



UK Human Functional Genomics Initiative

Guidelines for Collaboration with the Data Coordination Centre

Introduction to the UK Human Functional Genomics Initiative:

The UK Human Functional Genomics Initiative (FGx) is a multimillion-pound investment aimed at advancing our understanding of how genomic variation influences human physiology and disease. Coordinated from the University of Exeter and structured into thematic clusters, the FGx provides a central support hub for academics, clinicians and industry partners from across the UK's functional genomics community.

Background to the FGx Data Coordination Centre:

To support access to data and accelerate scientific breakthroughs in the field of functional genomics, the FGx has established a dedicated Data Coordination Centre (DCC) for enabling scientists to find, use and share data generated by the initiative and its external partners. The DCC will comprise a core team with collective expertise in bioinformatics, data science, data stewardship, and software engineering. This team is being strategically assembled so that the DCC can achieve its primary goal of developing a scalable data platform for linking and annotating datasets across diverse modalities for federated analyses. Through this process, the DCC also aims to improve metadata standards, disseminate workflows for standardised data cleaning and analysis, improve ease of data linkage, and upskill the community in data stewardship and the principles of FAIR (Findable, Accessible, Interoperable, and Reusable) data and open and reproducible science.

Working with the FGx Data Coordination Centre:

The development of our scalable data platform is an ongoing and adaptive process, and a key focus of the DCC is that this platform meet the core needs of the UK's functional genomics community. The DCC team recognises that regular community input, collaboration and feedback is therefore vital for the creation of a successful DCC data platform. As such, the DCC team have established a set of guidelines that they encourage collaborators to keep in mind when working with us:

- Implement standardised metadata collection for appropriate repository of a given data type.
- Maintain unique and consistent sample identifiers.
- Upload laboratory protocols to protocols.io, including links to training where available.
- Be able to share data to allow for platform development and implementation.
- Document and share reproducible (meta)data processing and analysis workflows.
- Be willing to be a beta user of the data platform and give feedback to shape its development.
- Engage with training available through the DCC and its wider partners.
- Join the FGx and DCC mailing lists to keep up to date with what is happening.
- Partake in surveys that the DCC distributes as a means to surface community needs.
- Be willing to participate in focus groups organised as part of the DCC.

The above set of guidelines remains under constant consideration by the DCC team, and it is possible that it may change over time as development the DCC data platform progresses and/or the needs of the community evolve. A log of major changes to this list will be maintained and any major updates to these guidelines will be communicated via the DCC mailing list.

Contacting the FGx Data Coordination Centre:

The FGx DCC team can be contacted via the following email address: dcc-fgx@exeter.ac.uk