Synthetic Genomics and Its Application to Viral Infectious Diseases

Timothy Stockwell (JCVI)

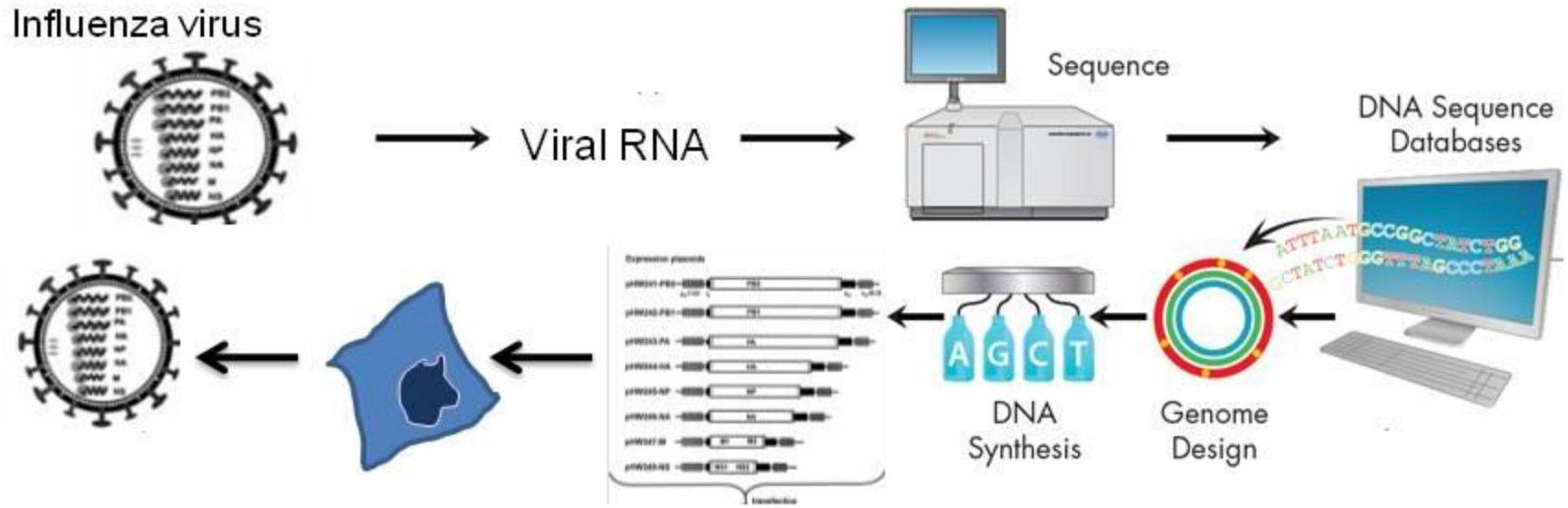
David Wentworth (JCVI)

J. Craig Venter™ INSTITUTE

Outline

- Using informatics to predict drift (strain selection)
- Synthetic Genomics: Preparedness (NIH/NIAID)
- Rapid Response to emerging viruses

  (BARDA/Novartis/SGVI)



J. Craig Venter INSTITUTE

Influenza vaccines could be great if:

- Improve strain prediction
  - -> Bioinformatics
- Existing vaccine candidates on the shelf
  - Synfluenza
- Speed production
  - Rapid response

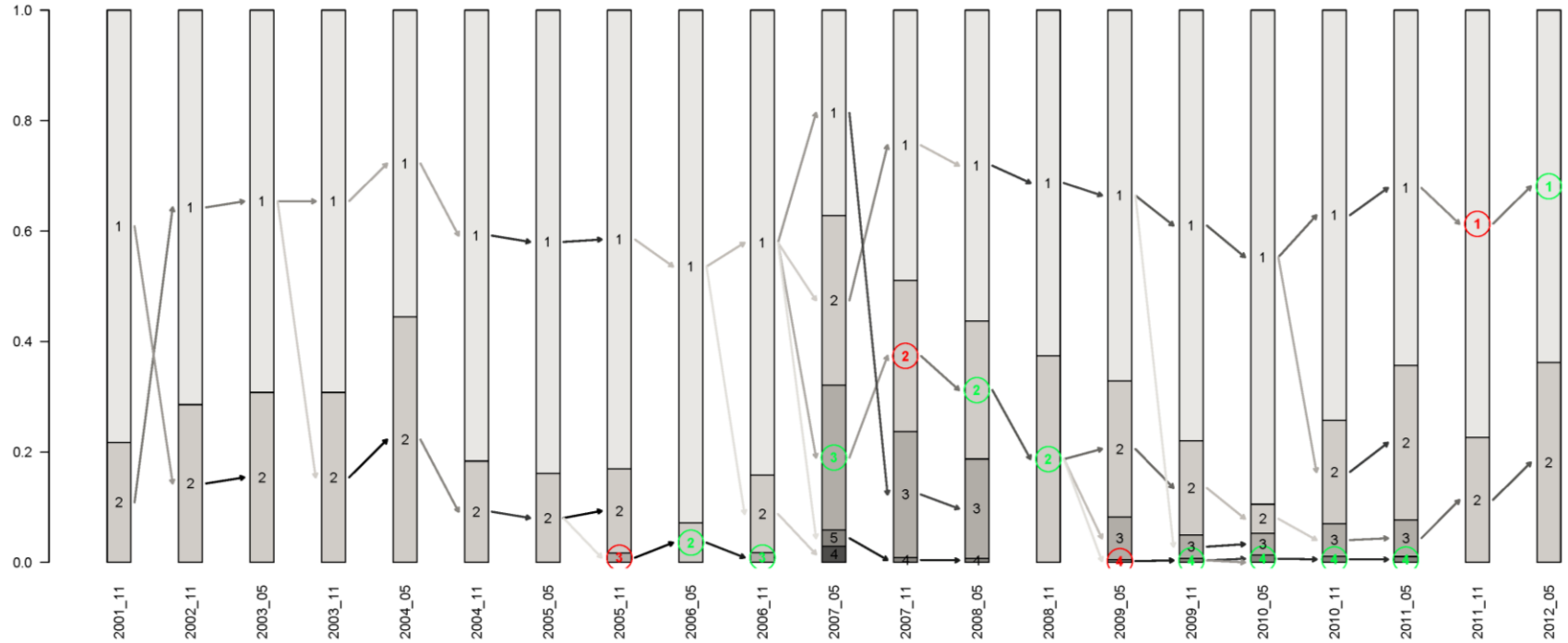
J. Craig Venter INSTITUTE

Vaccine Bioinformatics

- What should go into a vaccine?
  - Track the viral evolution
  - Determine/predict vaccine candidates protection
  - Combine the information

J. Craig Venter INSTITUTE

Tracking the flu over time



Analysis of FluB in September of 2012, showing the growth of a new group of viruses (2) that were different from the vaccine strain group (1).

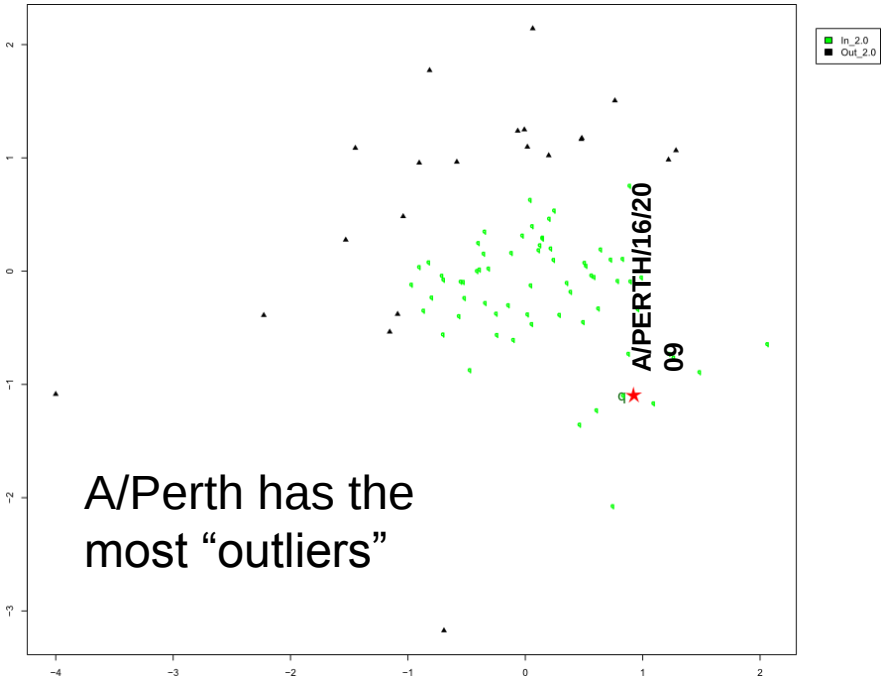
J. Craig Venter™ INSTITUTE

Vaccine Informatics

- Based on what types of flu viruses are circulating, and how the viruses are evolving, what should go into a vaccine?
- Track the flu
- Determine how well vaccine candidates might protect against the viruses in circulation
- Combine the information

