A population dynamics model of *Streptococcus pneumoniae* using genomics & Implementing a model translation and visualisation tool

Leonie Lorenz

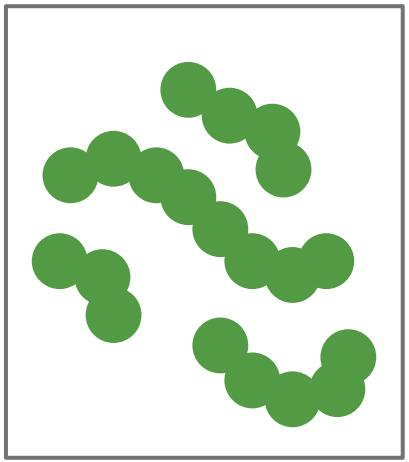
PhD Student | Lees group | EMBL-EBI

24.03.2025

EMBL-EBI

A population dynamics model of *Streptococcus pneumoniae* using genomics

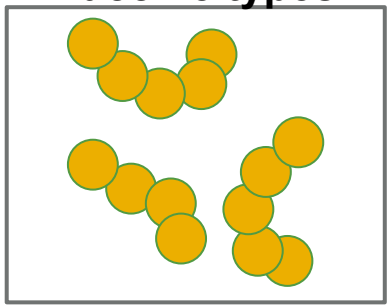
Streptococcus pneumoniae

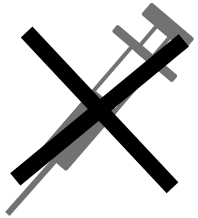


- Serotypically diverse bacteria with more than 100 serotypes
- Genetically diverse with hundreds of lineages/strains
- Major cause of diseases, including pneumonia and meningitis, especially in young children under 2 years of age and adults over 65 years of age
- Treatment: antibiotics
- Prevention: pneumococcal vaccines

