

Plant and Animal Genomes Conference, San Diego, 13th-18th January 2023

Functional Annotations of Animal Genomes (FAANG) Workshop, Friday 13th January 2023:

Implementing the Next Phase of FAANG (Round Table Discussion) Notes

Each Task Force (TF) was asked to discuss the following items to prepare a point or two that they view as highest priority to raise in the general session:

- A. Someone to handle communications and meetings coordination
- B. Someone to organize seminars to learn about new developments in your field
- C. Broadening scope by integrating expertise from other areas, e.g. model organisms/plants.
- D. Adding expertise that is currently missing in the Task Force and needed to make progress, e.g. by developing interactions with other TFs.
- E. Infrastructure and funding to generate data and submit data/metadata to the FAANG data portal
- F. Infrastructure to coordinate data visualization and meta-analysis across groups
- G. Coordinate and organize opportunities for training in new wet-lab or computational training within the task force or within FAANG.
- H. YOUR PRIORITY which is....

- The session started with taskforce leader introductions.
- Priorities were listed and discussed from an earlier discussion in the workshop:
 - MetaFAIR
 - Further reporting of mandatory pipelines and training.
 - Meeting in March: Looking to change to required containerised pipelines and a parametrised file to make these pipelines reproducible.
 - There is a gap in training and knowledge of track hubs. There is a need to find knowledgeable people to train others in using track hubs and to share knowledge.
 - FarmGTEX
 - Data generation, infrastructure and funding to generate data.
 - What kind of data is being generated? Was a major question.
 - Data is being generated for different species.
 - People within the FAANG community want to join FarmGTEX, but are not receiving emails. The group is planning to organise a sign-up sheet for people to join the task force.
 - Prediction
 - Data generation, infrastructure, funding to generate data, broadening scope by integrating expertise and training.
 - Three of the organisers for this section were present in the session.
 - A major goal is to incorporate farm data as routine practice and to share data across industry, which will be imperative to data access.
 - The question 'Can a national effort such as phenotyping be set up?'
 - The need to incorporate other areas of expertise that complement the expertise that is already present was identified. This new expertise included molecular geneticists and computer scientists.
 - HTP-DS Infrastructure

- A survey will become available in a month and will ask what people use to generate data.
 - Ontologies will be discussed with MetaFAIR to make descriptions easier and more usable.
 - A need to find a suitable method to collate data and make data more reusable in the future was identified.
 - Comparative Genomics
 - Welcoming new members.
 - An infrastructure to coordinate data visualisation is required, with tools to load multiple genomes and FAANG tracks needed.
 - A majority of this sort of work is already occurring in humans and mice. It was suggested if the community could work or make connections with the human and mice fields to gain knowledge on what pipelines they are using and how to integrate these into farm animal research.
 - Single Cell
 - Data generation and infrastructure.
 - Data generation has been identified as a big issue. Data standards are not well defined or established. Infrastructure has also been identified as a big issue, with sharing and storing the data generated.
 - The question, 'What are the standards we need to visualise this data?' was raised.
 - Training is needed for data submission.
- Recruiting is an issue within the FAANG area of research. The animal genome community faces challenges of people moving to human and biomedical research due to funding or other factors. More funding needs to be generated and/or obtained.
 - Opinion from an audience member:
 - The FAANG community needs to show future recruits a good future for their career within the field.
 - How do we attract students from not taking other animal scientist roles, such as nutritionists etc.
 - Animal genetics is a 'hard science' and is multi-factorial. How can the culture provide a better future for future scientists? Will funding be a key factor?
 - It is important to have undergraduates in the field early for research experience and fellowships for pre- and post-doctoral studies (in the USA only, so leaves out international students).
- Discussion points
 - Are additional task forces required?
 - Yes, perhaps regarding pipelines and epigenetics (linking variants and expression).
 - Pipeline task forces could fit within HTP-DS Infrastructure.
 - Important to share protocols, metadata and pipelines for datasets (e.g. single cell omics):
 - Should be tackled as a community.
 - Who will do what, and when?

- Groups don't necessarily describe their pipelines to the community very well. Euro-FAANG have a mandate to describe their pipelines and are doing it very well. The rest of FAANG need to follow suit.
- If working groups can put together recommendations for protocols and share these recommendations, people will be able to see what is required for each genomic technology. This will future-proof protocol recommendations.
- It was suggested that a common location needs to be assigned to place bioinformatics pipelines as they are not currently available unless accessed via papers.
- A guided decision tree for data generation in different technologies is being discussed in Euro-FAANG. It is not in use currently, but perhaps needed in the future as knowledge needs to be shared equally. Task forces are needed for each technology to guide other technologies, such as single-cell sequencing.
- Goals of large-scale projects and how to harmonise at a global FAANG scale:
 - Euro-FAANG will organise a global FAANG event at the conclusion of BovReg or PAG next year in 2024. Obtaining fully associated partners with Euro-FAANG, such as the way in which New Zealand is, would be a great way to diversify and become more global.
 - Does FAANG compliment AG2PI? How can this be moved forward? How does FAANG move in the future with partners and funding? As there is no core funding for FAANG presently.
 - The speed at which data is being generated and how to store this data is becoming a challenge.
 - FAANG needs to give updates on what has been achieved more clearly, so the community can understand what has been successful and what needs improvement.
 - A white paper could be useful in doing this, which could then be offered to the community. This paper could also work favourably for obtaining future funding. The paper could include:
 - What datasets have been generated, metrics for these datasets, sequences and variants identified and what protocols have been established.
 - Could possibly be a prospective instead of a paper? Colleagues are probably more interested in how they can access the data than how many variants etc. have been identified.
 - FAANG is about 80% complete, so it may be premature to publish a paper on what has been achieved so far.