

# EteRNA University - How to Build a Bot

26 Feb, 2014

Teacher: Jnicol

Boiled down lesson - Work in progress ([Raw original chat can be found here](#))

Hi everyone, Welcome to EteRNA Class Time!

In this class we will learn how to Build A Bot. A bot, in EteRNA, is a computer program that will automatically solve a puzzle. We want to pass the dot-bracket structure to the bot and expect an RNA sequence to be returned that will correctly fold into this structure.

For instance, if we pass the structure `(((((....))))` to the bot, we would expect a sequence like **CCCAAAAGGG** to be returned. Since this is an interactive class, lets open some windows you will need to participate.

Open the scripting interface window where the bot's code will be written,

<http://eterna.cmu.edu/web/script/create/>

I would like to use firedrake's suggestion and have the class try to work out the code for each task, and if needed, I will help out after a few minutes. The first task is to use the scripting interface window to input a structure (in the input box) and have that same structure string returned in the Result console.

jnicol: did anyone create an input box and display the structure?

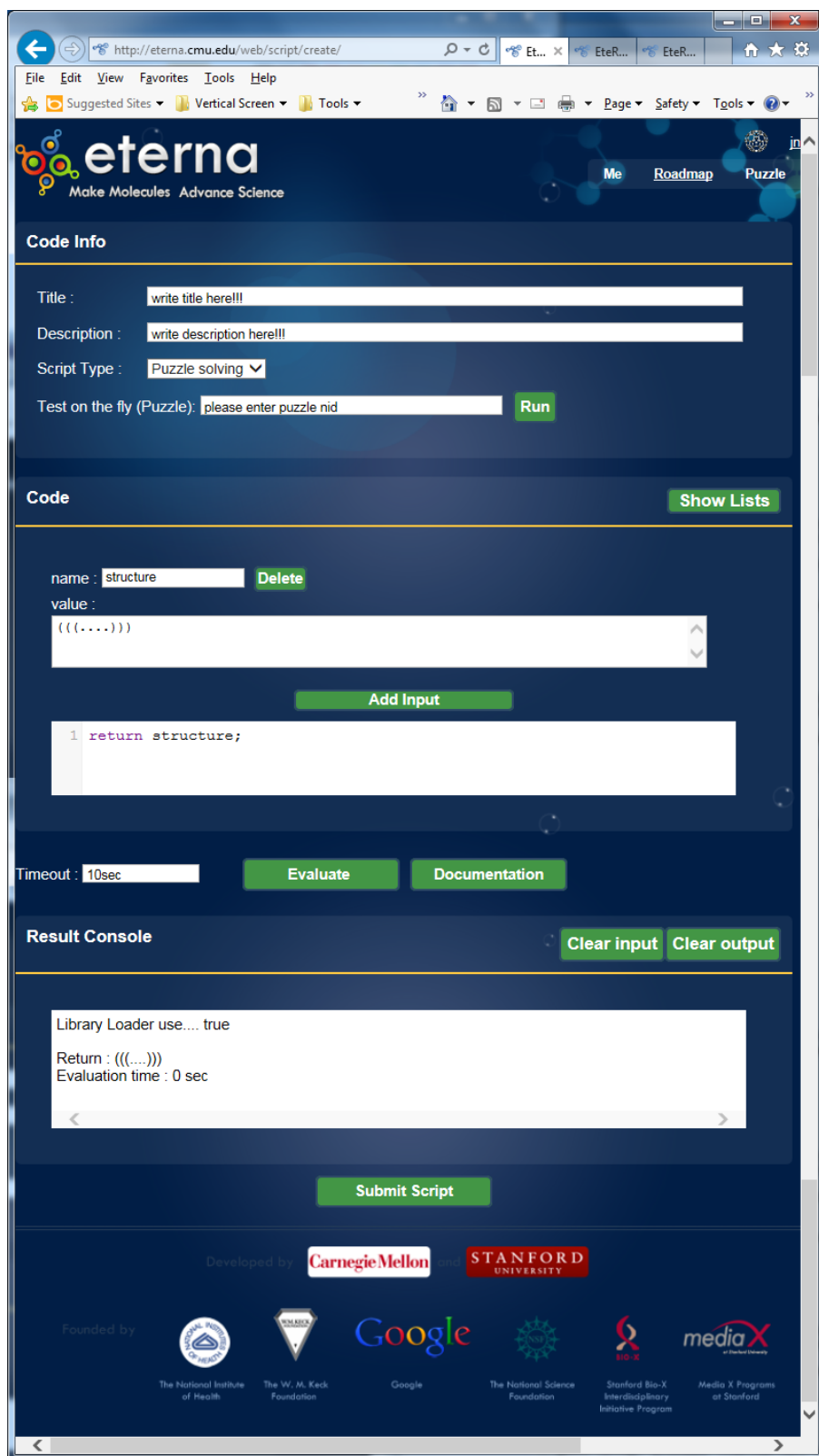
firedrake969: How do I read the input?

