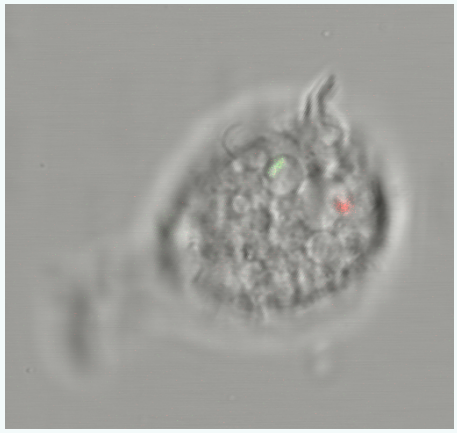
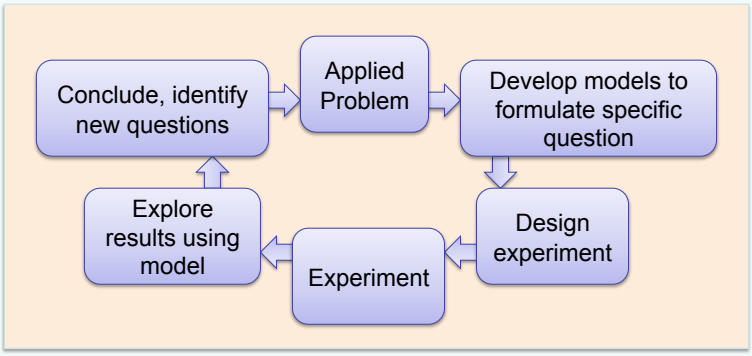
Disease Dynamics: From Equation To Experiment (And Back)



Julia Gog, DAMTP, University of Cambridge

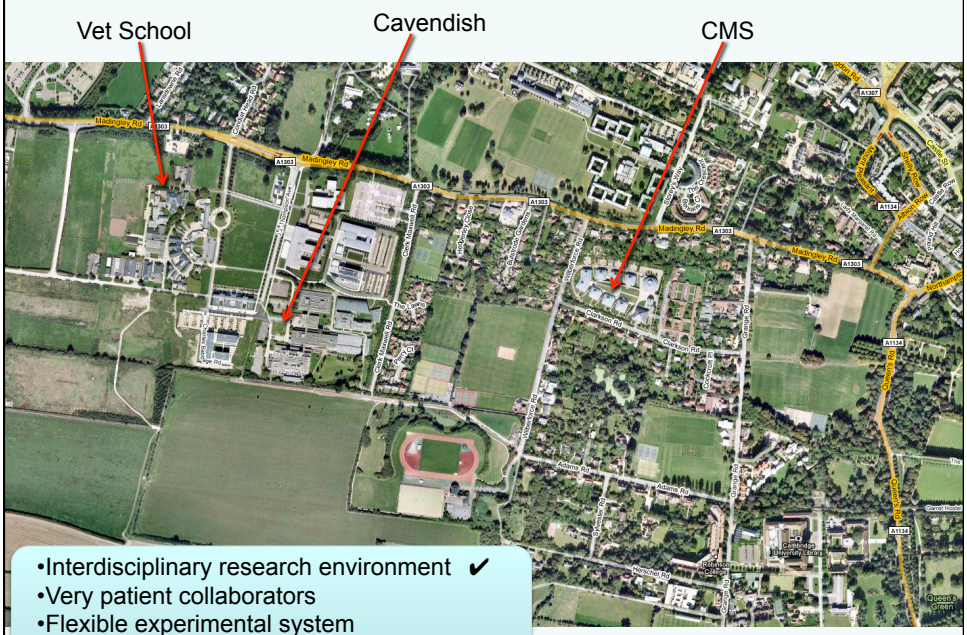
Iterative approach between models and experiment



But need:

- Interdisciplinary research environment
- Very patient collaborators
- Flexible experimental system

Research environment that supports interdisciplinary approach



- Interdisciplinary research environment ✔
- Very patient collaborators
- Flexible experimental system

