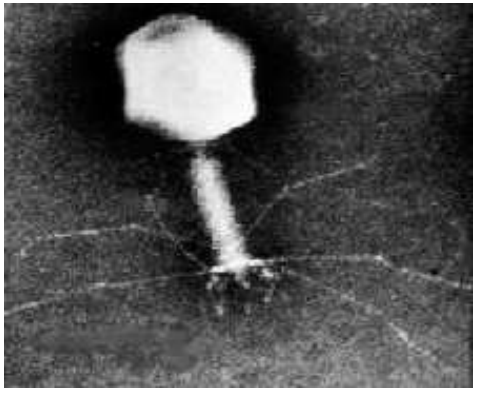
Modeling Marine Phage Ecology

Joseph M. Mahaffy

Nonlinear Dynamical Systems Group
Computational Sciences Research Center
Department of Mathematical Sciences
San Diego State University

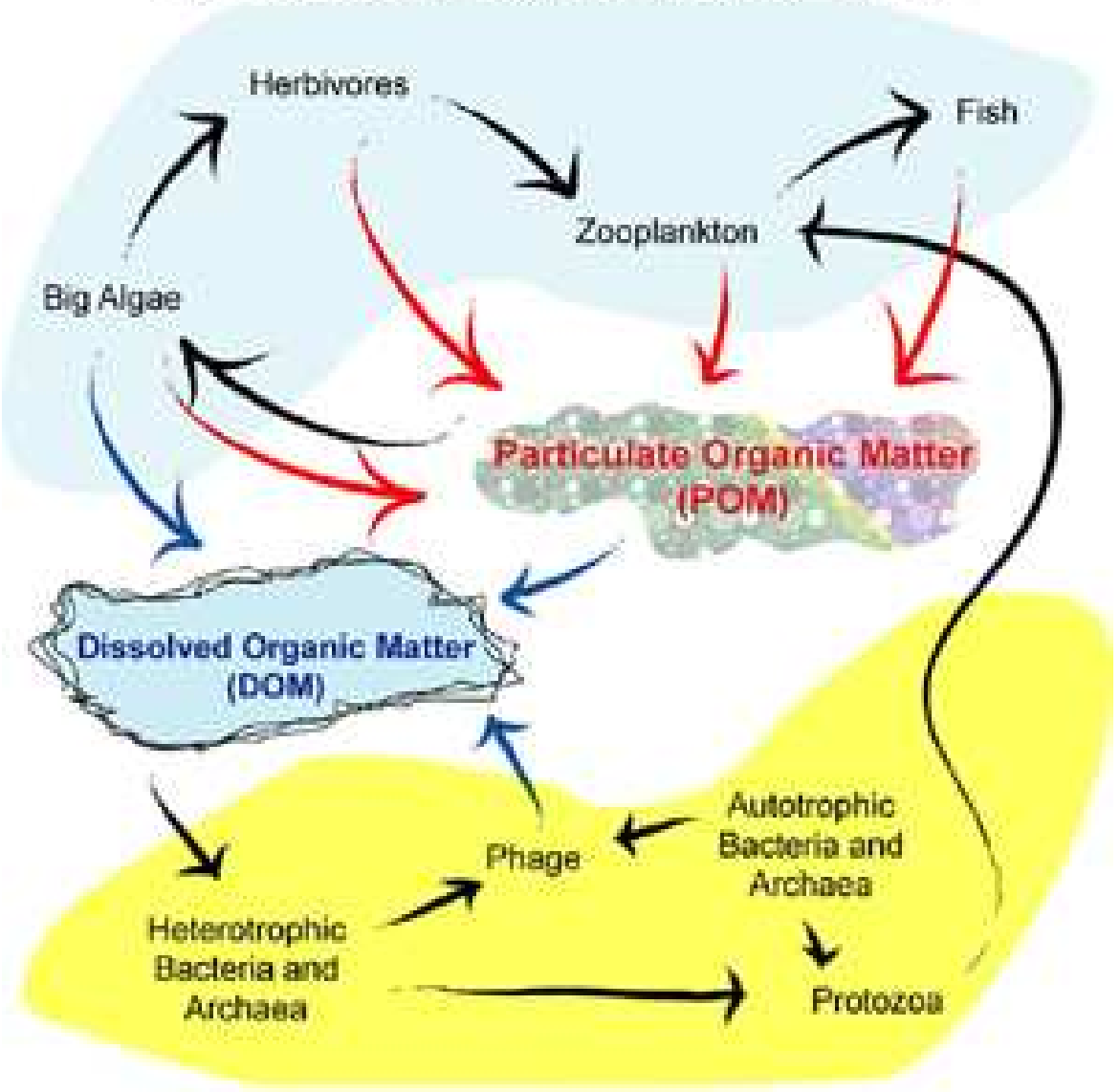
January 2006



Outline

- Introduction to Marine Phage
- Discuss Biological Experiments
- Contig Analysis
- Modeling Species Diversity
- Summarize Results
- Two Compartment Model
- Dynamic Model for Phage and Bacteria Interactions
- Lytic and Lysogenic Phage
- Results from the Models
- Future Directions and Conclusions

Classical Marine Food Web



Marine Microbial Food Web

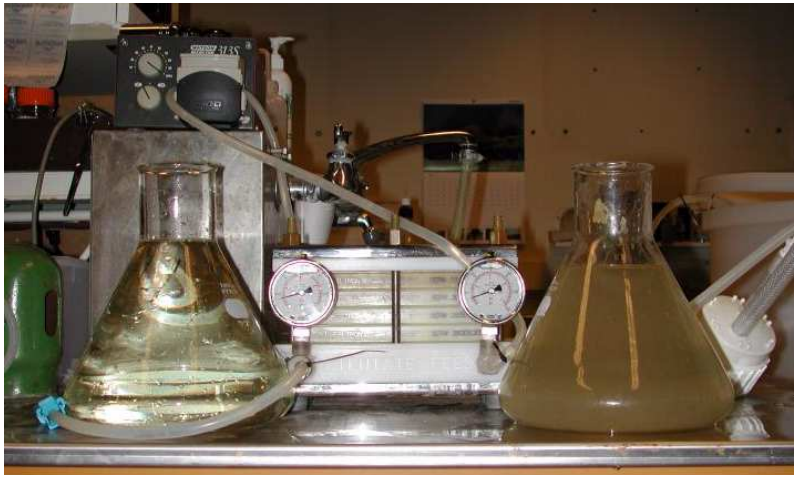
Biological Summary:

Marine Phage and Bacteria

- Estimated 1.2×10^{30} phage in the oceans
- Predominant biomass in oceans are bacteria (about 1.1×10^{13} kg of carbon)
  - Important players in global carbon cycling
  - Bacteria concentration $10^4 - 10^6$/ml
  - Phage concentration $10^5 - 10^7$/ml
- Bacterial half-life is approximately 24 hours
- About 50% of marine bacteria destroyed by phage
- Phage:Bacteria ratio is about 10:1 for many environments
- Phage are important for horizontal gene transfer
- Phage are important disease agents
  - Phage induce the toxin for cholera bacteria
  - Phage trigger the toxin for diphtheria
  - Phage genes affect virulence in Group A Streptococcus for rheumatic fever and toxic shock syndrome

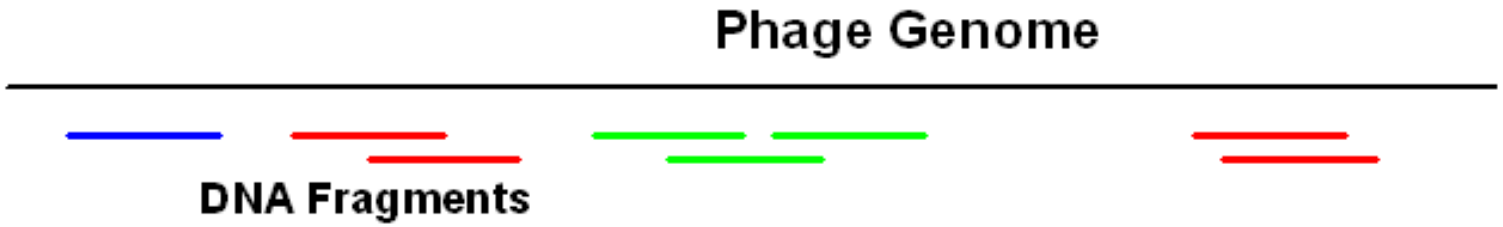
Biological Experiment

- Start with a 200 liter sample
- Filter water so only phage particles remain
- Extract the phage DNA
- Randomly break the DNA (Hydroshear)
- PCR amplify the DNA fragments
- Sequence about 1000 to create a shotgun sequence library (Linker-amplified shotgun libraries)
- Sequence lengths average 650 bp (used 663)
- Contig spectrum is obtained



What is a Contig?

- **Contigs** are contiguous sequences of DNA fragments
- An n-contig is an assembly of n overlapping DNA fragments
- An assembly is determined by 98% identity over at least 20 bp
- Below is a diagram showing a phage genome with a collection of fragments
- The diagram has one 1-contig, two 2-contigs, and a 3-contig



Experimental Contig Spectrum

- Scripp's Pier sample
  - 1021 one-contigs, 17 two contigs, 2 three contigs
- Mission Bay sample
  - 841 one-contigs, 13 two contigs, 2 three contigs
- Mission Bay Sediment Sample
  - 1152 one-contigs, 2 two contigs

Lander-Waterman Analysis - Single Genome

- Probability that two starting points on a genome of length $L = 50,000$ bp are not more than $x = 643$ (thus forming a contig) is

$$p = 1 - e^{-nx/L},$$

where n are the number of DNA fragments

- The probability that a random fragment is part of a q-contig is

$$w_q = qp^{q-1}(1-p)^2$$

a negative binomial distribution

- With n samples from the genome, the expected number of q-contigs is

$$c_q = nw_q$$

Modified Lander-Waterman Analysis

- Populations

- If there are M viral types each with populations of n_i, then the expected q-contigs observed are

$$c_q = \sum_{i=1}^{M} n_i w_{qi}$$

- Various forms of species distributions were tried and the best form for marine phages was the power law

$$n_i = ai^{-b} \qquad (1 \leq i \leq M)$$

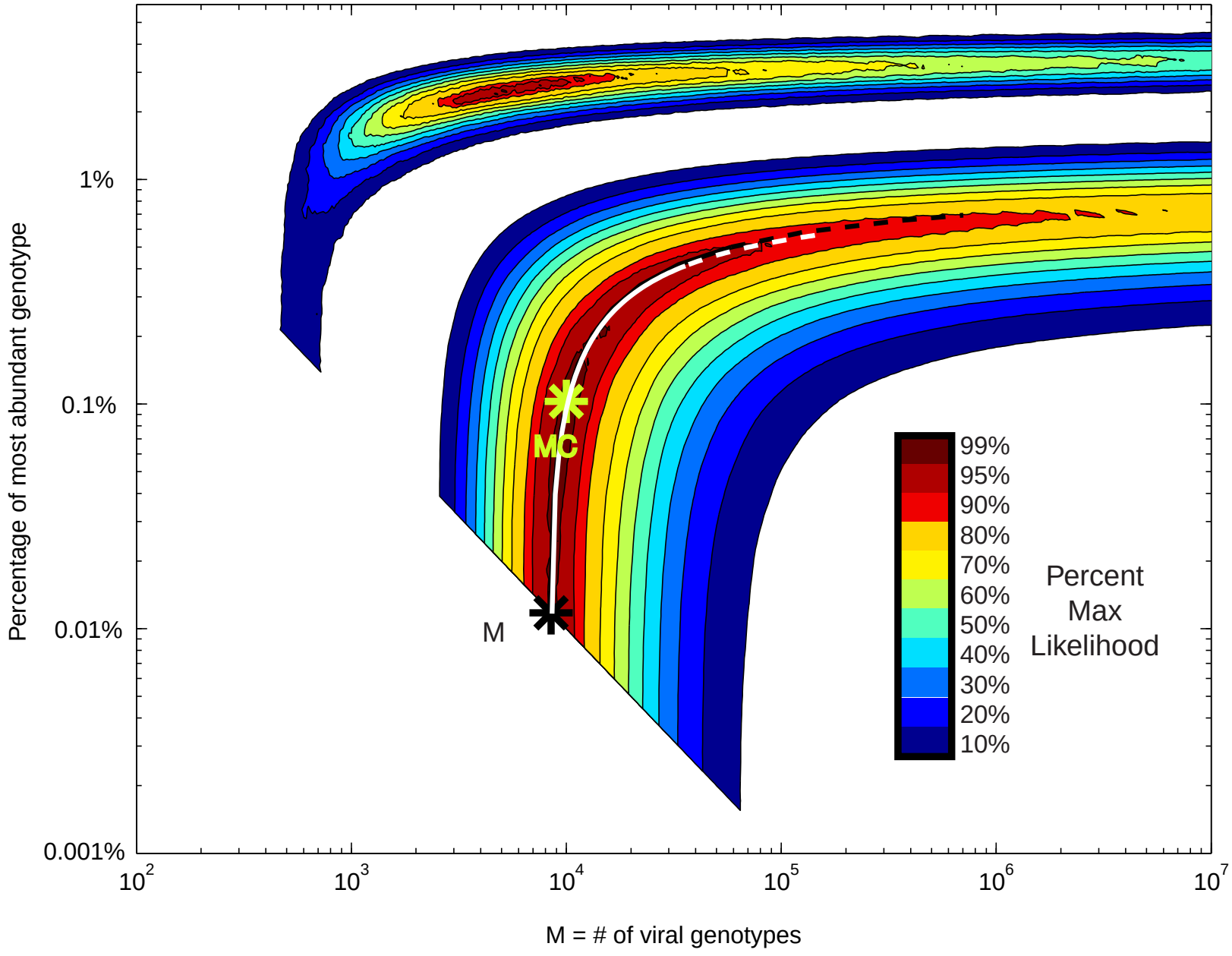
- Other population distributions tried included exponential