jnicol: when you name the input box

jnicol: use the same name in your program

firedrake969: So var input = [what?]

This is what the code window should look like,

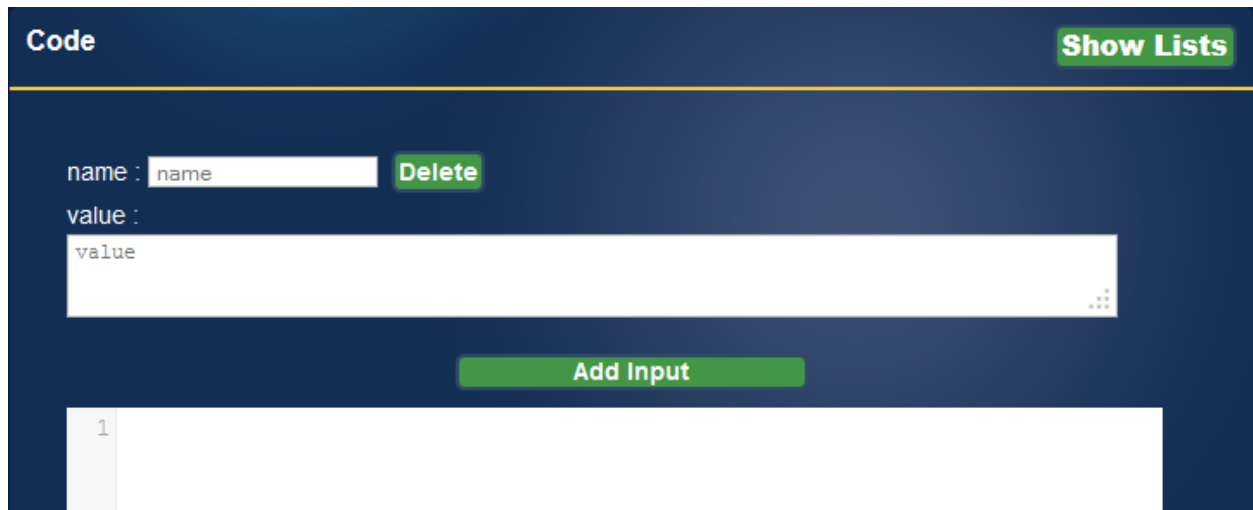


[http://amelior.com/eterna/classime/Class3\\_Task1.png](http://amelior.com/eterna/classime/Class3_Task1.png)

Here is how to do it. Click add input.



This will give you some fields where you can start making a program.



Big or small letter matters. If you write return sequence with big R, it will not become activated.

firedrake969: outln(structure);

jnicol: thats all you need for now!

jnicol: the return value will be read when we run test puzzles against the script

jnicol: so thats why we don't want to just use outln. Now, instead of returning the structure, lets return a sequence. For the second task, replace all ( with a C, all ) with a G and all . with an A.

Javascript has a function, replace to make this task easier.

Using **structure.replace('(', 'C');** will replace the first ( with C

Using **structure.replace(/[(]/g, 'C');** will replace all occurrences of ( with C

### Homework

An exercise after this class is to look up the javascript function replace to understand the syntax better.