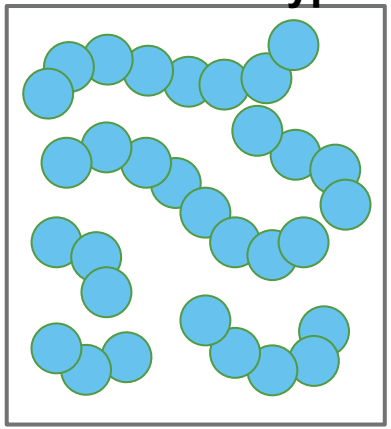
Pneumococcal vaccines

Vaccine types



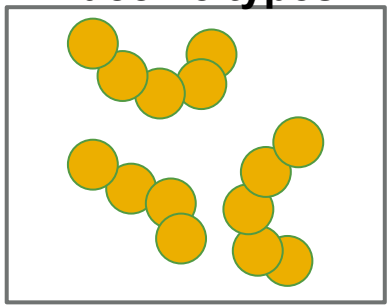


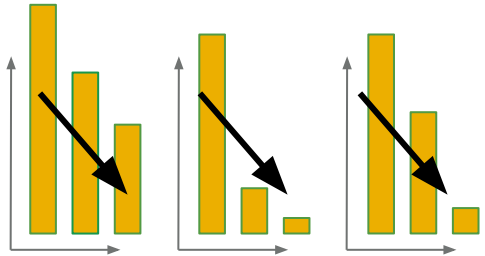
Non-vaccine types

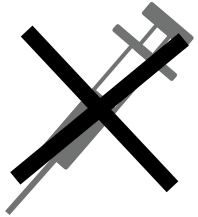


Pneumococcal vaccines

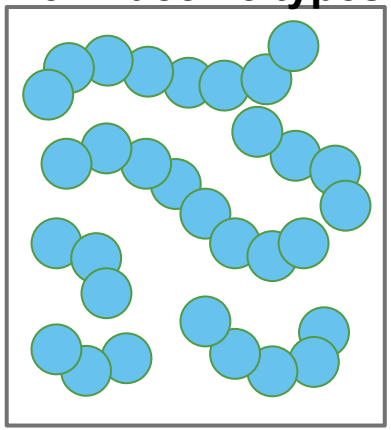
Vaccine types

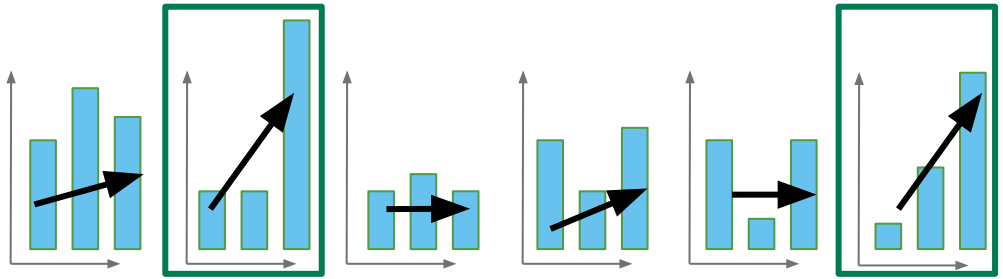






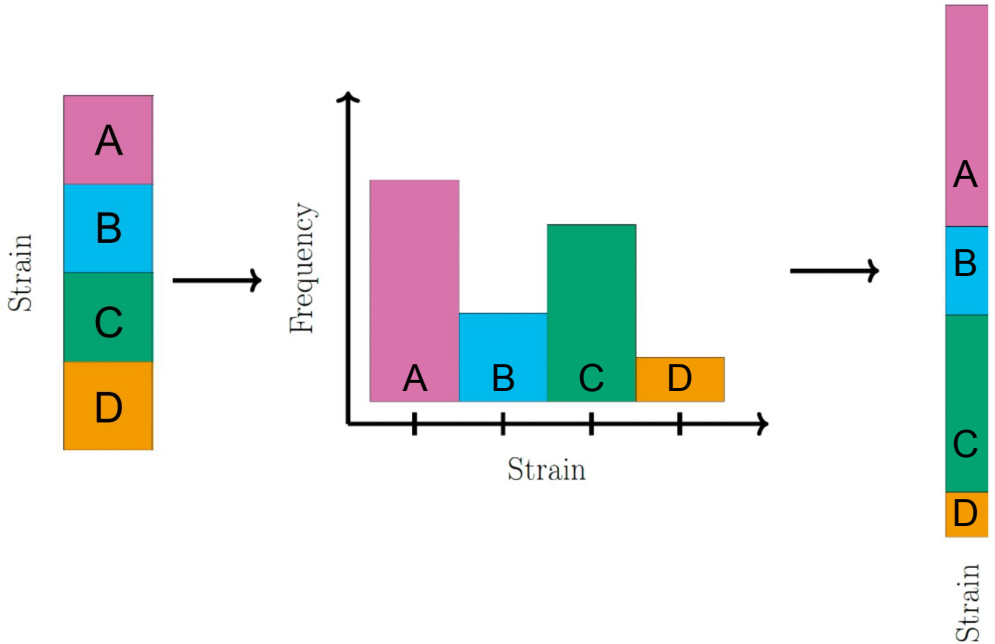
Non-vaccine types



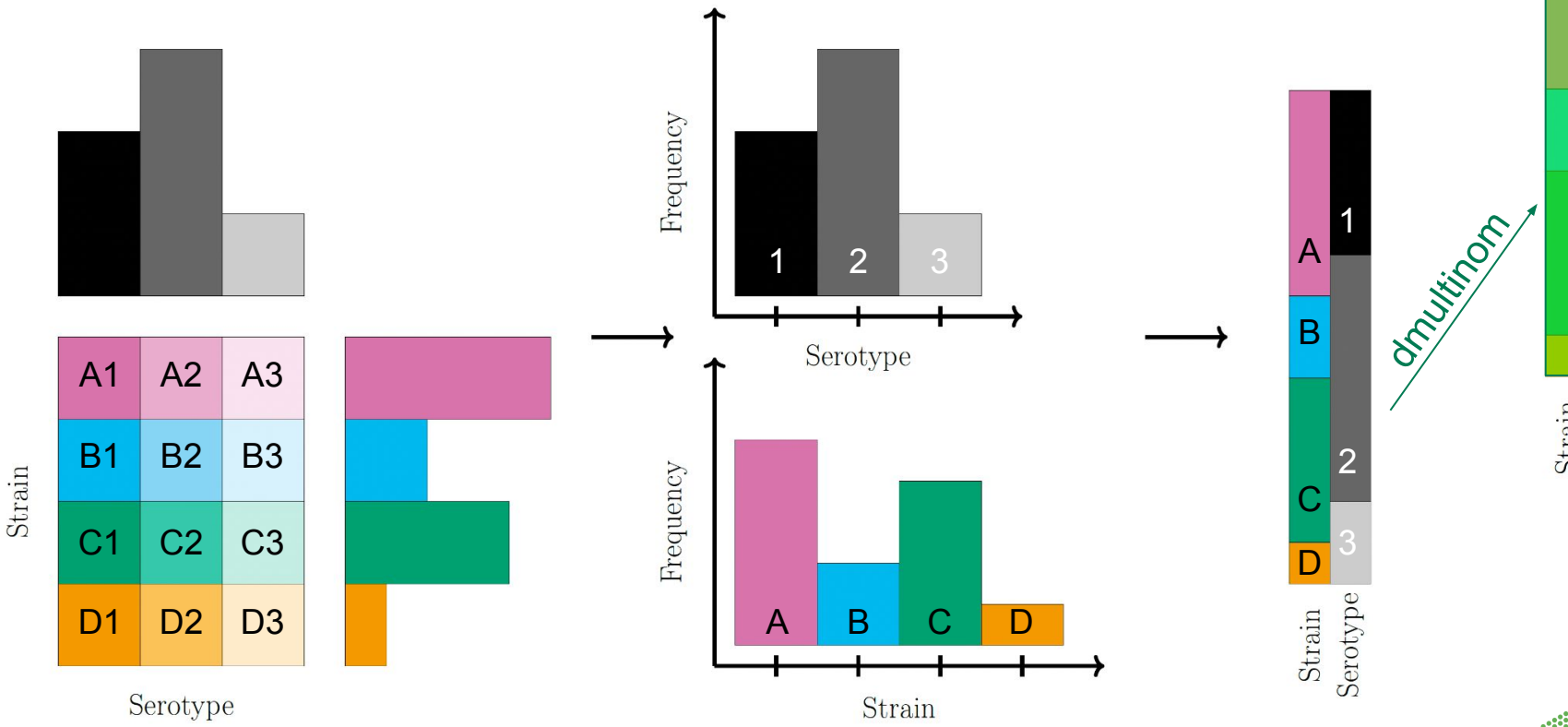


EMBL-EBI

Simple compartmental Model



Compartmental Model with Sub-compartments



Compartmental Model with Sub-compartments

Parameters:

- NFDS σ_f
- $prop_f$
- *vaccination v*
- migration *m*

$$P_{i,j}^{t+1} = B_{i,j}^{t} + M_{i,j}^{t}$$

Population — Births — Migration

$$B_{i,j}^{t} \sim Poisson(K \cdot (\frac{p_{i,j}^{t}}{\sum_{i=1}^{n}\sum_{j=1}^{l} p_{i,j}^{t}}) * (1 - \mathbf{m}))$$

$$p_{i,j}^{t} = (1+\sigma_{\mathbf{f}})^{\pi_{i,t}} * P_{i,j}^{t} * (1 - (v_{\text{time}}(t) * v_{\text{type}}(j) * \mathbf{v}))$$

$$\pi_{i,t} = \sum_{l \in G_i, l \in S} e_l - f_{l,t}$$

$$M_{i,j}^{t} \sim binomial(\hat{m}^t, m_{i,j})$$

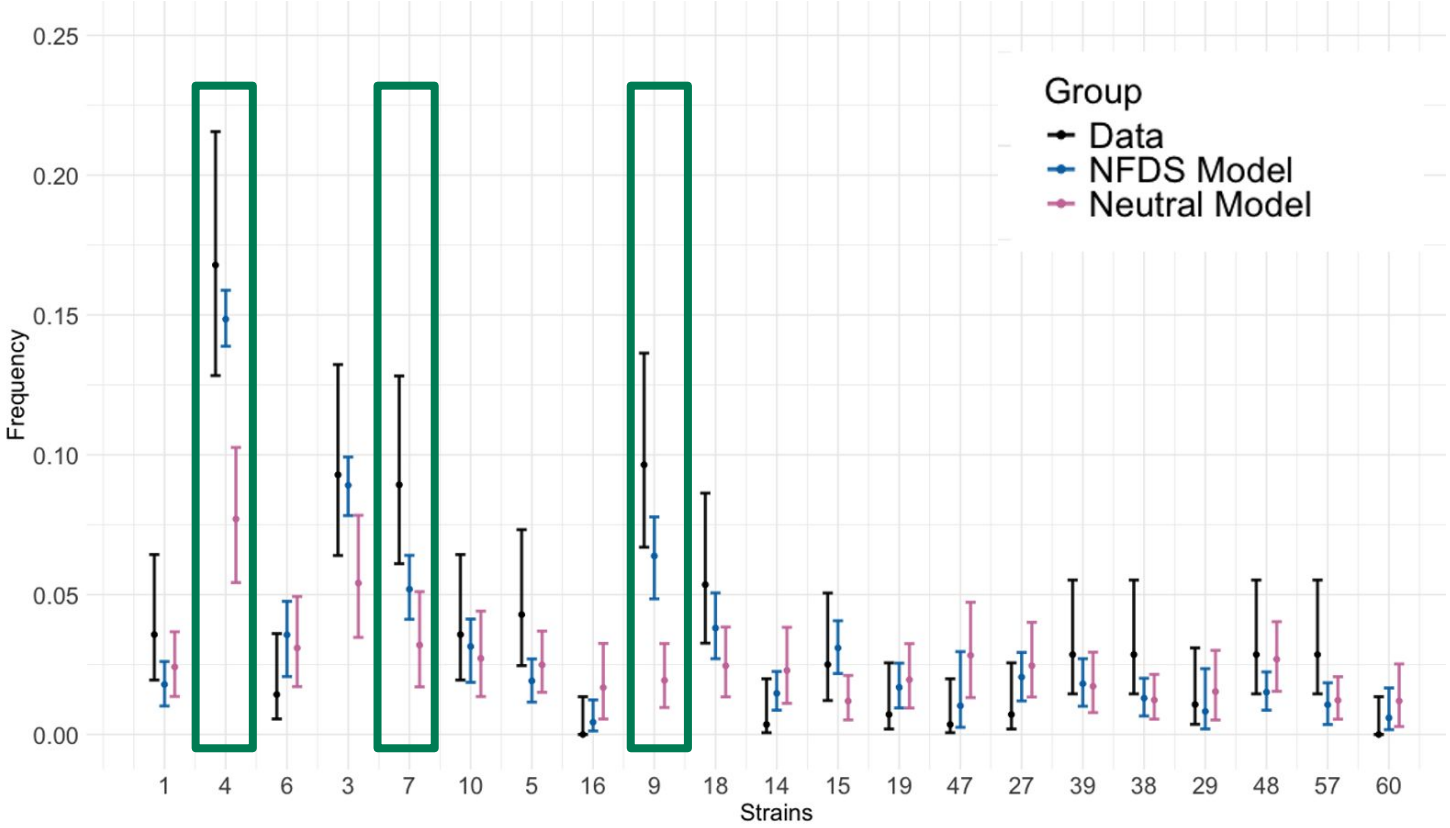
$$\hat{m}^t \sim \text{binomial}(K, \mathbf{m})$$

