LANDSCAPE GENETICS AND Gene flow in the banana pathogenic fungus *Mycosphaerella fijiensis*

Adrien Rieux, Fabien Halkett, Virginie Ravigné, Marie-Françoise Zapater, Luc Pignolet, Luc de Lapeyre de Bellaire and Jean Carlier

UMR BGPI : "Biology and genetics of plant / pathogen interactions"

Campus International de Baillarguet - 34 398 MONTPELLIER – France
Context of this study or why study gene flow?

Efficient and durable strategies of disease management should be defined in time and space taking into account epidemiology and evolutionary potential of pathogens.

Objective: To infer gene flow and dispersal processes in *Mycosphaerella fijiensis*
How to measure parameters related to gene flow?

**Exemple with two class of Methods**

**Direct methods**
- Disease gradient analysis
- Average dispersal distance

**Indirect methods**
- Population genetics
- Neutral molecular markers
- Genetic differentiation between populations analysis

Problems:
- Experiments costly and hard to realize
  (Ex: problems with source identification)

$$F_{ST} = \frac{Q_s - Q_T}{1 - Q_T}$$

Q_s : Probability of identity between 2 genes sampled in the same habitat

Q_T : Probability of identity between 2 genes sampled in the whole habitats
How to measure parameters related to gene flow?

*Back to the relation between $F_{ST}$ & dispersal ability*

A high migration (rate or distance) will homogenize allelic frequencies between populations and will lead to a decrease of the $F_{ST}$ parameter.

- **Objective**: Use $F_{ST}$ (easy to measure) in order to estimate parameters related to dispersal (rate, average distance…) (difficult to measure)

- **Method**: Use a model to explicitly link $F_{ST}$ and dispersal → dispersal parameters inference
How to measure parameters related to gene flow?

Exemple with two models

- Wright’s Island model
  - Population size: constant and identical
  - Mutation – genetic drift equilibrium
  - Equal contribution to the migrants pool
  - Migration rate / generation:

\[
F_{ST} \approx \frac{1}{1 + 4Nm} \quad \text{Migration rate} \\
1 + 4Nm \quad \text{Effective size}
\]

- Not realistic

- Isolation by distance Model (IBD)
  - Because of the limited migration in space, the probability of identity between genes is higher at short distance than at long distance.

\[
\text{Slope} = \frac{1}{2\pi D \sigma^2}
\]

\[
\frac{F_{ST}}{1 - F_{ST}} \quad \text{Ln(Distance)}
\]

- \(\sigma^2\) = average squared axial parent-offspring distance
- \(D = \) effective density of adults
The patho-system *Mycosphaerella fijiensis* - Banana plant: A good model to study gene flow in pathogenic fungus populations

- Aerial Ascomycete, Haploïd, Heterothallic
- Black leaf streak of Banana

→ Recent worldwide expansion (1970’) : replacement of *M. musicola* (yellow Sigatoka)
The patho-system *Mycosphaerella fijiensis* - Banana plant: A good model to study gene flow in pathogenic fungus populations

- Aerial Ascomycete, Haploïd, Heterothallic
- Black leaf streak of Banana
- 3 dispersal modes:
  - Infected plant material transport
  - Conidia dispersion (asexual reproduction)
  - Ascospores dispersion (sexual reproduction)

- Panmictic populations
- Relative demographic stability

Application of classical population genetics methods
Previous studies related to the populations genetic structure of *Mycosphaerella fijiensis*

- **Different scales & disease dissemination modes:**
  
  → High level of genetic diversity
  
  → Relative importance of the dispersal modes of *M. fijiensis* as a function of geographical scale

![Diagram showing the relationship between geographical scale and level of dispersion](image)

*Carlier, 2003*
Previous studies related to the populations genetic structure of *Mycosphaerella fijiensis*

- Within a production area (few 100 Km long)

**Costa Rica**

- Transect of 297 km
- 15 sites (26.4 km)
- 347 isolats / 7 PCR-RFLP and 9 microsatellites markers

**Cameroon**

- Transect of 270 km
- 10 sites (26.8 km)
- 287 isolats / 15 microsatellites markers
Previous studies related to the populations genetic structure of *Mycosphaerella fijiensis*

- Within a production area (few 100 Km long)

<table>
<thead>
<tr>
<th>Basic genetic analysis</th>
<th>Cameroon</th>
<th>Costa Rica</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene diversity, $H_E$</td>
<td>0.31 to 0.46</td>
<td>0.41 to 0.59</td>
</tr>
<tr>
<td>Genotypic diversity, $D_G$</td>
<td>~ 1.00</td>
<td>~ 1.00</td>
</tr>
<tr>
<td>Linkage disequilibrium, $r_D$</td>
<td>&lt; 0.02, NS</td>
<td>&lt; 0.04 NS</td>
</tr>
<tr>
<td>Differentiation between populations, $F_{st}$</td>
<td>0.014 to 0.26</td>
<td>0.012 to 0.17</td>
</tr>
</tbody>
</table>
Previous studies related to the populations genetic structure of *Mycosphaerella fijiensis*

