



# N.N. BURDENKO VORONEZH STATE MEDICAL UNIVERSITY

The reverse translation from English to  
DNA: the way how to encrypt messages  
in the genome

ABCDEFGHIH 



PAWEL ANDREJEW, DEPARTMENT OF CLINICAL MEDICINE (Л-601)



IRINA ILINA, DEPARTMENT OF PEDIATRICS (П-507)

SYNTHETIC BIOLOGY, SYNTHETIC CHROMOSOME, JCVI-syn1.0 clone sMmYCp235-1,  
MOLECULAR BIOLOGY, MOLECULAR GENETICS, BIOTECHNOLOGY, GENE,  
GENOME, BIOINFORMATICS, LINGUISTICS, ENGLISH LANGUAGE

# THE FIRST EVER CREATED SPECIES, CONTROLLED BY A SYNTHETIC GENOME



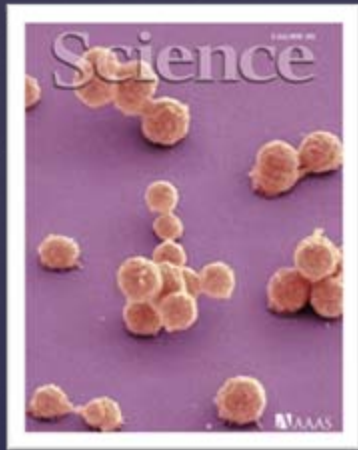
02.07.2010



ROCKWILL, MARYLAND, USA



INSTITUTE OF CRAIG VENTER



COVER of the SCIENCE journal:  
Electron micrograph of synthetic cells  
*Mycoplasma mycoides* JCVI-syn1.0  
(magnification ~25,000×)

THE YEAR 2010 WAS SIGNIFICANT FOR BIOLOGY  
IN THE LIGHT OF COMPLETE ARTIFICIAL  
SYNTHESIS OF GENOME AND ITS  
TRANSPLANTATION

[Gibson D. G., Glass J. I., Lartigue C., et al. Creation of a bacterial cell controlled by a chemically synthesized genome. *Science*. 2010;329:52–56. doi: 10.1126/science.1190719.]



A RESEARCH GROUP, LED BY THREE PROMINENT SCIENTISTS  
TRANSPLANTED DE NOVO SYNTHESIZED CHROMOSOME OF  
MYCOPLASMA MYCOIDES INTO MYCOPLASMA CAPRICOLUM



*Hamilton Othanel Smith*, Nobel prize laureate. Prominent for investigations of restriction enzymes. His studies revolutionized the methods of molecular biology and biotechnology

