Genomic tools to prevent accidental introductions of forest invasive alien species: The case of the Asian gypsy moth

Michel Cusson
Natural Resources Canada, Quebec City
- Background
- Measures taken to reduce risks of accidental introductions
- The identification challenge: a genomics-based solution
  - Individual samples
  - Bulk samples
- Determining the geographic origins of intercepted samples
  - Clues from mitochondrial genomes
  - Clues from nuclear genomes
- Working out the genetics of female flight capability
Forest pest (insect or pathogen) introduced (or presenting a high potential of being introduced) somewhere outside its natural range

Possible consequences:

- Rapid propagation and population outbreak
- Losses in wood volume (forest industry)
- Degradation of urban landscape
- Constraints imposed on exports
European gypsy moth (EGM)
(Spongieuse européenne/Bombyx disparate)

*Lymantria dispar dispar*

- Introduced from Europe in Massachusetts, 1869
- Wide host range, mainly broad-leaf trees
- Importance: commercial and urban forestry
- Losses & mngt costs: $3.2 billions/year in NA

European gypsy moth: regulated areas in North America

http://www.mda.state.mn.us/gmquarantine
Asian gypsy moth (AGM)
(Spongieuse asiatique)
*Lymantria dispar asiatica* & *L. d. japonica*
The Asian gypsy moth threat

AGM: name designating a complex of species and subspecies

- *Lymantria dispar asiatica*
- *Lymantria dispar japonica*
- *Lymantria umbrosa*
- *Lymantria postalba*
- *Lymantria albescens*

Present in Japan

Biological traits that are a source of concern

- Host range twice as broad as that of EGM (600 vs 300 hosts)
- Overwintering diapause of eggs completed earlier
- Females are flight-capable
- Eggs can be laid on inert surfaces (e.g. ships)
Reducing the risks of accidental introductions

**Within the North American continent**

- Public awareness campaigns about insects as “illegal immigrants”
- Establishment of regulated areas
- Networks of pheromone traps
- “Slow The Spread” (STS) program

**From outside North America**

- Certification program for incoming ships
- Networks of pheromone traps around ports
- Vessel inspections
- Pest identification
- Cleaning of infested vessels
- Eradication operations following introductions
Reducing the risks of accidental introductions

Within the North American continent

• Public awareness campaigns about insects as “illegal immigrants”
• Establishment of regulated areas
• Networks of pheromone traps
• “Slow The Spread” (STS) program

From outside North America

• Certification program for incoming ships
• Networks of pheromone traps around ports
• Vessel inspections
• **Molecular tools for pest identification**
• Cleaning of infested vessels
• Eradication operations following introductions
Procedures to limit the likelihood of AGM introductions

1. The CFIA (Canada) and APHIS (USA) require inspection and cleaning of North America-bound ships

Targeted countries: China, Russia, Japan, Korea
Procedures to limit the likelihood of AGM introductions

2. Second inspection by operators before entry of vessel into Canadian or US waters

3. Inspection of vessels in Canadian (CFIA) and US (CBP) ports

4. Infested vessels must leave the port and be cleaned
Problems associated with AGM identification

- Most common developmental stage found on ships: egg
- AGM eggs are impossible to distinguish from those of other gypsy moths, including EGM
- Molecular identification method used by the Canadian Food Inspection Agency (CFIA):
  - Takes over two days to run
  - Limited reliability
  - Limited scope

The CFIA wished to have a more rapid and reliable method for AGM identification
Development of a TaqMan assay for AGM diagnostics
Research Team – GAPP project, AGM component

NRCan-CFS
Michel Cusson
Don Stewart
Catherine Béliveau
Madjid Djoumad
Josyanne Lamarche
Audrey Nisole
Amélie Potvin
Sandrine Picq

U. Laval - IBIS
Roger Levesque
Luca Freschi

UBC
Richard Hamelin
Hester Yueh
Nicolas Feau
Dario Ojeda

CFIA
Cameron Duff
Reza Zahiri
Brittany Day
Dave Holden
Troy Kimoto

Genomic Applications Partnership Program
(GAPP)
A gypsy moth molecular ID tool: the CFIA’s wish list

The tool should be capable of:

1. Distinguishing EGM from insects of the AGM complex
2. Distinguishing AGM complex species/subspecies from one another
3. Distinguishing AGM and EGM from five other lymantriines presenting an important risk for Canada

*L. monacha* (“Nun moth”; conifer defoliator)

*L. mathura* (“Pink gypsy moth”; broad-leaf defoliator)

*L. lucescens* (“Lucescens tussock moth”; broad-leaf defoliator)

*L. xyлина* (“Casuarina moth”; broad-leaf defoliator)

*L. fumida* (“Red-bellied tussock moth”; defoliator of fir and larch)
Approach

Use genomics tools for the identification of “SNP” markers (“Single Nucleotide Polymorphism”)
Technology used for assay development: real-time PCR ("qPCR")

Principle of qPCR and TaqMan assays
Assay principle: discriminatory annealing

1. Discrimination through primers

Polymerase

<table>
<thead>
<tr>
<th>G</th>
<th>T</th>
<th>T</th>
<th>C</th>
<th>C</th>
<th>G</th>
<th>A</th>
<th>T</th>
<th>C</th>
</tr>
</thead>
</table>
A | G | A | T | C | T | C | A | G |
C | A | G | C | A | T | A | T | C |
T | C | T | G | G | C | A | A | G |
G | G | C | A | A | G | G | C | T | A | G | C | T |

Lymantria dispar asiatica

DNA (AGM)

2. Discrimination through probes

Probe

<table>
<thead>
<tr>
<th>G</th>
<th>T</th>
<th>T</th>
<th>C</th>
<th>C</th>
<th>G</th>
<th>A</th>
<th>T</th>
<th>C</th>
</tr>
</thead>
</table>
G | G | C | A | A | G | G | C | T | A | G | C | T |

Lymantria dispar asiatica

DNA

Probe

<table>
<thead>
<tr>
<th>G</th>
<th>T</th>
<th>T</th>
<th>C</th>
<th>C</th>
<th>G</th>
<th>A</th>
<th>T</th>
<th>C</th>
</tr>
</thead>
</table>
G | G | C | A | A | G | G | C | T | A | G | C | T |

Lymantria dispar dispar
Example of qPCR run targeting discrimination between *L. d. asiatica* and *L. d. dispar*
Search for marker genes

1. Search for existing *Lymantria* markers in public databases (e.g., NCBI, BOLD)

2. Amplification and sequencing of specific marker genes from our sample collection

3. Sequencing and assembly of mitochondrial genomes from several species and subspecies

4. Sequencing and assembly of nuclear genomes from: *L. d. dispar, L. d. asiatica, L. d. japonica* and *L. mathura.*
Choice of marker genes for assay development

In the end, the assay is based on two genes:

- Mitochondrial: COI (two regions)
- Nuclear: FS1

<table>
<thead>
<tr>
<th>5’ end</th>
<th>Barcoding region (“COI-5P”)</th>
<th>3’ end</th>
</tr>
</thead>
</table>

Cytochrome oxidase c subunit 1 (COI): 1531 bp
The AGM molecular assay:
built like a dichotomous identification key

Does the small animal have legs?

- Yes
  - Has it got wings?
    - No
      - Has it got a shell?
        - Yes
          - Is it active at night?
            - No
              - Worm
            - Yes
              - Snail
        - No
          - Yes
            - Centipede
    - Yes
      - Moth
      - Butterfly
- No
  - Yes
    - Spider
Molecular key for the identification of gypsy moths and other lymantriines of biosecurity concern

Stewart et al. PLoS ONE 2016
Is it AGM?

Unknown DNA sample

Does the sample have the MS of L. alboviresens?

Yes

L. alboviresens or L. postalba

No

Does the sample have the MS of other AGM species?

Yes

No

Does the sample have the MS of Asian L. dispar subspecies?

Yes

No

Does the sample have the MS of L. dispar asiatica?

Yes

No

L. dispar asiatica

L. dispar japonica

EGM assays

Stewart et al. PLoS ONE 2016
AGM assays → Is it EGM?