- Within a production area (few 100 Km long)

→ Differentiation between populations:

Costa Rica: $F_{st} = 0.012$ to $0.17$ (Costa Rica)
Cameroon: $F_{st} = 0.014$ to $0.26$ (Cameroon)

→ Clustering analysis (*STRUCTURE*, v2.2 — *Pritchard et al., 2000*)

Costa Rica: no structure detected
Cameroon: number of cluster $K = 3$
Previous studies related to the populations genetic structure of *Mycosphaerella fijiensis*

- Within a production area (few 100 Km long)

→ Differentiation between populations:

  - Costa Rica: $F_{st} = 0.012$ to $0.17$ (Costa Rica)
  - Cameroon: $F_{st} = 0.014$ to $0.26$ (Cameroon)

→ Clustering analysis

Cameroon: number of cluster $K = 3$
Previous studies related to the populations genetic structure of *Mycosphaerella fijiensis*

- Within a production area (few 100 Km long)

Isolation by distance analysis with discrete populations

\[ y = 0.0331x - 0.0675 \]

<table>
<thead>
<tr>
<th>Log distance (km)</th>
<th>( \frac{F_{ST}}{1 - F_{ST}} )</th>
<th>Genetic Differentiation</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>3.5</td>
<td>0.1</td>
<td>0.15</td>
</tr>
<tr>
<td>4</td>
<td>0.2</td>
<td>0.3</td>
</tr>
<tr>
<td>4.5</td>
<td>0.3</td>
<td>0.35</td>
</tr>
<tr>
<td>5</td>
<td>0.4</td>
<td>0.45</td>
</tr>
</tbody>
</table>

\[ \text{Cameroon} \]

\[ \text{Costa Rica} \]

\[ \text{P mantel} = 0.0268 \]

\[ \text{Direct method: } d \ll 1 \text{ Km} \]
Previous studies related to the populations genetic structure of *Mycosphaerella fijiensis*

- Within a production area (few 100 Km long)

**Conclusions of this first study**

- IBD Analysis on a distance too long in regard with dispersal capacity of the pathogen (geographical scale considered still too big)
  - Effects of barriers
  - Effects of demographic events

- Unrealistic genetic model (discrete populations)
Population analysis in *M. fijiensis* at a local scale (50x50 Km)

New 2D & 3D Sampling (Cameroon)

IBD in a continuous population analysis

Estimation of dispersal parameters

To delimit pathogen populations & to detect barriers to gene flow

Not presented today…New analysis still under construction
Population analysis in *M. fijiensis* at a local scale (50x50 Km)

**IBD in a continuous population analysis**

- Dispersal parameters estimation along a continuous population

- 90 sites of sampling: 1Km, 250 m, 50 m

- 2-6 isolates/sites genotyped with 20 microsatellites markers

- Continuous population IBD method (haploïds) *(Rousset, unpublished)*
IBD in a continuous population analysis

- No IBD signal detected along the transect

\[ b \approx \frac{1}{2\pi D \sigma^2} \]

\[ \sigma^2 = \text{average squared axial parent-offspring distance} \]

\[ D = \text{Density} \]

\[ F(1-F) = f (\ln(d)) \]

\[ y = 0.004x + 0.0097 \]
Population analysis in *M. fijiensis* at a local scale (50x50 Km)

**IBD in a continuous population analysis**

How to explain an absence of IBD signal? :

- High level of migration
- Lack of statistical power
- Method of IBD inappropriate to pathogenic fungus populations specificity

In particular: High effective size
Population analysis in *M. fijiensis* at a local scale (50x50 Km)

**IBD in a continuous population analysis**

- **Lack of statistical power?**

  → Individual-based simulation model, spatially-explicit, based on the coalescence theory (*Leblois et al. 2003*)

  → Simulation of our sampling

    (transect 33 km, X,Y,Z)

  ⇒ Statistical power related to our sampling:

    slope of $10^{-3}$ statistically detectable
Population analysis in *M. fijiensis* at a local scale (50x50 Km)

**IBD in a continuous population analysis**

- **High effective size**

  \[ \text{Slope} = \frac{1}{2\pi N \sigma^2} \]

  ![Graph showing effective size vs. \(\sigma\) with detectable slopes of \(10^{-3}\) and \(10^{-4}\)]
Conclusions and prospects...

- IBD model might not be adapted to described dispersal processes and gene flow in *M. fijiensis*
  
  → Test new methods to infer parameters related to dispersal processes (migration-selection models, estimation of dispersal curve, new direct methods ?…)

- Realize the spatialized genetic clustering analyse with a landscape genetics method (No *a priori* definition of populations)
  
  → Understand the genetic structure of *M. fijiensis* through an agricol landscape
  → Detect and define some eventual barriers to gene flow