ELIXIR Training activity Proposal

ELIXIR/GOBLET hackathon for metagenomic training material re-use

Object of the request: This training activity will be the first joint activity for GOBLET and ELIXIR since the publication of our joint strategy. Funds are requested to host a two-day workshop within an ELIXIR Node to work on the curation of a set of metagenomics training materials to ensure their reusability. It will bring together trainers and researchers working in the field of metagenomics from within Europe and beyond, not only to enhance the training network within the field of metagenomics, but also to further strengthen the relationship between ELIXIR and GOBLET. The output of the workshop will further be of benefit to the wider bioinformatics training community, by providing best practice for developing materials that can be widely re-used, and guidance on using this material to build training courses. We anticipate that the workshop would additionally be of benefit to EXCELERATE, specifically Work package 6 "*Marine metagenomic infrastructure services as driver for research and industrial innovation*". As one of the final outcomes will be a curated/collated set of training materials for metagenomic analysis, this WP will have an easily accessible resource from which to build appropriate training courses to support its activities. To ensure that the marine context is covered within the spectrum of this material, involvement of members of WP6 with training materials in this area is key.

Organisers: Teresa Attwood (UK/GOBLET), Pedro Fernandes (PT/GOBLET), Sarah Morgan (EBI/ GOBLET), Gabriella Rustici (UK/GOBLET), Celia van Gelder (NL/GOBLET) and Allegra Via (IT/GOBLET)

Motivation:

Metagenomics relies heavily on the use of NGS technologies to analyse microbial communities from a range of settings (*e.g.*, marine, environmental, agricultural, medical). Although there are training materials available in the area of NGS analysis, these are often developed and tested on model organism data. Often, however, metagenomic researchers work with microbial genomes that have not previously been, or are rarely, studied. This presents several significant challenges for both data management and bioinformatic analysis, a situation that is increasingly being noted in training course applications, where trainees are seeking guidance on working with non-model organism data.

A previous workshop (held in Cambridge in January 2015, supported by ELIXIR-UK) was successful in developing a resource for NGS training, based on the curation of existing training materials. These materials were provided by 25 trainers from across the world (but focused in Europe), and encompassed aspects of several NGS workflows (RNA-Seq, ChIP-Seq and variant analysis), but uniquely focused on the analysis of model organism data-sets.

The workshop proposed here would focus on the curation of training materials in the area of metagenomics, encompassing appropriate NGS workflows (genome assembly, annotation and analysis) and related subjects, such as phylogenetics and comparative genomics. To curate these training materials, the strategy developed during the previous workshop will be applied, aiming to produce a more generic guide to such approaches as a specific workshop outcome (as detailed above).

By hosting this workshop as a joint ELIXIR/GOBLET meeting, we will leverage international expertise to create a comprehensive collection of training materials and to create a networking opportunity for trainers from across the globe, whilst also enhancing and extending existing training and research collaborations.

Workshop objectives:
This action builds on work undertaken in a previous workshop held at the University of Cambridge in January 2015, where materials for NGS training were collected and curated:
(**https://bioinformatics.upsc.se/trainers/**)

- to develop a collection of metagenomics training materials that are consistently described, to enable sharing and re-use of these materials across the training community;
- to define an approach for the curation of training materials for re-use through training portals and repositories such as TeSS and GOBLET;
- to provide guidance and #n the form of a paper/documented process to be disseminated across ELIXIR nodes and beyond.

Workshop outcomes:

- a collated/curated set of metagenomics training materials that is accessible to the whole training community. **Date for delivery: 3 months post workshop**
- a working framework for curation of training materials (including best-practice guidelines) for trainers wishing to make their materials available via the TeSS/GOBLET portals. **Date for delivery: 6 months post workshop**
- a best-practice paper/report to include both the framework for material curation and guidance on the use of such material in the development of new training course, allowing dissemination of the approach to all ELIXIR nodes and a wider training audience. **Date for delivery: 6 months post workshop**

Workshop description:
Dates: 2-day workshop, dates TBC
Location: venue TBC, will be hosted by one of the organising ELIXIR Nodes.
Audience: open to all Nodes and GOBLET members. A mix of invited trainers with known expertise in metagenomic analysis, members of WP6 (marine metagenomics use case) who are actively involved in training, and an open invitation to other interested individuals through ELIXIR nodes and the GOBLET consortium.

Draft agenda

Day 1 (9:00 - 17:00)
Welcome and introductions (timings x presentation?)
Presentation of curation approach

Identify similarities/differences in teaching materials - assign groups
Coffee break
Group assignment – curation hackathon
Lunch
Group assignment – curation hackathon (cont.)