NFDS = Negative frequency-dependent selection,
Model based on Jukka Corander et al. “Frequency-dependent selection in vaccine-associated pneumococcal population dynamics”. (2017)

```
# identify genes that are under NFDS
pi_f_freq[,] <- if ((delta[i] <= prop_f * gene_no)) Genotypes[i,j] * (eq[i] - freq[i]) else 0
pi_f_genotypes[] <- sum(pi_f_freq[1:gene_no,i])

# Genotype specific probability to produce offspring
# those are the individuals' probabilities multiplied by the number of individual that have this genotype
probs[,] <- (1 + exp(sigma_f))^pi_f_genotypes[i] * Pop[i,j] * (1- (as.integer(time >= vacc_time) * vaccTypes[j] * v))

# generate the next generation based on the current one
y[,] <- if ((probs[i,j]/sum(probs[1:species_no,1:sero_no])) < 1)
  rpois(capacity * (probs[i,j] / sum(probs[1:species_no,1:sero_no])) * (1-exp(m)) ) else rpois(capacity * 1 *(1-exp(m)) )

# m is the migration rate
# fitness of individuals in the community is reduced by this rate
# determining migration number:
mig_num <- rbinom(capacity, exp(m))
Pop_mig[,] <- rbinom(mig_num, migVec[i,j])

## Core equation for population (assume constant size here):
update(Pop[,]) <- y[i,j] + Pop_mig[i,j]
update(Pop_tot[]) <- sum(y[i,]) + sum(Pop_mig[i,])

initial(Pop[,]) <- Pop_ini[i,j] # deterministic, user-based start value
initial(Pop_tot[]) <- sum(Pop_ini[i,])

```

Strain-serotype Model Fit to Massachusetts

Strain-serotype Model Fit to Massachusetts



Strain-serotype Model Fit - Parameter Comparison

Parameter	Massachusetts	Southampton	Nepal
neg. log likelihood	**-236.12234** (-239.9329 , -233.6487)	**−562.76** (-567.6667, -560.4743)	**-2415.2916** (-2419.1133, -2412.7752)
Strength of NFDS σ_f	**0.0345** (0.0144, 0.0685)	**0.0379** (0.0133, 0.0826)	**0.4670** (0.3474, 0.5386)
Proportion of genes under NFDS **prop$_f$**	**0.28835** (0.20115, 0.4566)	**0.1614** (0.1077, 0.3611)	**0.1708** (0.1584, 0.1967)
Migration rate **m**	**0.0133** (0.0087, 0.0214)	**0.0185** (0.0126, 0.0252)	**0.0792** (0.0648, 0.0953)
Vaccination selection strength **v**	**0.0814** (0.0627, 0.1046)	**0.0905** (0.0759, 0.1055)	**0.1766** (0.1405, 0.2159)

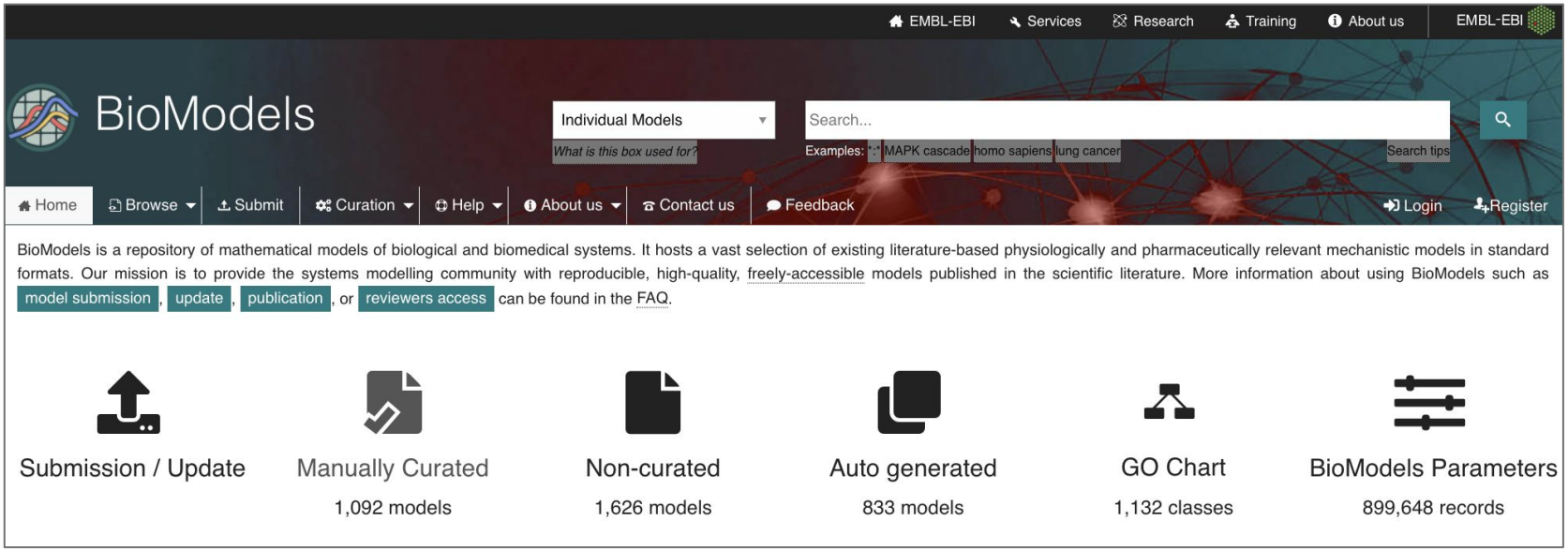
Benefits of implementing this model in odin