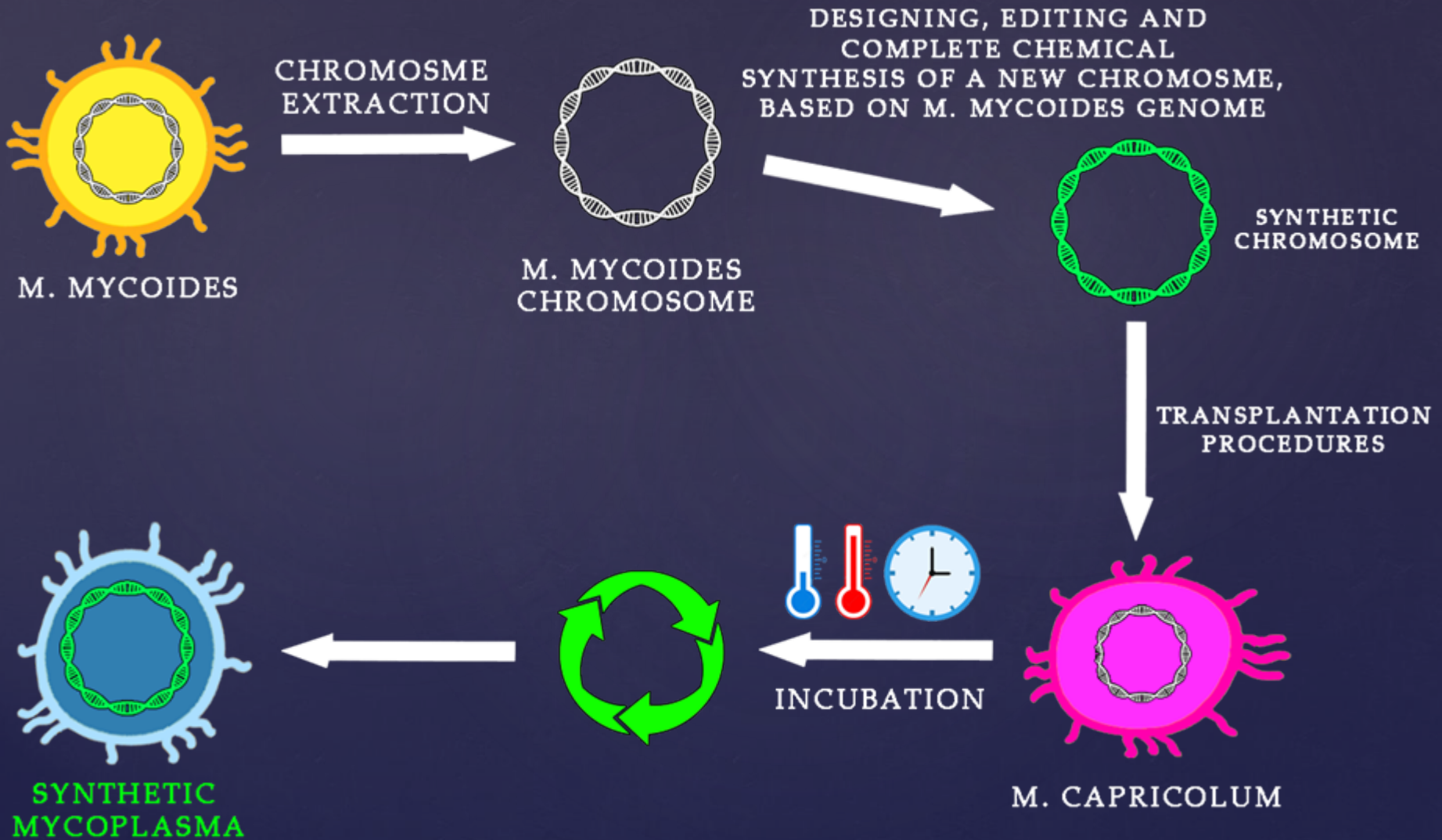


*Clyde A. Hutchison III*, prominent for investigations in sphere of site-directed mutagenesis. Participated in the determination of a complete genome sequence of phage  $\Phi$ X174. Made a contribution to the polymerase chain reaction discovery

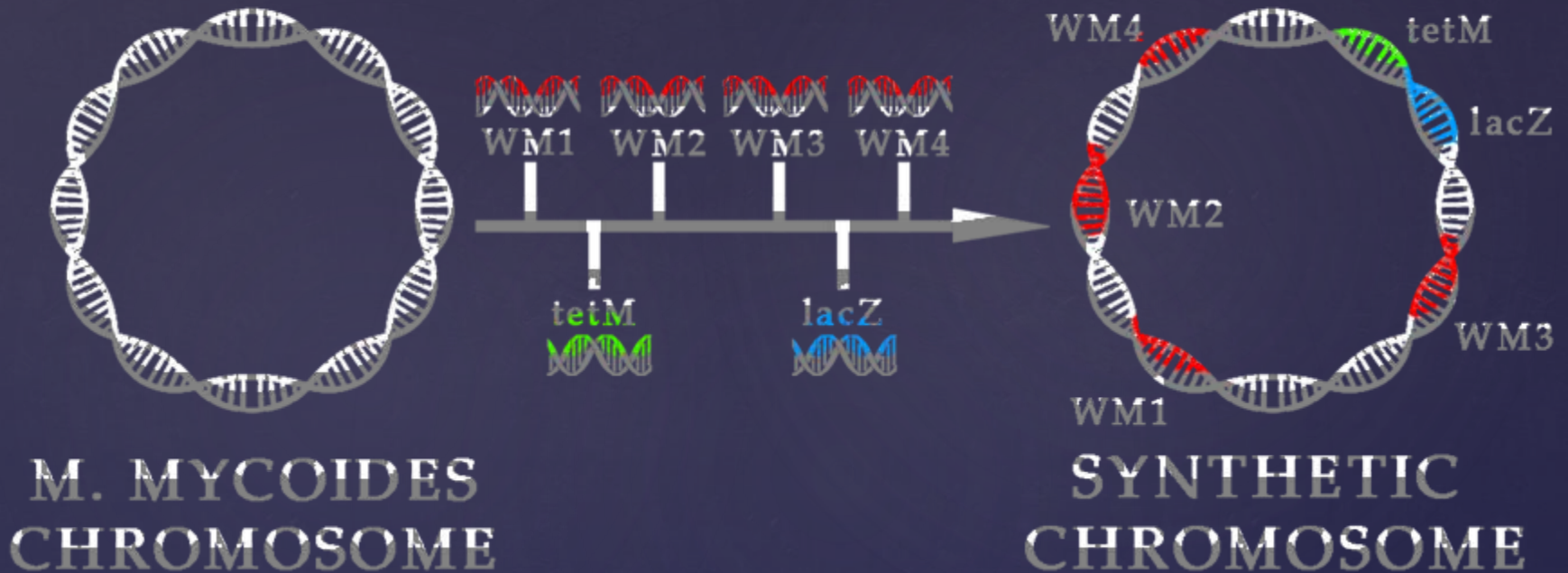


*John Craig Venter*, prominent for Human Genome Project and initial studies in synthetic biology

