A*STAR

Genome Institute of Singapore

# Regional Biosecurity Workshop

## May 28, 2007

**Edison Liu**
liue@gis.a-star.edu.sg
+65-6478-8007

# How might we deal with bioterrorist threats?

- Block access to dangerous agents
- Detect trafficking
- Detect outbreak
- Trace to source

### Category A

- Bacillus anthracis (anthrax)
- Clostridium botulinum toxin (botulism)
- Yersinia pestis (plague)
- Variola major (smallpox) and other related pox viruses
- Francisella tularensis (tularemia)
- Viral hemorrhagic fevers
  - Arenaviruses
    - LCM, Junin virus, Machupo virus, Guanarito virus
    - Lassa Fever
  - Bunyaviruses
    - Hantaviruses
    - Rift Valley Fever
  - Flaviruses
    - Dengue
  - Filoviruses
    - Ebola
    - Marburg

- Additional viral encephalitides
- West Nile Virus
- LaCrosse
- California encephalitis
- VEE
- EEE
- WEE
- Japanese Encephalitis Virus
- Kyasanur Forest Virus

### Category B

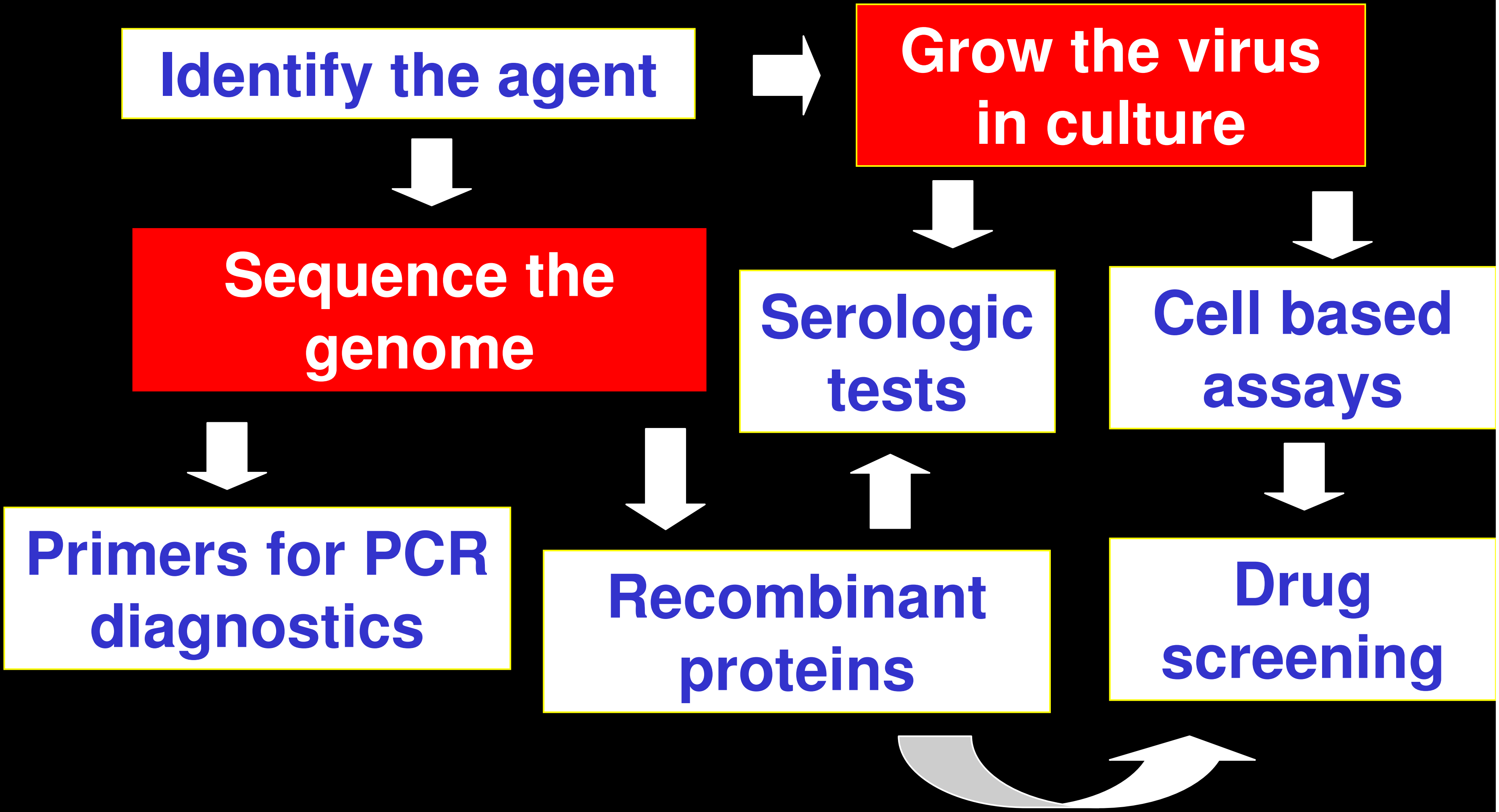
- Burkholderia pseudomallei
- Coxiella burnetii (Q fever)
- Brucella species (brucellosis)
- Burkholderia mallei (glanders)
- Chlamydia psittaci (Psittacosis)*
- Ricin toxin (from Ricinus communis)
- Epsilon toxin of Clostridium perfringens
- Staphylococcus enterotoxin B
- Typhus fever (Rickettsia prowazekii)
- Food and Waterborne Pathogens
  - Bacteria
    - Diarrheagenic E.coli
    - Pathogenic Vibrios
    - Shigella species
    - Salmonella
    - Listeria monocytogenes
    - Campylobacter jejuni
    - Yersinia enterocolitica)
  - ruses (Caliciviruses, Hepatitis A)
  - otozoa
    - Cryptosporidium parvum
    - Cyclospora cayatanensis
    - Giardia lamblia
    - Entamoeba histolytica
    - Toxoplasma
    - Microsporidia

