Article

Fernando Freire

December 13, 2018

Contents

1 Initializations

Script 1.0.1 (python)

```python
import requests, sys
import json
import numpy as np
import matplotlib.pyplot as plt

# The next line indicates to matplotlib to show the plots inline in the notebook
%matplotlib inline

def countBases(seq, bases = ['A', 'C', 'G', 'T']):
    """ Counts the number of occurrences of each base in the sequence seq
        There are two return values: the list of bases and the list of counts
    """
    counts = []
    for base in bases:
        counts.append(seq.count(base) + seq.count(base.lower()))
    return bases, counts

server = "https://grch37.rest.ensembl.org"
```

2 Lookup

Script 2.0.1 (python)

```python
ext = "/lookup/id"
headers={ "Content-Type" : "application/json", "Accept" : "application/json"}

# Get lookup data
r = requests.post(server+ext, headers=headers, data='{ "expand" : 1, "ids" :
 ↪  ["ENSG00000157764"] }')

if not r.ok:
    r.raise_for_status()
    sys.exit()

found = json.loads(r.text)
```

3 Get sequences

Script 3.0.1 (python)

```python
# Create list of ID
IDs = []
for transcripts in found.values():
    for transcript in transcripts["Transcript"]:
        IDs.append(transcript["id"])

dquotes_ids=json.dumps({"ids" : IDs})
print(dquotes_ids)

# Get sequence data
ext = "/sequence/id"
headers={ "Content-Type" : "application/json", "Accept" : "application/json"}

# Make request and check status
r = requests.post(server+ext, headers = headers, data = dquotes_ids)
if not r.ok:
    r.raise_for_status()
    sys.exit()

json_data = json.loads(r.text)
```

Output

```
{"ids": ["ENST00000496384", "ENST00000288602", "ENST00000479537", "ENST00000497784",
↪  "ENST00000469930"]}
```

4 Get oligos

Script 4.0.1 (python)

```python
count_oligos = {}
for record in json_data:
    oligo = record["seq"][102:105]
    if oligo in count_oligos:
        count_oligos[oligo] += 1
    else:
        count_oligos[oligo] = 1
print(count_oligos)
```

Output

```
{'AGT': 1, 'GGG': 1, 'AAA': 1, 'GGC': 1, 'GCC': 1}
```