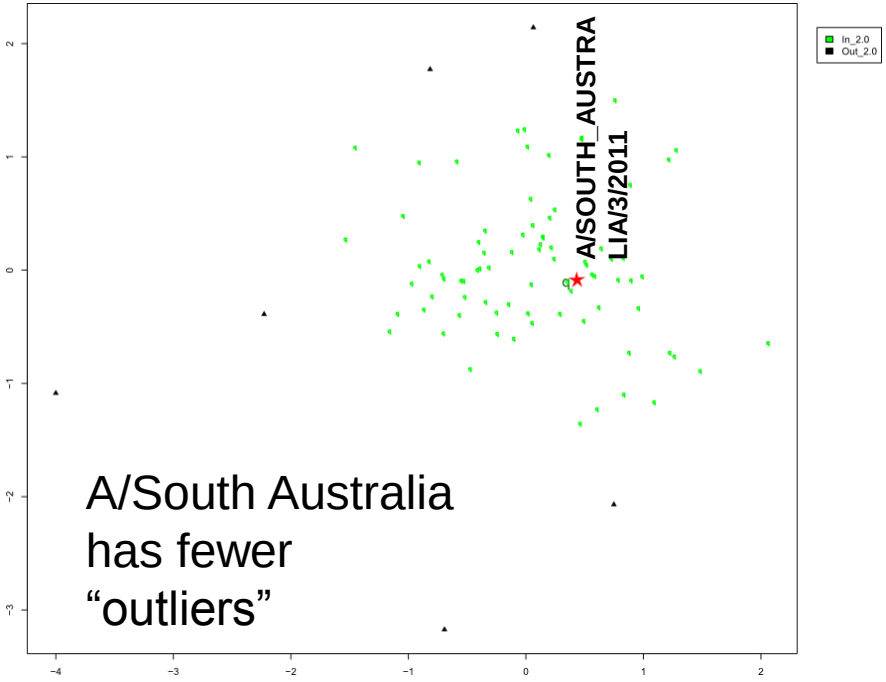
Predicting how well vaccines should work against other viruses

Antigenic distancing – based on "distances", try to "plot" everything



A/Perth has the most "outliers"

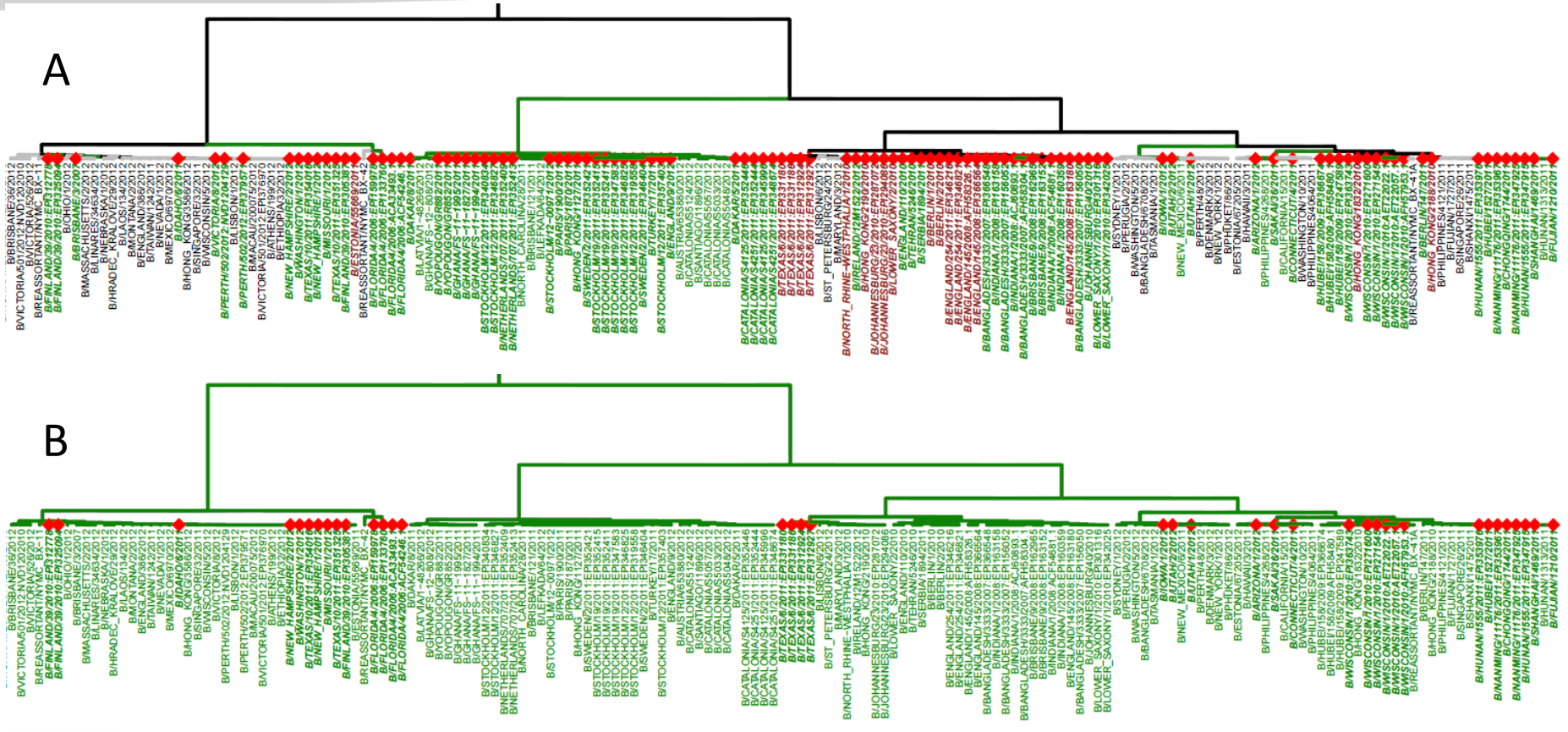
Then see if you can find a candidate virus that "covers" the most.



A/South Australia has fewer "outliers"

J. Craig Venter™ INSTITUTE

Putting it all together



Analysis of two vaccine candidates, the current vaccine at the time (A) and a alternative candidate (B) selected by algorithm as providing better predicted protection against circulating strains.

J. Craig Venter™ INSTITUTE

Synthetic Genomics Tools

RESEARCH ARTICLE

Complete Chemical Synthesis, Assembly, and Cloning of a *Mycoplasma genitalium* Genome

Daniel G. Gibson, Gwynedd A. Bend
Holly Baden-Tillson, Jayshree Zaveri, T
Mikkel A. Algire, Chuck Merryman, Lei
Clyde A. Hutchison III, Hamilton O. Sm

genome, we needed to establish convenient and reliable methods for the assembly and cloning of much larger synthetic DNA molecules.

Strategy for synthesis and assembly. The native 580,076-bp *M. genitalium* genome sequence (*Mycoplasma genitalium* G37 ATCC 33530 genomic sequence; accession no. L43967) (*3*) was partitioned into 101 cassettes of approximately 5 to 7 kb in length (Fig. 1) that were individually synthesized, verified by sequencing,

PNAS

One-step assembly in yeast of 25 overlapping DNA fragments to form a complete synthetic *Mycoplasma genitalium* genome

Daniel G. Gibson[a,1], Gwynedd A. Benders[b], Kevin C. Axelrod[a], Jayshree Zaveri[a], Mikkel A. Algire[a], Monzia Moodie[a], Michael G. Montague[a], J. Craig Venter[a], Hamilton O. Smith[b], and Clyde A. Hutchison III[b,1]

[a]The J. Craig Venter Institute, Synthetic Biology Group, Rockville, MD 20850 and [b]The J. Craig Venter Institute, Synthetic Biology Group, San Diego, CA 92121

Contributed by Clyde A. Hut

We previously reported *Mycoplasma genitalium* *myces cerevisiae* by reco ments to produce a 592-k demonstrating assembly lapping fragments in a sin greatly simplifies the ass synthetic and natural fra

Enzymatic assembly of DNA molecules up to several hundred kilobases

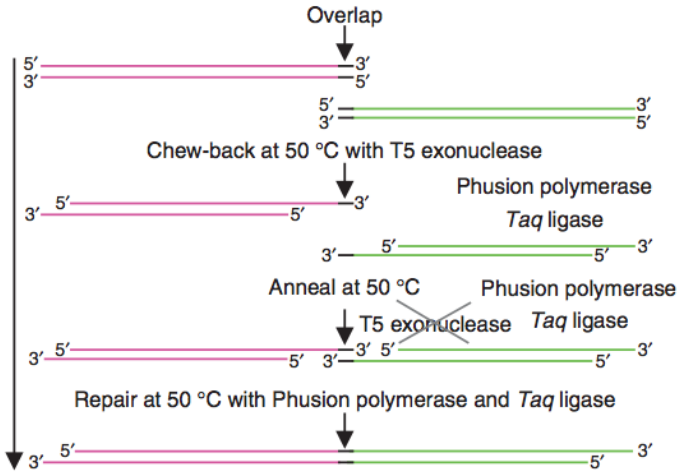
Daniel G Gibson[1], Lei Young[1], Ray-Yuan Chuang[1], J Craig Venter[1,2], Clyde A Hutchison III[2] & Hamilton O Smith[2]

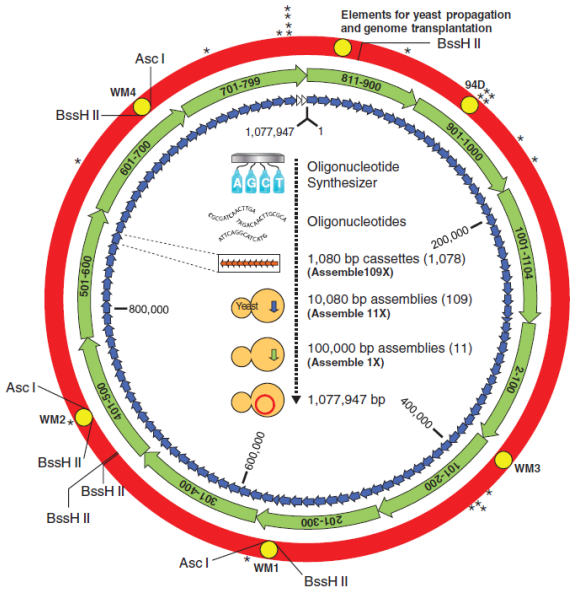
n method for

overlapping DNA molecules and then incubated at 50 °C for as few as 15 min (Online Methods). This approach dramatically simplifies the construction of large DNA molecules from constituent parts.

