Genomic approaches to the epidemiology and diagnostics of plant pathogens: tracking the global spread of agrobacteria

Alexandra J Weisberg

Postdoctoral Researcher, Chang and Grunwald labs
Department of Botany and Plant Pathology
Oregon State University, USA

crown gall on rose

Hairy root disease

graft union

cuttings

Nurseries and Orchards

Epidemiology, diagnostics, and the big questions

- How has modern genomics (DNA sequencing) revolutionized our ability to diagnose pathogens and track their movement?
  - Can we use these tools to better understand and diagnose disease?

- 2020: International year of plant health

- Epidemiology *-also timely*
  - How do pathogens spread?
  - What factors lead to disease prevalence?
  - What factors reduce incidence of disease?

Modern science is an interdisciplinary and collaborative effort

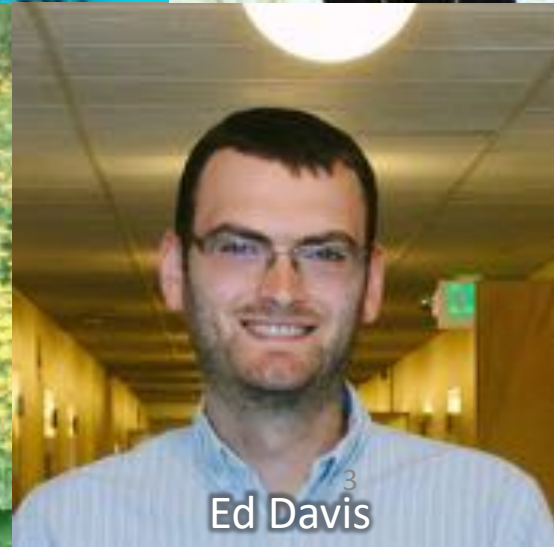Jeff Chang, Melodie Putnam, Nik Grunwald, Walt Ream, Joyce Loper, Marilyn Miller, Javier Tabima, Michael Belcher, Chih-Horng Kuo, Erh-Min Lai, Ed Davis
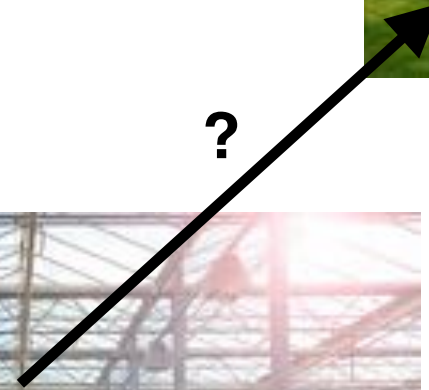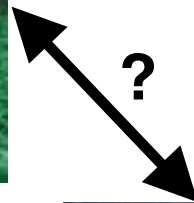
How do plant associated bacteria spread? How do mobile genetic elements affect plant/microbe symbiosis or pathogenesis?

Are strains shared between sites?
Are there multiple infection genotypes?
How long do genotypes persist?
Does horizontal gene transfer/plasmid conjugation play a role in infection?

How does this impact disease?
Other traits are beneficial
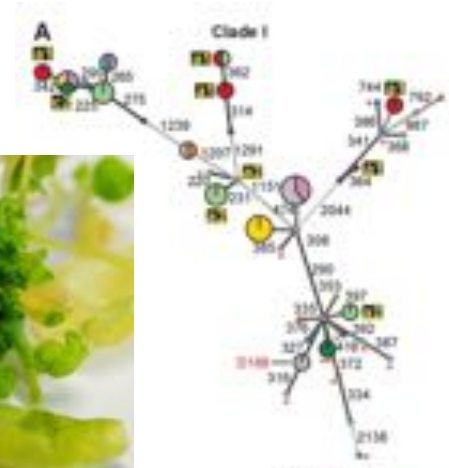
What role does human activity play?

Agrobacterium tumefaciens

Crown gall disease
Evolution of Ti plasmids
Global transmission patterns
Weisberg et al., 2020 Science

Rhodococcus fascians

Leafy gall disease
Tracking spread of _Rhodococcus_
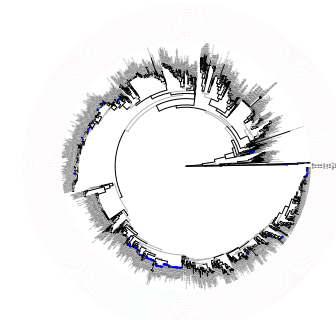**in nurseries**
Savory et al., 2017 eLife

Bradyrhizobium
Nitrogen fixation symbiosis
Characterizing ineffective "cheater" strains
Collaboration with Joel Sachs, UC Riverside

Streptomyces scabiei

Potato common scab disease
Collaboration with Chris Clarke, USDA

What we are interested in (and an outline)

- Evolution of *Agrobacterium tumefaciens*

- Evolution and emergence of Ti plasmids

- How the agro genotype/plasmid combination affects degree of pathogenicity and host range

  - Look to nature, what combinations work with what plants?
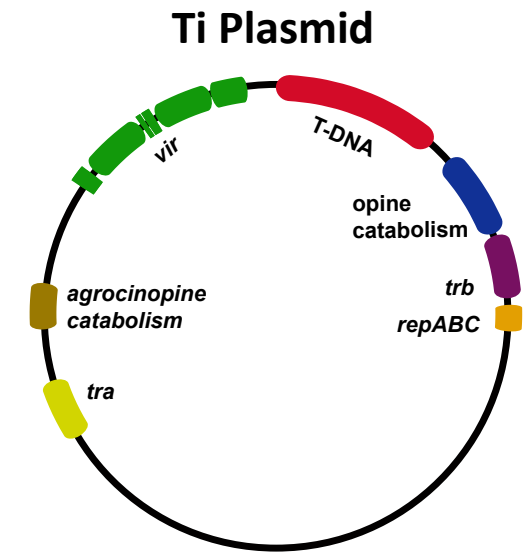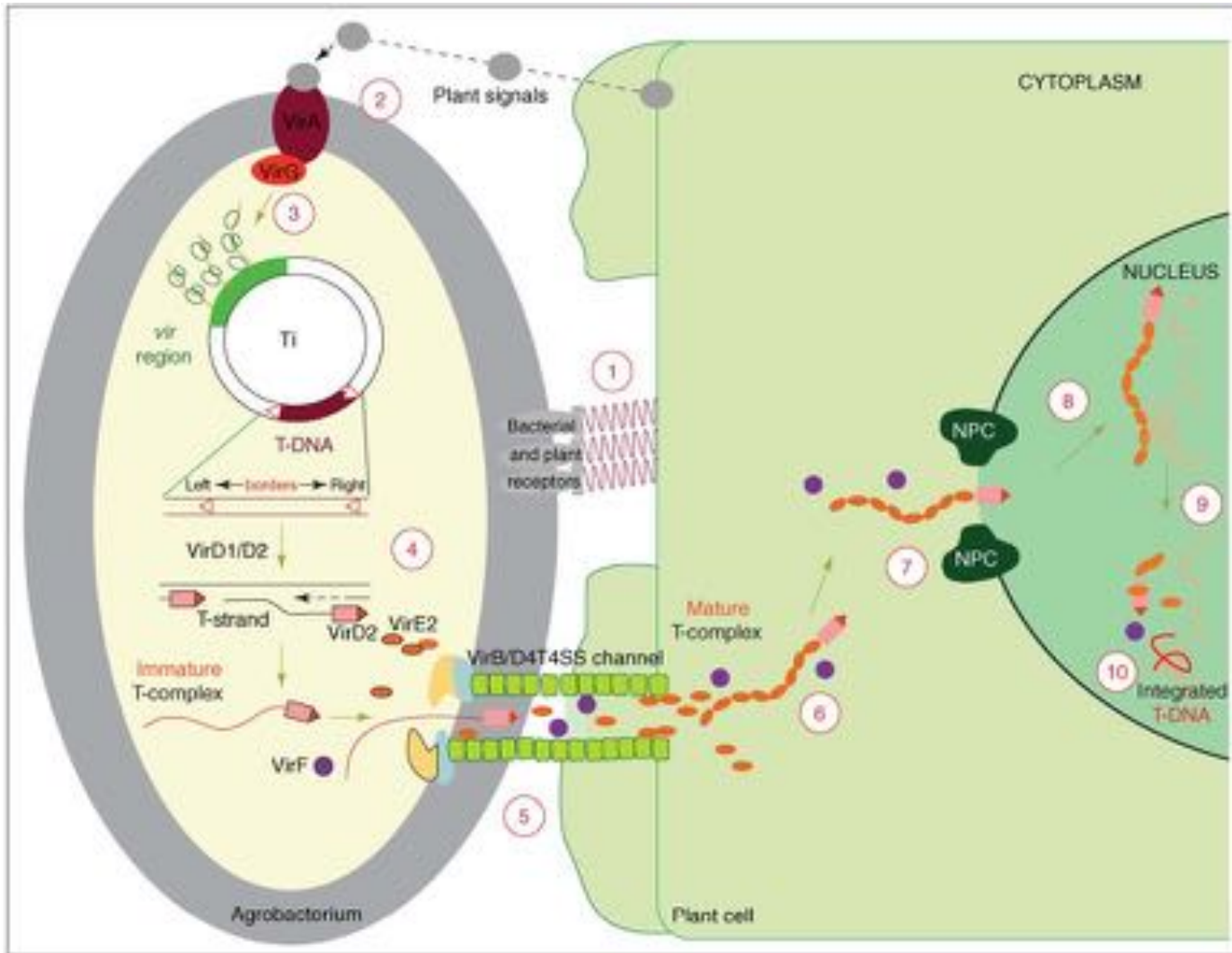
- The spread of *Agrobacterium*

