Australian Fgenesh++ Service Terms of Use and Policies

Version 1.0 (published 17 September 2021)

1. Definitions

1.1 Fgenesh++ Software

Fgenesh++ is a pipeline for automatic prediction of genes in eukaryotic genomes based on proprietary gene finding software. The software package includes both public software and data [1] as well as copyrighted software [2] which is owned by the company, Softberry Inc. Fgenesh++ software can NOT be used, copied or distributed without a license from Softberry Inc.

1.2 Fgenesh++ Service

The "Fgenesh++ Service" (the Service) is offered by the Australian BioCommons [3] and a number of its Operational Partners as part of their collective mission to provide access to digital techniques, data and tools, used by environmental, agricultural and biomedical researchers.

The Service includes access to a licensed copy of Fgenesh++ and relevant gene finding matrices that are installed on a virtual machine for a set time period. The majority of the systems administration tasks are taken care of by the Operators of the Service, which allows Australian-based research groups to upload genome assemblies and supporting evidence files to be annotated using the software.

"Fgenesh++ Service" (the Service) refers to the complete service that is provided to its users. This includes: (a) a web-based Reservation Booking System to enable users to book a timeslot for using the software, (b) obtaining relevant gene finding matrices from Softberry on behalf of the user, (c) creation of a virtual machine and installation of the Fgenesh++ software and relevant gene finding matrices on it, (d) user support from Australian Service Operators (and Softberry Inc. when required) for the time that access is enabled, (e) maintenance of the virtual machine and communications about the Service, (f) the creation and improvement of tutorials and documentation, and (g) the development and/or delivery of training materials and/or events for Australian audiences.

1.3 Fgenesh++ Instance

A "Fgenesh++ Instance" is a secure virtual machine with Softberry's Fgenesh++ software installed and with the requested gene finding matrices uploaded. Data owned by your research group or consortium is uploaded to the Instance which is available for an allotted period of time (generally 4 weeks). All instances are exclusively hosted on computational infrastructure provided by the Pawsey Supercomputing Centre in Perth, Australia

1.4 User Types

Due to the nature of how a user will access Fgenesh++ through the Australian Fgenesh++ Service, this document discerns between the following distinct user types and roles. Note that one individual may perform more than one of these roles.

User Type / Role	Responsibilities
Principal Investigator	Submits the request for access to the Fgenesh++ Service on behalf of a *bona fide* Research Group or Research Consortium. Must be affiliated with a recognised Australian academic or research institute (see Section 8) in a non-student capacity. Holds overall responsibility for the hosted Fgenesh++ Instance for the time period for which it is available. Is obliged to ensure all End Users adhere to all of the terms and conditions laid out in this document.
End User	Is authorised by the Principal Investigator to access and use the Fgenesh++ Instance to perform eukaryotic genome annotation. Obtaining a Pawsey Supercomputing Centre User account in order to use the Service Accessing the Reservation Booking System to reserve a timeslot in which to use the service, Uploading Data, performing the Genome Annotation Any End User along with their associated Principal Investigator is obliged to adhere to all of the terms and conditions laid out in this document.

2. Terms of Use

Interaction with any of:

- The Australian Fgenesh++ Service Support website: https://support.pawsey.org.au/documentation/display/US/The+Australian+BioCommons

- The Fgenesh++ Reservation Booking System https://reservation.pawsey.org.au/, or

- A Fgenesh++ Instance hosted through the Fgenesh++ Service.

Denotes agreement with the following terms:

2.1. Use of service

a. The Fgenesh++ Service is part of the Australian BioCommons suite of services that aim to provide fully subsidised access to publicly available services.

b. Principal Investigators affiliated with a <u>recognised Australian academic or research institute</u> (see Section 8) may choose to register for an account via the Fgenesh++ Service Sign Up on behalf of their Research Group or a Research Consortium in order to access the features described in our feature catalogue (Section 5).

c. End Users are authorised to interact with a Fgenesh++ Instance in accordance with the features described in our feature catalogue (Section 5).

d. Your access to the Fgenesh++ Service (no matter what role) is provided under the condition that you abide by the Fgenesh++ Service Acceptable Use Policy (Section 6).

e. You indemnify and hold the Fgenesh++ Service Operational Partners free and harmless in connection with your use of the Service.

f. You agree to acknowledge the use of the Fgenesh++ Service in your research outputs using the format of our Fgenesh++ Service Acknowledgement Statement (Section 7).