$$n_i = ae^{-ib} \qquad (1 \leq i \leq M),$$

logarithmic, log normal, and several others

- A Monte Carlo simulation was performed using a power law distribution with each pair of M and a values 150,000 times for a grid covering 100×500 parameter pairs for each of 3 data sets

Species Diversity



Summary of Species Diversity Analysis

All systems above were best fit by a power series distribution of species.

	% abundance a	evenness b	richness M	Shannon index
Monte Carlo				
Scripp's	1.9 ± 0.5	0.61 ± 0.06	2600 ± 800	7.4
MB	2.5 ± 0.5	0.70 ± 0.05	5100 ± 2100	7.8
MB Sed	0.1 ± 0.4	0.28 ± 0.45	10000 ± 6400	9.2
ML-W Model				
Scripp's	2.0 ± 4.5	0.64 ± 0.98	3300 ± 3000	7.6
MB	2.7 ± 5.5	0.73 ± 0.11	7000 ± 12000	8.0
MB Sed	0.012	0	8600	9.0

- Breitbart *et al* (2002) Genomic analysis of uncultured marine viral communities, PNAS **99**:14250-14255

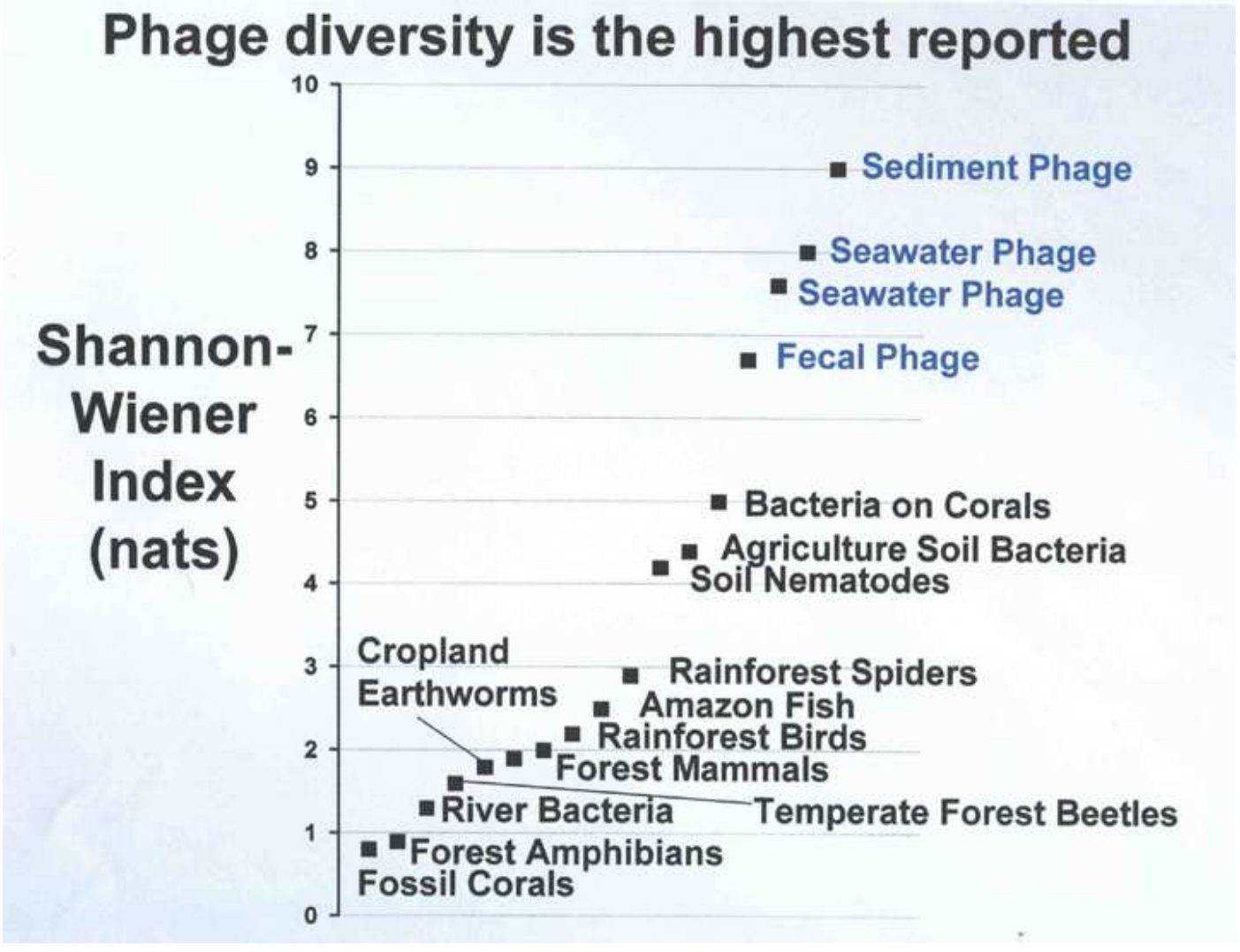
- Breitbart *et al* (2002) Diversity and population structure of a nearshore marine sediment viral community, Proc Royal Society B **271**:565-574

PHACCS -

Phage Communities from Contig Spectrum

- Our group has developed an online tool to access the biodiversity of uncultured viral communities
- Models community structure with modified Lander-Waterman algorithm
- Relative rank-abundance forms
  - Power law: $n_i = ai^{-b}, \quad 1 \leq i \leq M$
  - Logarithmic: $n_i = a(\log(i+1))^{-b}, \quad 1 \leq i \leq M$
  - Exponential: $n_i = ae^{-ib}, \quad 1 \leq i \leq M$
  - Broken stick: $n_i = \frac{N}{M} \sum_{k=i}^{M} \frac{1}{k}, \quad 1 \leq i \leq M$
  - Niche preemption: $n_i = Nk(1-k)^{i-1}, \quad 1 \leq i \leq M-1$ and $n_M = N(1-k)^M$
  - Lognormal (A more complicated popular ecological model)
- Most samples tested show Power law and Lognormal as best fits to contig spectrum, but number of species predicted is very different