Crown gall disease is a major problem for US agriculture

- masses of undifferentiated tissue
- millions $ in damages each year to nurseries, orchards, vineyards
- wide range of plant hosts
- Graft failure, reduced fruit yield/quality
- Constriction of the vascular tissue
- not always fatal but disfiguring for nursery plants
- no cure, plants must be destroyed

crown gall on rose

grapevine

Hairy root disease

graft union

cuttings

Nurseries and Orchards

Agrobacterium tumefaciens genetically modifies plants and causes crown gall disease
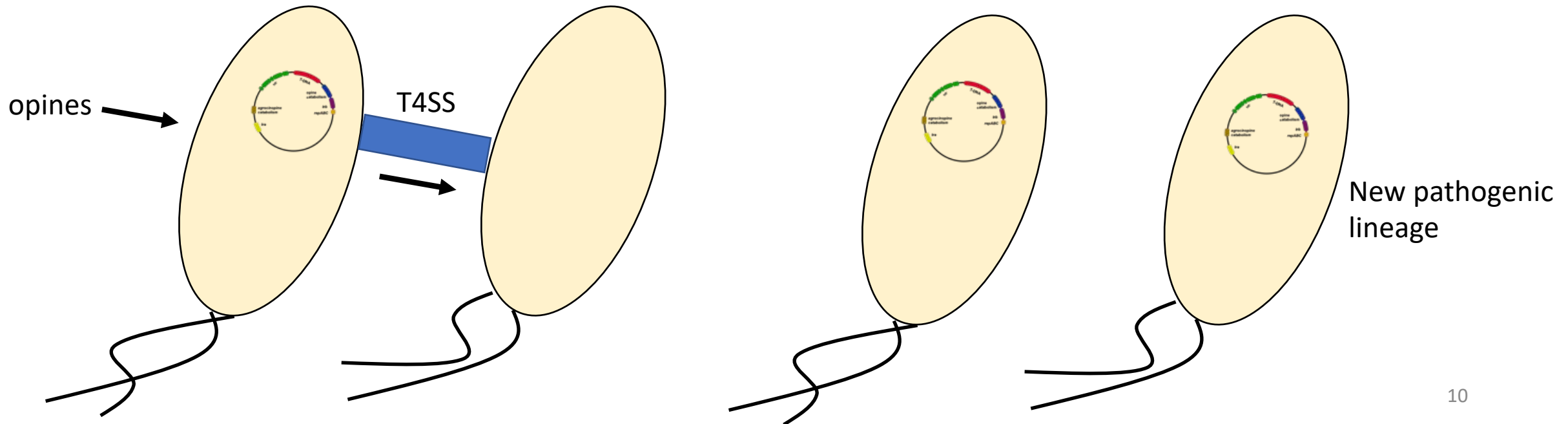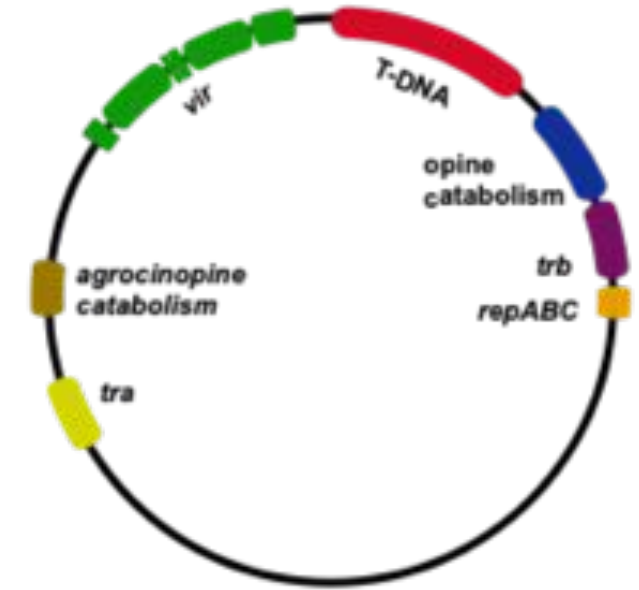

Tzfira and Citovsky, 2006

What's on a Ti plasmid?

- Essential elements:
  - T-DNA - oncogenes and opine synthase(s)
  - *virABCDE* operons
  - opine catabolism
  - plasmid replication (*repABC*) and conjugation (*tra/trb*)
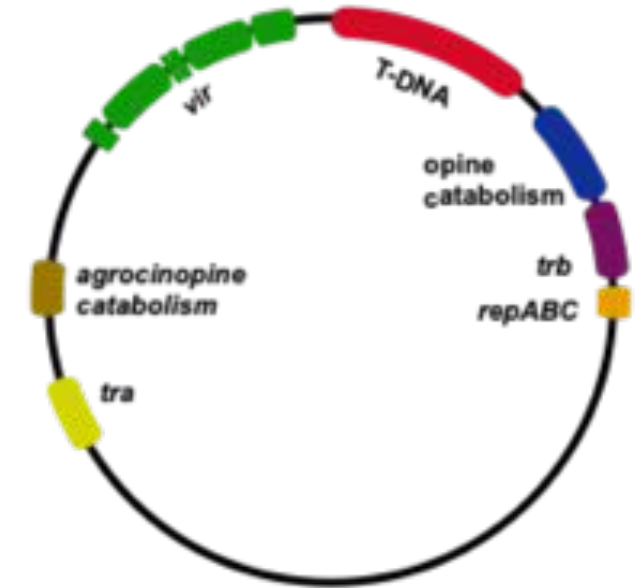
What's on a Ti plasmid?

- Essential elements:
  - T-DNA - oncogenes and opine synthase(s)
  - *virABCDE* operons
  - opine catabolism
  - plasmid replication (*repABC*) and conjugation (*tra/trb*)

What are inherent challenges to epidemiology studies in agriculture?

- Non pathogens have many features in common to pathogens
  - Ex. T3SS in commensals
- Genetic diversity of taxonomic groups is very high
  - Have to account for that in comparisons
- *Agrobacterium* - 4 polyphyletic lineages/biovars
  - More than 20 species-level groups (ANI)
- A conjugative plasmid confers virulence
  - Plasmids spread horizontally and create new pathogenic lineages
  - Tools for epidemiology are not applicable to plasmids
  - Virulence plasmids are genetically diverse –change rapidly

Source: news-medical.net

Experimental design

1. Ecological sampling

Melodie Putnam
OSU Plant Clinic

Larry Moore
Culture Collection

3. *De novo* **genome assemblies, genome analyses, chromosome (vertical) vs MGE (horizontal), phylogenetic trees, network analyses**
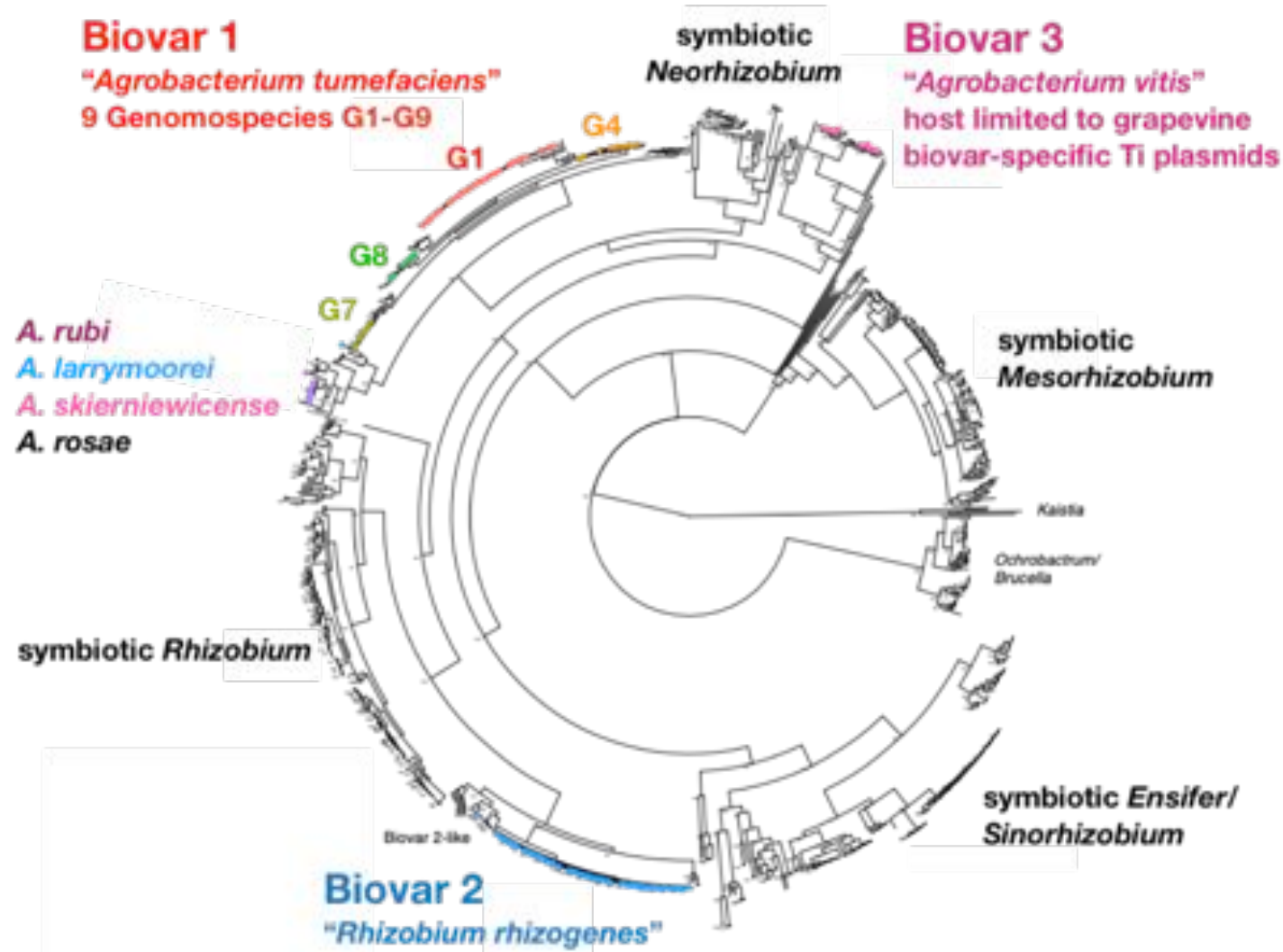
12