**With only a few exceptions, candidate bioterrorism agents look like standard infectious disease agents**

# How did we deal with a serious infectious threat?

- Organization
- Technologies
- Communications

# Singapore Battle Plan: SARS CoV
## April 2, 2003



Identify the agent

Grow the virus in culture

Sequence the genome

Serologic tests

Cell based assays

Primers for PCR diagnostics

Recombinant proteins

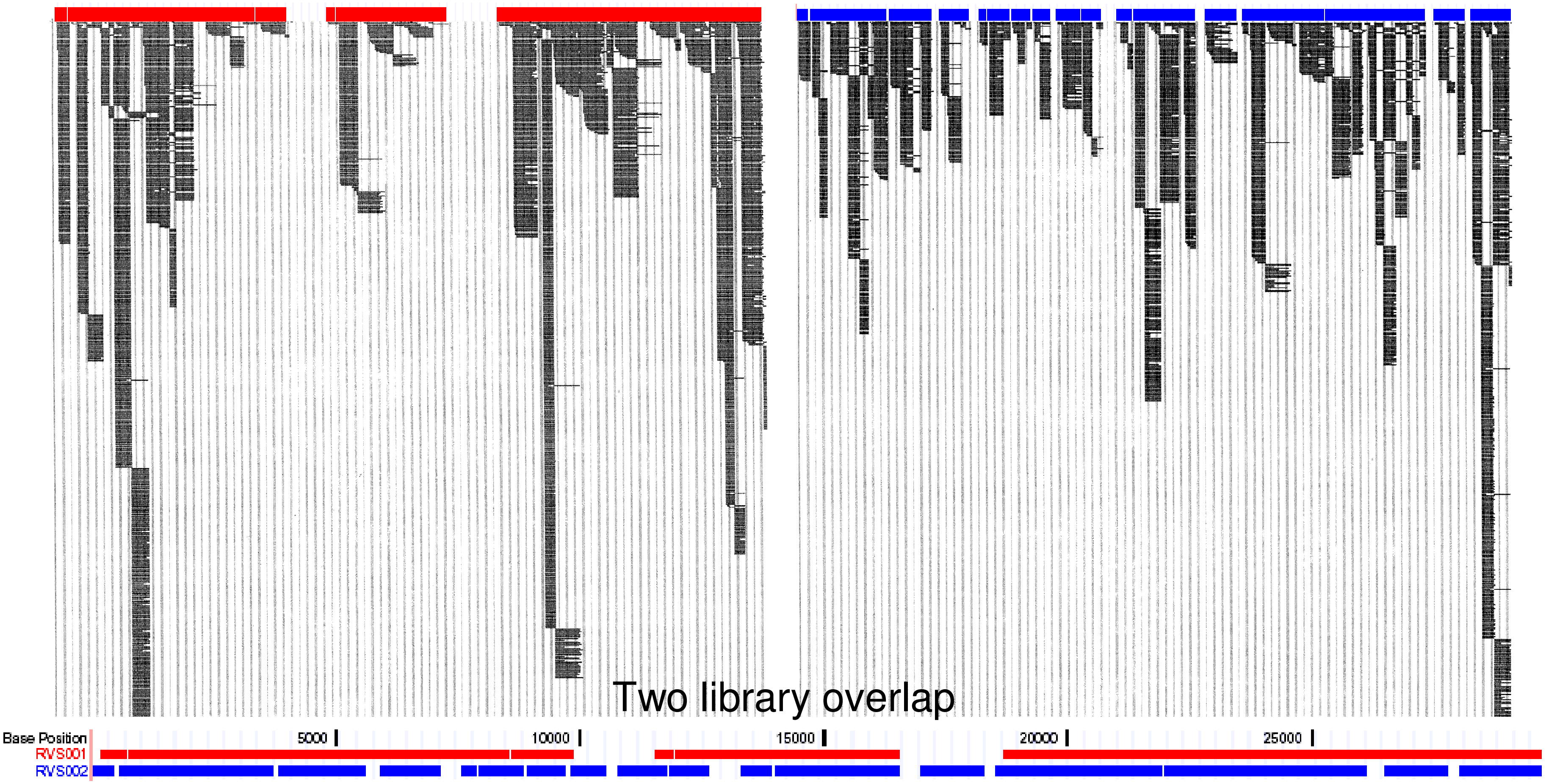
Drug screening

# Viral Genome Discovery

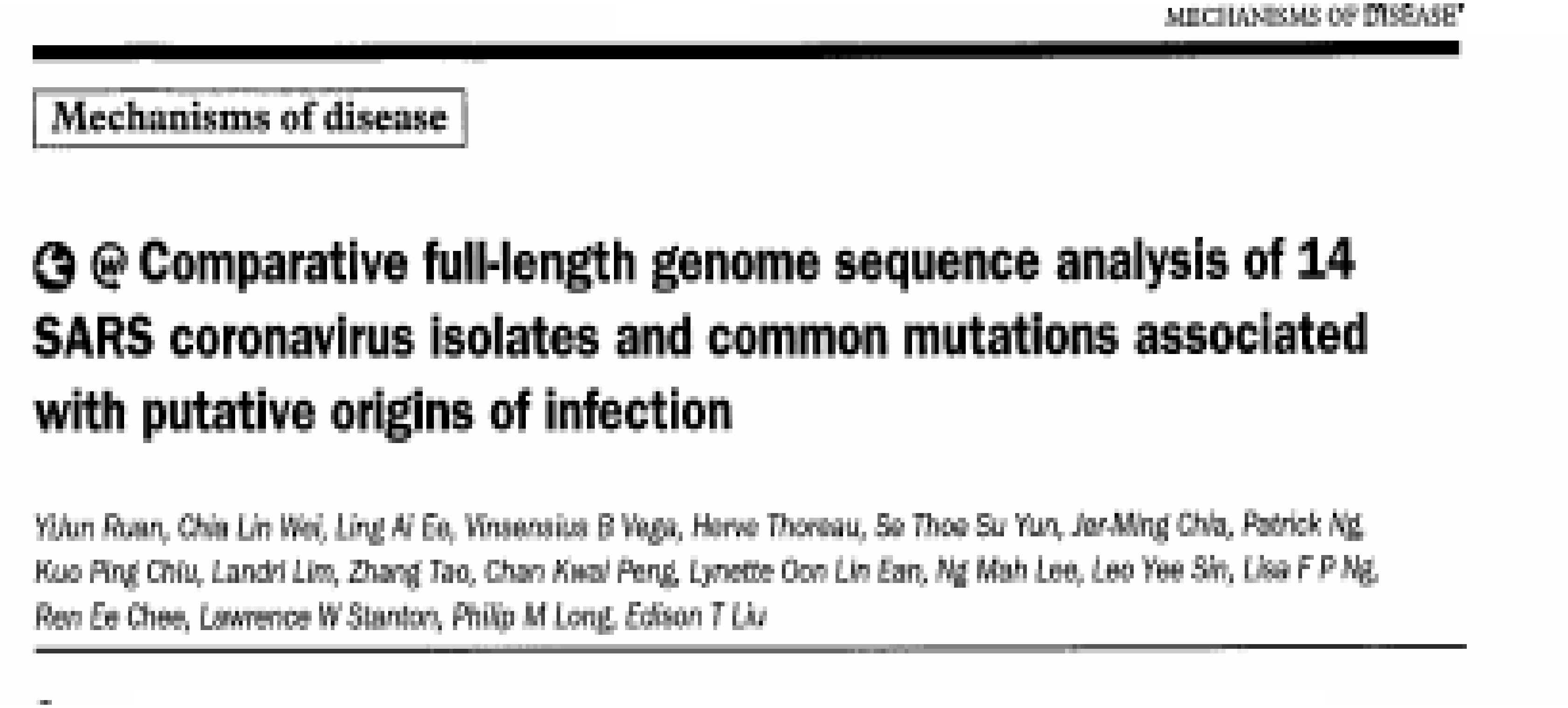
## SARS-Cov genome coverage by shotgun libraries

Taq I library (2000 reads)

Csp6I library (2000 reads)



# There is no distinction between good research and essential public health measures



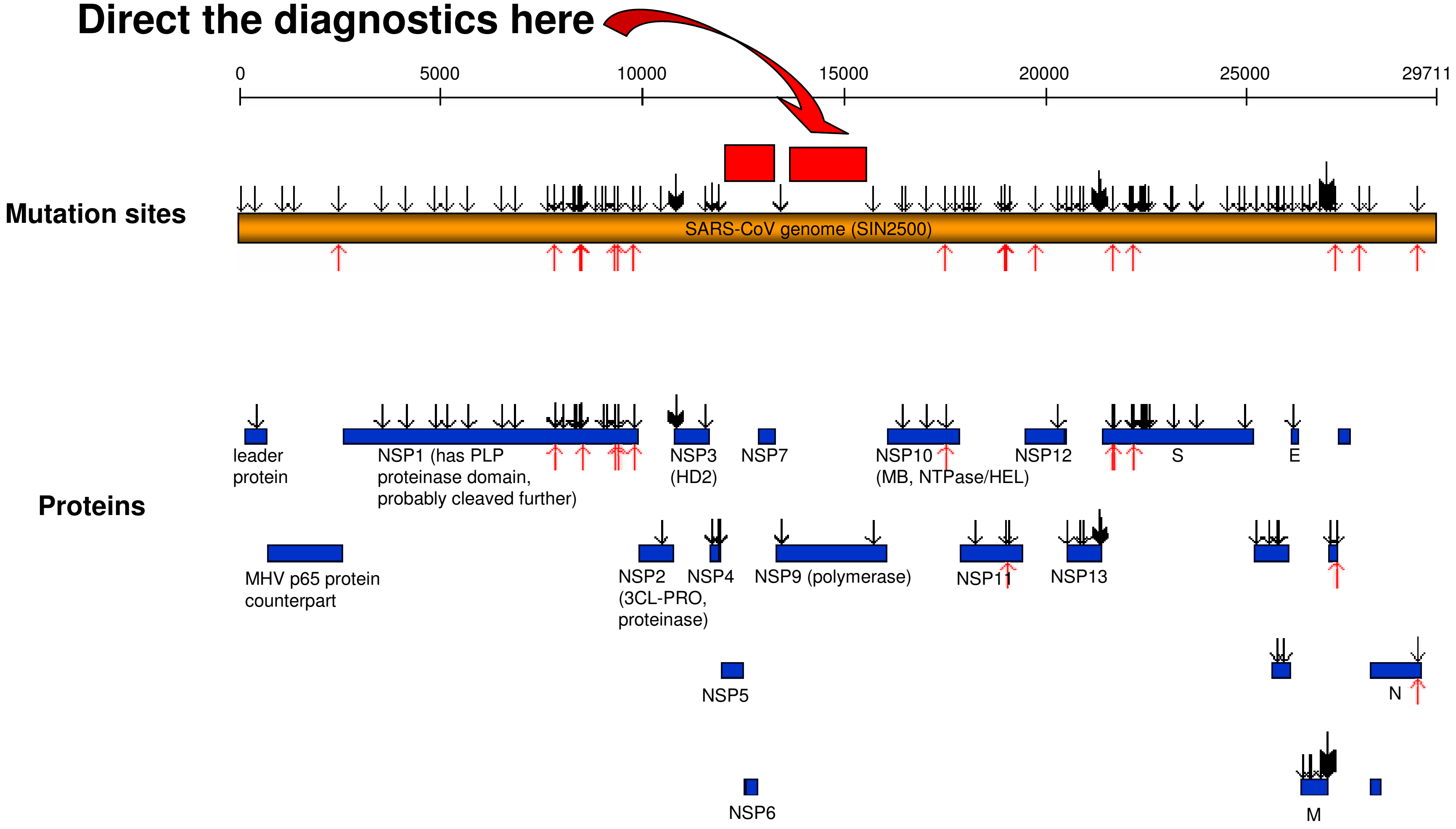
MECHANISMS OF DISEASE

Mechanisms of disease

@ Comparative full-length genome sequence analysis of 14 SARS coronavirus isolates and common mutations associated with putative origins of infection

*YiJun Ruan, Chia Lin Wei, Ling Ai Ee, Vinsensius B Vega, Herve Thoreau, Se Thoe Su Yun, Jer-Ming Chia, Patrick Ng, Kuo Ping Chiu, Landri Lim, Zhang Tao, Chan Kwai Peng, Lynette Oon Lin Ean, Ng Mah Lee, Leo Yee Sin, Lisa F P Ng, Ren Ee Chee, Lawrence W Stanton, Philip M Long, Edison T Liu*

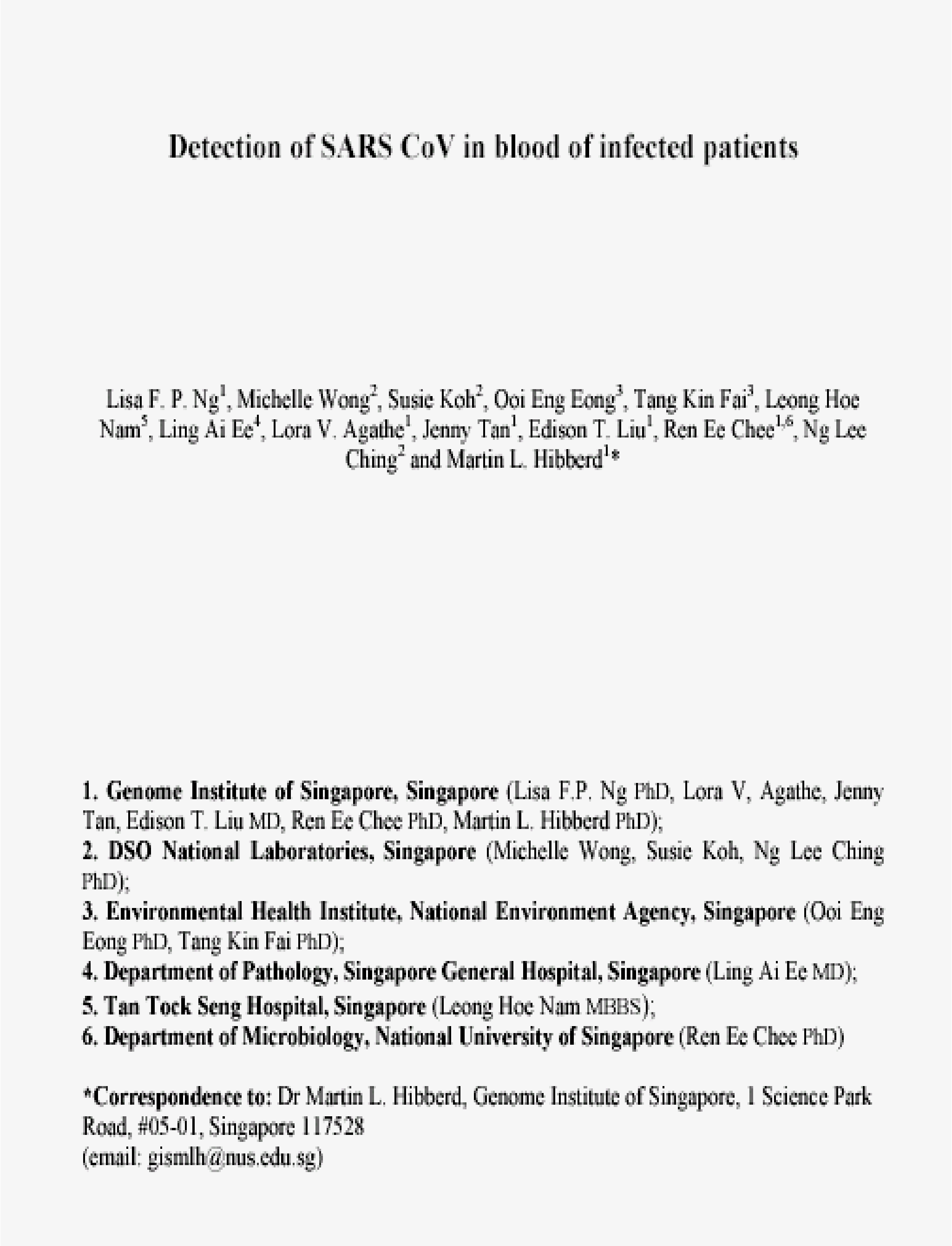
**Lancet. May 24;361(9371):1779-85, 2003.**