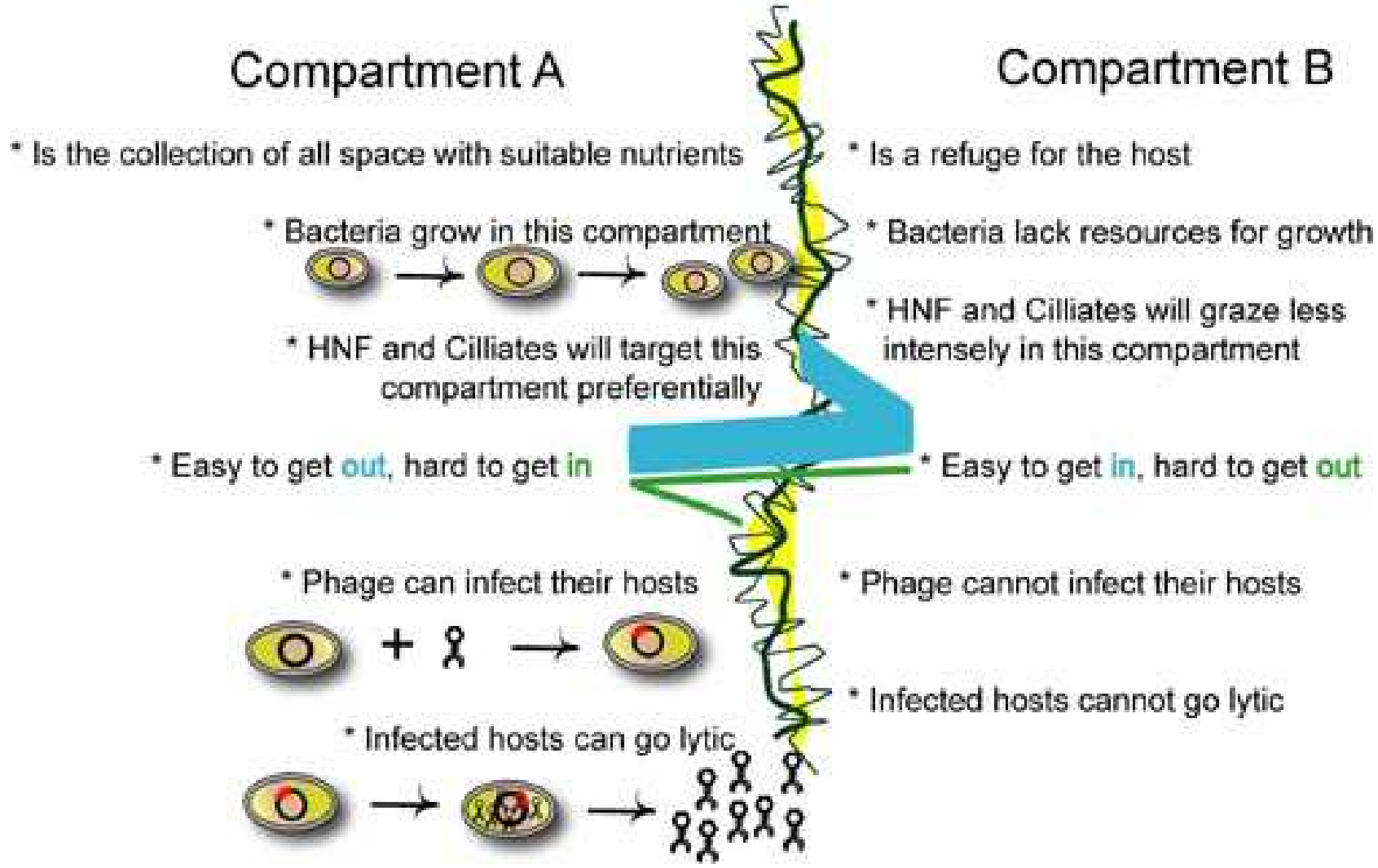
Shannon-Wiener Index



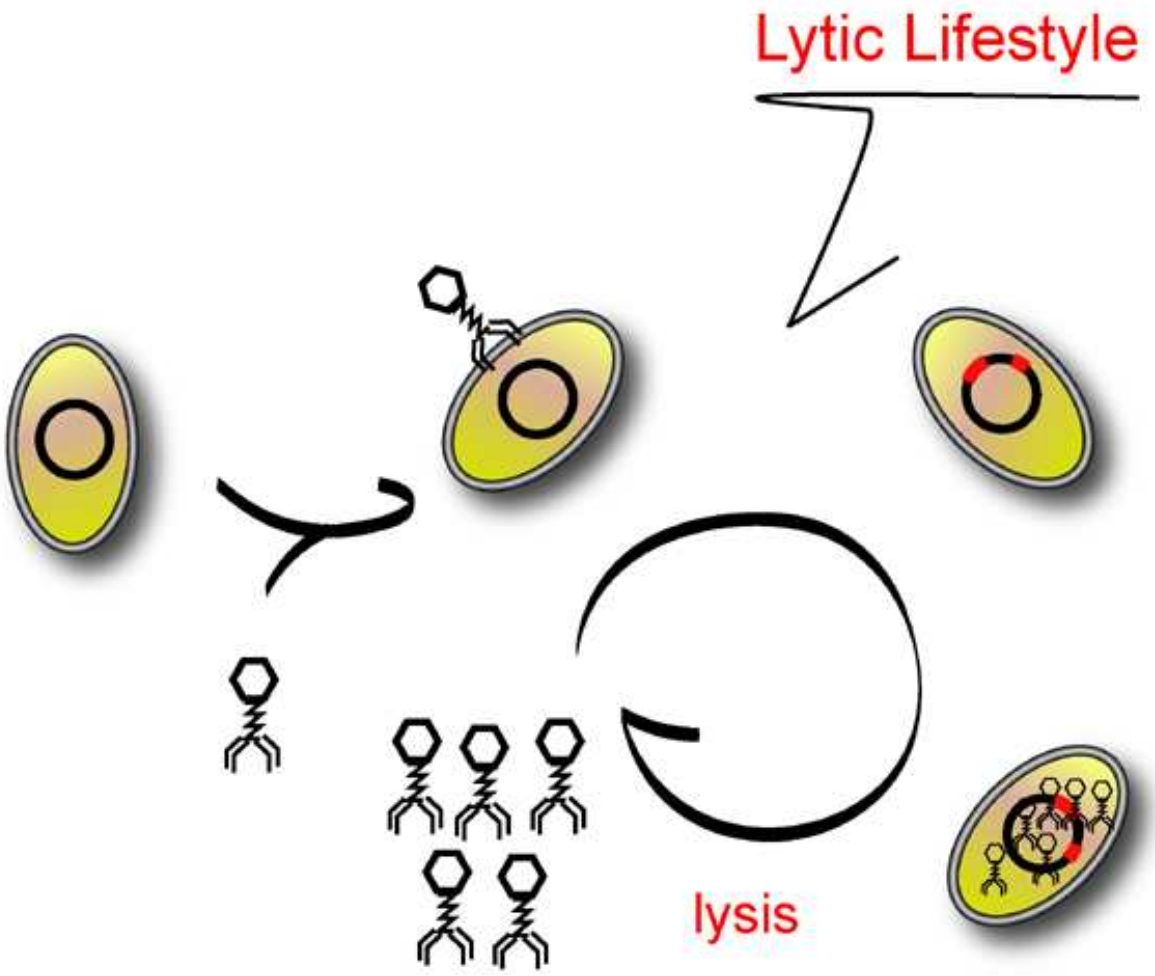
Modeling Directions and Assumptions

- Classical models based on chemostat
- Explain stable 10:1 ratio of phage to bacteria
- Ocean is a heterogeneous environment
- Create simplified single phage-host model, assuming no other interactions
- Assume this pair is roughly 1% of the total population (fairly abundant)
- Compare different strategies
  - Kill-the-winner
  - Lysogenic/lytic switch
- Narrow the parameter range

Two Compartments



Lytic Phage



Lytic Model-Phage Dynamics

$$
\begin{aligned}
\frac{dP_A(t)}{dt} &= -\gamma P_A + \beta\lambda I_A - \kappa P_A S_A \\
\frac{dI_A(t)}{dt} &= (r - g_A) I_A - \lambda I_A + \kappa P_A S_A \\
\frac{dS_A(t)}{dt} &= (r - g_A) S_A + m V_r (S_B - \alpha S_A) - \kappa P_A S_A \\
\frac{dS_B(t)}{dt} &= -g_B S_B - m V_r (S_B - \alpha S_A)
\end{aligned}
$$

Link to bifurcation

