REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
18-02-2014	R&D Status Report (Quarterly)	17-11-2013 to 16-02-2014

4. TITLE AND SUBTITLE
Construction of a Bacterial Cell that Contains Only the Set Of Essential Genes Necessary to Impart Life

5a. CONTRACT NUMBER
HR0011-12-C-0063

5b. GRANT NUMBER
NA

5c. PROGRAM ELEMENT NUMBER
NA

6. AUTHOR(S)
John Glass

Tony Yee

5d. PROJECT NUMBER
NA

5e. TASK NUMBER
NA

5f. WORK UNIT NUMBER
NA

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
J. Craig Venter Institute
9704 Medical Center Drive
Rockville, MD 20850

8. PERFORMING ORGANIZATION REPORT NUMBER
HR0011-12-C-0063.7

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
DARPA, MTO
Prog: Living Foundries ATGC
675 N Randolph St
Arlington, VA 22203

10. SPONSOR/MONITOR'S ACRONYM(S)
DARPA

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
NA

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited

13. SUPPLEMENTARY NOTES
NA

14. ABSTRACT

We have made significant progress on all fronts of the project. On the Top Down approach to genome minimization, we have used the Essential (E), Non-essential (N), and Impaired (I) gene categories to make steady progress with gene and gene cluster deletions. To date, we have removed approximately 283 kb from the Mycoplasma mycoides JCVI-syn1.0 genome by this method. The resultant 795 kb genome is viable and grows with a normal growth rate.

Our main effort has been focused on the Bottom Up approach. Results are substantial. We have continued work on the reduced genome design (RGD) based on the N/E/I categorization system and completed testing on all 8 segments from the RGD design. All segments are viable in a 7/8th wild type background. In several experiments, we have also tested genomes containing various combinations of minimized segments. We have multiple genomes with four or more minimized segments that are viable. The smallest genome to date is RGD 24678, at 696 kb.

The effort to modularize the genome is in progress. A tRNA module was constructed and inserted into the M. mycoides syn1.0 genome. The resulting genome is viable following transplantation. Construction of segments with the natural tRNA loci deleted is in progress.

Preliminary work aimed at genome complementation has continued. The use of a knock-in strain for re-inserting PCR amplified genes is under discussion as an approach.

15. SUBJECT TERMS
NA

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Tony Yee
UU	UU	UU	UU	9	**19b. TELEPHONE NUMBER** *(include area code)* 301 795 7133

J. Craig Venter INSTITUTE

Construction of a Bacterial Cell that Contains Only the Set of Essential Genes Necessary to Impart Life

Report Title:	R&D Status Report (Quarterly)
Report Number:	HR0011-12-C-0063.7
Reporting Period:	November 17, 2013 to February 16, 2014
Contract No.:	HR0011-12-C-0063
Performing Organization:	J. Craig Venter Institute 9704 Medical Center Drive Rockville, MD 20850 USA
Author of Report:	John Glass, Anthony Yee
Principal Investigators:	Hamilton Smith, Clyde Hutchison

Abstract

We have made significant progress on all fronts of the project. On the Top Down approach to genome minimization, we have used the Essential (E), Non-essential (N), and Impaired (I) gene categories to make steady progress with gene and gene cluster deletions. To date, we have removed approximately 283 kb from the *Mycoplasma mycoides* JCVI-syn1.0 genome by this method. The resultant 795 kb genome is viable and grows with a normal growth rate.

Our main effort has been focused on the Bottom Up approach. Results are substantial. We have continued work on the reduced genome design (RGD) based on the N/E/I categorization system and completed testing on all 8 segments from the RGD design. All segments are viable in a 7/8th wild type background. In several experiments, we have also tested genomes containing various combinations of minimized segments. We have multiple genomes with four or more minimized segments that are viable. **The smallest genome to date is RGD 24678, at 696 kb.**

The effort to modularize the genome is in progress. A tRNA module was constructed and inserted into the *M. mycoides* syn1.0 genome. The resulting genome is viable following transplantation. Construction of segments with the natural tRNA loci deleted is in progress.