Day 2 (9:00 - 17:00)
Group assignment – curation hackathon (cont.)
Presentation/discussion of curation hackathon
Lunch
Defining generic approach to curation
Discussion on course design methods

Budget: (€50,000 cap)
No. of participants: 30 max (including organisers)
Direct workshop costs:

- Venue hire: €500-1,000
- Refreshments: coffee breaks x 4; lunch x 2 €1,000 for both; dinner x 1 €800
- Hotel accommodation for 2 nights €6,000
- Cost of transfers

Participant travel costs:

- 20 Europeans €530 per person: €10,600
- 10 outside Europe €1,800 per person: €18,000

Grand total: €37,400

Contacts

Working part of the document:

Prospective trainers to invite:
Titus Brown or **Adina Chuang Howe** (Iowa State University)
Harvard Medical School representative - Radhika Khetani or colleague
Alex Mitchell (EMBL-EBI)/Rob Finn (EMBL-EBI) - **well worth asking Alex (I'm sure he's done training in this area)**
Australian delegate(s) from Bioplatforms Australia
Paul Greenfield, CSIRO (Australia)
Brazilian delegate(s) from Belo Horizonte/Brasilia - Gabriel Fernandes, Brasilia
Micro B3 consortium - Frank Oliver Glockner?
Bas Dutilh or Sacha van Hijum (NL) - I have to check with them first, see also below

Is your Node well connected within the metagenomic community. Do you know of people that would like to attend this workshop?
(please provide your answer after your country)

Spain
University of Malaga (ES) is collaborating on a metagenomic analysis project with LNCC in Brazil, looking at the respiratory tract of healthy and diseased pigs. As part of this project, a number of tools have been developed that could be of use. Additionally, it is possible that people involved in this project may be interested in attending.

Greece
The Hellenic Centre for Marine Research (GR) is member of the Greek node and its Institute of Marine Biology and Genetics has performed metagenomics analysis of the largest to date submarine volcanic ecosystems dataset. People to contact:
Antonios Magoulas (Director) magoulas@hcmr.gr
Georgios Kotoulas (Research Director) kotoulas@hcmr.gr
Anastasis Oulas (Research fellow) oulas@hcmr.gr

ITALY
CNR (National Research Council) - Istituto di Biomembrane e Bioenergetica and Institute of Biomedical Technologies (Research Area: Molecular Biodiversity)

- Bruno Fosso (Research fellow) b.fosso@ibbe.cnr.it
- Monica Santamaria (Researcher) m.santamaria@ibbe.cnr.it

University of Florence, Dept. of Neuroscience, Psychology, Pharmacology and Children's Health

- Duccio Cavalieri (PI) duccio.cavalieri@fmach.it

University of Trento, Laboratory of Computational Metagenomics Nicola Segata (PI) nicola.segata@unitn.it

Switzerland

There are various SIB members involved in metagenomics projects, but I will have to identify the one(s) available and willing to participate. Once this project is accepted, and you tell me how many people from each node can/should attend, I will send out the invitation within the SIB.

Portugal

Portugal has run a Metagenomics training course in 2013 (15 Participants)
http://gtpb.igc.gulbenkian.pt/bicourses/2013/AWGSMD13/
with Nick Loman, Chris Quince, Anders Andersson and Josh Quick, and has planned a revised version for 2015 (TBA).
Pedro Fernandes (pfern@igc.gulbenkian.pt)

In a brainstorming meeting held here in March 2014 for COST action Seqahead, specific training in microbial genomics and metagenomics was discussed by Nick Loman (Birmingham)
http://gtpb.igc.gulbenkian.pt/bicourses/TMNDA/
Nick referred to the investment in the CLIMB infrastructure, which in the meantime became operational, http://www.climb.ac.uk/, which has its own ongoing training activity.
I believe it would be interesting to consider their activity in this scope; maybe we can interact with them using our contacts via ELIXIR-UK? **Andreas Heger**, MRC CGAT, Oxford, UK was here for that meeting. Even if CLIMB is not in connection with ELIXIR, I guess it makes sense to connect with their staff for training purposes.

UK

Mick Watson/Bert Overduin Edinburgh Genomics might be interested, and would know some interested people. Vicky Schneider (TGAC) <Vicky.Schneider@tgac.ac.uk> is well connected with the metagenomics community. Alex Mitchell (EBI) - see above.

Finland

CSC provides support only in marker gene metagenomics (like 16S rRNA NGS). There are research groups in Finland who do full-scale metagenomics, but they are not involved in training and will hence not have training materials to curate.

