Version control for cell engineering – new processes & tools for more open and trustworthy engineering biology

Lenka Pelechova[*], Leanne Hobbs[+], Nadia Rostami[^], Natalie Partridge[~]

[*] LP is a postdoctoral social scientist working at the Interdisciplinary Computing and Complex Biosystems (ICOS) research group.

[+] LH is a research software engineer working at ICOS

[^] NR is a senior engineer biologist working at ICOS

[~] NP is a doctoral student studying at the Policy, Ethics and Life Sciences (PEALS) research centre.

The commoditization of synthetic DNA and gene-editing tools are rapidly accelerating while fuelling the growth of "designer" biological cell lines. In turn, these scientific advances are at the root of the creation of new companies and new biologically derived consumer products, which are set to impact most areas of our lives.

At the same time, there are deep gaps in communication of engineering biology, since only a relatively small proportion of the information generated during *the process of cell engineering* is being made available at the time of publication. Therefore, new processes, tools and policies are needed to substantially improve traceability, transparency, and—hopefully - trustworthiness in engineering biology.

Open Publication:

With this in mind, we recently published[1] in open access form "CellRepo", which is an online strain and species database that allows monitoring, organising and tracking of

[1] Tellechea-Luzardo, et al. *Versioning biological cells for trustworthy cell engineering*. **Nat Commun** 13, 765 (2022). https://doi.org/10.1038/s41467-022-28350-4

engineered cell lines. Our approach to a more open and transparent biotechnology practice draws strongly from (open source) software engineering and, more specifically, from version control repositories (VCRs). VCRs are at the heart of the software development life-cycle of projects large and small. VCRs store programmers' files including the history of changes made to these files throughout the development, either by individuals or teams, and are automatically maintained and tracked by the system. VCRs enable traceability, backtracking (if necessary) as well as branching new versions during software development projects. At the same time, version control radically increases software teams' productivity and creates a more open ecosystem of innovation around software projects. Free software and open-source projects depend heavily on version control systems. Those mentioned above are some of the benefits we are helping to bring into engineering biology.

Free Software:

Alongside our publication, we released a free-to-use cloud-based version control app for engineered microorganisms, that allows the recording and tracking of changes in engineered cell lines in a manner akin (albeit not identical) to software engineering VCR. This global repository of engineered organisms allows researchers to more reliably reproduce experiments involving engineered cell lines due to access to the complete record of every genetic modification made to the engineered organisms. Availability of such information, that is beyond the scope of traditional publications, will dramatically increase efficiency in the laboratories and promote collaborative genetic engineering efforts.

Open Laboratory Processes:

To supplement the publication and the cloud-based software (https://cellrepo.ico2s.org), we also included an open-source novel molecular barcoding protocols for laboratories. In addition, we demonstrated the wide-ranging applicability of our approach by providing barcoding protocols for six of the most important and diverse microbial species that are used in both academia and industry (https://cellrepo.ico2s.org/got).

Open Public Communication and Policy:

Importantly, by making engineering biology research more collaborative and transparent from the onset, our work contributes towards UKRI's Responsible Research and Innovation (RRI) framework as it improves research practices by ensuring innovation is carried out with inclusivity and openness at its core. With our proposed system, biological engineering evolves from an unconnected and solitary endeavour by laboratories in academia or industry to a multi-actor inclusive process in which researchers, citizens, policy makers, funders and NGOs can scrutinise the engineering of biological organisms at every step of the bio-design process.

We believe that it is important to make the process of science as open and transparent as possible as this can impact the public attitudes towards a discipline and towards science in general. Past research has shown that publics' concerns about engineering biology are often placed *on the process of research* rather than on better understanding of the technical specifics. In this way transparent research is regarded as an important component of public trust. To strengthen this aspect of our research, we also contributed to a short policy brief[2] in which we summarised the digital opportunities to promote openness and traceability in genetic engineering for policy makers.

[2] Partridge, N. (2021). *Towards greater transparency: Digital opportunities to promote traceability in genetic engineerin*g. **Zenodo**. https://doi.org/10.5281/zenodo.5343358