Considerations for environmental risk assessment of gene drive organisms: moving from theory to practice

Keith R Hayes, Geoffrey Hosack, Jeffrey Dambacher, Jessica Ford, Adrien Ickowicz, Scott Foster & David Peel
### NASEM (2016) step by step guide

**BOX 5-1: Example Activities to be Performed during Each Testing Phase**

#### Phase 0: Research Preparation
- Develop a Target Product Profile
- Identify and plan for regulatory requirements
- Use models to inform standards, thresholds of acceptance, and study design
- Establish site-selection criteria (if research includes phase 2-4 trials)
- Identify risk assessment needs
- Identify appropriate confinement and containment strategies

#### Phase 1: Laboratory-Based Research
- Acquire required laboratory regulatory approvals
- Develop containment strategies
- Develop mitigation strategies
- Detect and measure off-target effects
- Optimize design of guide RNAs (when using CRISPR/Cas9-based gene drive)
- Utilize an optimized endonuclease with high cutting efficiency and accuracy
- Optimize for the use of homology-directed repair versus non-homologous end joining to maximize precision of editing
- Evaluate effects on organismal fitness in the presence of the gene drive
- Evaluate gene drive stability over multiple generations
- Mark gene drive organisms
- Use quantitative and computational methods
- Set baseline population-level effects

#### Phase 2: Field-Based Research
- Acquire site-specific regulatory approvals
- Validate efficacy
- Validate population-level effects
- Estimate impact on selected non-targets

#### Phase 3: Staged Environmental Release
- Acquire site-specific regulatory approvals
- Conduct monitoring and surveillance for efficacy
- Conduct monitoring and surveillance for harms

#### Phase 4: Post-Release Surveillance
- Acquire regulatory approvals
- Conduct monitoring and surveillance
- Measure impact

**Key Activities:***
- Use models to inform standards, thresholds of acceptance and study design
- Evaluate effects on organismal fitness
- Evaluate gene drive stability over multiple generations
- Validate population effects
- Estimate impact of selected non-targets

---

Keith Hayes, NAS-ILSIRF, July 2017: Slide 2 of 15
Levels and sources of uncertainty

Source: Speigelhalter and Riesch (2011)
## Identifying risk assessment needs

<table>
<thead>
<tr>
<th>Community concern</th>
<th>Potential endpoint</th>
</tr>
</thead>
<tbody>
<tr>
<td>Could humans get sick from eating animals bitten by transgenic mosquitoes</td>
<td>Toxicity or allergenicity</td>
</tr>
<tr>
<td>Could other diseases emerge from the decrease of mosquitoes</td>
<td>Persistence of transgenic mosquitoes</td>
</tr>
<tr>
<td>Will vector competency of transgenic mosquitoes be altered</td>
<td>Probability of enhanced transmission capacity</td>
</tr>
<tr>
<td>How can you monitor persistence of flying animals</td>
<td>Persistence of transgenic mosquitoes</td>
</tr>
<tr>
<td>Sterility not complete and males are able to reproduce and persist</td>
<td>Persistence of transgenic mosquitoes</td>
</tr>
<tr>
<td>Transgenic mosquitoes have different resistance to insecticide</td>
<td>Probability of enhanced insecticide resistance</td>
</tr>
<tr>
<td>Sterility will affect all mosquito species</td>
<td>Persistence of transgenic mosquitoes</td>
</tr>
<tr>
<td>Sterile males will effect non-target species</td>
<td>Persistence of transgenic mosquitoes</td>
</tr>
</tbody>
</table>
Level 1: Unavoidable stochasticity

Source: Thomas et al. (2013)
Level 2: Direct elicitation for limited information

\[ n = 0 \]
Level 2: Bayesian learning with evidence of absence

$n = 300, y = 0$

$n = 1000, y = 0$
Level 3: Models

**STATISTICAL MODELS**
- **REALISM**
- **GENERALITY**
- **PRECISION**

**MECHANISTIC MODELS**
- **REALISM**
- **GENERALITY**
- **PRECISION**

**QUALITATIVE MODELS**
- **REALISM**
- **GENERALITY**
- **PRECISION**

**A** Real world process represented by a statistical model: regression model, time series model, etc...
- Input data
- Response variables

**B** Real world process represented by a set of difference or differential equations: populations, biogeochemical...
- Input data
- Response variables

**C** Real world process represented graphically: loop analysis, qualitative mathematical models.
- Input data
- Input variables

**D** Real world process treated as unknown
- Input variables
- Response variables

**E** Real world process represented by a statistical model or a mechanistic model.
- Response variables
- Observation model
- Hyper-parameters

**Note:**
- Input variables
- Response variables
- Observation model
- Hyper-parameters
- Machine learning techniques
Level 3: Ecosystem effects model structure uncertainty

Source: Hayes et al. (2014)
Level 4: Recognised inadequacies

Fig. 1. Divergence between laboratory and field mosquitoes over time is determined by differences in environmental conditions (climate, nutrition, microbial exposure), and population size. Laboratory colonies will rapidly suffer a dramatic loss of allelic diversity and fitness, while field populations may change over time, while responding and adapting to varying conditions.

Source: Aguilar et al. (2005)
Level 4: Indirect elicitation Human biting rate

See Hosack et al. (2017) for methods
Level 5: Unknown inadequacies

The "unknown unknowns"

- we don’t know what we don’t know
- there is no risk assessment solution
- increasing complexity of the analysis not an appropriate response

Defensive governance regime would

- conduct rigorous analysis of alternative solutions with stakeholders
- perform systematic hazard analysis
- conduct rigorous post-release monitoring
Model assisted monitoring design
References


Thank You

CSIRO Data61

Keith R. Hayes

t  +61 3 6232 5260
e  Keith.Hayes@csiro.au
w  http://people.csiro.au/H/K/Keith-Hayes