Exonucleases that recess double-stranded DNA from 5′ ends will not compete with polymerase activity. Thus, all enzymes required for DNA assembly can be simultaneously active in a single isothermal reaction. Furthermore, circular products can be enriched as they are not processed by any of the three enzymes in the reaction. We optimized a 50 °C isothermal assembly system using the activities of the 5′ T5 exonuclease (Epicentre), Phusion DNA polymerase (New England Biolabs (NEB)) and *Taq* DNA ligase (NEB) (**Fig. 1**). *Taq* DNA polymerase (NEB) can be used in place of Phusion DNA polymerase (data not shown), but the latter

Gibson Assembly





Synthesis of DNA fragments in yeast by one-step assembly of overlapping oligonucleotides

Daniel G. Gibson*

The J. Craig Venter Institute, Synthetic Biology Group, 9704 Medical Center Drive, Rockville, MD 20850, USA

Received January 1, 2009; Revised August 1, 2009; Accepted August 4, 2009

ABSTRACT

Here it is demonstrated th
myces cerevisiae can take
least 38 overlapping singl
otides and a linear double-
transformation event. Thes
overlap by as few as 20 b
as 200 nucleotides in lengt
scheme for assembling
oligonucleotides could be a
synthetic DNA molecules.

RESEARCH ARTICLE

Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome

Daniel G. Gibson,[1] John I. Glass,[1] Carole Lartigue,[1] Vladimir N. Noskov,[1] Ray-Yuan Chuang,[1] Mikkel A. Algire,[1] Gwynedd A. Benders,[2] Michael G. Montague,[1] Li Ma,[1] Monzia M. Moodie,[1] Chuck Merryman,[1] Sanjay Vashee,[1] Radha Krishnakumar,[1] Nacyra Assad-Garcia,[1] Cynthia Andrews-Pfannkoch,[1] Evgeniya A. Denisova,[1] Lei Young,[1] Zhi-Qing Qi,[1] Thomas H. Segall-Shapiro,[1] Christopher H. Calvey,[1] Prashanth P. Parmar,[1] Clyde A. Hutchison III,[2] Hamilton O. Smith,[2] J. Craig Venter[1,2]*

We report the design, synthesis, and assembly of the 1.08–mega–base pair *Mycoplasma mycoides* JCVI-syn1.0 genome starting from digitized genome sequence information and its transplantation into a *M. capricolum* recipient cell to create new *M. mycoides* cells that are controlled only by the synthetic chromosome. The only DNA in the cells is the designed synthetic DNA sequence, including "watermark" sequences and other designed gene deletions and polymorphisms, and mutations acquired during the building process. The new cells have expected phenotypic properties and are capable of continuous self-replication.

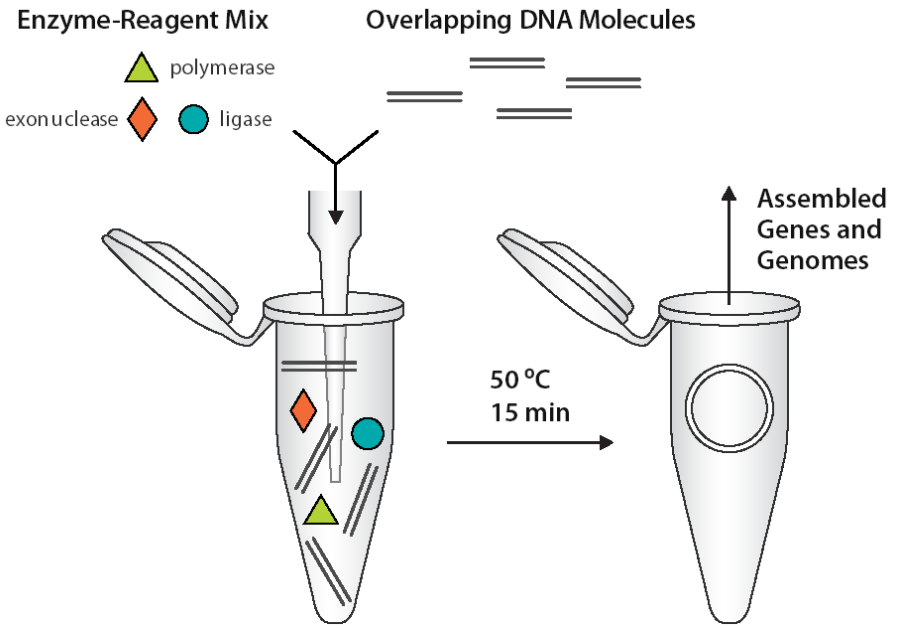
crude *M. mycoides* or *M. capricolum* extracts, or by simply disrupting the recipient cell's restriction system (*8*).

We now have combined all of our previously established procedures and report the synthesis, assembly, cloning, and successful transplantation of the 1.08-Mbp *M. mycoides* JCVI-syn1.0 genome, to create a new cell controlled by this synthetic genome.

Synthetic genome design. Design of the *M. mycoides* JCVI-syn1.0 genome was based on the highly accurate finished genome sequences of two laboratory strains of *M. mycoides* subspecies *capri* GM12 (*8*, *9*, *11*). One was the genome donor used by Lartigue *et al*. [GenBank accession CP001621] (*10*). The other was a strain created by transplantation of a genome that had been cloned and engineered in yeast, YCpMmyc1.1-Δ*typeIIres* [GenBank accession CP001668] (*8*). This project was critically dependent on the accuracy of these sequences. Although we believe that both finished *M. mycoides* genome sequences are reliable, there are 95 sites at which they differ. We

J. Craig Venter™
INSTITUTE

Rapid *in vitro* recombination of ssDNA and dsDNA



Enzymatic assembly of DNA molecules up to several hundred kilobases

Daniel G Gibson[1], Lei Young[1], Ray-Yuan Chuang[1], J Craig Venter[1,2], Clyde A Hutchison III[2] & Hamilton O Smith[2]

overlapping DNA molecules and then incubated at 50 °C for as few as 15 min (Online Methods). This approach dramatically simplifies the construction of large DNA molecules from constituent parts.

Exonucleases that recess double-stranded DNA from 5′ ends will not compete with polymerase activity. Thus, all enzymes required for DNA assembly can be simultaneously active in a single isothermal reaction. Furthermore, circular products can be enriched as they are not processed by any of the three enzymes in the reaction. We optimized a 50 °C isothermal assembly system using the activities of the 5′ T5 exonuclease (Epicentre), Phusion DNA polymerase (New England Biolabs (NEB)) and *Taq* DNA

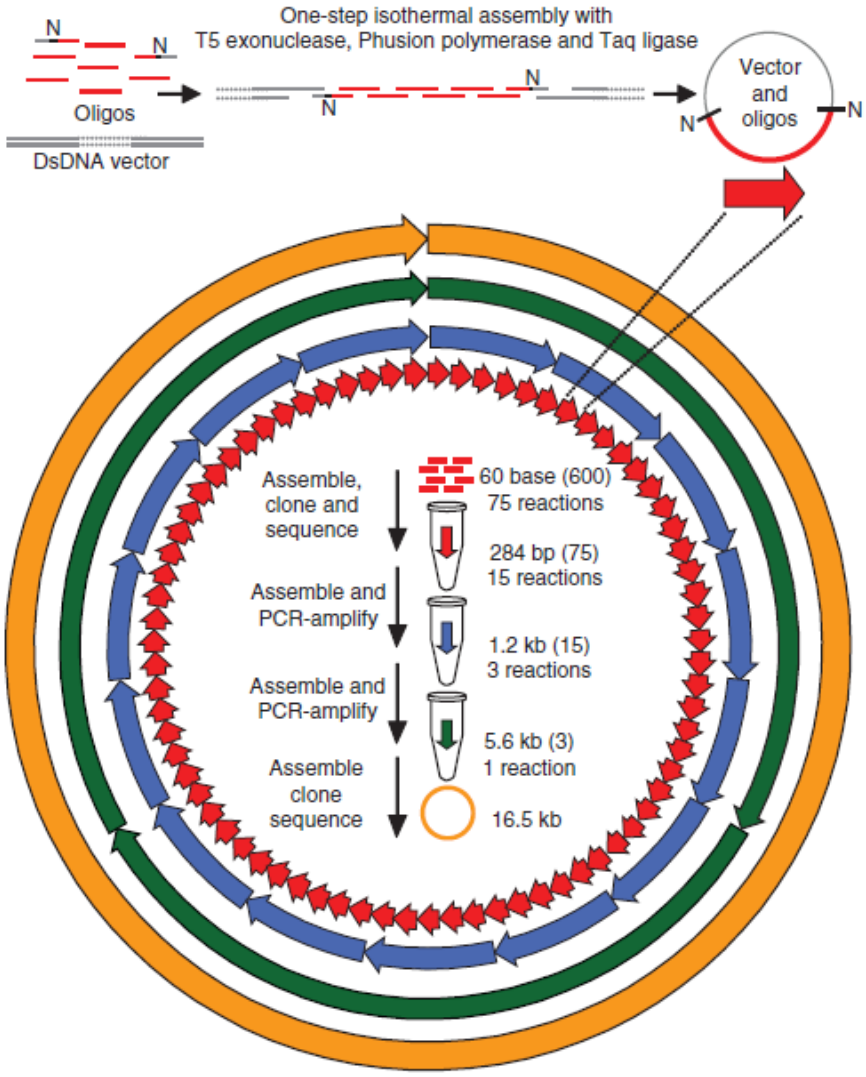
Chemical synthesis of the mouse mitochondrial genome

Daniel G Gibson[1], Hamilton O Smith[2], Clyde A Hutchison III[2], J Craig Venter[1,2] & Chuck Merryman[1]

We describe a one-step, isothermal assembly method for synthesizing DNA molecules from overlapping oligonucleotides. The method cycles between *in vitro* recombination and amplification until the desired length is reached. As a

and estimate that one individual could reconstruct the entire 16.3-kb molecule in just 5 d (**Supplementary Fig. 1**).