Direct the diagnostics here

**There are 129 places in the SARS genome different amongst the 14 SARS isolates.**

**Want diagnostics that will work on all isolates.**

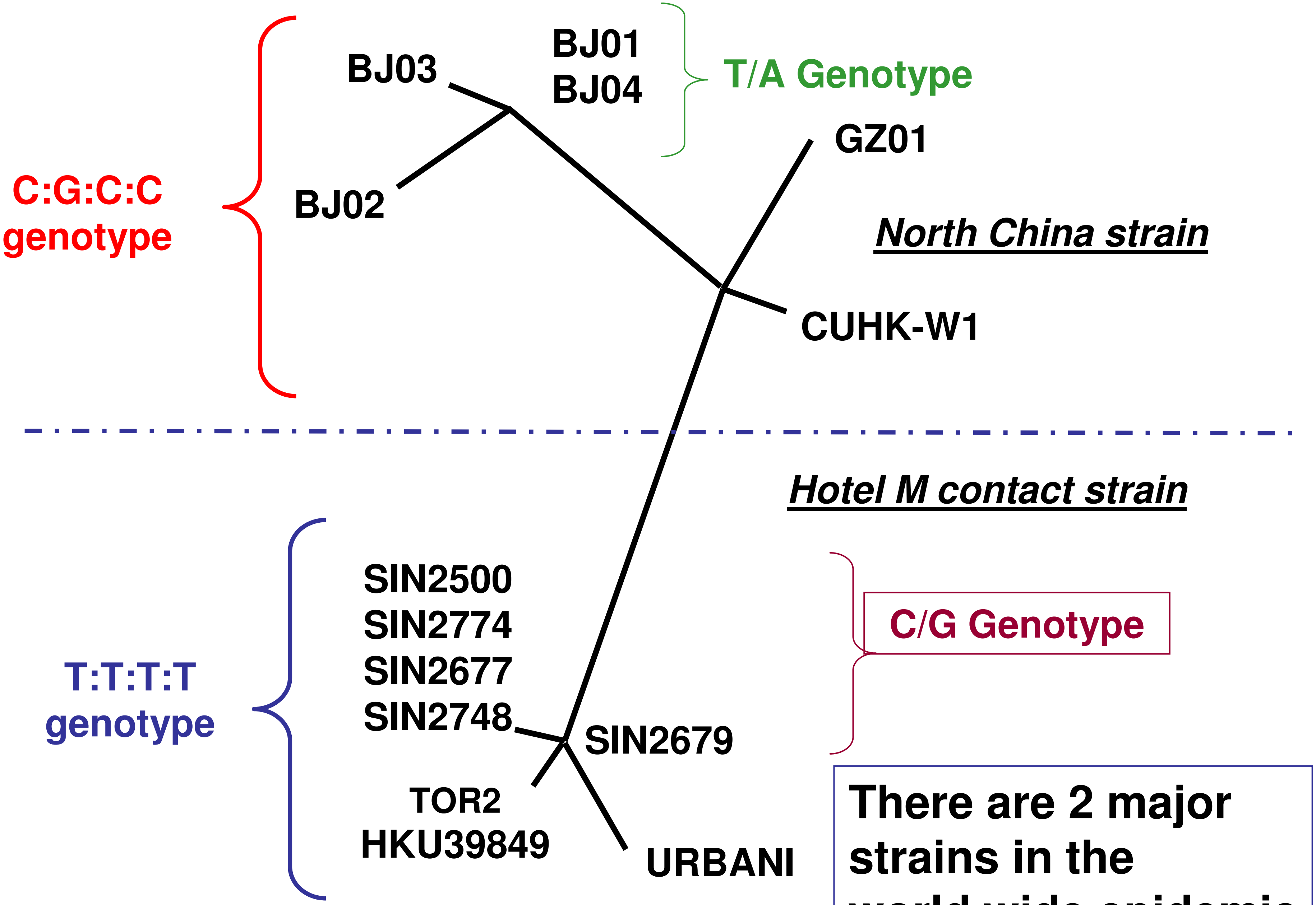


# Detection of SARS CoV in blood of infected patients

Lisa F. P. Ng[1], Michelle Wong[2], Susie Koh[2], Ooi Eng Eong[3], Tang Kin Fai[3], Leong Hoe Nam[5], Ling Ai Ee[4], Lora V. Agathe[1], Jenny Tan[1], Edison T. Liu[1], Ren Ee Chee[1,6], Ng Lee Ching[2] and Martin L. Hibberd[1]*

1. **Genome Institute of Singapore, Singapore** (Lisa F.P. Ng PhD, Lora V, Agathe, Jenny Tan, Edison T. Liu MD, Ren Ee Chee PhD, Martin L. Hibberd PhD);
2. **DSO National Laboratories, Singapore** (Michelle Wong, Susie Koh, Ng Lee Ching PhD);
3. **Environmental Health Institute, National Environment Agency, Singapore** (Ooi Eng Eong PhD, Tang Kin Fai PhD);
4. **Department of Pathology, Singapore General Hospital, Singapore** (Ling Ai Ee MD);
5. **Tan Tock Seng Hospital, Singapore** (Leong Hoe Nam MBBS);
6. **Department of Microbiology, National University of Singapore** (Ren Ee Chee PhD)

***Correspondence to:** Dr Martin L. Hibberd, Genome Institute of Singapore, 1 Science Park Road, #05-01, Singapore 117528
(email: gismlh@nus.edu.sg)



C:G:C:C genotype

BJ03

BJ01
BJ04

T/A Genotype

BJ02

GZ01

*__North China strain__*

CUHK-W1

*__Hotel M contact strain__*

T:T:T:T genotype

SIN2500
SIN2774
SIN2677
SIN2748

SIN2679

C/G Genotype

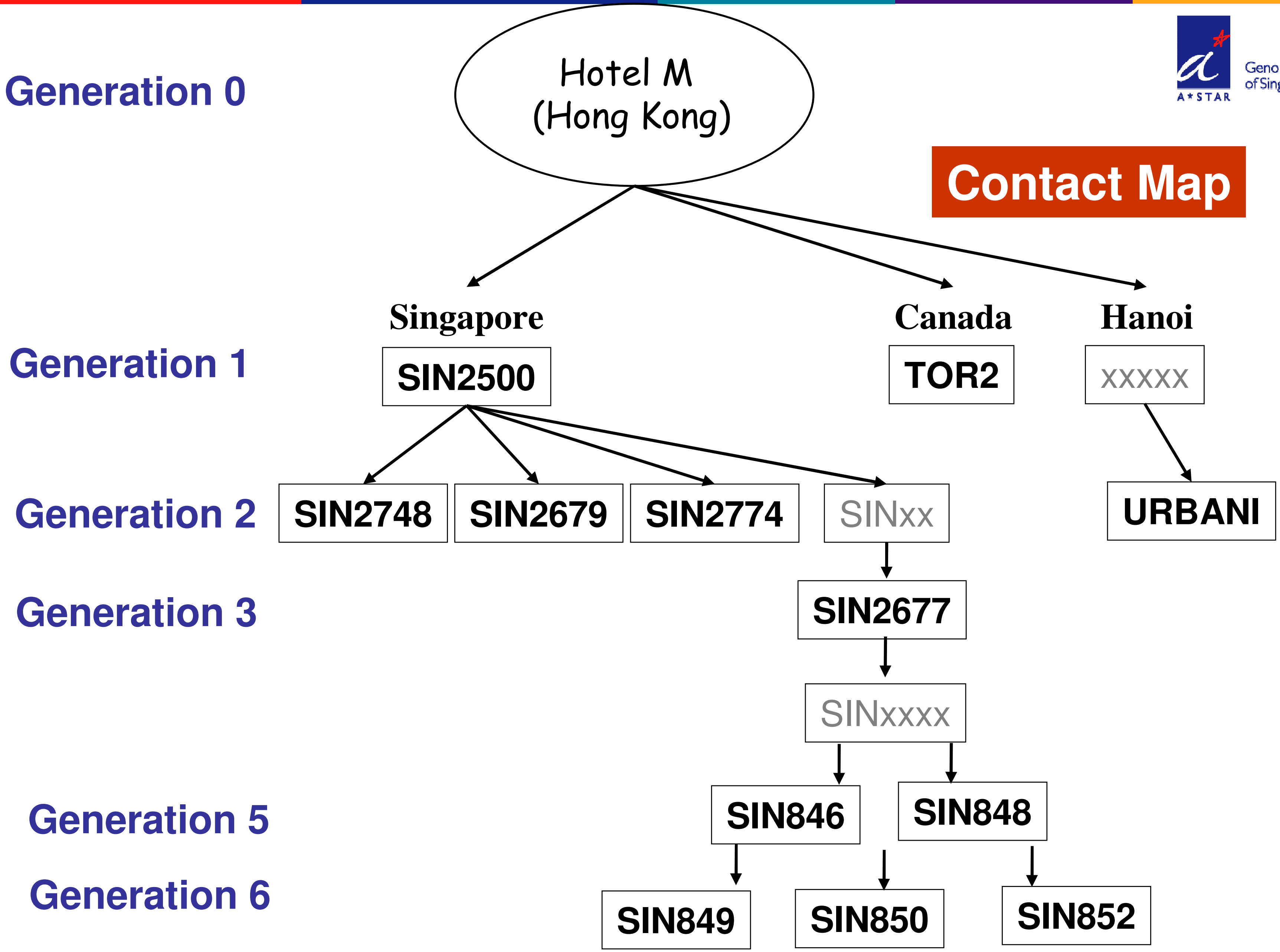
TOR2
HKU39849

URBANI

**There are 2 major strains in the world wide epidemic**

Genome Institute of Singapore

# Contact Map



**Generation 0**

Hotel M (Hong Kong)