# THE CONCEPTION OF EXPERIMENTAL WORK



# DESIGN OF THE SYNTHETIC CHROMOSOME

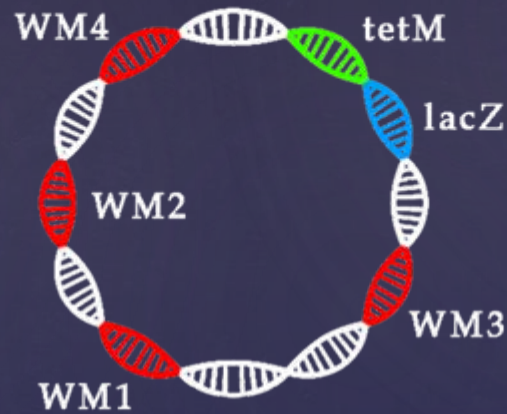
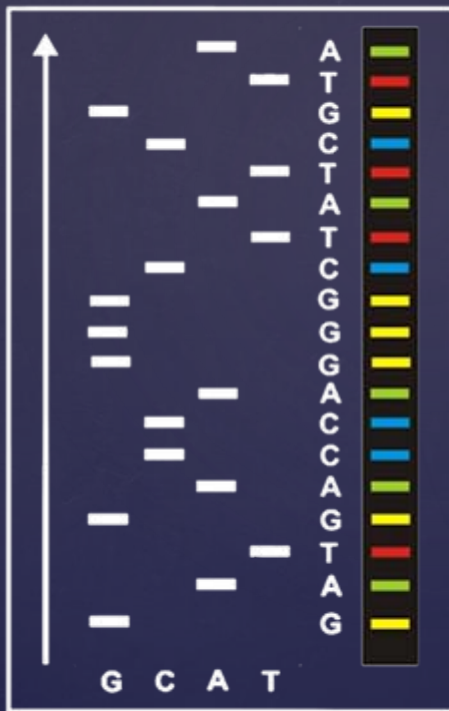


SYNTHETIC CHROMOSOME CONTAINS 6 EXTRA REGIONS: 4 WATERMARKS WITH DECODED MESSAGES IN ENGLISH LANGUAGE, TETRACYCLINE RESISTANCE GENE (*tetM*) AND BETA-GALACTOSIDASE GENE (*LacZ*)

In order to verify the success of the experimental work, the donor genome was edited by inserting the nucleotide sequence of **4 watermarks** into different loci of the genome, **antibiotic (tetracycline) resistance gene (tetM)** and **beta-galactosidase (lacZ) gene**. Thus, the efficiency of transplantation could be evaluated using three approaches:

1. Control sequencing of the synthetic genome and identification of watermarks
2. Testing viability of bacterial culture, treating their growth medium with antibiotic.
3. Turning the colony bright blue in the presence of an organic compound X-gal, metabolizing by a product of lacZ gene.

# 1. CONTROL SEQUENCING OF THE SYNTHETIC GENOME AND IDENTIFICATION OF WATERMARKS



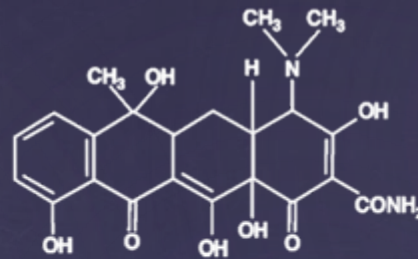
SYNTHETIC CHROMOSOME WITH WATERMARKS



INTACT CHROMOSOME (NO WATERMARKS)



## 2. TESTING VIABILITY OF BACTERIAL CULTURE, TREATING THEIR GROWTH MEDIUM WITH ANTIBIOTIC



TETRACYCLINE  
ANTIBIOTIC



SYNTHETIC CELLS ARE  
TETRACYCLINE-RESISTANT  
(tetM-positive)

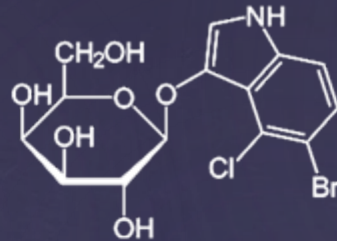


COLONY LYSIS  
(tetM-negative INTACT CELLS)





### 3. TURNING THE COLONY BRIGHT BLUE IN THE PRESENCE OF AN ORGANIC COMPOUND X-GAL, METABOLIZING BY A PRODUCT OF lacZ GENE



X-gal, metabolized  
by lacZ gene product  
(beta-galactosidase)



SYNTHETIC CELLS METABOLIZE  
X-gal (lacZ-positive),  
THE COLONY BECOMES BLUE



INTACT CELLS (lacZ-negative)  
DO NOT METABOLIZE X-gal,  
THE COLONY DOES NOT TURN BLUE



These three approaches confirmed success in transplantation, proclaiming that the first ever species was created by human. The linguistic interest of this great scientific breakthrough lies in **watermarks**, which were inserted into the donor genome. Each additional sequence encodes a message in **English**.

#### THE AIM OF THE STUDY:

- To biologically and linguistically analyze the adapt of English alphabet to a genetic code
- Mapping the exact localization of 4 watermarks in synthetic genome (WMs loci have never been published)
- Deciphering encoded data (**complete nucleotide sequence of the watermarks has never been published**; decoded data is **partially** available)



# MATERIALS AND METHODS

1. Analysis of C. Venter's book "**Life at the speed of light**" in order to get **cryptography** between English alphabet and a specific trinucleotide code, used to encipher it.
2. Building the **tagging sequences** (TSs) for pairwise alignment in order to identify **watermarks** in the synthetic chromosome. TSs were created without punctuation marks, except commas and spaces as .txt files and converted in FASTA format by **MEGA-X (Version: 10.1.7) tool**.
3. Dividing the sequence of synthetic chromosome in 5 fragments (Fs) in order to simplify and speed up the pairwise sequence alignment, using **DNASTAR Lasergene 17.01.1 (2020) MegAlign Pro**.
4. PSA (Pairwise Sequence Alignment) of Fs and TSs, completed by **DNASTAR Lasergene 17.01.1 (2020) MegAlign Pro** with **MAUVE** algorithm.
5. An extra PSA was completed by **DNASTAR Lasergene 17.01.1 (2020) MegAlign Pro**, using the **Smith-Waterman** DNA alignment algorithm [Matrix: "NUC44"; Gap penalty: 10; Gap extension penalty: 1].
6. Deciphering the watermarks codon by codon, moving to the 5'-end and 3'-end.
7. Overview of decoded watermarks.