- Defining variables as vectors / matrices
- Readability of the model code (looks like equations)
- Reproducibility when fitting the model
- Having a likelihood and using MCMC
- It’s fast

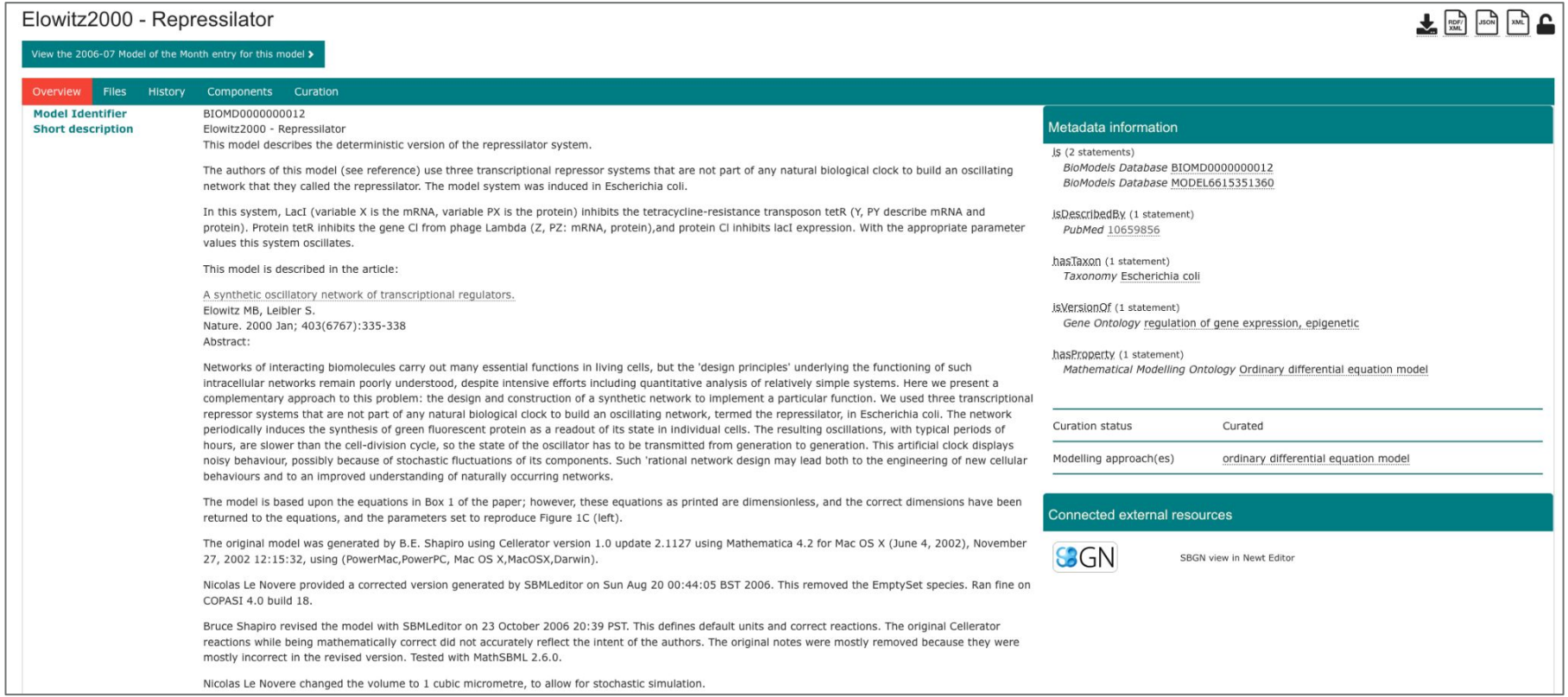
—> Reusability

Implementing a model translation and visualisation tool

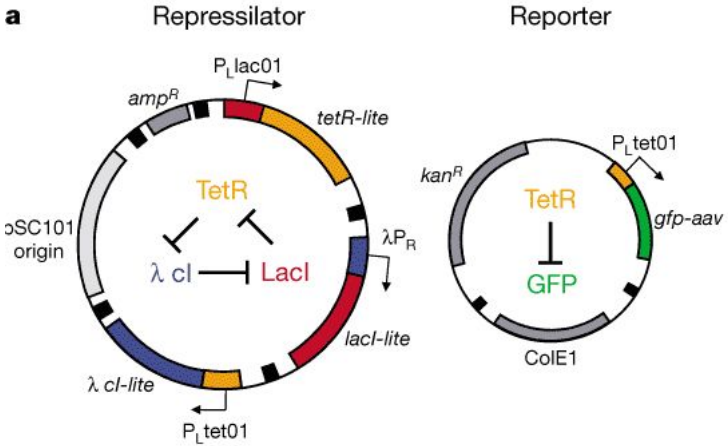
Model Accessibility



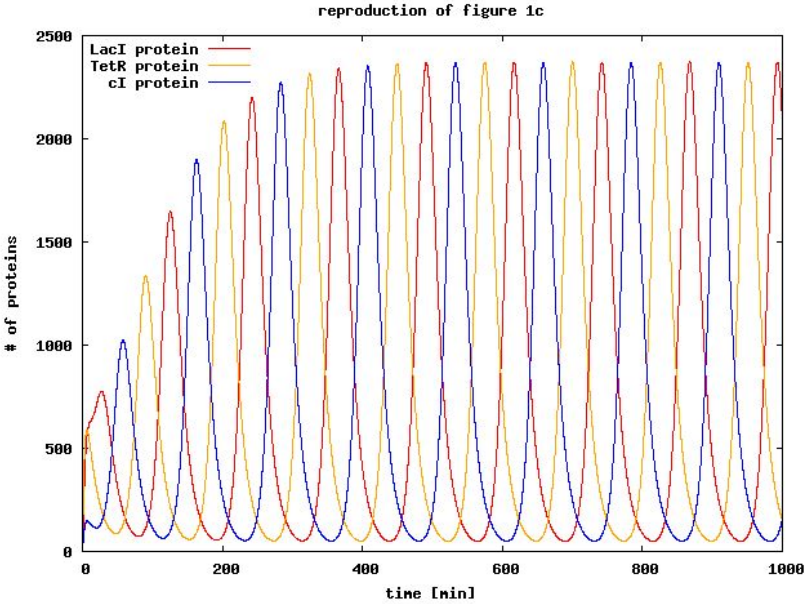
Model Accessibility



System Biology Models



Adapted from Elowitz and Leibler (2000)
DOI: 10.1038/35002125



Taken from BioModels
https://www.ebi.ac.uk/biomodels/BIOMD0000000012#Curation

EMBL-EBI

The pipeline

BioModels

ODE Models in SBML

SBML Systems Biology Markup Language

```
        </lambda>
      </math>
        </functionDefinition>
  </listOfFunctionDefinitions>
  <listOfUnitDefinitions>
    <unitDefinition id="length" name="length">
      <listOfUnits>
        <unit exponent="1" kind="metre" multiplier="1" scale="0"/>
      </listOfUnits>
    </unitDefinition>
    <unitDefinition id="area" name="area">
      <listOfUnits>
        <unit exponent="2" kind="metre" multiplier="1" scale="0"/>
      </listOfUnits>
    </unitDefinition>
    <unitDefinition id="volume" name="volume">
      <listOfUnits>
        <unit exponent="1" kind="litre" multiplier="1" scale="-3"/>
      </listOfUnits>
    </unitDefinition>
    <unitDefinition id="time" name="time">
      <listOfUnits>
        <unit exponent="1" kind="second" multiplier="86400" scale="0"/>
      </listOfUnits>
    </unitDefinition>
    <unitDefinition id="substance" name="substance">
      <listOfUnits>