We recently described a one-step, isothermal *in vitro* recombination system capable of joining overlapping double-stranded DNA molecules up to hundreds of kilobases long[7]. The assembly reaction mixture in this system contains three separate enzymes (T5 exonuclease, Phusion polymerase and *Taq* ligase) that work in harmony to join multiple DNA fragments. In a typical reaction the assembly is accomplished in as few as 15 min. This method is robust and amenable to automation. For these reasons, we adapted it for assembly beginning at the oligo level. We optimized several parameters including the number of oligos used in a single reaction, their length, the amount of overlap, orientation, oligo concentration in the reaction, reaction temperature and reaction time (**Supplementary Tables 1–9** and



J. Craig Venter™ INSTITUTE

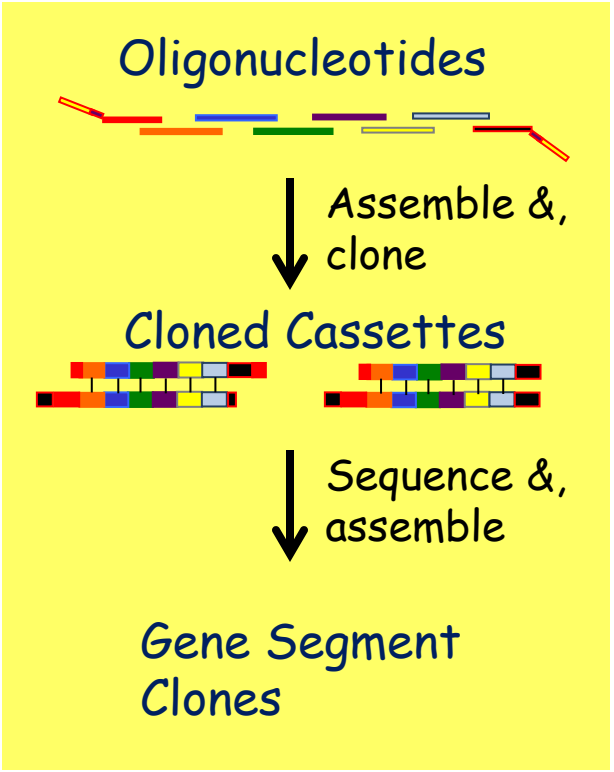
Synfluenza Project Details

NIAID project to create ~1000 HA's and NA's

- 12 host subtype combinations
- Span sequence diversity (past 5 years)
  - Human – H1N1pdm, H1N1, H3N2, Influenza B
  - Avian – H5N1, H7N3, H7N7, H9N2
  - Swine – H1N1, H1N2, H3N1, H3N2

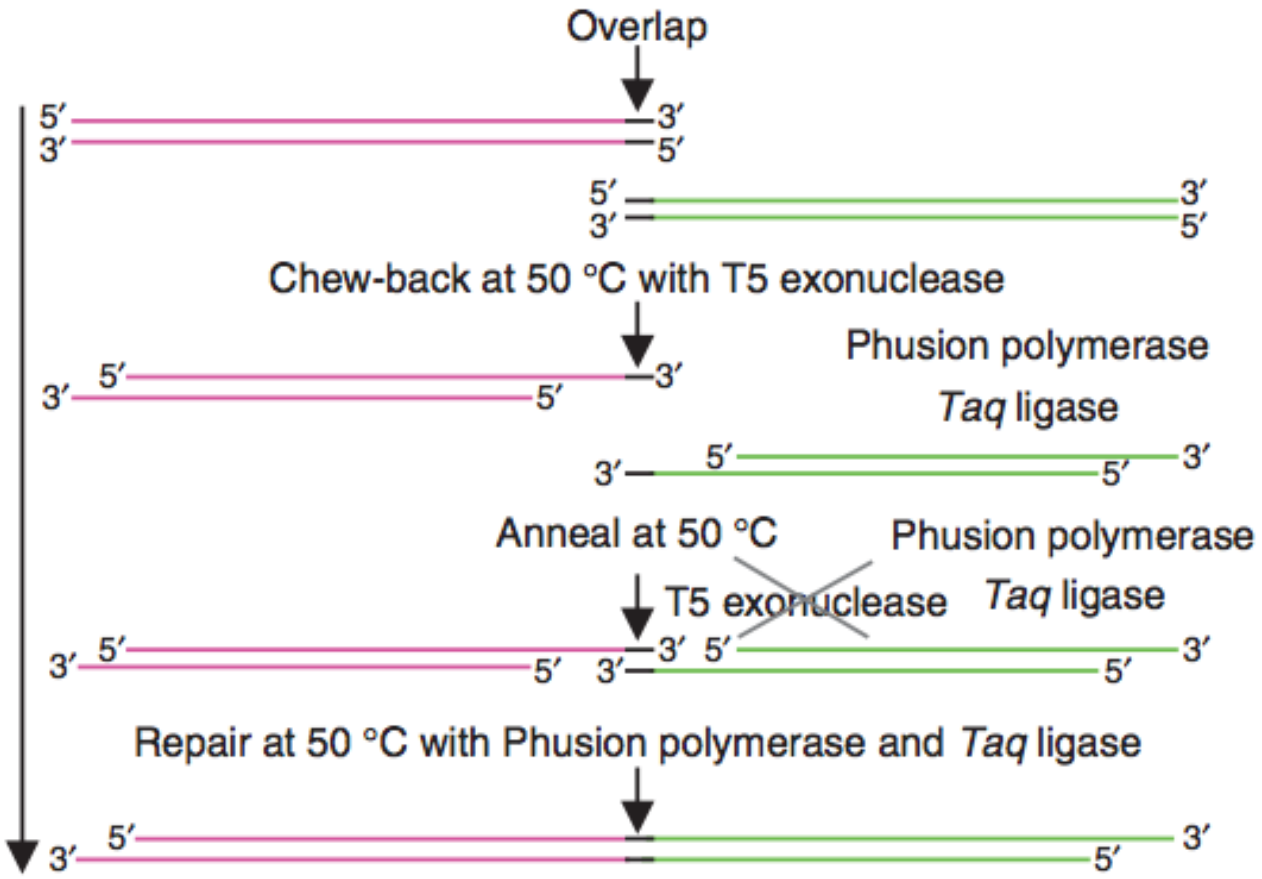
Algorithms to maximize reuse of oligos/cassettes and minimize costs

- Each molecule made from 7 (HA) or 5 (NA) cassettes (~350bp)
  - Each cassette is made from 8 oligos (~65 bp)
- Designs based on GenBank sequences with consensus UTRs



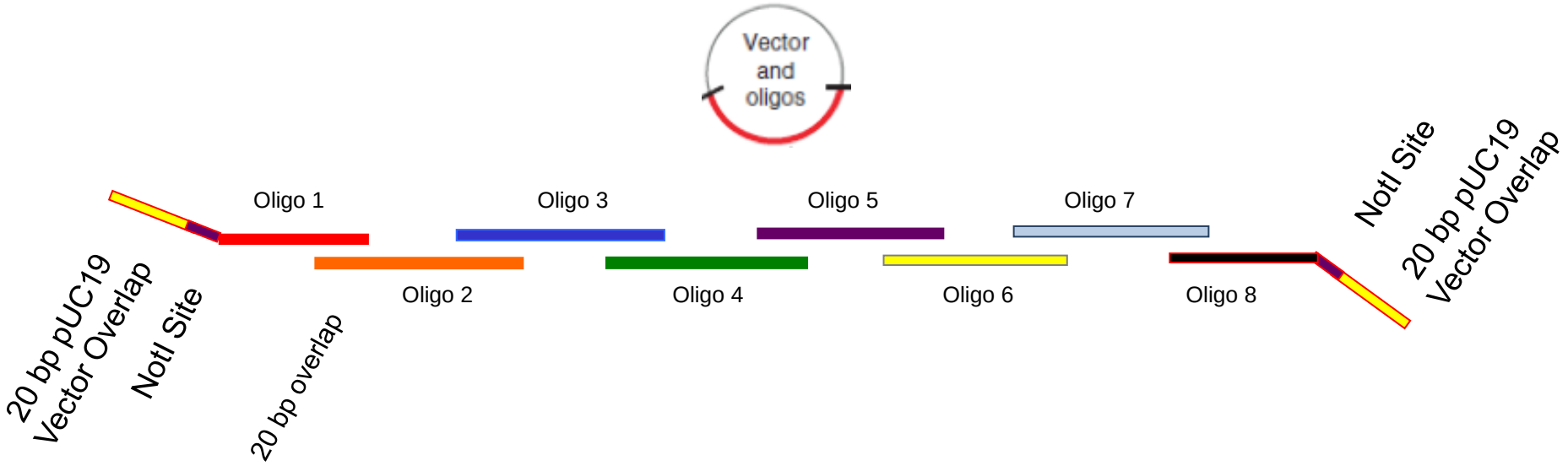
J. Craig Venter INSTITUTE

Gibson Assembly



J. Craig Venter™ INSTITUTE

Assembling a Cassette



Oligos are 59-72 bp

Cassettes are 335-401 bp

J. Craig Venter INSTITUTE

Assembling a Flu Molecule

HA (7 Cassettes)
1716-1885 bp (ungapped)

