THE WORKFLOW HUB

A REGISTRY FOR LIFE SCIENCE WORKFLOWS

http://www.workflowhub.eu

Carole Goble[1], Finn Bacall[1], Stian Soiland-Reyes[1], Stuart Owen[1], Alan Williams[1], Ignacio Eguinoa[2], Bert Droesbeke[2], Hervé Ménager[3], Laura Rodriguez Navas[4], Kiran K Telukunta[2], Flora D'Anna[2], José Mª Fernández[4], Alban Gaignard[5], Ambarish Kumar[6], Maciek Bak[7], Ivan Topolsky[8], Salvador Capella[4], Michael R. Crusoe[9] and Frederik Coppens[2].

The Workflow Hub

A new registry that is Workflow Management System (WfMS) agnostic.

Workflows may remain in their native repositories in their native forms.

Workflows makers are custodians of their own workflows.

Holds related supplementary information: test and example data, SOPs, presentations etc

Workflows can be crated into collections

Metadata is generated by and harvested from WfMS. Galaxy, Snakemake and Nextflow are working with the Hub to seamlessly support packaging and registration.

Metadata standards

Bioschemas.org   about a workflow

COMMON WORKFLOW LANGUAGE   describe the workflow

Workflow RO-Crate   organise & package the workflow components

Yellow pages
Spaces
Teams
Organizations
People

Assets
Data files
SOPs
Workflows
Publications
Documents
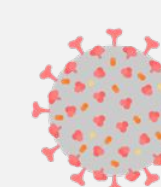Collections

Activities
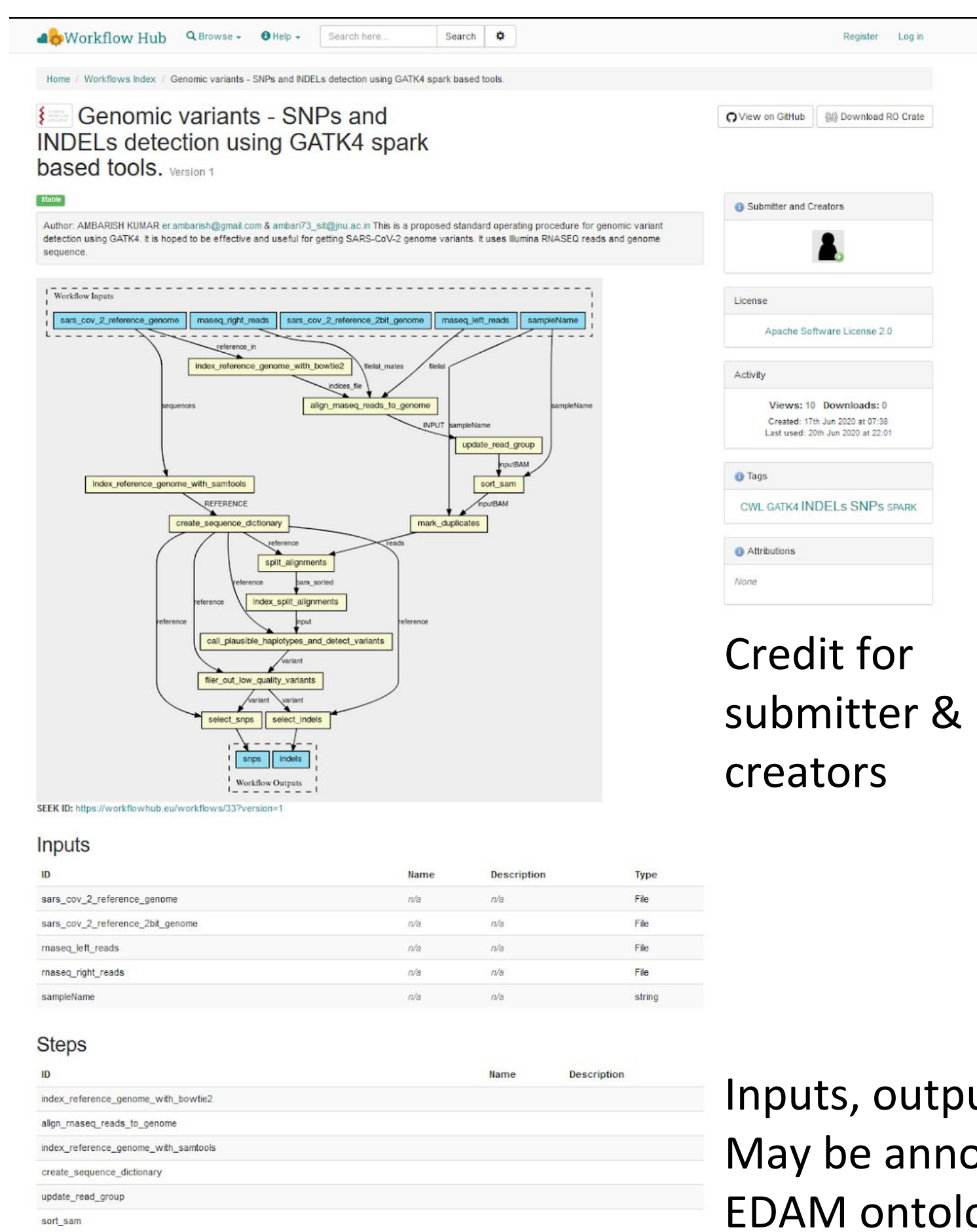Presentations
Events

Samples
Organisms

A Workflow Zoo…

Computational processing using workflows has taken hold in the life sciences. There are 280+ Workflow Management Systems in operation [1]. Workflow registries and repositories typically cater for specific workflow management systems and there is little standardisation of the metadata. But just like data, workflows should be FAIR - Findable, Accessible, Interoperable, Reusable [2].

Responding to COVID-19   Workflow Hub

The Hub is due for Beta release in September 2020. However, the team fast-tracked to open in April 2020 to take COVID-19 workflows. Consequently the Hub is being perpetual developed in the open. So we are rubbing off the rough edges with our users. The Hub is registered on the EU Covid-19 Data Portal (https://www.covid19dataportal.org/).

Workflows presented as CWL or in native image and form

Credit for submitter & creators

Inputs, outputs and steps May be annotated using the EDAM ontology

GitHub   → Visit source

Workflows can be viewed at their github or at their source – e.g. Galaxy instance.

Workflow RO-Crate

Workflows can be uploaded / downloaded as Research Objects

Or registered as collections of files

The Hub was established during the COVID-19 Virtual Biohackathon April 5-11 2020.

The Hub is an instance of the FAIRDOM-SEEK commons infrastructure [3].

How do I get involved in Workflow Hub?

- Join the Workflow Hub Club - an open community that works together to develop the Hub.

- Contribute your Workflows – the Hub Club will help you. https://about.workflowhub.eu/

References
[1] https://s.apache.org/existing-workflow-systems
[2] Goble C et al FAIR computational workflows. Data Intelligence 2(2020), 108–121. doi: 10.1162/dint_a_00033
[3] Wolstencroft K et al SEEK: a systems biology data and model management platform, BMC Systems Biology, 9:33, 2015. doi:10.1186/s12918-015-0174-y.

Affiliations:
[1]The University of Manchester, [2]VIB-UGent Center for Plant Systems Biology, [3]Institut Pasteur, [4]Barcelona Supercomputing Center, [5]Université de Nantes, [6]Jawaharlal Nehru University, [7]University of Basel, [8]ETH Zurich, [9]The Common Workflow Language.