ORIGINAL ARTICLE

Whole-Genome Sequencing and Social-Network Analysis of a Tuberculosis Outbreak

Jennifer L. Gardy, Ph.D., James C. Johnston, M.D., Shannan J. Ho Sui, Ph.D.,
Victoria J. Cook, M.D., Lena Shah, M.Sc., Elizabeth Brodkin, M.D.,
Shirley Rempel, R.N., Richard Moore, Ph.D., Yongjun Zhao, D.V.M.,
Robert Holt, Ph.D., Richard Varhol, M.Sc., Inanc Birol, Ph.D., Marcus Lem, M.D.,
Meenu K. Sharma, Ph.D., Kevin Elwood, M.D., Steven J.M. Jones, Ph.D.,
Fiona S.L. Brinkman, Ph.D., Robert C. Brunham, M.D.,
and Patrick Tang, M.D., Ph.D.

jclub

07/20/2017

JL

Preface (Subjective)

- Not a technically rigorous/intense publication (not a single p-value)
- Interest in the integration of data

Mycobacterium Tuberculosis

- Important bacteria, spread through air, that infects lungs, brain, and spine, causing infectious disease, tuberculosis

- Possible symptoms: pain, cough, fever, weight loss, death!

- Smear positive and smear negative (main types)

- While it is increasingly uncommon in the 21st century, those who have TB suffer from long antibiotic courses, as well as a high potential for spreading the disease

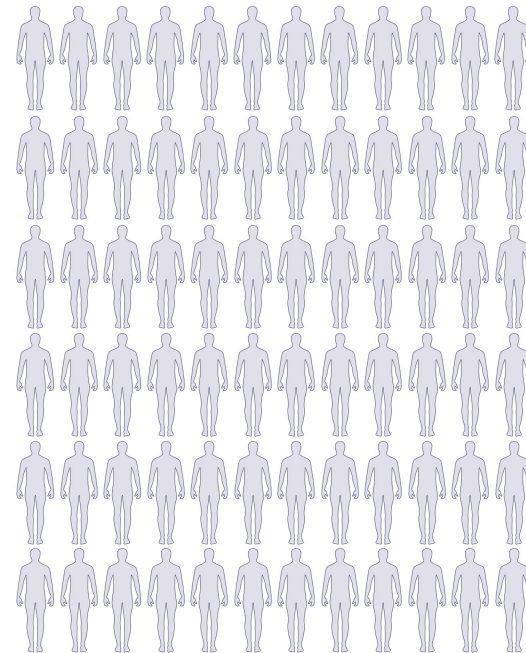- Control programs for TB exist, even in developed countries

Mycobacterium Tuberculosis
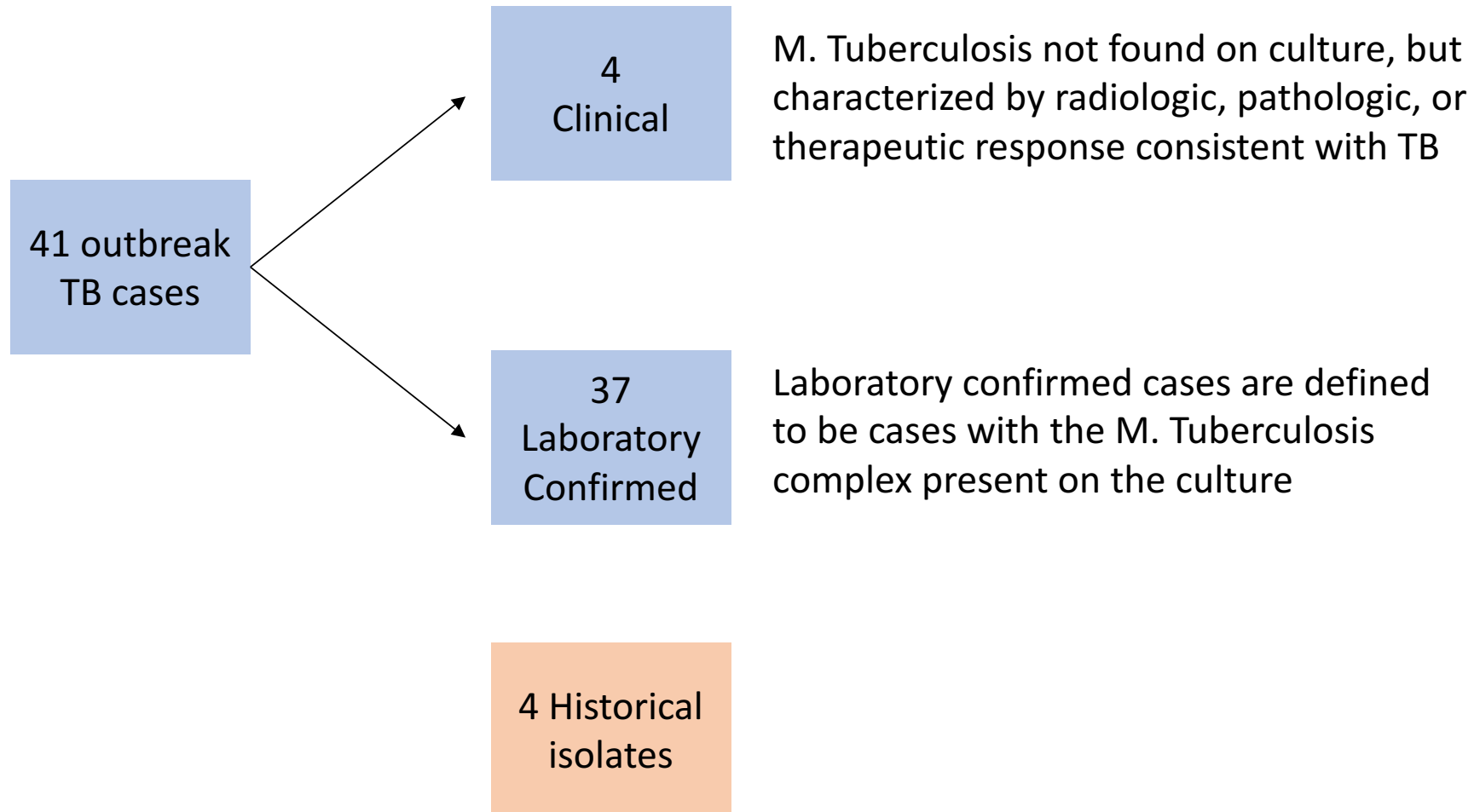# *Outbreak in British Columbia, Canada*

2007 National Average ᵀ
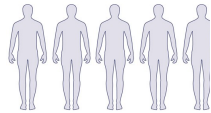
2007 British Columbia community outbreak ᵀ

Cases of Tuberculosis in one region of B.C. between May 2006 to December 2008.