Very patient collaborators

Me

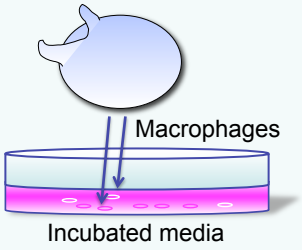
Dr Clare Bryant
Department of Veterinary Medicine
University of Cambridge

Dr Pietro Cicuta
Cavendish Laboratory
University of Cambridge

- Interdisciplinary research environment ✔
- Very patient collaborators ✔
- Flexible experimental system

The experimental system

Much of the experimental work in this talk was done by **Dr Alicia Murcia**, PhD co-supervised by Bryant and Gog



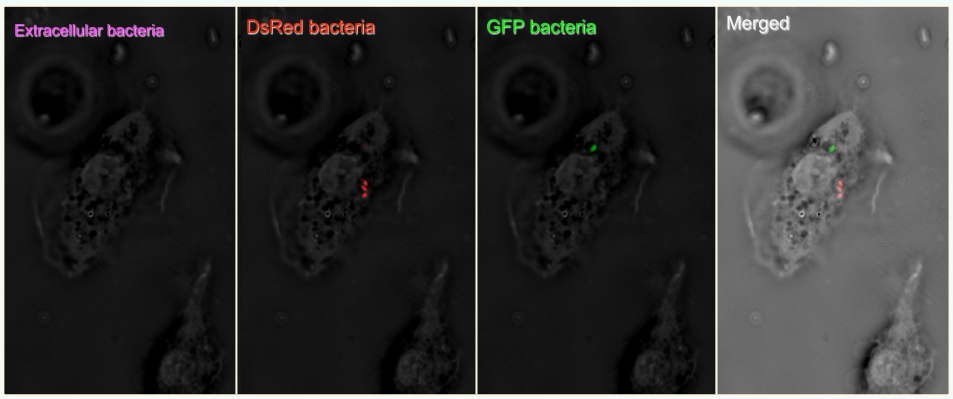
Add salmonella bacteria

modified to express red or green fluorescence

EITHER: Fix at some set timepoint, wash, stain, count etc.
OR: Do live imaging

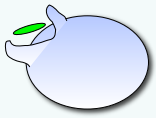
- Interdisciplinary research environment ✔
- Very patient collaborators ✔
- Flexible experimental system ✔

Experimental approach

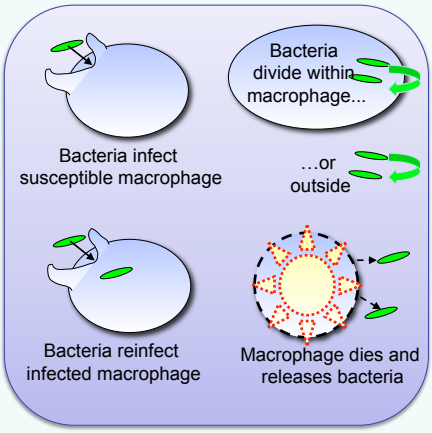


Why study the dynamics of *Salmonella*-macrophage interactions?

- This strain causes Typhoid-like illness in mice
- Widely used experimental system
- Used to identify virulence proteins
- Has led to the identification of vaccine candidates
- But… basic understanding of the events involved in Salmonella infection are surprisingly not well characterised.



Preliminary attempts to construct a model



Can construct a mathematical model with relatively simple assumptions:

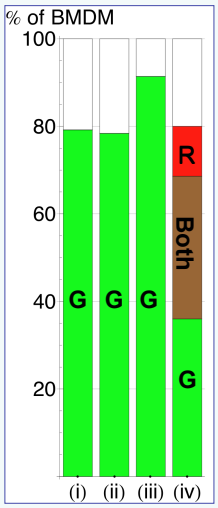
$$\begin{aligned} B'_{ex} &= dB_{in} + g_{ex}B_{ex} - bB_{ex}(C_s + C_i) \\ B'_{in} &= bB_{ex}(C_s + C_i) - dB_{in} + g_{in}B_{in} \\ C'_s &= -bC_sB_{ex} \\ C'_i &= -dC_i + bC_sB_{ex} \end{aligned}$$