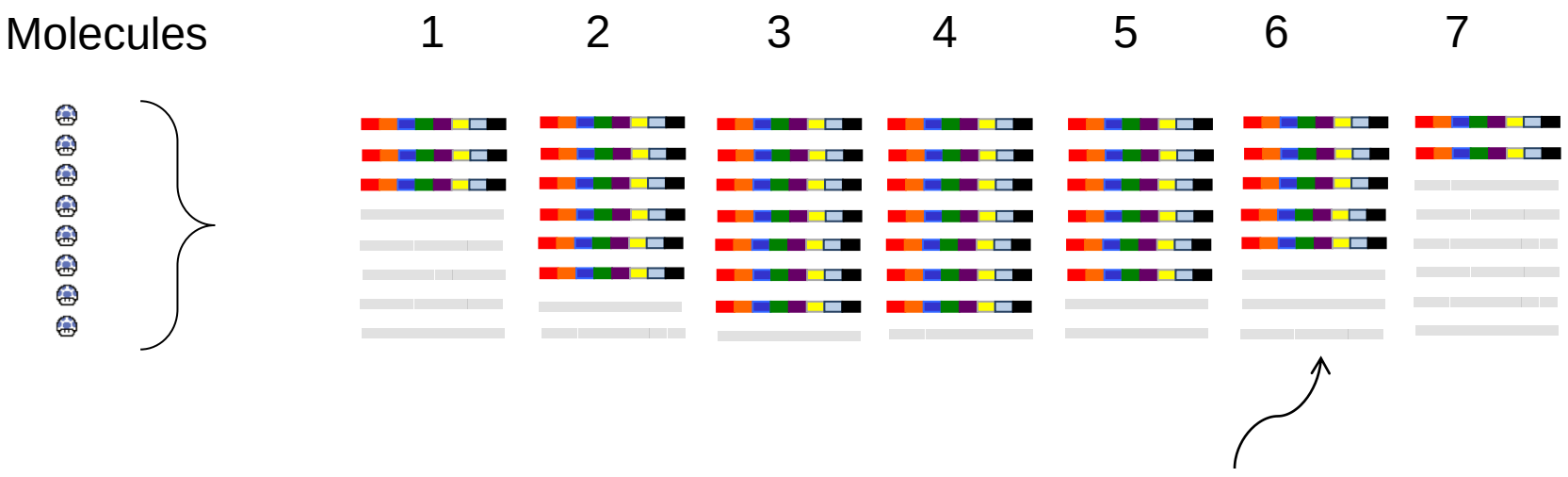
NA (5 Cassettes)
1374-1560 bp (ungapped)

40 bp overlap

J. Craig Venter INSTITUTE

Cassette Design

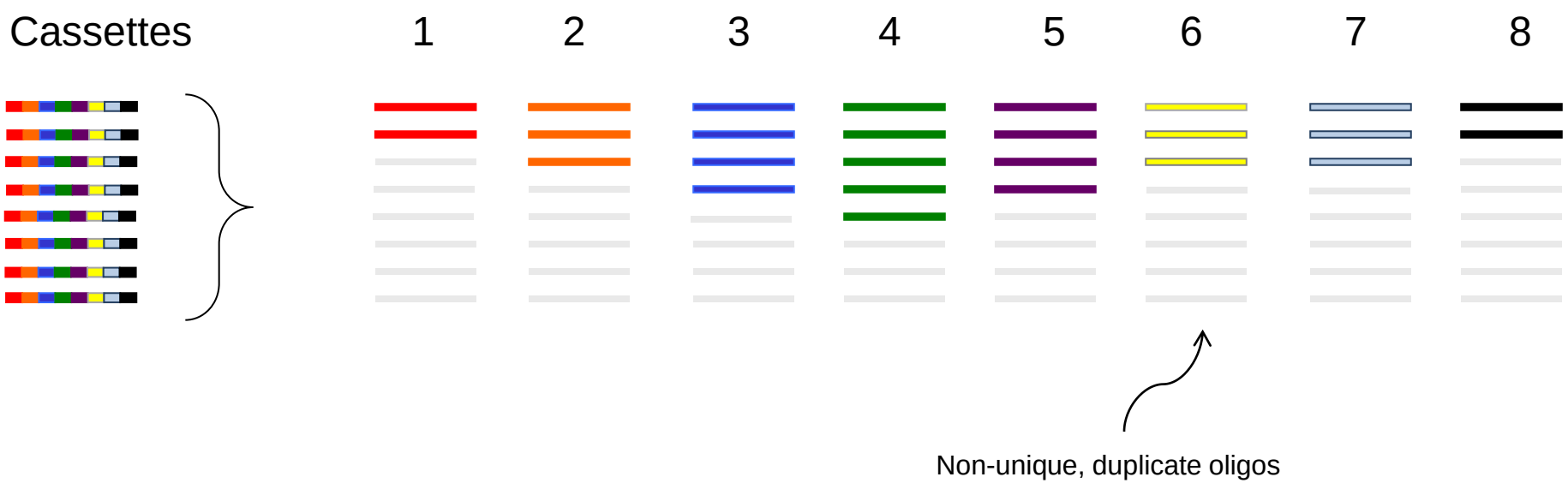
Only 1 copy of each unique cassette is made for each Host, Subtype, Segment & Position (e.g. Avian H5N1 HA)



J. Craig Venter INSTITUTE

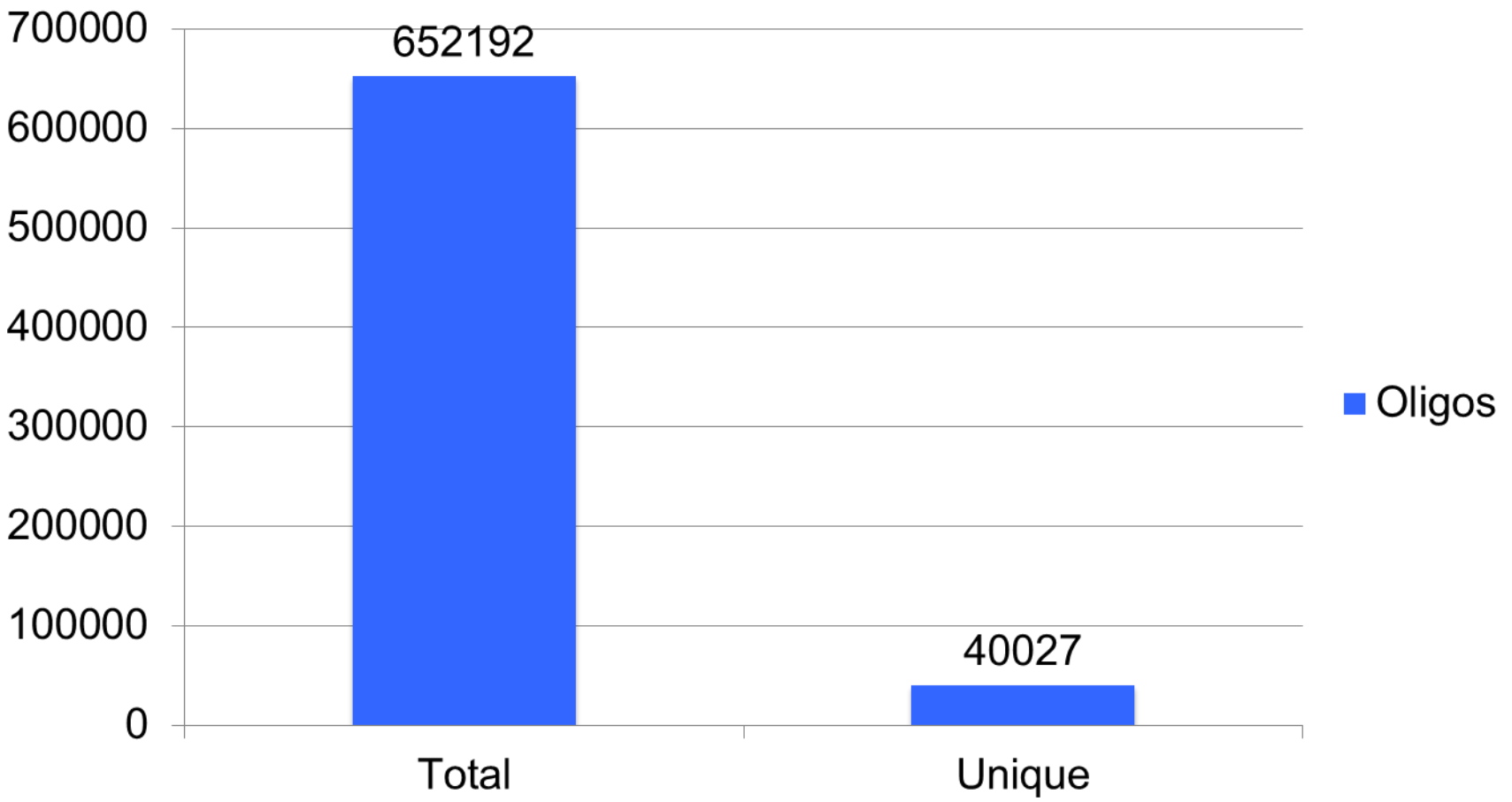
Oligo Design

Only 1 copy of each unique oligo is made for each Host, Subtype, Segment, Cassette, & Position