**Generation 1**

**Singapore**

**SIN2500**

**Canada**

**TOR2**

**Hanoi**

xxxxx

**Generation 2**

**SIN2748**

**SIN2679**

**SIN2774**

SINxx

**URBANI**

**Generation 3**

**SIN2677**

SINxxxx

**Generation 5**

**SIN846**

**SIN848**

**Generation 6**

**SIN849**

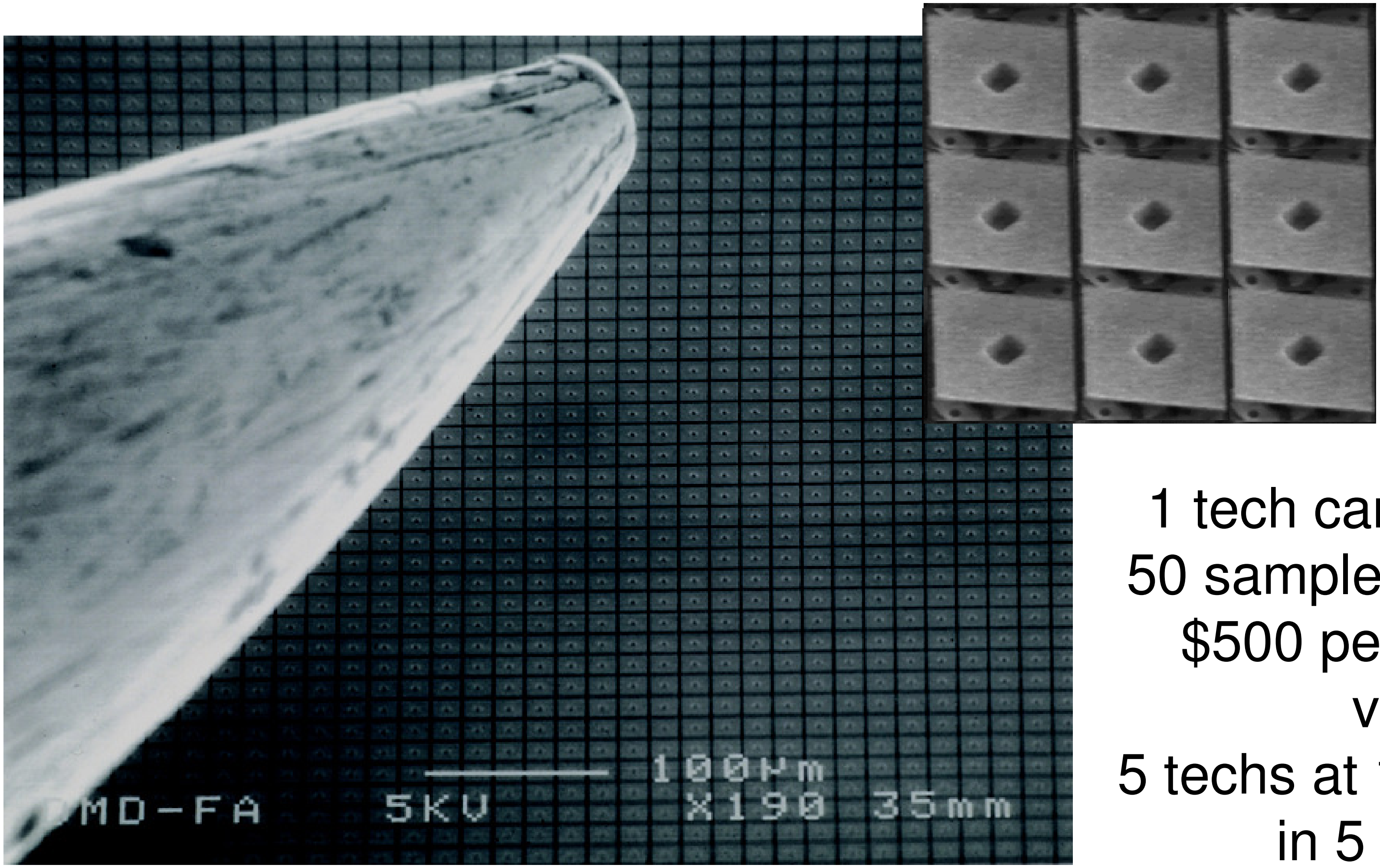
**SIN850**

**SIN852**

# Assignment of SARS CoV strains assists in contact mapping and contagion monitoring

Need: to develop a fast and inexpensive genotyping approach

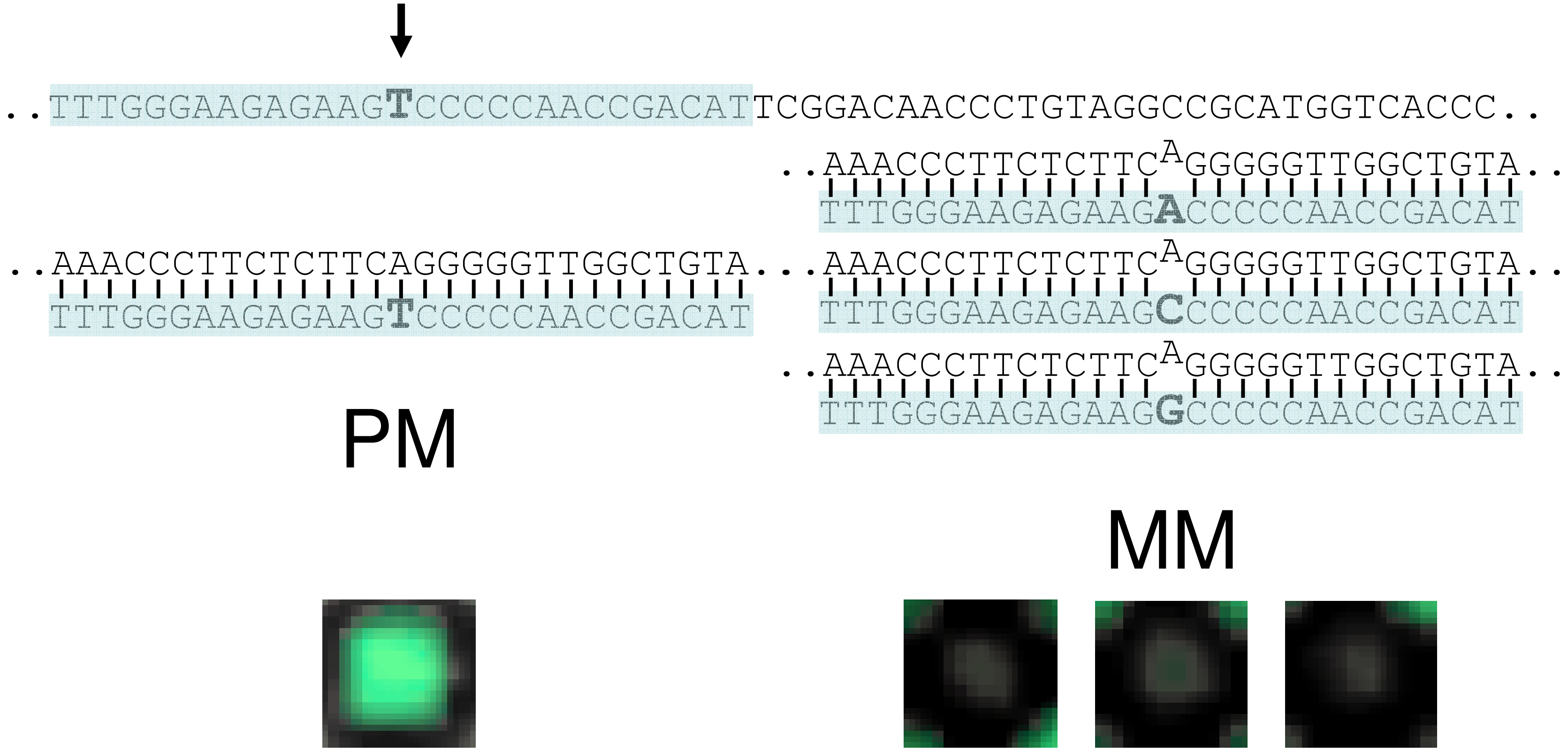
## 385,000 probes to sequence the SARS genome



1 tech can process
50 samples in 5 days
$500 per sample
vs
5 techs at 10 samples
in 5 days

# Sequencing by Hybridization

A*STAR
Genome Institute of Singapore



..TTTGGGAAGAGAAG**T**CCCCCAACCGACATTCGGACAACCCTGTAGGCCGCATGGTCACCC..