Preliminary work aimed at genome complementation has continued. The use of a knock-in strain for re-inserting PCR amplified genes is under discussion as an approach.

Sponsored by
Defense Advanced Research Projects Agency
Microsystems Technology Office (MTO)
Program: Living Foundries: Advanced Tools and Capabilities for Generalizable Platforms (ATCG)
Issued by DARPA/CMO under Contract No: HR0011-12-C-0063

Approved for Release; Distribution is Unlimited

Table of Contents:

Summary

The goal of the project is to create a cell that contains only the set of genes that are essential for life under ideal laboratory conditions. We are working to minimize *Mycoplasma mycoides* JCVI-syn1.0 (the synthetic version of *Mycoplasma mycoides* subsp *capri*) using two approaches:

- Top Down: remove genes and clusters of genes one (or a few) at a time, proceeding only if the reduced strain is viable, with a reasonable growth rate
    - We previously reported the results of a transposon study was conducted and allowed us to categorize genes as Essential (E), Non-essential (N), or Impaired (I)
        - This categorization scheme has been the basis of most of our subsequent work
    - The *M. mycoides* genome has been reduced to 795 kb using the Tandem Repeat Endonuclease Cleavage (TREC) strategy
    - By making gene cluster deletions in the 1/8th genome segments and assembling the segments into a genome, a 790 kb genome was made
        - The genome was viable, but had a doubling time of 82 minutes
- Bottom Up: design a reduced genome based on our best Tn5 gene disruption and deletion data (RGD), and synthesize it
    - Synthesis from oligonucleotides
        - All 1/8th genome molecules have been tested and found to be viable
        - Repairs to Segment 1 were completed and verified
        - Viable genomes containing multiple RGD segments have been constructed
        - **The smallest genome to date is 696 kb.**

The initial modularization experiments are progressing. A 5.3 kb tRNA module containing the 30 tRNA genes plus the necessary promoters and terminators was constructed and sequence verified. The module was inserted into the genome and found to be viable.

Work has continued on the genome complementation front. In order to progress as quickly as possible, we will predict deleted genes that may be problematic, PCR amplify the genes and assemble them into a cassette, and insert them into a knock-in strain for transplantation.

Introduction

The goal of this research project is to build a minimal bacterial cell that contains only the genes that are required for life in ideal laboratory conditions. The pursuit of a minimized cell is critical to the advancement of biology, both as a pathway for understanding the basic requirements for replication and as a chassis for creating an optimized platform for any number of possible applications.

We previously reported that the *Mycoplasma mycoides* JCVI-syn1.0 genome was successfully reduced from 1078 kb to 779 kb; however, while the 779 kb genome was viable, the growth rate

was far too slow to allow follow up experiments at an acceptable pace. Using the N, E, I gene categories, the genome has been reduced to 795 kb, but with a normal doubling time.

Synthesis of a newly designed reduced genome (RGD, 539 kb) was completed. All 8 of the 1/8th genome segments tested have proven to be viable. Assembly of genomes with combinations of RGD segments has been in progress since the last report. We now have several genomes that contain up to 5 RGD segments, including the smallest genome to date at 696 kb.

A preliminary tRNA module was designed, constructed, and introduced into the *M. mycoides* syn1.0 genome and found to be viable. Synthesis of segments with the natural tRNA loci removed is in progress.

Experiments on a transposon-based genome complementation system are being conducted.

Methods, Assumptions and Procedures

TOP DOWN APPROACH

The plan here was to start with the full size 1078 kb *M. mycoides* JCVI-syn1.0 synthetic genome. We have continued to use the TREC strategy to make iterative deletions in the mycoplasma genome. Targeting the N category genes and clusters is proving to be effective (further discussed in the Results and Discussion section). We have made a series of strains that are progressively reduced with little to no reduction in growth rate.

