Kireet Kundavaram

4/27/2015

Rough Draft

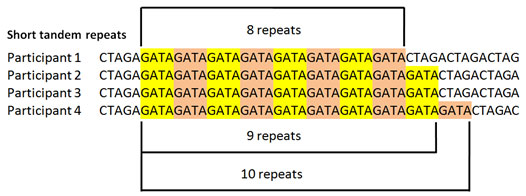
**Analysis of specific tandem repeats across an array of cyanobacteria**

**Introduction**

Bacteria constitute a large domain of prokaryotic microorganisms. There are approximately 5x1030 bacteria on Earth, that’s a biomass greater than that of all the animals and plants combined [2]. They are typically a few micrometres in length, and come in all shapes including but not limited to: spheres, spirals and rods. Bacteria, arguably one of the first lifeforms to appear on Earth, are present in almost every habitat imaginable. They have the capacity to inhibit water, hot springs, soil, radioactive waste and inside the Earth’s crust [1]. Cyanobacteria is a phylum of bacteria that obtain their energy through photosynthesis. The name originates from the Greek word kyanós meaning blue. They get this name because they are often referred to as blue-green algae. Cyanobacteria by producing gaseous oxygen as a byproduct of photosynthesis, are thought to have converted the early reducing atmosphere into an oxidizing one, and dramatically changing the composition of life forms on Earth by stimulating biodiversity. According to endosymbiotic theory, the chloroplasts found in plants and eukaryotic algae evolved from cyanobacterial ancestors via endosymbiosis [3]



Of the various kinds of sequences found in DNA, repeated sequences are of great importance to us. There are many kinds of repeated sequences such as: CRISPERSs, short and very short dispersed repeats, clustered repeats, and Tandem repeats. Clustered regularly interspaced short palindromic repeats also known as CRISPERs are segments of prokaryotic DNA containing short repetitions of base sequences with regular gaps separating the sequence. Short disperses repeats also known as SDR’s are repeated sequences in the range of in the range of 20-200 nt. Clustered repeats are repeated sequences that are generally repeated in clusters through the sequence. Finally Tandem repeats are short oligomeric units repeated one after another. Tandem repeats occur in DNA when a pattern of one or more nucleotides is repeated and the repetitions are directly adjacent to each other [4].



The focus of this experiment is going to be on tandem repeats. More specifically the "TTGGTCATTGGTCA" repeat observed to occur in many cyanobacteria. The “ATTCGATTCG” tandem repeat will then be testes against the cyanobacteria that test positive for the "TTGGTCATTGGTCA" tandem repeat. 43 different cyanobacteria were tested for the presence of "TTGGTCATTGGTCA" but only 15 showed up as positive for the sequence. And of the 15 that were positive for the "TTGGTCATTGGTCA" tandem repeat only 11 were also positive for the “ATTCGATTCG” tandem repeat.

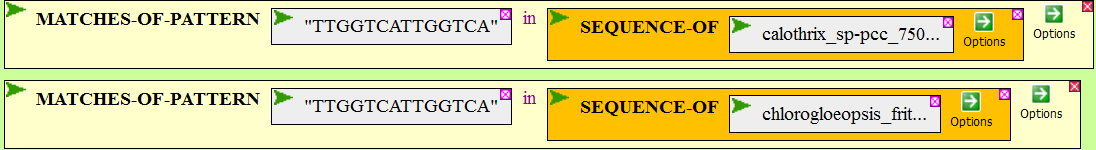


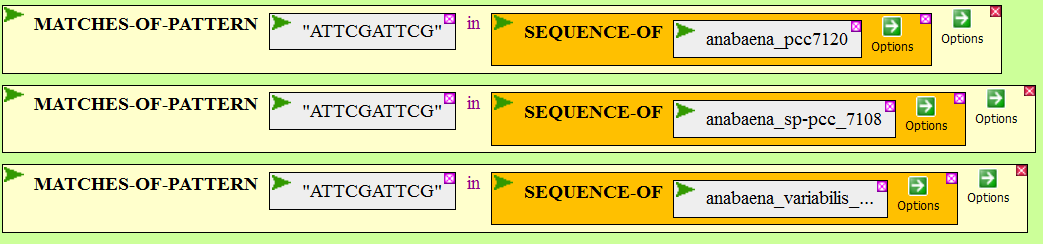
**Methods**

The tool tandem repeat finder was used initially to comb through dozens of potential cyanobacteria [6]. BioBike is a cloud-based, through-the-web programmable bioinformatics platform [12]. The PhAnToMe and cyanoBIKE version of BioBIKE were used to create and analyze the tandem repeats in 15 different cyanobacteria. The cyanobacteria used are:

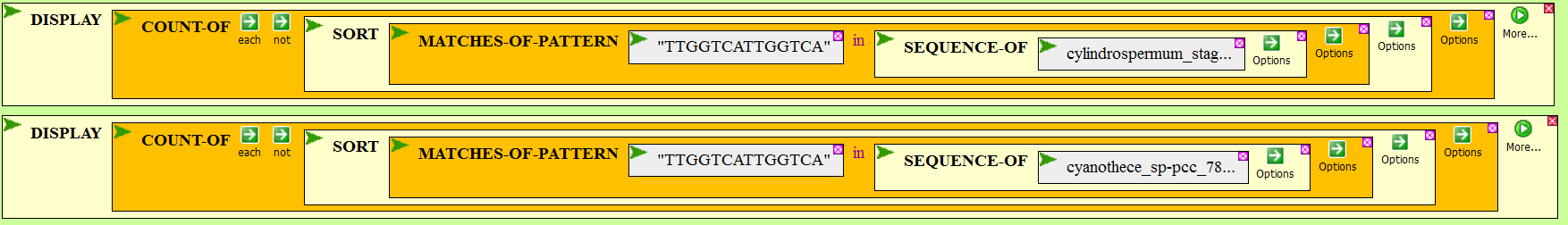
* anabaena\_cylindrica\_pcc\_7122
* anabaena\_sp-pcc\_7108
* chlorogloeopsis\_fritschii\_pcc\_6912
* cyanothece\_sp-pcc\_7425
* cylindrospermum\_stagnale\_pcc\_7417
* anabaena\_pcc7120
* anabaena\_variabilis\_atcc29413
* calothrix\_sp-pcc\_7103
* calothrix\_sp-pcc\_7507
* chroococcidiopsis\_thermalis\_pcc\_7203
* crinalium\_epipsammum\_pcc\_9333
* cyanothece\_atcc51142
* stanieria\_cyanosphaera\_pcc\_7437
* cyanothece\_sp-pcc\_7424
* cyanothece\_sp-pcc\_7822

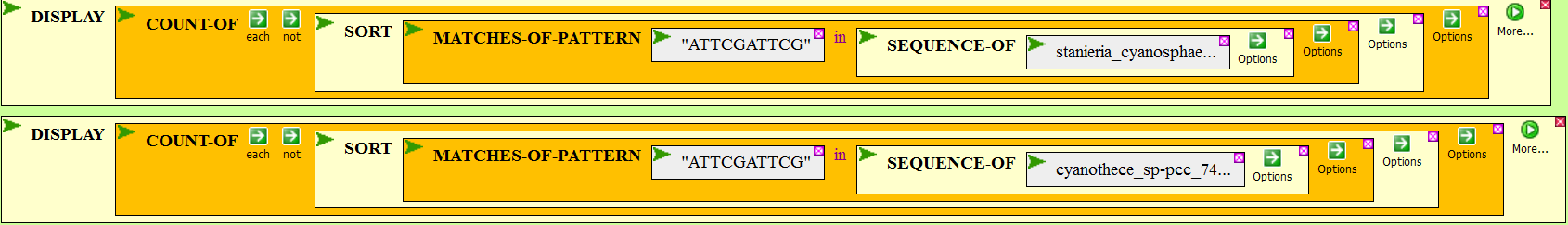
They were first tested to find the "TTGGTCATTGGTCA" tandem repeat. Then if the repeat successfully found then they would be tested for the “ATTCGATTCG” tandem repeat.





After that the number of time each tandem repeat occurred was calculated and recorded.





**Results**

The result of the following experiment was out of all the cyanobacteria used only 4 did not share both of the tandem repeats. The results of the experiment are recorded below. The blue table indicates if the corresponding tandem repeats were present in the tested cyanobacteria while the green table records the number of time each of the tandem repeats occur.