France

There are research groups in France involved in marine metagenomics: Christophe Caron in Roscoff and Olivier Collin in Rennes. Claudine Medigue in Paris studies prokaryotic metagenomes. Christine Gaspin in Toulouse also develops metagenomics approaches.

Slovenia

ELIXIR-SI is a member of the National Institute of Biology - NIB (department Marine Biology Station) actively participated in Ocean Sampling Day, which is dedicated to metagenomics of microbiota present in the oceans http://www.microb3.eu/osd. ELIXIR-SI (NIB in collaboration with Faculty of Medicine in Ljubljana and Maribor and other members) has particular interest for metagenomic of different planktonic organisms (*e.g.*, toxic species of phytoplankton and zooplankton in general). We would like to participate in training and we are willing to participate in field sampling, either at our region or abroad; we can also offer our facilities. There is great interest in the Adriatic region to incorporate metagenomics approaches in marine research. Contact: Andreja Ramšak (andreja.ramsak@nib.si).

Czech Republic

There are a number of groups involved mainly in microbial metagenomics, and the representatives of these groups are also involved in teaching (Baldrian, Uhlik, Pospisek). Czech Republic will find a representative of these groups to attend the hackathon.

Denmark

Jon Ison is interested in participating in this workshop (jison@ebi.ac.uk).

The Netherlands

I expect we can offer both teachers as well as participants from NL.

- Bas E. Dutilh co-founded the Netherlands Metagenomics Platform (NMP, http://www.metagenomics.nl) with Sacha van Hijum (also mentioned above as prospective teachers). The NMP has been active since 2013, organising meetings with specific interest topics such as marine metagenomics or viral metagenomics. Moreover, NMP also co-organises courses/workshops (with NBIC/BioSB or The Netherlands Institute of Ecology, NIOO). Courses cover the basic bioinformatic metagenomics skill-set (e.g. taxonomic and functional profiling, statistics, experimental design), but also more specific topics (e.g. comparative metagenomics, metagenome assembly). A yearly recurring course on Metagenomic Approaches and Data Analysis has been running since 2013 (part of the NBIC/BioSB course portfolio): http://biosb.nl/education/course-portfolio/course-metagenomics-approaches-and-data-analysis/. In 2014, we organised a 5-day course & symposium in collaboration with NBIC/BioSB and NIOO (sponsored by AllBio): Environmental Metagenomics Course & Symposium, 15-19 September 2014, Wageningen, NL (https://nioo.knaw.nl/en/metagenomics-course and https://nioo.knaw.nl/en/metagenomics-symposium).
- Many research groups active in metagenomics participate in the Netherlands Metagenomics Platform (NMP):
  - Utrecht University (UU) and its Utrecht Bioinoformatics Center (UBC, http://ubc.uu.nl/community/allgroups) have a strategic alliance with The Netherlands Institute for Sea Research (NIOZ) that includes a strong focus on Bioinformatics for Marine Metagenomics (http://pers.uu.nl/universiteit-utrecht-gaat-participeren-in-nwo-instituut-nioz).
  - The Netherlands Institute for Sea Research (NIOZ)
  - Wageningen UR
  - Radboud University Nijmegen / Radboudumc
  - University of Amsterdam / Institute for Biodiversity and Ecosystem Dynamics (IBED)
  - Free University Amsterdam (VU) / Academic Center for Dentistry Amsterdam (ACDA)
  - The Netherlands Institute of Ecology (NIOO)
  - (… and more)

TrCG that would like to attend the workshop:

Alex Upton (ES)
Pedro Fernandes (PT)
Rita Hendricusdottir (UK, as an observer)
Brane Leskošek (SI)
Marian Novotny (CZ)

Sarah Morgan (EBI)
Celia van Gelder (NL, as observer, but also depending on date/time and on other NL representatives)
Gabriella Rustici (no TrCG but happy to participate) - will bring UoC perspective.
Terri Attwood (UK)
Andreja Ramšak (SI)

FInal Report

Summary

ELIXIR Training Platform and GOBLET (Global Organisation for Bioinformatics Learning, Education & Training) organised a joint workshop for metagenomics training materials re-use, on 7-8 April 2016 at EMBL-EBI in Hinxton, UK.
The aim was to collect a comprehensive set of metagenomics training materials and define a common approach for their curation. Such materials can then be exchanged and re-used via online portals and repositories, such as GOBLET and TeSS.
The worskhop brought together 19 participants form ELIXIR Nodes (Denmark, EMBL-EBI, Italy, Netherlands, Portugal, Slovenia, and UK) and from Germany, Australia, Brasil, Canada and USA. The participants identified existing training materials available for each metagenomics workflow and described them, using a common set of descriptors.
The Training Platform will soon release the curated materials on GitHub.

Attendee list