..AAACCCTTCTCTTCAGGGGGTTGGCTGTA..
TTTGGGAAGAGAAG**T**CCCCCAACCGACAT

PM

..AAACCCTTCTCTTC$^{A}$GGGGGTTGGCTGTA..
TTTGGGAAGAGAAG**A**CCCCCAACCGACAT

..AAACCCTTCTCTTC$^{A}$GGGGGTTGGCTGTA..
TTTGGGAAGAGAAG**C**CCCCCAACCGACAT

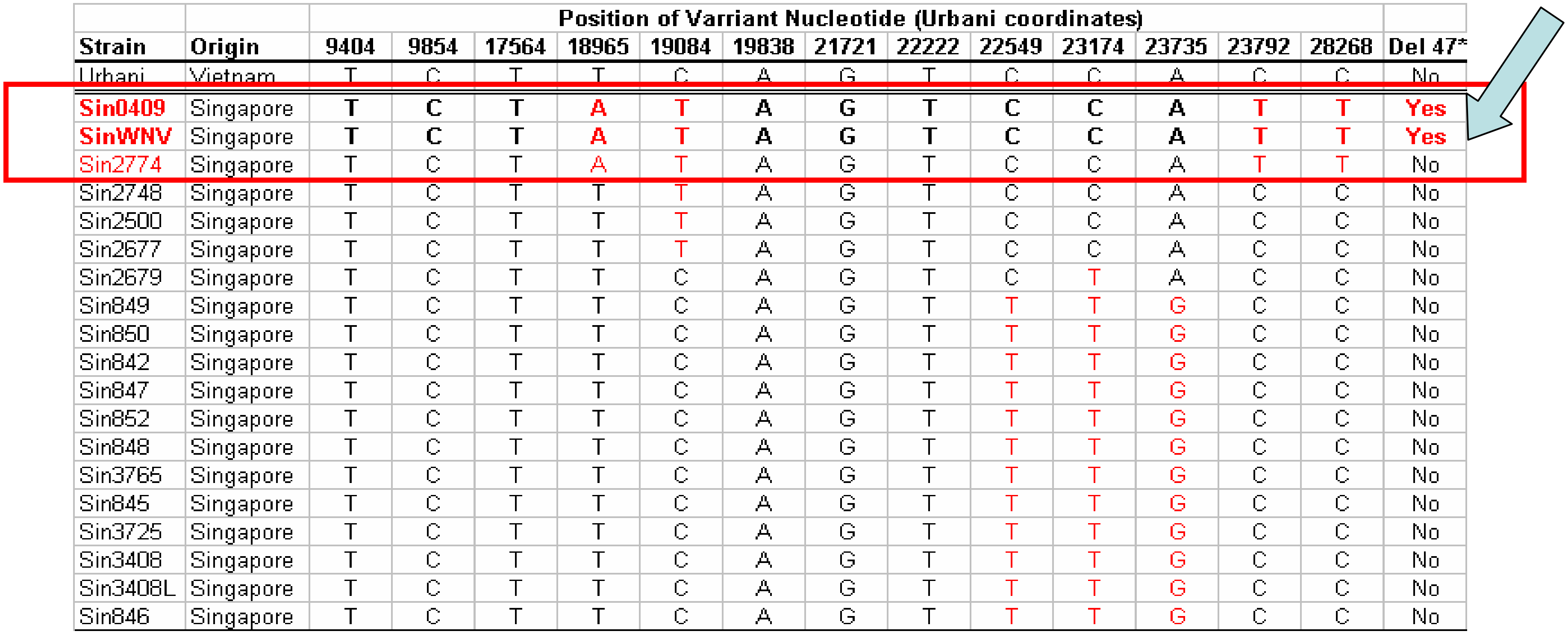
..AAACCCTTCTCTTC$^{A}$GGGGGTTGGCTGTA..
TTTGGGAAGAGAAG**G**CCCCCAACCGACAT

MM

# Lab-acquired SARS case

- A 27 year old graduate student tested positive for SARS in September 2003
- No known contacts with SARS patients or travel to active SARS areas
- Student worked on West Nile Virus in BSL3 lab that works with SARS-CoV
- GIS sequenced the virus directly from patient sample using array-based method

# Contact tracing of a field sample



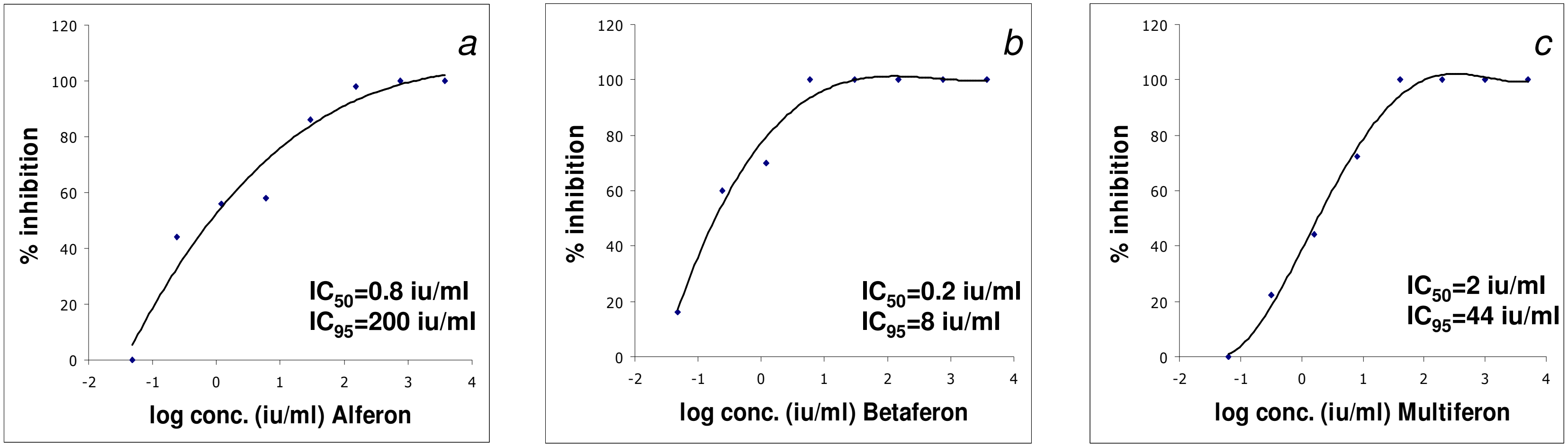
| Strain | Origin | Position of Varriant Nucleotide (Urbani coordinates) | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 9404 | 9854 | 17564 | 18965 | 19084 | 19838 | 21721 | 22222 | 22549 | 23174 | 23735 | 23792 | 28268 | Del 47* |
| Urbani | Vietnam | T | C | T | T | C | A | G | T | C | C | A | C | C | No |
| **Sin0409** | Singapore | **T** | **C** | **T** | **A** | **T** | **A** | **G** | **T** | **C** | **C** | **A** | **T** | **T** | **Yes** |
| **SinWNV** | Singapore | **T** | **C** | **T** | **A** | **T** | **A** | **G** | **T** | **C** | **C** | **A** | **T** | **T** | **Yes** |
| Sin2774 | Singapore | T | C | T | A | T | A | G | T | C | C | A | T | T | No |
| Sin2748 | Singapore | T | C | T | T | T | A | G | T | C | C | A | C | C | No |
| Sin2500 | Singapore | T | C | T | T | T | A | G | T | C | C | A | C | C | No |
| Sin2677 | Singapore | T | C | T | T | T | A | G | T | C | C | A | C | C | No |
| Sin2679 | Singapore | T | C | T | T | C | A | G | T | C | T | A | C | C | No |
| Sin849 | Singapore | T | C | T | T | C | A | G | T | T | T | G | C | C | No |
| Sin850 | Singapore | T | C | T | T | C | A | G | T | T | T | G | C | C | No |
| Sin842 | Singapore | T | C | T | T | C | A | G | T | T | T | G | C | C | No |
| Sin847 | Singapore | T | C | T | T | C | A | G | T | T | T | G | C | C | No |
| Sin852 | Singapore | T | C | T | T | C | A | G | T | T | T | G | C | C | No |
| Sin848 | Singapore | T | C | T | T | C | A | G | T | T | T | G | C | C | No |
| Sin3765 | Singapore | T | C | T | T | C | A | G | T | T | T | G | C | C | No |
| Sin845 | Singapore | T | C | T | T | C | A | G | T | T | T | G | C | C | No |
| Sin3725 | Singapore | T | C | T | T | C | A | G | T | T | T | G | C | C | No |
| Sin3408 | Singapore | T | C | T | T | C | A | G | T | T | T | G | C | C | No |
| Sin3408L | Singapore | T | C | T | T | C | A | G | T | T | T | G | C | C | No |
| Sin846 | Singapore | T | C | T | T | C | A | G | T | T | T | G | C | C | No |

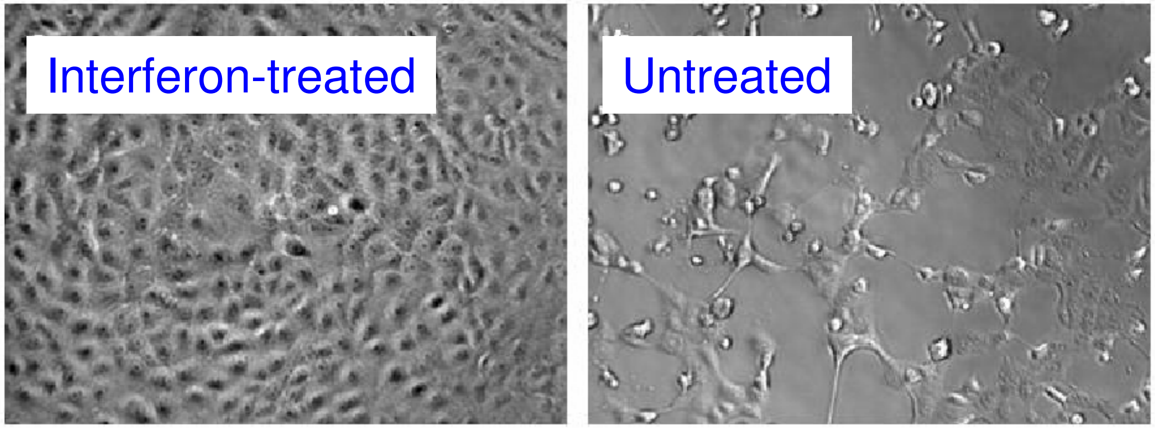
# Towards a Treatment for SARS:

- Identify effective drugs

    Vero cell assay

# Interferons inhibit SARS in plaque assay





SARS-infected Vero cells

| **Antiviral** | **Source** | **Highest concentration tested** | **Inhibition of cytopathic effect** |
|---|---|---|---|
| ***Interferons*** | | | |
| Interferon α2a (Roferon®) | Roche | 100,000 iu/ml | No |
| Interferon α2b (Intron A®) | Schering Plough | 500,000 iu/ml | No |
| **Interferon αn1 (Wellferon®)** | **Glaxosmithkline** | **500,000 iu/ml** | **Yes** |
| **Interferon αn3 (Alferon®)** | **Hemispheryx** | **10,000 iu/ml** | **Yes** |
| Interferon β1a (Rebif®) | Serono | 500,000 iu/ml | No |
| **Interferon β1b (Betaferon®)** | **Schering AG** | **100,000 iu/ml** | **Yes** |
| ***Nucleoside analogs*** | | | |
| Acyclovir | Faulding | 1000 μg/ml | No |
| Amantadine (Symmetrel®) | Novartis | 1000 μg/ml | No |
| Cymevene (Ganciclovir) | Roche | 50,000 μg/ml | No |
| Foscarnet (Foscavir®) | Astrazeneca | 8000 μM | No |
| **Ribavirin** | **ICN Pharma** | **10,000 μg/ml** | **Yes**** |

