

Minutes Steering 13 December 2016 (8 AM CDT)

On call: CKTuggle, MWatson, HCheng, GPlastow, SWhite, GTosser-Klopp.

Regrets: MGroenen, LClarke, PHarrison, HZhou, AArchibald, EGiuffra

Minutes from 8 Nov 2016 call- approved

Announcements

Reports from Standing Committee Chairs

- Animals/Samples/Assays:

Davis group gave presentation on CHIP-seq- some information on website.

- Metadata and Data Sharing:

Email report from Laura: There continues to be small iterations on the experimental metadata standards and various Roslin people use them to submit data. These have all been announced on the metadata list. There are three submissions labeled as FAANG in the system at the moment. None of them meet the experimental metadata standards. We will be contacting the submitters to discuss this.

We hope to have the first version of the sample data tables for our portal ready for PAG.

We will be recruiting a person to help with this in the new year. If you know of possible candidates, do let me know.

Metadata committee will be meeting at PAG at 4 PM Monday – room is Eaton at Town & Country.

- Bioinformatics and Data Analysis:

Mick Watson reported on BDA groups (Thanks (!!)) for the written notes, Mick!

RNA (LeI Eory)

Focused on 3 main areas

- Annotation / transcript discovery
- Promoter / TSS discovery using CAGE
- Gene expression quantification

Transcript discovery based on Illumina

- Tested mappers BWA, TopHat, STAR, HISAT, GMAP
- Tested “assemblers” Cufflinks, StringTie, Ensembl
- Current choice: STAR+Cufflinks

Transcript discovery based on PacBio

- The group reports limited data availability
- Plans comparison of PacBio and Illumina

Small RNA

- Planned but no activity

A LincRNA hackathon has taken place

- Based on STAR+Cufflinks
- FEELnc used to identify LincRNA

CAGE pipeline

- Based on BWA and CAGEr
- (no further details)

Gene expression quantification

- Reference based: RSEM (using ENCODE pipeline)
- Reference-free: Kallisto
- (no further details)

Structure and 3D (Sylvain Foissac)

Focusing on three assays

- DNase-Seq
- ATAC-Seq
- Hi-C

Preliminary pipelines have been developed for Hi-C and ATAC-Seq

Have had problems with larger number of chromosomes in livestock genomes of human and mouse

QC metrics now include insert size, implemented in HiC-Pro

Feedback has been provided to the ASA committee RE experimental protocols

Data used in this work has been the FR-AgENCODE pilot project data (pig, cow, chicken, goat)

- Communication:

-call December 6:

1. Further discussion/organization of future FAANG meetings

a. PAG 2017 FAANG Workshop

b. Friday afternoon discussions on FAANG proposals to AFRI

c. ISAG 2017. Preparations are unknown to those on call.

Chris will speak to Jim Reecy about organization of

Sequencing/Annotation Workshop to offer help from FAANG. Elisabetta

asked whether COST would be holding a meeting at ISAG. Laura

indicated it was possible now that the EC has provided clarity on budgets for the next period.

2. Discuss funding opportunities website/other means to communicate + collaborate

a. website update for COST and AG2P groups: The homepage now has clickable icons for the NSF AG2P and the COST FAANG-Europe projects. The Funding opportunities link is now renamed Funding and Projects and is reorganized with opportunities listed briefly on left and funded project at right. We will add the two pilot projects and others as information provided (Australia, Genome Canada).

b. Twitter records on homepage of FAANG website: too little traffic?

c. Connecting the EU and US training efforts: Elisabetta is discussing with Carl and Bo Thomsen the development of a list of reference labs. They are planning to work with the DCC to identify those labs submitting samples and data. It was discussed that bioinformatics labs that have submitted bioinformatic pipelines would also be good reference labs for training, and that the DCC could also potentially help with identifying those who have submitted such pipelines.

Old Business

1. Monthly newsletters? *Any volunteers to provide content will be welcome.*

2. Further discussion on working together to foment increased opportunities for FAANG-related projects H2020, FFAR, etc.

- see notes above; the email comments on too little traffic was echoed on the call; this will be re-thought.

New business

None

Future Steering Committee calls- frequency and time of day to discuss

It was agreed that calls as follows would suffice for the first half of 2017:

Early Feb

Early April

Early June

Further, those on the call were comfortable with the approximate time of mid-morning US Central time.

Chris will send out a Doodle, looking for 8, 9 or 10 AM US Central times in the first week or so of Feb, April, and June.