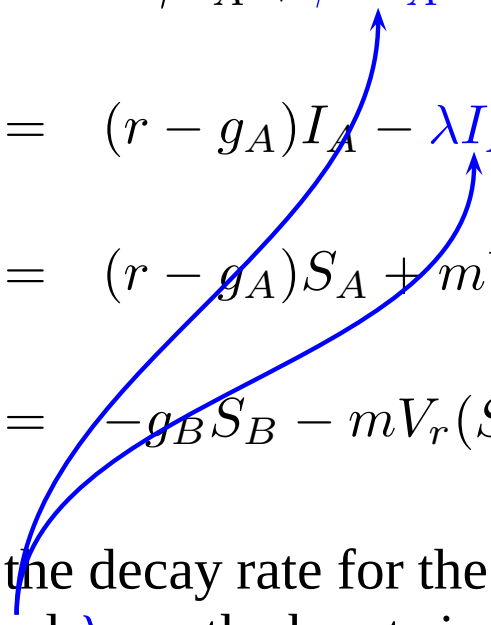
Lytic Model-Phage Dynamics

$$
\begin{aligned}
\frac{dP_A(t)}{dt} &= -\gamma P_A + \beta\lambda I_A - \kappa P_A S_A \\
\frac{dI_A(t)}{dt} &= (r - g_A) I_A - \lambda I_A + \kappa P_A S_A \\
\frac{dS_A(t)}{dt} &= (r - g_A) S_A + m V_r (S_B - \alpha S_A) - \kappa P_A S_A \\
\frac{dS_B(t)}{dt} &= -g_B S_B - m V_r (S_B - \alpha S_A)
\end{aligned}
$$

The parameter γ is the decay rate for the phage.

Lytic Model-Phage Dynamics

$$
\begin{aligned}
\frac{dP_A(t)}{dt} &= -\gamma P_A + \beta\lambda I_A - \kappa P_A S_A \\
\frac{dI_A(t)}{dt} &= (r - g_A) I_A - \lambda I_A + \kappa P_A S_A \\
\frac{dS_A(t)}{dt} &= (r - g_A) S_A + m V_r (S_B - \alpha S_A) - \kappa P_A S_A \\
\frac{dS_B(t)}{dt} &= -g_B S_B - m V_r (S_B - \alpha S_A)
\end{aligned}
$$



The parameter γ is the decay rate for the phage.
The parameters β and λ are the burst size and rate of lysis for lytic phage emerging from infected bacteria.