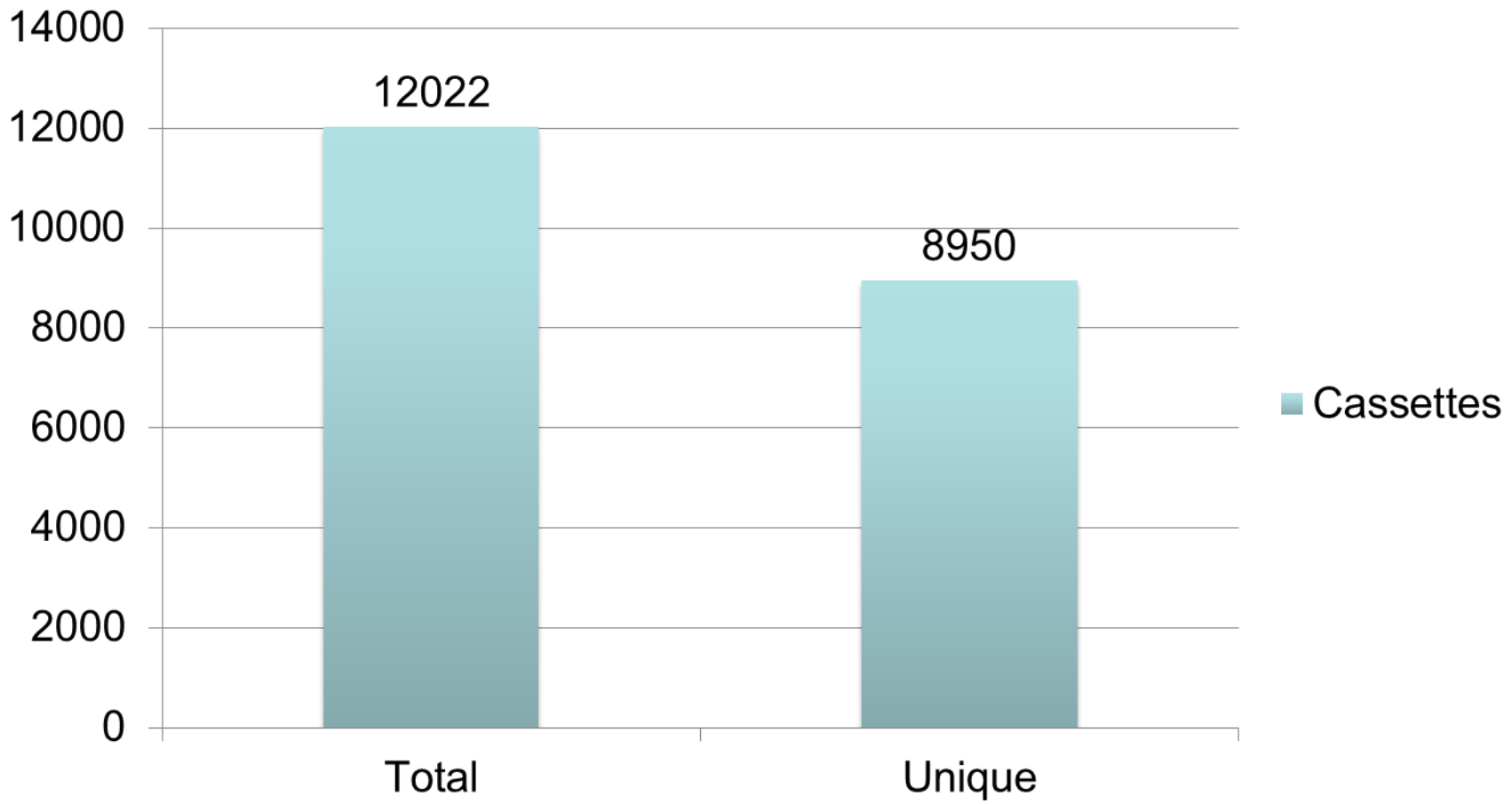
J. Craig Venter INSTITUTE

Oligo Savings



J. Craig Venter INSTITUTE

Cassette Savings (Initial 1000 HA & NA)



J. Craig Venter™ INSTITUTE

HA's and NA's Constructed Via Automated DNA Synthesis and Assembly

Designed Sequence

HA, NA Genes

Order/Synthesize Oligonucleotides

Assembly reaction

Biomek FX

Hamilton µStar

µFill

Cloning

E. coli transformation

Colony picking

QPix

µFill

Template production

Biomek FX

µFill

Sequencing reaction

Biomek FX

ABI9700 Thermal Cycler

Sequencing

ABI3730

Select clones

Biomek FX

Iterative assembly and amplification

µFill

~13 kb per 384-well oligo plate

J. Craig Venter INSTITUTE

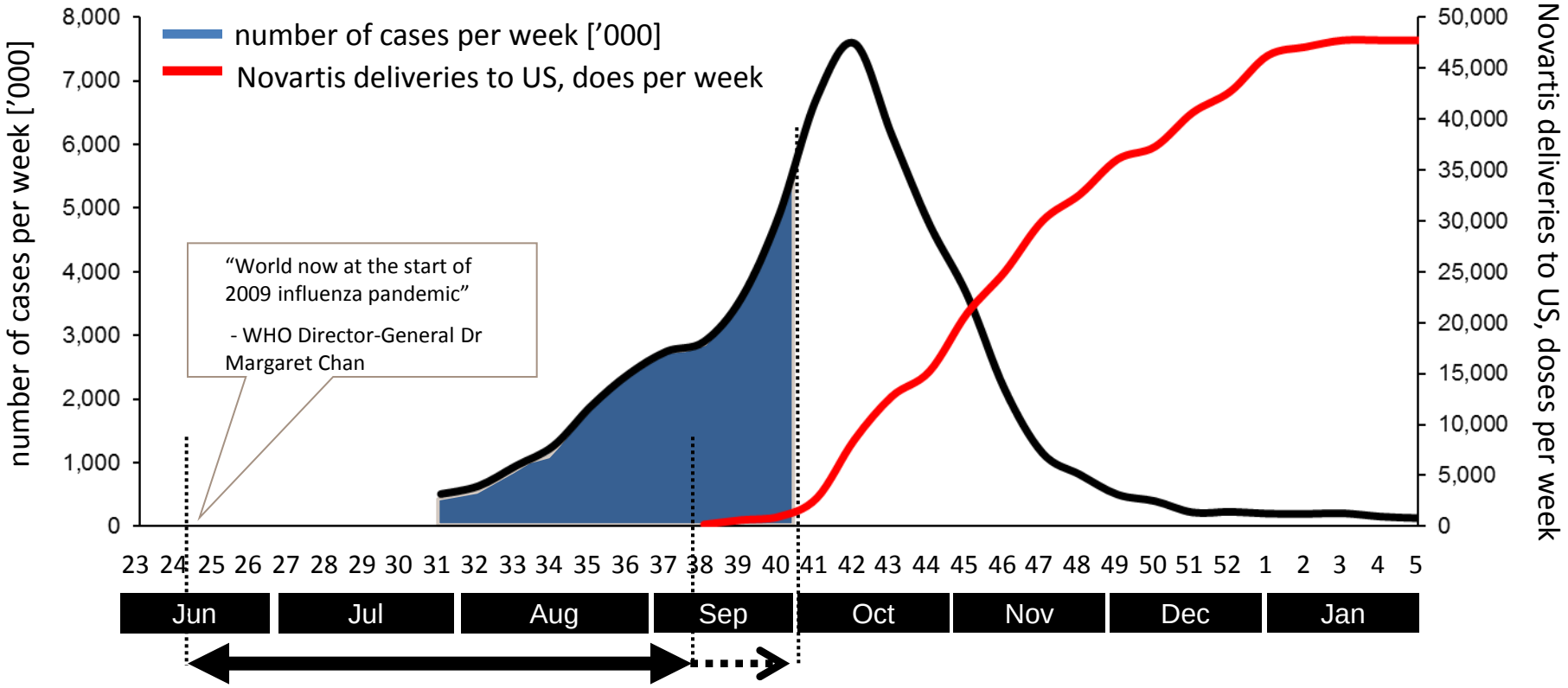
Project Breakdown

Host	Subtype	Segment	Molecules	Unique Cassettes	Unique Oligos	Intial 1000 Molecules	Intial 1000 Unique Cassettes	Intial 1000 Unique Oligos
AVIAN	H5N1	HA	992	2982	5913	289	1629	4318
AVIAN	H5N1	NA	874	1848	3729	322	1287	3111
AVIAN	H7N3	HA	84	232	815	16	108	586
AVIAN	H7N3	NA	36	101	408	11	53	286
AVIAN	H7N7	HA	28	128	564	14	95	492
AVIAN	H7N7	NA	31	103	478	12	60	349
AVIAN	H9N2	HA	273	1167	3822	148	906	3427
AVIAN	H9N2	NA	160	568	2446	101	470	2297
HUMAN	FLUB	HA	363	659	1158	13	85	348
HUMAN	FLUB	NA	487	602	1030	64	240	567
HUMAN	H1N1	HA	829	1528	2220	92	441	947
HUMAN	H1N1	NA	849	1065	1546	63	238	549
HUMAN	H1N1PDM	HA	3103	2149	2636	171	519	977
HUMAN	H1N1PDM	NA	2860	1259	1557	121	297	514
HUMAN	H3N2	HA	1058	1660	2322	142	609	1181
HUMAN	H3N2	NA	1050	1330	1762	187	576	1043
PORCINE	H1N1	HA	88	378	1685	42	282	1493
PORCINE	H1N1	NA	81	255	1082	40	180	929
PORCINE	H1N2	HA	67	290	1452	36	241	1380
PORCINE	H1N2	NA	72	226	1071	37	181	1009
PORCINE	H3N1	HA	3	14	111	2	14	111
PORCINE	H3N1	NA	2	10	80	2	10	80
PORCINE	H3N2	HA	69	319	1233	41	260	1139
PORCINE	H3N2	NA	63	216	907	36	169	796