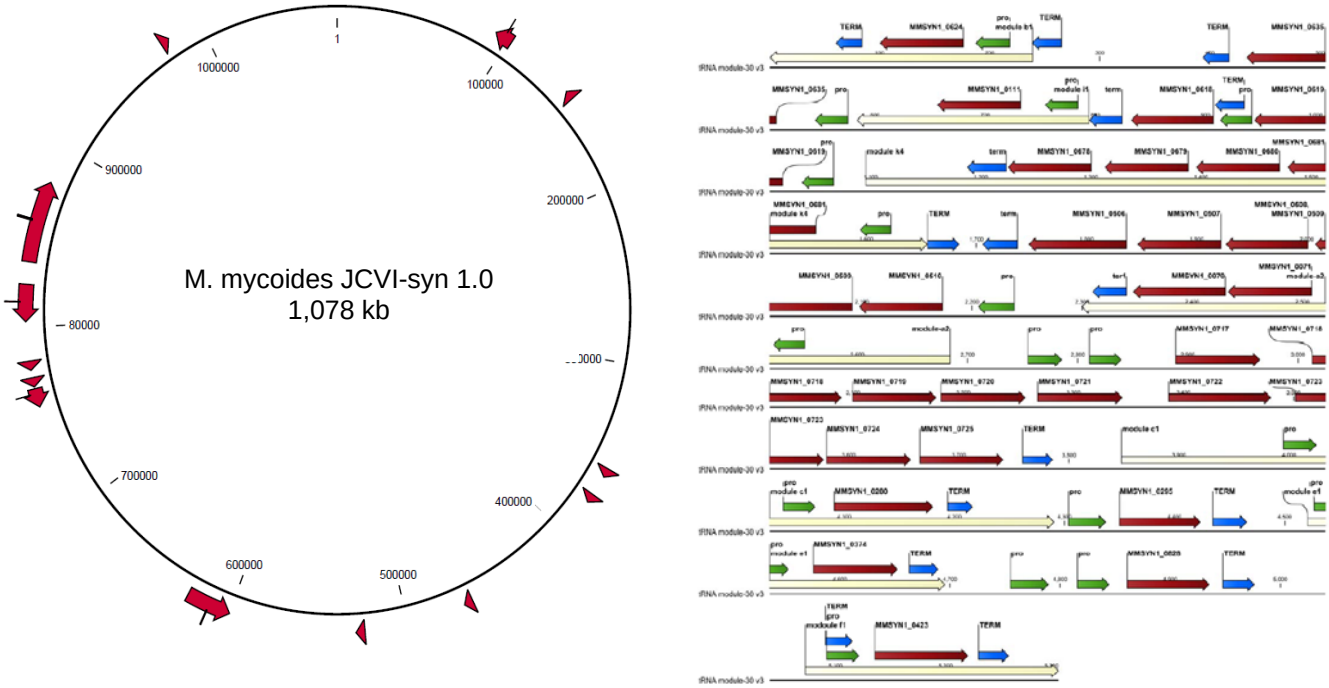
BOTTOM UP APPROACH

Synthesis from oligonucleotides: A new eight piece genome design was completed using the N, E, I gene classification system (RGD, 539 kb). All 8 have been found to support cellular life. We have continued testing combinations of reduced genome segments. We now have several genomes that contain up to 5 RGD segments.

MODULARIZATION

To test gene modularization, we have organized the 30 tRNA genes of *M. mycoides* into a single contiguous module. The module contains the coding regions, as well as the promoters and terminators needed for regulation. The tRNA genes are naturally distributed around the genome in 13 loci.

Figure 1



(a) Natural distribution of tRNA genes in M. mycoides. *The tRNA gene clusters have been enlarged in Fig.1(a) to show the direction of transcription. The M. mycoides JCVI-syn1.0 genome has 8 single tRNA genes and 5 clusters of 2 to 9 genes, for a total of 30.*

(b) tRNA module design. *The 30 tRNA genes have been relocated into a single module. (Green arrows represent promoters. Red arrows show tRNA genes. Blue arrows are terminators.)*

Each of the 13 loci was synthesized by PCR using syn1.0 as the template, cloned in E. coli and then joined together into a single cassette with appropriate yeast markers. The cassette was inserted into syn1.0 to replace the largest cluster of 9 tRNAs at 10 o'clock on the genome map. The resulting genome is viable after transplantation. We will synthesize segments with the 12 other tRNA loci around the genome removed from the design.