41 outbreak TB cases

4 Clinical — M. Tuberculosis not found on culture, but characterized by radiologic, pathologic, or therapeutic response consistent with TB

37 Laboratory Confirmed — Laboratory confirmed cases are defined to be cases with the M. Tuberculosis complex present on the culture

4 Historical isolates

Dual approach to understanding the outbreak

Establishing a Community Network

- Contact Tracing
- Social-Network Questionnaire

Genomic Analysis from VNTR and WGS

- Hierarchical clustering
- Bayesian MCMC phylogenetic inference
- Maximum-likelihood phylogenetic inference

Social Network Approach

Contact Tracing

- Involves identifying the people who are infected and who they have come in contact with.

- Here, a bias is shown towards patient MT0009, a pediatric case, an unlikely source.

- Potentially, an incomplete network is present

MT0009
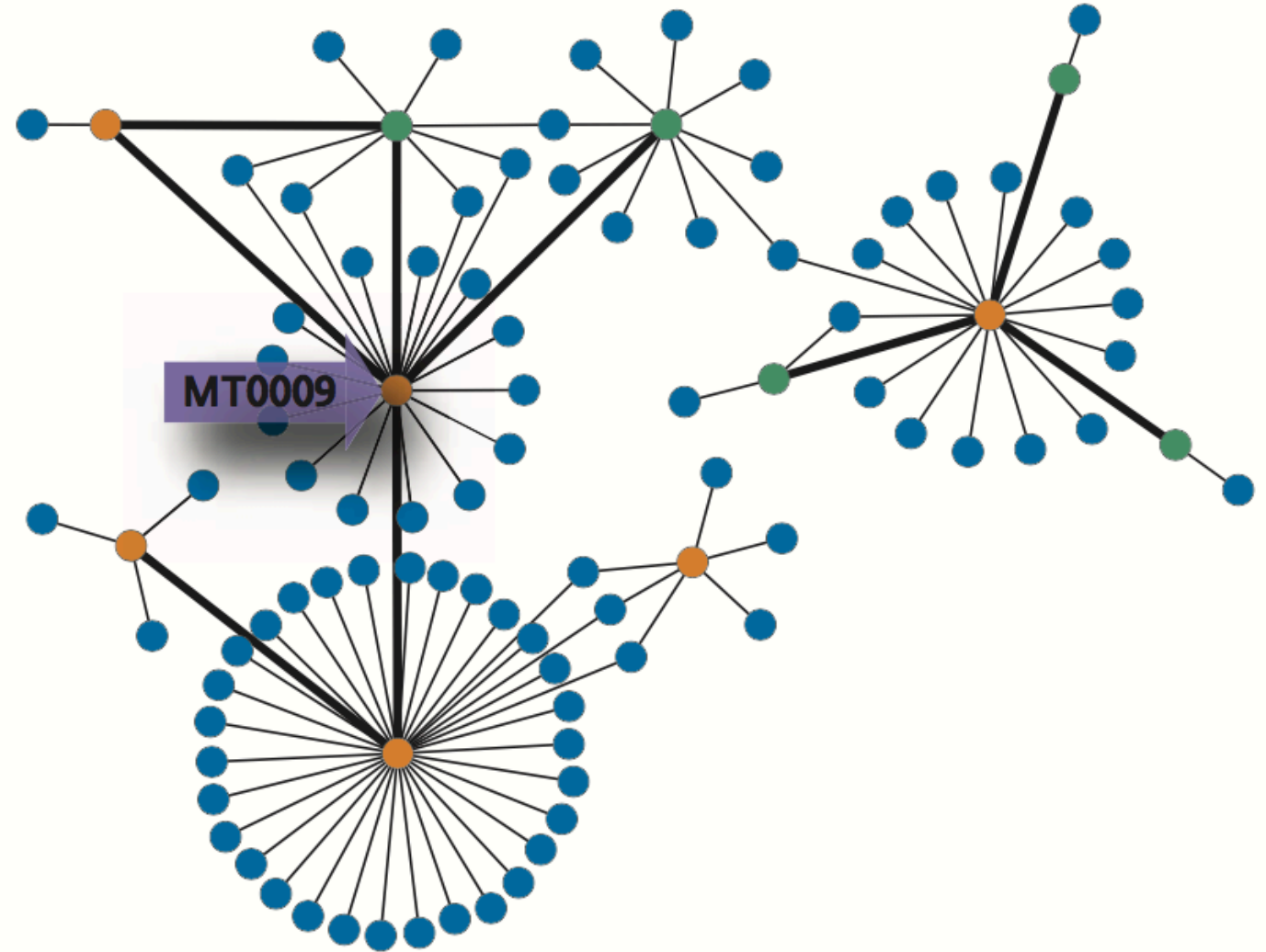
Social-Network Questionnaire

Example topics included:
- Medical History
- Places of aggregation
- Drug and Alcohol History