# 1. ANALYSIS OF THE “LIFE AT THE SPEED OF LIGHT

The chapter 8 “**Synthesis of the M. mycoides Genome**” has a **key** for deciphering the nucleotide cryptography, used in watermarks. The series of characters, including complete English alphabet of 26 letters; newline; space; numbers (0-9); various punctuation marks, such as point, comma, hyphen etc; mathematical signs, such as plus, minus and some additional symbols were represented by the developed trinucleotide code.

## THE CRYPTOGRAPHY BETWEEN WATERMARKS AND ENGLISH ALPHABET, PUNCTUATION MARKS AND ADDITIONAL CHARACTERS

- A is encoded by TAG trinucleotide; B – AGT; C – TTT; D – ATT; E – TAA; F – GGC; G – TAC; H – TCA; COMMA (,) – GTG; POINT (.) – CGA; SPACE ( ) – ATA; 1 – CTT; 2 – ACT; 3 – AAT; @ - TCG etc.
- Total number of encoded characters – **64**, corresponding to the real number of DNA-codons.
- There is a **total mismatch** between codification of the English alphabet and one-letter code of amino acids. The only exception is in position **12**: letter L of JCVI-syn1.0 alphabet corresponds to leucine (Leu, L). Mismatch example: letter G in JCVI-syn1.0 alphabet is encoded by TAC, which in the real genetic code means M (Met, methionine) etc.
- **The table of cryptography** was created to make building the tagging sequences more convenient . It also was used for further decoding.

## 2. BUILDING THE TAGGING SEQUENCES (TSs) FOR PAIRWISE ALIGNMENT

Identification of watermarks in the synthetic genome required building the sequences to complete pairwise alignment. Because of the fact the whole sequences of watermarks remained **unpublished**, 4 specific tagging sequences (TSs) were created in order to identify corresponding fragments of each watermark in the chromosome. The TSs were built according to the only available data from “The life at the speed of light”. The chapter 8, “Synthesis of *M. mycoides* genome”, **fragmentally** provides some **insights** of what was encoded in each watermark. Thus, among encrypted data in WM<sub>1</sub> there is “Craig Venter Institute”; the quote from James Joyce: “To live, to err, to fall, to triumph, to recreate life out of life” in WM<sub>2</sub>; the quote of Robert Oppenheimer: “See things not as they are, but as they might be” in WM<sub>3</sub> and Richard Feynman’s quote: “What I cannot build, I cannot understand” in WM<sub>4</sub>. Considering these insights from the chapter 8 we hypothesized that “Craig Venter Institute” and the quotes above are suitable to be such tagging sequences for PSA. Using the cryptographic features from Table 1 TSs were created with no punctuation marks, except comas and spaces as .txt files and converted in FASTA format.



THE FILES CONVERTATION WAS COMPLETED BY MEGA-X (V:10.1.7)

### 3. DIVIDING THE SEQUENCE OF SYNTHETIC CHROMOSOME

Complete sequence of synthetic *Mycoplasma mycoides* JCVI-syn1.0 clone sMmYCp235-1 was found in NCBI Nucleotide database (**GenBank: CP002027.1**) and downloaded in FASTA. The sequence was divided in **5 fragments**, according to the schematic representation of watermarks in original paper "Creation of a bacterial cell controlled by a chemically synthesized genome". There are:

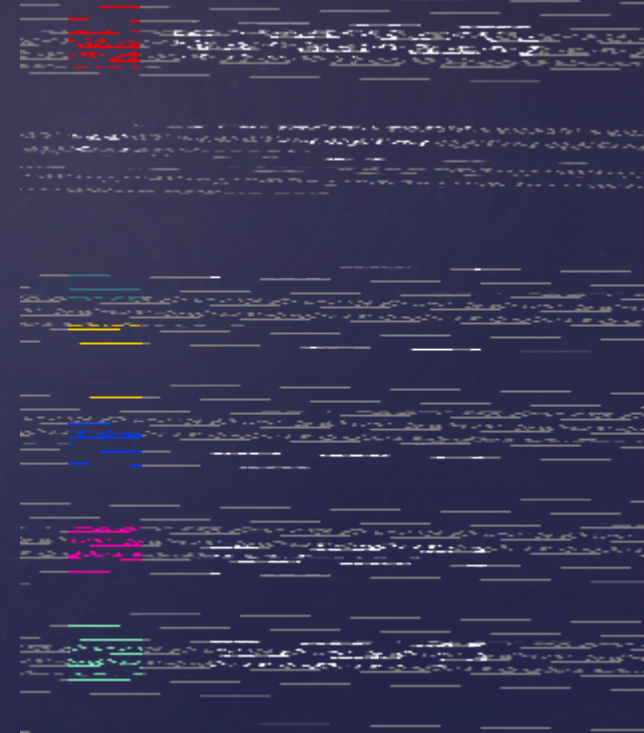
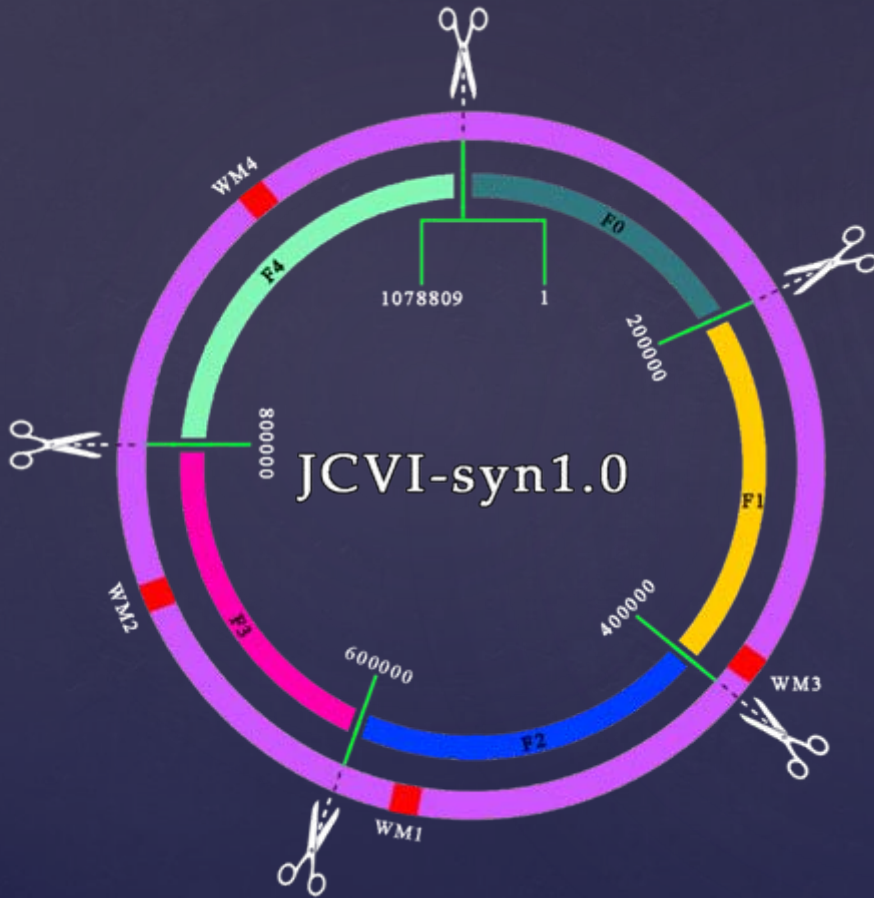
- **F<sub>0</sub>** (1→200000 nucleotides)
- **F<sub>1</sub>** (200001→400000)
- **F<sub>2</sub>** (400001→600000)
- **F<sub>3</sub>** (600001→800000)
- **F<sub>4</sub>** (800001→1078809)

This manipulation was completed in order to **simplify** and **speed up** the pairwise sequence alignment.



The sequence was fragmented, using DNASTAR Lasergene 17.01.1 (2020) MegAlign Pro

# Fragmentation of the complete sequence of synthetic *Mycoplasma mycoides* JCVI-syn1.0 clone sMmYCp235-1



**F<sub>0</sub>** (1→200000 nucleotides); **F<sub>1</sub>** (200001→400000); **F<sub>2</sub>** (400001→600000);  
**F<sub>3</sub>** (600001→800000); **F<sub>4</sub>** (800001→1078809)

#### 4. PSA (MAUVE ALGORITHM); 5. CONTROL PSA (SMITH-WATERMAN ALGORITHM). COMPLETED BY



DNASTAR Lasergene 17.01.1 (2020) MegAlign Pro

1. The first WM was identified via alignment of  $F_2$  and  $TS_1$ :

5'-TTTCTATAGCTGTACATATTGTAATGCTGATAACTAATACTGTGCGCTTGCAGCTGGAT  
CCTGATAA-3' → "Craig Venter Institute" ([565527](#)→[565592](#)).

CONTROL PSA WAS  
COMPLETED BY CLUSTAL  
SOFTWARE

2.  $WM_2$  was found via alignment of  $F_3$  and  $TS_2$ :

5'-

TGACGTAACCTGTTGTAATGACGTAACTACTATGACGTGGCTAGAACAACCTGACGT  
TGACTACTGTCCCAAACATCATGACGTCTATAATTTCTATAATAGTGATAAAACCTGG  
GCTAACGTTCCCTGACGTGGCAACCTGGGCTAA-3' → "To live, to err, to fall, to triumph,  
to recreate life out of life", James Joyce ([726295](#)→[726492](#)).

3.  $WM_3$  was identified via  $TS_3$ :

5'-

GCTTAATAAATATGATCACTGTGCTACGCTATATGCCGTTGAATATAGGCTATATGATC  
ATAACATATATAGCTATAAGTGATAAGTTCCTGAATATAGGCTATATGATCATAACAT  
ATACAACCTGTACTCATGAATAAGTTAA-3' → "See things not as they are, but as they  
might be", - the quote of Robert Oppenheimer ([390264](#)→[390407](#)).

