



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

Deliverable 2.2: Project Website

Lead beneficiary: Institute for Research in Biomedicine (IRB Barcelona)

Dissemination level: PUBLIC

Due date: 01/01/2016

Actual submission date: 18/01/2016

Copyright© 2015-2018 The partners of the MuG Consortium



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 676556.

Document history

Version	Author(s)	Partner	Date	Comments
0.1	Anna Montras	IRB Barcelona	08/01/2016	Description of website structure and contents circulated to consortium for additional inputs
0.2	Anna Montras	IRB Barcelona	12/01/2016	Updated contents and screenshots added
1.0	Anna Montras	IRB Barcelona	13/01/2016	Approved by consortium.

Table of contents

1	INTRODUCTION	5
2	WEBSITE BASICS.....	5
2.1	DOMAIN AND MAINTENANCE	5
2.2	DESIGN.....	5
3	WEBSITE STRUCTURE AND CONTENTS	7
3.1	HOME	7
3.2	ABOUT MUG.....	8
3.2.1	<i>Objectives</i>	10
3.2.2	<i>Key Innovations</i>	11
3.2.3	<i>Pilot cases</i>	11
3.2.4	<i>Publications</i>	12
3.2.5	<i>Milestones</i>	13
3.2.6	<i>Deliverables</i>	14
3.3	PEOPLE.....	15
3.4	NEWS AND EVENTS	16
3.5	MUG VIRTUAL RESEARCH ENVIRONMENT	16
3.6	CONTACT	18
3.7	PRIVATE AREA	18
4	FUTURE WORK.....	20

Executive summary

The MuG project website has been launched in January 2016 and its registered domain is www.multiscalegenomics.eu.

The present report contains the details of the MuG website development process, including website structure, contents and visual image. Screenshots of the website are included in the report.

The document also reports on additional features that have been foreseen for the website and that will be gradually implemented in future upgrades, as they are linked to project progress (including newsletters, development of the MuG virtual research environment framework, etc.).

1 INTRODUCTION

The MuG web structure is based on the following approach:

1. **MuG public website:** containing public information for dissemination and communication purposes.
2. **MuG Virtual Research Environment portal:** the gate to the services that will be the main results of the MuG project.
3. **Private area:** internal communication will be handled using the BaseCamp project management tool.

The public website has been officially launched in January 2016. In its design, the website follows the visual image defined by the project logo. The consortium sought professional advice in the design of the logo in order to create a powerful brand for MuG that will enhance communication and give visibility to the project from the early months.

2 WEBSITE BASICS

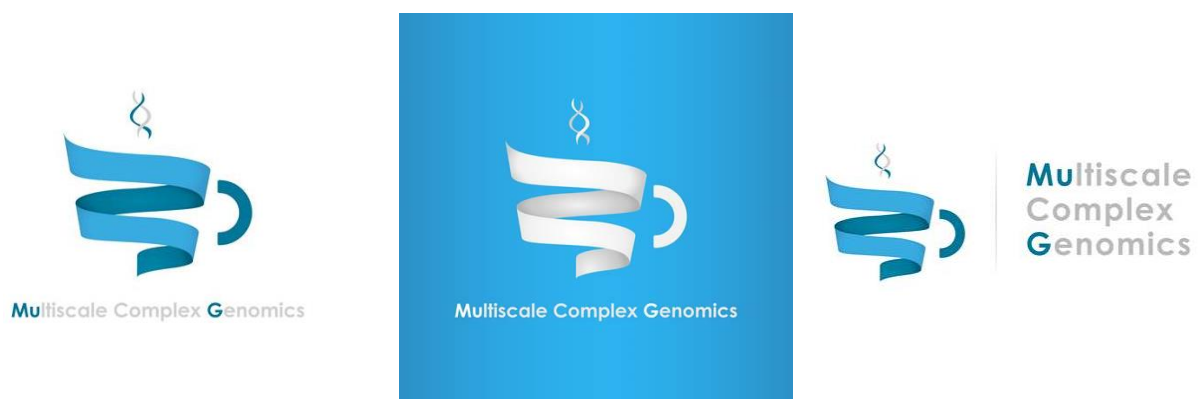
2.1 Domain and maintenance

The MuG consortium partners agreed on the registration of the domain www.multiscalegenomics.eu for the MuG project.