2.2 Data hosting and analysis

a. The Fgenesh++ Service is designed to support the automated annotation of your genomes. If you choose to upload and host data in the Service you shall ensure that you comply with applicable laws, including, but not limited to, privacy, data protection and intellectual property rights. It is the responsibility of the Principal Investigator and End User of a Fgenesh++ Instance to comply with these laws.

b. The Fgenesh++ Service is not designed to handle sensitive data, including potentially re-identifiable human-derived data. You are prohibited from using the Service to store or process sensitive data, including but not limited to personally-identifying genetic, genomic or medical data.

c. The Fgenesh++ Service is designed for short-term access and does not provide any long term storage for any data, including your input genome assembly and output genome annotation data.

d. The license conditions stipulate that one named user can access the software at any one time. In order to meet this condition, access to the Fgenesh++ Service is strictly controlled and administered through the Reservation Booking System system to ensure the number of licensed users accessing the software is not breached at any time. Therefore access to an End User is granted for a period of 4 weeks that has been booked through the Reservation Booking System. End Users will be expected to set their own reminder alerts at the time of reservation booking. At the end of the reserved time period, the Instance will be deleted.

e. If an Instance is not required for the entire 4 week period, the End User is requested to notify the system administrator when they have finished using the Instance.

f. It is the responsibility of the Project Investigator to ensure End Users have read, acknowledged and adhered to these Terms of Use.

2.3 Sharing / Downloading content

a. Project Investigators and any End Users associated with a Fgenesh++ Instance are the collective owners of any new data (content) generated through the use of the Fgenesh++ pipeline.

b. Project Investigators and End Users can choose to delete their content at any time.

c. It is the sole responsibility of the End User to download a copy of all output files off the Instance prior to its termination at the end of the booked availability period. These files can NOT be recovered by the system administrator once the instance is terminated.

d. It is expressly forbidden for any End User or Project Investigator to copy or download the Fgenesh++ software off the Instance.

2.4 Termination

a. The Fgenesh++ Service Operational Partners reserve the right, without notice, at its sole discretion and without liability:

i. to alter, delete or block access to content that it deems to be inappropriate, and

ii. to restrict or remove access where it considers that your use of the Fgenesh++ Service interferes with its operations or violates these Terms of Use or applicable laws.

2.5 Change of terms

These Terms of Use are subject to change at any time and without notice. Updates or modifications to the Terms of Use will be communicated on the support webpage and via email to current or upcoming users of the Service via their registered accounts.

2.6 Service maintenance and downtime

a. The Fgenesh++ Service is underpinned by resources at Pawsey Supercomputing Centre which is subject to scheduled and occasional maintenance downtime for updates that may impact the Service. Users will be notified prior to occasional scheduled maintenance downtime via email communication at the address provided by the user when joining the Fgenesh++ Service.

b. In the event of a security breach, Fgenesh++ Service may need to perform urgent maintenance resulting in Service downtime at a time other than the scheduled maintenance windows. In this case, Fgenesh++ Service will make

every effort to give the user/s as much advance notice as is possible to limit disruption.

If you have any questions or comments with respect to the Fgenesh++ Service, or if you are unsure whether your intended use is in line with these Terms of Use, or if you seek permission for a use that does not fall within these Terms of Use, please contact us via help@pawsey.org.au.

3. Fgenesh++ Service Privacy Policy

The Fgenesh++ Service is committed to protecting your privacy. The Fgenesh++ Service Operational Partners do not collect, process, retain or share personal data other than for the purposes described below.

