Modelling Epidemics

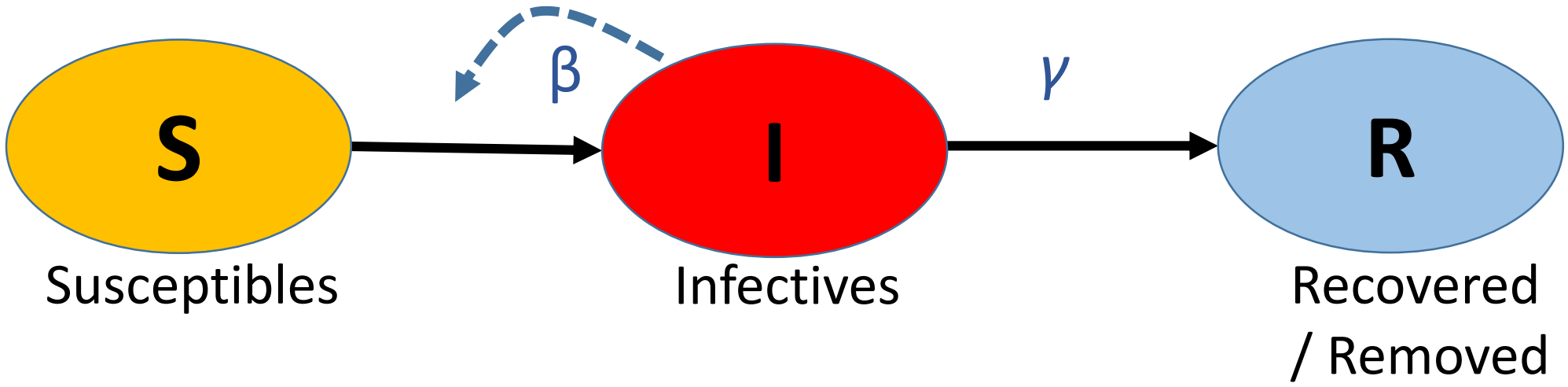
Lecture 6: Heterogeneous deterministic epidemiological models with different levels of susceptibility and infectivity

Andrea Doeschl-Wilson

Overview

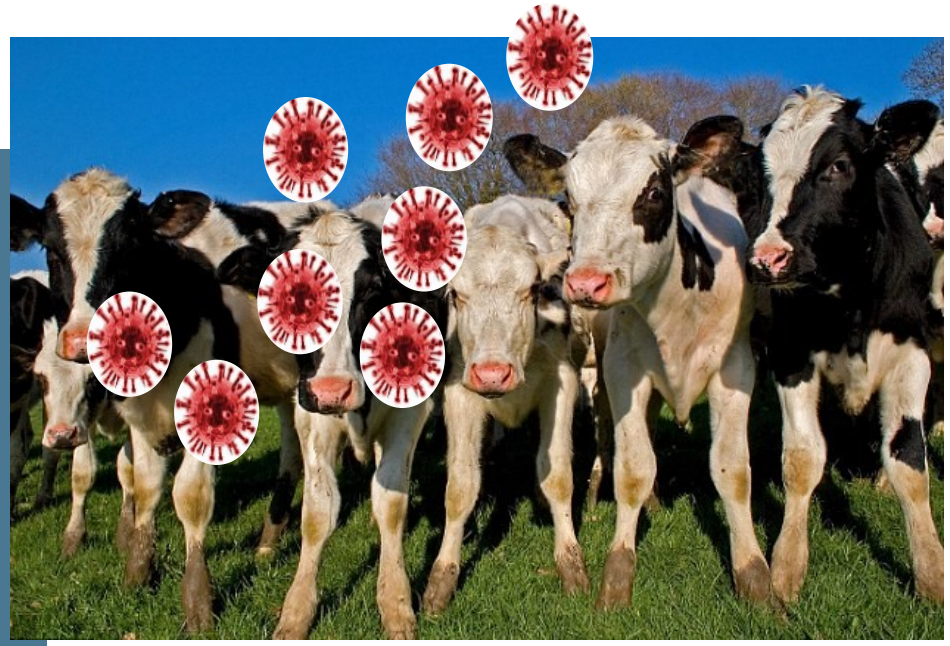
- Implicit assumptions in basic SIR models
- Sources of heterogeneity
- Incorporating genetic heterogeneity
- Case 1: The SIR model for two levels for susceptibility
- Case 2: Difference in susceptibility confers difference in infectivity
- More complex cases of (genetic) heterogeneity

Recap: the basic SIR model



- Assumes that all individuals have the same
  - Risk of becoming infected (susceptibility)
  - Risk of transmitting the infection (infectivity)
  - Duration of infectious period
- Dynamics is fully described by 2 model parameters β and γ and initial proportions of individuals in each compartment

But: not all individuals are the same!



Do these individuals have the same risk of becoming infected or transmitting infections?

Potential sources of host heterogeneity

1. Differences in host environment

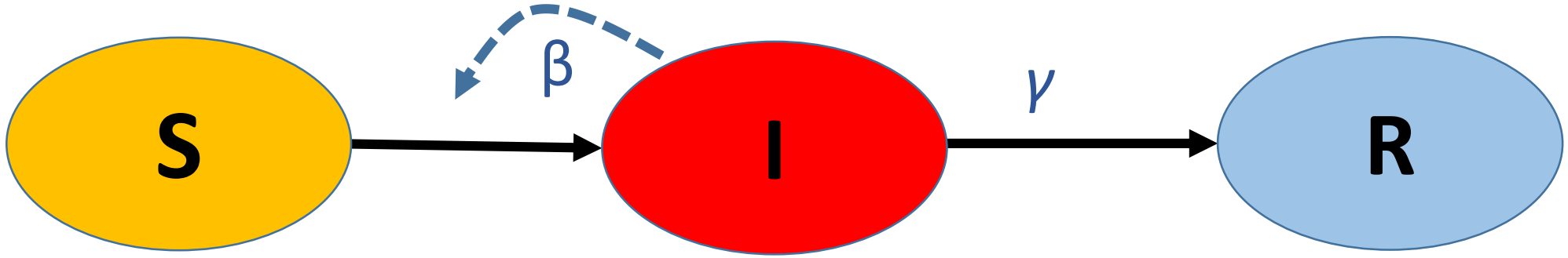
- Rearing system
  - Different rearing systems associated with e.g. different space, nutrition or other environmental factors may affect epidemiological parameters
- Climatic conditions
  - E.g. higher infection pressure for liverfluke or footrot in wet and mild conditions
- Pathogen diversity
  - E.g. exposure to different pathogen strains
  - Co-infections

Potential sources of host heterogeneity (cont.)

2. Differences in host characteristics

- Age
  - Age dependent immunity or contact behaviour may affect transmission and recovery rate
    - e.g. maternal antibodies provide temporal immunity after birth
  - Large body of literature for age-structured epidemiological models
- Immune status
  - E.g. vaccinated / non-vaccinated; previously exposed / not exposed
  - Large body of literature on impact of vaccination on epidemiology
- Host genetics
  - Relatively few studies despite much evidence for genetic heterogeneity
  - Focus in this lecture

How does the host genetics affect the epidemiology?



- Host genetics may affect the
  - Risk of becoming infected (susceptibility): impact on β
  - Risk of transmitting the infection (infectivity): impact on β
  - Ability to interfere with the pathogen life-cycle within the host: impact on duration of infectious period 1/γ

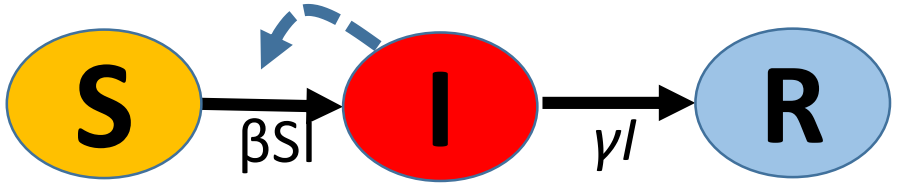
How to incorporate genetic heterogeneity into epidemiological models

- Which host trait is most likely under genetic influence?
  - (e.g. susceptibility, infectivity, time to death / recovery)
- What is the genetic architecture of this trait?
  - Trait primarily controlled by single gene or locus (Major gene model)
    - E.g. gene affecting binding protein
    - Discrete genotypes
  - Trait controlled by many genes, each with a small effect (Polygenic model)
    - Majority of cases
    - Continuous genotypes
  - Different genetic architectures lend themselves to different modelling approaches

How to incorporate genetic heterogeneity into epidemiological models

- Which host trait is most likely under genetic influence?
  - (e.g. susceptibility, infectivity, time to death / recovery)
- What is the genetic architecture of this trait?
  - **Trait primarily controlled by single gene or locus (Major gene model)**
    - E.g. gene affecting binding protein
    - **Discrete genotypes**
  - Trait controlled by many genes, each with a small effect (Polygenic model)
    - Majority of cases
    - Continuous genotypes
  - Different genetic architectures lend themselves to different modelling approaches