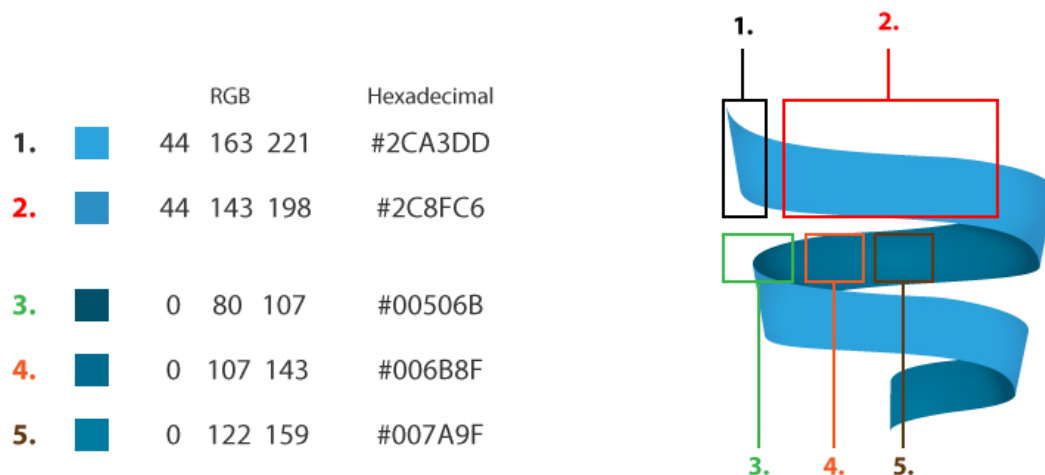
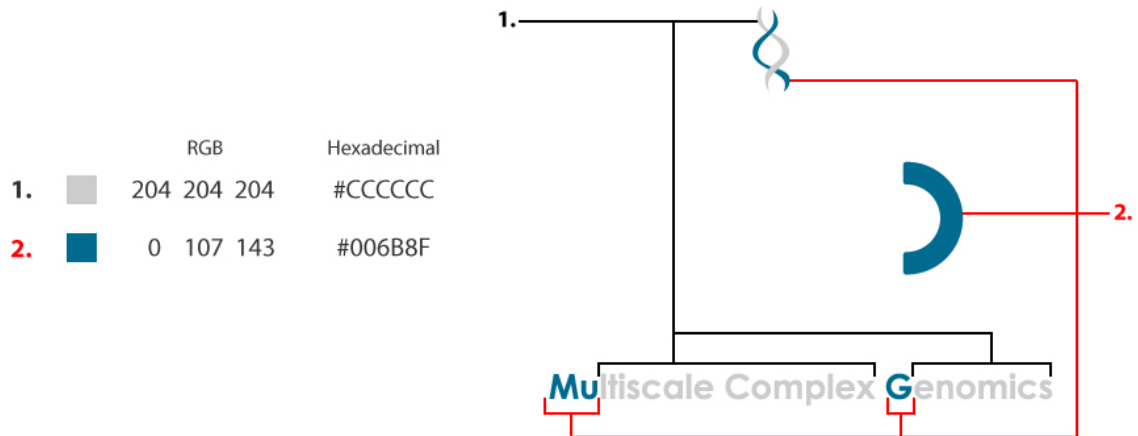
IRB Barcelona, as coordinator, has registered and administers the domain. Website updates and maintenance are the main responsibility of the coordinator. Inputs from all beneficiaries will be required in generating web contents and keeping the information up to date with the latest developments within the project and in the respective fields involved.

2.2 Design

The website design is based on the concept logo developed for the MuG project, which defines the visual image of the MuG project.



The colour codes that define the design presented above, which are the base for the project website image, are as follows:



Project beneficiaries are encouraged to use the MuG project logo and add the link to www.multiscalegenomics.eu in all possible communications and in the websites of their institutions and departments.

3 WEBSITE STRUCTURE AND CONTENTS

The URL of the website is <http://www.multiscalegenomics.eu>

The MuG website is online since January 2016 and will be progressively upgraded to include new features and improve its functionality.

The public area of the MuG website is composed of a home page that shortly introduces the project. From the main page the different options of the menu can be accessed:

1. Home: includes a description of the project motivation and goals and quick link to latest info.
2. About MuG
 - Objectives
 - Innovations
 - Pilot cases
 - Publications
 - Milestones
 - Deliverables
3. Partners
4. News and events
5. MuG VRE
6. Contact
7. Private Area (link providing access to restricted contents)

Some of the above-described functionalities, including links to a downloadable version of the periodic newsletter, links to publications and to public deliverables will be incorporated progressively.

3.1 Home

In the “Home” section, the motivation and main goals of the project are highlighted, focusing in the challenges imposed by outstanding advance in the field of genomics.

From the Home page links to different information sources and social media are provided, with the aim of giving a quick overview about the project’s latest developments:

- **Latest news,**
- **Project brochure:** This section includes a link to a downloadable version, in pdf, of the project brochure developed for dissemination purposes (deliverable 2.1).
- **Twitter feed**

Contents:

Multiscale complex Genomics - Exploring genome beyond sequence

3D/4D genomics represents one of the biggest challenges for biology and biomedicine in the next decade. The MuG project will provide tools to integrate the navigation in genomics data from sequence to 3D/4D chromatin dynamics data.

Understanding how genome is organized in the space and how this affects gene regulation would be instrumental to fully understand the time-dependent connection between genome and phenotype.

Heterogeneity of data, disconnection among databases and lack of standardization in analysis and simulation tools, are threatening to hold back advances in this field.

MuG responds to the latest computational challenges of 3D/4D genomics by bringing this community closer to the HPC and Big Data world and providing a suitable set of **tools** and **infrastructure**.

A screenshot of the home page is presented below:

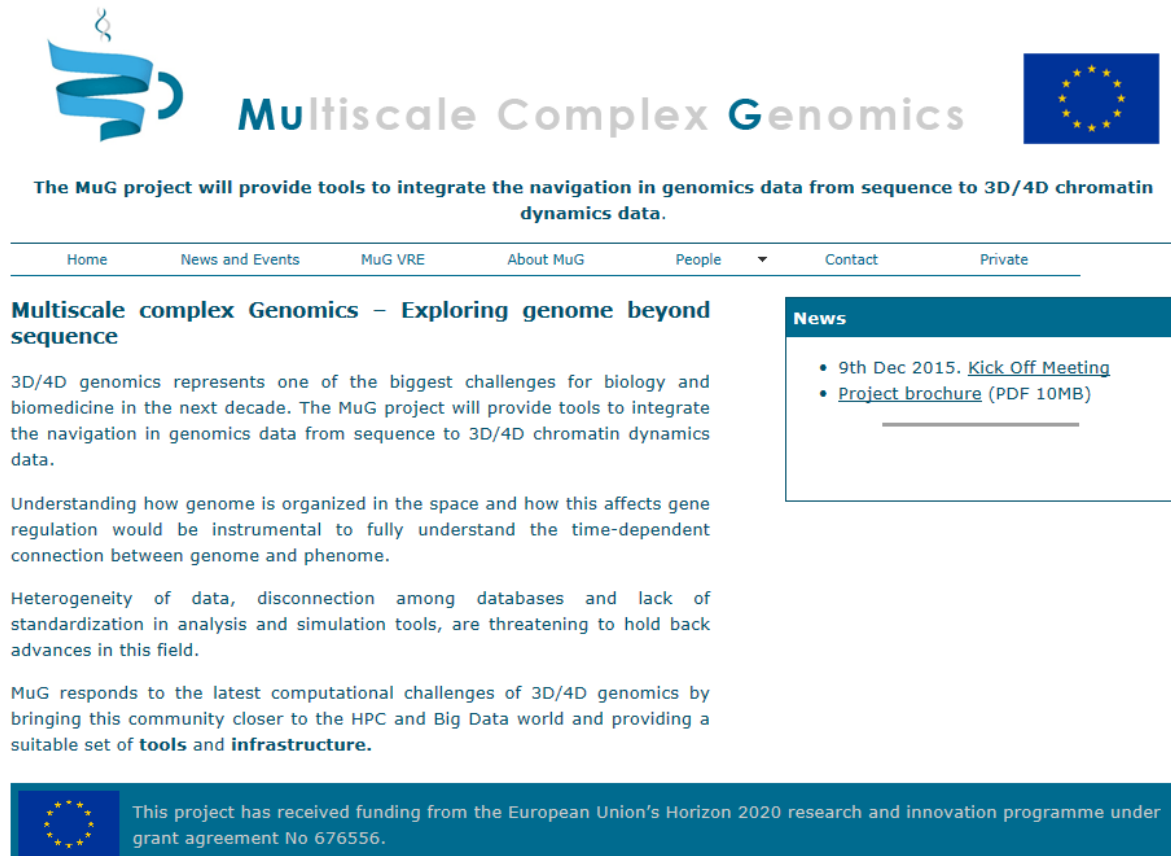


Figure 1: Screenshot of the www.multiscalegenomics.eu home page. Includes basic project description as well as link to main news and Twitter feed.

3.2 About MuG

This section is composed of different sub-sections, and compiles all information related to the MuG project.

The result is shown in the following screenshots.



The MuG project will provide tools to integrate the navigation in genomics data from sequence to 3D/4D chromatin dynamics data.

[Home](#) [News and Events](#) [MuG VRE](#) [About MuG](#) [People](#) [Contact](#) [Private](#)

MuG at a Glance

Coordinator: Dr. Modesto Orozco (IRB Barcelona)

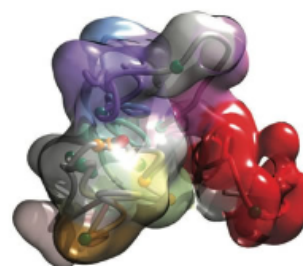
Start date: 1st November 2015

Project duration: 36 months

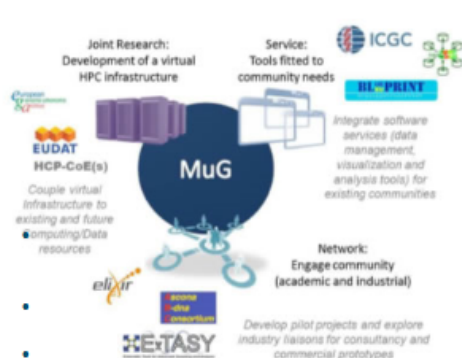
Total Project Cost: 2.961.163,75 €

Research efforts in 3D and 4D genomics are experiencing exponential growth. However, the current lack of standardization in analysis and simulation tools is threatening to become a bottleneck.

MuG responds to the latest computational challenges of 3D/4D genomics by approaching this community to the HPC world and providing a suitable set of tools and infrastructure.



The Challenge of 3D and 4D genomics



3D and 4D genomics represent one of the greatest challenges for biology and biomedicine in the next decade.

Understanding how genome is organized in the space and how this affects gene regulation would be instrumental to fully understand the time-dependent connection between genome and phenotype.

The vision of the MuG project lays on the achievement of the following challenges:

- To prevent a potential collapse in 3D/4D genomics due to inability to manage the computational problems originated by massive sequencing experiments and simulations.
- To nucleate the 3D/4D community and bring it closer to the HPC and Big Data world.
- To provide an integrated three-dimensional picture of the genome

Objectives

- To stimulate and engage the active and growing biology community to interact with the HPC and Big Data world through the development of user oriented solutions. This is to be achieved by defining standards in software and data and creating a sustainable, HPC-optimized infrastructure to foster research in this field.
- To develop, benchmark, standardize, and integrate in common workflows multi-physics simulation and analysis tools used in 3D/4D genomics.
- To create a toolbox of HPC ready software for higher-scale simulations and for structural analysis of experimental data.
- To create a common data infrastructure, generate a common database repository, generate the required data ontology, and integrate 3D/4D data with structural and 1D genomics databases.
- To create the first ever multi-resolution genome-browser with new visualization tools able to move from 1D genomics, to low- and even high-resolution levels.
- Through pilot studies to prove the strength of the developed methodology: improved understanding of chromatin structure and dynamics and other biological relevant topics.