The Script Documentation List says:

**Lib.replace(sequence, index, target)**

return sequence which character at index in sequence replaced to target

ex) **Lib.replace("AGCU", 2, "A") = "AGAU";**

firedrake969: How does the second replace thing work? It seems the same as the first besides the `/[()/g`

firedrake969: instead of just `"(`

jnicol: every occurrence of `"(` will be replaced

firedrake969: And the second one?

jnicol: for the second

jnicol: the first one will just replace the first occurrence of `"(`

firedrake969: Why does it have `/g` at the end?

jnicol: `g` means global

jnicol: Since `(` and `.` are special characters in javascript, a way to escape them is by enclosing them in `[]`

firedrake969: And how do I get the letter of a a string knowing its position within a string? jnicol: `structure[0]`

firedrake969: WHen I put `/[()/g` into the first arg of `replace()`, it turns red, like there's an error

jnicol: are you using `outln` or `return`?

firedrake969: `return`

firedrake969: `structure.replace(/[()/g, "C");structure.replace(/[./]g, "A");structure.replace(/[()/g, "G");return structure;`

jnicol: very nice!

firedrake969: It gives me the original structure though

jnicol: I came up with code similar, `return structure.replace(/[()/g, 'C').replace(/[()/g, 'G').replace(/[./]g, 'A');`

firedrake969: For some reason, mine won't work

jnicol: `structure` doesn't change, only what is returned from the `replace` call

firedrake969: So should I give it a new variable that can change?

firedrake969: `var newStruct = structure;newStruct.replace(/[()/g, "C");newStruct.replace(/[./]g, "A");newStruct.replace(/[()/g, "G");return newStruct;`

jnicol: you would need to do something like `sequence=structure.replace(...`

firedrake969: Still returns the original structure

jnicol: then `return sequence;`

firedrake969: `var sequence = structure;sequence.replace(/[()/g, "C");sequence.replace(/[./]g, "A");sequence.replace(/[()/g, "G");return sequence;`

firedrake969: Still gives the original structure

jnicol: `newstruct=structure.replace(/[()/g,'C'); return newstruct;`

firedrake969: Oddly, that works.

firedrake969: OH ok thanks

firedrake969: I know

jnicol: not oddly :)

firedrake969: I figured out what happened

jnicol: you need to assign the return value from the `replace` to a variable

firedrake969: `var sequence = structure;sequence = sequence.replace(/[()/g, "C");sequence = sequence.replace(/[./]g, "A");sequence = sequence.replace(/[()/g, "G");return sequence;`

Code

Show Lists

name :

Delete

value :

Add Input

```

1 var sequence = structure;sequence = sequence.replace(/[(]/g, "C");
2 sequence = sequence.replace(/[.]/g, "A");
3 sequence = sequence.replace(/[)]/g, "G");
4 return sequence;

```

jnicol: ok, good

firedrake969: I just didn't reset it

jnicol: yes

firedrake969: yay I passed my first puzzle

Result Console

Clear input Clear output

```

Library Loader use.... true
===== Test on the fly results with puzzle(20102).... =====
-- Puzzle title : Hairpin Series
-- Puzzle creator : Sneh
Script's Return : CCCCCAAAAAGGGGG
Script's Folded : (((((.....))))))
Puzzle secstruct : (((((.....))))))
Result : PASS
Evaluation time : 0 sec

```

firedrake969: I did Test on the fly with ID 20102

**Code Info**

Title :

Description :

Script Type :

Test on the fly (Puzzle):

jnicol: so you get CCCAAAAGGG as a return?

**Result Console**

```
Library Loader use.... true  
  
Return : CCCAAAAGGG  
Evaluation time : 0 sec
```

firedrake969: Yeah

jnicol: Lets take a step back and see what we have.

firedrake969: Well, 5C and 5xG

jnicol: We have input a structure and returned a sequence that we hope will fold into that structure

firedrake969: Because I had 5x( and 5x)

jnicol: Congratulations, you have just created your first bot!

jnicol: To prove this, we can enter some puzzle ids in Test On The Fly input box and see that our bot can solve them.

jnicol: If we order the challenge puzzles by number of solvers, the first and easiest puzzle is

<http://eterna.cmu.edu/web/puzzle/20111/>

jnicol: take the puzzle id number, 20111, and put it in the Test on the Fly input box and click run.