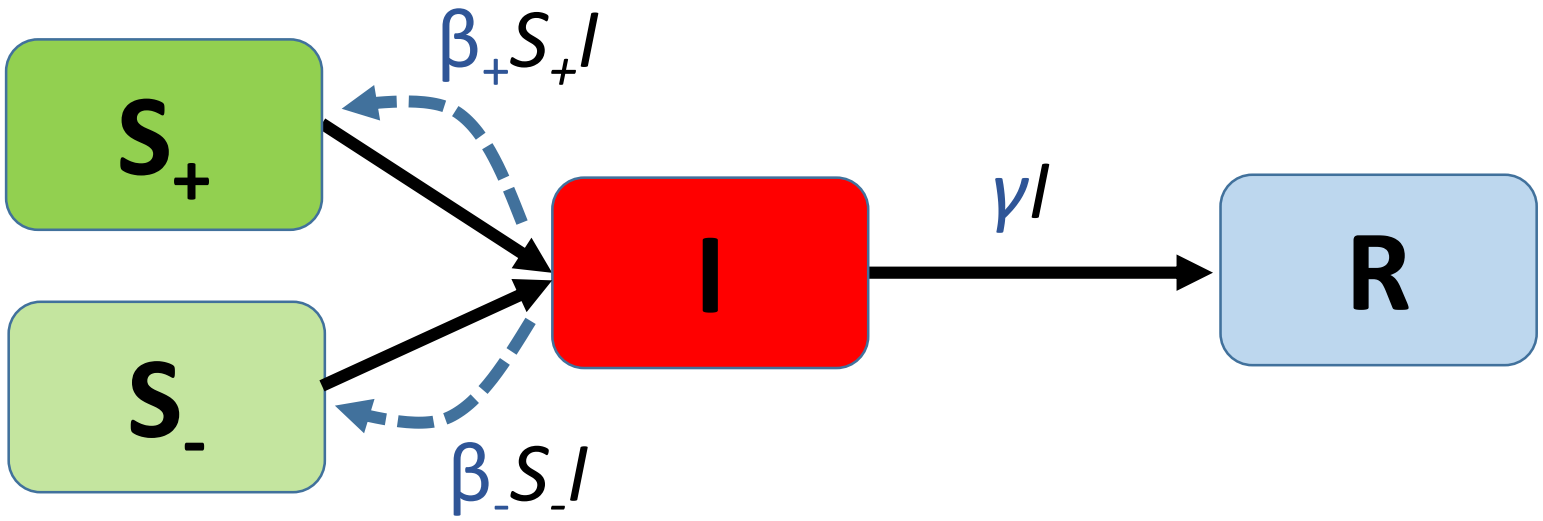
A (very) simple genetic epidemiological model



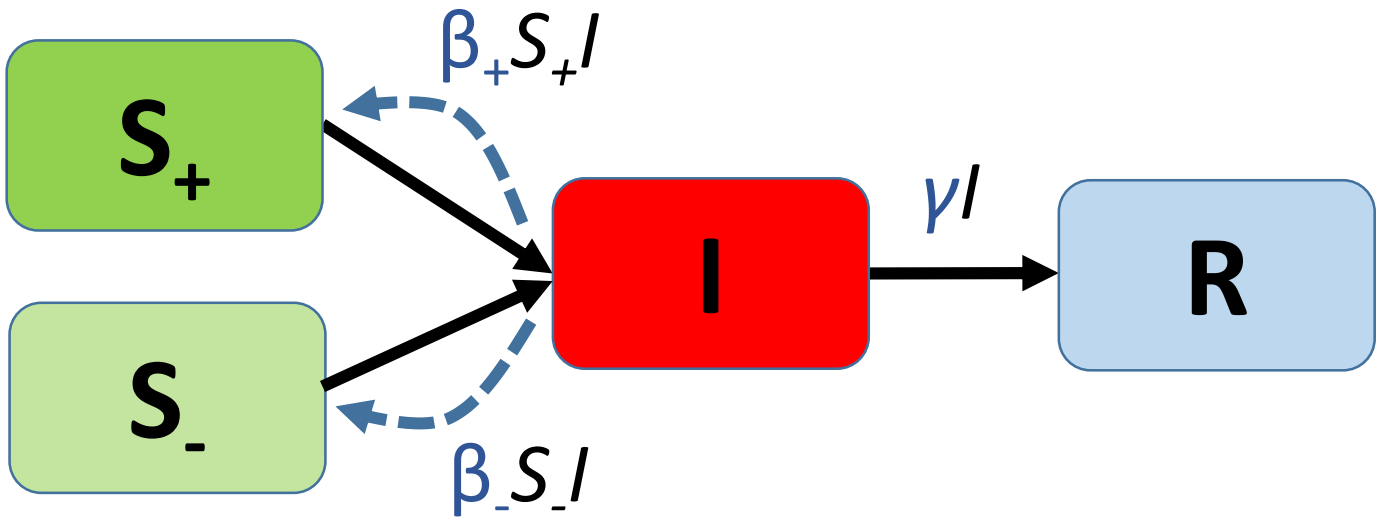
- Assume SIR epidemiological model in a closed population with homogeneous contact structure
- Assume differences in susceptibility only
- For simplicity assume major gene affects susceptibility in a dominant manner, i.e. one copy of the unfavourable allele is sufficient to increase susceptibility (**AA, Aa, aa**)
  - 2 genotypes: **S+** and **S-** representing groups with high and low susceptibility, respectively

A (most) simple genetic epidemiological model

- 2 genotypes: **S+** and **S-** representing high and low susceptibility
- 2 transmission rates: $\beta_+ > \beta_-$



Model equations



ODEs describing the dynamics:

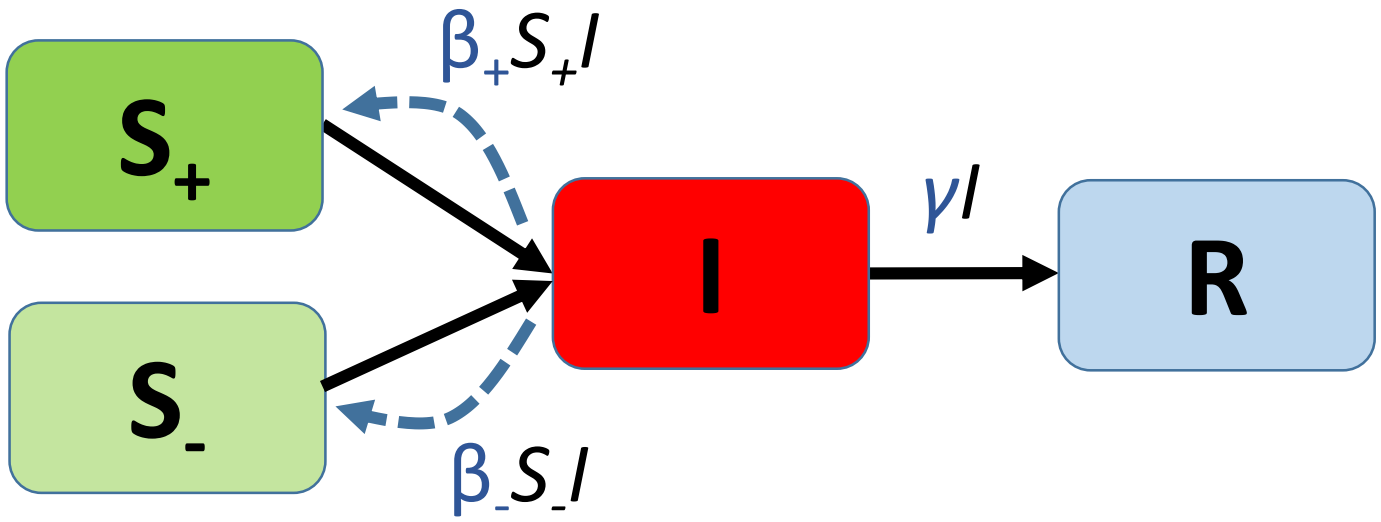
$$\frac{dS_+}{dt} = -\beta_+ S_+ I$$

$$\frac{dS_-}{dt} = -\beta_- S_- I$$

$$\frac{dI}{dt} = (\beta_- S_- + \beta_+ S_+) I - \gamma I$$

$$\frac{dR}{dt} = \gamma I$$

Model equations



Initial conditions:

Specify $S_+(0)$, $S_-(0)$, I(0) and R(0) such that:

- Prop. of individuals with low susceptibility = **p**
- $S_+(0) + S_-(0) + I(0) + R(0) = 1$

e.g. $S_-(0) = \frac{p(N-1)}{N}$; $S_+(0) = \frac{(1-p)(N-1)}{N}$; $\mathrm{I}(0) = \frac{1}{N}$; $\mathrm{R}(0) = 0$