4. The last watermark ( $WM_4$ ) was detected via  $TS_4$ :

5'-

GTCTCATAGTGAATACTGATATTTTAGTGCTGCCGTTGAATAAGTTCCTGAACATTGT  
GATACTGATATTTTAGTGCTGCCGTTGAATATCCTGCATTTAACTAGCTTGATAGTGCA



## 6. DECIPHERING THE WATERMARKS CODON BY CODON, MOVING TO THE 5'-END AND 3'-END



After WMs were identified, their complete sequences were decoded, moving to the 5'-end and to the 3'-end until the beginning of a senseless nucleotide context (?,?,fgjhdnf?>smkl31:":F or something like that)

# RESULTS: 4 WMs IN JCVIsyn-1.0 WERE IDENTIFIED AND DECIPHERED

## WATERMARK 1

5'-

TTAACTAGCTAAGTTCGAATATTTCTATAGCTGTACATATTGTAATGCTGATAACTAA  
TACTGTGCGCTTGACTGTGATCCTGATAAATAACTTCTTCTGTAGGGGTAGAGTTTTATT  
TAAGGCTACTCACTGGTTGCAAACCAATGCCGTACACTACTAGCTTGATCCTTGGTTCG  
GTCATTGGGGGATATCTCTTACTAATAGAGCGGCCTATCGCGTATTCTCGCCGGACCC  
CCCTCTCCCACACCAGCGGTGTAGCATCACCAAGAAAATGAGGGGAACGGATGAGG  
AACGAGTGGGGGCTCATTGCTGATCATAATGACTGTTTATATACTAATGCCGTCAACT  
GTTTGCTGTGATACTGTGCTTTCGAGGGGCGGGAGATTCGTTTTTTGACATACATAAATAT  
CATGACAAAACAGCCGGTCATGACAAAACAGCCGGTCATAATAGATTAGCCGGTGA  
CTGTGAAACTAAAGCTACTAATGCCGTCAATAAATATGATAATAGCAACGGCACTGA  
CTGTGAAACTAAAGCCGGCACTCATAATAGATTAGCCGGAGTCGTATTCATAGCCGG  
TAGATATCACTATAAGGCCCAGGATCATGATGAACACAGCACCCACGTCGTCGTCGGA  
GTTTTTTTGCTGCGACGTCTATACCACGGAAGCTGATCATAAATAGTTTTTTTGCTGCG  
GCACTAGAGCCGGACAAGCACACTACGTTTGTAATAACATCGTTCCGAATTGTAAAT  
AATTAATTCGTATTTAAATTATATGATCACTGGCTATAGTCTAGTGATAACTACAAT  
AGCTAGCAATAAGTCATATATAACAATAGCTGAACCTGTGCTACATATCCGCTATAC  
GGTAGATATCACTATAAGGCCCAGGACAATAGCTGAACCTGTGCTACATATCCGCTATAC  
TAGCTTGACTGTGGTCGGTTTTTTTGCTGCGACGTCTATACGGAAGCTCATAACTATAA  
GAGCGGCACTAGAGCCGGCACACAAGCCGGCACAGTCGTATTCATAGCCGGCACTC  
ATGACAAAACAGCGGCGCGCCTTAAGCT-3'.

◦The **largest** WM. It is localized in the region: **565506 → 566573**

◦Consisting of **1068** nucleotides

◦Encoding **356** characters

# DECIPHERED WM1

Q2>EJ. CRAIG VENTER INSTITUTE 2009

[PARAGRAPH]

ABCDEFGHIJKLMNOPQRSTUVWXYZ

[PARAGRAPH]

0123456789#@)(-+ \ = / < ; > \$ & } { \* } " [ % ! ' . ,

[PARAGRAPH]

SYNTHETIC GENOMICS, INC.

[PARAGRAPH]

```
<!DOCTYPE HTML><HTML><HEAD><TITLE>GENOME
```

```
TEAM</TITLE></HEAD><BODY><A
```

```
HREF=HTTP://WWW.JCVI.ORG/>THE JCVI</A><P>PROVE YOU'VE
```

```
DECODED THIS WATERMARK BY EMAILING US <A
```

```
HREF=MAILTO:MROOSTIZ@JCVI.ORG>HERE!</A></P></BODY></HTML
```

```
>F5+ERS
```

## ENCODED DATA

- Involved organizations and year of origin: Craig Venter institute, Synthetic Genomics, 2009
- Contains the key for deciphering (the series of used characters)
- HTML-code for emailing after decoding the watermark

# WATERMARK 2

5'-TTAACTAGCTAACAACTGGCAGCATAAAACATATAGAACTACCTGCTATAAGTGAT  
ACAACCTGTTTTTCATAGTAAAACATACAACGTTGCTGATAGTACTCCTAAGTGATAGCTTA  
GTGCGTTTAGCATATATTGTAGGCTTCATAATAAGTGATATTTTAGCTACGTAACATAAATA  
AACTAGCTATGACTGTACTCCTAAGTGATATTTTCATCCTTTGCAATACAATAACTACTAC  
ATCAATAGTGCGTGATATGCCTGTGCTAGATATAGAACACATAACTACGTTTGCTGTTTTTC  
AGTGATATGCTAGTTTTCATCTATAGATATAGGCTGCTTAGATTCCCTACTAGCTATTTCTG  
TAGGTGATATACGTCCATTGCATAAGTTAATGCATTTAACTAGCTGTGATACTATAGCAT  
CCCCATTCCCTAGTGCATATTTTCATCCTAGTGCTACGTGATATAATTGTACTAATGCCTGT  
AGATAATTTAATGCCTGGCTCGTTTGTAGGTGATAATTTAGTGCCTGTAAAACATATACCT  
GAGTGCTCGTTGCGTGATAGTTCGTTTCATGCATATACAACCTAGGCTGCTGTGATATGGTC  
ACTGCCCTTACTGTGCTACATATTACTGCGAGGGGGATGACGTATAAACCTGTTGTAAGT  
GATATGACGTATATAACTACTAGTGATATGACGTATAGGCTAGAACAACGTGATATGAC  
GTATATGACTACTGTCCCAAACATCAGTGATATGACGTATACTATAATTTCTATAATAGT  
GATAAATAAACCTGGGCTAAATACGTTCCCTGAATACGTGGCATAAACCTGGGCTAACGA  
GGAATACCCATAGTTTAGCAATAAGCTATAGTTCGTCATTTTTTAAGGCGCGCCTTAACTA  
GCT-3'