Example questions included:
- Who are your closest friends or the people you hung out with the most in the last year?
- Do you share drugs with them? (If yes, which drugs and how are they shared)
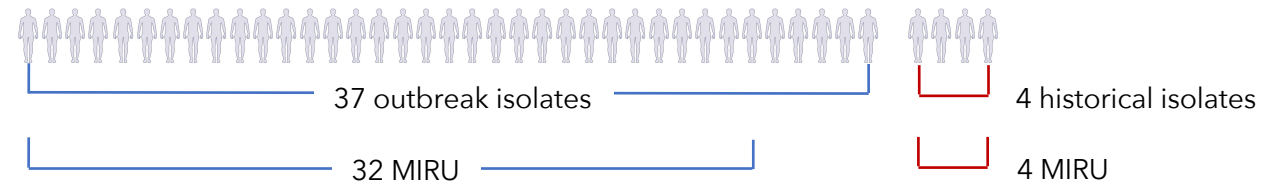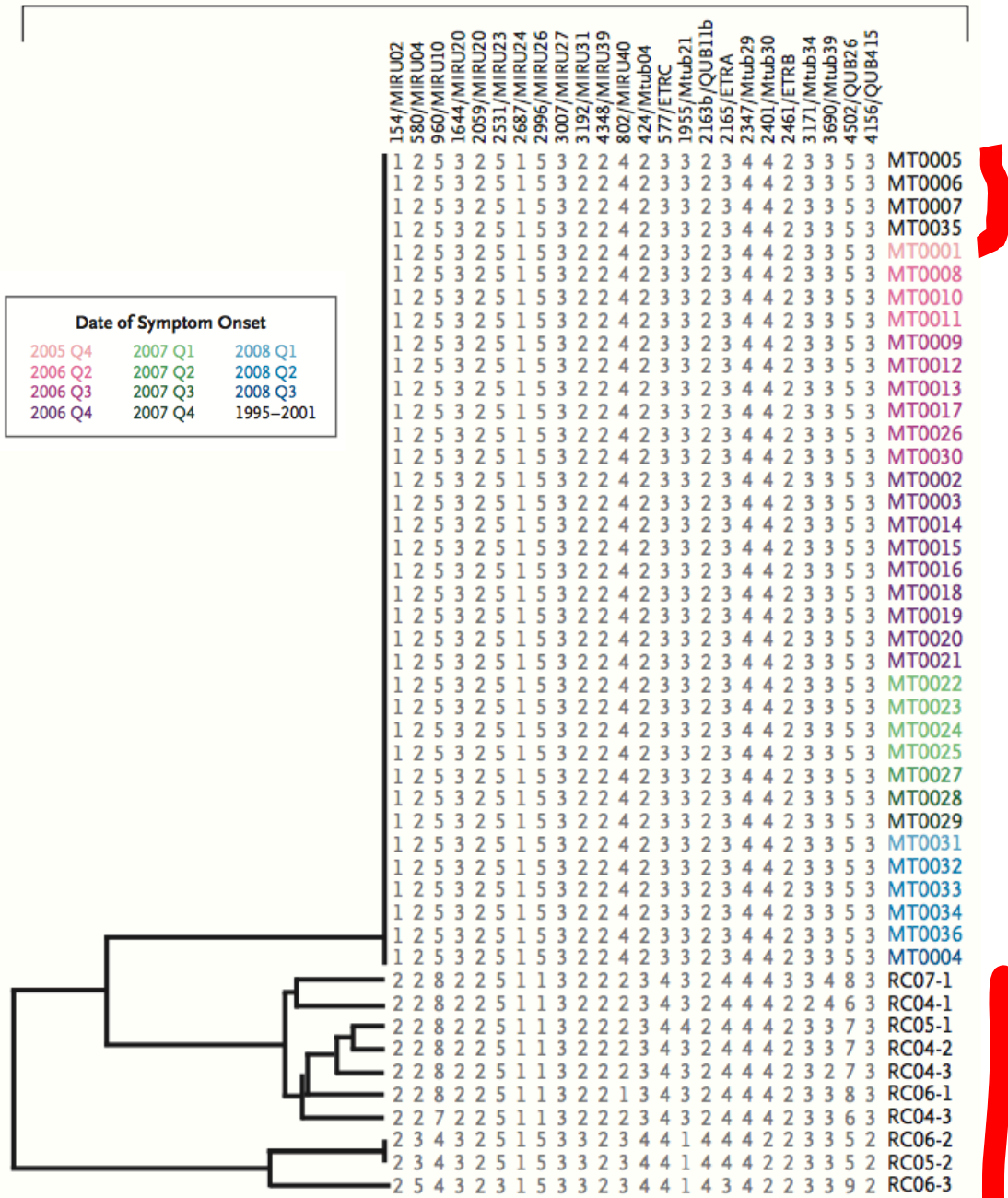- Are you infected with HIV?

Table 1. Characteristics of 41 Patients with Laboratory-Confirmed or Clinical Tuberculosis in a British Columbia Community from May 2006 through December 2008, According to Tuberculosis Lineage.*

Characteristic	Lineage A (N=13)	Lineage B (N=19)	Not Known (N=9)
Age — yr			
Median (range)	37 (21–55)	33 (1–71)	38 (4–52)
Mean (range)	36 (21–55)	34 (1–71)	32 (4–52)
Female sex — no. (%)	6 (46)	13 (68)	4 (44)