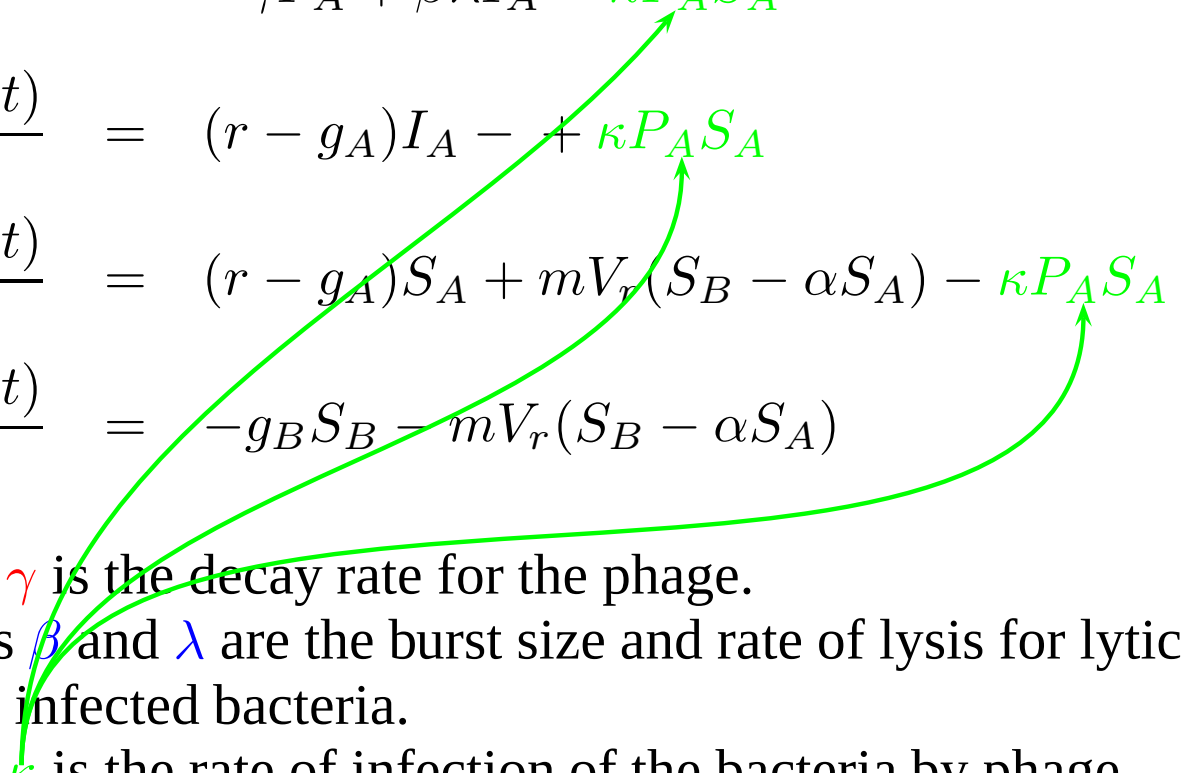
Lytic Model-Phage Dynamics

$$\frac{dP_A(t)}{dt} = -\gamma P_A + \beta\lambda I_A - \kappa P_A S_A$$

$$\frac{dI_A(t)}{dt} = (r - g_A) I_A - + \kappa P_A S_A$$

$$\frac{dS_A(t)}{dt} = (r - g_A) S_A + m V_r (S_B - \alpha S_A) - \kappa P_A S_A$$

$$\frac{dS_B(t)}{dt} = -g_B S_B - m V_r (S_B - \alpha S_A)$$



The parameter γ is the decay rate for the phage.
The parameters β and λ are the burst size and rate of lysis for lytic phage emerging from infected bacteria.
The parameter κ is the rate of infection of the bacteria by phage.

Lytic Model-Bacteria Dynamics

$$
\begin{aligned}
\frac{dP_A(t)}{dt} &= -\gamma P_A + \beta\lambda I_A - \kappa P_A S_A \\
\frac{dI_A(t)}{dt} &= (r - g_A) I_A - \lambda I_A + \kappa P_A S_A \\
\frac{dS_A(t)}{dt} &= (r - g_A) S_A + m V_r (S_B - \alpha S_A) - \kappa P_A S_A \\
\frac{dS_B(t)}{dt} &= -g_B S_B - m(S_B - \alpha S_A)
\end{aligned}
$$

The marine bacteria are divided among active infected (I_A) and susceptible (S_A) and inactive susceptible (S_B).

Lytic Model-Bacteria Dynamics

$$\begin{aligned}
\frac{dP_A(t)}{dt} &= -\gamma P_A + \beta\lambda I_A - \kappa P_A S_A \\
\frac{dI_A(t)}{dt} &= (r - g_A) I_A - \lambda I_A + \kappa P_A S_A \\
\frac{dS_A(t)}{dt} &= (r - g_A) S_A + m V_r (S_B - \alpha S_A) - \kappa P_A S_A \\
\frac{dS_B(t)}{dt} &= -g_B S_B - m(S_B - \alpha S_A)
\end{aligned}$$

The parameter r is the growth rate for the bacteria.

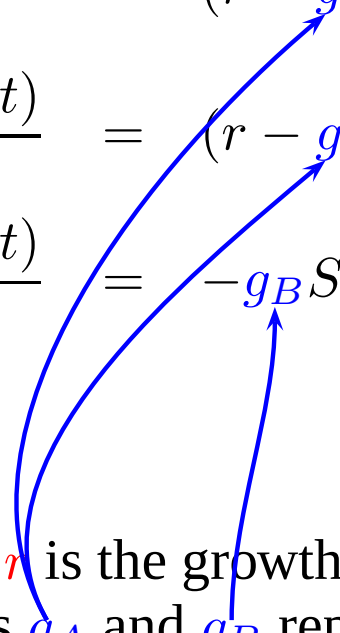
Lytic Model-Bacteria Dynamics

$$\frac{dP_A(t)}{dt} = -\gamma P_A + \beta\lambda I_A - \kappa P_A S_A$$

$$\frac{dI_A(t)}{dt} = (r - g_A) I_A - \lambda I_A + \kappa P_A S_A$$

$$\frac{dS_A(t)}{dt} = (r - g_A) S_A + m V_r (S_B - \alpha S_A) - \kappa P_A S_A$$

$$\frac{dS_B(t)}{dt} = -g_B S_B - m(S_B - \alpha S_A)$$



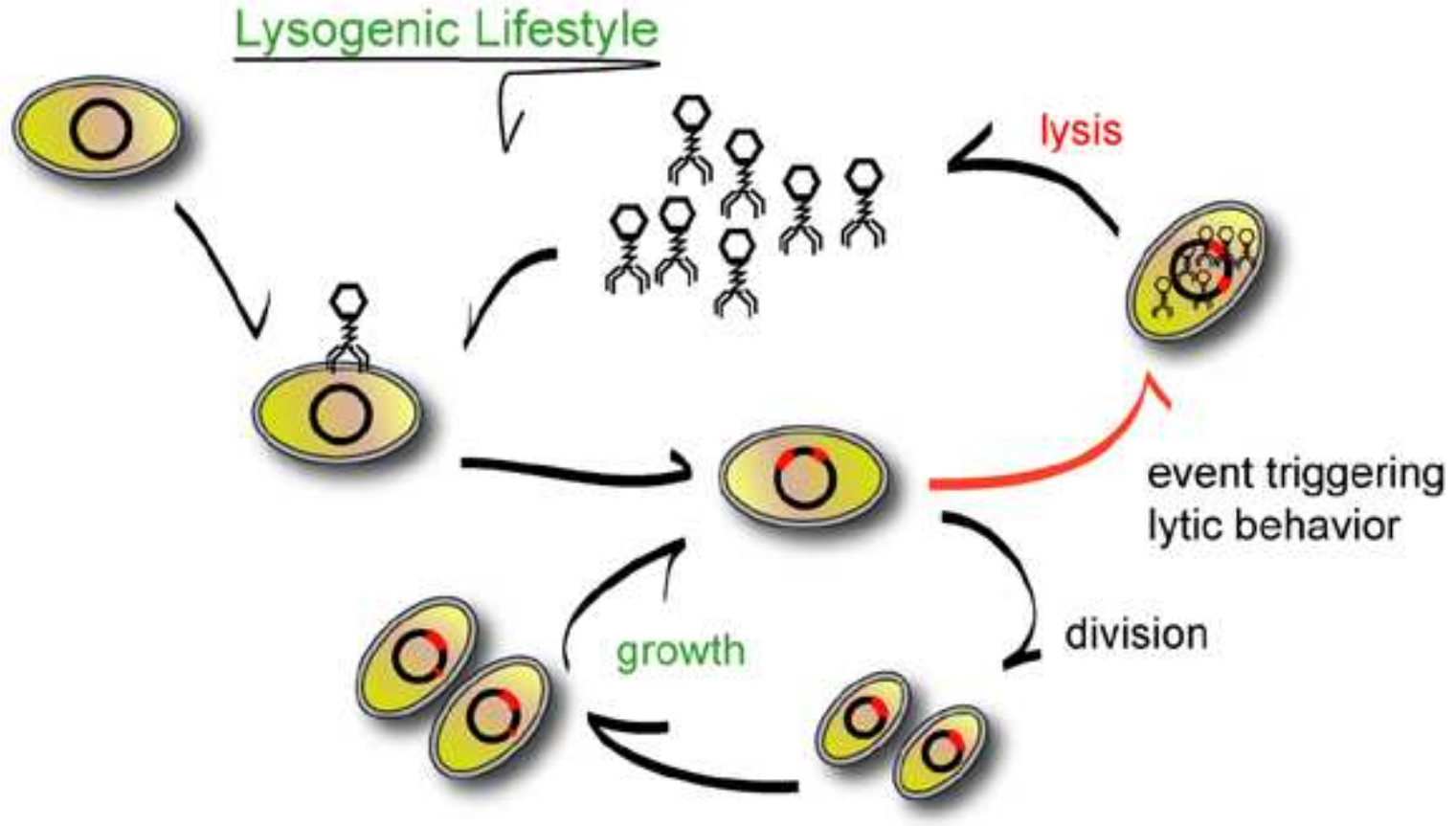
The parameter r is the growth rate for the bacteria.
The parameters g_A and g_B represent the grazing of the protists on the bacteria.

