Migration and evolution of
Phytophthora ramorum

Niklaus J. Grünwald
Horticultural Crops Research Laboratory
USDA Agricultural Research Service
Outline

1. Introduction
2. Migration
   - Markers systems
   - US population
   - Global Population
3. Evolution
4. Lessons learned
Sudden oak death

Photos courtesy Marin County Fire Dept.
SOD distribution
West coast shipments resulted in ...

- 1.6 million potentially infected plants shipped
- estimated about 11,000 plants infected with *P. ramorum* sold
- *P. ramorum* detected in 175 infested sites in over 20 states
Trace-forwards and positive detections across the U.S. July 2004

Source: USDA APHIS
Accelerated dispersal from nursery crops to continental US

Urban or natural forests

Nursery industry

Occasional jump to nursery crops
2. Migration: US nurseries

• Do US nursery populations show genetic diversity and population structure?
• Infer major migration pathways and potential sources of recent migrants?
• Evidence for sexual reproduction?
European and North American *P. ramorum* isolates form three distinct genetic groups

– By AFLP, microsatellites, and mitochondrial DNA

Ivors et al. Mol. Ecol. 2006

Three clonal lineages of *P. ramorum*

- **EU1**
  - initially limited to Europe
  - now regularly found in US nurseries
  - A1 mating type (A2 rare in Belgium)

- **NA1**
  - limited to North America
  - responsible for forest epidemics
  - most common lineage in US nurseries
    - A2 mating type

- **NA2**
  - US nurseries
  - A2 mating type

Grünwald et al. Phytopathology 2009
Microsatellites
Figure 1: Subjective view of the changing relative importance of different molecular markers. The horizontal axis indicates time. At each time point, the vertical axis corresponds to the total use of molecular markers. If more than one molecular marker is used at a given time point, its relative importance is reflected by its proportion on the vertical axis. AFLP, amplified fragment length polymorphism; RAPD, randomly amplified polymorphic DNA; RFLP, restriction fragment length polymorphism; SNP, single nucleotide polymorphism.
What are microsatellites?

- Microsatellites (also known as SSR – Simple Sequence Repeats)

  Mononucleotide SSR (A)\textsuperscript{11}
  
  AAAAAAAAAAAA

  Dinucleotide SSR (GT)\textsuperscript{6}
  
  GTGTGTGTGTGT

  Trinucleotide SSR (CTG)\textsuperscript{4}
  
  CTGCTGCTGCTG

  Tetranucleotide SSR (ACTC)\textsuperscript{4}
  
  ACTCACTCACTCACTC
Microsatellites

- Simple Sequence Repeats (SSRs)
- consist of repeating units of 1–6 bp length:

- Highly polymorphic
- codominant
Microsatellites

Fragment size (bp)
Microsatellites

What are microsatellites?

- **Homozygous**

  ...CGTAGCCTTGCATCCTTCTCTCTCTCTCTCTATCGGTACTACGTGG...
  ...CGTAGCCTTGCATCCTTCTCTCTCTCTCTCTATCGGTACTACGTGG...

  5’ flanking region  microsatellite locus  3’ flanking region

- **Heterozygous**

  ...CGTAGCCTTGCATCCTTCTCTCTCTCTCTCT...CTCTCTCTCTCTCTCTCTCTCTATCGGTACTACGTGG...
  ...CGTAGCCTTGCATCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTATCGGTACTACGTGG...
Microsatellites

How do microsatellites mutate?

- Replication Slippage
- Unequal crossing-over during meiosis
Microsatellites

Models of Mutation

- Stepwise Mutation Model (SMM)

  two alleles that differ by one repeat are more closely related than alleles that differ by many repeats
When two alleles are identical in state but not identical by descent, is known as homoplasy.
Microsatellites

Applications

- Forensics
- Diagnosis and Identification of Human Diseases
- Population Studies
- Conservation Biology
Multilocus Genotyping

DNA extraction facilitated by FastPrep tissue homogenizer

Microsatellite & AFLP preparation in 96 well plates

Lineages reported on Grunwald web page

Samples run through sequencer and electropherogram data analyzed in GeneMapper
Genotyping of US nursery isolates

Genotyped 279 isolates collected between 2004 and 2007 from nurseries in 19 states using microsatellite primers from Prospero et al. 2007 and Ivors et al. 2006 (PrMS6, PrMS9C3, PrMS39, PrMS43, PrMS45, 18, and 64)

NA1: 228 isolates, 53 multilocus genotypes

NA2: 17 isolates, 2 multilocus genotypes

EU1: 34 isolates, 2 multilocus genotypes

Distribution of lineages in US nurseries by state

Two distinct groups within nursery NA1 isolates

Probability of membership for each isolate in one of two inferred groups:

Two major migration pathways:
Two major migration pathways:
Distribution of NA1 groups by state

Sampling year 2004 unless indicated otherwise
USDA APHIS confirmed trace-forwards for 2004
Populations established?

A. Georgia

B. Virginia

C. Texas

D. Connecticut
NA1 isolates clustered into two groups, one containing isolates from Connecticut, Oregon, and Washington and the other isolates from California and the remaining states.

This structure suggests migration to Connecticut from Oregon or Washington (2004) and from California to other examined states (mainly in 2004), consistent with USDA APHIS findings.
2. Migration: Global patterns
Hypothesis on global migration
http://oregonstate.edu/~grunwaln/phytophthora.php

Phytophthora ramorum Multilocus Genotyping Database

Filter selections

<table>
<thead>
<tr>
<th>Host Genus</th>
<th>Host Species</th>
<th>Country</th>
<th>State/Prov</th>
<th>County/Town</th>
<th>Year</th>
<th>Contact</th>
<th>Lineage</th>
<th>Mating Type</th>
<th>Sort by</th>
<th>Filter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any</td>
<td>Any</td>
<td>Any</td>
<td>Any</td>
<td>Any</td>
<td>Any</td>
<td>Any</td>
<td>Any</td>
<td>Any</td>
<td>Any</td>
<td>any</td>
</tr>
</tbody>
</table>

Results (23)

<table>
<thead>
<tr>
<th>Grunwald ID</th>
<th>Original ID</th>
<th>Host Genus</th>
<th>Host Species</th>
<th>Country</th>
<th>State/Prov</th>
<th>County/Town</th>
<th>Year Isolated</th>
<th>Contact</th>
<th>Lineage</th>
<th>Mating Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>GR-07-003</td>
<td>07-83-37-A</td>
<td>-</td>
<td>USA</td>
<td>OR</td>
<td>-</td>
<td>2007</td>
<td>N. Ostebræus</td>
<td>EU1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>GR-07-004</td>
<td>07-83-37-A</td>
<td>-</td>
<td>USA</td>
<td>OR</td>
<td>-</td>
<td>2007</td>
<td>N. Ostebræus</td>
<td>NA1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>GR-07-005</td>
<td>07-83-37-A</td>
<td>-</td>
<td>USA</td>
<td>OR</td>
<td>-</td>
<td>2007</td>
<td>N. Ostebræus</td>
<td>EU1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>GR-07-012</td>
<td>07-331-16-A</td>
<td>Pini</td>
<td>japonica</td>
<td>USA</td>
<td>OR</td>
<td>2007</td>
<td>N. Ostebræus</td>
<td>EU1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>GR-07-013</td>
<td>07-331-16-B</td>
<td>Pini</td>
<td>japonica</td>
<td>USA</td>
<td>OR</td>
<td>2007</td>
<td>N. Ostebræus</td>
<td>EU1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>GR-07-014</td>
<td>07-331-16-C</td>
<td>Pini</td>
<td>japonica</td>
<td>USA</td>
<td>OR</td>
<td>2007</td>
<td>N. Ostebræus</td>
<td>EU1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PR-07-002</td>
<td>07-83-37-B</td>
<td>-</td>
<td>USA</td>
<td>OR</td>
<td>-</td>
<td>2007</td>
<td>N. Ostebræus</td>
<td>NA1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PR-07-003</td>
<td>07-83-37-B</td>
<td>-</td>
<td>USA</td>
<td>OR</td>
<td>-</td>
<td>2007</td>
<td>N. Ostebræus</td>
<td>EU1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PR-07-004</td>
<td>07-83-37-C</td>
<td>-</td>
<td>USA</td>
<td>OR</td>
<td>-</td>
<td>2007</td>
<td>N. Ostebræus</td>
<td>EU1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PR-07-005</td>
<td>07-83-37-D</td>
<td>-</td>
<td>USA</td>
<td>OR</td>
<td>-</td>
<td>2007</td>
<td>N. Ostebræus</td>
<td>EU1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PR-07-006</td>
<td>07-83-37-B</td>
<td>-</td>
<td>USA</td>
<td>OR</td>
<td>-</td>
<td>2007</td>
<td>N. Ostebræus</td>
<td>NA1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PR-07-007</td>
<td>07-83-37-C</td>
<td>-</td>
<td>USA</td>
<td>OR</td>
<td>-</td>
<td>2007</td>
<td>N. Ostebræus</td>
<td>EU1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PR-07-008</td>
<td>07-83-37-D</td>
<td>-</td>
<td>USA</td>
<td>OR</td>
<td>-</td>
<td>2007</td>
<td>N. Ostebræus</td>
<td>EU1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PR-07-026</td>
<td>07-1161-1C</td>
<td>Rhododendron</td>
<td>lapidosides</td>
<td>USA</td>
<td>OR</td>
<td>2007</td>
<td>N. Ostebræus</td>
<td>NA1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PR-07-027</td>
<td>07-1161-1E</td>
<td>Rhododendron</td>
<td>lapidosides</td>
<td>USA</td>
<td>OR</td>
<td>2007</td>
<td>N. Ostebræus</td>
<td>NA1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PR-07-026</td>
<td>07-1161-1A</td>
<td>Rhododendron</td>
<td>lapidosides</td>
<td>USA</td>
<td>OR</td>
<td>2007</td>
<td>N. Ostebræus</td>
<td>NA1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PR-07-027</td>
<td>07-1161-1B</td>
<td>Rhododendron</td>
<td>lapidosides</td>
<td>USA</td>
<td>OR</td>
<td>2007</td>
<td>N. Ostebræus</td>
<td>NA1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PR-07-026</td>
<td>07-1161-1C</td>
<td>Rhododendron</td>
<td>lapidosides</td>
<td>USA</td>
<td>OR</td>
<td>2007</td>
<td>N. Ostebræus</td>
<td>NA1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PR-07-026</td>
<td>07-1161-1D</td>
<td>Rhododendron</td>
<td>lapidosides</td>
<td>USA</td>
<td>OR</td>
<td>2007</td>
<td>N. Ostebræus</td>
<td>NA1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PR-07-027</td>
<td>07-1161-1E</td>
<td>Rhododendron</td>
<td>lapidosides</td>
<td>USA</td>
<td>OR</td>
<td>2007</td>
<td>N. Ostebræus</td>
<td>NA1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PR-07-027</td>
<td>07-1161-1F</td>
<td>Rhododendron</td>
<td>lapidosides</td>
<td>USA</td>
<td>OR</td>
<td>2007</td>
<td>N. Ostebræus</td>
<td>NA1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PR-07-027</td>
<td>07-1161-1G</td>
<td>Rhododendron</td>
<td>lapidosides</td>
<td>USA</td>
<td>OR</td>
<td>2007</td>
<td>N. Ostebræus</td>
<td>NA1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PR-07-063</td>
<td>07-1240-1</td>
<td>Camellia</td>
<td>USA</td>
<td>OR</td>
<td>-</td>
<td>2007</td>
<td>N. Ostebræus</td>
<td>NA1</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Notes: EU1, NA1, NA2 indicate lineage groups.
3. Evolution of *P. ramorum*

- Do we see DNA sequence variation in nuclear genes? Within lineages? Among lineages?

- Are the three lineages recent (i.e. a consequence of isolation following introduction) migrants from same location?

- What about sex?
Sequenced five nuclear loci in a diverse sample of isolates:

- beta-tubulin

- trp1 (tryptophan biosynthesis)

- gwEuk.30.30.1, predicted protein containing a glycosyl transferase domain (intron with Pr9C3 microsatellite)

- 2 putative effectors (PrAvh120 & PrAvh121)

→ When 2+ heterozygous sites, cloned gene to obtain haplotypes
Relationships among lineages differed among genes, suggesting historical recombination between lineages.

Outgroups: Ph – *P. hibernalis*, Pf – *P. foliorum*, Pl – *P. lateralis*
Using the coalescent to age lineages

The coalescent uses stochastic processes, based on a population genetic model, to approximate the ancestry of a sample of DNA sequences going backwards in time by evaluating all possible mutational pathways back to a common ancestor.

Using the programs genetree and recom, we obtained estimates of the time to the most recent common ancestor (TMRCA) of each gene and the ages of mutations in the genes.
Present 22 individuals
Present

22 individuals
18 ancestors
Present

22 individuals
18 ancestors
16 ancestors

Time
Present

22 individuals
18 ancestors
16 ancestors
14 ancestors
Present

22 individuals
18 ancestors
16 ancestors
14 ancestors
12 ancestors

Time
22 individuals
18 ancestors
16 ancestors
14 ancestors
12 ancestors
9 ancestors
Present: 22 individuals, 18 ancestors, 16 ancestors, 14 ancestors, 12 ancestors, 9 ancestors, 8 ancestors.
Present

22 individuals
18 ancestors
16 ancestors
14 ancestors
12 ancestors
9 ancestors
8 ancestors
8 ancestors

Time
Present

22 individuals
18 ancestors
16 ancestors
14 ancestors
12 ancestors
9 ancestors
8 ancestors
8 ancestors
7 ancestors
7 ancestors
5 ancestors

Time
Present

22 individuals
18 ancestors
16 ancestors
14 ancestors
12 ancestors
10 ancestors
8 ancestors
7 ancestors
5 ancestors

Time

22 individuals
18 ancestors
16 ancestors
14 ancestors
12 ancestors
10 ancestors
8 ancestors
7 ancestors
5 ancestors
Present

22 individuals

18 ancestors

16 ancestors

14 ancestors

12 ancestors

10 ancestors

9 ancestors

8 ancestors

8 ancestors

7 ancestors

7 ancestors

5 ancestors

5 ancestors

3 ancestors

3 ancestors
Present

22 individuals
18 ancestors
16 ancestors
14 ancestors
12 ancestors
10 ancestors
8 ancestors
8 ancestors
7 ancestors
7 ancestors
5 ancestors
5 ancestors
3 ancestors
3 ancestors
3 ancestors
3 ancestors
Present

22 individuals
18 ancestors
16 ancestors
14 ancestors
12 ancestors
9 ancestors
8 ancestors
8 ancestors
7 ancestors
7 ancestors
5 ancestors
5 ancestors
3 ancestors
3 ancestors
3 ancestors
3 ancestors
2 ancestors

Time
Present

22 individuals
18 ancestors
16 ancestors
14 ancestors
12 ancestors
10 ancestors
9 ancestors
8 ancestors
7 ancestors
6 ancestors
5 ancestors
4 ancestors
3 ancestors
2 ancestors
1 ancestor

Time
Present

Time

Most recent common ancestor (MRCA)
A diagram illustrating the concept of mutation over time. The sequence TCGAGGTATTAAC is shown with a mutation indicating a change from T to T at a specific point in the sequence. The timeline is labeled as 'Time' with the 'Present' at the top. The diagram also indicates the 'Most recent common ancestor (MRCA)'.
Present

TCGAGGTATTAAC

TCTAGGTATTAAC

Time

Mutation

Most recent common ancestor (MRCA)
Most recent common ancestor (MRCA)

Time

Present

TCAGGTATAAAC
TCTAGGTATAAAC
C
Present

Ti

Most recent common ancestor (MRCA)

TCGAGGTATTAAC
TCTAGGTATTAAC
TCGAGG CATTAAC

Time
Most recent common ancestor (MRCA)
Present

Most recent common ancestor
(MRCA)
Present

Time

Most recent common ancestor (MRCA)

TCGAGGTATTAAC
TCTAGGTATTAAC
TCGAGGCATTAAC
TCTAGGTGTTAAC

TCGAGGTATTAAC
TCTAGGTATTAAC
TCGAGGCATTAAC
TCTAGGTGTTAAC
Present

Time

Most recent common ancestor (MRCA)

TCGAGGTATTAAC
TCTAGGTATTAAC
TCGAGGCATTAAC
TCTAGGGTTAACC
TCGAGGTATTAGC
Most recent common ancestor (MRCA)
Most recent common ancestor (MRCA)
**Present**