Key Innovations

1. A scalable computational infrastructure for multi-scale complex genomics, defining a new scheme for interaction between the 3D/4D genomics community and the HPC world
2. Modelling of chromatin re-organizations aimed at improving the knowledge on cell biology and predicting links to disease.
3. Connect 1D to 3D/4D genomics data, allowing for the first time to navigate across data from different sources and scales in a single "complex" model.
4. Visualization of complex genomic machinery with a 4D multiresolution picture.

User Driven

The utility of MuG will be demonstrated through the development of [3 pilot projects](#) that will have early access to the MuG Virtual Research Environment services, providing valuable feedback and thus helping to develop the data management, computational infrastructure and analysis modules.

Understanding the Chromatin reorganization during senescence

Multi-resolution dynamic map of Yeast chromatin

Physical properties of DNA and protein-DNA complexes

Impacts

ACADEMIA AND RESEARCH INFRASTRUCTURES

Genomics and personalized medicine research community
Biosimulation and 3D Genomics community
HPC for Life Sciences initiatives
Users of European Infrastructures for HPC/Big Data
Genomics research community
Computer Science research community

LIFE SCIENCES INDUSTRY

Sequencing instrument vendors
Biotech/Pharma and Public Health



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 676556.

Figure 2: The "About MuG" section provides the main information about the project (<http://www.multiscalegenomics.eu/about.htm>).

The contents of this section have been created based on the following compilation of information about the project.

The content is subject to modifications as the results become available and with the aim to improve communication with the end user community.

3.2.1 Objectives

Contents:

- **To stimulate and engage the active and growing biology community** to interact with the HPC and Big Data world through the development of user oriented solutions. This is to be achieved by defining standards in software and data and creating a sustainable, HPC-optimized infrastructure to foster research in this field.
- To develop, benchmark, standardize, and integrate in common workflows multi-physics **simulation and analysis tools** used in 3D/4D genomics.
- To create a **toolbox of HPC ready software** for higher-scale simulations and for structural analysis of experimental data.
- To create a **common data infrastructure**, generate a common database repository, generate the required data ontology, and integrate 3D/4D data with structural and 1D genomics databases.

- To create the first ever **multi-resolution genome-browser** with new visualization tools able to move from 1D genomics, to low- and even high-resolution levels.
- Through **pilot studies to prove** the strength of the developed methodology: improved understanding of chromatin structure and dynamics and other biological relevant topics.

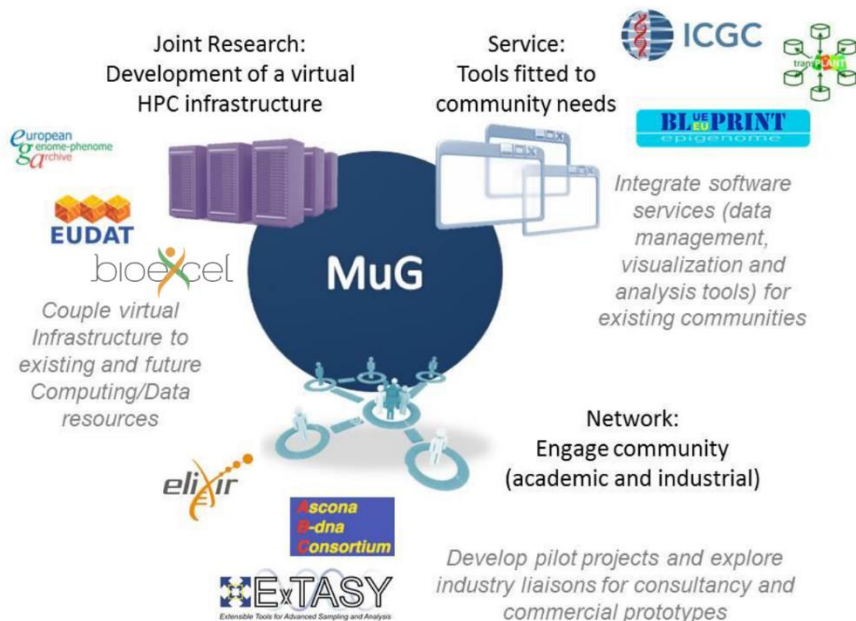


Figure 3: The figure summarized the vision of the MuG Project.

3.2.2 Key Innovations

Contents:

MuG goes beyond the state of the art in different fields:

1. Connection of **1D to 3D/4D genomic data** in a single data paradigm. This allows MuG to navigate through different scales using the appropriate model at each scale.
2. **Visualization** of complex genomic machinery with a 4D multiresolution picture. MuG will provide an integrative way to simulate entire chromatin at different levels of resolution
3. **Modelling of chromatin re-organizations**, linked to different needs and/or pathological state of the cells.
4. **New paradigm of relationship in the 3D/4D genomics community**, including a new scheme of interaction with the HPC world.

3.2.3 Pilot cases

Contents:

The utility of MuG will be demonstrated through the development of **3 pilot projects**, which will have early access to VRE services, providing valuable feedback and thus helping to develop the data management, computational infrastructure and analysis modules.

1. Senescence

Contents

CNRS-IGH will use the MuG VRE to study the changes of nuclear organization that occur during senescence and guide the computational efforts to develop and adapt the VEW to the real needs of biologists working on 3D genomics.