J. Craig Venter™ INSTITUTE

Influenza Vaccine: The Need for Faster Vaccine Development

The 2009 H1N1 pandemic confirmed everyone's fears –close to 40% of cases occurred in a time when no meaningful vaccine quantities were available



Data provided courtesy of Phil Dormitzer at Novartis Vaccines & Diagnostics. Source: source is: http://www.cdc.gov/h1n1flu/estimates_2009_h1n1.htm and http://www.cdc.gov/flu/weekly/index.htm; As of Jan16, 2010 the CDC estimated that about 57 million people are infected with 2009 H1N1. weekly data on influenza positive tests reported to CDC by U.S. WHO/ NREVSS collaborating laboratories applied to CDC estimate to arrive at the weekly estimate for number of cases in the US.

J. Craig Venter INSTITUTE

Speeding Vaccine Seeds

A BARDA-funded collaboration between Novartis,* Synthetic Genomics Vaccines Inc. *(SGVI)/J. Craig Venter Institute (JCVI)

- Rapidly synthesize flu gene segments (HA and NA)
- Rescue recombinant viruses with optimized flu backbone

Milestone 1 (Sept. 2011): Demonstrate virus rescue within 7 days of receiving HA and NA sequence information

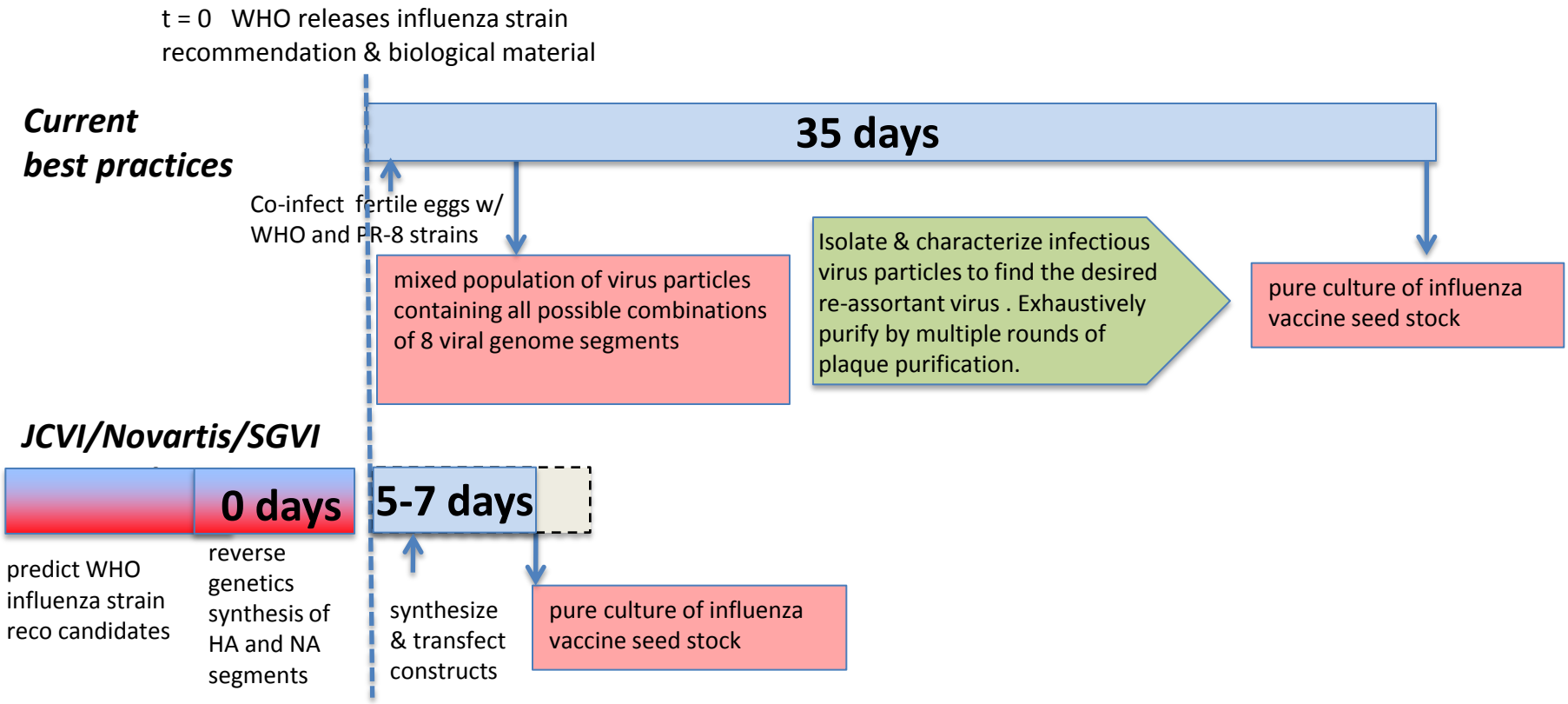
Status – Milestone surpassed

We were able to confirm rescue of an H7N9 virus within 5 days of initiating the process

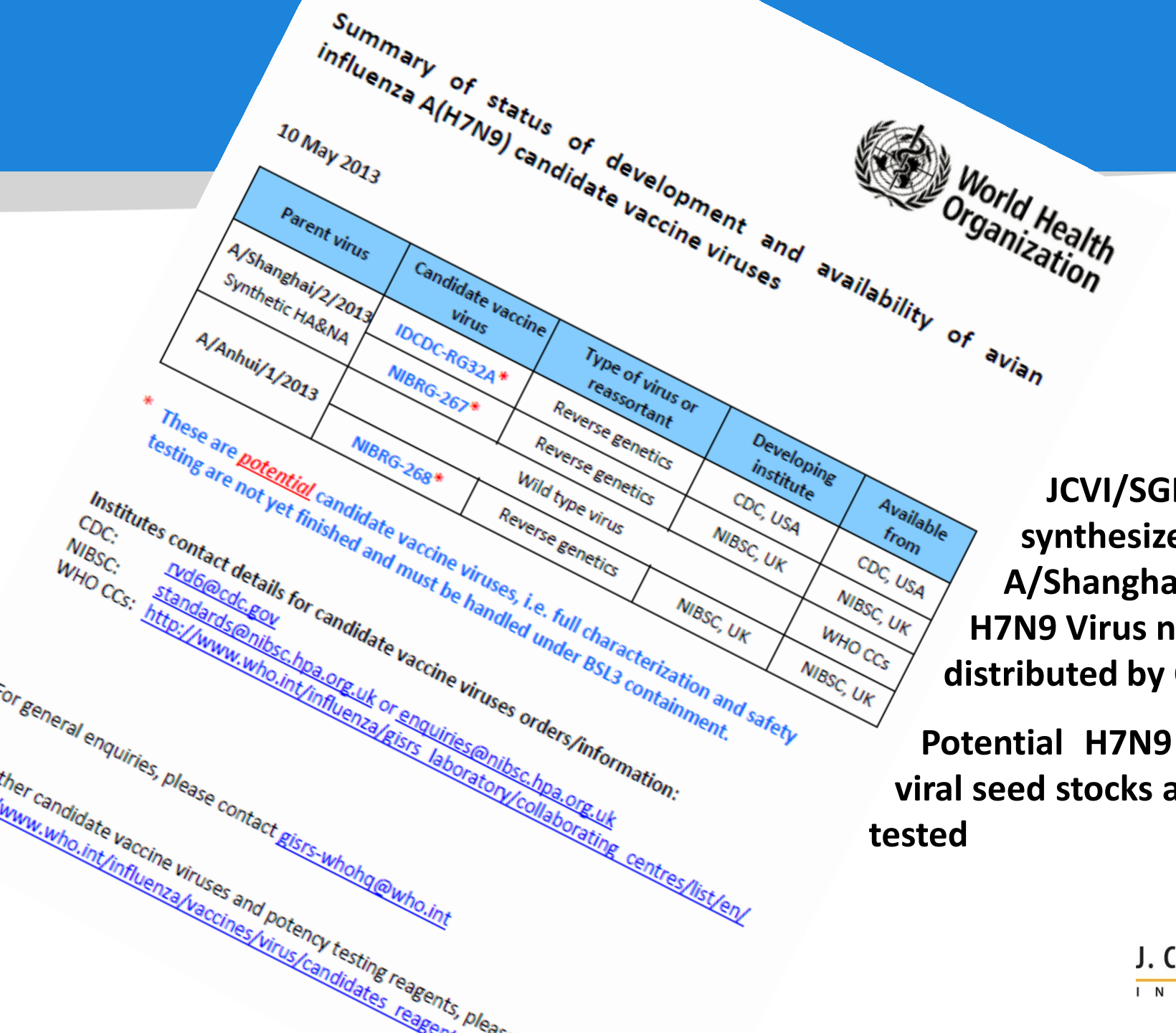
Slide Modified from Peter Mason, Novartis

J. Craig Venter™ INSTITUTE

Accelerating Flu Vaccine Development



J. Craig Venter INSTITUTE



Summary of status of development and availability of avian influenza A(H7N9) candidate vaccine viruses

World Health Organization

10 May 2013

Parent virus	Candidate vaccine virus	Type of virus or reassortant	Developing institute	Available from
A/Shanghai/2/2013 Synthetic HA&NA	IDCDC-RG32A*	Reverse genetics	CDC, USA	CDC, USA
	NIBRG-267*	Reverse genetics	NIBSC, UK	NIBSC, UK
A/Anhui/1/2013		Wild type virus		WHO CCs
	NIBRG-268*	Reverse genetics	NIBSC, UK	NIBSC, UK

* These are potential candidate vaccine viruses, i.e. full characterization and safety testing are not yet finished and must be handled under BSL3 containment.