ODEs describing the dynamics:

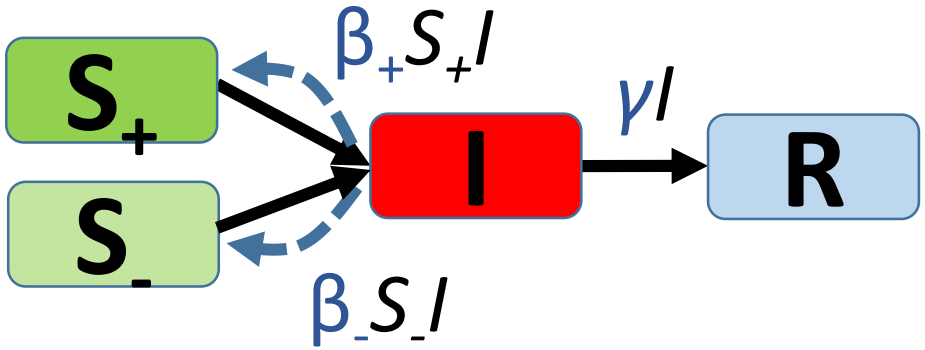
$$\frac{\mathrm{d}S_+}{dt} = -\beta_+ S_+ \mathrm{I}$$

$$\frac{\mathrm{d}S_-}{dt} = -\beta_- S_- \mathrm{I}$$

$$\frac{\mathrm{d}I}{dt} = (\beta_- S_- + \beta_+ S_+)\mathrm{I} - \gamma \mathrm{I}$$

$$\frac{\mathrm{d}R}{dt} = \gamma \mathrm{I}$$

The basic reproductive ratio R_0

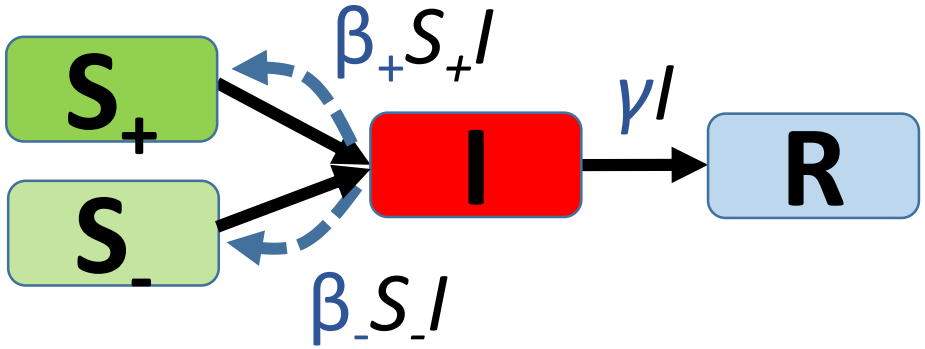


Definition: Basic reproductive ratio R_0 (same as for homogeneous populations)
The average number of individuals an infectious individual will infect, assuming that the rest of the population is susceptible

- **Assume $\beta_+ = \beta$; $\beta_- = \varepsilon\beta, with\ 0 \leq \varepsilon \leq 1$**
- An average infected individual
  - infects $\beta_- S_- + \beta_+ S_+ = \varepsilon\beta p + \beta(1-p)$ susceptible individuals per day
  - is infectious for a period of 1/γ days
  - will thus generate $(p\varepsilon\beta + (1-p)\beta)$ x 1/γ new infections over its lifetime

➢ $$R_0 = \frac{p\,\varepsilon\beta + (1-p)\beta}{\gamma}$$

The Threshold Phenomenon



Imagine a scenario where I_0 infectives are introduced into a susceptible population.

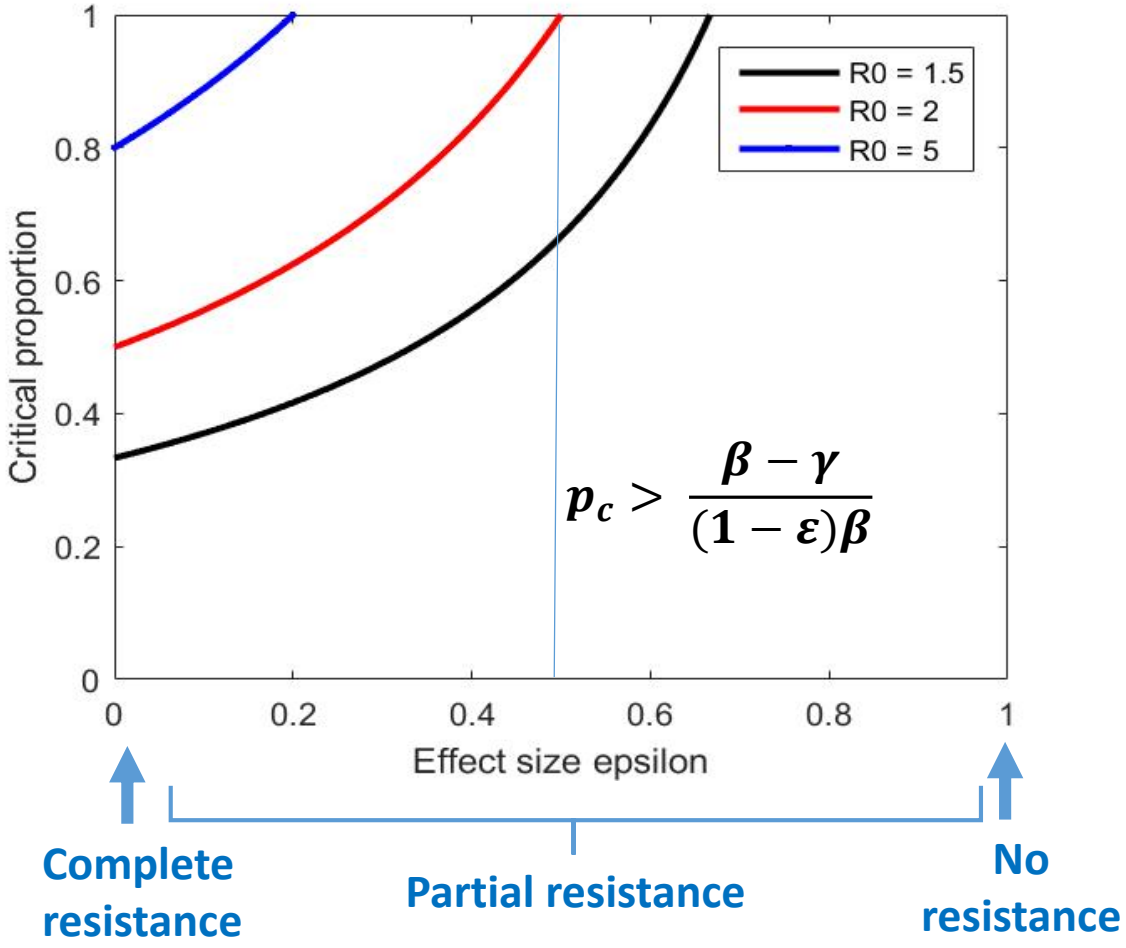
Will there be an epidemic?

- Similar approach as for the homogeneous SIR model:
- Epidemic will not occur if $R_0 < 1$ (or $\frac{dI}{dt} \leq 0$)
- One can show (see tutorial!) that this is the case iff:

the critical proportion $\boldsymbol{p_c}$ of individuals with low susceptibility exceeds the threshold