3.1 Usage data

a. In order to monitor and enhance Fgenesh++ Service, the Fgenesh++ Service Operational Partners track and retain information about your interaction with the Service (i.e. IP addresses, cookies, pages visited, etc).

b. The Fgenesh++ Service Operational Partners may share log files of interactions with the service in aggregation with our Operational Partners.

c. The Fgenesh++ Service Operational Partners may share debugging information with the Softberry [4] developers to improve their systems. When the Fgenesh++ Service Operational Partners do, the Fgenesh++ Service Operational Partners aggregate and anonymise this data to protect your privacy.

d. For promotional reasons and reporting to funding bodies, the Fgenesh++ Service periodically creates public usage reports. When the Fgenesh++ Service Operational Partners do, the Fgenesh++ Service Operational Partners aggregate and anonymise this data to protect your privacy.

3.2 Registration data

a. If you choose to register an account with Fgenesh++ Service, the Fgenesh++ Service Operational Partners collect registration data (i.e. email address, institutional affiliation and public name).

b. Fgenesh++ Service Operational Partners collects and stores your personal information for the administration of your use of Pawsey services to support your research.

c. Fgenesh++ Service may contact registered account holders using the email address they provide for operational reasons, including:

    i. Checking user account security;

    ii. Communicating service downtime;

    iii. Communicating service improvements;

      iv.   Communicating actions the Fgenesh++ Service Operational Partners are taking with respect to data quotas, retention periods, and compute resource usage;

      v.   Responding to support requests;

      vi.   Surveying existing and planned use in order to improve and promote the Service;

      vii.   Discussing breaches of these Terms of Use, and;

      viii.   Communicating updates to Fgenesh++ Service policies.

d. Fgenesh++ Service may contact registered account holders using the email address they provide for promotional reasons, such as seeking consent and content for newsletters, annual reports or media releases by the Australian BioCommons, operational partners and funding partners.

3.3 Data you upload and share

a. You are prohibited from using Fgenesh++ Service to host or process sensitive data, including but not limited to potential re-identifiable or identifiable personal data.

b. The Fgenesh++ Service Operational Partners may examine the data you share through the service for alignment with our Fgenesh++ Service Acceptable Use Policy.

3.4 Support requests

a. Support is offered for users accessing the Fgenesh++ Service with two access points depending on your question:

      i.   For technical support:

          1.   Users can submit a request with help@pawsey.org.au. This includes support with: booking the service, acquisition of username or login, accessing the service, uploading the matrices, time extensions or changes to the reservation period. When you submit a support request to help@pawsey.org.au, the Fgenesh++ Service Operational Partners collect your email address and may ask for personal data as part of the support process.

          2.   Data collected as part of a support request is stored at CSIRO resources controlled by CSIRO as the lead agent of the Pawsey Supercomputing Centre ("Pawsey").

      ii.   For support related to bioinformatics and annotation:

          1.   The nominated End User can contact Softberry, Inc. at softberry@softberry.com during the duration of their access to the Instance.

3.5 Data location

a. User support data managed by Fgenesh++ Service Operational Partners may be stored outside Australia.

b. All other data is stored and processed in Australia by our Operational Partners, and will not be provided to any third party unless

c. You choose to share content through the service, and;

d. The Fgenesh++ Service Operational Partners are required to do so by law.

3.6 Modifying and removing personal data

a. The personal information collected by Fgenesh++ Service for registered account holders includes name, organisation and email address associated with the Fgenesh++ Instance. Optimally the Fgenesh++ Service Operational Partners may also store a user's account name and GitHub account name.

b. Registered account holders can modify collected information or request the removal of personal information at any time with effect for the future by emailing help@pawsey.org.au.

4. Fgenesh++ Service Operational Partners

The Australian BioCommons operates Fgenesh++ Service in collaboration with and underpinned by resources provided by, Pawsey Supercomputing Centre and Bioplatforms Australia. The service is underpinned by computational resources provided by the Pawsey Supercomputing Centre.