Generating a diverse dataset to provide a framework for comparison

- >150 diverse *Agrobacterium* strains

- 1927-2016

- >50 plant host species

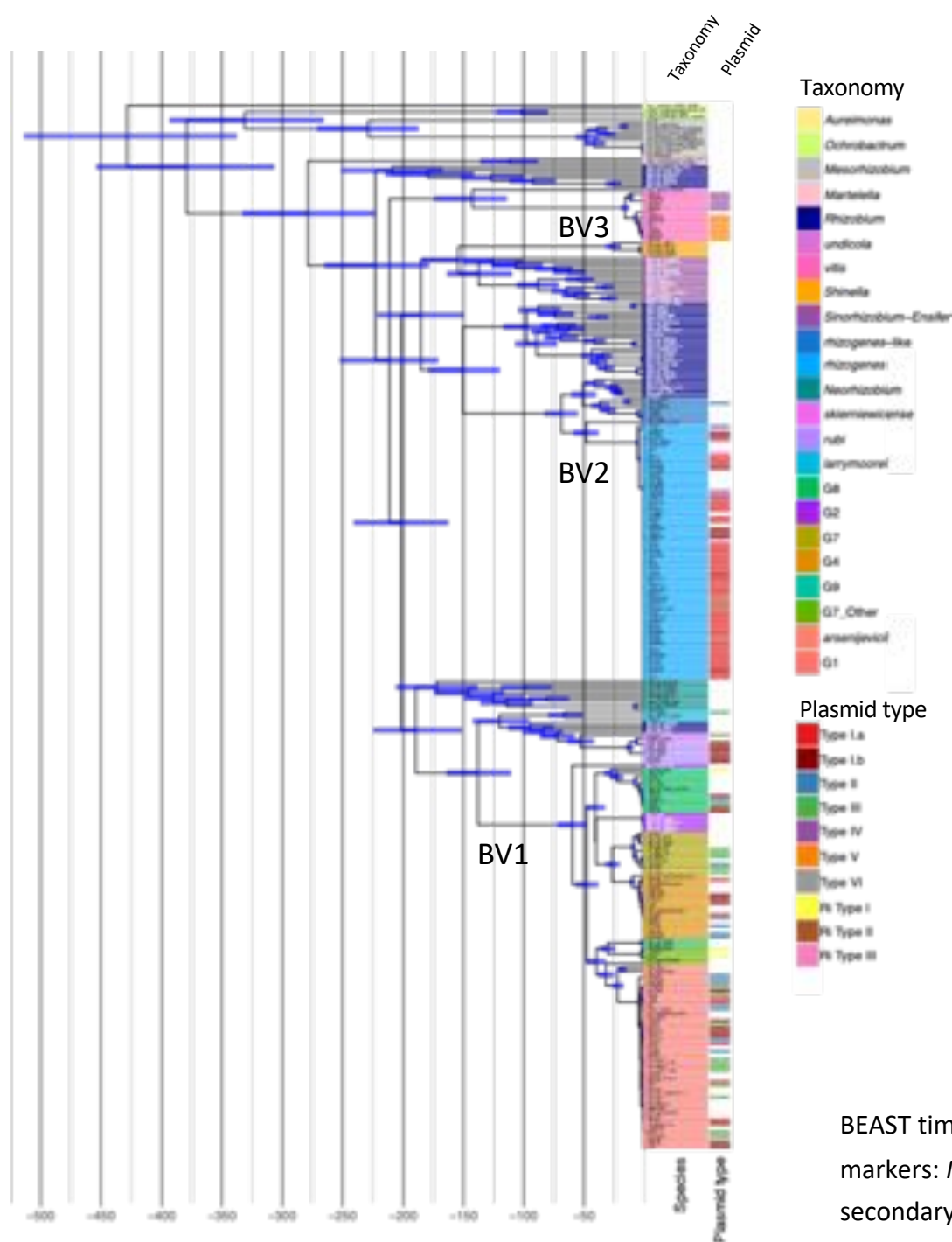- Also >60 contemporaneous strains from nurseries

Agrobacteria comprises three polyphyletic clades interspersed with Rhizobium

MLSA phylogeny of 24 housekeeping genes (AutoMLSA)

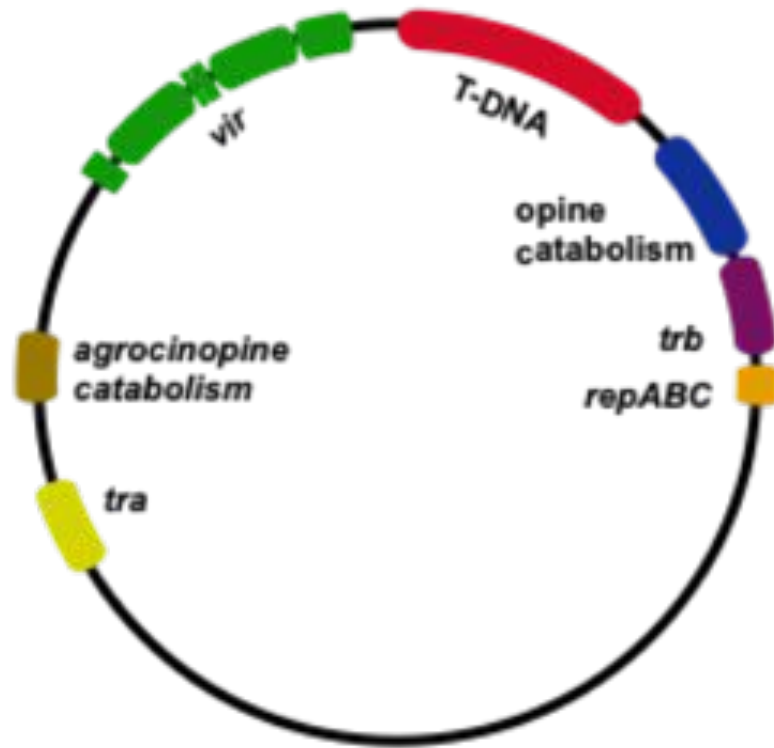Whole genome-based approaches (ANI, POCP) are consistent with these groupings

Biovar 1 diversified earliest, biovars 2 and 3 are recent

Biovar 1: ~59 Mya
Biovar 2: ~1.6 Mya (bottleneck)
Biovar 3: ~16 Mya

BEAST time tree, 8 neutrally evolving genes

markers: *Mesorhizobium* 250 Mya, *Sinorhizobium* 150 Mya

secondary markers from Chikri-Adeeb & Chikri 2016

Ti/Ri Plasmids drive the phenotype

Ti plasmids are diverse, highly recombinogenic

- Essential elements:
  - T-DNA - oncogenes and opine synthase(s)
  - *virABCDE* operons
  - opine catabolism
  - plasmid replication and conjugation

