

---

# **PyKmers Documentation**

***Release 1.0***

**Alessio Zanga, Emanuele Cavenaghi**

**Feb 25, 2019**



---

## Contents:

---

<b>1</b>	<b>Getting Started with PyKmers</b>	<b>1</b>
1.1	Dependencies . . . . .	1
1.2	Build and Install PyKmers . . . . .	1
<b>2</b>	<b>pykmers package</b>	<b>3</b>
2.1	Module contents . . . . .	3
<b>3</b>	<b>Indices and tables</b>	<b>5</b>
	<b>Python Module Index</b>	<b>7</b>



---

## Getting Started with PyKmers

---

### 1.1 Dependencies

PyKmers relies on **pygfa library** to load GFA files. You can build this library downloading the source code from [this repository](#). In order to build and install pygfa, you need **python3**, **pip3** and **setuptools**.

```
python setup.py bdist_wheel
pip install dist/pygfa*.whl
```

### 1.2 Build and Install PyKmers

Once pygfa is installed, you can build and install PyKmers from [here](#).

```
python setup.py bdist_wheel
pip install dist/pykmers*.whl
```



### 2.1 Module contents

#### PyKmers Module

Collection of functions to extract k-mers from GFA file.

Copyright (C) 2019 by Alessio Zanga <alessio.zanga@outlook.it> Emanuele Cavenaghi <cavenaghi.emanuele@gmail.com>

`pykmers.pykmers.com(base)`

Return complement base of given base

**Parameters** `base` – Input base

**Returns** Complement base

`pykmers.pykmers.get_kmers_from_file(file, kmer_length, sorted=False, debug=False)`

Find all the kmers that corresponds at the given sequence

**Parameters**

- **file** – Input GFA file
- **kmer\_length** – Kmer length
- **sorted** – True to sort the kmers in lexicographical order
- **debug** – True to show debug info

**Returns** List of kmers

`pykmers.pykmers.get_kmers_from_string(sequence, k)`

Find all the kmers that corresponds at the given sequence

**Parameters**

- **sequence** – Sequence
- **k** – Kmer length

**Returns** List of kmers

`pykmers.pykmers.get_paths_from_graph(graph)`

Find paths in graph from root to leaf

**Parameters** `graph` – Graph

**Returns** List of path in the graph

`pykmers.pykmers.get_sequence_from_path(graph, path)`

Find the sequence that corresponds to the path without consider the overlap

**Parameters**

- `graph` – Graph
- `path` – Path

**Returns** Sequence that corresponds to the path

`pykmers.pykmers.get_subgraphs_from_graph(graph)`

Divide graph in connected components

**Parameters** `graph` – Graph

**Returns** List of subgraphs

`pykmers.pykmers.load_graph_from_file(file)`

Load a MultiGraph instance of Networkx from GFA file using pygfa library

**Parameters** `file` – Input GFA file

**Returns** MultiGraph instance

`pykmers.pykmers.rev_and_com(sequence)`

Return reverse and complement sequence of given sequence

**Parameters** `base` – Input sequence

**Returns** Reverse and complement sequence



## CHAPTER 3

---

### Indices and tables

---

- `genindex`
- `modindex`
- `search`



**p**

`pykmers.pykmers`, 3



## C

`com()` (in module `pykmers.pykmers`), 3

## G

`get_kmers_from_file()` (in module `pykmers.pykmers`), 3

`get_kmers_from_string()` (in module `pykmers.pykmers`), 3

`get_paths_from_graph()` (in module `pykmers.pykmers`), 4

`get_sequence_from_path()` (in module `pykmers.pykmers`), 4

`get_subgraphs_from_graph()` (in module `pykmers.pykmers`), 4

## L

`load_graph_from_file()` (in module `pykmers.pykmers`), 4

## P

`pykmers.pykmers` (module), 3

## R

`rev_and_com()` (in module `pykmers.pykmers`), 4