jnicol: You should get the following result,

http://eterna.cmu.edu/web/script/create/

File Edit View Favorites Tools Help

Suggested Sites Vertical Screen Tools

Find: replace Previous Next Options

**eterna**  
Make Molecules Advance Science

Me Roadmap Puzzle

### Code Info

Title :

Description :

Script Type :

Test on the fly (Puzzle):

### Code

name :

value :

```
1 return structure.replace(/([ ]/g, 'C').replace(/([ ]/g, 'G').replace(/([ ]/g, 'A').replace(/([ ]/g, 'U')
```

Timeout :

### Result Console

```
Library Loader use.... true
===== Test on the fly results with puzzle(20111).... =====
-- Puzzle title : Simple Hairpin
-- Puzzle creator : Sneh
Script's Return : CCCCCAAAAAGGGGG
Script's Folded : (((((.....))))))
Puzzle secstruct : (((((.....))))))
Result : PASS
Evaluation time : 0.001 sec
```

[http://amelior.com/eterna/classtime/Class3\\_Task2.png](http://amelior.com/eterna/classtime/Class3_Task2.png)

jnicol: If we actually Submit Script, we get an additional button, See Results of Player Puzzles, which randomly tests many puzzles. As an example, lets use Meechl's O Christmas Tree script, <http://eterna.cmu.edu/web/script/3223005/>

Click on the See Results of Player Puzzles, and we see that many puzzles are tested,

<http://amelior.com/eterna/classtime/MeechIsScriptResults.png>

This script solves 50% of the player puzzles tested and is a great example of how GC pairs can be flipped to solve puzzles. A good exercise would be to read through Meechl's code to see how this was accomplished. The next step for our bot is initialize the sequence with alternating GC/CG pairs. The scripting interface a very useful function for this task, getPairmap.

firedrake969: What does `do{for}` do?

```
firedrake969: i.e. do(condition){for(condition){}
```

jnicol: it is a loop to repeat itself when the sequence does not fold properly

Lets see what **getPairmap** can do. Copy this code into the scripting interface window. (Delete previous code)

```
rnastruc = new RNA(structure);  
return rnastruc.getPairmap(structure);
```

Enter (((...))) in the structure input box and click **Evaluate**

firedrake969: Oh, so you can make a new RNA design an object?

jinicol: there are many functions the devs have provided for the RNA design, yes

firedrake969: Ok, so what does `getPairmap()` do?

jinicol: You should see this result: Return : 9,8,7,,,,,2,1,0 from the code example

jinicol: Notice the bases that Puzzlemaker labels as 1-10, javascript labels as 0-9

inicol: We see from the results that base 0 is paired to base 9, base 1 to 8 and base 2 to 7.

Bases 3,4,5,6 do not have pairs. If you open puzzlemaker, it maybe easier to see and grab the structure

<http://eterna.cmu.edu/game/puzzlemaker/>

firedrake969: I input this

firedrake969: ((((((.....((((((.....((((((.....((((((.....((((((.....)))))))))....

```
firedrake969: And it returned some bases that were 2-in-a-row unpaired.
```

firedrake969: nvm. I see

firedrake969: But what about the trailing RNA unpaired strand?

jinicol: i get

69,68,67,66,65,,64,63,62,61,60,,59,58,57,56,55,,54,53,52,51,50,,49,48,47,46,45,,44,43,42,41,4

0,39,,,35,34,33,32,31,30,28,27,26,25,24,22,21,20,19,18,16,15,14,13,12,10,9,8,7,6,4,3,2,1,0

firedrake969: Same

firedrake969: It's 74 bases long...

firedrake969: but it only returns 70

firedrake969: And cuts off the last four bases

jnicol: ah, looks like trailing unpaired bases has a bug

jnicol: we need to report that to jee

jnicol: I like to use an array to hold the sequence for each base.

jnicol: The reason is because if I set sequence = "CCCCAAAGGG";

firedrake969: You can't access it easily or change it

jnicol: I cannot easily update a base, for example I cannot use sequence[3] = 'G'; With an array, I can update each base individually. The following code is an example of how an array can be used to hold and initialize the sequence.

```
rnastruc = new RNA(structure);
pairmap = rnastruc.getPairmap(structure);
sequence = Array();
for (i=0; i<structure.length; i++) {
    sequence[i] = 'A';
}
return sequence.join("");
```

jnicol: You can use this code as a starting point to generate a sequence with alternating GC pairs. Instead of initializing every base to A, first test the pairmap, then decide if an A, G or C should be used.