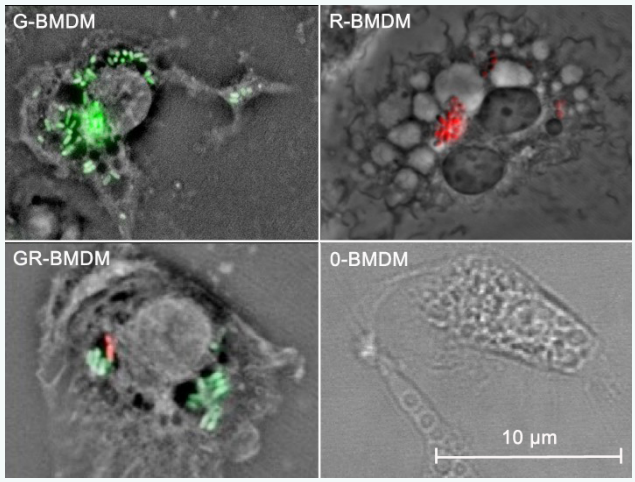
This is not intended for fitting to data, but the process of constructing identifies gaps where we can't decide between possible assumptions.

Question from preliminary experiments and model

- *Are infected macrophages susceptible to further infection events or re-infection?*
- Use imaging techniques to distinguish intra and extracellular bacteria
  - (i) sequential challenge with two different bacterial strains
  - (ii) microscopy investigating intra-cellular localisation of multiple bacteria
  - (iii) live imaging to directly observe a reinfection event.