```

Import BioModels into odin

SBML file

```
      <math xmlns="http://www.w3.org/1998/Math/MathML">
        <lambda>
          <bvar>
            <ci> z </ci>
          </bvar>
          <bvar>
            <ci> beta_t </ci>
          </bvar>
          <bvar>
            <ci> S </ci>
          </bvar>
          <bvar>
            <ci> I </ci>
          </bvar>
          <bvar>
            <ci> N </ci>
          </bvar>
          <apply>
            <divide/>
            <apply>
              <times/>
              <ci> z </ci>
              <ci> beta_t </ci>
              <ci> S </ci>
              <ci> I </ci>
            </apply>
            <ci> N </ci>
          </apply>
        </lambda>
      </math>
          </functionDefinition>
  </listOfFunctionDefinitions>
  <listOfUnitDefinitions>
    <unitDefinition id="length" name="length">
      <listOfUnits>
        <unit exponent="1" kind="metre" multiplier="1" scale="0"/>
      </listOfUnits>
    </unitDefinition>
    <unitDefinition id="area" name="area">
      <listOfUnits>
        <unit exponent="2" kind="metre" multiplier="1" scale="0"/>
      </listOfUnits>
    </unitDefinition>
    <unitDefinition id="volume" name="volume">
      <listOfUnits>
        <unit exponent="1" kind="litre" multiplier="1" scale="-3"/>
      </listOfUnits>
    </unitDefinition>
    <unitDefinition id="time" name="time">
      <listOfUnits>
        <unit exponent="1" kind="second" multiplier="86400" scale="0"/>
      </listOfUnits>
    </unitDefinition>
    <unitDefinition id="substance" name="substance">
      <listOfUnits>
```