5. Fgenesh++ Service Feature Catalogue

5.1 Fgenesh++ Service features and quotas by account type

	Principal Investigator	End User
Request Fgenesh++ Instance	✓	✗
Request Matrix/ices	✗	✓
Request support from Softberry Inc.	✗	✓
Request more computational resources	✗	✓
Add/delete data	✗	✓
Annotate data	✗	✓
View	✓	✓
Export (download) Data	✗	✓
Data upload Quota	Up to 5 TB as needed for one week. Ongoing quota cap at 2 TB.	
Data retention period	Fgenesh++ Service Operational Partners will ensure that all data uploaded to an Instance (within the data quota limits outlined above) is available for the time that the instance (virtual machine) exists. It is the sole responsibility of the End User to download a copy of all output files off the Instance before its termination at the end of the booked availability period. These files can NOT be recovered by the system administrator once the instance is terminated.	

Principal Investigators must be affiliated with a recognised Australian academic or research institute (see Section 8), and register using an email address from our list of Australian research institutions.

5.2 Best effort operations

The Fgenesh++ Service Operational Partners realise that the data and annotations you host with Fgenesh++ Service are important to you. For that reason, the Fgenesh++ Service Operational Partners commit to making the best efforts to keep the Service as reliable and responsive as possible and to keep your data intact.

Our operational staff work normal Australian West Coast office hours (i.e. Australian Western Standard Time). During those hours the Fgenesh++ Service Operational Partners aim to triage tickets, tool requests and reference data requests within one business day. Resolution of tickets, requests, outages and issues will take longer outside normal office hours, on weekends, and on public holidays.

The Fgenesh++ Service runs research-grade infrastructure contributed by the Fgenesh++ Service Operational Partners. Fgenesh++ Service's reliability and capacity may be impacted if partner infrastructure unexpectedly becomes unavailable.

5.3 Service capacity

The Fgenesh++ Service is underpinned by significant compute and storage capabilities, allowing you to scale analysis beyond your desktop computer. The Fgenesh++ Service Operational Partners periodically project capacity requirements, taking into account growth in users, jobs and community requests for new types of analysis that might need specialised hardware.

5.4 Service lifetime

The service will continue for the lifetime of the Australian BioCommons, which is currently funded until June 2023.

6. Fgenesh++ Service Acceptable Use Policy

You may use the Fgenesh++ Service to analyse and annotate life science data for the purposes of research.

a. You agree to interact with the Fgenesh++ Service Community (defined as other users of the service, technical support staff, operational staff, and those involved in the governance of the service) in a responsible and respectful manner at all times, consistent with the Australian BioCommons Code of Conduct.

b. The following uses of the Fgenesh++ Service are explicitly forbidden:

1. Copying or downloading the Fgenesh++ software off an Instance;

2. The deliberate or reckless creation, transmission, storage, downloading, or display of any offensive or menacing images, data, metadata or other material, or any data capable of being resolved into such images or material;

3. Use which constitutes an infringement of any intellectual property rights of another person or group;

4. The creation of duplicate user accounts and/or double registration by a single user for the purposes of intentionally receiving increased quota on the Fgenesh++ Service;

5. Attempt to circumvent the appropriate authentication or access control measures, security or restrictions on the use of a registered account, including the unauthorised distribution or use of tools of any kind for compromising security;

6. The storage or processing of any sensitive data, including but not limited to potentially personally identifying genetic, genomic or medical data;

7. Communications which would be actionable under the law of defamation;

8. Use for private or business purposes other than scientific research, and;

9. The deliberate or reckless undertaking of activities which seek to, or result in:

    i. the imposition of (or attempted imposition of) an unreasonable burden on the Fgenesh++ Service Systems Administration team;

    ii. corruption of, or disruption to, data stored on the Fgenesh++ Service, or to the data of another user;

    iii. disruption to other users, and/or;

    iv. introduction or transmission of malware into the Fgenesh++ Service's systems.

c. Other uses of the Fgenesh++ Service are only permitted with the written permission of the Fgenesh++ Service Manager (contact via help@pawsey.org.au).

7. Fgenesh++ Service Acknowledgment Statement

The Australian Fgenesh++ Service is provided by the Australian Biocommons in collaboration with and underpinned by resources provided by Pawsey Supercomputing Centre and Bioplatforms Australia.