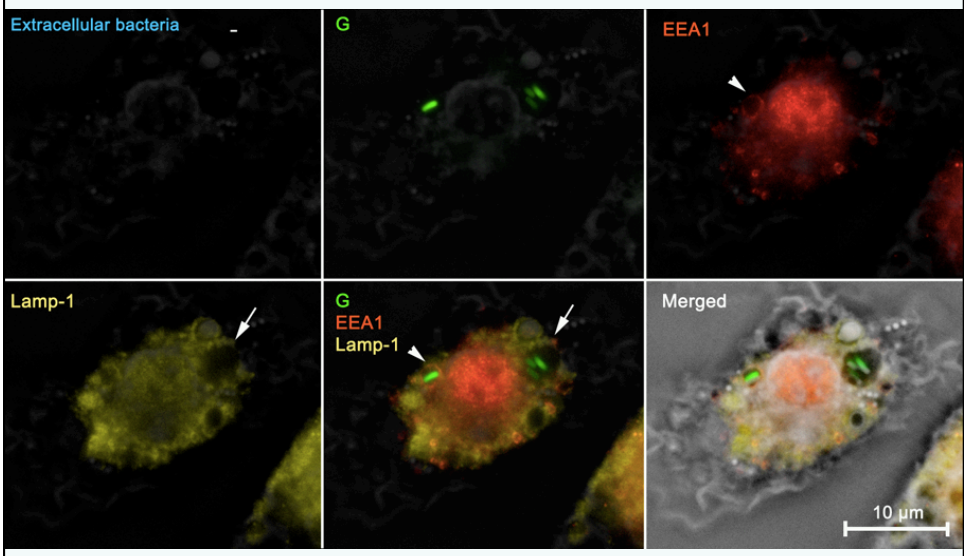
Sequential infection



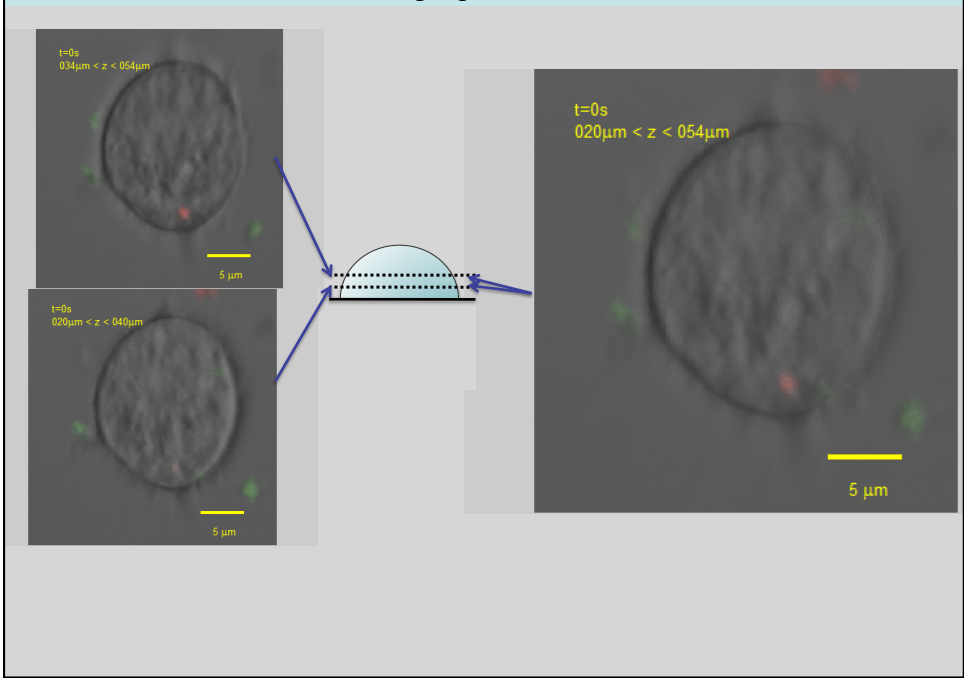


i) G, 30m
ii) G, 30m, wash, 30m
iii) G, 30m, wash, G, 30m
iv) G, 30m, wash, R, 30m

Really reinfection? Or both colours at same time?



Live imaging of re-infection



Answered one question, several more raised…

- Some cells resistant to infection: is there more than one population of cells?
- What is the infection rate?
- What is the probability of infection when bacterium encounters a cell (thought to be close to one)?
- Is an infected cell more or less susceptible to a second infection event? (relative reinfection rate)

Infection with red and green are NOT independent

	No R	R	Tot
No G	**301** (271) 	**39** (69) 	*340*
G	**97** (127) 	**63** (33) 	*160*
Tot	*398*	*102*	*500*

Count observed
(expected from totals)

$p < 10^{-8}$
from Fisher's Exact Test (one-tailed)

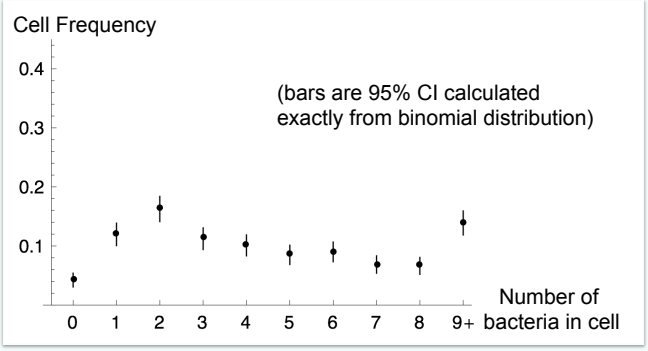
Candidate explanation	*Does it work?*
New cells grow	*Probably not, too fast*
Cells killed by infection	*Maybe, but seems too fast*
Some cells more susceptible	*Maybe, but real news if so*
Infection renders cells susceptible	*Plausible*

Modelers job: come up with simple models to explore these options, and identify an experiment that will distinguish between them

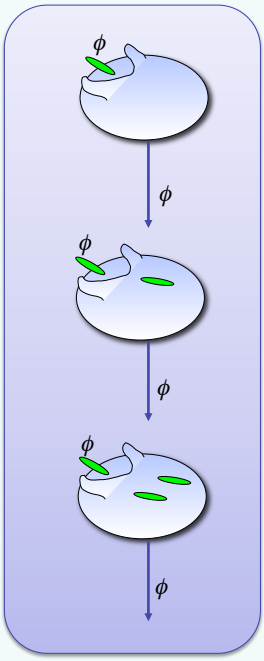
Next generation of experiments

Cout the number of bacteria per cell, use very early timepoint, and vary the MOI used.

1500 cells counted per figure, and classified according to the number of intracellular bacteria:



Construct models to reflect the data resolution (i.e. the model outputs distributions of bacteria per cell). Combine models and experiments to answer some of the questions above.