- 5 available sequenced plasmids are all different
- Hypothesized plasmids evolved by recombination:
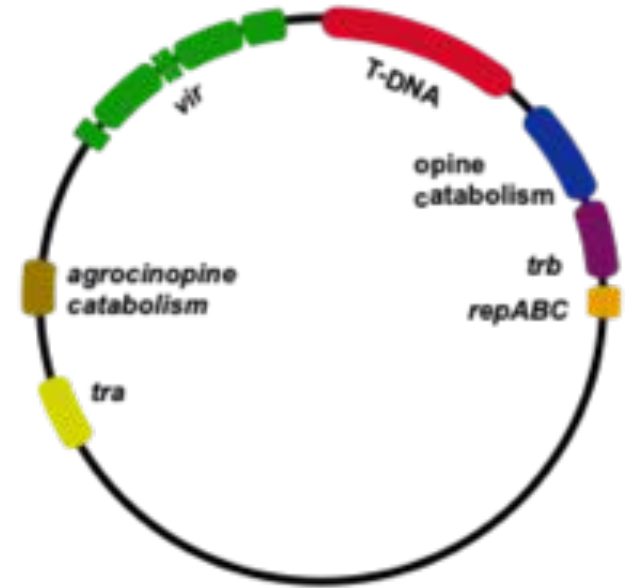
"Frequent horizontal gene transfer makes the construction of phylogenetic trees for ... pTi/pRi plasmids ... practically impossible. Even if thousands of Agrobacterium genomes were available, it might still be impossible to establish phylogenetic trees" (Chen and Otten 2017)
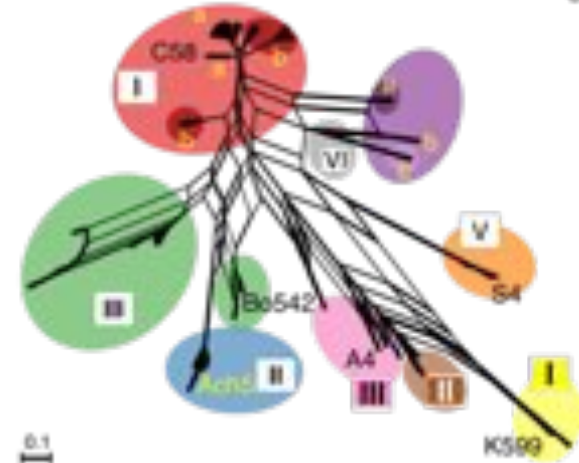
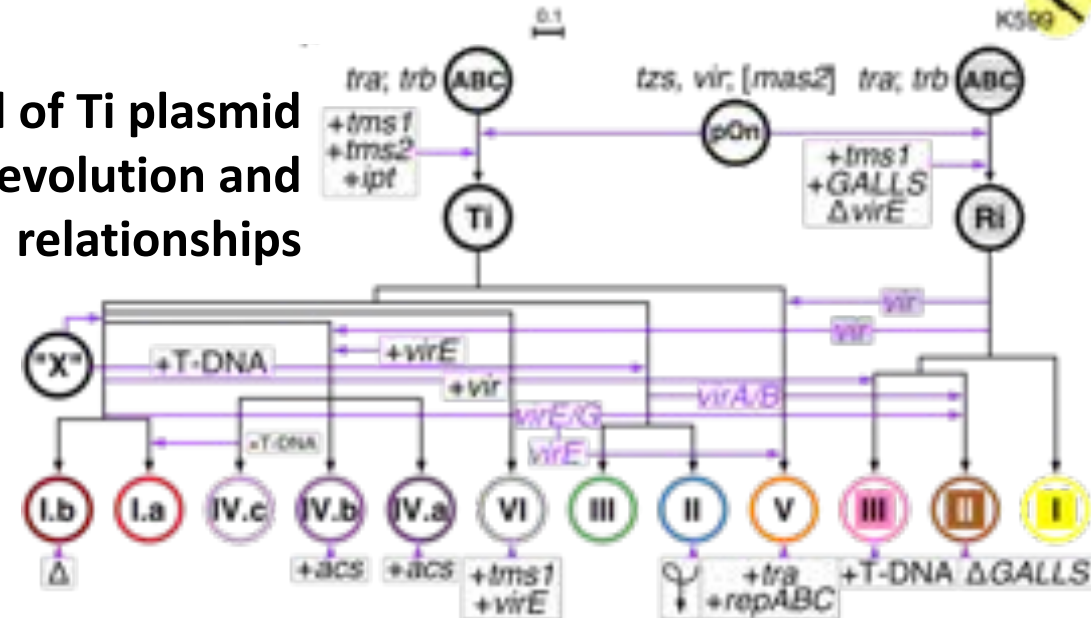# Classification and evolution of the Ti plasmids
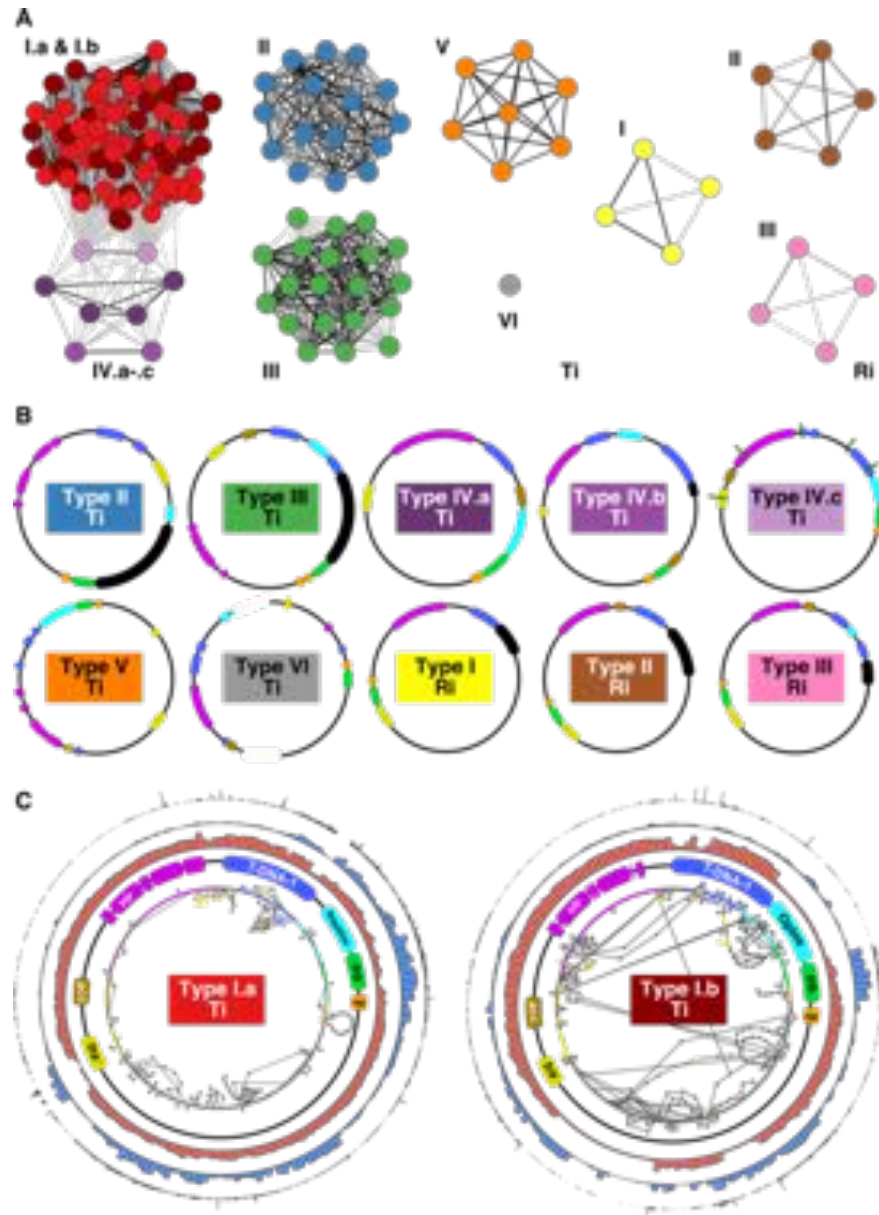
Categorized Ti plasmids using K-mers, gene content, phylogeny, structure
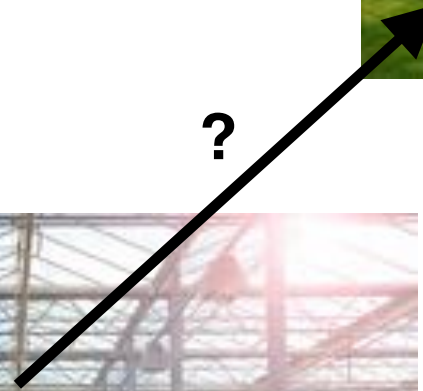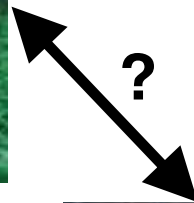
Ti plasmids represent 6 types Ri plasmids have 3 types

Model of Ti plasmid evolution and relationships

Gene synteny Networks to visualize variation/indels

Weisberg et al., 2020 Science 368 (6495) eaba5256

How does *Agrobacterium* spread? How do virulence plasmids impact the spread of disease and creation of new lineages?
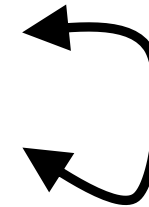

? ?

Are strains shared between nurseries?

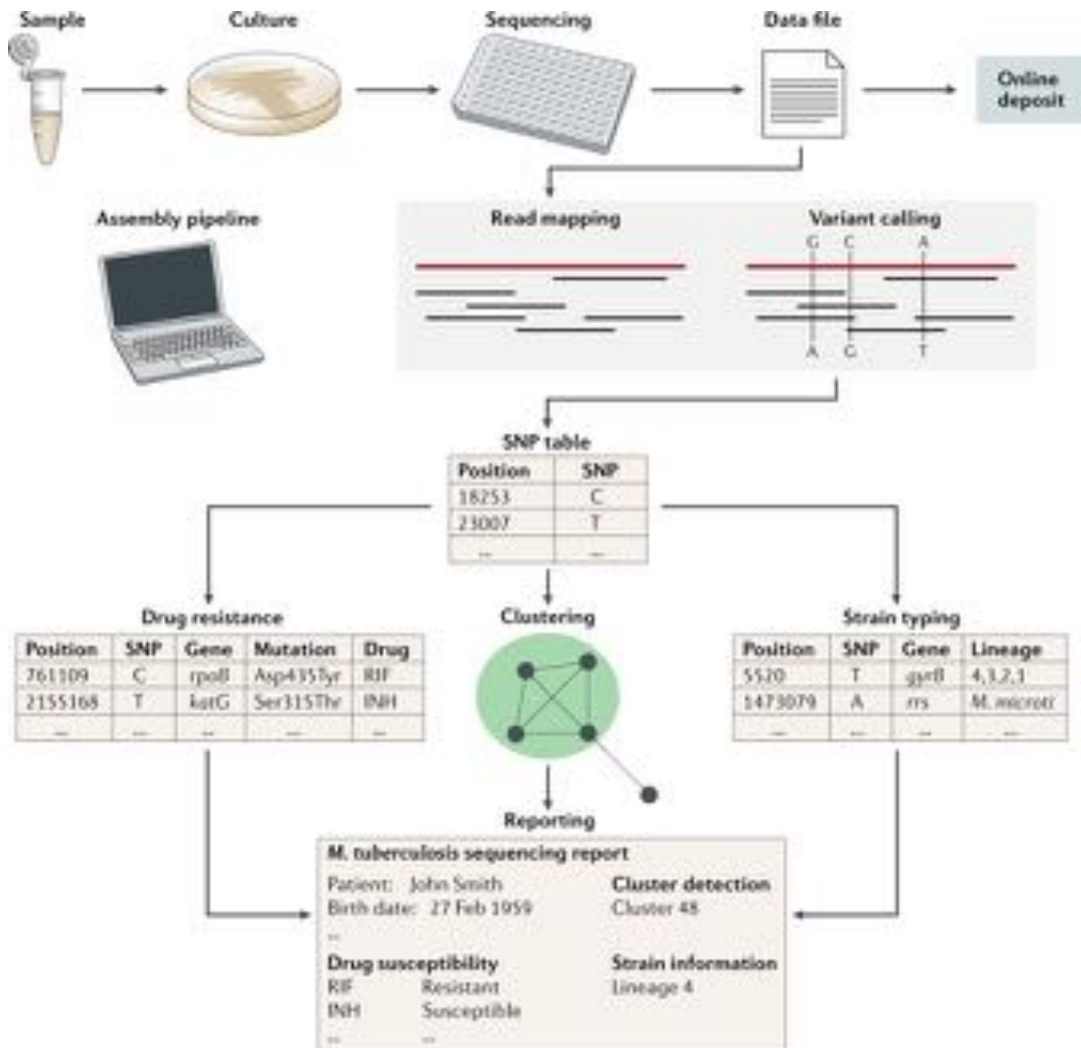Are there multiple infection genotypes?

How long do genotypes persist?

Does Ti/Ri plasmid conjugation play a role in infection?