Institutes contact details for candidate vaccine viruses orders/information:
CDC: rvd6@cdc.gov
NIBSC: standards@nibsc.hpa.org.uk or enquiries@nibsc.hpa.org.uk
WHO CCs: http://www.who.int/influenza/gisrs_laboratory/collaborating_centres/list/en/

For general enquiries, please contact gisrs-whohq@who.int

ther candidate vaccine viruses and potency testing reagents, plea
www.who.int/influenza/vaccines/virus/candidates_reagen

JCVI/SGI/Novartis synthesized A/Shanghai/2/2013 H7N9 Virus now being distributed by CDC

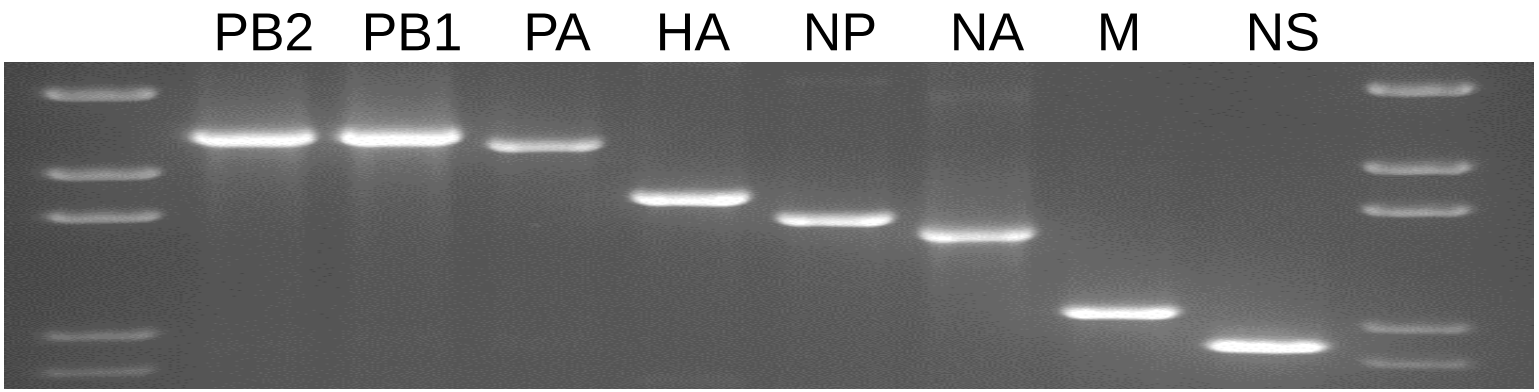
Potential H7N9 vaccine viral seed stocks are being tested

J. Craig Venter INSTITUTE

Emerging Viral Genome Synthesis

- Identified unique H7N9 virus in people in late March
- Novel subtype for humans
  - Antigenic shift -> Pandemic potential
- Sequence of first viruses available April 1
- 135 cases to date

- Synthesized the H7N9 genome (Wentworth, A/Anhui/1-JCVI.1/2013)

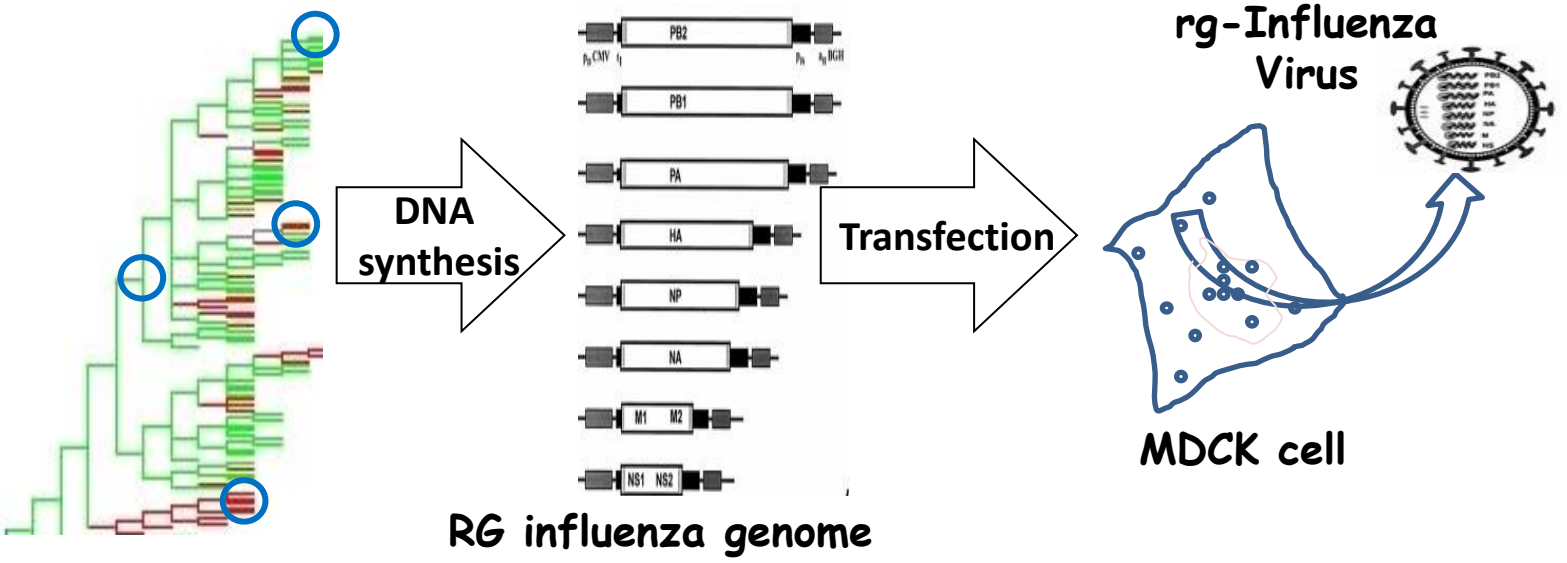


- Research and Experimental Live Attenuated AIV Production

- Bat influenza
- Coronaviruses: MERS, HKU1
- Morbillivirus
- Rhinovirus

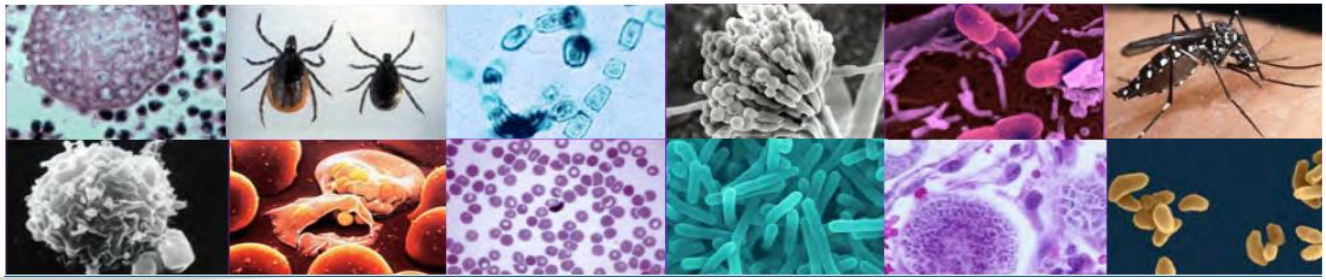
J. Craig Venter™ INSTITUTE

Summary



- **Synthetic genomics** - create gene segments (BARDA/Novartis) or pre-existing gene segments could be used (synfluenza)
- **Rescue vaccine pre-seeds -** 6:2 vaccine seeds (TIV, LAIV)
  - Pre-existing stocks ?
- Engineered complete genomes as LAIVs?

J. Craig Venter INSTITUTE



These projects* have been funded with federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services through the Genomic Sequencing Centers for Infectious Diseases.

* Vaccine Informatics funded by HHS.

National Institute of Allergy and Infectious Diseases
Leading research to understand, treat, and prevent infectious, immunologic, and allergic diseases.

J. Craig Venter™ INSTITUTE

BROAD INSTITUTE

UNIVERSITY of MARYLAND SCHOOL of MEDICINE 200 Years

INSTITUTE FOR GENOME SCIENCES

Synfluenza Summary

- Purpose:
  - Develop a technical capability to generate and stockpile synthetic DNA encoding influenza gene segment, which could be used to produce virus seeds stocks.
- Deliverable
  - Library of ~1000 sequence verified HA & NA genes
    - Available through the Biodefense and Emerging Infections Research Resource Program (BEI)
- Synthetic gene segment generation
  - Gibson in-vitro assembly
  - Assembly uses automated robotic systems
  - Enables construction of an extensive library of influenza genes
    - Potential to use cassettes in the future for new viruses
- Library of clones
  - Vaccine seeds
  - Diagnostics
  - Basic Research

J. Craig Venter™ INSTITUTE