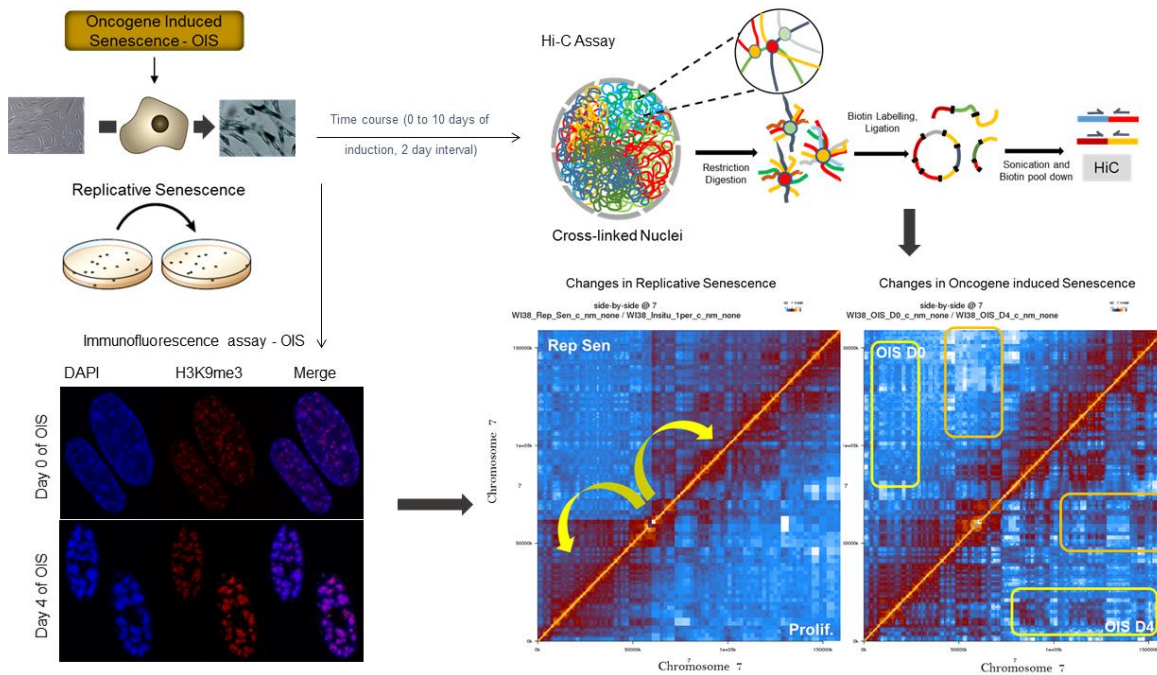


Figure 4: The illustration depicts the research activities carried out by the group of Giacomo Cavalli, at IGH-CNRS, which constitute one of the 3 use cases on which MuG will be demonstrated.

2. Multi-resolution dynamic map of Yeast chromatin

Contents:

This pilot project will merge efforts from different partners to create a genome visualizer, providing biologists with a full range of resolution information on the DNA, from nucleosome to the entire chromosome.

3. Physical properties of DNA and protein-DNA complexes

Contents:

Using high throughput simulation tools embedded in the VRE, transcription factor binding and DNA bending will be studied.

3.2.4 Publications

As the project progresses, publications made by the partners in the project research framework will be listed in this section.

3.2.5 Milestones

The achievement of the different milestones will be updated throughout the project.

Milestone No	WP	Title	Achievement date
MS1	1	First annual report submitted	
MS2	1	Second annual report submitted	
MS3	1	Third annual and final reports submitted	
MS4	2	Publication of project factsheet	
MS5	2	Setting the project website	
MS6	2	Initial exploitation plan	
MS7	2	Final exploitation plan	
MS8	2	First training workshop	
MS9	2	Second training workshop	
MS10	3	JavaScript-based browser for 1D genomic data taking inputs 3D models and genome annotations	
MS11	3	A working version of the 3D genome browser connecting 1D and 2D panels	
MS12	3	A working version of the 3D genome browser connecting 1D, 2D and 3D panels	
MS13	4	Establishment of open source software development infrastructure	
MS14	4	Publication of public data APIs	
MS15	4	Deployment of components to public and private clouds	
MS16	4	Online documentation and training materials available	
MS17	5	Early prototypes of the computational infrastructure	
MS18	5	User support tools available	
MS19	5	Programming models release	
MS20	5	Computational infrastructure final version	
MS21	6	Production of the first 3D/4D models of protein-DNA complexes	
MS22	6	4D data from the ABC project processed to generate database of DNA sequence-flexibility relationship	
MS23	6	Tools to predict DNA sequence-flexibility relationship	
MS24	6	Comparison of software to calculate and analyse FISH data	
MS25	6	Comparison of software to calculate and analyse 3C data	
MS26	6	Comparison of different methodologies to add protein-induced perturbations into predictions of DNA sequence-flexibility relationship	
MS27	7	Production of 3D genome maps of normal cells as well as senescent cells	
MS28	7	Obtention of yeast Hi-C and FISH data	
MS29	7	ABC data entered into the MuG 3D/4D databases	
MS30	7	ABC data analysed and DNA sequence-dependent flexibility data obtained	
MS31	7	MuG databases and visualisers to identify TFs with binding coupled to DNA topological constraints	
MS32	7	Simulations of selected DNA-TFs complexes	
MS33	7	Analysis of TFs 3D/4D data and correlation with ABC predictions	

3.2.6 Deliverables

The public deliverables will be published in the website at the end of the project. Publication of the deliverables in the website will be subject to embargo time of results destined to publication.