# Applying this model to Dengue

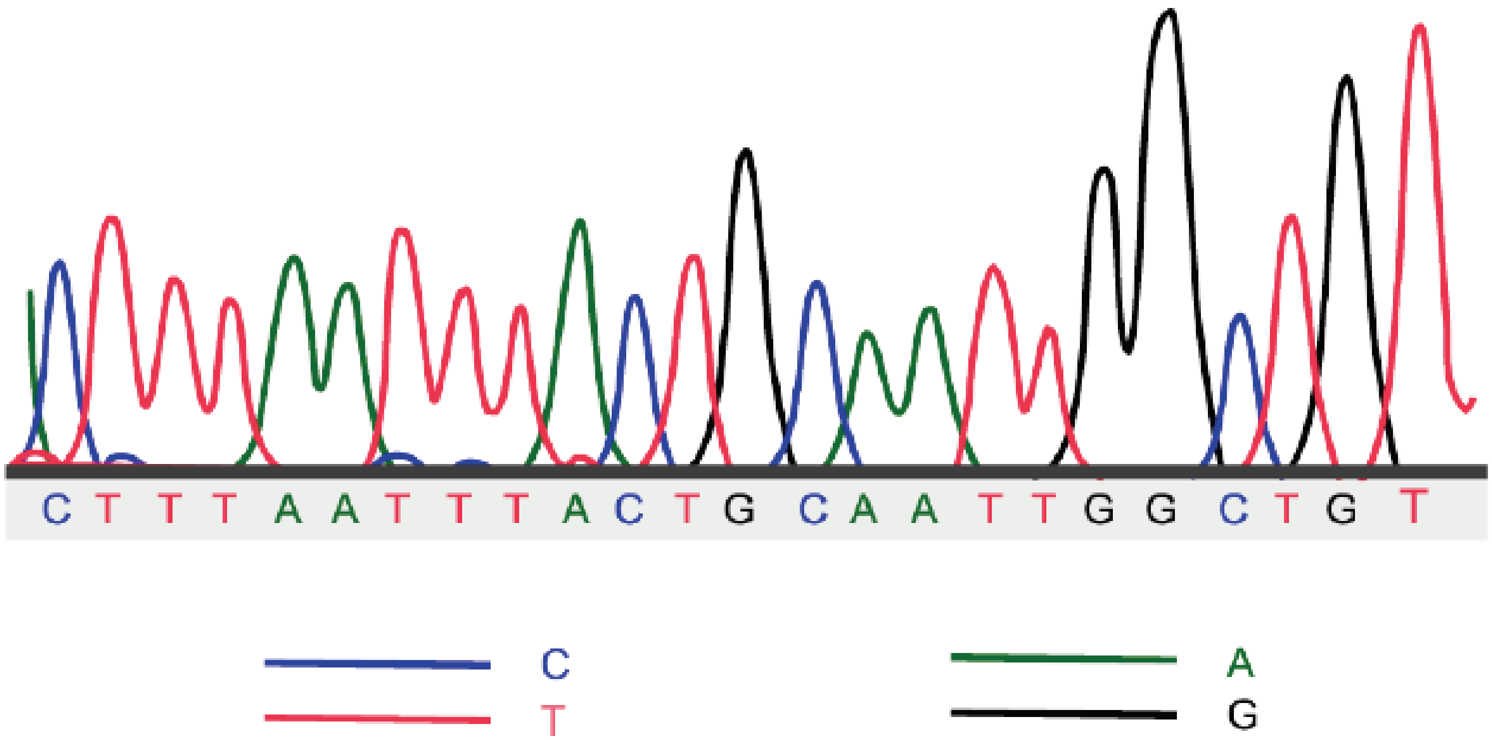
**MOH (TTSH), NEA, GIS, STN, DSO, Novartis**

- Capture cases (MOH – Tan Tock Seng Hospital)
- Prepare patient samples (STN)
- Grow and characterize the virus (DSO, NEA)
- Sequence virus and expression analysis (GIS)
- Find a treatment (Novartis)

- Block access to dangerous agents

- Monitor comprehensively
- Detect trafficking
- Detect outbreak
- Use advanced technologies
- Trace to source
- Integrated and coordinated information dissemination

# The Fundamental Technology



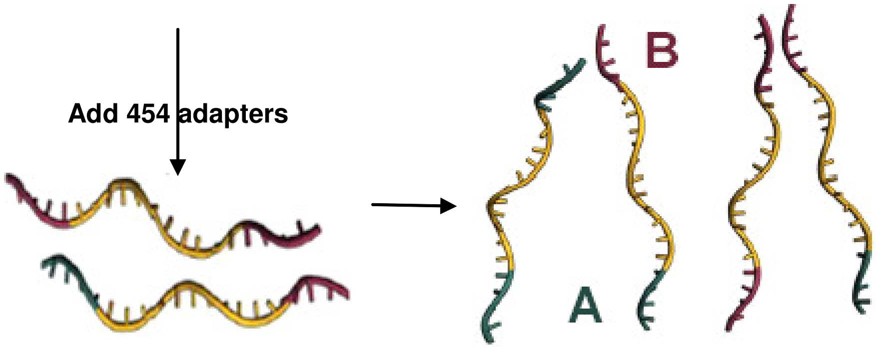
## “Sanger” DNA sequencing

# Multiplex Sequencing: Mix and Match Technology

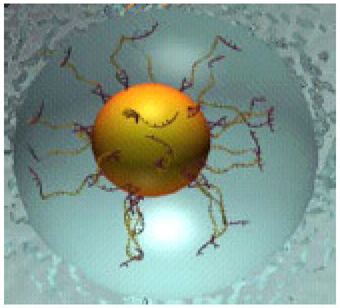
Genome Institute of Singapore

**PCR'd DNA**

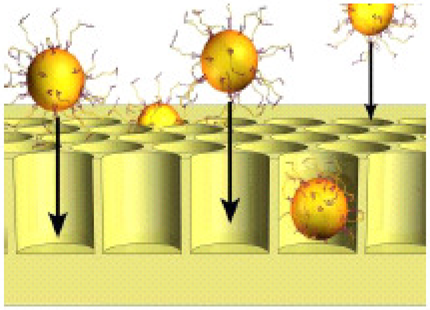
**Add 454 adapters**



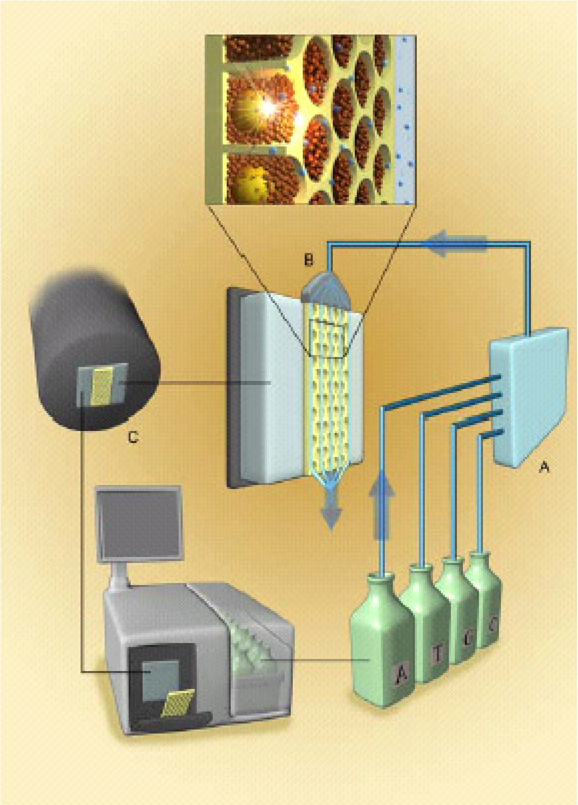
1) Prepare Adapter Ligated ssDNA Library



2) Clonal Amplification on 28 μ beads



3) Load beads and enzymes in PicoTiterPlate™



4) Perform Sequencing by synthesis

454 LIFE SCIENCES

# Sequencing Technologies: Current and Immediate Future
## Sequencing a bacterial genome (2 Mb)

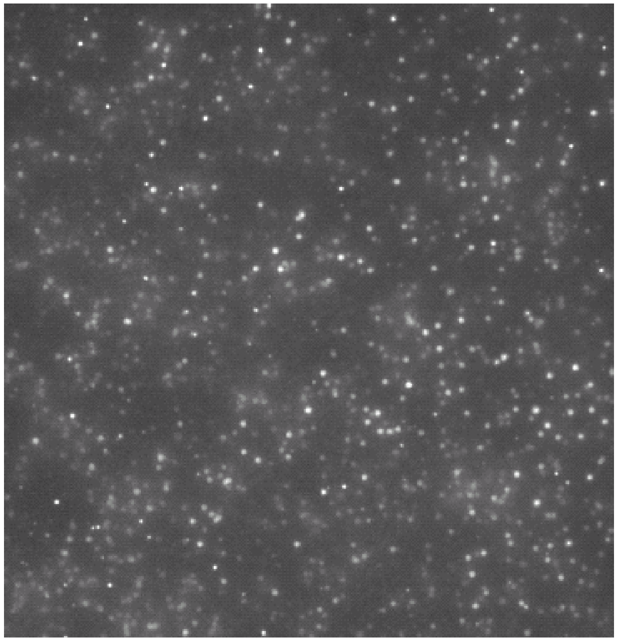
| | "Sanger"<br>ABI 3730 | Roche FLX | Solexa 1G | (Projected)<br>ABI SOLiD |
|---|---|---|---|---|
| Read length (bp) | 750 | 250 | 30-40 | 25-50 |
| Multiplexity (reads) | 400K | 400K | 40M | 100M |
| Total throughput (bp) | 10M | 100M | 1-1.6B | 2-3B |
| Time cost per run | 8 hr | 8 hr | 3 days | 3 days |
| Reagent cost per base in pennies | 0.1 | 0.01 | 0.001 | 0.0001 |
| Cost sequencing H. influenza genome (2 Mb) | $2,000 | $200 | $ 11 | $ 5 |
| Time per bacterial Genome per machine | 1.5 hours | 0.2 hours | 12 minutes for 10 | 6 minutes for 10 |

We can now sequence all possible
samples of microbes in regional
laboratories and
develop a forensic database
of infectious agents

# Discovering microbes in the environment

# A pilot project

Metagenomic Analyses of an Uncultured Viral Population from Human Feces (a GIS-SDSU collaboration)



**500g fresh fecal matter from a 33 yr healthy male**

↓

**Purify & concentrate virions**

↓ ↓

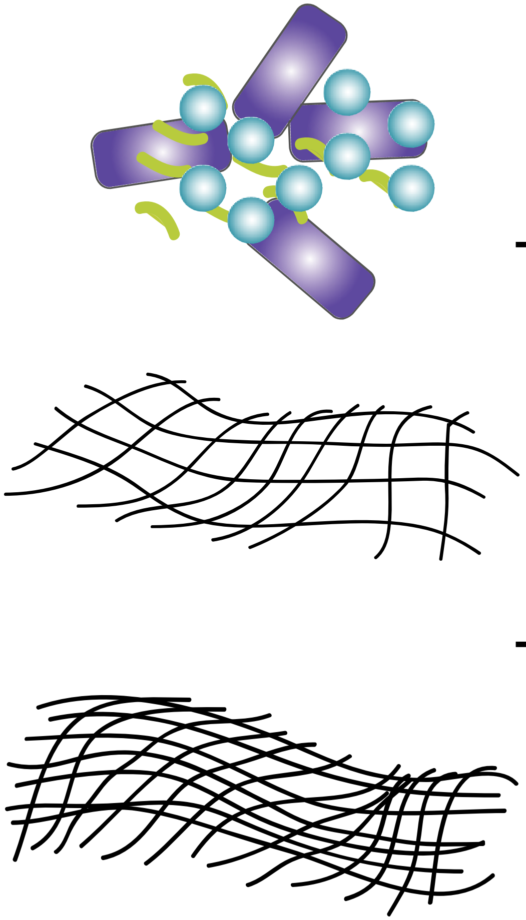
**DNA viral library** **RNA viral library**

