ANNOUNCING …

GP-Write's Genome Design Software Challenge   Jef Boeke and Joel Bader

A cash prize for the best software!

If you wanted to redesign a genome for a specific purpose (e.g. pervasive recoding to enable resistance to naturally occurring viruses), how would you do it? Well, step one would be to go to a computer and design it using software.

The problem: Software for genome design is research-ready but not production-ready. The BioStudio platform, for example, was designed with yeast in mind but has limitations that prevent ready use for mammalian-scale genomes. BioStudio was built atop existing open source components, which permitted rapid development. Now, however, there are difficulties as these components are becoming outdated and deprecated.

Our idea: We will a launch a challenge for the global GP-Write community to develop robust open source software for designing entire chromosomes and genomes across well-annotated genomes. We plan to release a detailed specification for what the open source code can do by early 2020. We have a potential sponsor for this activity, who will offer a cash prize to the best entry.

In addition to requiring a permissive free-and-open-source license for the software and its dependencies, we will also require command-line tools for sophisticated users implementing batch edits (recoding, watermarking), a GUI for fine editing and visualization, and mechanisms for converting a genome design into a DNA order sheet.

We will set a deadline for initial testing of code by mid-late 2020 in order to identify (3?) finalists, who will be invited to the next GP-Write meeting for a "bakeoff". The finalists will be given identical hard drives containing a genome of interest and associated annotation files (e.g. gff3 for input/output) and a specific task or set of global genome wide editing tasks to be accomplished (e.g. convert all TAG stop codons to TAA and insert a loxPsym sequence 3 bp downstream of each stop codon). Each will be given a chance to show off the user interface and speak about capabilities in a special session.

The runtime and precision of the output will be measured, ideally in a real-time live meeting session. The software packages will be assessed based on a combination of adherence to requirements, accuracy, memory footprint, run time performance, command line functionality and GUI ease-of-use, and additional features.