$$\boldsymbol{p_c > \frac{\beta - \gamma}{(1-\varepsilon)\beta}}$$

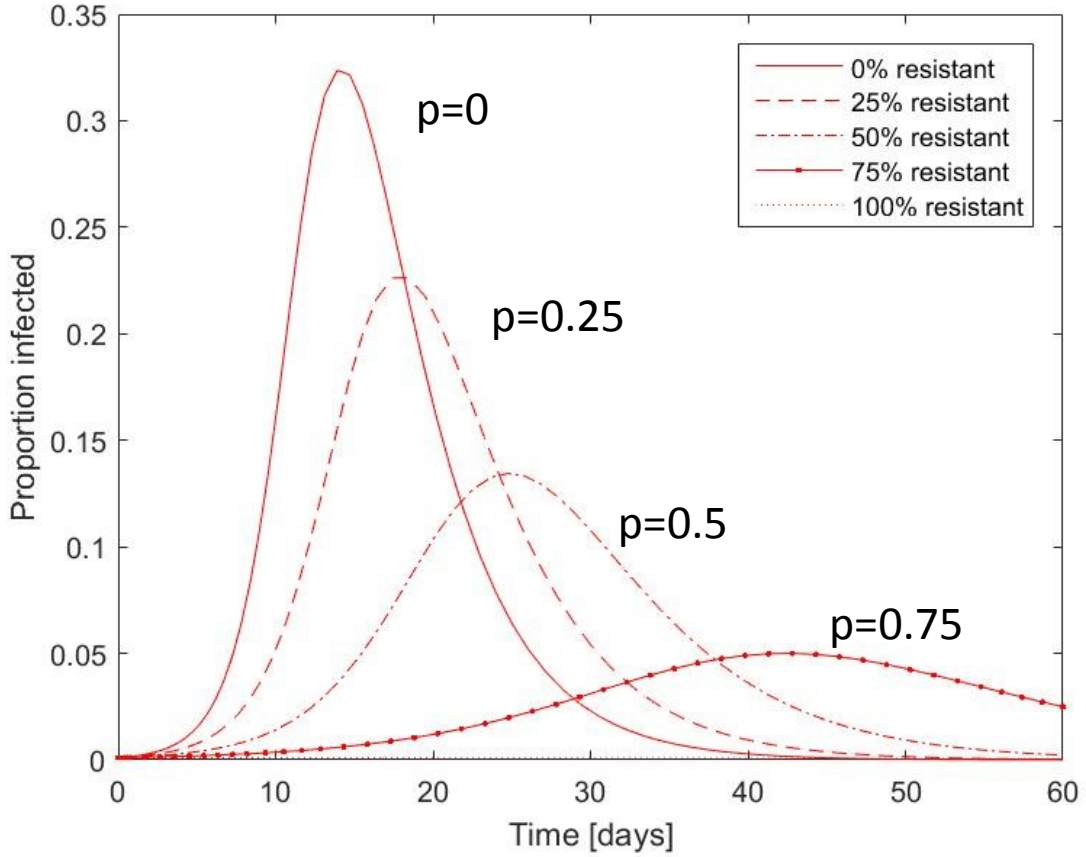
Critical proportion to achieve R_0<1 depends on the level of resistance ε as well as on epidemiological parameters



- More virulent diseases (higher R_0*) require higher proportion of individuals with partial resistance
- The critical proportion increases non-linearly with effect size ε
- Partial resistance may not be sufficient to eliminate the infection
- How does this inform strategies for disease control (e.g. vaccination or genetic selection)?

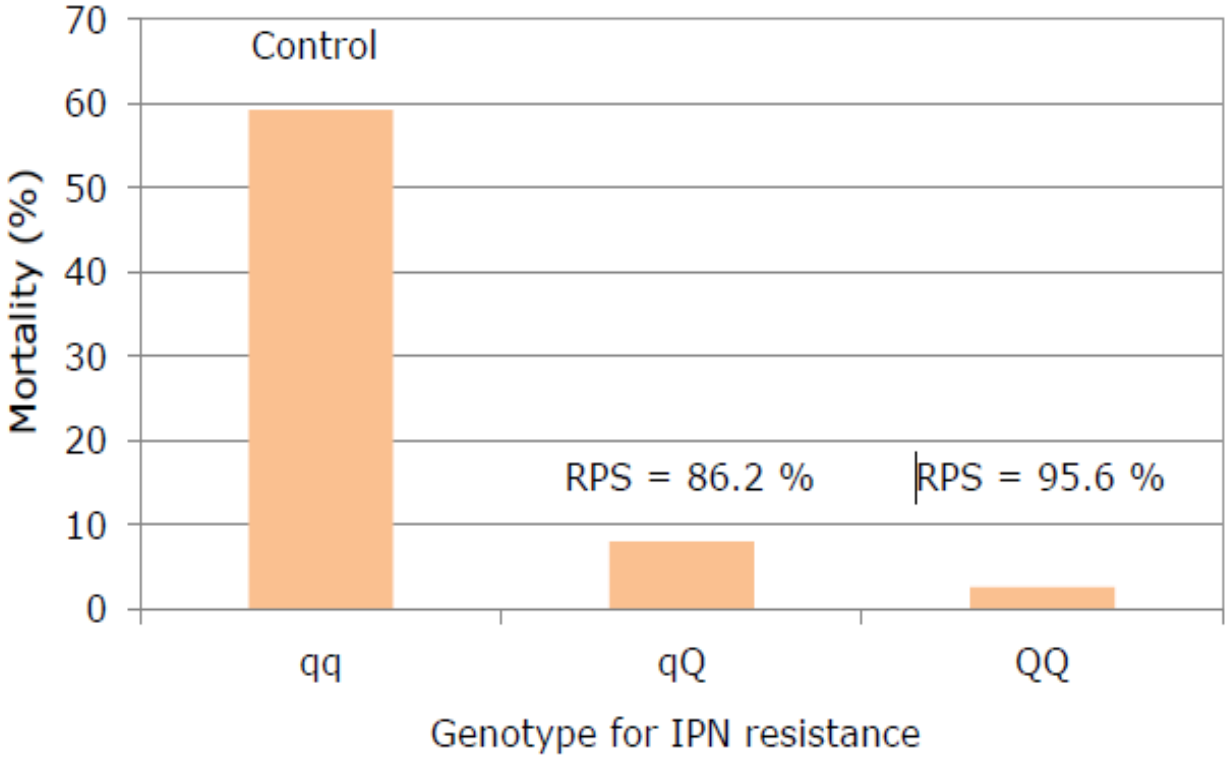
* R_0 in homogeneous population with $R_0 = \beta/\gamma$

Impact of population structure on epidemiological characteristics



- Higher proportion p of individuals with low susceptibility implies weaker, but more prolonged epidemics
- See tutorial for more in depth exploration of dynamic properties of this model

Example: IPN in Atlantic salmon



- Resistance largely controlled by a single gene (Houston et al., 2009; Moen et al. 2009)
- Presence of Q allele confers high resistance
  - Two genotypes:
    S_- (QQ and qQ) and S_+ (qq)
- Disease control through dissemination of salmon eggs with higher resistance in progress

Source: Aquagen newsletter 2014

IPN questions of interest

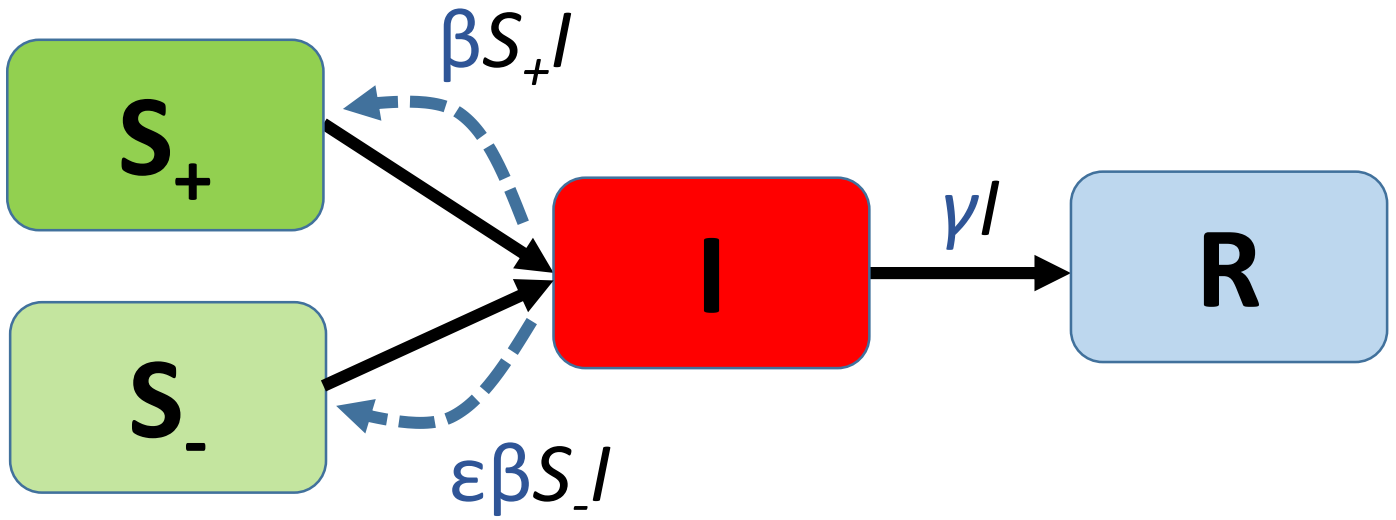
Assuming that the Q allele confers differences in susceptibility only:

What proportion of resistant individuals is required to prevent an IPN outbreak?

Solution: critical proportion of resistant fish for preventing IPN outbreak:

Assuming that the Q allele confers differences in susceptibility only:

What proportion of resistant individuals is required to prevent an IPN outbreak?