Canadian-born — no. (%)	13 (100)	19 (100)	9 (100)
Tuberculosis type — no. (%)			
Pulmonary	10 (77)	12 (63)	6 (67)
Pleural	3 (23)	4 (21)	3 (33)
Extrathoracic	0	1 (5)	0
Disseminated	0	2 (11)	0
Sputum smear positive for AFB — no. (%)	5 (38)	8 (42)	1 (11)
Culture-positive tuberculosis — no. (%)	13 (100)	19 (100)	5 (56)
Treatment outcome — no. (%)			
Death	1 (8)	3 (16)	0
Cure or treatment completed	11 (85)	16 (84)	8 (89)
Treatment not completed	0	0	1 (11)
Relapse	1 (8)	0	0
Risk factor — no. (%)			
Alcohol use	11 (85)	13 (68)	3 (33)
Crack cocaine use	7 (54)	12 (63)	6 (67)
Crack cocaine or alcohol use	11 (85)	16 (84)	7 (78)
HIV seropositivity	0	0	0
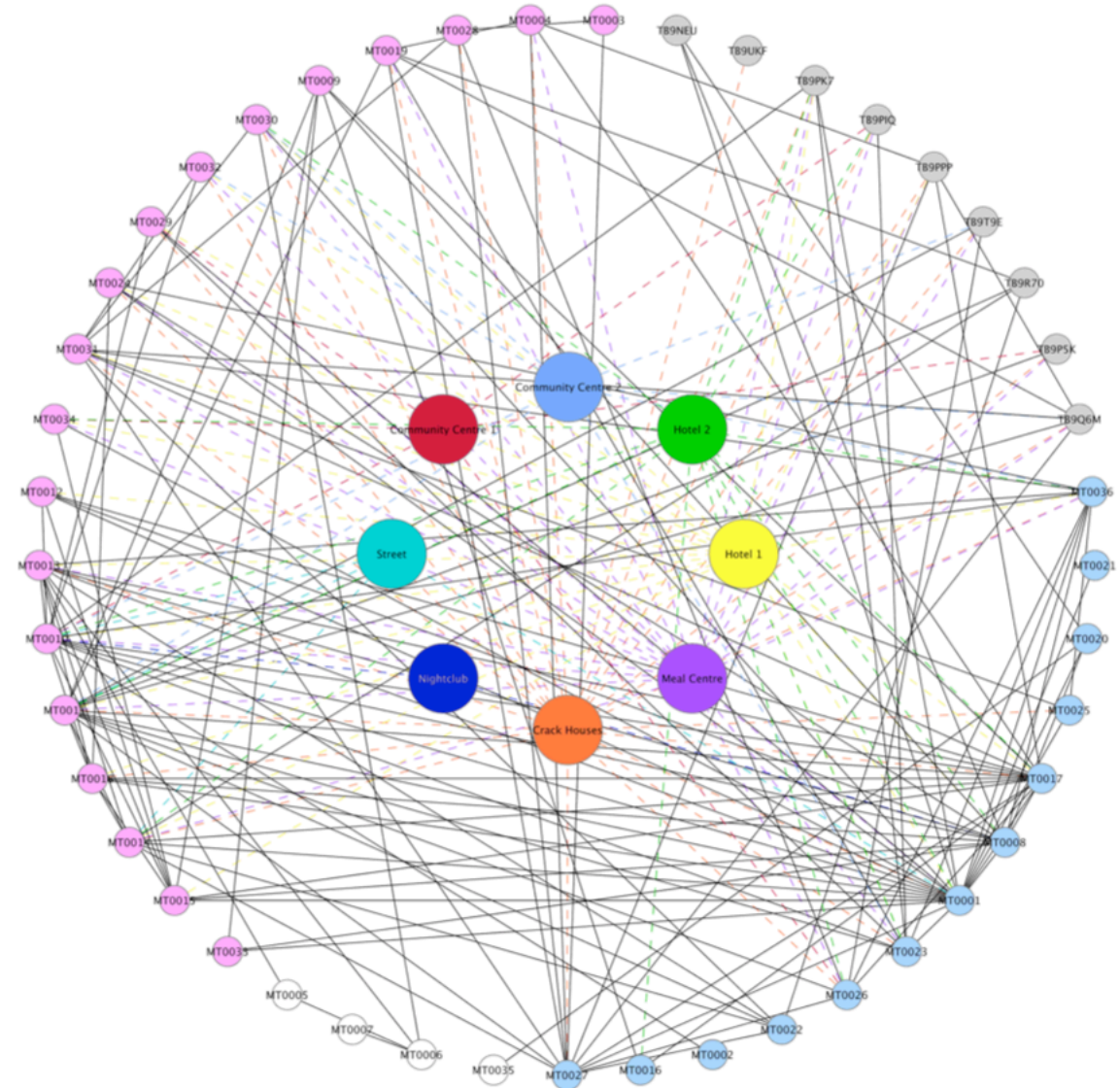
Genomics Approach