https://www.ebi.ac.uk/biomodels/BIOMD0000000982

SBMLtoOdin

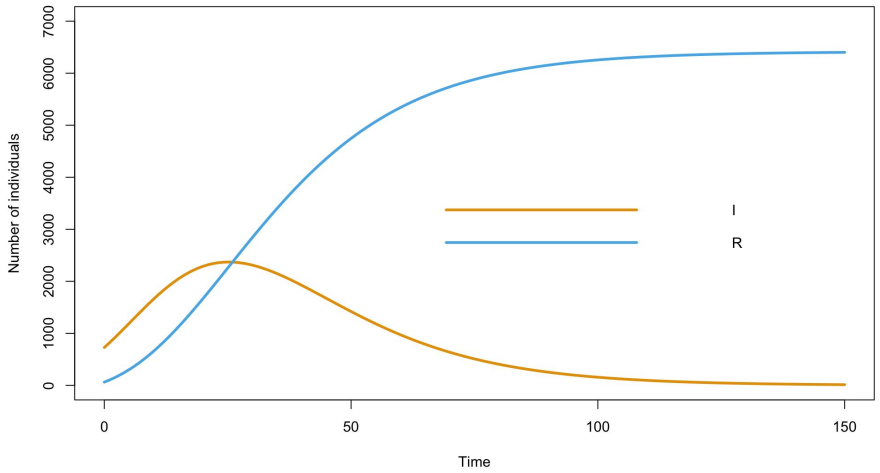
Automatically generated code by SBMLtoOdin

```
1
2  initial(Susceptible) <- Susceptible_init
3  Susceptible_init <- user(32680000)
4  initial(Infected) <- Infected_init
5  Infected_init <- user(728)
6  initial(Removed) <- Removed_init
7  Removed_init <- user(62.0000000000001)
8  delta <- Trigger_Stage_1 * 0.025 + Trigger_Stage_2 * 0.042 + Trigger_Stage_3 * 0.05
9  p <- Trigger_Stage_1 * 0.0784 + Trigger_Stage_2 * 0.045 + Trigger_Stage_3 * 0.0466
10 z <- Trigger_Stage_1 * 0.4374 + Trigger_Stage_2 * 0.3914 + Trigger_Stage_3 * 0.4047
11 N <- Infected + Removed + Susceptible
12 I_total <- Infected + Removed
13 beta <- beta_0 * (1 - p)^( t)
14 deriv(Susceptible) <- 0 - Malayasia * z * beta * Susceptible * Infected / N
15 deriv(Infected) <- 0 + Malayasia * z * beta * Susceptible * Infected / N - Malayasia * delta * Infected
16 deriv(Removed) <- 0 + Malayasia * delta * Infected
17 beta_0 <- user(0.4114)
18 Trigger_Stage_1 <- user(0)
19 Trigger_Stage_2 <- user(0)
20 Trigger_Stage_3 <- user(1)
21 Malayasia <- 1
```