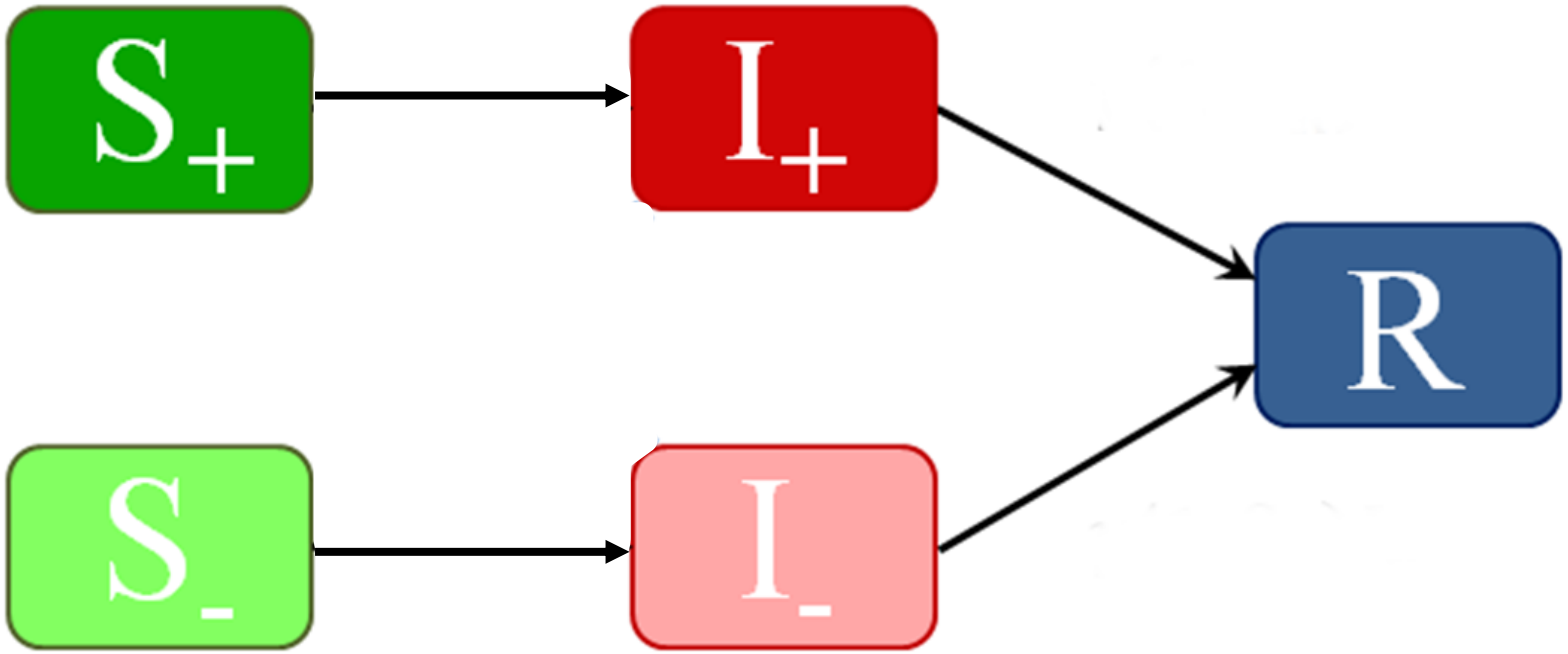


- Infer estimates for ε, β, γ from data
- Apply $p_c > \frac{\beta - \gamma}{(1-\varepsilon)\beta}$

IPN questions of interest

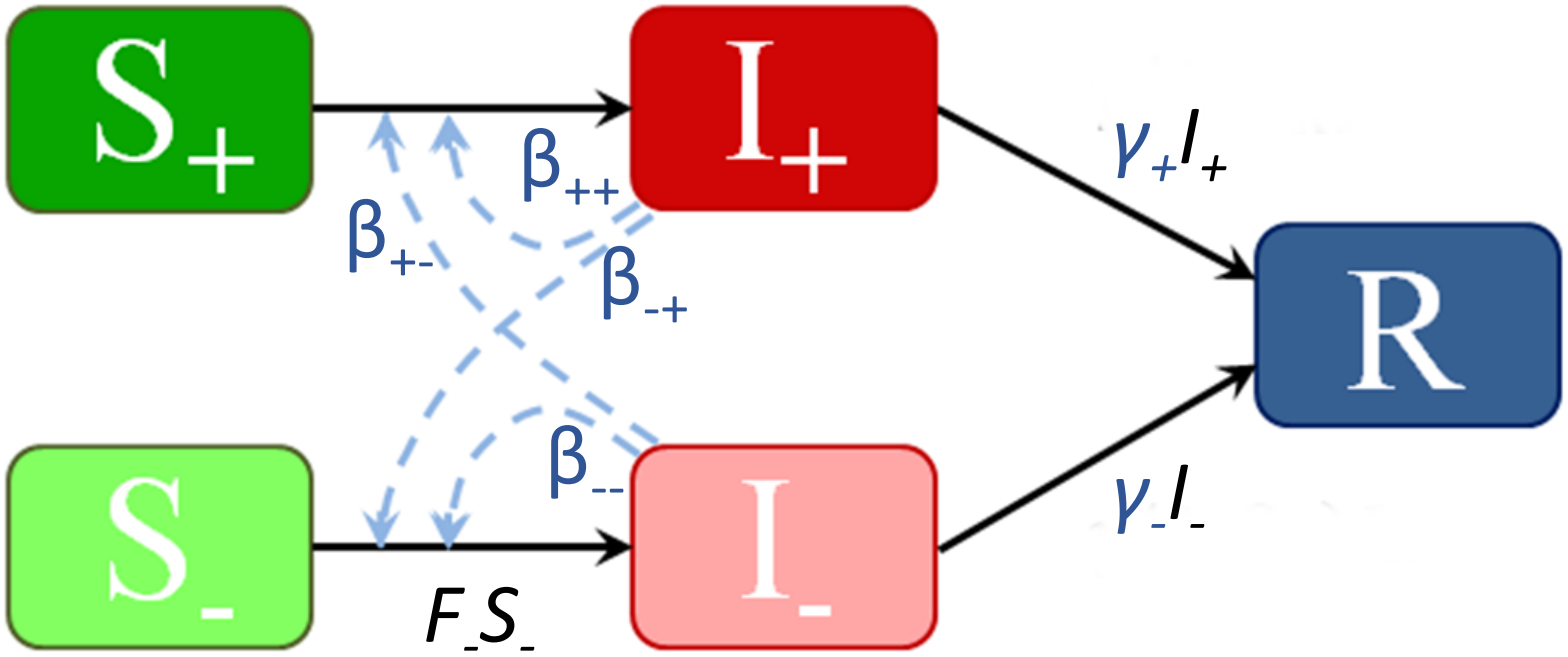
- Are resistant fish also less infectious?
- If so, how does this affect the critical proportion of resistant individuals required to prevent an outbreak?

