, complete for 95% of accepted ([APC](https://biodiversity.org.au/nsl/services/APC)) Australian angiosperm species.

The GAP phylogenomics project will adopt a target capture approach (a.k.a. bait capture, target enrichment, exon capture) using a common set of hundreds of nuclear markers, with plastid markers recovered as off-target reads. This will maximise the chance of resolving the backbone of the phylogeny (to genus level) to provide a phylogenetic framework for continental-scale systematic, ecological and evolutionary studies. Future work (i.e. Stages 2 and 3 above) will focus on more detailed studies of specific clades, which may require the use of clade-specific marker sets.

Stage 1 of the GAP project will be undertaken in collaboration with the Plant and Fungal Tree of Life project ([PAFTOL](https://www.kew.org/science/our-science/projects/plant-and-fungal-trees-of-life)). PAFTOL aims to estimate a genus-complete phylogeny of the world’s angiosperms (*c*. 14,000 genera) using the [Angiosperms-353](https://academic.oup.com/sysbio/article/68/4/594/5237557) target capture nuclear bait set. PAFTOL is in the final two years of its program and current funds are insufficient to complete more than 50–60% of genera. The GAP – PAFTOL collaboration ensures at least 95% coverage of Australian genera, using Australian species exemplars. In stage 1, target capture sequencing will use the Angiosperms-353 bait set to leverage global datasets generated by PAFTOL and others, and to maximise re-use of GAP data. We are currently undertaking a pilot study comparing the [Angiosperms-353](https://academic.oup.com/sysbio/article/68/4/594/5237557) and Waycott lab target capture nuclear bait sets. The outcomes of this pilot will assist in the design of Stages 2 and 3.

**How can I contribute to GAP Phylogenomics?**

For participation in AAToL Stage 1, we seek contributing teams to provide taxonomically verified and vouchered samples for Australia's plant genera in the form of DNA shotgun libraries suitable for target sequence capture. The cost of target capture and high throughput sequencing will be covered by Bioplatforms and can only be carried out by a Bioplatforms facility (e.g. Australian Genome Research Facility - AGRF). Teams may be institutions, or groups of collaborating researchers. The team lead must be based at an Australian institution.

The aim of Stage 1 is to generate target capture sequence data for all 2081 Australian native angiosperm genera. Sequencing for c. 450 of those genera (using Australian native species) is complete or underway by PAFTOL and/or the GAP pilot study. Sequencing of the remaining c. 1630 genera will require 18 x 96-well plates of DNA libraries. To achieve this goal in a time and cost-efficient manner, participating teams are kindly requested to contribute a minimum of **one 96-well plate of libraries by 30 June 2020**.

Sample selection will be facilitated through a mediated process between contributing teams and GAP's Sample Coordinator to ensure the aims of AAToL stage 1 are met. Teams are expected to cover a proportion of samples which may not align with their research priorities, so called “orphan taxa”, to ensure that all Australian genera are covered. In the EOI form below, teams will be asked to provide details of the strength of their collections and specific taxonomic expertise in their team which will guide prioritisation of sampling efforts across different teams. In cases of contributing teams with overlapping research interests, GAP will encourage collaboration between teams to maximise synergies and minimise duplication of effort.

**What are the sample requirements for GAP Phylogenomics?**

All plant samples must comply with the requirements of the Nagoya Protocol, and be vouchered by a specimen identified to species level that is accessioned in a public herbarium. Suitable plant material for preparation of libraries includes silica dried tissue, most herbarium material, and fresh material. Unsuitable material includes material fixed in ethanol or FAA or dried herbarium material that has been treated with alcohol or other preservative chemicals (e.g. using mercuric chloride). Wherever possible, the type species of the genus or a close relative is prioritised, unless the type is not native to Australia. Only one species per genus will be sequenced in the AAToL Stage 1.

Samples to be provided to AAToL for sequencing must be DNA-extracted, sheared to c. 350 bp fragments, and Illumina sequencing libraries prepared as co-investment from the contributing team. Library preparation is expected to cost up to $60 per sample (actual cost to be determined). The Project Manager, Mabel Lum ([mlum@bioplatforms.com](file:///C%3A%5CUsers%5CMabel%20Lum%5CDropbox%20%28BPA%29%5CBPA%20General%5CAustralian_Plants%5CCommunications%5CEOI%5C2.%20Phylogenomics%5C3.%20EOI_Nov2019_v2%5Cmlum%40bioplatforms.com)) can provide further information on sample quality requirements. For interested participants that lack access to suitable facilities for DNA extraction and/or library preparation, please contact Mabel Lum to discuss options for outsourcing.