- The smallest WM. It is localized in the region:  
**725653 → 726558**
- Consisting of **909** nucleotides
- Encoding **303** characters

# DECIPHERED WM2

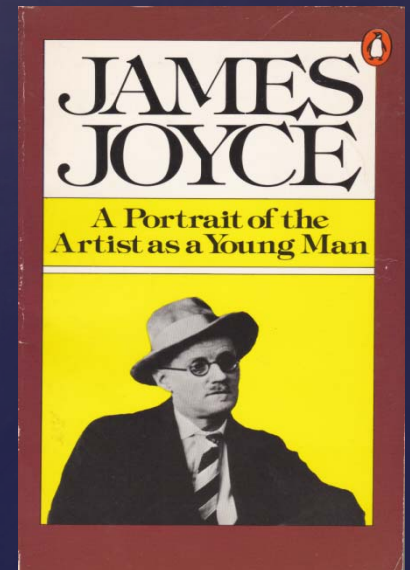
Q2>EMIKKEL ALGIRE, MICHAEL MONTAGUE, SANJAY VASHEE, CAROLE LARTIGUE, CHUCK MERRYMAN, NINA ALPEROVICH, NACYRA ASSAD-GARCIA, GWYN BENDERS, RAY-YUAN CHUANG, EVGENIA DENISOVA, DANIEL GIBSON, JOHN GLASS, ZHI-QING QI.

[PARAGRAPH]

“TO LIVE, TO ERR, TO FALL, TO TRIUMPH, TO RECREATE LIFE OUT OF LIFE.” – JAMES JOYCE

## ENCODED DATA

- Names of the scientists, participated in the project (n=13)
- The quote from the James Joyce's
- **“A Portrait of the Artist as a Young Man”**



COVER OF THE BOOK

# 3

5'-

TTAACTAGCTAATTTAACCATATTTAAATATCATCCTGATTTTCACTGGCTCGTTGCGTGA  
TATAGATTCTACTGTAGTGCTAGATAGTTCTGTACTAGGTGATACTATAGATTTTCATAGAT  
AGCACTACTGGCTTCATGCTAGGCATCCCAATAGCTAGTGATAGTTTAGTGCATACAACG  
TCATGTGATACAACGTTGCTGGCTGTAGATAACAACGTCGTATTCTGTAAGTGATACAATA  
GCTATTGCTGTGCATAGGCCTATAGTGGCTGTAAGTACTAGTGATATCACGTAACAACCATAT  
AAGTTAGATTTAATGCCCTGACTGAACGCTCGTTGCGTGATAGTTTAGGCTCGTTGCAT  
ACAAGTGTGATTTTCATAAAACAACGTGATAATTTAGTGCTAGATAAGTTCCGCTTAGCA  
AGTGATAGTTTCCGCTTGACTGTGCATAGTTCGTTTCATGCGCTCGTTGCGTGATAAACTAG  
GCAGCTTCACAAGTACTGATAATTTAATTGCTGATATTGCTGGCTGTCTAGTGCTAGTGATCAT  
AGTGCGTGATAGTTTAAGCTGCTCTGTTTTAGATATCACGTGCTTGATAATGAACTAACT  
AGTGATACTACGTAGTTAACTATGAATAGGCCTACTGTAAATTCAATAGTGCGTGATATT  
GAACTAGATTCTGCAACTGCTAATATGCCGTGCTGCACGTTTGGTGATAGTTTAGCATGC  
TTCAGTATAATAAATATGGTAGTTGTAAGTACTGCGAATAGGGGGAGCTTAATAAATATG  
ATCACTGTGCTACGCTATATGCCGTTGAATATAGGCTATATGATCATAACATATATAGCT  
ATAAGTGATAAGTTCCTGAATATAGGCTATATGATCATAACATATACAAGTACTCATG  
AATAAGTTAACGAGGA-3'

- Localized in the region: **389493 → 390413**
- Consisting of **921** nucleotides
- Encoding **307** characters

# DECIPHERED WM3

Q2>ECLYDE HUTCHISON, ADRIANE JIGA, RADHA KRISHNAKUMAR, JAN MOY, MONZIA MOODIE, MARVIN FRAZIER, HOLLY BADEN-TILSON, JASON MITCHELL, DANA BUSAM, JUSTIN JOHNSON, LAKSHMI DEVI VISWANATHAN, JESSICA HOSTETLER, ROBERT FRIEDMAN, VLADIMIR NOSKOV, JAYSHREE ZAVERI.