MIRU-VNTR Tree

Date of Symptom Onset

2005 Q4	2007 Q1	2008 Q1
2006 Q2	2007 Q2	2008 Q2
2006 Q3	2007 Q3	2008 Q3
2006 Q4	2007 Q4	1995–2001

37 outbreak isolates

4 historical isolates
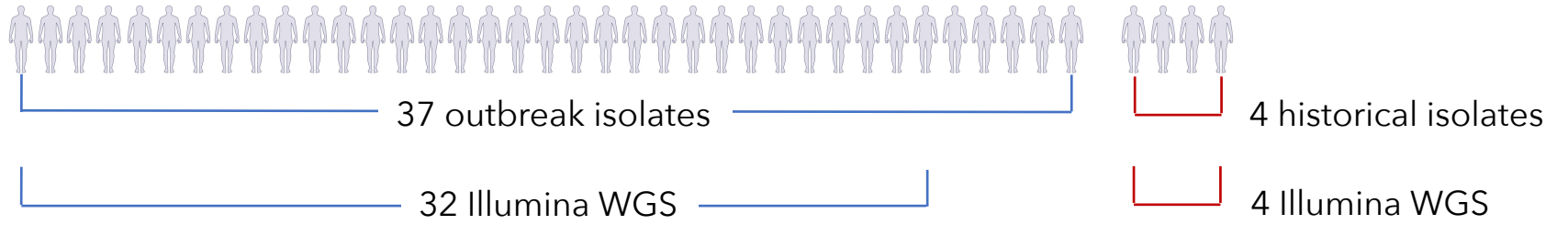
32 MIRU

4 MIRU

36 MIRU VNTR

Mycobacterium interspersed repetitive units (multiple locus VNTR)

Low resolution way to determine the outbreak as clonal
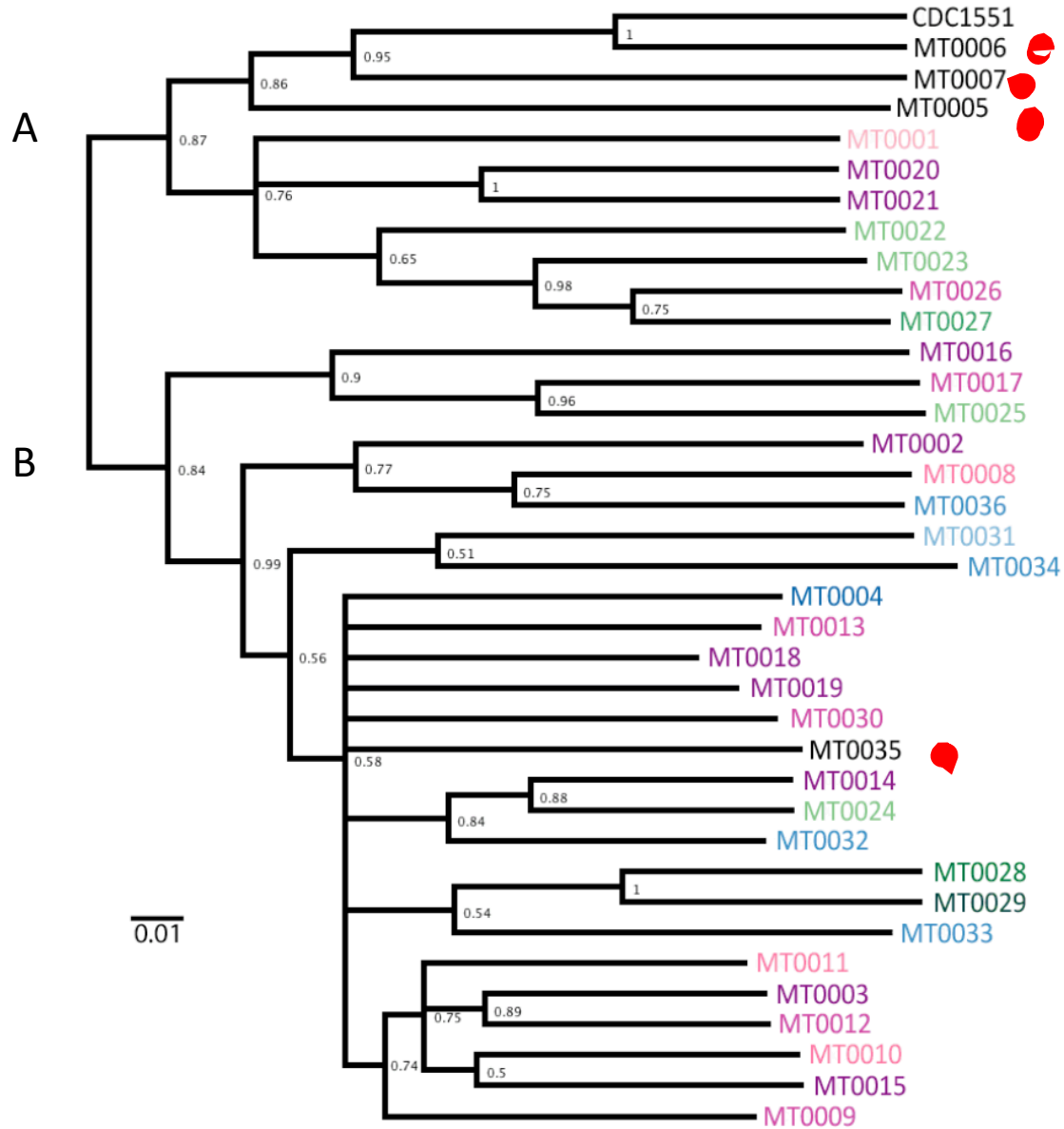
12