A more general genetic epidemiological SIR model

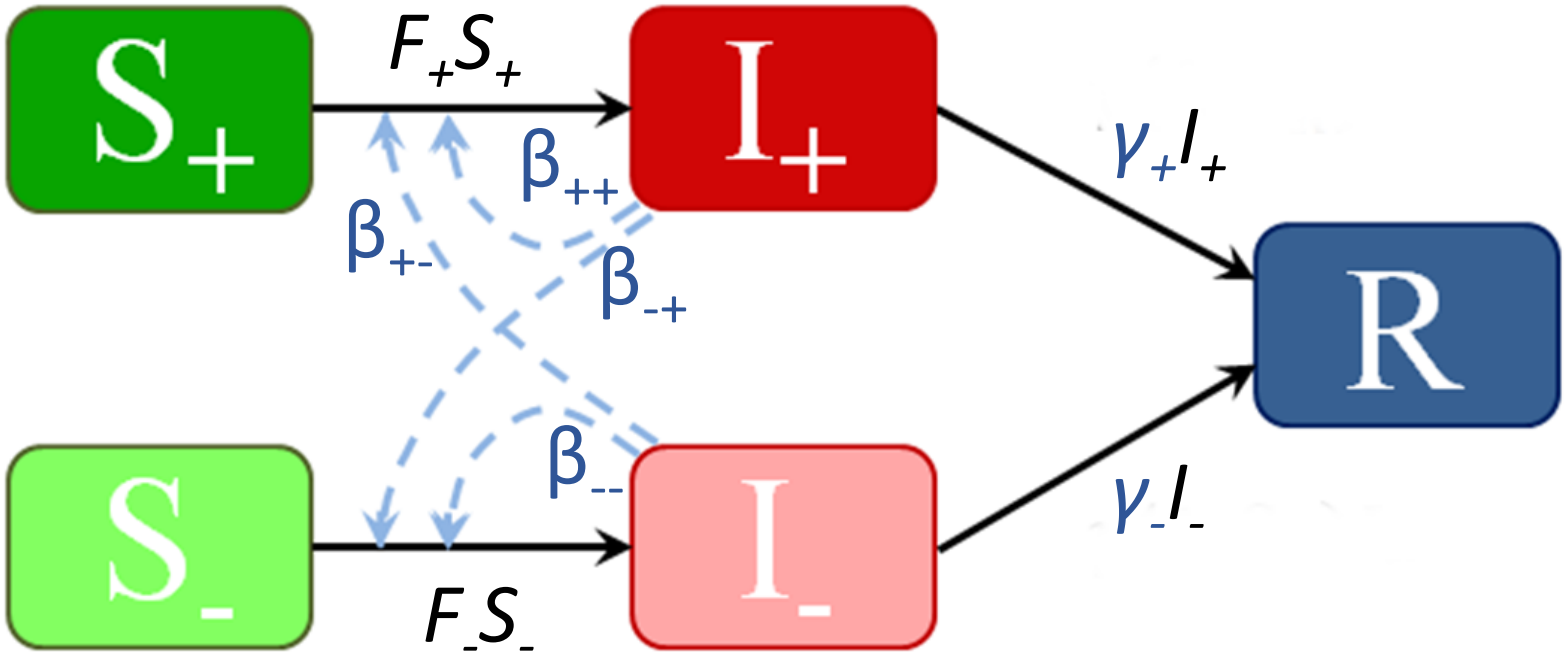


Susceptible Infected Dead

A more general genetic epidemiological SIR model



A more general genetic epidemiological SIR model



Force of infection:

$$F_+ = (\beta_{++} I_+ + \beta_{+-} I_-)/\mathrm{N}$$

$$F_- = (\beta_{-+} I_+ + \beta_{--} I_-)/\mathrm{N}$$

Can you write the model equations?

- See tutorial

The Who Aquires Infection From Whom (WAIFW) matrix

Instead of one transmission parameter β we now get a matrix:

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_{++} & \beta_{+-} \\ \beta_{-+} & \beta_{--} \end{pmatrix}$$

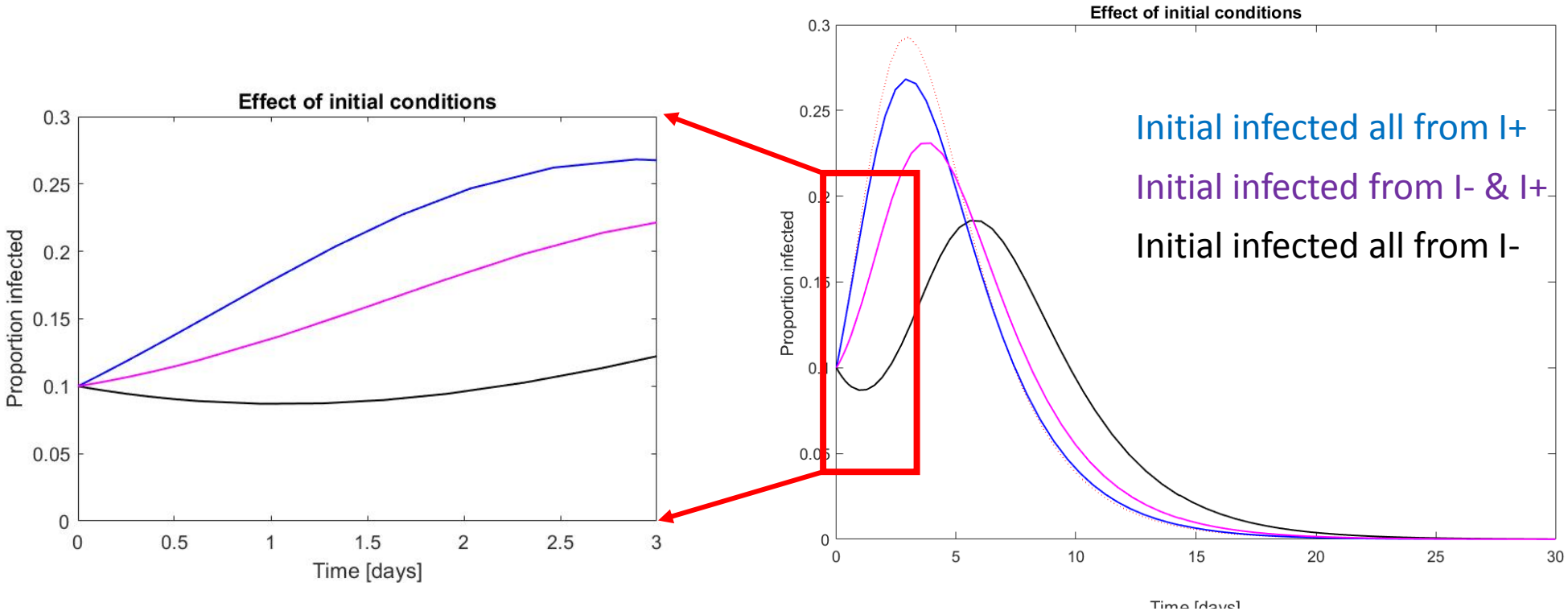
where β_{ij} means infection from individual j to individual i

What do we know about the transmission parameters β in the case of IPN if

- Reduced susceptibility confers no difference in infectivity?
- Reduced susceptibility allele confers higher / lower infectivity?

See tutorial!

Initial dynamics: Effect of initial conditions

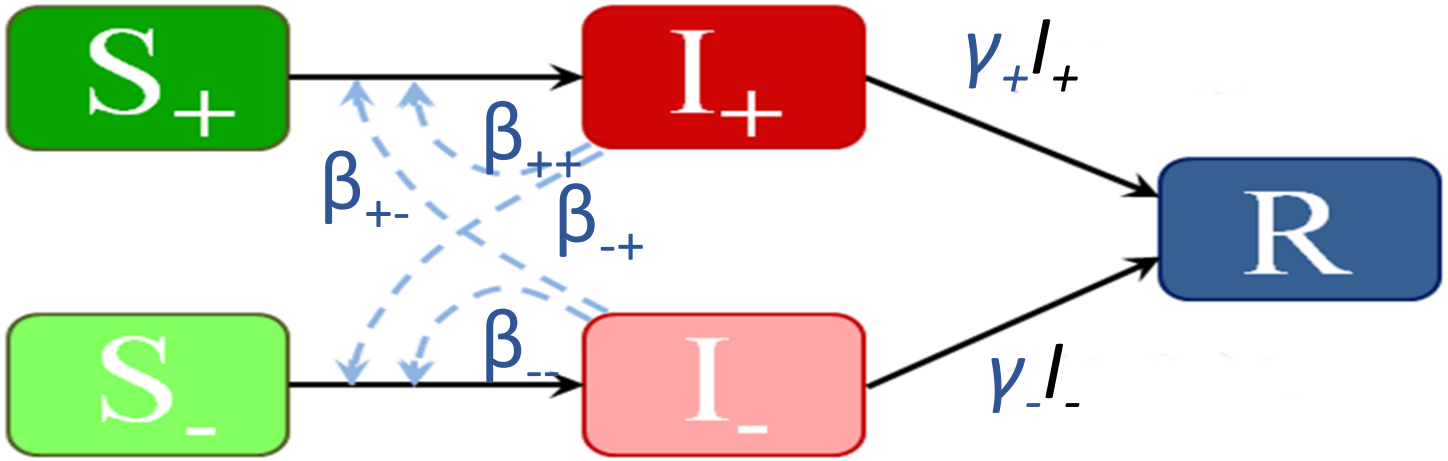


- The early dynamics depend on who is seeding the infection
- The effect of the initial seed fades over time
- R_0 can't be associated with the early stage of the epidemics

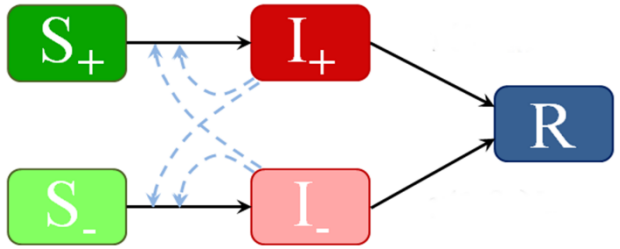
The basic reproductive ratio R_0

Definition: Basic reproductive ratio R_0 for heterogeneous populations
*The average number of secondary cases arising from an **average** infected individual in an entirely susceptible population, **once initial transients have decayed**.*

- Threshold characteristics of R_0 still valid, i.e. no epidemic if $R_0<1$
- How to calculate this number?