firedrake969: rnastruc = new RNA(structure);rnaPairmap = rnastruc.getPairmap(structure);var baseArray = [];for (var i = 0; i < sequence.length; i++) { baseArray[i] = "A";}

firedrake969: Did something slightly different

jnicol: I like your initiative!

firedrake969: What's the advantage of using var newArray = Array(); rather than [whatever] = [];

firedrake969: It seems the same, as both set up an empty array

jnicol: not sure, I usually use =Array()

firedrake969: Ok. So it's just a matter of style?

jnicol: yes, coding preference

### **There is no correct code, only a correct result**

Next task is to create a sequence with alternating GC pairs

firedrake969: For any structure, or explicitly?

jnicol: any structure that is input

firedrake969: So would one approach be to separate the getPairmap into a list somehow?

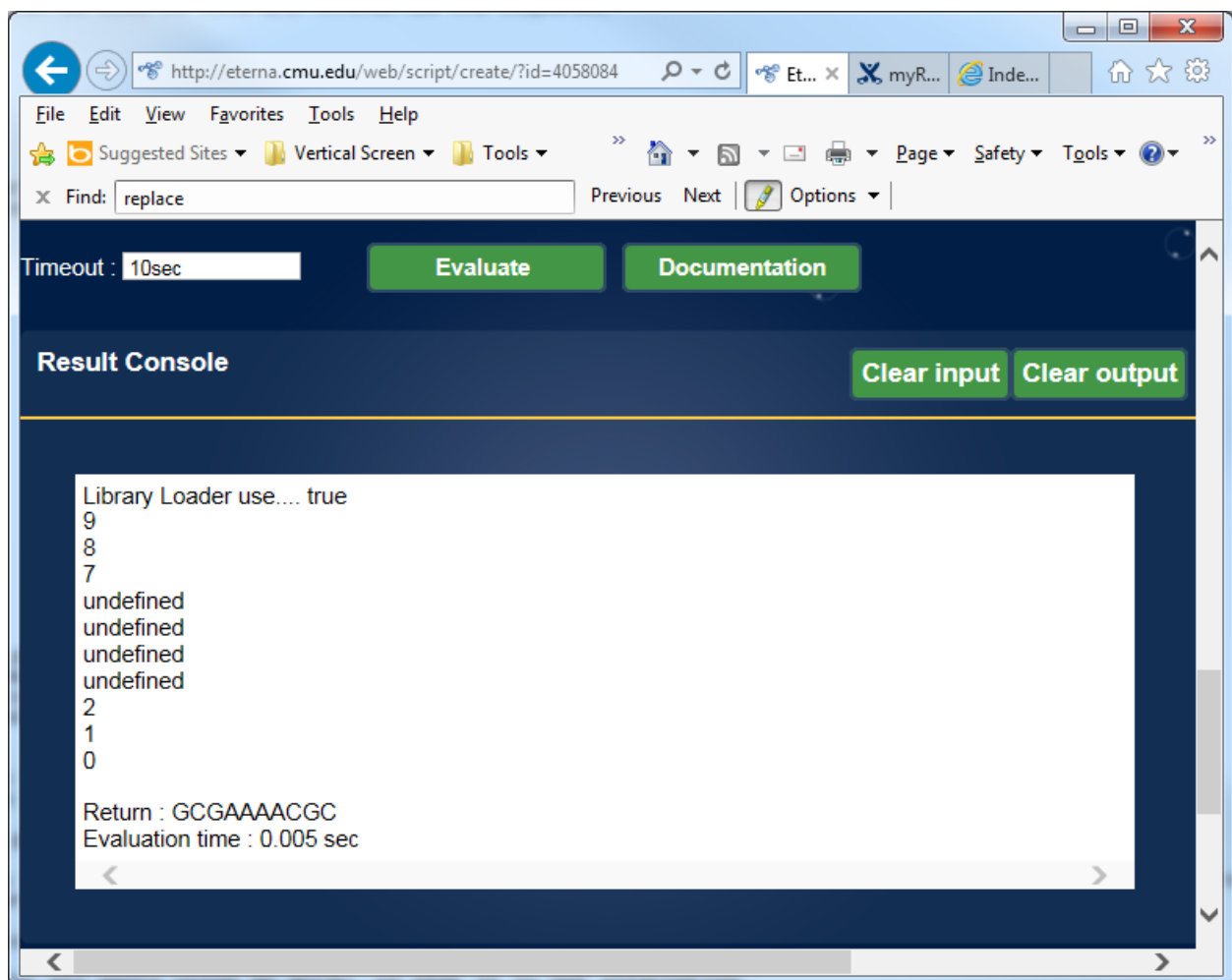
jnicol: correct

firedrake969: It's kind of hard because it's just ,, rather than , ,

firedrake969: no spaces or anything in between  
jnicol: try this

```
rnastruc = new RNA(structure);  
pairmap = rnastruc.getPairmap(structure);  
sequence = Array();  
for (i=0; i<structure.length; i++) {  
  outln(pairmap[i]);  
}  
return sequence.join("");
```

It shows what the pairmap can be compared to:



[http://amelior.com/eterna/classtime/Class3\\_Task3.png](http://amelior.com/eterna/classtime/Class3_Task3.png)

firedrake969: How does it handle the undefineds and skip the commas?  
jnicol: commas are not part of the array, they just display when converted to a string

```
jnicol: if (pairmap[i] == undefined) { sequence[i] = 'A'; } for instance
```

firedrake969: Ok and is it undefined or "undefined" (a string)?

jnicol: the way the dev programmed getPairmap, it is undefined (without quotes), like null

Here is the code I came up with for this task:

```
mastruc = new RNA(structure);
pairmap = rnastruc.getPairmap(structure);