Ti

Typical epidemiology approaches track outbreaks using whole genome SNPS

- Compare against a common reference genome

- Identify single nucleotide polymorphisms (SNPs)

- Cluster strains into genotypes based on pairwise SNP differences (0 -10 or 25 SNP threshold)

- Track outbreaks by comparing genotypes

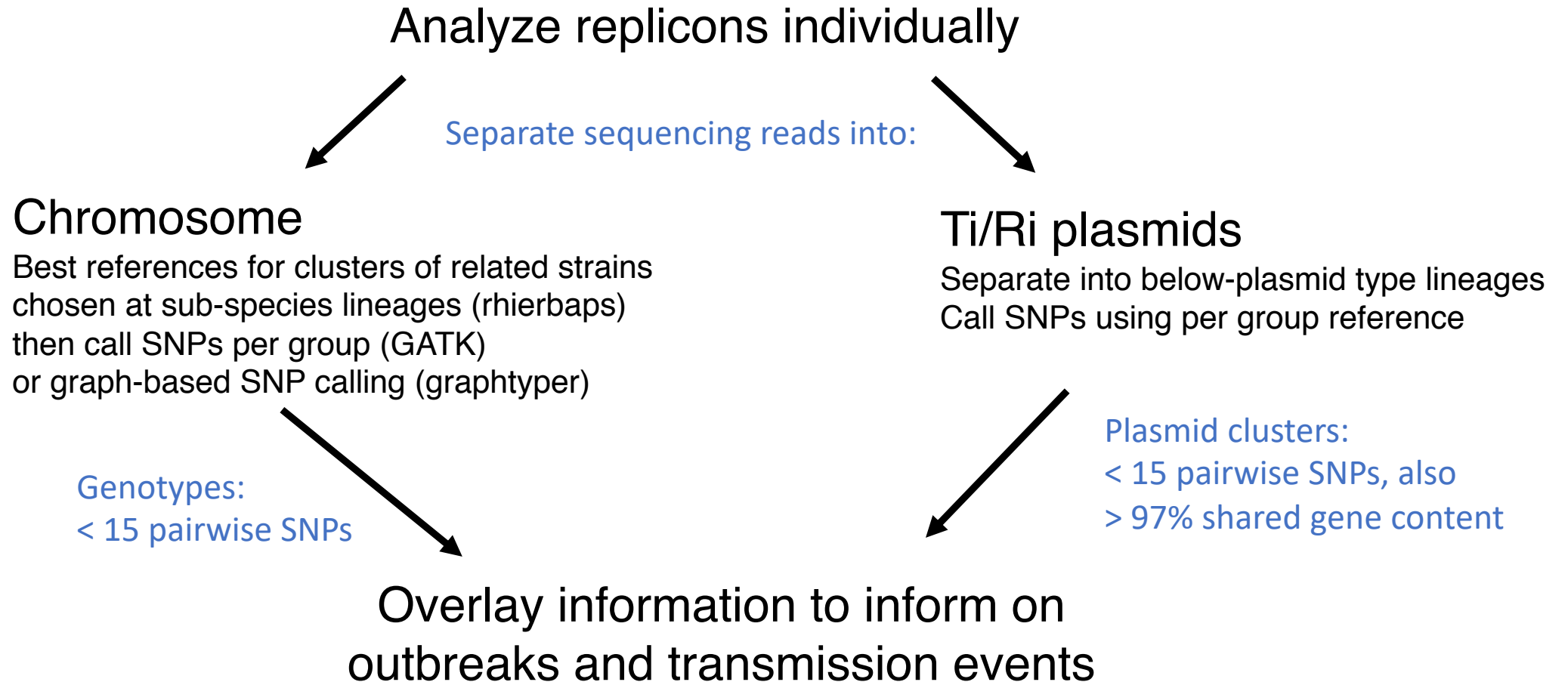Meehan et al 2019

Roetzer et al., 2013
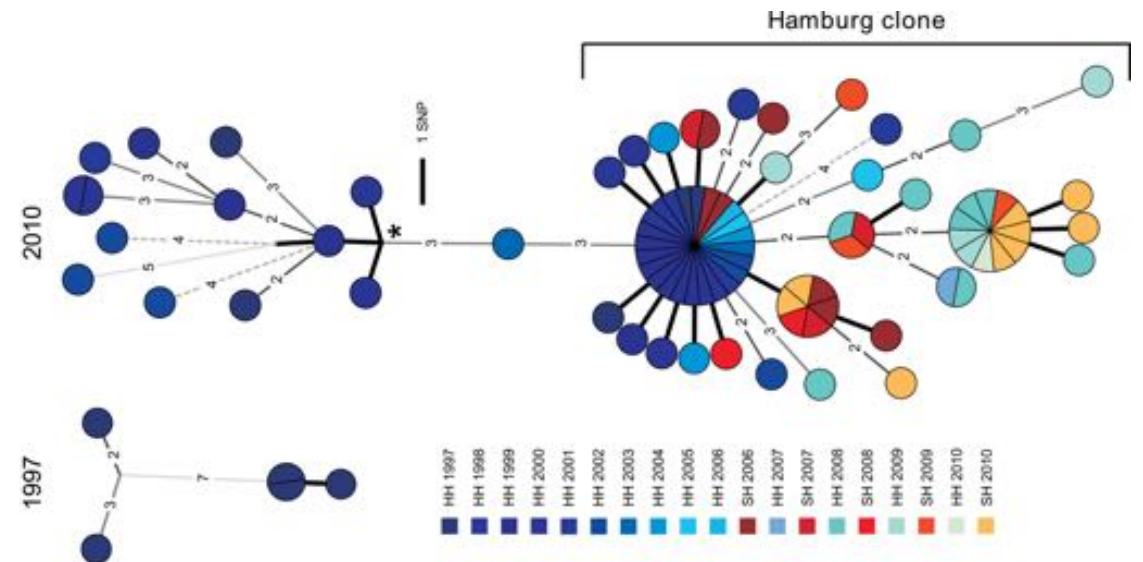
Why studying the epidemiology of agrobacteria is difficult

- *Agrobacterium* has high genetic diversity
  - 4 polyphyletic lineages/biovars
  - > 20 species-level groups (ANI)
  - No single common reference genome (leads to noise)
    - Poor reference choice = greater # of false positives, negatives

- A conjugative plasmid confers virulence
  - Plasmids spread horizontally and create new pathogenic lineages
  - Tools for epidemiology are not applicable to plasmids
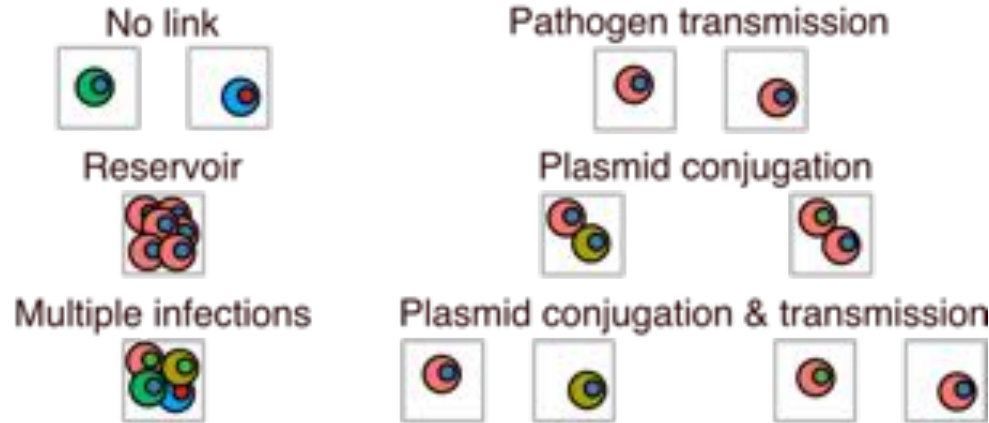 - Plasmids change as well

How did we approach this?

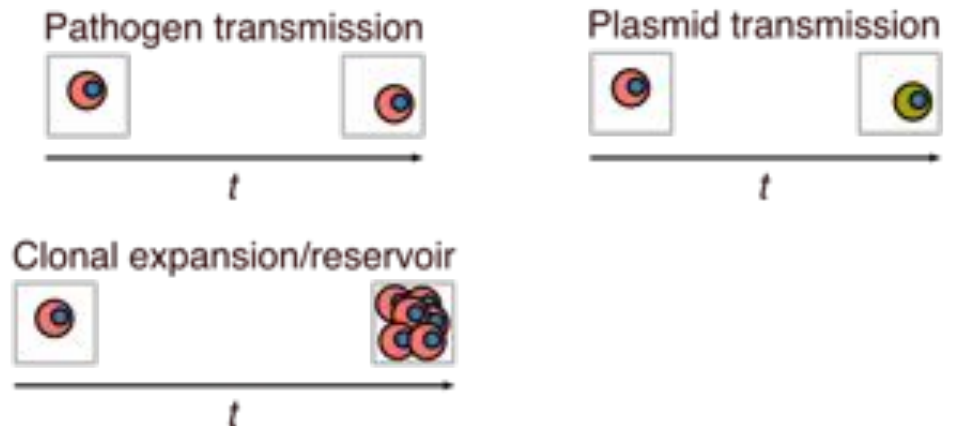- Model transmission of chromosome and plasmids as uncoupled replicons

Analyze replicons individually

Separate sequencing reads into:

Chromosome
Best references for clusters of related strains
chosen at sub-species lineages (rhierbaps)
then call SNPs per group (GATK)
or graph-based SNP calling (graphtyper)

Ti/Ri plasmids
Separate into below-plasmid type lineages
Call SNPs using per group reference

Genotypes:
< 15 pairwise SNPs

Plasmid clusters:
< 15 pairwise SNPs, also
> 97% shared gene content

Overlay information to inform on
outbreaks and transmission events

Predicted epidemiological patterns

Spatial patterns

No link

Reservoir

Multiple infections

Pathogen transmission

Plasmid conjugation

Plasmid conjugation & transmission

Temporal patterns

Pathogen transmission

Plasmid transmission

Clonal expansion/reservoir

Transmission of chromosome and plasmid modelled independently

Chromosome & plasmid in phase -> transmission

Chromosome & plasmid out of phase -> conjugation occurred

Cannot draw conclusions on direct vs indirect

Growers were deidentified
S# = state; C# = country; N# = nursery/location

23

Epidemiological patterns
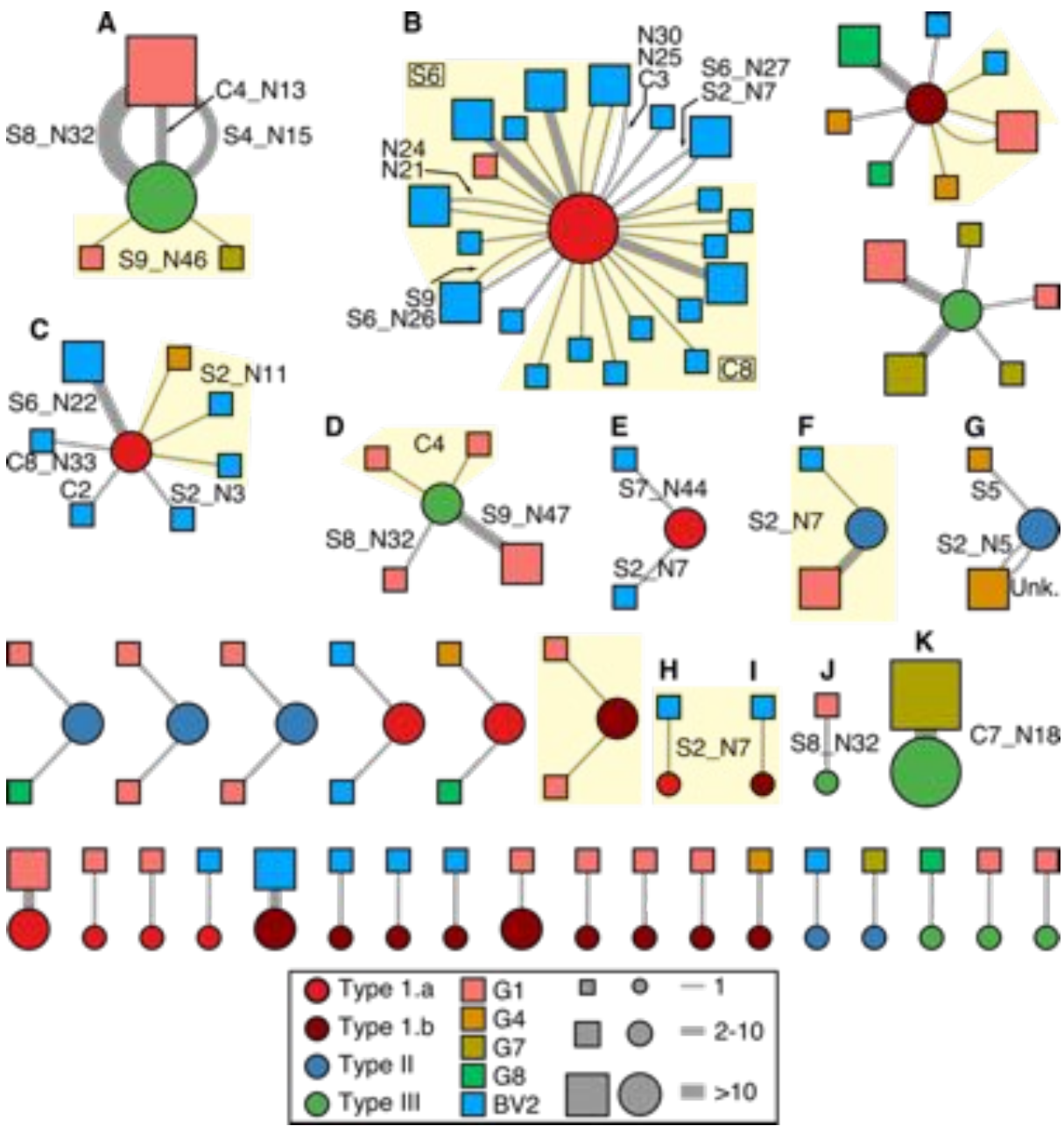
Two networks were combined:

Circle: plasmid cluster
(≤15 SNPs; ≥97% genes)

Square: strain genotype
(≤15 SNPs)

Edge: strain at same location

Colored according to classification
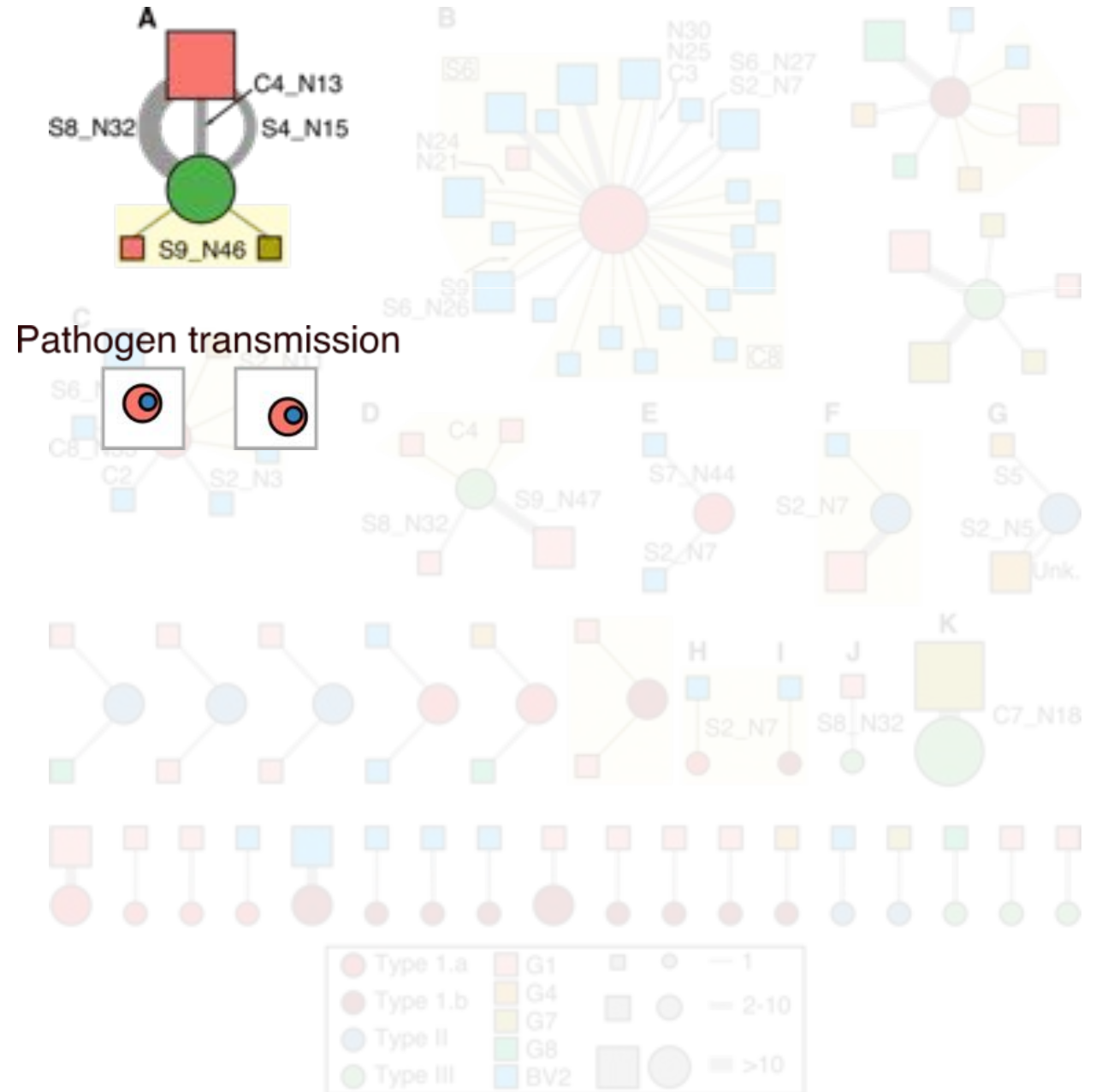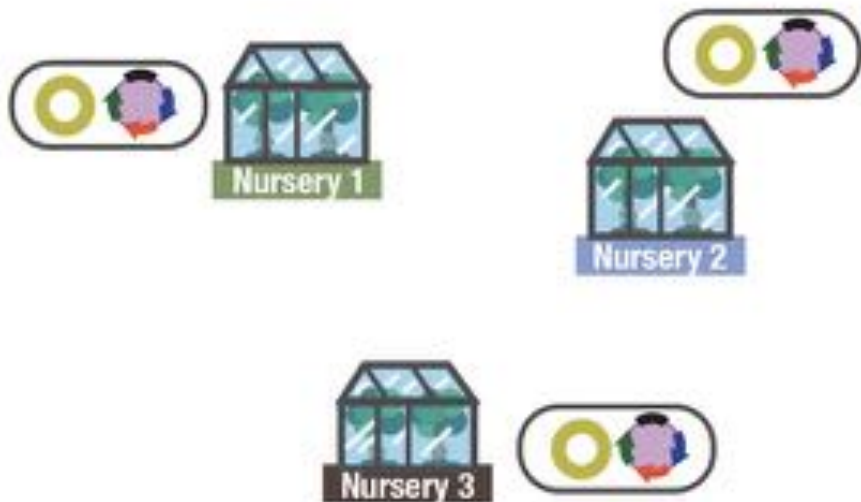(genomospecies, plasmid type)