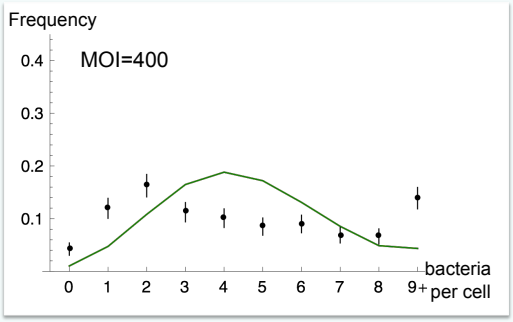
p_i – proportion of cells with i intracellular bacteria
ϕ –infection rate

$$
\begin{array}{rcl}
\dot{p}_0 & = & -\phi p_0 \\
\dot{p}_1 & = & \phi p_0 - \phi p_1 \\
\dot{p}_2 & = & \phi p_1 - \phi p_2 \\
\dot{p}_3 & = & \phi p_2 - \phi p_3 \\
\ldots & \ldots & \ldots \\
\dot{p}_{9+} & = & \phi p_8
\end{array}
$$

This is the simplest type of this model: a constant infection rate on each cell, regardless of how many bacteria are contained already. This leads to a Poisson distribution….

…however gives a very bad fit…

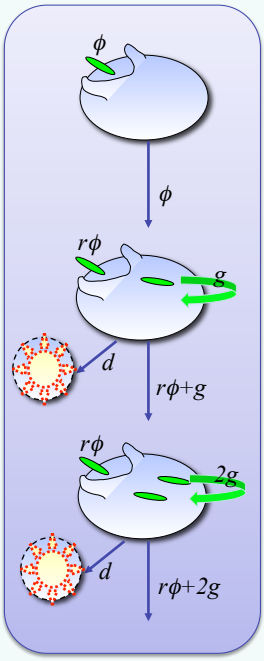
Simplest model – a bad fit



Poisson model best fit (by maximum likelihood) predicts many cells with intermediate number of bacteria, whereas we observe more at both extremes.
What have we missed?

Plausible factors that might explain the difference:
- ***Intracellular bacterial growth:*** could explain why some heavily infected?
- ***Reinfection:*** at different rate than first infection
- ***Death of infected cells:*** would explain an excess of uninfected cells
- ***Heterogeneity of macrophages:*** some more easily infected than others

Build further models, explore to untangle effects of the factors above



p_i – proportion of cells with i intracellular bacteria
ϕ –infection rate for uninfected cells
$r\phi$–reinfection rate (so r is relative rate)
g – intracellular growth rate of bacteria
d – death rate of infected cells

$$
\begin{aligned}
\dot{p}_0 &= -\phi p_0 \\
\dot{p}_1 &= \phi p_0 - (r\phi + g + d)p_1 \\
\dot{p}_2 &= (r\phi + g)p_1 - (r\phi + 2g + d)p_2 \\
\dot{p}_3 &= (r\phi + 2g)p_2 - (r\phi + 3g + d)p_3 \\
\cdots &\quad \cdots \quad \cdots \\
\dot{p}_{9+} &= (r\phi + 8g)p_8 - dp_{9+}
\end{aligned}
$$

$$
\begin{aligned}
\dot{p}_0 &= -\phi p_0 \\
\dot{p}_1 &= \phi p_0 - (r\phi + g + d)p_1 \\
\dot{p}_2 &= (r\phi + g)p_1 - (r\phi + 2g + d)p_2 \\
\dot{p}_3 &= (r\phi + 2g)p_2 - (r\phi + 3g + d)p_3 \\
\dots &\quad \dots \quad \dots \\
\dot{p}_{9+} &= (r\phi + 8g)p_8 - dp_{9+}
\end{aligned}
$$

Can produce some solutions for various parameter combinations

$$p_i = \frac{e^{-\phi t}}{i!}\left(\frac{1-e^{-gt}}{g}\right)^i \left(\prod_{j=0}^{i-1}(\phi + ig)\right)$$

$$p_i = \frac{(\phi t)^i}{i!}e^{-\phi t} \qquad p_i = \frac{r^{i-1}}{(r-1)^i(i-1)!}e^{-\phi t}\gamma\left(i, \phi t(r-1)\right)$$

