

Whole Genome Alignment Visualization

Sarah Weintraub, Aidan Pecorale

1 Introduction

2 One-sentence description

We aim to build a tool for simple genomic alignment visualization that can be used as a standalone tool as well as implemented into complex assembly pipelines.

3 Project Type

Visualization / Application

4 Audience

Who is the audience for this project? How does it meet their needs? What happens if their needs remain unmet?

Scientists looking at genomic data is the audience - specifically, this tool will be implemented into a genomic assembly pipeline that was published in Nature communications and has been used by researchers across disciplines. Right now, the pipeline just does not provide visualizations as the tools used when published do not currently work with the updates implemented.

5 Approach

5.1 Details

What is your approach?

We plan to use either a Sankey or chord chart to make our visualization happen. We will build it in d3 using two template genomes that we know the consensus for so that we can accurately test our tool.

5.2 Evidence for Success

Why do you think it will work?

There are similar tools that exist, we do not think it is a difficult problem, but not many people have attempted it due to little overlap of skills.

6 Best-case Impact Statement

In the best-case scenario, what would be the impact statement (conclusion statement) for this project? [?, ?]

The best case scenario for this project is to build a tool that is implementable for genomic pipelines to confirm assembly

7 Major Milestones

8 Obstacles

8.1 Major obstacles

8.2 Minor obstacles

9 Resources Needed

What additional resources do you need to complete this project?

We're not sure, but if the genomes are too big to run on our computers, we can use Turing to align/visualize them.

10 5 Related Publications

List 5 major publications that are most relevant to this project, and how they are related (sample citation [?]).

<https://peerj.com/articles/cs-116/>
<https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-021-04556-z>
<https://academic.oup.com/bioinformatics/article/37/3/413/5885081>
<https://www.frontiersin.org/articles/10.3389/fgene.2020.00292/full>
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3371849/>

11 Define Success

What is the minimum amount of work necessary for this work be publishable?

Aligning two genomes that show consensus will replace the outdated perl tool.