Whole-Genome Sequencing Tree

37 outbreak isolates

4 historical isolates

32 Illumina WGS

4 Illumina WGS

204 SNPs for 36 isolates

Hierarchical clustering of SNPs suggests two
simultaneous outbreaks
Lineage A, Lineage B
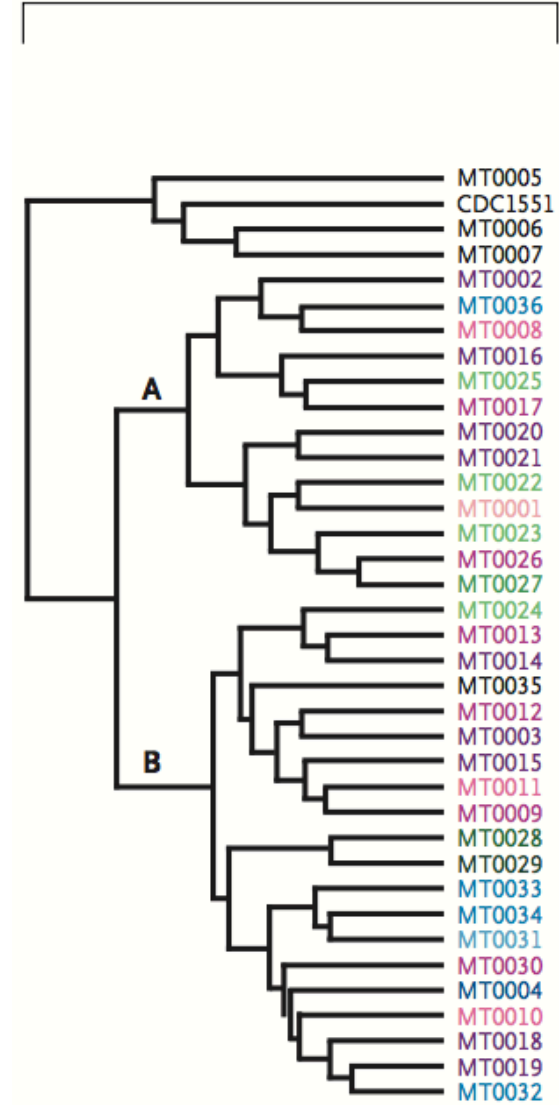
Date of Symptom Onset

2005 Q4	2007 Q1	2008 Q1
2006 Q2	2007 Q2	2008 Q2
2006 Q3	2007 Q3	2008 Q3
2006 Q4	2007 Q4	1995–2001

13

Bayesian MCMC Phylogenetic Inference

Whole-Genome Sequencing Tree

Date of Symptom Onset

2005 Q4	2007 Q1	2008 Q1
2006 Q2	2007 Q2	2008 Q2
2006 Q3	2007 Q3	2008 Q3
2006 Q4	2007 Q4	1995–2001

14

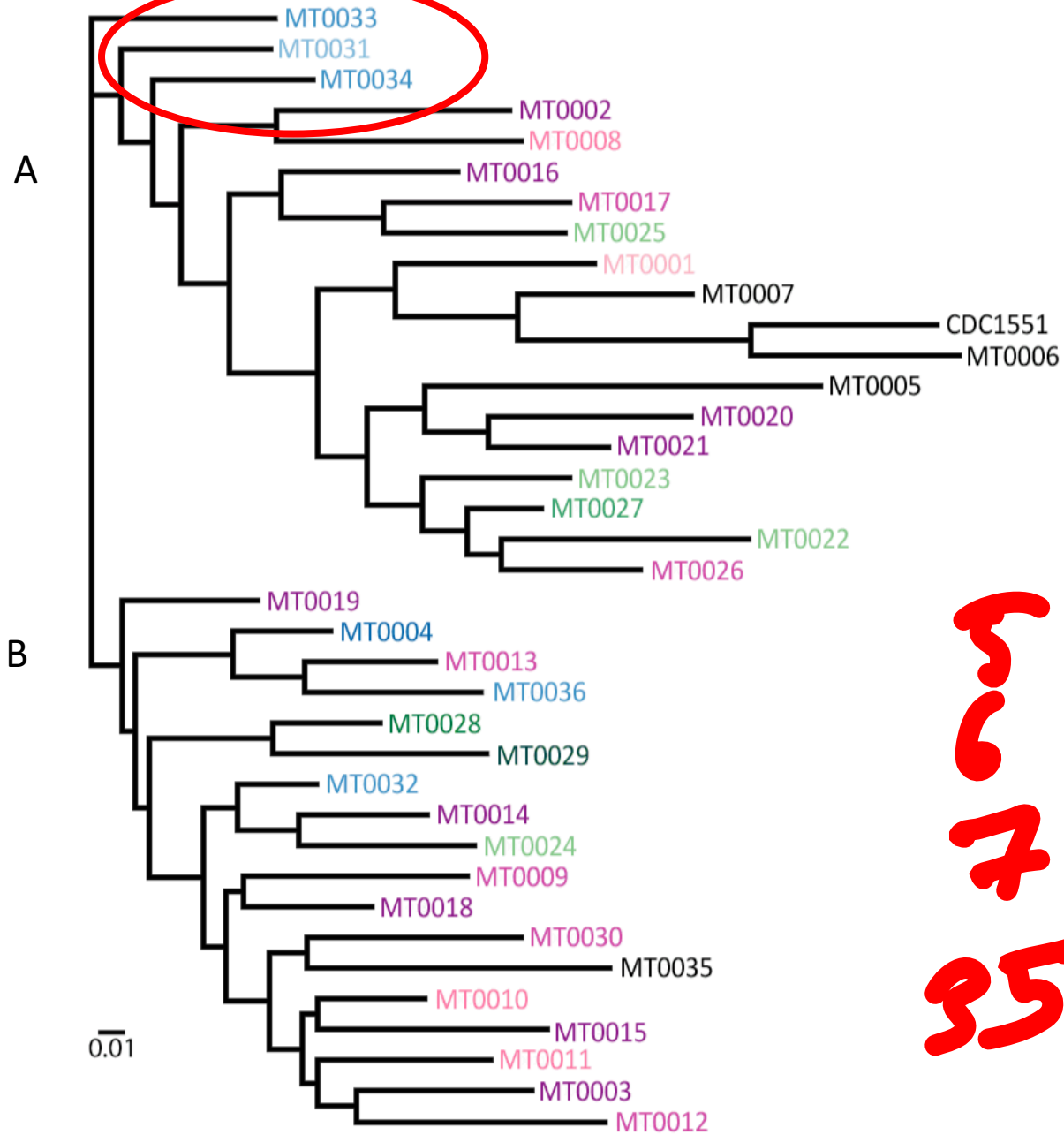# Maximum Likelihood Phylogenetic Inference