↓ ↓

500 clones seq 500 clones seq

# Viral Genome Discovery

## Isolating viruses from original virus-containing samples



→ Bacterial and big particles

100 μm filter

→ virus → Density gradient

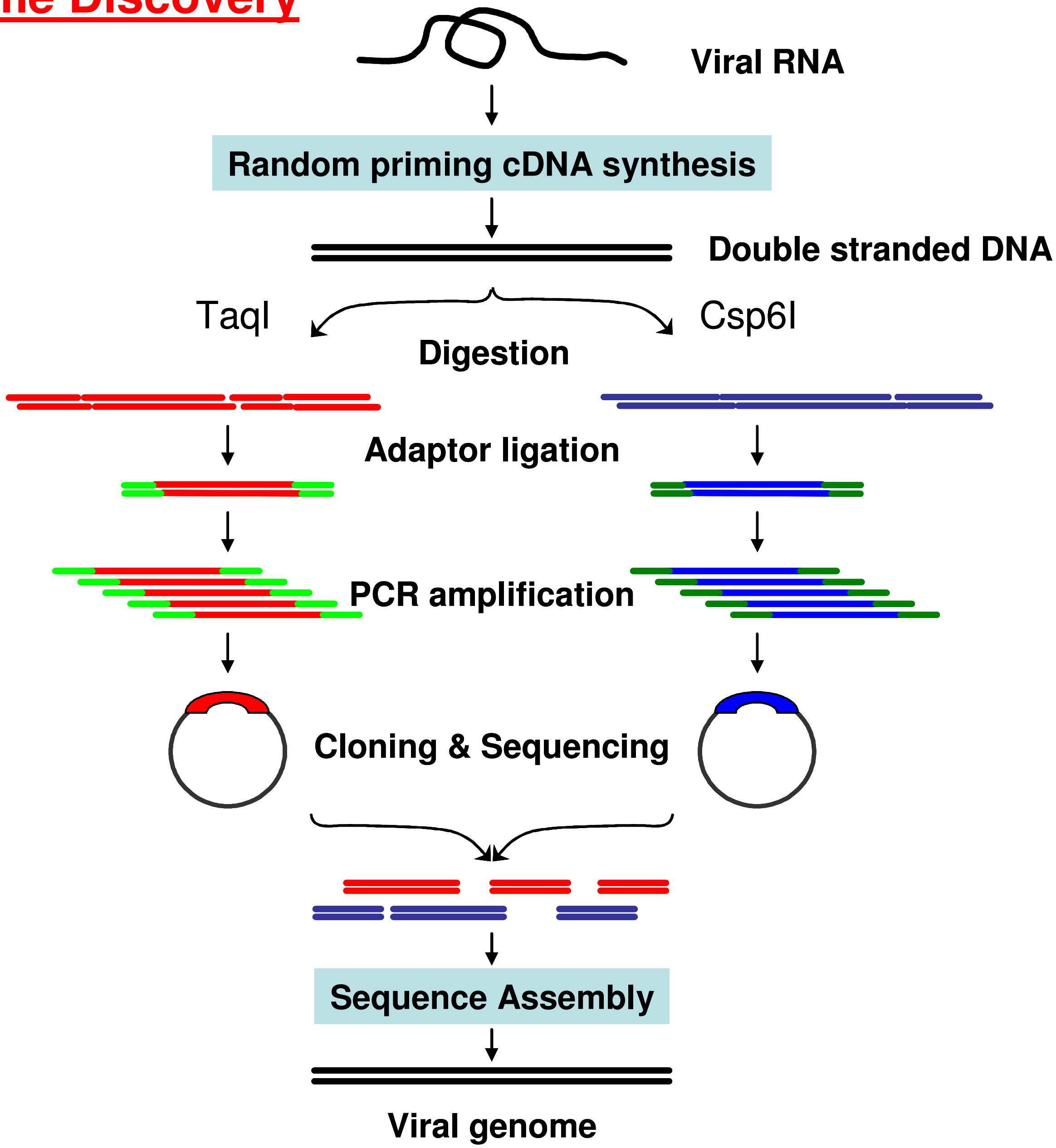
100 kD filter

→ Free nucleotides

# Viral Genome Discovery

Genome Institute of Singapore

A*STAR



Viral RNA

Random priming cDNA synthesis

Double stranded DNA

TaqI

Digestion

Csp6I

Adaptor ligation

PCR amplification

Cloning & Sequencing

Sequence Assembly

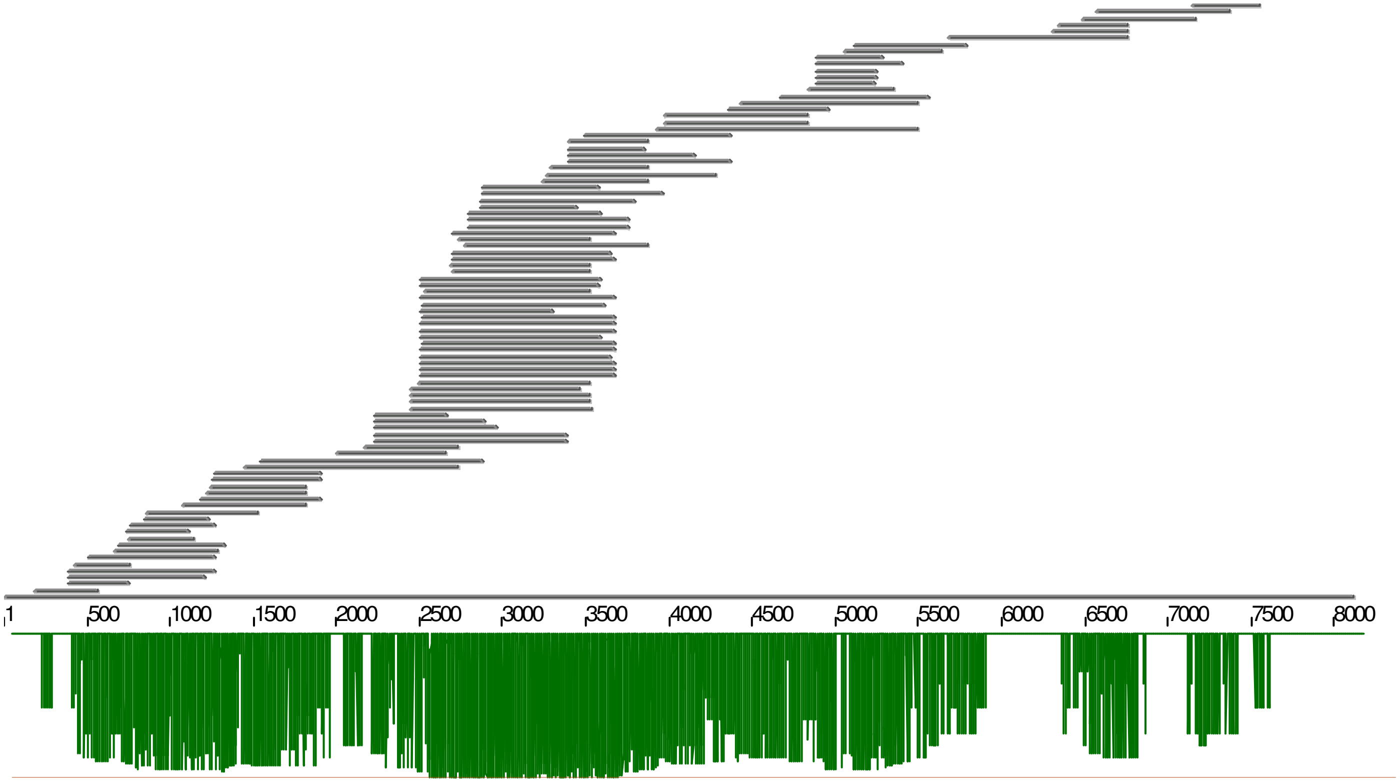
Viral genome

# Viral Genome Discovery

A*STAR

Genome Institute of Singapore

## Assemble a whole viral genome from metagenome sequence data



**The 300 RNA viral clones aligned to the pepper mild mottle virus genome**

# Viral Genome Discovery

Genome Institute
of Singapore

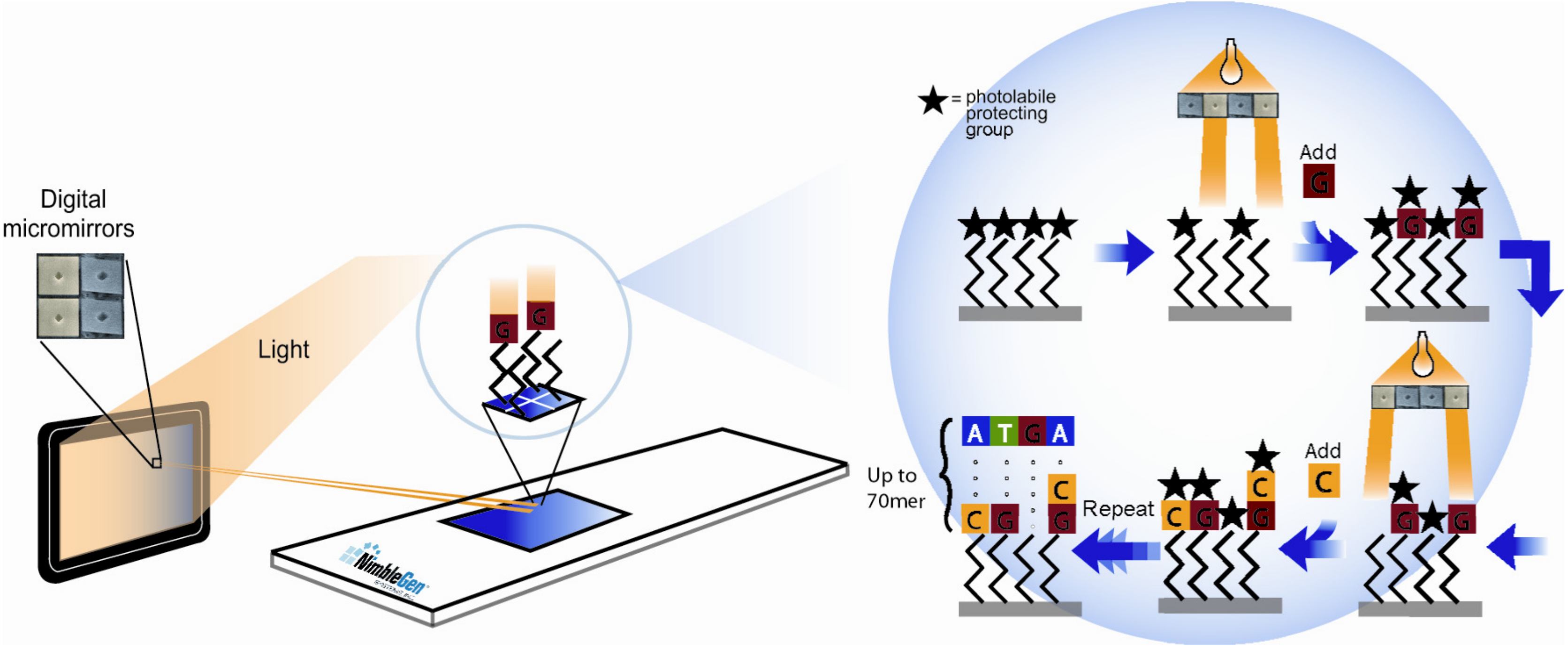
| library ID | total clone | known sequence | | | unknown sequences | |
|---|---|---|---|---|---|---|
| | | virus-like | | non-virus | virus-like | novel |
| | | plant | animal | | | |
| library 1 | 10576 | 9934 | 27 | 165 | 110 | 340 |
| library 2 | 13572 | 7106 | 701 | 3348 | 269 | 2148 |
| library 3 | 12621 | 8004 | 7 | 4351 | 11 | 248 |
| total | 36769 | 25044 | 735 | 7864 | 390 | 2736 |