**What in-kind contributions are expected from the contributing teams?**

Please see below for a list of in-kind contributions that are expected from the contributing teams. Bioplatforms investments are also listed below. For further information, please contact the Project Manager, Mabel Lum <[mlum@bioplatforms.com](file:///C%3A%5CUsers%5CMabel%20Lum%5CDropbox%20%28BPA%29%5CBPA%20General%5CAustralian_Plants%5CCommunications%5CEOI%5C2.%20Phylogenomics%5C3.%20EOI_Nov2019_v2%5Cmlum%40bioplatforms.com)>.

|  |  |  |  |
| --- | --- | --- | --- |
| **Steps** | **Responsibility**  | **Cost covered by** | **Notes** |
| Sample metadata | Researcher to provide all the required metadata. Note: Sample metadata will only be required if your EOI is accepted.  | Researcher  | We cannot accept samples with no accompanying metadata.We cannot accept samples that do not comply with the Nagoya Protocol.  |
| **Steps** | **Responsibility**  | **Cost covered by** | **Labs with expertise** |
| DNA extraction | Researchers or Bioplatforms facilities | Researchers  | AGRF, most herbaria and universities |
| Library prep | Researchers or Bioplatforms facilities | Researchers | AGRF, several other labs (contact Project Manager) |
| Library quantification | Researchers  | Researchers.Researchers to confirm quantity and quality prior to submission. QC will be performed by Bioplatforms facilities  | AGRF, several other labs (contact Project Manager) |
| Library QC | Bioplatforms facilities | Bioplatforms  | Bioplatforms facilities |
| Target capture hybridisation (including provision of baits)  | Bioplatforms facilities | Bioplatforms  | Bioplatforms facilities |

|  |  |  |  |
| --- | --- | --- | --- |
| **Steps** | **Responsibility**  | **Cost covered by** | **Notes** |

|  |  |  |  |
| --- | --- | --- | --- |
| Sequence data generation | Bioplatforms facilities | Bioplatforms | Bioplatforms facilities |

|  |  |  |  |
| --- | --- | --- | --- |
| Raw data management | Bioplatforms facilities and Project Manager | Bioplatforms | Raw data can only be accessed from the Bioplatforms Data Portal. |
| Bioinformatics | Researchers |  Researchers |  |
| Data analysis | Researchers |  Researchers |  |
| Training/ knowledge transfer | Researchers |  Researchers |  |

**How can I be part of GAP Phylogenomics?**

All participating teams are required to read and sign the GAP Collaborative Agreement. These can be accessed in the links below.

* [Collaborative Agreement](https://www.genomicsforaustralianplants.com/wp-content/uploads/2019/04/GenomicsAustPlants_Collaborative_Agreement_final.pdf)
* [SCHEDULE 1: Data Policy](https://www.genomicsforaustralianplants.com/wp-content/uploads/2019/04/SCHEDULE_1_GenomicsAustPlants_Data_Policy_final.pdf)
* [SCHEDULE 2: Communications Policy](https://www.genomicsforaustralianplants.com/wp-content/uploads/2019/04/SCHEDULE_2_GenomicsAustPlants_Communications_Policy_final.pdf)

**GAP Data Policy and Data Access**

The GAP Data Policy and Data Access is outlined [here.](https://www.genomicsforaustralianplants.com/wp-content/uploads/2019/04/SCHEDULE_1_GenomicsAustPlants_Data_Policy_final.pdf) Consortium members who have agreed and signed the [GAP Collaborative Agreement](https://www.genomicsforaustralianplants.com/agreements-policies/) will be given access upon registration at the Data Portal.

Briefly, 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.

The “restricted-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. After this phase, the data will become publicly available, as outlined in the Data Policy.

**GAP communication and publication policy**

The GAP Communication Policy is outlined [here.](https://www.genomicsforaustralianplants.com/wp-content/uploads/2019/04/SCHEDULE_2_GenomicsAustPlants_Communications_Policy_final.pdf)

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

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

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

**Who can I contact for further information?**

If you would like to get involved and have any further questions, please email the Phylogenomics Lead and/or the Project Manager to discuss further.