Lytic Model-Bacteria Dynamics

$$
\begin{aligned}
\frac{dP_A(t)}{dt} &= -\gamma P_A + \beta\lambda I_A - \kappa P_A S_A \\
\frac{dI_A(t)}{dt} &= (r - g_A) I_A - \lambda I_A + \kappa P_A S_A \\
\frac{dS_A(t)}{dt} &= (r - g_A) S_A + m V_r (S_B - \alpha S_A) - \kappa P_A S_A \\
\frac{dS_B(t)}{dt} &= -g_B S_B - m (S_B - \alpha S_A)
\end{aligned}
$$

The parameter r is the growth rate for the bacteria.
The parameters g_A and g_B represent the grazing of the protists on the bacteria.
The parameter m is the migration rate of the bacteria between Compartments A and B with the scaling for volume V_r, and α represents the fraction not adhering to nutrients.

Lysogenic Phage



Lysogenic Model

$$
\begin{aligned}
\frac{dP_A(t)}{dt} &= -\gamma P_A + \beta\lambda I_A - \kappa P_A S_A \\
\frac{dI_A(t)}{dt} &= (r - g_A - \lambda) I_A + \kappa P_A S_A + m V_r (I_B - \alpha I_A) \\
\frac{dI_B(t)}{dt} &= -g_B I_B - m(I_B - \alpha I_A) \\
\frac{dS_A(t)}{dt} &= (r - g_A) S_A - \kappa P_A S_A + m V_r (S_B - \alpha S_A) \\
\frac{dS_B(t)}{dt} &= -g_B S_B - m(S_B - \alpha S_A)
\end{aligned}
$$

Lysogenic Model

$$
\begin{aligned}
\frac{dP_A(t)}{dt} &= -\gamma P_A + \beta\lambda I_A - \kappa P_A S_A \\
\frac{dI_A(t)}{dt} &= (r - g_A - \lambda) I_A + \kappa P_A S_A + m V_r (I_B - \alpha I_A) \\
\frac{dI_B(t)}{dt} &= -g_B I_B - m(I_B - \alpha I_A) \\
\frac{dS_A(t)}{dt} &= (r - g_A) S_A - \kappa P_A S_A + m V_r (S_B - \alpha S_A) \\
\frac{dS_B(t)}{dt} &= -g_B S_B - m(S_B - \alpha S_A)
\end{aligned}
$$

The only difference in this lysogenic model for the marine environment is that the $r > \lambda$, so the infected bacteria survive long enough to migrate to Compartment B.

Parameters

- Many parameters are difficult to measure
- Growth, burst size, and lysis timing vary with conditions
- Phage decay rates vary widely in the literature

Constraints

- Need approximately 10:1 phage to bacteria ratio
- Turnover of bacteria about 24 hour
- Limited range on many parameters in the literature

Simulation of Models

First the lytic model was fit to reasonable parameters.

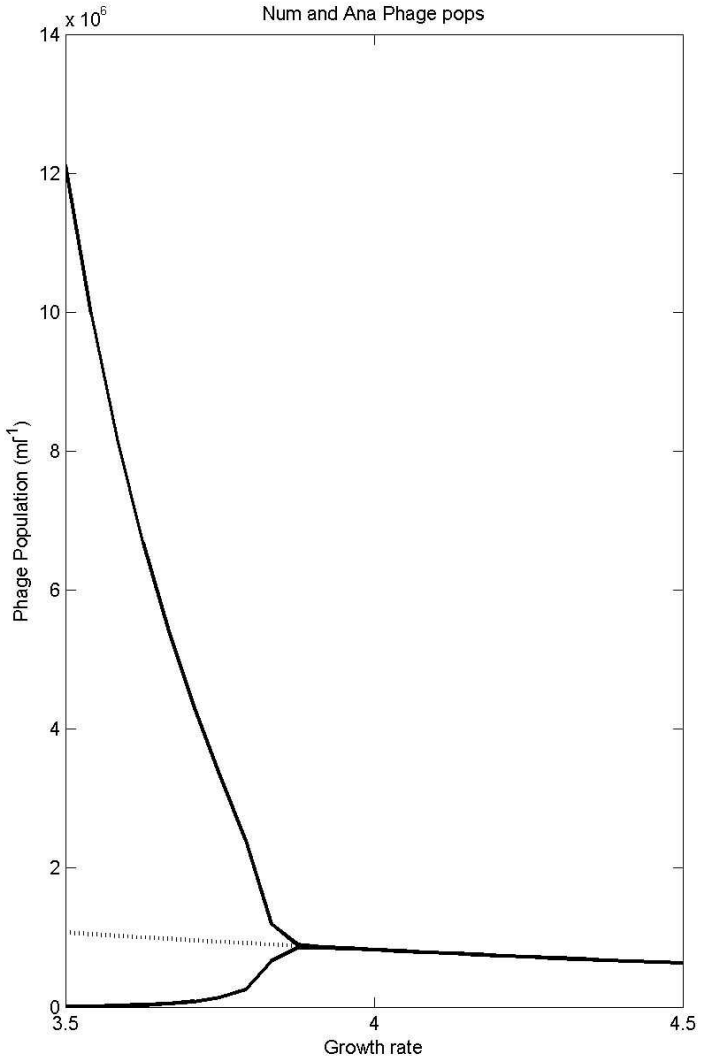
	Lytic	Lysogenic	Lysogenic
Changes	-	λ	r, α, V_r
Phage	7.46×10^5	7.46×10^5	8.06×10^5
Bacteria	5.88×10^4	1.14×10^5	5.10×10^4
Ratio	14.4:1	5.93:1	15.8:1
% Inactive	81 %	90 %	88 %
% Infected	10.3 %	62.2 %	94.8 %
Turnover	24.2 hr	47.3 hr	25.3 hr
Behavior	Stable	Stable	Stable