GENOME COMPLIMENTATION

We have begun initial experiments aimed at enabling genetic complementation to restore desirable phenotypes to deleted strains, for example, RGD. A system capable of quickly adding deleted genes back into a genome would be a powerful tool to help de-convolute growth-retarding synthetic effects.

We are initiating an approach in which we will predict the problematic genes, assemble them into a cassette and insert them into a knock-in strain to test for restored viability.

Results and Discussion

MODULARIZATION

We have moved ahead with modularization of the tRNA loci. Construction was difficult because the complex secondary structures of the tRNAs interfered with the synthesis process. The module is complete, sequenced, and inserted into the genome. The genome containing the tRNA module is viable, and we will proceed by synthesizing the genome with the natural loci removed from the design.

TOP DOWN APPROACH

Iterative deletions using the TREC based approach are making steady progress toward a minimal genome. A table outlining the progress to date is shown below. Since the last reporting period, strain D13 has been tested and found to be viable with a good qualitative growth rate (quantitative growth rate evaluation has not been performed).

Table 1

Strains	DT(min)	Genome Size (bp)	# of Genes Deleted
syn1.0	64	1,078,809	0
syn1.0D6 RE		1,062,183	17
DISs		1,048,690	31
D1		979,083	68
D2		969,069	74
D3		944,159	90
D4		931,710	97
D5		923,647	102
D6	67	908,931	108
D7		877,942	135
D8		866,271	155
D9	64	844,265	173
D10	65	828,901	181
D11		816,807	194
D12		805,506	201
D13		~794,666	209
D14*		~784,000	214

** Planned*

Sequential deletion of genes and clusters of genes is not the way that we ultimately expect to pursue a minimal genome design. While we are making steady progress in this approach, the greater outcome of performing the top down deletions is the generation of information regarding unanticipated interactions between elements of the genome that are difficult or impossible to identify using the Bottom Up approach. De-convolution of synthetic effects will be difficult work, which we hope to delay as long as possible with information gained from the top down approach.