… but better numerically to just do this:

$$\dot{\underline{\mathbf{p}}} = \underline{\underline{\mathbf{A}}}(\phi, r, g, d).\underline{\mathbf{p}}$$

$$\underline{\mathbf{p}}_t = Exp(\underline{\underline{\mathbf{A}}}t).\underline{\mathbf{p}}_0$$

Make two populations:

Proportions of macrophages of each type

$$a\underline{\mathbf{p}}_t(\phi\rho, r, g, d) + (1-a)\underline{\mathbf{p}}_t(\phi, r, g, d)$$

One type is more easily infected: $\rho > 1$

Finally, condition on cell being alive:

$$\text{final} \quad \hat{p}_i(\phi, r, g, d) = \frac{p_i}{\sum_j p_j}$$

Count a total of N cells, and observe counts c_i (number of macrophages containing i bacteria. Then likelihood of observation given our parameters is just multinomial:

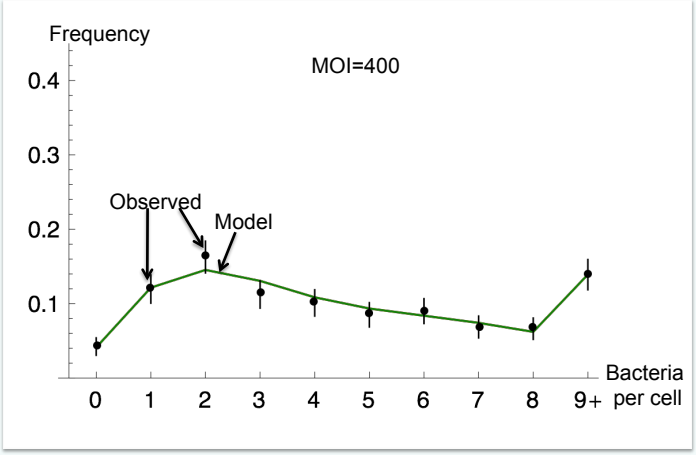
$$P\left((c_0, c_1, \ldots, c_n)|(p_0, p_1, \ldots, p_n)\right) = \frac{N!}{\prod_i c_i!} \prod_i {p_i}^{c_i}$$

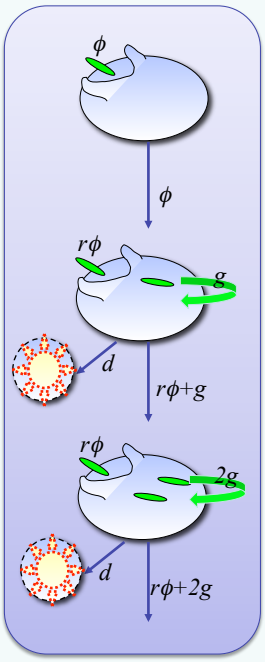
Log likelihood: $$LL = Log\left(\frac{N!}{\Pi_i c_i!}\right) + \sum_i c_i Log(p_i)$$

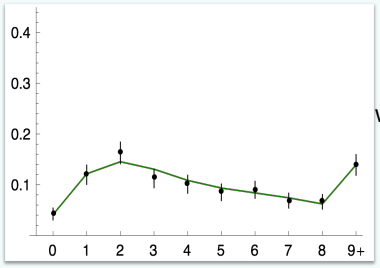
Want to maximise LL over the parameters.

Parameters only affect p, so only need to maximise second term.

Full model: better fits







Can get excellent fits with the full model: but are all the model components needed?

Make all 16 combinations of different components:
- Death of infected bacteria (or not, $d=0$)
- Different reinfection rate (or not, $r=1$)
- Two populations or one ($a=0$)
- Growth of intracellular bacteria (or not, $g=0$)

Use **AIC** to select 'best' model(s) based combination of how good the fit is, and model parsimony.