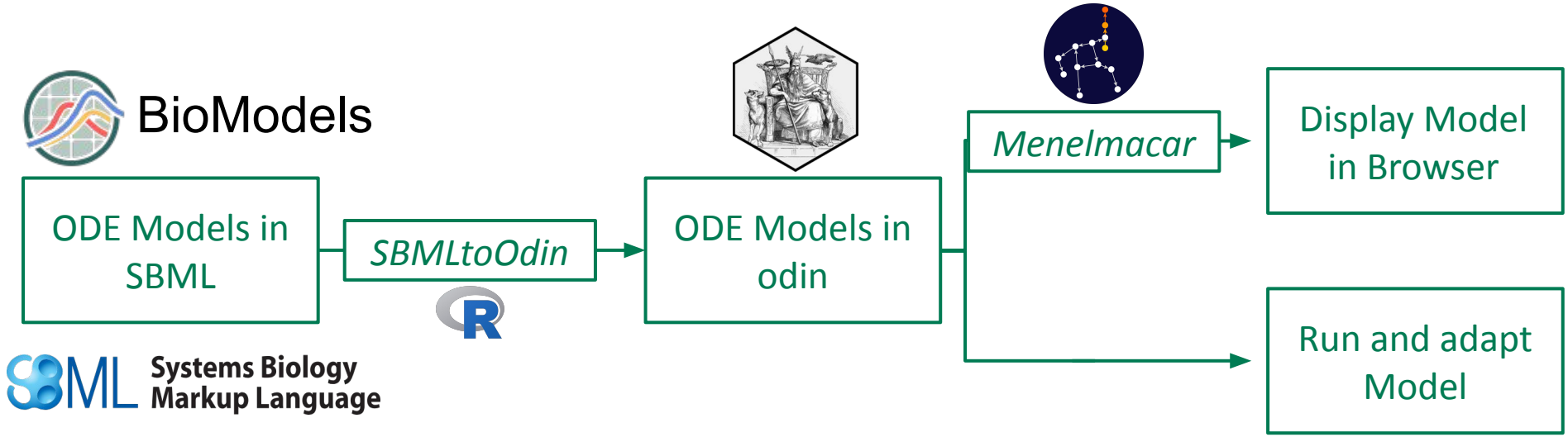
odin

Import BioModels into odin

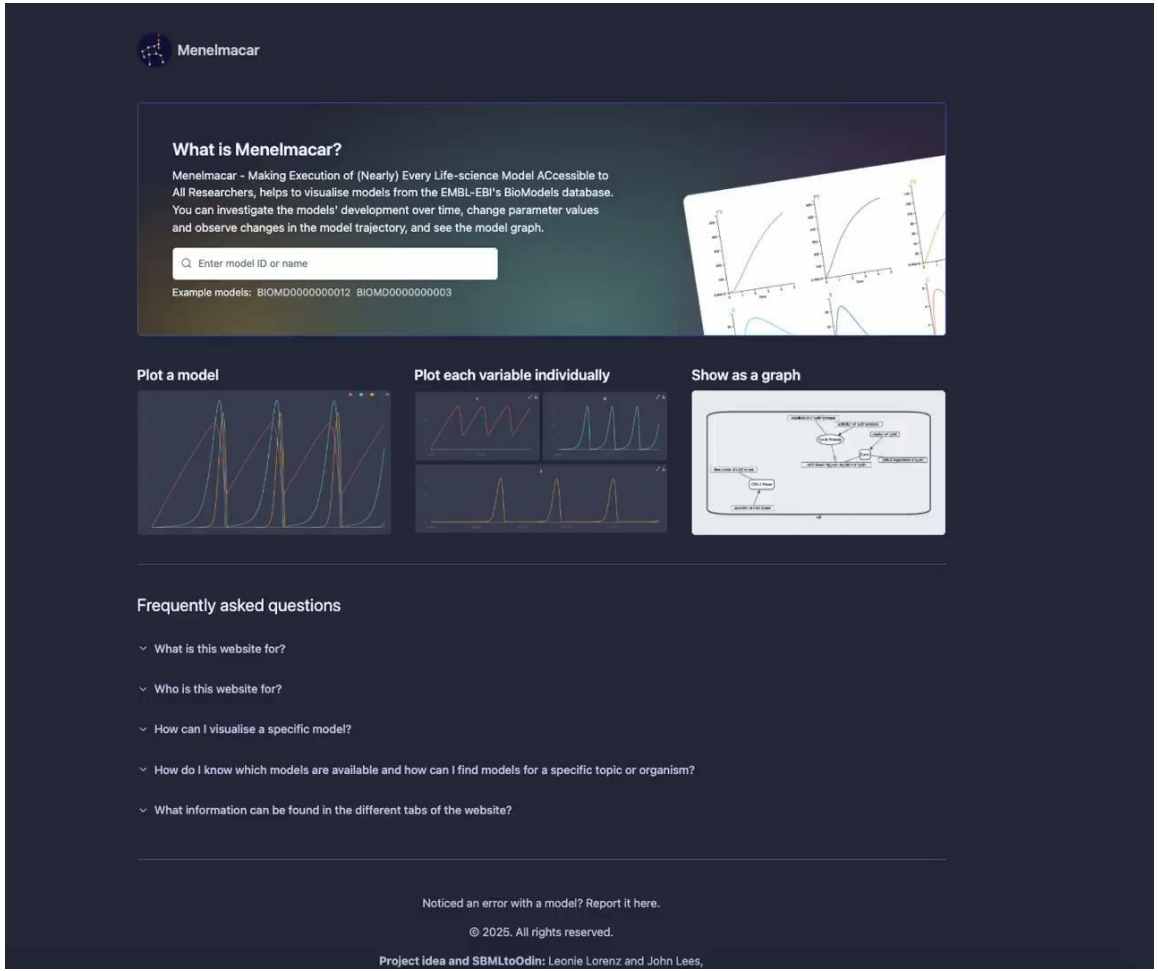
```
# SIR model example (Malaysia)
importSBMLfromBioModels("BIOMD0000000982","../LawModel.R")

model_generator <- odin::odin("../LawModel.R") # creates model
LawModel <- model_generator$new() #creates object, with parameters if needed
LawModel_res <- LawModel$run(0:150) # runs model forward in time, for specified time
# plot simulation
matplot(LawModel_res[, 1], LawModel_res[, c(3,4)], xlab = "Time", ylab = "Number of individuals",
        type = "l", col = c(I = "#E69F00", R = "#56B4E9"), lty = 1, ylim = c(0,7000), lwd = 3)
legend(x = 50, y =4000, lwd = 3, col = c("#E69F00", "#56B4E9"), legend = c("I", "R"), bty = "n")
```



The pipeline



Menelmacar = Making Execution of (Nearly) Every Life-science Model ACcessible to All Researchers


Menelmacar
What is Menelmacar?
Menelmacar - Making Execution of (Nearly) Every Life-science Model ACcessible to All Researchers, helps to visualise models from the EMBL-EBI's BioModels database. You can investigate the models' development over time, change parameter values and observe changes in the model trajectory, and see the model graph.
Enter model ID or name
Example models: BIOMD0000000012 BIOMD0000000003
Plot a model
Plot each variable individually
Show as a graph
Frequently asked questions
What is this website for?
Who is this website for?
How can I visualise a specific model?
How do I know which models are available and how can I find models for a specific topic or organism?
What information can be found in the different tabs of the website?
Noticed an error with a model? Report it here.
© 2025. All rights reserved.
Project idea and SBMLtoOdin: Leonie Lorenz and John Lees,

Code availability

- SBMLtoOdin: https://github.com/bacpop/SBMLtoOdin
- Menelmacar website:
  - Code: https://github.com/bacpop/odinviewer
  - Website: biomodels.bacpop.org
- NFDS Model: https://github.com/bacpop/compartmental-models-odin-dust

THANK YOU

Pathogen Informatics and Modelling Group

John Lees
Joel Hellewell
Jacqueline Toussaint
Hsueh-Chien (Raymond) Cheng
Samuel Horsfield
Johanna von Wachsmann
Víctor Rodríguez Bouza
Tanushree Tunstall
Daniel Anderson
Daria Frolova
Gherard Batisti Biffignandi
Ines Sanchez-Hombria
Andrea Sanchez Serrano
Alireza Tajmirriahi

Collaborators

Nick Croucher
Henrik Salje
Stephanie Lo
Stephen Bentley
Rahuman S. Malik Sheriff
Gomoku Studio :
Andrea Epifani, Zeqing Lu

UNIVERSITY OF CAMBRIDGE
EMBL-EBI