Maximum Likelihood Phylogenetic Inference

Whole-Genome Sequencing Tree
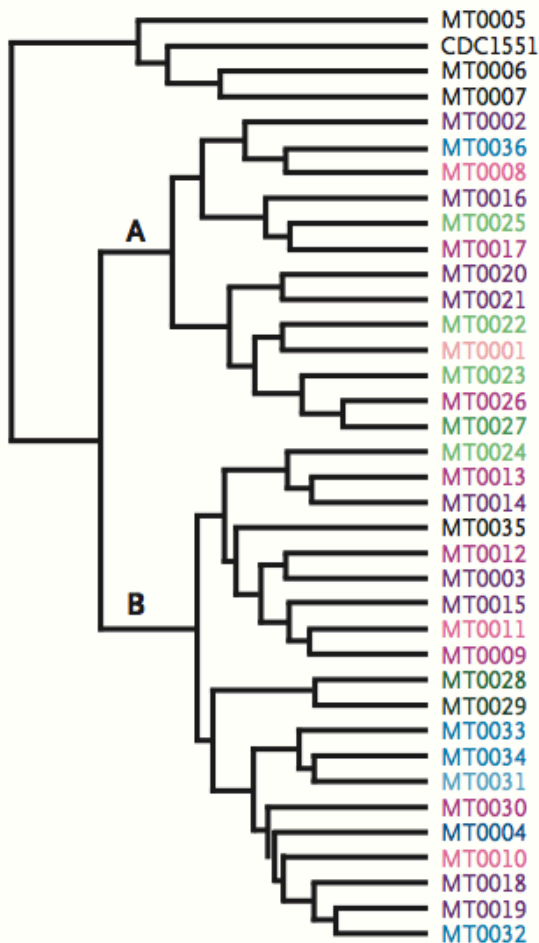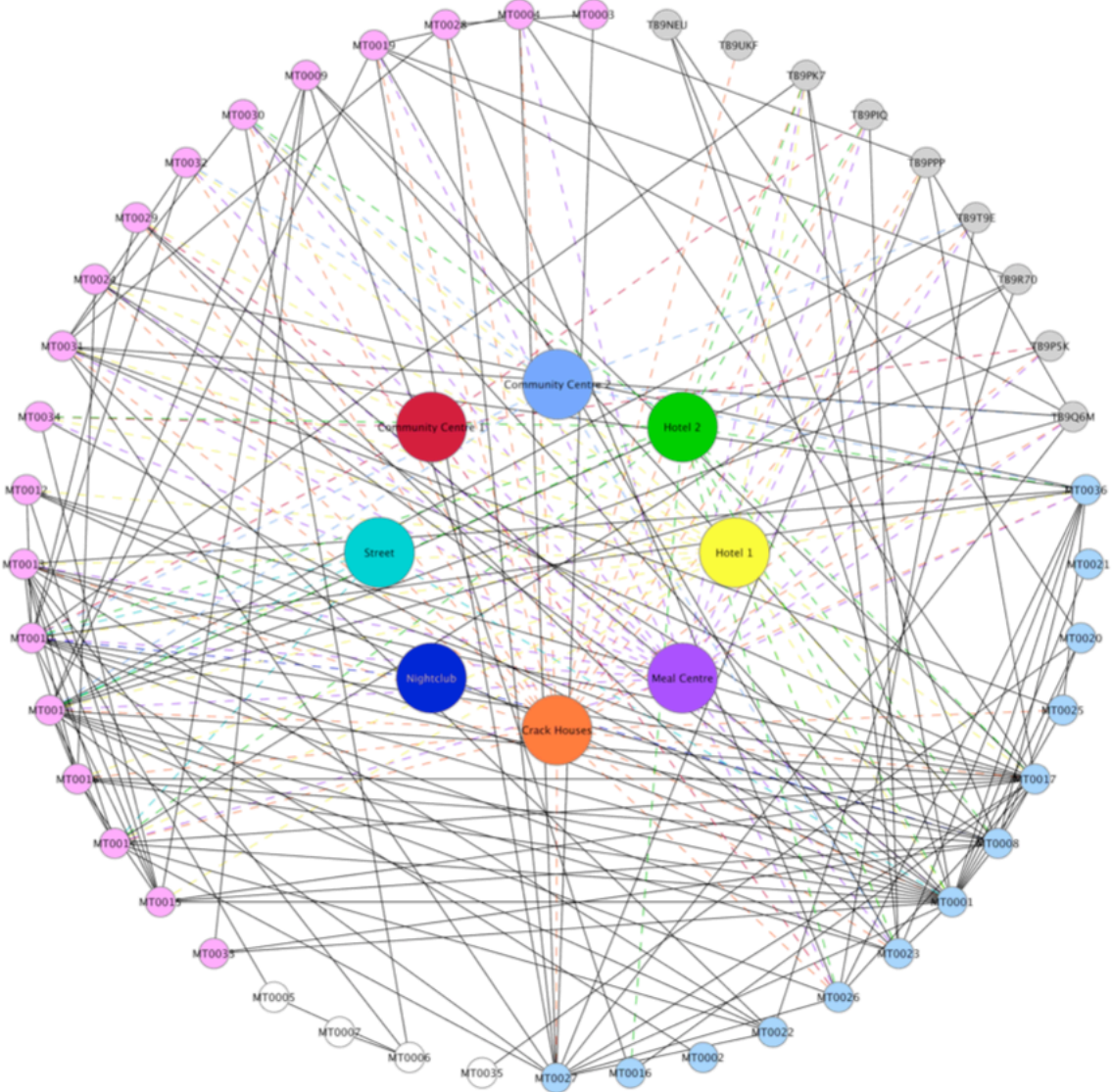
Overlay the social network on to the phylogenetic tree
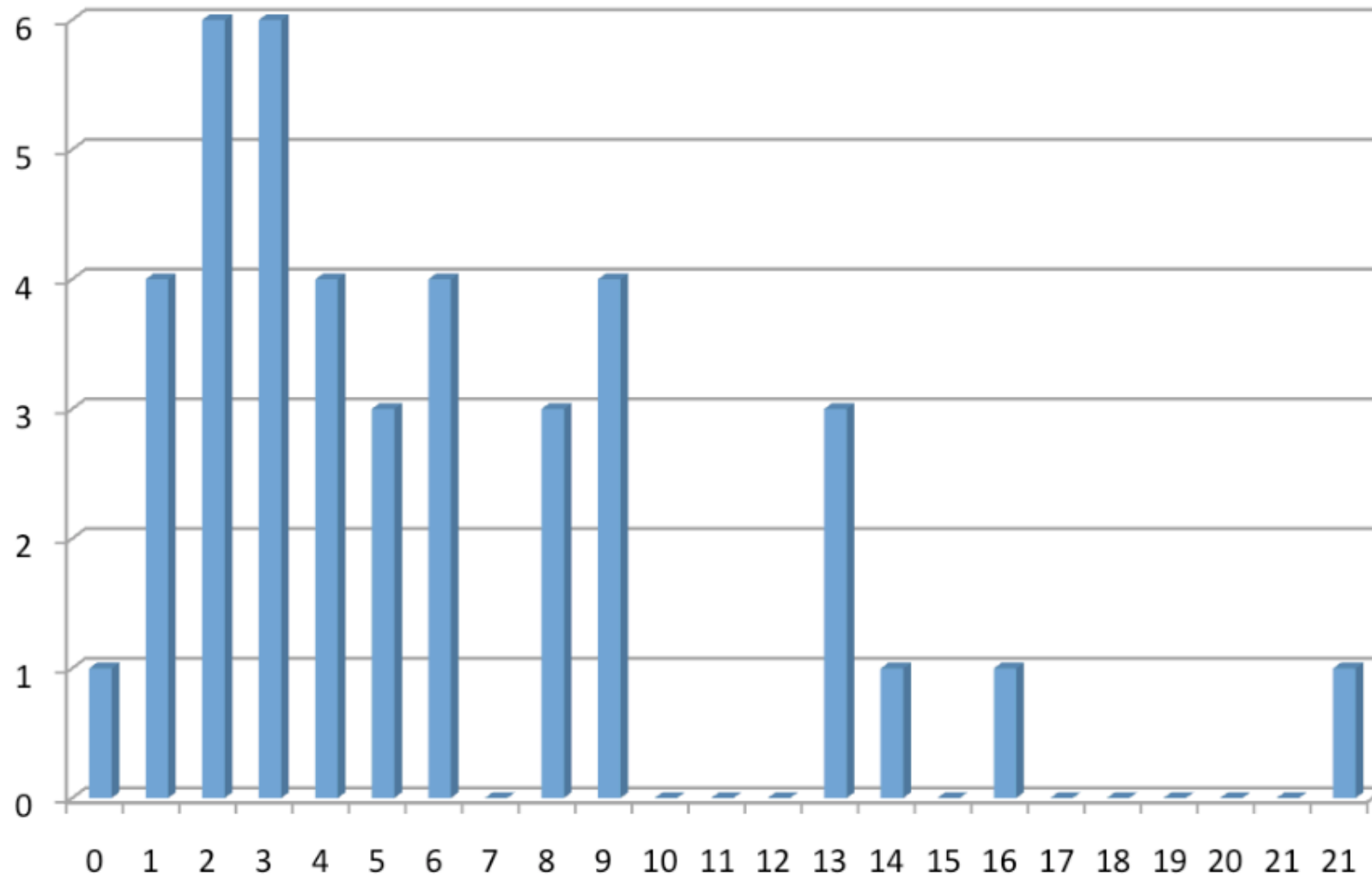
Date of Symptom Onset