Scaled to count

BV2   I.a

7 indicative of strain/plasmid transmission

≥2 edges between strain + plasmid

Example shows 2 countries, 3 locations

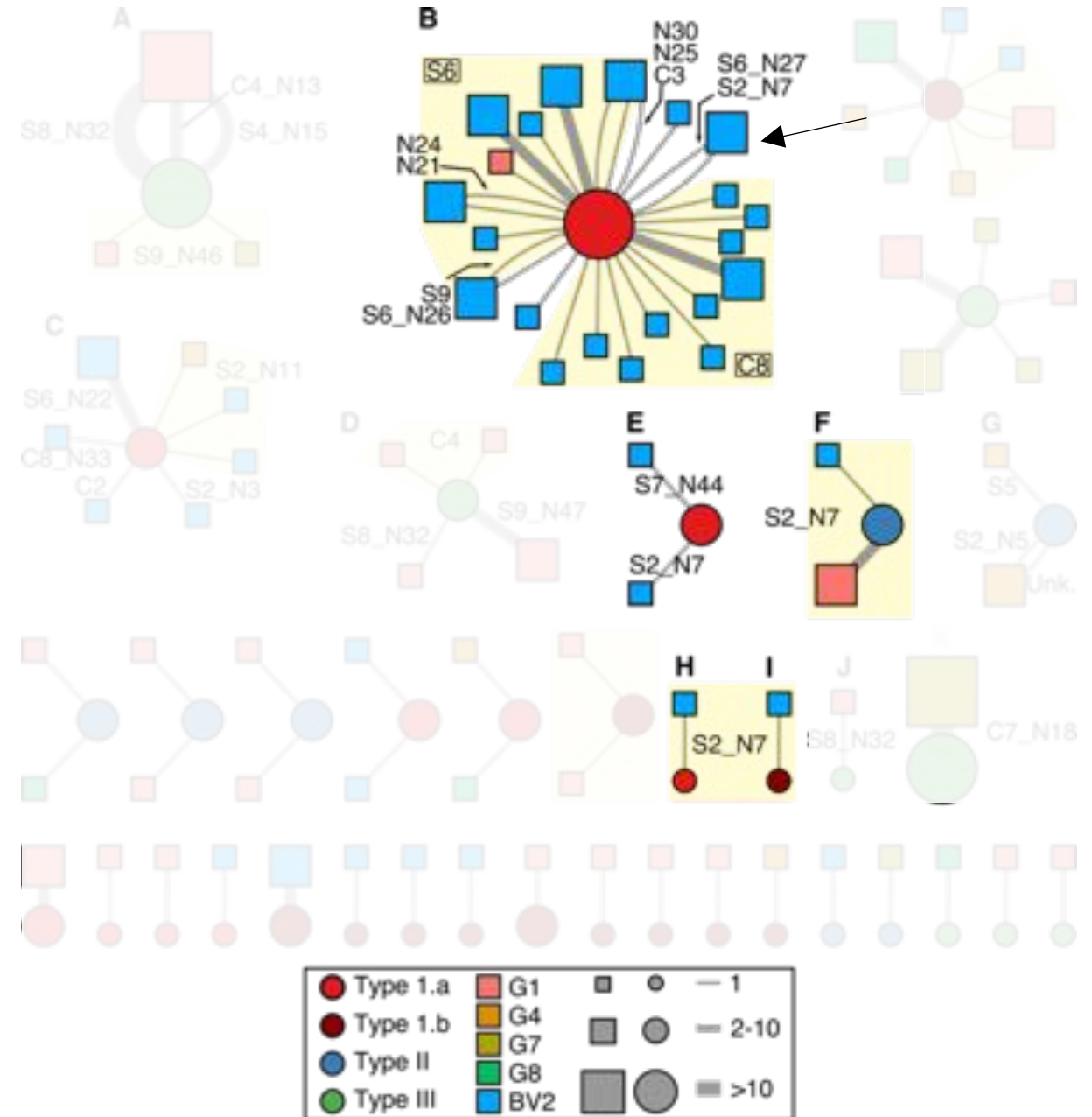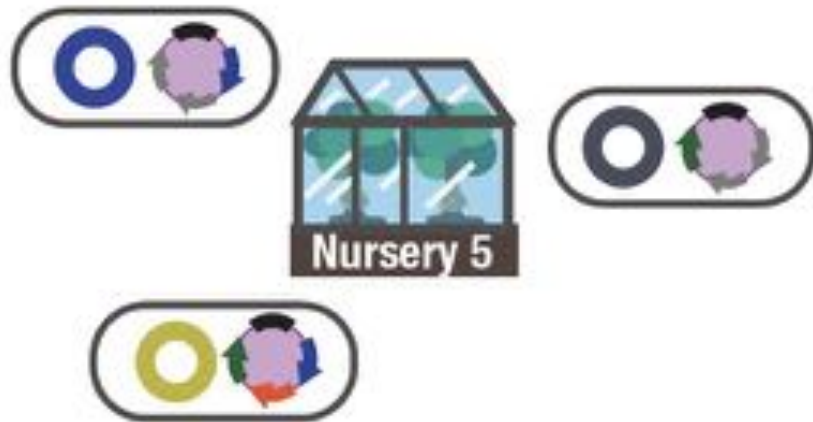Potential common source; one is a supplier

Some nurseries have multiple infections

One particularly unfortunate grower (S2_N7) has 7 unique genotype/plasmid combinations.

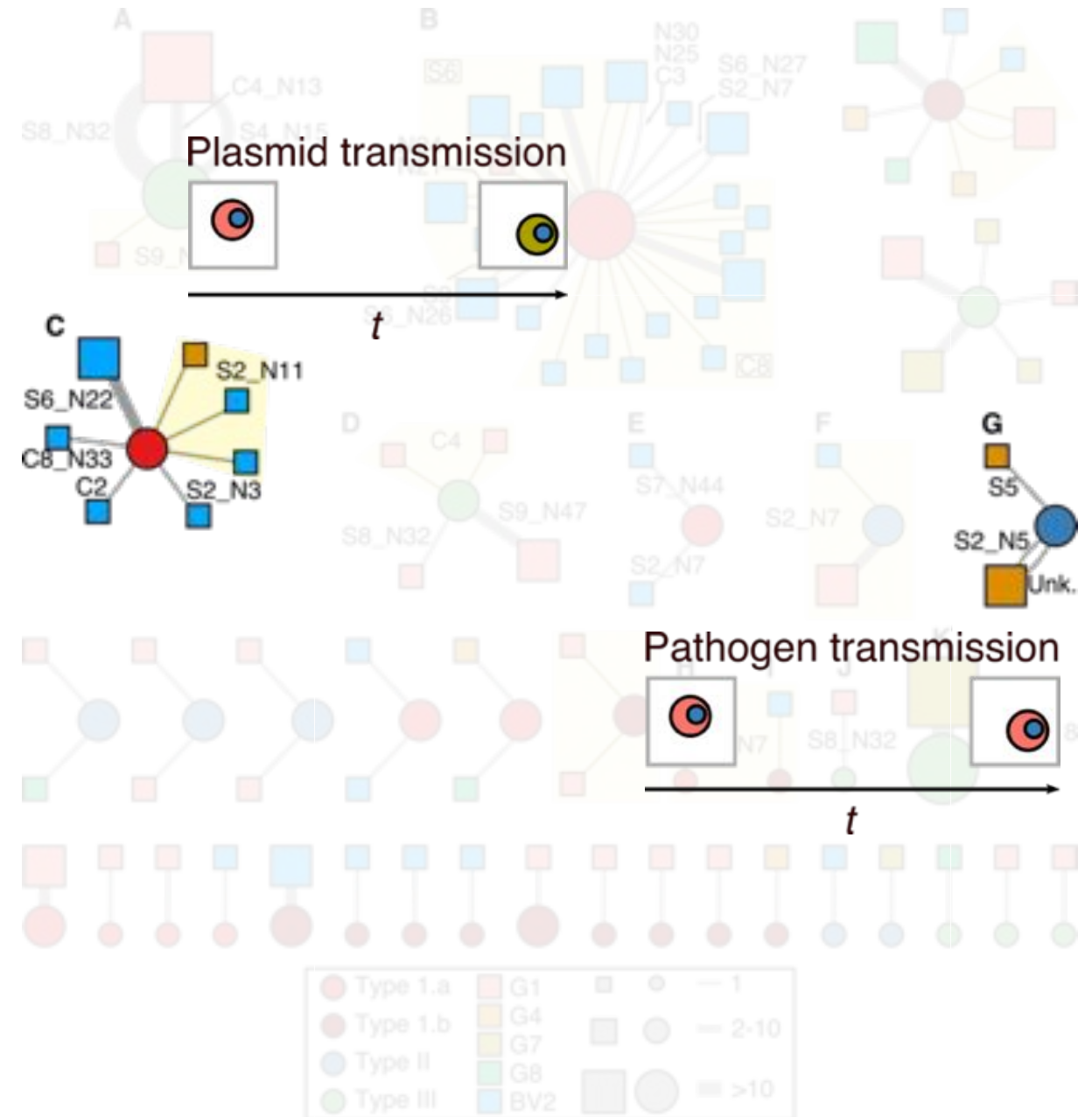Multiple infections co-occurring and multiple introductions of pathogens

