Clunies Ross Street

Black Mountain ACT 2601

Australia

11 June 2021

To Nature Communications Editorial Team,

In the submitted article we are discussing how machine learning and laboratory automation can be used for rational design of small biological parts. We have used Gaussian Process Regression for prediction of relative transcription initiation rates of bacterial ribosome binding sites and the Upper Confidence Bound-based Bandit algorithm for recommendation of genetic designs to be tested in vitro. We have integrated the ML algorithms with laboratory automation and high-throughput processes, creating a robust workflow for the design of custom RBSs. By employing these techniques, we were able to increase the reliability and reproducibility of results and increase the confidence in the design process. Using our workflow, we generated a novel library of diverse RBSs with a wide range of expression levels. Notably, a high number of these sites demonstrate translation initiation rates equalling or exceeding the currently known strong RBSs. Additionally, this work elucidated some design guidelines, including the favourable level of difference between two sequences and efficient numerical representation of the analysed DNA/RNA sequence.

We have previously contacted one of the Editors, Ross Cloney, who after reading our abstract has invited us to submit our manuscript to the journal.

We believe that our manuscript presents novel work that shows how even a small interdisciplinary team of synthetic biologists and machine learning experts can generate excellent results and we hope that the Editorial Team will find it interesting.

Best wishes,

The Authors