- Does the sample have the MS of *L. dispar dispar* (EGM)?
  - Yes
    - *L. dispar dispar*
  - No
    - *L. dispar dispar* (NN: flightless)

- Does the sample show signs of Asian introgression?
  - No
    - *L. dispar dispar* (NN: flightless)
  - Yes
    - Is it homozygous?
      - Yes
        - *L. dispar dispar* (AA: homozygous)
      - No
        - *L. dispar dispar* (AN: heterozygous)

OTLS assays
Use of the FS1 marker to address the issue of Asian introgression into *L. d. dispar*

- Some specimens identified as *L. d. dispar* using mt markers may, in fact, have flight-capable females, due to hybridization near the subspecies geographical boundaries.

- Such insects have been identified in central Asia, Siberia and Lithuania (Keena et al. 2007, 2008).

- In spite of their *L. d. dispar* mt signature, these insects are as much a biosecurity concern as *bona fide* AGM.
A nuclear marker, FS1, has Asian and North American alleles that can be used to diagnose introgression.
Design of FS1 probes for each allele

Homozygous FS1 - NN

Homozygous FS1 - AA

Heterozygous FS1 - NA

Stewart et al. PLoS ONE 2016
FS1 genotype of 30 *Lymantria dispar* specimens

Molecular key for the identification of gypsy moths and other lymantriines of biosecurity concern

Is it AGM?
- Unknown DNA sample
  - Does the sample have the MS of *L. oleracea*?
    - Yes
      - *L. oleracea* or *L. passalina*
    - No
      - Does the sample have the MS of other AGM species?
        - Yes
          - *L. unipuncta*
        - No
          - Does the sample have the MS of Asian *L. dispar* subspecies?
            - Yes
              - *L. dispar dispar* is the MS of *L. dispar*.
            - No
              - Does the sample show signs of Asian introgression?
                - Yes
                  - *L. dispar dispar* (AA, heterozygous)
                - No
                  - *L. dispers dispers* (A, homozygous)

Is it EGM?
- Does the sample have the MS of *L. dispar dispar* (EGM)?
  - Yes
    - *L. monacha*
  - No
    - Does the sample have the MS of *L. juminda*?
      - Yes
        - *L. juminda*
      - No
        - Does the sample have the MS of *L. mothura*?
          - Yes
            - *L. mothura*
          - No
            - *L. xylina*

Is it another invasive *Lymantria* species?
- Does the sample have the MS of other invasive *Lymantria*?
  - Yes
    - *L. monacha*
  - No
    - Does the sample have the MS of *L. juminda*?
      - Yes
        - *L. juminda*
      - No
        - Does the sample have the MS of *L. xylina*?
          - Yes
            - *L. xylina*
          - No
            - *L. monacha* (ineffective)

Molecular key for the identification of gypsy moths and other lymantriiines of biosecurity concern

Is it OTLS?

L. monacha

Yes

Is it L. monacha?

No

L. fumida

Is it L. fumida?

No

L. xyлина

Yes

Is it L. xyлина?

No

L. lucescens

Yes

Is it L. lucescens?

No

NTS
Development of a multigene assay for detection of AGM in bulk pheromone trap samples
Why a multigene bulk assay for AGM detection?

- Original TaqMan assay (Stewart et al. 2016) was designed for analysis of individual egg samples

- GM monitoring programs using pheromone traps in unregulated areas (e.g. BC) can generate large numbers of moths

- Current procedures rely on the analysis of a subsample of these moths; AGM specimens could be missed

- Detection of a single AGM in a large background of EGM, using qPCR-TaqMan technology, poses a special challenge

- Some of the COI-based probes designed for the single-sample assay were not appropriate for detection in bulk samples
Challenge and approach adopted

- **Problem:** discrimination in some of the original COI-based assays (e.g., “Duplex 1B”) was provided solely by the probe.

  To avoid “drowning” the AGM signal in a “sea” of EGM signals, we need to use discriminatory primers (as opposed to probes), which could not always be designed using COI-based SNPs.

- **Solution:** use other informative SNPs identified through a comparison of full mt genome sequences; confirm inter-individual SNP consistency through resequencing of multiple specimens.