The basic reproductive ratio R_0



Definition: Basic reproductive ratio R_0 for heterogeneous populations

*The average number of secondary cases arising from an **average** infected individual in an entirely susceptible population, **once initial transients have decayed**.*

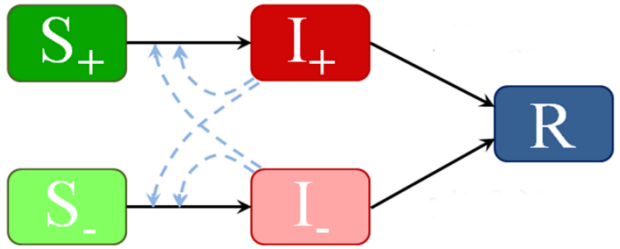
- Assume a susceptible population in which a proportion p is from category S_- and (1-p) from category S_+
- Expected number of secondary cases **arising from from I_+** category:

$$R_0^+ = (\beta_{++}(1-p) + \beta_{-+}\, p)/\gamma_+$$

- Expected number of secondary cases **arising from I_-** category:

$$R_0^- = (\beta_{+-}(1-p) + \beta_{--}\, p)/\gamma_-$$

The next generation matrix M and R_0



Definition: Basic reproductive ratio R_0 for heterogeneous populations

*The average number of secondary cases arising from an **average** infected individual in an entirely susceptible population, **once initial transients have decayed**.*

- The entries of the **Next Generation Matrix M**, i.e. m_{ij} refer to the number of secondary type i cases produced by an infected individual of type j

$$\begin{array}{c} \\ S_+ \\ S_- \end{array} \begin{array}{c} \begin{array}{cc} I_+ & I_- \end{array} \\ \begin{pmatrix} \frac{\beta_{++}(1-p)}{\gamma_+} & \frac{\beta_{+-}(1-p)}{\gamma_-} \\ \frac{\beta_{-+}p}{\gamma_+} & \frac{\beta_{--}p}{\gamma_-} \end{pmatrix} \end{array}$$

- R_0 is the largest eigenvalue of M
- R_0 is bounded by R_0^+ and R_0^-

Calculating Eigenvalues

For a 2×2 matrix: $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$:

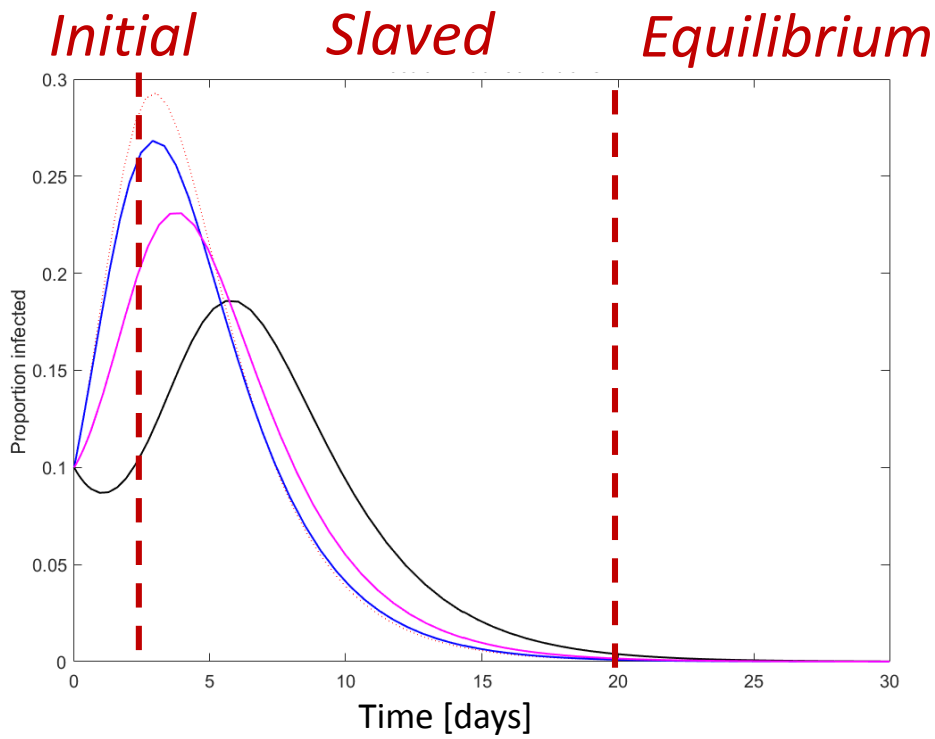
The largest eigenvalue is

$$R_0 = \frac{a + d + \sqrt{(a+d)^2 - 4(ad - bc)}}{2a}$$

It can be calculated in R using *eigen(M)*

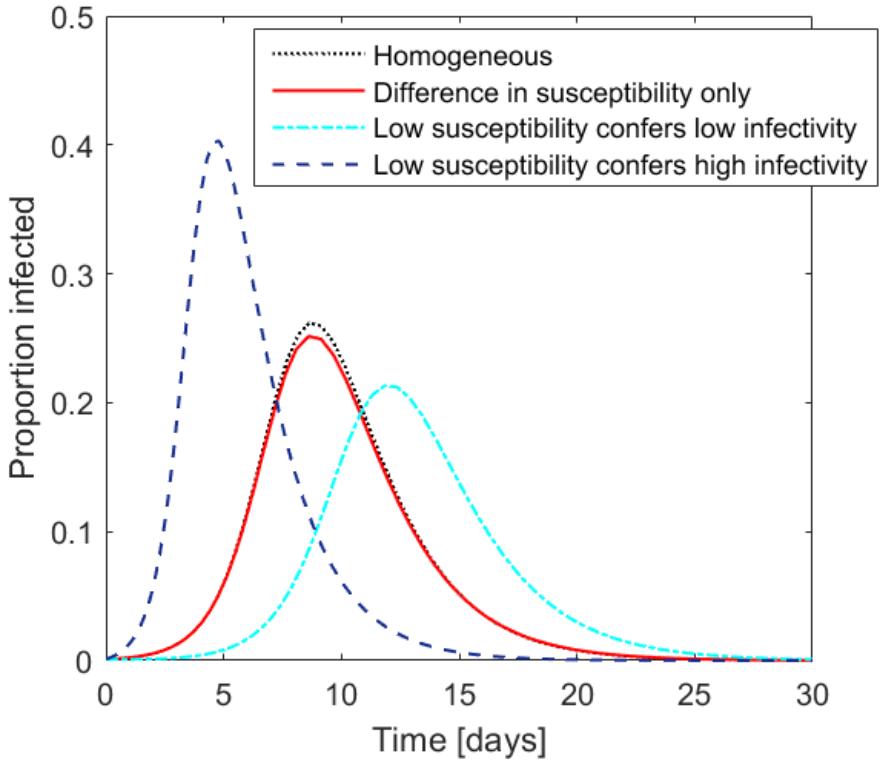
Dynamic properties

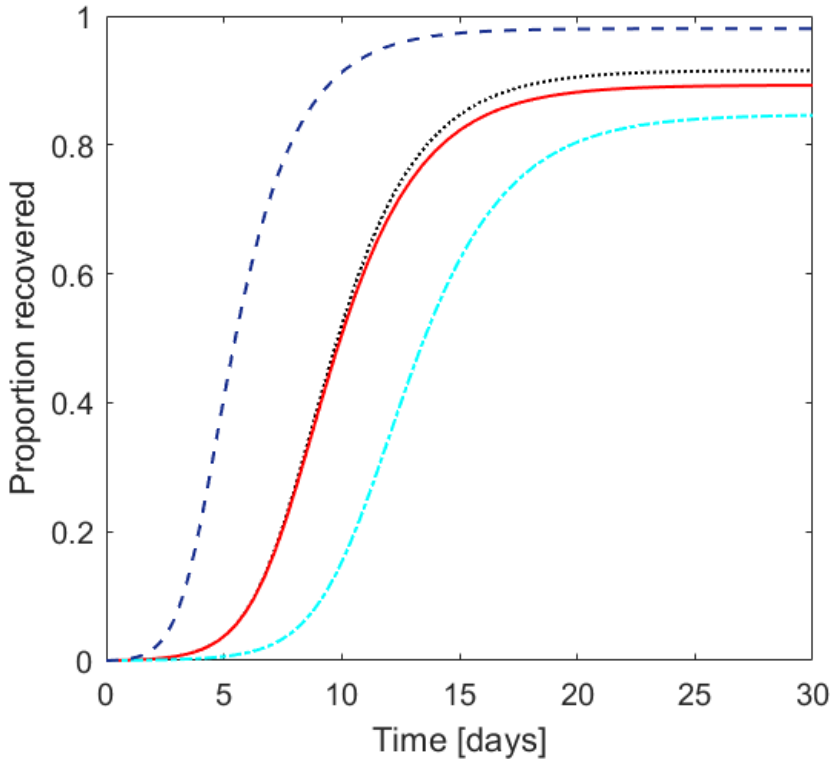
- Epidemics characterised by 3 phases:
  1. Initial phase: depends on who seeds the epidemic
  2. ‘Slaved’ phase: depends on R_0
  3. Equilibrium phase: Infection always dies out, but prop. of infected varies