Deliverable No	Title	Link
D1.1	Project Handbook	
D1.2	Quality plan	
D2.1	Project fact sheet	
D2.2	Project website	
D2.3	Plan for dissemination and use of knowledge	
D2.4	Monitoring of the plan for dissemination and use of knowledge	
D2.5	Project's outputs	
D2.6	Initial exploitation plan based on market knowledge	
D2.7	Final exploitation plan based on market knowledge	
D2.8	Report on the MuG training programme	
D3.1	A critical evaluation of the problems on data structure the browser has to solve	
D3.2	A browser-track that implements and connects all the 1D data from a genome or a genomic domain	
D3.3	A browser-track that implements and connects all the 2D data from a genome or a genomic domain	
D3.4	Preliminary browser-track that implements and connects all the 3D data from a genome or genomic domain	
D3.5	Tailored implementation of browser to support pilot projects in WP7, including the interface with the Yeast genome viewer	
D3.6	A 3D genome browser connecting all the 1D-2D-3D tracks. The browser will allow different renderings of the upload coordinates	
D3.7	A first prototype of analysis tools for mining of the data provided by the browser	
D4.1	A requirement specification document with the data types, processing and data models needed	
D4.2	Data Management plan	
D4.3	Pipelines design and implementation	
D4.4	Database and ETL design and implementation	
D4.5	Data access API specification and implementation	
D4.6	Benchmarks and documentation	
D5.1	Computational infrastructure set-up	
D5.2	Computational infrastructure components implementation	
D5.3	Final computational infrastructure and its components	
D6.1	Design of computational architecture of software modules	
D6.2	Software tools of protein-DNA interactions	
D6.3	Software tools linking multiresolution structural and simulation data	
D6.4	Software tools on 3C interaction data	
D6.5	Software tools on analysis of single-cell experiments	
D7.1	Report on the use of MuG VRE on the senescence project	
D7.2	Report on the use of MuG VRE on the integration of whole yeast genome data	
D7.3	Report on the use of MuG VRE on the integration of DNA simulation data	

3.3 People

Several areas of expertise and world-leading institutions and researchers in their respective research fields join efforts to bring this project to life. This section is devoted to presenting the different institutions involved in the MuG project and their background and key expertise.

A general view of the whole consortium is provided, including a group picture taken during the project kick-off meeting at IRB Barcelona facilities in December 2015.

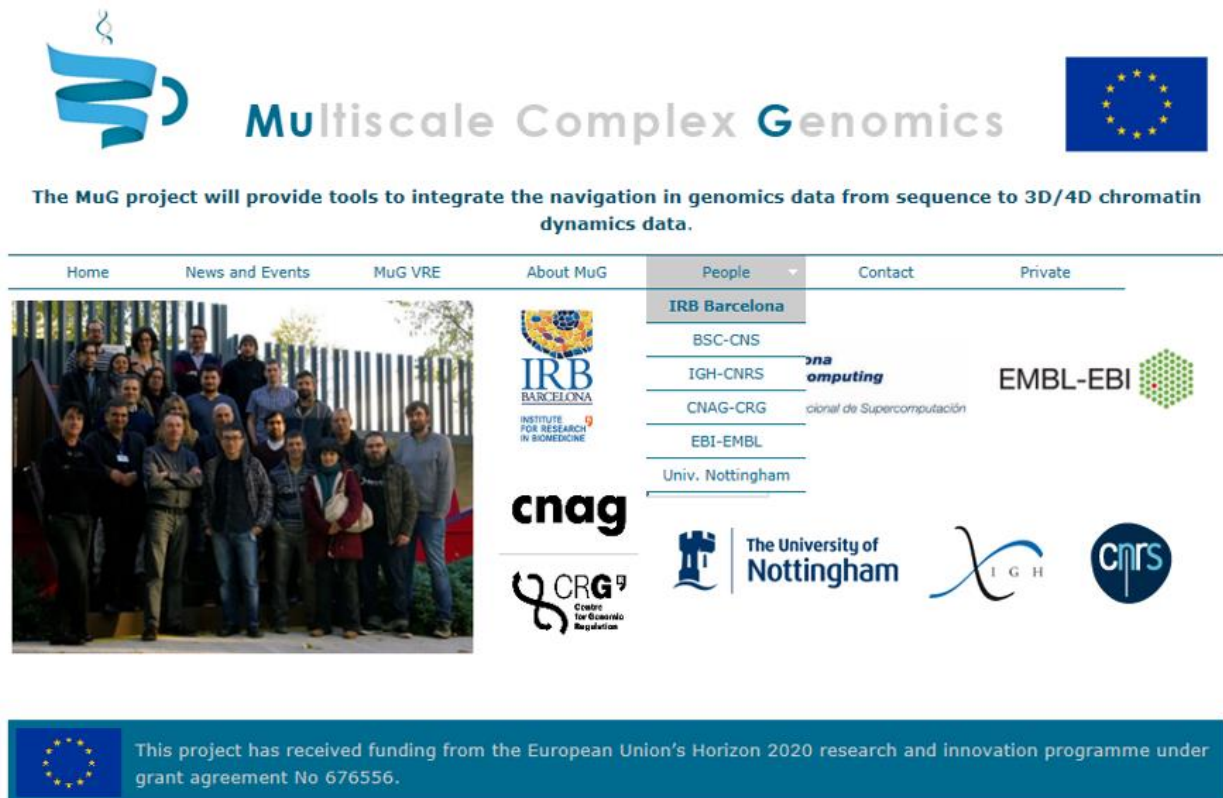


Figure 5: Information about the members of the consortium, including information about involved groups and PIs and contributed expertise description (<http://www.multiscalegenomics.eu/partners.htm>)

From the partners section, a menu provides access to additional information for each partner: official logos of each one of the 6 MuG project beneficiaries, **links to their websites** and detailed information about each institution and the PIs /research groups involved in MuG.