When papers acknowledge the Australian Fgenesh++ Service, they support continued research infrastructure investments by institutions and governments. When you use the Australian Fgenesh++ Service to support your publication or project, please cite its use using the following acknowledgement statement and web address:

"This work was supported by the Australian Fgenesh++ Service (biocommons.org.au/fgenesh-plus-plus) provided by the Australian BioCommons and the Pawsey Supercomputing Centre."

8. List of Australian Research Institutions

The Fgenesh++ Service is for exclusive use by publicly funded Australian researchers, as described in our feature catalogue. Access to the features and quotas is available to Principal Investigators registered using an email address from one of the Australian Research Institutions listed in the table below. The Fgenesh++ Service Operational Partners acknowledge that this list may not be exhaustive. Please contact us (help@pawsey.org.au) to request the addition of your research institution to this list.

Australian research institution	Email domains
ANZAC Research Institute	@anzac.edu.au
Australia New Zealand Gynaecological Oncology Group	@anzgog.org.au
Australian Catholic University	@myacu.edu.au, @acu.edu.au
Australian Centre for Heart Health	@australianhearthealth.org.au
Australian Institute of Aboriginal and Torres Strait Islander Studies	@aiatsis.gov.au
Australian Institute of Marine Science	@aims.gov.au
Australian Museum	@nma.gov.au
Australian National University	@anu.edu.au
Australian Red Cross	@redcrossblood.org.au
Baker Heart and Diabetes Institute	@baker.edu.au, @bakeridi.edu.au
Bond University	@bond.edu.au
Burnet Institute	@burnet.edu.au
Calvary	@calvarymater.org.au
Cancer Council NSW	@nswcc.org.au
Cancer Council Queensland	@cancerqld.org.au
Cancer Council Victoria	@cancervic.org.au
Centenary Institute of Cancer Medicine and Cell Biology	@centenary.org.au
Central Adelaide Local Health Network Incorporated	@sa.gov.au

Central Australian Aboriginal Congress Aboriginal Corporation	@caac.org.au
Central Queensland University	@cqu.edu.au
Centre for Eye Research Australia	@cera.org.au
Charles Darwin University	@cdu.edu.au
Charles Sturt University	@csu.edu.au
Children's Medical Research Institute	@cmri.org.au
CSIRO	@csiro.au
Curtin University	@curtin.edu.au, @student.curtin.edu.au, @postgrad.curtin.edu.au
Deakin University	@deakin.edu.au
Ear Science Institute Australia Incorporated	@earscience.org.au
Edith Cowan University	@ecu.edu.au
Federation University Australia	@federation.edu.au
Flinders University	@flinders.edu.au
Florey Institute of Neuroscience and Mental Health	@florey.edu.au
Griffith University	@griffith.edu.au, @griffithuni.edu.au
Health Support Queensland	@health.qld.gov.au
Heart Research Institute	@hri.org.au
Hudson Institute of Medical Research	@hudson.org.au
Institute for Breathing and Sleep	
James Cook University	@jcu.edu.au, @my.jcu.edu.au
La Trobe University	@latrobe.edu.au, @students.latrobe.edu.au
Ludwig Institute for Cancer Research	@ludwig.edu.au
Macquarie University	@mq.edu.au, @students.mq.edu.au, @hdr.mq.edu.au

Mater Research	@mater.org.au
Melanoma Institute Australia	@melanoma.org.au
Melbourne Health	@mh.org.au
Menzies School of Health Research	@menzies.edu.au
Metro South Hospital and Health Service	@health.qld.gov.au
Monash University	@monash.edu, @student.monash.edu
Murdoch Childrens Research Institute	@mcri.edu.au
Murdoch University	@murdoch.edu.au, @student.murdoch.edu.au, @ccg.murdoch.edu.au
Museum Victoria	@museum.vic.gov.au
National Acoustic Laboratories	@nal.gov.au
National Ageing Research Institute	@nari.edu.au
National Institute for Aboriginal and Torres Strait Islander Health Research	@lowitja.org.au
NSW Health Pathology	@health.nsw.gov.au
Peter MacCallum Cancer Centre	@petermac.org
Queensland University of Technology	@qut.edu.au
RMIT University	@rmit.edu.au, @student.rmit.edu.au
Royal Brisbane and Women's Hospital Foundation	@rbwhfoundation.com.au
South Australian Health and Medical Research Institute	@sahmri.com
South Australian Museum	@samuseum.sa.gov.au
Southern Cross University	@scu.edu.au, @student.scu.edu.au
St Vincent's Institute of Medical Research	@svi.edu.au
St Vincent's Health Australia	@svha.org.au