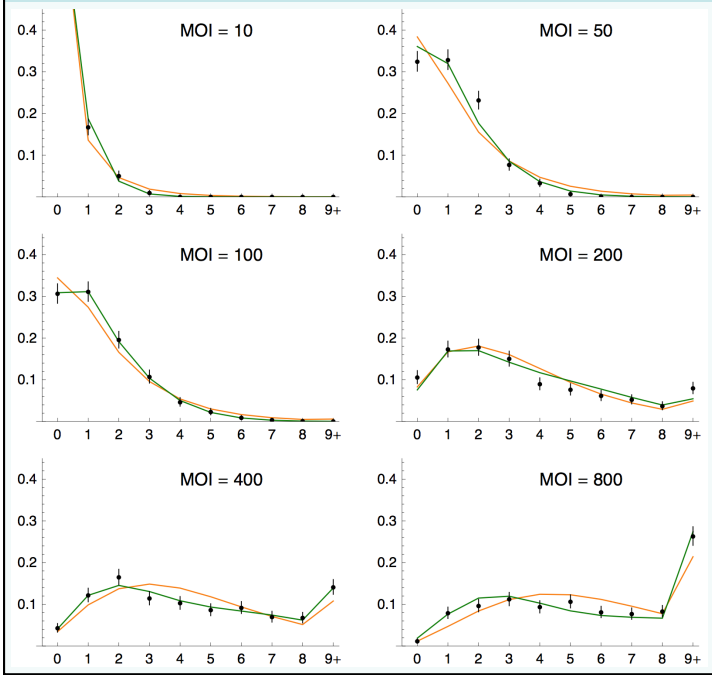
AIC = 2 x Log Likelihood – 2 x number of parameters

Model selection: which features are supported?

	Fitted	ΔAIC	weight	g	r	d	a	ρ
One population	-	1918.5	0	-	(1)	-	-	-
	d	1920.5	0	-	(1)	$< 10^{-8}$	-	-
	r	951.7	0	-	1.98	-	-	-
	r, d	953.7	0	-	1.98	$< 10^{-8}$	-	-
	g	391.5	0	1.1×10^{-3}	(1)	-	-	-
	g, d	393.5	0	1.1×10^{-3}	(1)	$< 10^{-8}$	-	-
	g, r	385.6	0	1.2×10^{-3}	0.88	-	-	-
	g, r, d	387.6	0	1.2×10^{-3}	0.88	$< 10^{-8}$	-	-
Two populations	-	10.2	0	-	(1)	-	0.47	2.92
	d	12.2	0	-	(1)	$2. \times 10^{-5}$	0.47	2.91
	r	12.1	0	-	0.99	-	0.47	2.94
	r, d	11.3	0	-	0.83	1.2×10^{-3}	0.51	2.87
	g	7.7	0.02	9.4×10^{-5}	(1)	-	0.46	2.85
	g, d	9.7	0.01	9.4×10^{-5}	(1)	$< 10^{-8}$	0.46	2.85
	g, r	0	0.71	2.5×10^{-4}	0.84	-	0.48	3.
	g, r, d	2.	0.26	2.5×10^{-4}	0.84	$< 10^{-8}$	0.48	3.

Support for:
- ***Two*** populations (instead of one)
- ***Growth*** Intracellular bacteria (during 10 mins..)
- ***Reinfection*** at lower rate than first infection ($r < 1$)
- ***No death*** of infected cells (on this timescale)

10 minutes, different MOIs



Allow infection rate ϕ to vary for different MOIs but all other parameters same (total 11 parameters)

Maximum likelihood fit to full dataset (total 3000 cells).

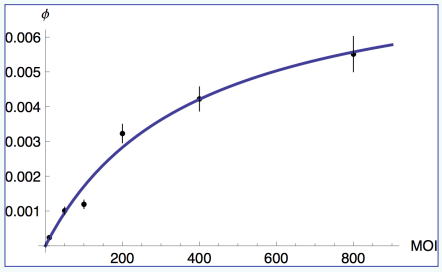
Orange – best fit for one population model

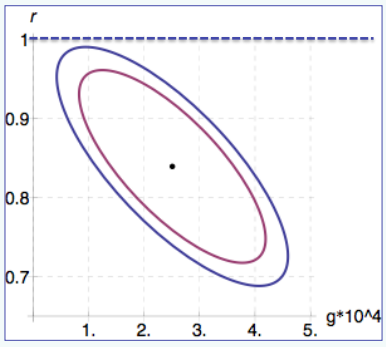
Green – best fit for two populations model