**Time**

Most recent common ancestor (MRCA)

tti

TCGAGGTATTAAC
TCTAGGATTAAC
TCGAGGCATTAAC
TCTAGGTGTTAAC
TCGAGGTATTAGC
TCTAGGTATCAAC

* ** *
**Population 1** **Population 2** **Population 3**

**Present**

Most recent common ancestor (MRCA)

**Time**

```
TCGAGGTATTAAC
TCTAGGTTATTAAC
TCGAGGCATTAAC
TCTAGGTGTTAAC
TCGAGGTATTAGC
TCTAGGTATCAAC
*   ** * *
```
Mutation ages

– Many lineage-specific mutations (open bars), which tend to be younger than shared mutations (filled bars).

Coalescent with recombination:
Mutation ages

– Many lineage-specific mutations (open bars), which tend to be younger than shared mutations (filled bars).

– btub\textsuperscript{64} was the youngest mutation shared between lineages (age = 0.11 when time to the most recent common ancestor, TMRCA, is scaled to 1.0).

Coalescent with recombination:
Coalescent analysis indicates recombination in several genes

\[\text{P. ramorum}\] lineages are descended from a sexually reproducing population

<table>
<thead>
<tr>
<th></th>
<th>btub</th>
<th>trp1</th>
<th>gwEuk.30.30.1</th>
<th>PrAvh120</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NO RECOMBINATION</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean TMRCA</td>
<td>1.27</td>
<td>1.48</td>
<td>NA</td>
<td>1.03</td>
</tr>
<tr>
<td>(\theta) estimate</td>
<td>2.05</td>
<td>3.68</td>
<td>NA</td>
<td>4.00</td>
</tr>
<tr>
<td><strong>WITH RECOMBINATION</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean TMRCA</td>
<td>2.64</td>
<td>1.49</td>
<td>2.49</td>
<td>1.22</td>
</tr>
<tr>
<td>Mean recombination events</td>
<td>6.96</td>
<td>0.00</td>
<td>4.03</td>
<td>0.96</td>
</tr>
<tr>
<td>(\theta) estimate</td>
<td>1.90</td>
<td>3.70</td>
<td>4.44</td>
<td>3.90</td>
</tr>
<tr>
<td>(\rho) estimate</td>
<td>2.88</td>
<td>0.00</td>
<td>1.90</td>
<td>0.88</td>
</tr>
</tbody>
</table>
Recombination events

A. gwEuk.30.30.1

B. PrAvh121
**P. ramorum** clonal lineages are ancient

- first *btub* mutation was 11% of the TMRCA
- lineages have probably been separated for about one tenth of their evolutionary history
- **ASSUME**: synonymous substitution rate between $2 \times 10^{-9}$ and $7 \times 10^{-9}$ mutations per site per year based on the literature for plants, animals, and fungi,
  - Estimate history of the three lineages (TMRCA) to be between 1.5 and 5.4 million years old.
- Most recent mutation shared between lineages a minimum of 165,000 years old (11% of 1.5 million years).
Conclusions

- *P. ramorum* lineages are descended from a sexually reproducing population.

- The lineages have been isolated for a significant period of time, on the order of 100,000s of years.

- The introductions of the lineages to US and Europe were likely from three geographically distinct source populations.
Hypothesis on global migration
Lessons learned

- Sample populations not single isolates
  - *P. ramorum* at least three global migrations
  - Populations in East show clonal divergence: established?

- Routine genotyping should be integral to any crop biosecurity efforts
  - routine isolate submission for genotyping

- Ideally, genotyping should be conducted in one laboratory (or in a well defined, narrow collaboration)

- Sequence to identify other novel taxa:
  - *P. pseudosyringae, P. foliorum*, etc.
Ignazio Carbone
Don Givens
Gary Chastagner
Stéphan Brière
Phytophthora genome

Funding
• FNRI
• OAN
• ODA
• NNCRC
• USFS PSW
• J. Frank Schmidt

Grünwald lab:
• Caroline Press
• Erica Goss
• Meg Larsen
• Virginia McDonald
• Val Fieland

Kenny Rolfe
Stephanie Bollmann
Kim Henslee
Karan Fairchild

Isolates

Thank you
Questions?