Swinburne University of Technology	@swin.edu.au, @student.swin.edu.au
Sydney Local Health District	@health.nsw.gov.au
Telethon Kids Institute	@telethonkids.org.au
The Asbestos Diseases Research Institute	@adri.org.au
The Bionics Institute of Australia	@bionicsinstitute.org
Queensland Institute of Medical Research	@qimr.edu.au, @qimrberghofer.edu.au
The Garvan Institute of Medical Research	@garvan.org.au
The George Institute for Global Health	@georgeinstitute.org
The Olivia Newton-John Cancer Research Institute	@onjcri.org.au
The Prince Charles Hospital	@health.qld.gov.au
The Sax Institute	@saxinstitute.org.au
The University of Adelaide	@adelaide.edu.au
The University of New England	@une.edu.au
The University of Newcastle	@newcastle.edu.au
The University of Notre Dame Australia	@nd.edu.au
The University of Queensland	@uq.edu.au
Walter and Eliza Hall Institute	@wehi.edu.au
Torrens University Australia	@torrens.edu.au
Townsville Hospital and Health Service	@health.qld.gov.au
Trans Tasman Radiation Oncology Group	@trog.com.au
University of Canberra	@canberra.edu.au
University of Melbourne	@unimelb.edu.au, @student.unimelb.edu.au
University of New South Wales	@unsw.edu.au, @ad.unsw.edu.au, @student.unsw.edu.au

University of South Australia	@unisa.edu.au, @mymail.unisa.edu.au
University of Southern Queensland	@usq.edu.au
University of Sydney	@sydney.edu.au, @uni.sydney.edu.au
University of Tasmania	@utas.edu.au
University of Technology Sydney	@uts.edu.au, @student.uts.edu.au
University of the Sunshine Coast	@usc.edu.au, @student.usc.edu.au, @research.usc.edu.au
University of Western Australia	@uwa.edu.au, @student.uwa.edu.au, @research.uwa.edu.au
University of Wollongong	@uow.edu.au, @uowmail.edu.au
VCS Foundation	@vcs.org.au
Victor Chang Cardiac Research Institute	@victorchang.edu.au
Victoria University	@vu.edu.au
Wesley Medical Research	@wesleyresearch.com.au
Western Sydney Local Health District	@health.nsw.gov.au
Western Sydney University	@westernsydney.edu.au
Woolcock Institute of Medical Research	@woolcock.org.au

---

[1] Fgenesh++ package uses the following public software and data:

- NCBI BLAST+ blastp or BLAST blastall, bl2seq programs;
- NCBI NR database (non-redundant protein database);
- NCBI RefSeq database.

[2] Fgenesh++ package includes the following copyrighted software:

- Fgenesh - an accurate and fast HMM-based gene prediction program;
- Fgenesh+ - gene prediction program that uses similar protein information;
- gene finding parameters for Fgenesh and Fgenesh+;
- est_map - program for mapping known mRNAs to genome (genome alignment with splice sites identification) and mapping a set of available ESTs to improve the gene prediction accuracy and add 5' and 3'- noncoding sequences;

- prot_map - mapping protein database to genomic sequences;
- perl scripts.

[3] The Australian BioCommons initiative is hosted by The University of Melbourne.

[4] [Softberry, Inc](). own the software Fgenesh++. The Australian BioCommons currently hold a license for the use of Fgenesh++ to supply access to bona fide Australian researchers via the Australian Fgenesh++ Service.