



Multiscale
Complex
Genomics



Multiscale Complex Genomics



Project Acronym: MuG

Project title: Multi-Scale Complex Genomics (MuG)

Call: H2020-EINFRA-2015-1

Topic: EINFRA-9-2015

Project Number: 676556

Project Coordinator: Institute for Research in Biomedicine (IRB Barcelona)

Project start date: 1/11/2015

Duration: 36 months

Milestone 13: Establishment of open source software development infrastructure

Lead beneficiary: European Molecular Biology Laboratory

Dissemination level: PUBLIC

Due date: 30/06/2016

Actual achievement date: 29/06/2016

Copyright© 2015-2018 The partners of the MuG Consortium





Document History

Version	Contributor(s)	Partner	Date	Comments
0.1	Mark McDowall	EMBL-EBI	20/06/2016	First draft
0.2	Andrew Yates	EMBL-EBI	23/06/2016	Revision, minor comments
0.3	Mark McDowall	EMBL-EBI	25/06/2016	Added conclusions
1.0	Mark McDowall	EMBL-EBI	29/06/2016	Final version

EXECUTIVE SUMMARY

The MuG consortium will provide access to code generated that has been generated to partners in an open and centralised fashion. Github is considered a good approach to allow partners to distribute their tools and pipelines. To that end a GitHub organisation has been set up and is accessible via the MuG website - <https://github.com/Multiscale-Genomics>.

INTRODUCTION

As part of the MuG consortium there is going to be the development of a number of tools and pipelines. Making this code openly available and version controlled to maximise its usefulness and adoption by others. GitHub is seen as an ideal way to collate the tools and make them accessible.

ACCESS

The repository is located on GitHub (<https://github.com/Multiscale-Genomics>). Members of the consortium are able to create new repositories for their tools and development work. Those outside of the consortium are able to clone the repositories and use the code. If there are changes made by a user outside of the consortium then a pull request should be made. It is then up to the owner of the repository to accept or reject the changes.

There is a matching GitHub.io site (<https://multiscale-genomics.github.io>). This provides a list of the repositories, the descriptions and primary language used.

REPOSITORY STRUCTURE

While each tool is different and can follow different development pathways dependent on the developer(s) all repositories will contain a README.md file detailing:

- Overview of the repository
- Dependences on other software or hardware
- How to install the code
- Examples about how to use the software
- Authors and Acknowledgements
- Description of the consortium and link to the main website

CONCLUSION

The use of GitHub as a key part of the software development infrastructure should ensure that the tools developed are openly accessible, versioned and should ensure their long term availability. Adopting standard practices in relation to documentation and versioning should ensure that the tools are easy to install, use and develop by those that are inside and outside of the consortium.