Dynamic properties

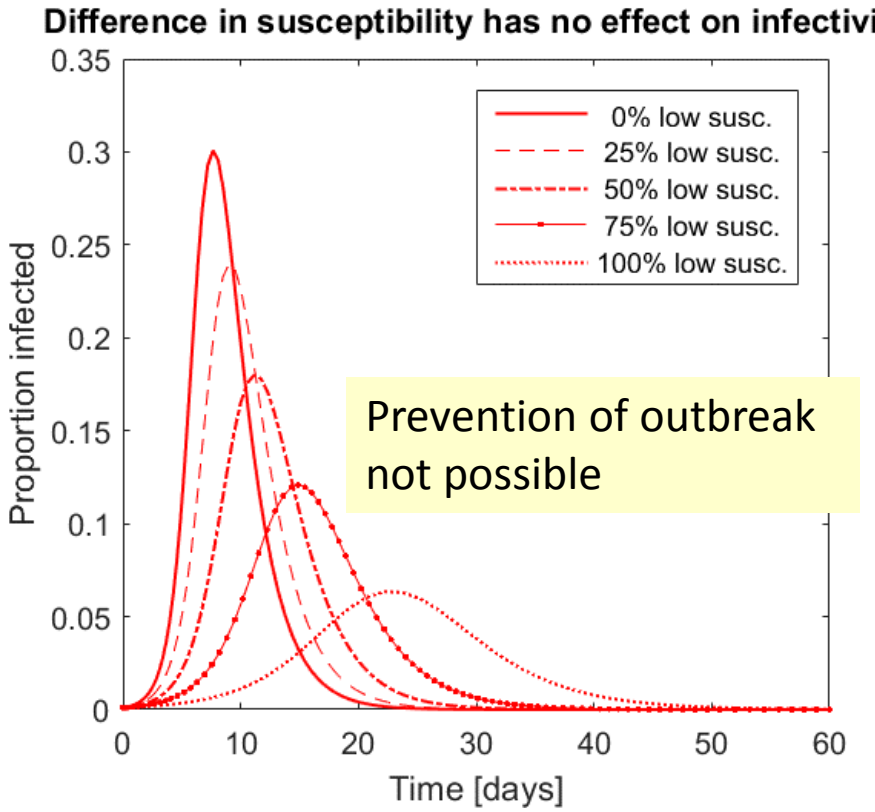
- Similar profile shapes as for homogeneous populations
- But different epidemiological characteristics (peak prevalence, total prop. infected etc.)
  - See tutorial for in depth analysis

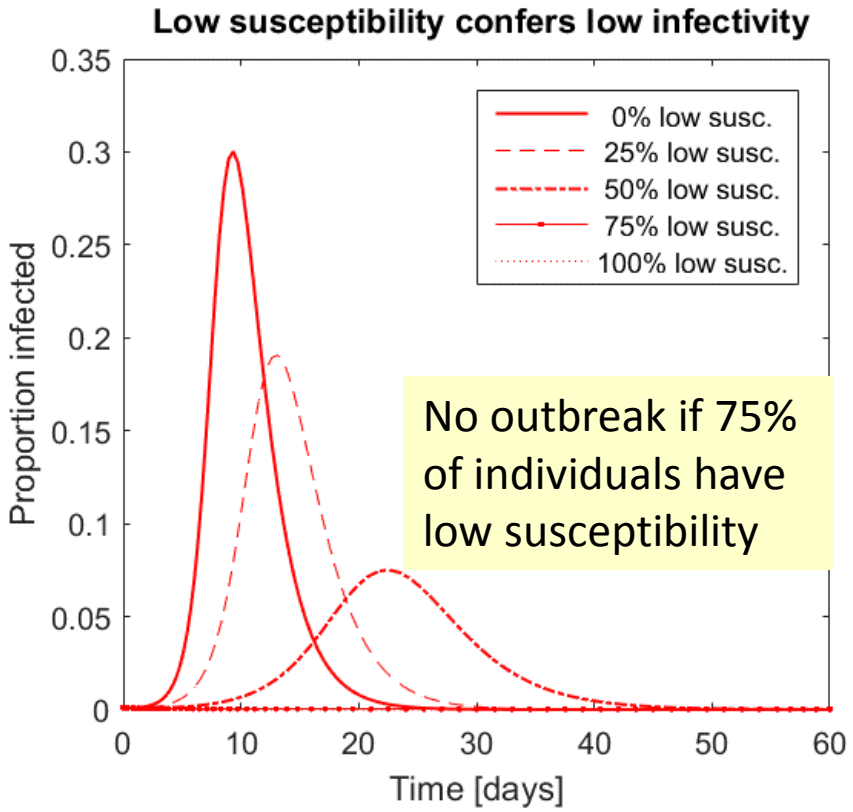




Critical proportion of individuals required to achieve $R_0<1$

- Depends on the relationship between susceptibility & infectivity
  - See tutorial for in depth analysis





Adding complexity

- The model can be easily extended to more discrete classes, e.g.
  - more genotypes (AA, aA, aa) (See Doeschl-Wilson et al. 2018)
  - Susceptibility and infectivity controlled by different genes (see e.g. Anche et al. 2015, Biemans et al. 2017)
  - More sources of heterogeneity (recovery rate)
- For discrete classes the theory is well established, although the analysis (e.g. calculating R_0, equilibria, etc.) becomes more difficult with increasing complexity

General properties of epidemiological models in heterogeneous populations with discrete structures (e.g. genoytpes)

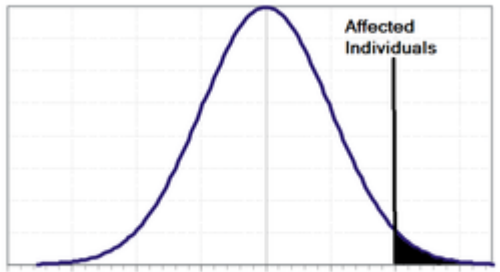
- The single transmission parameter β is replaced by a WAIFW matrix
- R_0 is the largest Eigenvalue of the next generation matrix
  - Derived in Diekmann et al. 1990; Hesterbeek 2002
- R_0 in heterogeneous populations is often larger than if the structures were ignored (i.e. all individuals have same average transmission rate)
  - Harder to eliminate the epidemic in highly heterogeneous populations!

But what if epidemiological parameters follow a continuous distribution?

- **Trait primarily controlled by single gene or locus (Major gene model)**
  - E.g. gene affecting binding protein
  - **Discrete genotypes**
- Trait controlled by many genes, each with a small effect (Polygenic model)
  - Majority of cases
  - Continuous genotypes
- Different genetic architectures lend themselves to different modelling approaches

But what if epidemiological parameters follow a continuous distribution?

- Trait primarily controlled by single gene or locus (Major gene model)
  - E.g. gene affecting binding protein
  - Discrete genotypes
- **Trait controlled by many genes, each with a small effect (Polygenic model)**
  - **Majority of cases**
  - **Continuous genotypes**
  - **Different modelling approach required!**
    - **See next lecture**



Summary

- Epidemic models including heterogeneity are more realistic
  - Ignoring heterogeneity can produce wrong predictions
- Heterogeneity is relatively easy to implement if represented by discrete classes
  - Individual parameters are replaced by matrices
- The basic reproductive number R_0 is differently defined and calculated for models of heterogeneous populations
  - R_0 found by an eigenvalue approach
- Heterogeneity can be harnessed by control strategies
  - E.g. mixing different types of individuals, targeted treatment, genetic selection, …

References

Diekmann, O., J. A. P. Heesterbeek, and J.A.J. Metz. "On the definition and the computation of the basic reproduction ratio R 0 in models for infectious diseases in heterogeneous populations." *Journal of mathematical biology* 28.4 (1990): 365-382.

Diekmann, O., J. A. P. Heesterbeek, and M. G. Roberts. "The construction of next-generation matrices for compartmental epidemic models." *Journal of the Royal Society Interface* (2009): rsif20090386.

Heesterbeek, J. A. P. "A brief history of R0 and a recipe for its calculation."*Acta biotheoretica* 50.3 (2002): 189-204.

Keeling, Matt J., and Pejman Rohani. Modeling infectious diseases in humans and animals. Princeton University Press, 2008.

Anche, M. T., M. C. M. de Jong, and P. Bijma. "On the definition and utilization of heritable variation among hosts in reproduction ratio R0 for infectious diseases." *Heredity* 113.4 (2014): 364-374.

Biemans, F., Jong, M. C., & Bijma, P. (2017). A model to estimate effects of SNPs on host susceptibility and infectivity for an endemic infectious disease. *Genetics Selection Evolution*, *49*(1), 53.

Doeschl-Wilson A. et al. (2018). New opportunities for genetic disease control: beyond disease resistance. In *11th World Congress on Genetics Applied to Livestock Production*.