The MuG project will provide tools to integrate the navigation in genomics data from sequence to 3D/4D chromatin dynamics data.

Home About MuG People News and Events Private Contact

Institute for Research in Biomedicine of Barcelona (IRB Barcelona)



Modesto Orozco



Institute for Research in Biomedicine
Barcelona Institute for Science and Technology
Parc Científic de Barcelona
C/ Baldri Reixac 10
08028 Barcelona - Spain

The Institute for Research in Biomedicine of Barcelona (IRB Barcelona) is an independent, non-profit research centre engaged in basic and applied biomedical science. It was created by the Government of Catalonia and the University of Barcelona in October 2005. The quality of research performed at IRB has been recognized by the Severo Ochoa Distinction for Excellence, awarded by the Spanish Ministry of Science and Innovation. Research is conducted by 23 groups, organized into 5 interdisciplinary programmes. Each programme includes a diverse faculty of researchers with overlapping scientific interests. The research areas have been structured to be highly interactive, fostering multidisciplinary by combining modern molecular biology, classical genetics, biochemistry, structural biology, bioinformatics, systems biology, advanced microscopy, and genomic and proteomic approaches.

Figure 6: Excerpt from IRB Barcelona description page. Similar pages with information of each project beneficiary can be accessed from the menu in the “People” section.

3.4 News and events

This section will be essential to keep the visitors updated about the latest developments within the MuG project as well as providing information about all upcoming relevant events.

To this end, this will be periodically updated with related scientific conferences, training events, consortium meetings, etc.

MuG periodic **newsletters** will also be downloadable from this section of the website.

3.5 MuG virtual research environment

This section of the website shall become the gate to the MuG virtual research environment services:

- Software tools:
 - Genome browser: successive releases of the browser will be made throughout the project, until the final version, connecting 1D, 2D and 3D panels becomes finally available in M36.
 - Tools of protein-DNA interactions
 - Tools linking multiresolution structural and simulation data

- Tools on 3C interaction data
- Tools on analysis of single-cell experiments
- User support tools
- Documentation and training materials

At present, the section describes the service activities that will be developed and deployed in the framework of the project.

Contents:

MuG develops **tools and services** fitted for community needs as well as the **computing infrastructure** necessary to support the VRE services.

- **Infrastructure provisioning:** user accounts, user certificates and assignment of a gateway for access to the European data and computing infrastructure to enable simulations on distributed supercomputing resources.
- **End-user interfaces:** interactive and programmatic interfaces will be provided
- **Interoperability,** adopting standards implementation to ensure alignment with the activities on e-infrastructures.
- **Deployment, maintenance and support** to guarantee the quality of the platform.
- **Documentation and training** to ensure the modules of this VRE can be correctly used by users and developers.

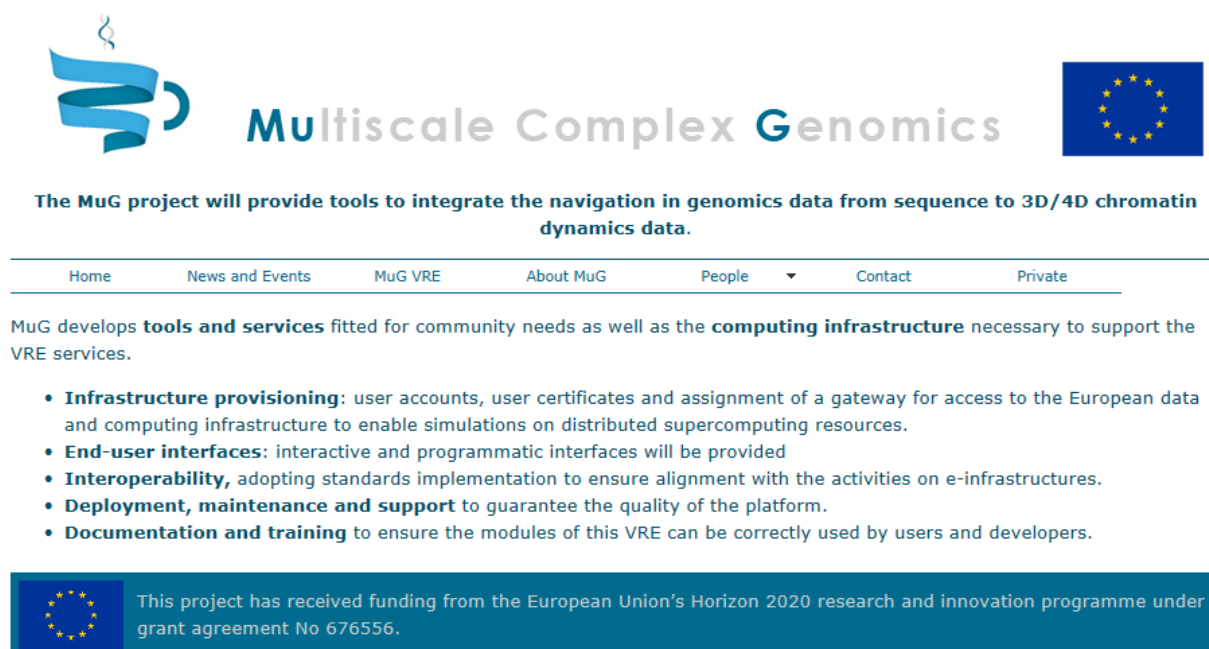



Figure 7: MuG Virtual Research Environment will become the portal to the MuG services as project results become available. <http://www.multiscalegenomics.eu/MuGVRE.htm>


3.6 Contact

MuG has established a number of communication channels, including the use of social media, to enhance communication with target audiences. Links are included in this section of the website.

- **e-mail** requests will be managed by the coordinator through the following e-mail address, which is available on the website: irbmultiscalegenomics@irbbarcelona.org
- **Twitter**: a dedicated twitter account (@MuG_genomics) has been created as the main means to improve communication through social media.
- **YouTube**: a youtube channel will be used for any future releases of video related to the project.





Multiscale Complex Genomics





The MuG project will provide tools to integrate the navigation in genomics data from sequence to 3D/4D chromatin dynamics data.

Home News and Events MuG VRE About MuG People ▼ Contact Private

 irbmultiscalegenomics@irbbarcelona.org

 https://www.youtube.com/channel/UCzDZ3b80eck24S_m2oKq6-g?view_as=public

 https://twitter.com/MuG_genomics

 This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 676556.

3.7 Private area

The private area, for restricted use of the members of the MuG consortium, is accessible from a link in the public website:

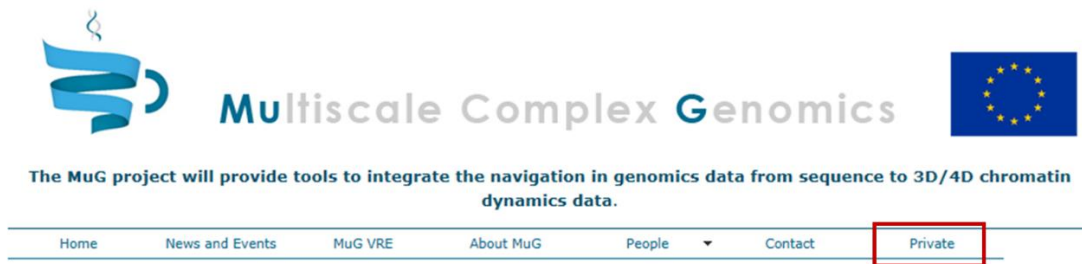
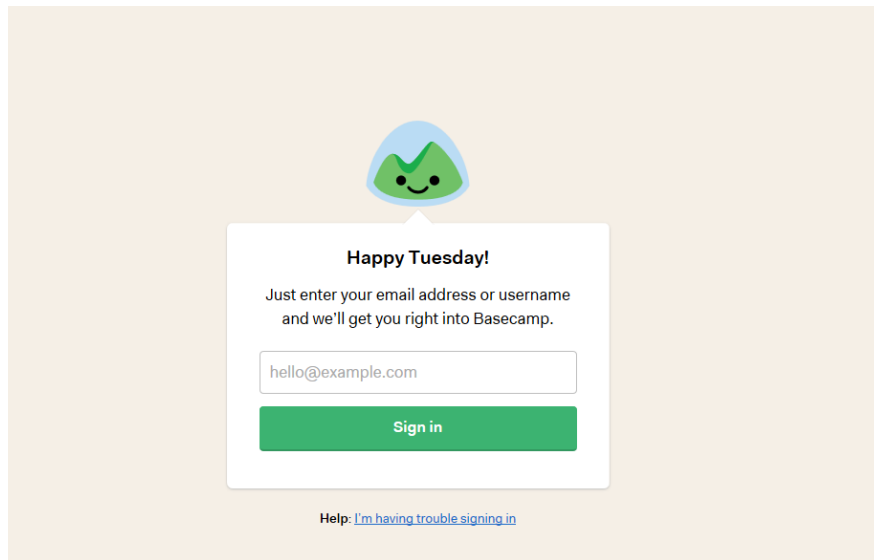


Figure 8: Link to the Project restricted area.



For the means of private communication and exchange of information among the MuG consortium partners, a link is available on the MuG website to a private account in an online project management platform (Basecamp), to which all project beneficiaries have been granted access.

The easy-to-use tool **Basecamp** has been selected to collect documentation generated during the lifetime of the project. This internal use tool will be essential for daily internal communication among consortium members.

The objective is maintain a constant contact between partners and also to have all the information: deliverables, working documents, etc. collected on the same platform. This tool also allows the members to create discussions, set to-do lists, and keep assignments and relevant files / documents / images and any related schedules in the project.

4 FUTURE WORK

The current version of the website complies with the function of dissemination and communication about the MuG project activities.

Periodic updates, with the latest information about the project (newsletters, related events, social media, etc.) will be progressively incorporated into the public site.

As the results of the project become available and the infrastructure of the virtual research environment becomes available, the MuG VRE portal will be accessible from the link provided in the public site, as defined in the present document.

Implementation of Google Analytics is also foreseen in order to monitor the visits and extract relevant information of the visitors' profiles via the Google Analytics platform. The collected data will give insights into the website traffic and valuable information valuable towards evaluation of the key performance indicators and development of the exploitation plan.