2005 Q4   2007 Q1   2008 Q1
2006 Q2   2007 Q2   2008 Q2
2006 Q3   2007 Q3   2008 Q3
2006 Q4   2007 Q4   1995–2001

Transitivity of the Social Network. The degree of an individual case – the number of direct social contacts with other cases – is shown along the x axis, while the number of cases with that degree is plotted on the y axis.

Transmission Rules

0*) Transmission can only occur within a lineage (A or B)
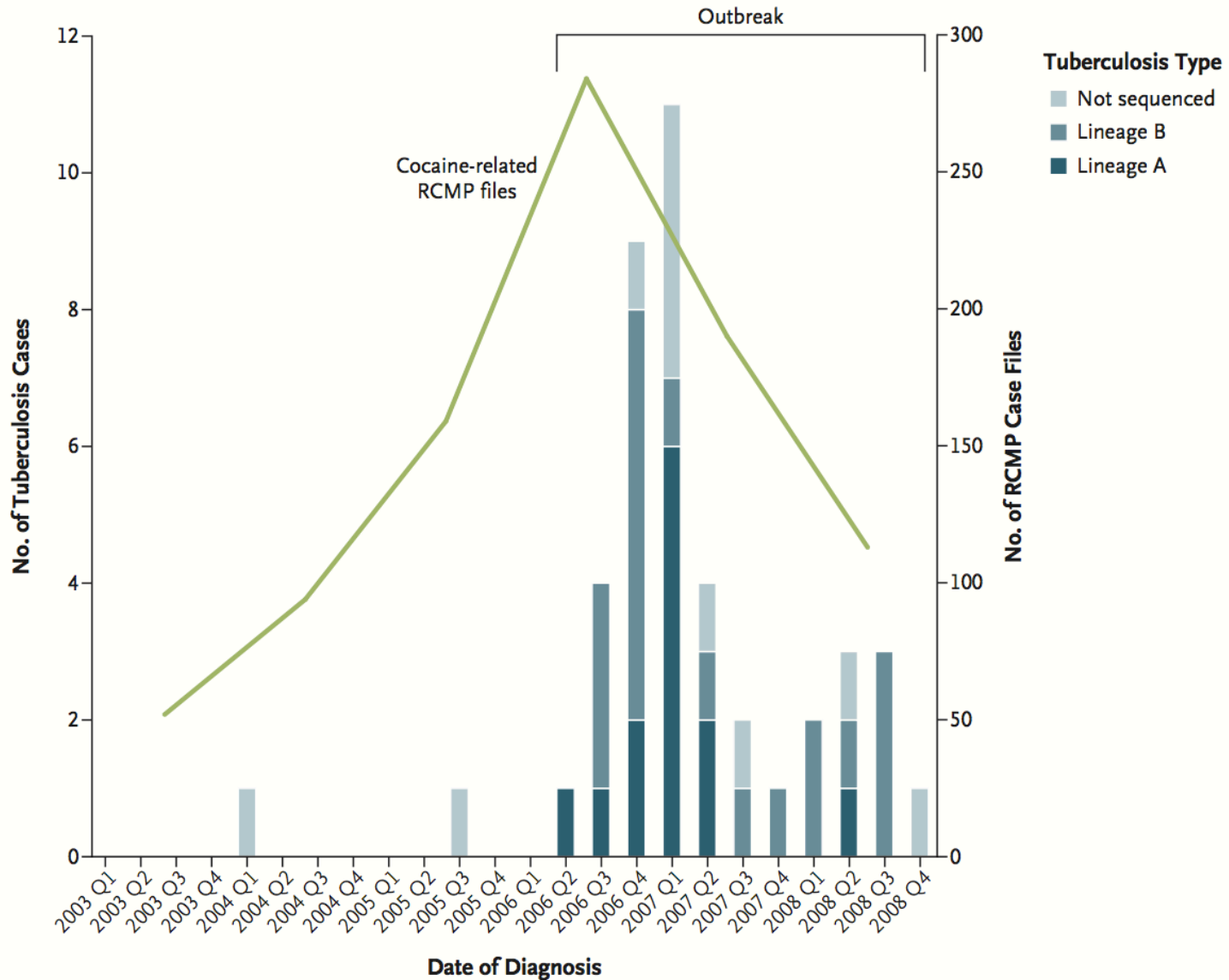1) Transmission moves forward in time
2) A person with only a single previously infected contact must have been infected by that person
3) In cases where a person has multiple infected, upstream contacts, smear-positive contacts are more likely to transmit than smear-negative contacts

Whole-Genome Sequencing and Social-Network Analysis
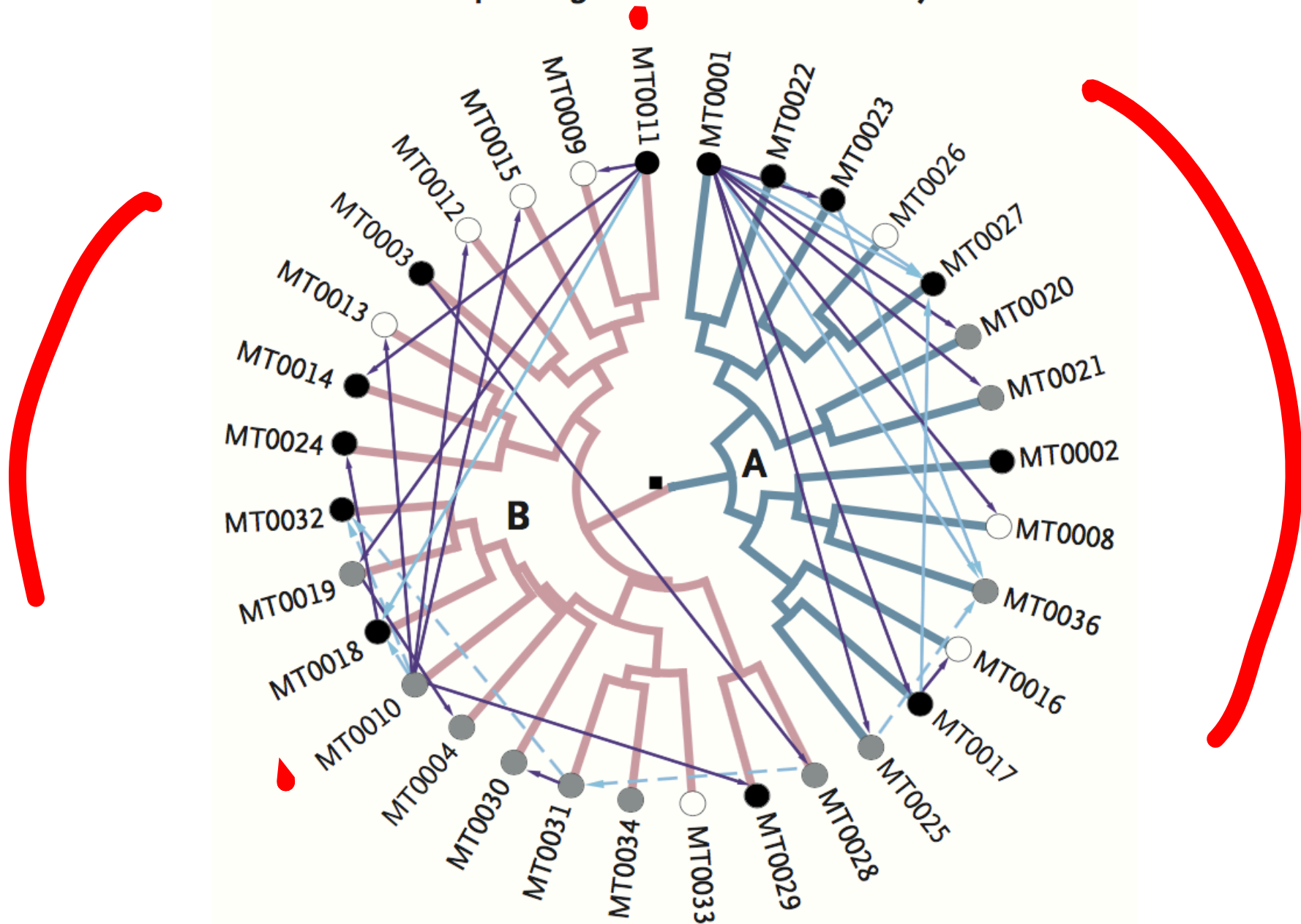
The True (Unproven) Cause?

The simultaneous reappearance of two extant lineages suggests that a social or environmental factor, not a genetic change in the organism, most likely triggered the outbreak.

In conclusion

- Low resolution VNTR suggests a clonal outbreak
- High resolution Illumina sequencing suggests lineages A, B
- Social network questionnaire allows for creation of network between isolates
- Integration of social network and phylogeny allows for transmission network
- The main 'selling' point (indicated by the authors) of using WGS is to allow for a more detailed approach in analyzing the social network (i.e. separation into two lineages)