# How do we then monitor microbes in the environment ?

- Detect and identify all animal-infecting viruses.
- Detect and identify selected bacteria (respiratory disease focus).
- Discover novel viruses
- Detect presence of co-infections of multiple pathogens
- Sensitive, accurate diagnosis based on *in silico* predictions of hybridization footprint

# Nimblegen Technology

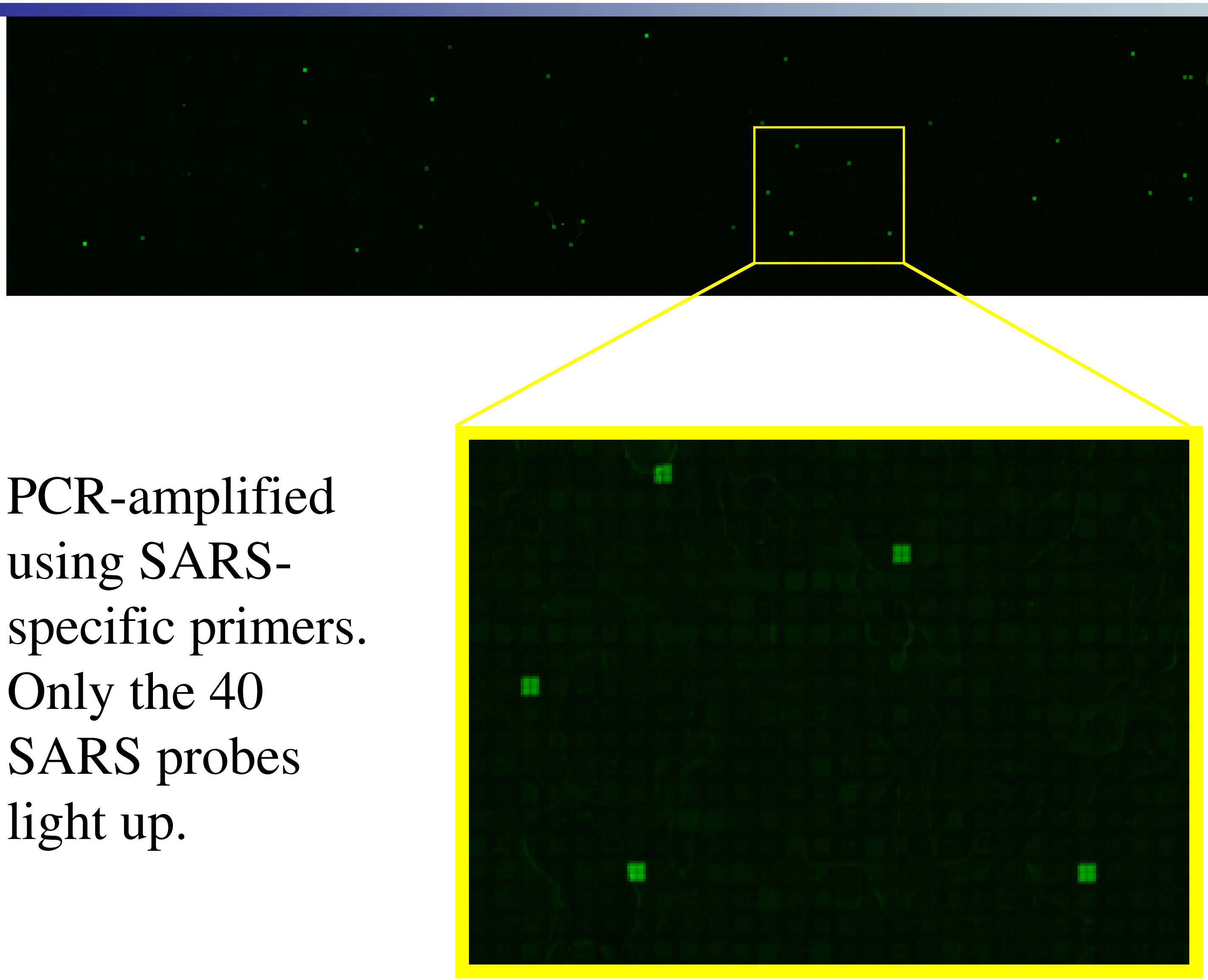
- ~380,000 probes
- Maskless photolithography



# The Pilot Chip

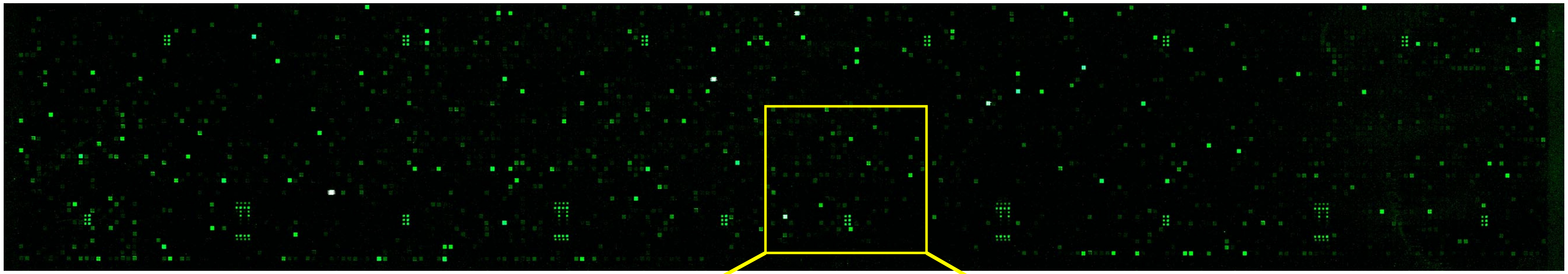
- **10,722 oligos covering 850 viruses**
- **204 oligos targeting human control genes (immune response)**
- **7-fold replication of all probes**

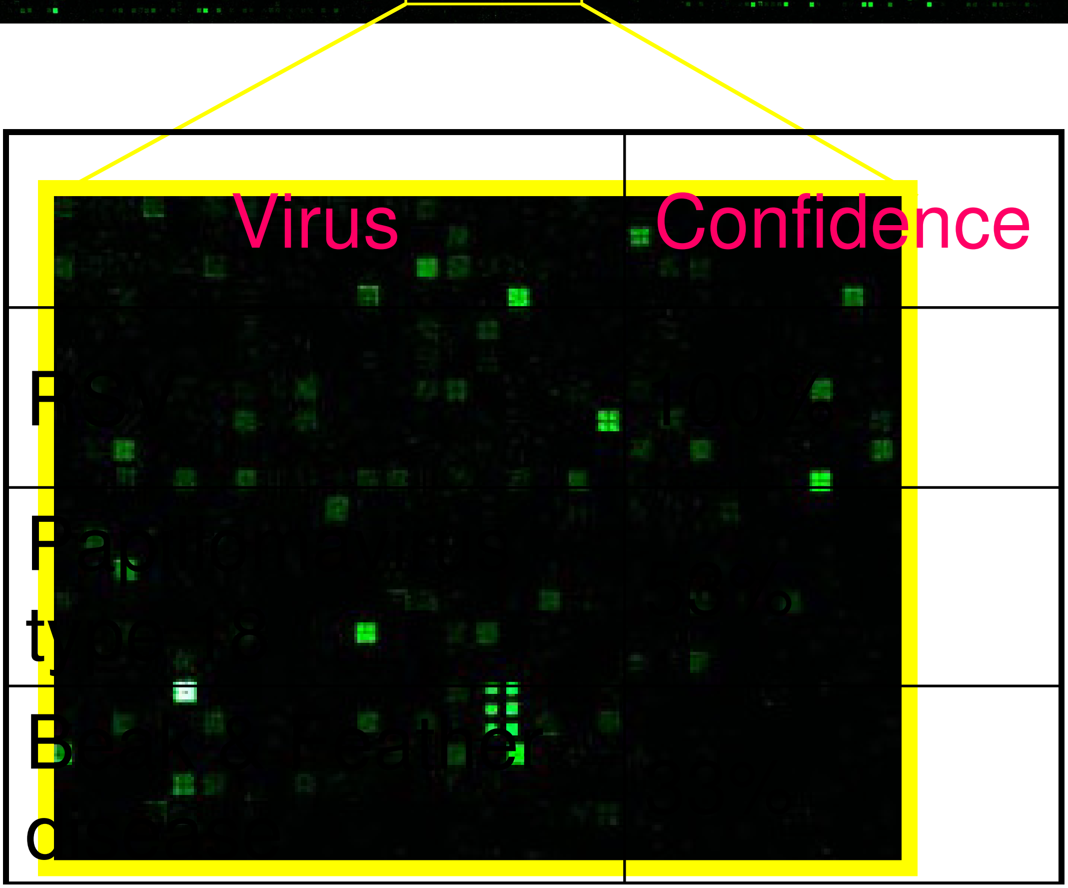
# SIN850 SARS cDNA



PCR-amplified using SARS-specific primers. Only the 40 SARS probes light up.

# RSV + mycoplasma cDNA (from ATCC)





PCR-amplified using random primers. Analysis of signal intensities identify presence of RSV with 100% confidence.

A*STAR Genome Institute of Singapore

- Block access to dangerous agents

- Monitor comprehensively
- Detect trafficking
- Detect outbreak
- Use advanced technologies
- Trace to source
- Integrated and coordinated information dissemination

A*STAR Genome Institute of Singapore

# Regional Biosecurity Workshop

## May 28, 2007

***Edison Liu***

*liue@gis.a-star.edu.sg*

***+65-6478-8007***