BOTTOM UP APPROACH

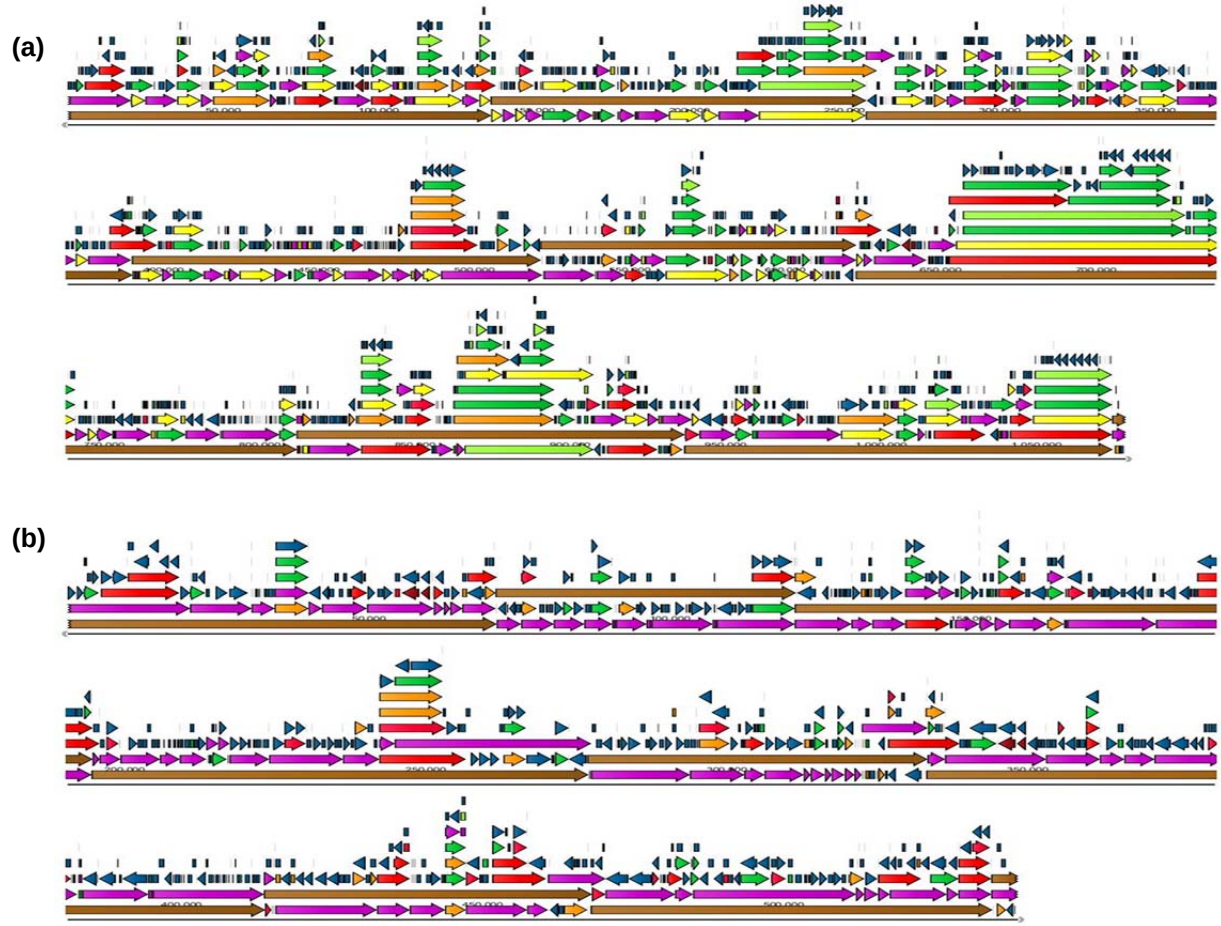
A new genome design (RGD) was completed using the N, E, I classification system. The design essentially involves removing all of the N (non-essential) genes and then checking to see that we have not disturbed the promoters and terminators necessary for expression of the remaining E (essential), and I (impaired growth if deleted) genes.

Figure 2, below, shows the contrast between (a) the *M. mycoides* JCVI-syn1.0 genome and (b) the new genome design.

Figure 2

(a)

(b)



(a) The 1,078 kb M. mycoides JCVI-syn1.0 genome, showing the genes retained in the new genome design (purple arrows)*. Note that genes in the new design are interspersed around the genome, comingled with genes that were removed from the design;* ***(b) The new design of a minimized, 539 kb M. mycoides genome****. Contrast the spacing between the retained genes (purple) with the spacing in (a).*

If a lethal deletion (red) falls entirely within a region deleted in the design (yellow) then there is an inconsistency between the design and the deletion data. Where the data from different deletions disagrees, green trumps red because failure to get a transplant is a negative result that could result from some problem with that particular transplantation reaction.

The table below shows the length of each 1/8th genome Not I segment in the *M. Mycoides* JCVI-syn1.0 genome, and the newly designed RGD genome. The new design represents a 50% reduction in total genome length.

Table 2

Not I Fragment #	*M. Mycoides* JCVI-syn1.0 Length (bp)	RGD Designed Length (bp)	(RGD1)/(*M. Mycoides* JCVI-syn1.0)
1	140,739	75,732	0.54
2	120,912	49,888	0.41
3	133,208	73,958	0.56
4	131,623	82,531	0.63
5	101,708	56,501	0.56
6	189,357	80,747	0.43
7	124,976	54,482	0.44
8	137,887	66,717	0.48
Total	1,080,410	540,566	
Overlaps	-1,601	-1,601	
Genome Length	1,078,809	538,955	0.50

In this reporting period, we have made a great deal of exciting progress on the RGD genome. All 8 of the RGD segments are viable; however, segment 1 grew slowly in a 7/8th wild type background.