2 indicative of long-term persistence

C) C2 collected in 1964; others 1995-2009
Plasmids have ≤ 2 SNPs

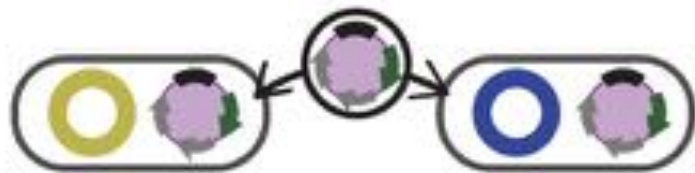G) Strains (7 SNPs) + plasmids (0 SNPs) sampled 60 years apart
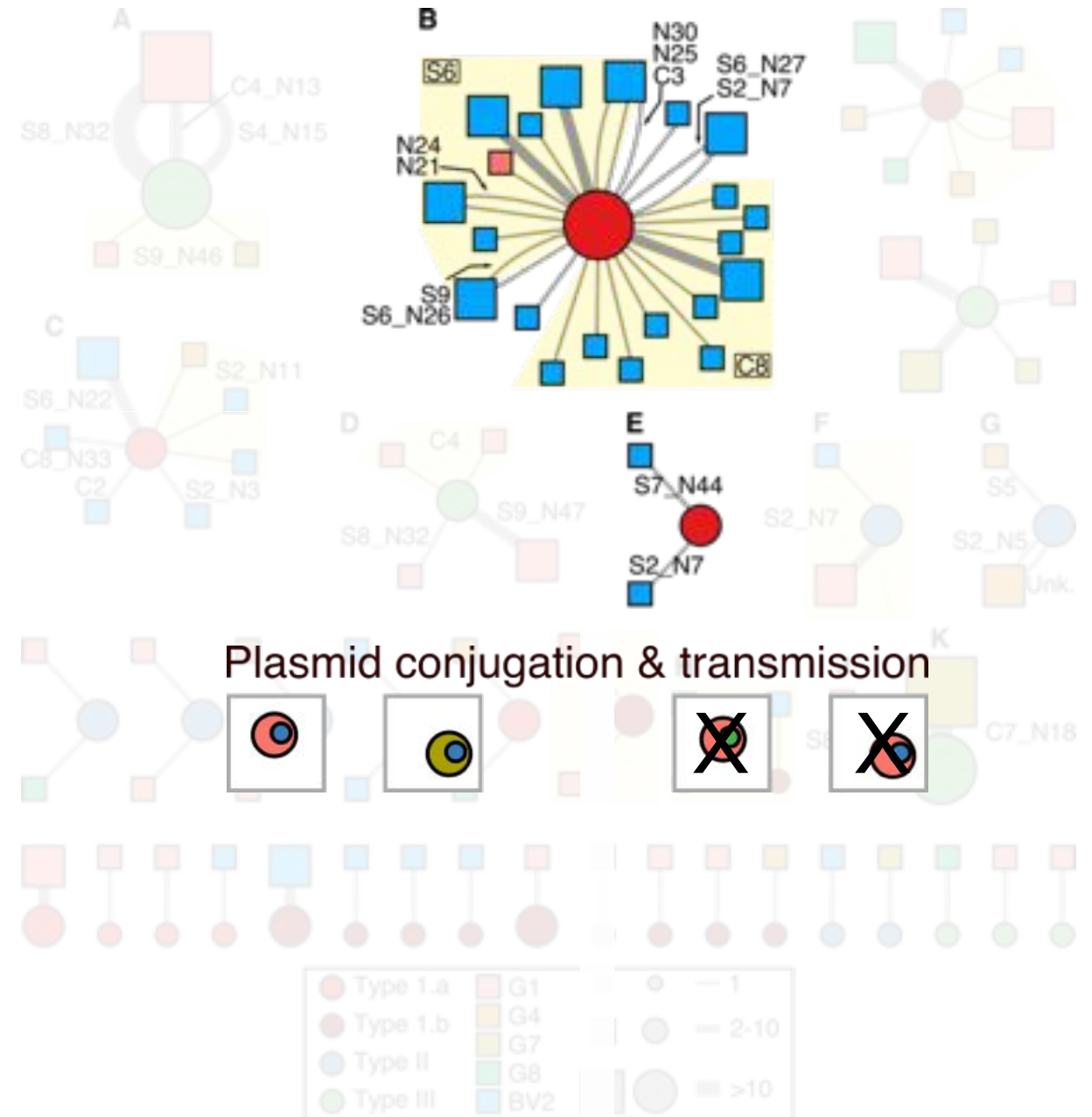
15 suggesting plasmid conjugation causes spread

Plasmids (circle) as central hub = associated with multiple strain genotypes

B) 1 plasmid lineage (≤ 13 SNPs) -> 21 genotypes
3 states & 3 countries

E) 1 plasmid lineage (2 SNPs) -> 2 genotypes (22,000 SNPs)
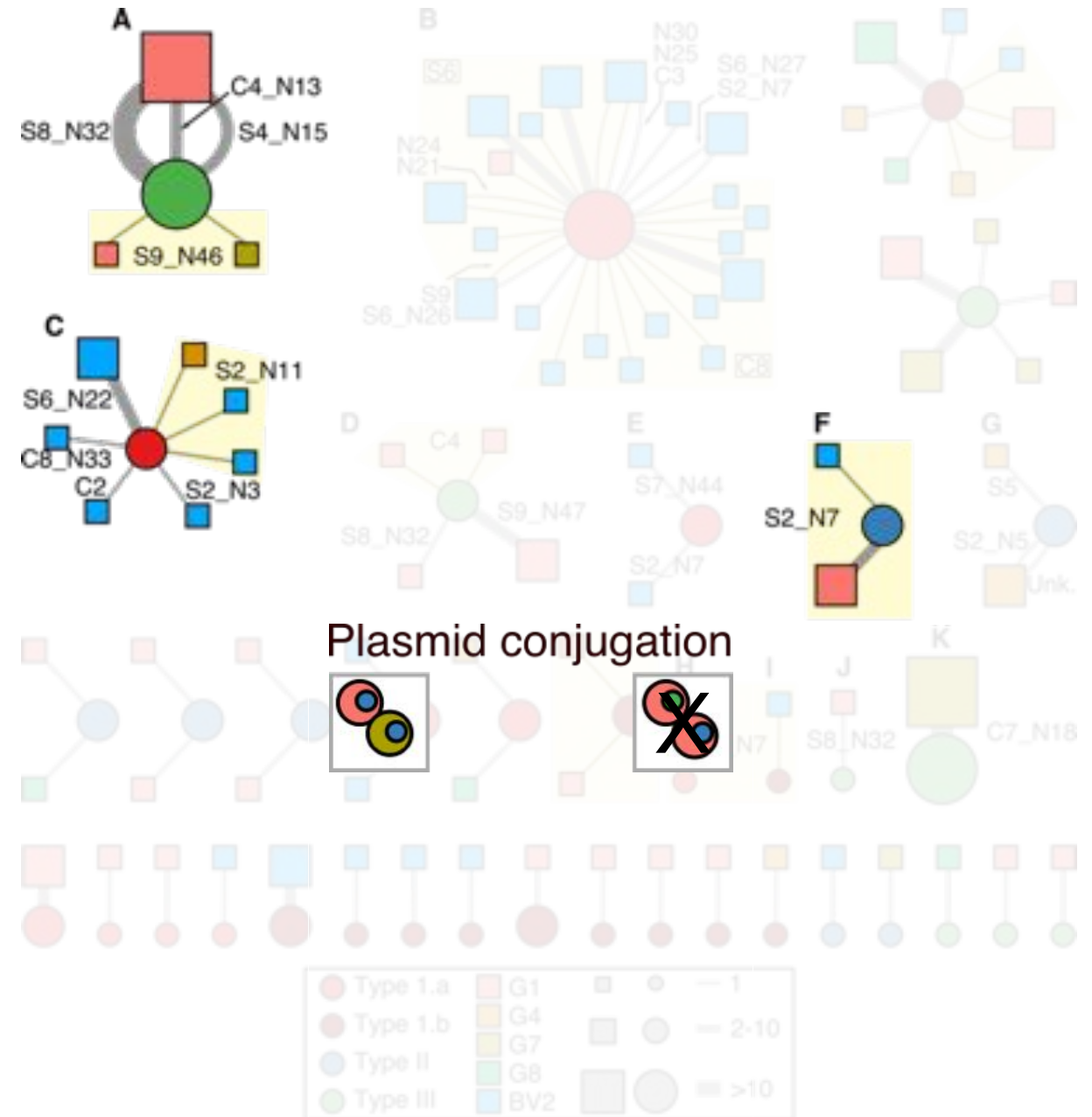
Where? When?
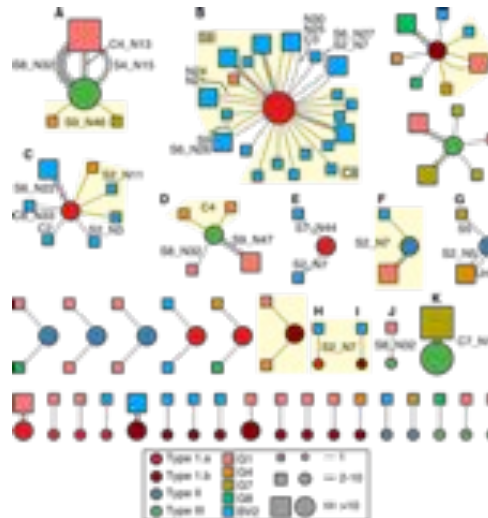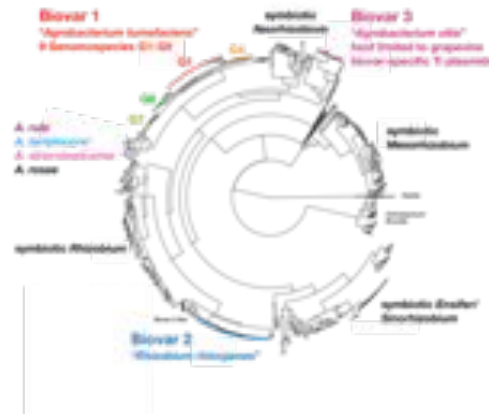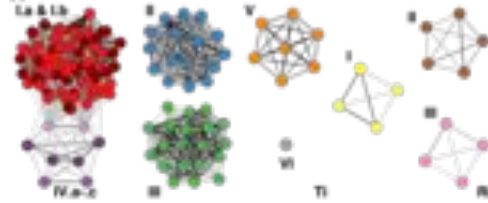
Plasmid conjugation & transmission

3 consistent with plasmid conjugation at a single location

Plasmids (circle) as central hub = associated with multiple strain genotypes & collected from same location (highlighted in yellow)

Summary

- WGS is extremely <u>cost effective</u>
  - Already being used to track foodborne pathogens
  - Time is ripe for implementing in agriculture

- Applied in a unique way to track pathogens and their mobile genetic elements

- Clarified relationships of agrobacteria + rhizobia

- Overcame challenges inherent to plasmids
  - Classified Ti/Ri plasmids + modeled their evolution
  - Used as a framework for epidemiology

- HGT of plasmids diversifies pathogens, plasmids, and promotes persistence and spread of disease

- Human activities (plant production industry) escalate the process

Acknowledgements

Jeff Chang Lab
Edward Davis II
Chih-Feng Wu
Michael Belcher

Niklaus Grunwald Lab
Javier F Tabima

Joyce Loper Lab
Max Kohen
Brenda Shaffer

OSU Plant Clinic
Melodie Putnam

Larry Moore
Marilyn Miller
Walt Ream
Erh-Min Lai
Chih-Horng Kuo

Twitter: @weisbergaj

Department of Botany and Plant Pathology
USDA ARS

Funding
USDA SCRI
NSF GRFP to Ed Davis
USDA NIFA Fellowship to A. Weisberg

31