|  |  |  |
| --- | --- | --- |
| Cyanobacteria | "TTGGTCATTGGTCA" matches | “ATTCGATTCG” matches |
| #$anabaena\_cylindrica\_pcc\_7122 | ✔ | ✔ |
| #$anabaena\_pcc7120 | ✔ | ✖ |
| #$anabaena\_sp-pcc\_7108 | ✔ | ✔ |
| #$anabaena\_variabilis\_atcc29413 | ✔ | ✖ |
| #$calothrix\_sp-pcc\_7103 | ✔ | ✔ |
| #$calothrix\_sp-pcc\_7507 | ✔ | ✔ |
| #$chlorogloeopsis\_fritschii\_pcc\_6912 | ✔ | ✔ |
| #$chroococcidiopsis\_thermalis\_pcc\_7203 | ✔ | ✖ |
| #$crinalium\_epipsammum\_pcc\_9333 | ✔ | ✔ |
| #$cyanothece\_atcc51142 | ✔ | ✔ |
| #$stanieria\_cyanosphaera\_pcc\_7437 | ✔ | ✖ |
| #$cyanothece\_sp-pcc\_7424 | ✔ | ✔ |
| #$cyanothece\_sp-pcc\_7425 | ✔ | ✔ |
| #$cyanothece\_sp-pcc\_7822 | ✔ | ✔ |
| #$cylindrospermum\_stagnale\_pcc\_7417 | ✔ | ✔ |

|  |  |  |
| --- | --- | --- |
| Cyanobacteria | "TTGGTCATTGGTCA" hits | “ATTCGATTCG” hits |
| #$anabaena\_cylindrica\_pcc\_7122 | 34 | 11 |
| #$anabaena\_pcc7120 | 28 | 0 |
| #$anabaena\_sp-pcc\_7108 | 22 | 5 |
| #$anabaena\_variabilis\_atcc29413 | 30 | 0 |
| #$calothrix\_sp-pcc\_7103 | 12 | 16 |
| #$calothrix\_sp-pcc\_7507 | 36 | 6 |
| #$chlorogloeopsis\_fritschii\_pcc\_6912 | 33 | 15 |
| #$chroococcidiopsis\_thermalis\_pcc\_7203 | 62 | 0 |
| #$crinalium\_epipsammum\_pcc\_9333 | 39 | 12 |
| #$cyanothece\_atcc51142 | 6 | 9 |
| #$stanieria\_cyanosphaera\_pcc\_7437 | 7 | 0 |
| #$cyanothece\_sp-pcc\_7424 | 15 | 11 |
| #$cyanothece\_sp-pcc\_7425 | 4 | 9 |
| #$cyanothece\_sp-pcc\_7822 | 18 | 15 |
| #$cylindrospermum\_stagnale\_pcc\_7417 | 96 | 10 |

**Citations**

1. Fredrickson JK, Zachara JM, Balkwill DL, Kennedy D, Li SM, Kostandarithes HM, Daly MJ, Romine MF, Brockman FJ (2004). ["Geomicrobiology of high-level nuclear waste-contaminated vadose sediments at the Hanford site, Washington state"](http://www.ncbi.nlm.nih.gov/pmc/articles/PMC444790). *Applied and Environmental Microbiology* **70** (7): 4230–41
2. C.Michael Hogan. 2010. [*Bacteria*. Encyclopedia of Earth. eds. Sidney Draggan and C.J.Cleveland, National Council for Science and the Environment, Washington DC](http://www.eoearth.org/article/Bacteria?topic=49480)
3. [Schopf, J. W.](http://en.wikipedia.org/wiki/J._William_Schopf" \o "J. William Schopf) (2012) ["The fossil record of cyanobacteria. In Ecology of Cyanobacteria II (pp. 15-36). Springer Netherlands"](https://books.google.co.nz/books?id=4oJ_vi27s18C&pg=PA17&dq=%22Rusting+of+the+Earth%22&hl=en&sa=X&ei=V80JVZOTA8_i8AXpiYIo&ved=0CCIQ6AEwAQ#v=onepage&q=%22Rusting%20of%20the%20Earth%22&f=false) In: Brian A. Whitton (Ed) *Ecology of Cyanobacteria II: Their Diversity in Space and Time*, Springer Science & Business Media.
4. [Tandem Repeat](http://www.nlm.nih.gov/cgi/mesh/2011/MB_cgi?mode=&term=Tandem+Repeat) at the US National Library of Medicine [Medical Subject Headings](http://en.wikipedia.org/wiki/Medical_Subject_Headings) (MeSH)
5. <http://www.nature.com/scitable/topicpage/tandem-repeats-and-morphological-variation-40690>
6. <https://tandem.bu.edu/trf/trf.html>
7. <http://www.people.vcu.edu/~elhaij/bnfo301-15/Units/Genome-analysis/repeat-search.pdf>
8. <http://www.people.vcu.edu/~elhaij/bnfo301-15/Units/Genome-analysis/tandem-repeats-2014.pdf>
9. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC148217/>
10. <http://mic.sgmjournals.org/content/144/5/1189.short>
11. <http://www.biomedcentral.com/1471-2180/1/2/>
12. <http://biobike.csbc.vcu.edu/>