sequence = Array();
nextbase = 'G';
for (i=0; i<structure.length; i++) {
  outln(pairmap[i]);
  if (pairmap[i] == undefined) {
    sequence[i] = 'A';
  }
  else if (pairmap[i] > i) {
    sequence[i] = nextbase;
    if (nextbase == 'G') { nextbase = 'C'; }
    else { nextbase = 'G'; }
    sequence[pairmap[i]] = nextbase;
  }
}
return sequence.join("");
```

jinicol: Notice that I put in an **outln(pairmap[i]);** statement, which is just for debugging purposes. **outln** is probably the best tool we have for now to find out what is not working as we expect it to.

firedrake969: (based on some of that code I've seen)

```
firedrake969: nnastruc = new RNA(structure);pairmap = nnastruc.getPairmap(structure);var
bases = []var nextBase = "G"for (var i = 0; i < structure.length; i++) { if (pairmap[i] == undefined)
{ bases[i] = "A"; } else if (nextBase == "G") { bases[i] = "G"; nextBase = "C"; } else if
(nextBase == "C") { bases[i] = "C"; nextBase = "G"; }}outln(bases);
```

firedrake969: Sorry for the horrible spacing :P

jinicol: Make sure you have semicolons ; at the end of each line, they all did not display

firedrake969: Odd

firedrake969: Oh, OK, fixed.

firedrake969: Do I included ; on the end of if and for?

jinicol: no

firedrake969: Thanks for the reminder

firedrake969: My code on this structure

$$(((((((.( (((((.( ((((((.( ((((((.( ((((((.( ((((((.( ...)))))))))...)$$

firedrake969: gives

G,C,G,C,G,A,C,G,C,G,C,A,G,C,G,C,G,A,C,G,C,G,C,A,G,C,G,C,G,A,C,G,C,G,A,A,A,C,G,C,  
G,C,G,C,G,C,G,C,G,C,G,C,G,C,G,C,G,C,G,C,G,C,G,C,G,C,G,C,A,A,A,A

jnicol: looks good!

firedrake969: Going to implement something that checks if it works

jnicol: this might be a good place to stop, thats the next step for our bot. I actually made a script that I was hoping to get up to tonight, but it can wait until next week to explain it.

<http://eterna.cmu.edu/web/script/4058084/>