Parameter fits

Fitted Parameters		
a	0.52	$(0.47 - 0.57)$
g	2.51	$(0.74 - 4.28) \times 10^{-4}$
r	0.84	$(0.71 - 0.97)$
ρ	3.	$(2.76 - 3.24)$
ϕ_{10}	2.34	$(1.93 - 2.74) \times 10^{-4}$
ϕ_{50}	1.01	$(0.9 - 1.11) \times 10^{-3}$
ϕ_{100}	1.19	$(1.07 - 1.32) \times 10^{-3}$
ϕ_{200}	3.23	$(2.96 - 3.5) \times 10^{-3}$
ϕ_{400}	4.22	$(3.87 - 4.57) \times 10^{-3}$
ϕ_{800}	5.51	$(5. - 6.02) \times 10^{-3}$

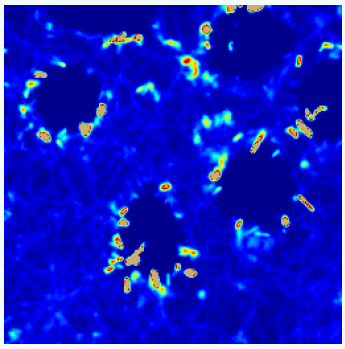
The parameters were estimated by maximum likelihood (range in brackets give 95% interval)

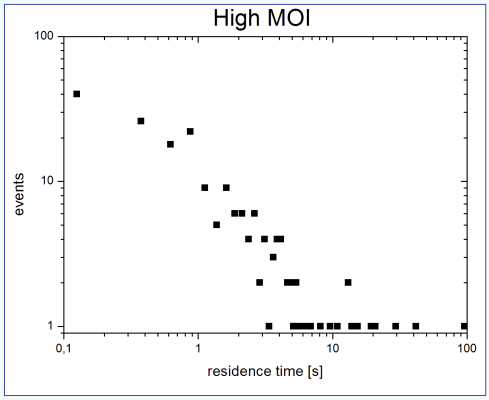




Probability of infection and infection rate

Analysis of “sticking times”.
Long tail - seems to correspond to infection.
From this, overall probability of infecting cell is < 5%.





Can combine this with swimming model for encounter rate to give an infection rate: is in same ballpark as infection model estimate

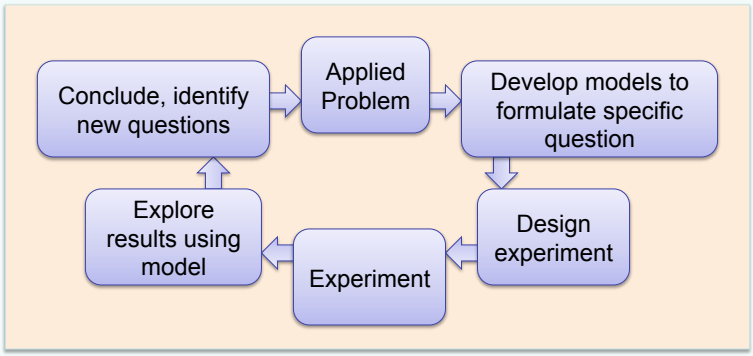
Conclusions

- Strong evidence that re-infection occurs
- Re-infection rate lower than initial infection
- Best fit to data if we assume two populations
- Probability of infection seems low

Preliminary

- Have identified two different types of macrophages
  - But don’t correspond to proportions suggested by model
  - DO appear to have differing susceptibility
  - Needs to be replicated and explored further
- Different story at 20 minutes
  - Death of heavily infected macrophages,
  - but strange time-dependency

Iterative approach between models and experiment



Acknowledgements

- Alicia Mucia
- Clare Bryant
- Pietro Cicuta
- Natan Osterman
- TJ McKinley
- Mark Shepperd
- Olivier Restif
- Duncan Maskell
- James Wood
- Bin Wei
- Piero Maestroeni
- CIDC