This was due to the deletion of genes 31 and 33 (conserved hypothetical genes of unknown function). In this period, we have completed work to re-insert these two genes. The repaired segment 1 has been confirmed to be viable.

In addition to testing the individual segments in a wild type background, we have begun testing combinations of viable, minimized segments. The genomes are produced via 8-piece assembly in yeast, and transplanted. In contrast, individual segments are tested by swapping into a 7/8th landing pad strain. Viable genomes are listed below.

[Note: the strain nomenclature used below lists the reduced segments in that particular genome; wild-type segments are not listed (example: "RGD 1234" contains reduced segments 1-4 and wild-type segments 5-8)]

- RGD 26 (previously reported)
- RGD 268 (previously reported)
- RGD 2678
- RGD 24678, 696 kb
- RGD 25678, 698 kb
- RGD12478: This genome produces the expected PCR confirmation banding pattern, but is not correct when examined by restriction and CHEF gel analysis

Assembly of RGD 12345678 has been completed, but the genome was not viable. A repeat of the experiment is in progress. In addition, we are using new transposon studies and complementation to debug the genome.

Conclusions

Tasks from the Statement of Work:

Task 1: **Complete a detailed global Tn5 transposon mutagenesis insertion map.**
The Tn5 transposon insertion map was submitted with the initial quarterly report.

Due: Month 6; Status - complete

Task 2: **Delete up to 27 large gene clusters**
We have reduced the genome size of M. mycoides JCVI-syn1.0 from 1079 kb to 779 kb through the deletion of some 30 clusters, representing a ~30% reduction. We are moving forward with an 848 kb genome because it grows at a higher rate than the smaller 779 kb version.

Due: Month 12; Status – complete

Task 3: **Construct a preliminary modular map of the genome**
The design of a modular map of the genome is complete was presented in the May 2012 quarterly report

Due: Month 12; Status – complete

Task 4: **Make new transposon insertion map. Identify non-essential small 2-4 gene clusters. Delete small clusters.**
A new transposon study was performed and previously reported. A table showing the resulting N/E/I categorization system was presented in May 2013 (Deliverable 1). We continue to work toward a minimal cell chassis using multiple simultaneous approaches.

Due: Status – in progress. We propose delaying the completion of this task until the testing of the RGD is complete.

We have several transposon studies in progress at the present time, on genomes RGD 24678, RGD 1 and RGD 3. We believe that the synthetic effects are tied to segments 1 and 3. This should give us a clear picture of the scope of debugging that will be required for the construction of a minimal cell. We anticipate that the current transposon work will be completed by mid-March 2014.

Once the RGD is tested and viable, we will use Tn5 mutagenesis to discover if the classifications of any of the remaining genes have changed and to determine and report which additional genes might be removed (Deliverable 2).

Planned Activities for the Next Reporting Period

1. Continue the Top Down minimization of our synthetic genome (near term) and conduct a new transposon study on the minimized genome (within approximately 4-6 months)
2. Complete testing the RGD genome as part of our Bottom Up strategy.
3. Continue with verification and testing of the tRNA gene module.
4. Continue development of genome complementation strategies

Program Financial Status

In Process & Completed Tasks	Planned Expend	Actual Expend (Cumulative to Date)	% Budget Completion	At Completion	Latest Revised Estimate	Remarks
Task 1	$305,646	$305,646	100%	$305,646	$305,646	Completed
Task 2	$826,256	$798,351	97%	$798,351	$826,256	Completed
Task 3	$43,487	$43,487	100%	$43,487	$43,487	Completed
Task 4	$1,100,000	$774,630	70%	N/A	$1,100,000	N/A
Cumulative	$2,275,389	$1,922,114	85%	N/A	$2,275,389	N/A

There is no management reserve or unallocated resources. The financial data presented is current through Dec. 2013.

Based on the currently authorized work:

- Is current funding sufficient for the current fiscal year? Yes
- What is the next fiscal year funding requirement at current anticipated levels? $1,214,151.00
- Have you included in the report narrative any explanation of the above data and are they cross-referenced? Not applicable; current funding is sufficient for the current fiscal year.