

# Gfa\_parser for genome graph viewing with path name and color

By Xuewen Wang

2021-01-05<sup>th</sup>

Email: [xwwang@uga.edu](mailto:xwwang@uga.edu), [xwang.kib@gmail.com](mailto:xwang.kib@gmail.com)

## Contents

1. Tool gfa_parser .....	1
2. Function .....	1
3. Options.....	2
4. Usage.....	2
5. Output.....	2
6. Demonstration with testing data.....	2
7. Viewing in Bandage.....	2

## 1. Tool gfa\_parser

gfa\_parser.pl

Version: build20200105; Currently it is a perl version.

Source: see attachment.

## 2. Function

Parse the graph file in .gfa. and prepare for viewing paths in graph. The Bandage can't display the path information from the original graph in .gfa.

The output could be load into Bandage for viewing graph in colorized path and path names for each node ;

for help:

```
perl gfa_parser.pl
```

for running:

```
perl /home/xuewen/Downloads/pan_genome/scripts/gfa_parser.pl -g gfa_graph_file [-s  
short_path_name_T|F] -o [gfa_graph_file.csv for Bandage]
```

### 3. Options

- g Name of the input file, in .gfa format;
- s T will display the path name in a short name in P1, P2, ...;  
F will display the full name in the original graph; default F
- o Name of output file in .csv format; default in input\_file.csv

A file (.sat1) summary of paths and nodes are generated. The .sat1 file will have the correlation between short path names and original path names.

The output is in the same directory as the input file. Alternatively, the output file can be specified in the -o parameter.

### 4. Usage

```
perl gfa_parser.pl -g tifchr02_sim_gap.xmfa.gfa [options]
```

### 5. Output

a .gfa.csv for tool Bandage

A file (.sat1) summary of paths and nodes in the input graph

### 6. Demonstration with testing data

```
perl $scriptdir/gfa_parser.pl -g tifchr02_sim_gap.xmfa.gfa -s T
```

-s T will show the path in short name P1, P2; otherwise -s F will use original paths names which may be very long.

Input:

tifchr02\_sim\_gap.xmfa.gfa

Output:

tifchr02\_sim\_gap.xmfa.gfa.csv

tifchr02\_sim\_gap.xmfa.gfa.csv.sat1

### 7. Viewing in Bandage

Load the .gfa file

e.g. :

tifchr02\_sim\_gap.xmfa.gfa

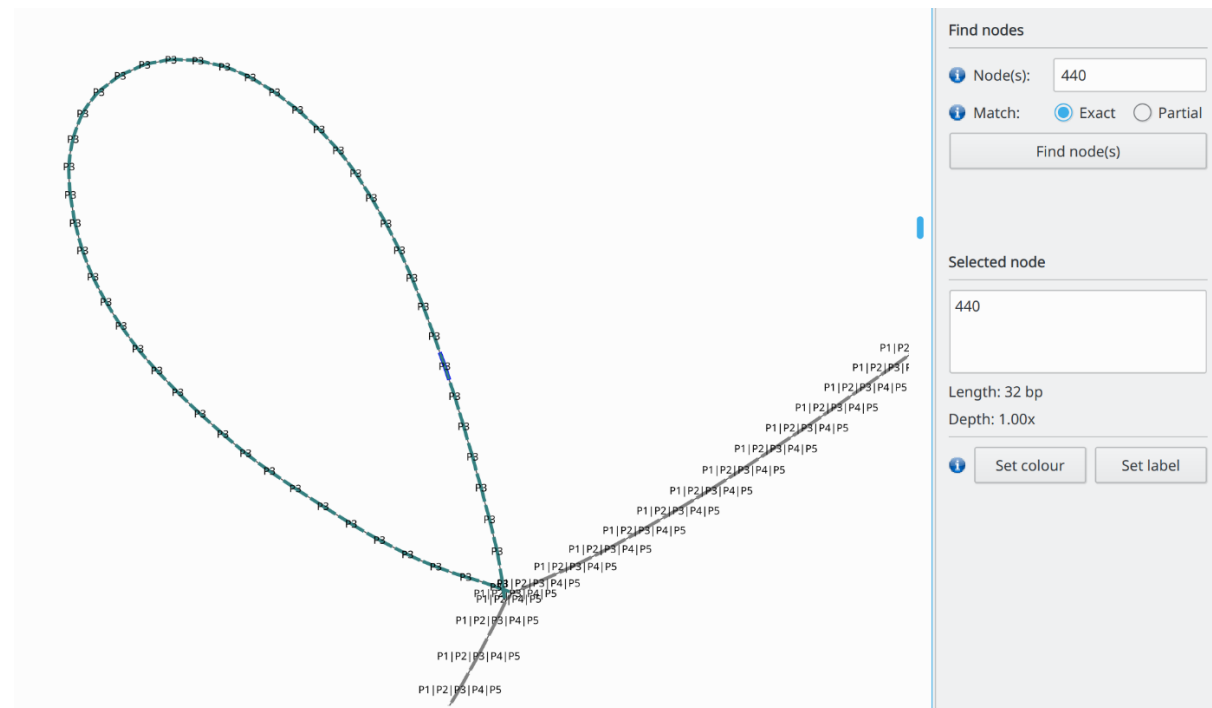
Load the .csv from this tool:

tifchr02\_sim\_gap.xmfa.gfa.csv

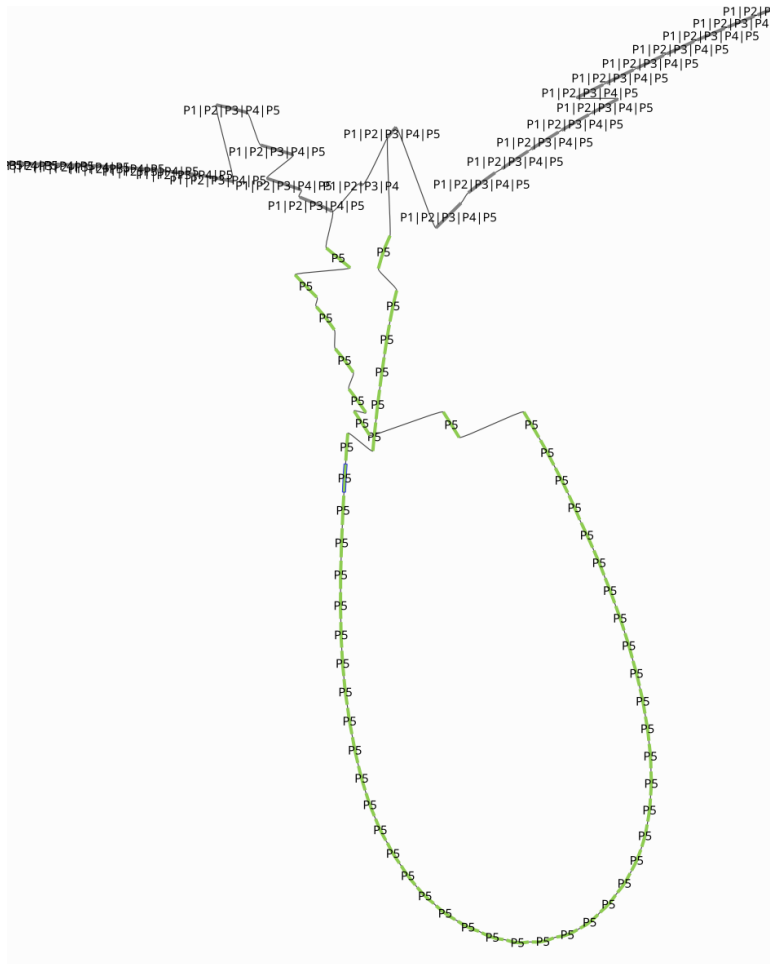
Then Bandage will display the graph with path names and path in colors:

e.g.

Go the node “440”



Go the node “1360”:



Go the node “2530”