Phylogenomics Lead: Prof Darren Crayn [darren.crayn@jcu.edu.au](file:///C%3A%5CUsers%5CMabel%20Lum%5CDropbox%20%28BPA%29%5CBPA%20General%5CAustralian_Plants%5CCommunications%5CASBS%20conference_Nov2019%5Cdarren.crayn%40jcu.edu.au)

Project Manager: Dr Mabel Lum [mlum@bioplatforms.com](file:///C%3A%5CUsers%5CMabel%20Lum%5CDropbox%20%28BPA%29%5CBPA%20General%5CAustralian_Plants%5CCommunications%5CASBS%20conference_Nov2019%5Cmlum%40bioplatforms.com)

Sampling co-ordinator: Lalita Simpson lalita.simpson@my.jcu.edu.au

**Phylogenomics working group members:**

|  |  |
| --- | --- |
| Prof Michelle Waycott | michelle.waycott@adelaide.edu.au |
| Dr Daniel Murphy  | Daniel.Murphy@rbg.vic.gov.au |
| Dr Hervé Sauquet  | herve.sauquet@gmail.com |
| Dr Katharina Nargar | Katharina.Nargar@csiro.au |
| Dr Kelly Shepherd | kelly.shepherd@dbca.wa.gov.au |
| Dr Matthew Barrett | matthew.barrett@dbca.wa.gov.au |
| Dr Anna Syme (GAP bioinformatician) | Anna.Syme@rbg.vic.gov.au |
| Dr Chris Jackson (GAP bioinformatician) | Chris.Jackson@rbg.vic.gov.au |

Please see pages 6-8 for instructions for submitting your EOI.



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

**Australian Angiosperm Tree of Life (AAToL)**

Stage 1: Expression of Interest

**Instructions**

1. Access the online Google Sheets workbook “GAP Phylogenomics EOI STAGE 1”. <https://tinyurl.com/ve6u3yp>.
2. Within the workbook, the MASTER spreadsheet contains the list of Australian angiosperm genera and instructions for completing the spreadsheet. The dummy team spreadsheets outline the sample requirements.
3. Each team will be provided an exclusive-access team spreadsheet to develop their EOI (contact the Project Manager to arrange this).
4. Address all the criteria listed on pages 6 – 8 of this document. Provide a short 1 paragraph answer for each question. Answers to all the criteria should not exceed 2 pages.
5. Complete the spreadsheet tab and submit this document to the Project Manager by COB Friday 13th March 2020.
6. If your EOI is accepted, everyone in the team will be asked to read and sign the [GAP Collaborative Agreement.](https://www.genomicsforaustralianplants.com/agreements-policies/)

**Deadline:** COB Friday, 13th March 2020.

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

**Contact information:**

Darren Crayn | GAP Phylogenomics lead | [darren.crayn@jcu.edu.au](file:///C%3A%5CUsers%5CMabel%20Lum%5CDropbox%20%28BPA%29%5CBPA%20General%5CAustralian_Plants%5CCommunications%5CEOI%5C2.%20Phylogenomics%5C3.%20EOI_Nov2019_v2%5Cdarren.crayn%40jcu.edu.au)
Mabel Lum | Project Manager | mlum@bioplatforms.com

**Team lead (must be based at an Australian institution)**

|  |  |
| --- | --- |
| Name: |  |
| Affiliation: |  |
| Contact details: | Email:Phone: |
| Conflict of interest to declare: |  |

**Team members (please insert more rows if required)**

|  |  |
| --- | --- |
| Name | Affiliation |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

**List of criteria to be addressed:**

3. Please indicate how each team member will contribute to the project. Please also indicate team members for which the project will facilitate training in genomic lab skills, bioinformatics and phylogenomic data analysis or who will contribute training/knowledge transfer.

1. Please describe which plant groups are particularly well represented/curated in your team/institutions.

2. Please indicate which taxonomic groups your team is particularly well positioned to contribute material to GAP and why. Please provide a summary here and further sample details in the online Google Sheets workbook “GAP Phylogenomics EOI STAGE 1”. <https://tinyurl.com/ve6u3yp>.

**List of criteria to be addressed:**

1. If your team/institution is able to provide in-kind and cash contributions, please list them below.

Additionally, please provide letters of support from the participating host institutions to the value of the indicated in-kind and/or cash co-contributions indicated below.

Please feel free to list any other additional contributions you would like to make to the GAP Phylogenomics.

|  |  |  |  |
| --- | --- | --- | --- |
| **Item** | **Contributions** | **Name** | **Institution**  |
| **In-kind (yes/no)**  | **Cash (yes/no).** **If yes, please list amount.** |
| DNA extraction |  |  |  |  |
| Library prep |  |  |  |  |
| Bioinformatics |  |  |  |  |
| Data analysis |  |  |  |  |
| Training/ knowledge transfer |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |