Bioinformatics Core Facility

Summary

The Bioinformatics facility comprising two Bioinformatics experimental officers and an academic lead aims to support research across the Faculty of Health and Medical Sciences by providing advice in experimental design and costing for research bids, and skills in data processing, handling and analysis. This document outlines: 1) the terms of reference for the facility, 2) the skills and resources available, and 3) access to the facility, 4) costing and types of services, 5) the prioritisation of tasks, 6) Contact of the head of the facility.

1. Terms of reference

Staff members of the Bioinformatics facility will:

- Provide experimental design, data processing and handling, and analysis support to researchers accessing the Bioinformatics facility as outlined in Section 3.
- Distribute and maintain analysis pipelines, software packages and software containers including associated documentation for routine (frequently requested) analyses.
- Provide training to researchers performing routine tasks/analyses and/or where there is a demand for a given topic, provide half-day workshops.
- Assess developments within the field of high throughput data analysis (e.g. processing standards of single-cell RNA-seq) via regular literature searching and attendance of conferences. Significant changes, changes likely to have an impact on any existing or future pipelines/SOPs,will be reported back to researchers accessing the facility via written reports.
- Liaise with faculty IT to maintain the required computing and storage hardware and software of the Bioinformatics server(s).

Bioinformatics core facility is not responsible for storing or backing up users' data. Researchers must seek their own long-term solution for the storage and management of their data according to what they define in their data management plan. Data will be moved to the Bioinformatics server(s) once the required tasks have been completed and will be deleted upon project (task) completion.

The head of the Bioinformatics facility will:

- Act as line manager for the staff members of the Bioinformatics facility.
- Regularly meet with the staff members of the Bioinformatics facility to ensure the needs of the team and the faculty are met.
- Provide guidance on the development of pipelines and SOPs that are of interest and/or strategic priority to the faculty.
- Consult and liaise with researchers and members of the Bioinformatics facility to address identified issues in the provision of data analysis e.g. issues in the prioritising of work and/or misunderstandings.
- Provide costings or facilitate costings for grant bids which require support from the Bioinformatics facility.
- Oversee and manage the budgets associated with the Bioinformatics facility.
- Oversee and manage the software licences associated with the Bioinformatics facility.

- Oversee the 'billing' for work carried out to external (to the faculty) partners/contractors.
- Prepare reports on the activity of the Bioinformatics facility, as requested by HoS.
- Provide justification for staff additions and/or further training of staff members of the facility.
- Liaise with faculty IT to ensure the maintenance of the required computing and storage hardware and software of the Bioinformatics server(s).

2. Bioinformatic resources and skills available

Resources available to the Bioinformatics facility include:

- Two FT staff members.
- 8 Linux servers comprising 2 - 64 CPUs, 1-128GB RAM and 16 -15000GB of storage, with this capacity increasing over time, based on the contributions from research grants.

Experienced in processing of data from:

- DNA- and RNA-seq, including Genome assembly, GWAS, SNP analysis, differential expression, peak identification, metagenomics
- Microarrays, including differential expression, ChIP-on-chip, genotyping/SNP
- Metabolomics data from mass spectrometry instruments
- Phenomic data from Biolog
- Clinical measurement data

Analytical skills:

- Exploratory data analysis (incl. PCA, descriptive statistics, histograms, boxplots)
- General Data mining (incl. Clustering, multidimensional scaling, artificial neural networks, fuzzy logic)
- Data visualization (incl. custom made and standard visualisation plots)
- Statistical analyses (incl. mixture models, ANOVA, Kaplan-meier estimate, power calculations)
- General sequence analysis and manipulation (alignment, motif search)
- Machine learning and classification (incl. short time-series, linear models)
- 'Systems analysis' (flux balance analysis of genome scale metabolic networks (GSMNs), and general modelling and simulation using tools such as SurreyFBA, QSSPN)

Note: The skills highlighted above are not exhaustive and the team welcomes the chance to expand and enhance their skills to support the faculty. If you would like to discuss your own data/required skills please do get in touch at bioinformatics@surrey.ac.uk.

3. Access

Requests for bioinformatics support can be sent to bioinformatics@surrey.ac.uk, where the team endeavour to respond to emails within 24 hours during the working week. Note that requests to support a new project should come from the PI, where information regarding the source of funding and allocation to the facility (if applicable) is required. This information is used to prioritise requests (see below), and facilitates the monitoring of the facility's activities and budget.

All publications or presentations that include work that has been supported by the bioinformatics facility must acknowledge the facility according to the following format: "I/We thank the Bioinformatics facility at University of Surrey; [name of staff member] for [tasks/analyses performed]".

4. Costs and Types of Services

The bioinformatics provides the following types of services

- *Consulting advice*. Consulting and guidance regarding bioinformatics and data analysis to FHMS staff. This includes sharing software, training, advice running supported software, support designing analysis plans, troubleshooting analyses. In this modality we do not run analyses on behalf of the user. First 8h free, after that charged according to the nominal technician rate.
- *Projects*. The planning and execution of detailed analysis plans. The facility elicits and conducts the analyses agreed in the analysis plan on behalf of the principal investigator. In the end of the project, the facility provides the following items: i) raw data and processed results in file format, ii) standard figures, and iii) methods in the SMART* format. Writing papers, hand-drawing figures and formatting tables are outside the remit of the facility. The payment of the cost is made by a journal transfer to the FHMS or included in the budget of a new proposal, in which case execution will resume once the proposal is granted. First 8h free.
- *Clinics*. Online or face-to-face sessions providing consulting advice to FHMS staff and post-graduate students. First 1h free, after that charged according to the nominal technician rate.
- *Training*. Online workshop sessions on how to use the pipelines and packages developed by the core facility. Free for FHMS staff and students.

5. Prioritising Requests

The Bioinformatics facility prioritises and categorises requests based on the following definitions:

I. Request related to a funded project work, highest priority. Where funds have been directly allocated to the Bioinformatics facility, either as a percentage ofthe team's time included in a grant or directly costed and paid for via work to external (non- RCUK) funders.
II. General technical advice, medium priority. Requests related to general technical advice, such as how to run specific software and/or which statistical analysis is the most appropriate, will receive a timely response (within 24 hours duringthe working week).

III. Request related to a future project: research proposal support, medium priority. Power

calculations, Preliminary data analysis and Experimental design, where the outputs of such work will directly support the submission of a bid. Timing of such requests should be no less than 4 weeks in advance of bid submission. The costing of these requests follows the project scheme

To ensure fair management, allowing for all groups to receive efficient and effective service, the team require that PIs communicate with the team so that:

1) The team are aware that funds have been awarded and thus a percentage of the team's time needs to be allocated to the project.
2) In advance, the team are aware of the likely date that the data will become available, project starts and/or set deadlines for given objectives to be completed.

Those planning on submitting a bid for a project that will rely on bioinformatics analysis/expertise should discuss their requirements at least four weeks ahead of grant submission. This step will ensure data can be handled effectively and results are generated within submission deadline. Failure to engage with the bioinformatics team, to discuss future projects and/or experiment, may result in the team's inability to analyse the data.

6. Contact

For further information, queries on your requests, as well as input and feedback on service provided, please contact head of the bioinformatics core facility at: a.coutoalves@surrey.ac.uk.