Markers used for discriminating AGM species from EGM:

**CytB:** *L. dispar asiatica/L. dispar japonica*

**ND1:** *L. umbrosa*

**COI:** *L. albescens/L. postalba*
1. Spiking experiment

- **A**
  - L. *albescens* x L. *albescens* + 10⁶ L. *dispar*

- **B**
  - L. *umbrosa* x L. *umbrosa* + 10⁶ L. *dispar*

- **C**
  - L. *dispar asiatica* (LDA) x LDA + 10⁶ L. *dispar*
2. Actual test with moth legs

- **Principle:** 1 AGM leg in 100 Ldd legs *or* 100 Ldd legs alone
- Four different sources of Ldd used: CFL, NDL, VIC and DEN
- Four AGM taxa tested: *Ldae, Ldj, L. umbrosa, L. albescens*
- Total of 8 assays (4 w/4wo AGM legs) run in triplicates
- Each assay run in triplex (FAM, Cy5 and HEX probes)
L. dispar mitochondrial genome sequencing and analysis
Search for markers to identify geographic origins
Comparative genomics analysis revealed many new potentially informative SNPS

Phylogenetic analysis reveals the presence of a novel *L. dispar* variant

Genotyping-by-sequencing-derived SNPs to identify the geographic origins of gypsy moth samples
PCA analysis applied to 2194 neutral SNPs shows strong population structure

96 individuals, 8 populations (6 females et 6 males/pop)

As few as 48 SNPs allow successful assignment of moths to original populations
Sampling locations across the GM’s range: 2017 data set

1122 moths, 71 sites, 23 countries
2018 sampling campaign: increasing coverage of GM’s range

- Add > 15 populations (10-30 males/location; yellow dots)

- Hybrid zone (Russia, Estonia, Kyrgyzstan)
- Asian port regions (Japan, Russian Far East)
- New European locations (Spain, Portugal)

Legend:
- L. d. dispar
- L. d. asiatica
- L. d. japonica
The genetics of flight capability in female gypsy moth
BioSurveillance of Alien Forest Enemies (BioSAFE)

Flight genetics of gypsy moths

Richard Hamelin
Julien Prunier
Isabelle Giguère
Sandrine Picq
Gwylim Blackburn
Ilga Porth
Michel Cusson

Genome Canada
Genome Québec
Genome British Columbia
GM females phenotyped for flight capability

Map showing the distribution of three species of L. d. dispar, L. d. asiatica, and L. d. japonica. The areas marked as flight-capable and flight-incapable are highlighted with green and red, respectively. Locations CT, Li, Gr, Rb, M, BJ, Rm, and Ja are labeled on the map. Photos of Melody Keena and Nathan Havill are also included.
Candidate SNPs for flight capability: isofemale lines

Flight ability

- 5 Flying
- 4
- 3 Gliding
- 2
- 1 Flightless

P

G1

G2

G5
GWAS and outlier analyses of AGM female flight ability

- **GWAS**: Flight strength analysis showing a significant difference between flight-incapable and flight-capable groups.
- **Outlier analysis**: Using $\beta$ statistics to identify outliers in the flight ability data.
- **Functional analysis**: Integrating GWAS results with other functional data to understand the underlying mechanisms.

The map highlights regions with a high number of flight-incapable (red) and flight-capable (green) individuals, with specific populations labeled (e.g., C (40), Rb (30)).
Conclusions

• Genomics-based marker discovery approaches enable the development of rapid and accurate diagnostic assays for FIAS

• Genome-wide-derived markers can be used to assign unknown specimens to their country/region of origin

• Current research is focusing on the identification of the genomic determinants of specific invasive traits, e.g., flight capability
## Acknowledgements

### Collaborators

<table>
<thead>
<tr>
<th>Name</th>
<th>Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Melody Keena</td>
<td>USFS</td>
</tr>
<tr>
<td>Nathan Havill</td>
<td>USFS</td>
</tr>
<tr>
<td>Hannah Nadel</td>
<td>APHIS</td>
</tr>
<tr>
<td>Ken Dewar</td>
<td>U. McGill</td>
</tr>
</tbody>
</table>

### Funding

<table>
<thead>
<tr>
<th>Fund</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genome Canada</td>
</tr>
<tr>
<td>Genome British-Columbia</td>
</tr>
<tr>
<td>Genome Québec</td>
</tr>
<tr>
<td>Genomics Research and Development Initiative</td>
</tr>
</tbody>
</table>