[PARAGRAPH]

“SEE THINGS NOT AS THEY ARE, BUT AS THEY MIGHT BE.”

## ENCODED DATA

- Names of the scientists, participated in the project (n=15)
- Robert Oppenheimer's quote

# WATERMARK 4

5'-

TTAACTAGCTAATTTTCATTGCTGATCACTGTAGATATAGTGCATTCTATAAGTCGCTCC  
CACAGGCTAGTGCTGCGCACGTTTTTCAGTGATATTATCCTAGTGCTACATAACATCAT  
AGTGCGTGATAAACCTGATACAATAGGTGATATCATAGCAACTGAACTGACGTTGCAT  
AGCTCAACTGTGATCAGTGATATAGATTCTGATACTATAGCAACGTTGCGTGATATTTT  
CACTACTGGCTTGACTGTAGTGCATATGATAGTACGTCTAACTAGCATAACTAGTGAT  
AGTTATATTTCTATAGCTGTACATATTGTAATGCTGATAACTAGTGATATAATCCAAC  
AGATAGTCCTGAACTGATCCCTATGCTAACTAGTGATAAACTAACTGATACATCGTTC  
CTGCTACGTGATAGCTTCACTGAGTTCATACATCGTCGTGCTTAAACATCAGTGATAA  
CACTATAGAGTTCATAGATACTGCATTAAGTGTGATATGACTGCAAATAGCTTGACG  
TTTTGCAGTCTAAACAACGTGATAATTCTGTAGTGCTAGATACTATAGATTTCCCTGCT  
AAGTGATAAGTCTACTGATTTACTAATGAATAGCTTGGTTTTGGCATACACTGTGCGCT  
GCACTGGTGATAGCTTTTCGTTGATGAATAATTTCCCTAGCACTGTGCGTGATATGCTA  
GATTCTGTAGATAGGCTAAATTCGTCTACGTTTGTAGGTGATAGTTTAGTTGCTGTAAC  
TAATATTATCCCTGTGCCGTTGCTAAGCTGTGATATCATAGTGCTGCTAGATATGATAA  
GCAAATAATAGAGTCGAGGGGGAGTCTCATAGTGAATACTGATATTTTAGTGCTGCC  
GTTGAATAAGTTCCTGAACATTGTGATACTGATATTTTAGTGCTGCCGTTGAATATCC  
TGCATTTAACTAGCTTGATAGTGCATTCGAGGAATACCCATACTACTGTTTTTCATAGCT  
AATTATAGGCTAACATTGCCAATAGTGCGGCGCGCCTTAACTAGCT-3'

- oLocalized in the region: **958641** → **959684**
- oConsisting of **1044** nucleotides
- oEncoding **348** characters



# DECIPHERED WM4

Q2>ECYNTHIA ANDREWS-PFANNKOCH, QUANG PHAN, LI MA, HAMILTON SMITH, ADI RAMON, CHRISTIAN TAGWERKER, J CRAIG VENTER, EULA WILTURNER, LEI YOUNG, SHIBU YOOSEPH, PRABHA IYER, TIM STOCKWELL, DIANA RADUNE, BRIDGET SZCZYPINSKI, SCOTT DURKIN, NADIA FEDOROVA, JAVIER QUINONES, HANNA TEKLEAB.

[PARAGRAPH]

“WHAT I CANNOT BUILD, I CANNOT UNDERSTAND.” – RICHARD FEYNMANF5+ERS

## ENCODED DATA

- Names of the scientists, participated in the project (n=18)
- Richard Feynman's quote

# CONCLUSION

THIS STUDY REVEALED:

- EXACT LOCI OF WATERMARKS (WM1: 565506 → 566573; WM2: 725653 → 726558; WM3: 389493 → 390413; WM4: 958641 → 959684).

As far as we know, this data is absolutely exclusive and it has been never published before.

- COMPLETE NUCLEOTIDE SEQUENCE OF THE WMs

WATERMARKS WERE FULLY DECODED. THEY CONTAIN:

- THE **KEY** FOR DECIPHERING AND HTML CODE FOR EMAILING
- **46** NAMES OF SCIENTIST, INVOLVED IN THE PROJECT
- **4** QUOTATIONS
- TOTAL NUMBER OF NUCLEOTIDES, COMPOSING 4 WATERMARKS = **3942**
- THE LARGEST IS WM1 (**1069** NTs), THE SMALLEST IS WM2 (**909** NTs)



ON THE 90<sup>TH</sup> ANNIVERSARY  
OF THE DEPARTMENT OF  
FOREIGN LANGUAGES  
**THANK YOU FOR  
ATTENTION!**

ON THE 10<sup>TH</sup>  
ANNIVERSARY OF  
SYNTHETIC MYCOPLASMA



К 90-ЛЕТИЮ КАФЕДРЫ ИНОСТРАННЫХ ЯЗЫКОВ. ON THE 90<sup>TH</sup> ANNIVERSARY  
OF THE DEPARTMENT OF FOREIGN LANGUAGES. ZUM 90-JÄHRIGE JUBILÄUM  
DER FAKULTÄT FÜR FREMDSPRACHEN. DEBIDO AL 90 ANIVERSARIO DEL  
DEPARTAMENTO DE LENGUAS EXTRANJERAS. SUR LE 90E ANNIVERSAIRE DE LA  
FACULTÉ DES LANGUES ÉTRANGÈRES.