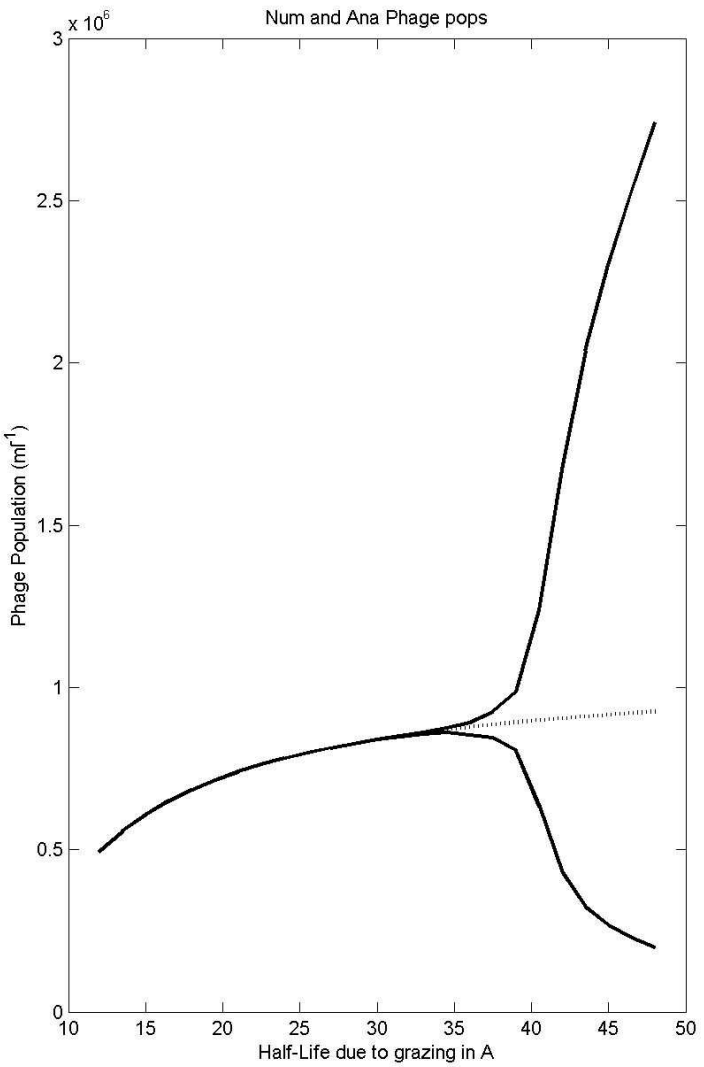
Compartment B acts like a refuge with most bacteria there.
Lysogeny results in many more infected bacteria.

Parameter Sensitivity

1. The growth parameter r had the greatest effect
2. Rate of lysis λ of bacteria by phage
3. Parameter α representing fraction of bacteria available to diffuse into Compartment B
4. Grazing by protists g_A in Compartment A
5. ...
6. Minimal effects by g_B, m, and κ

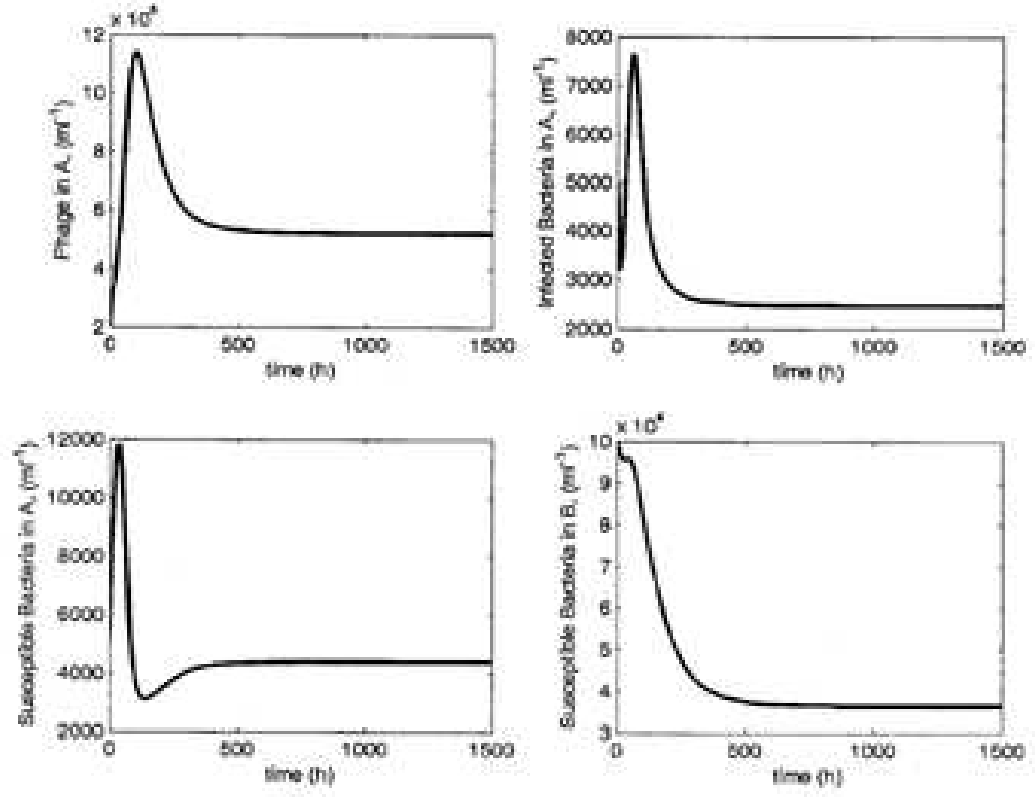
Bifurcation Study (Lytic Model) - r and g_A





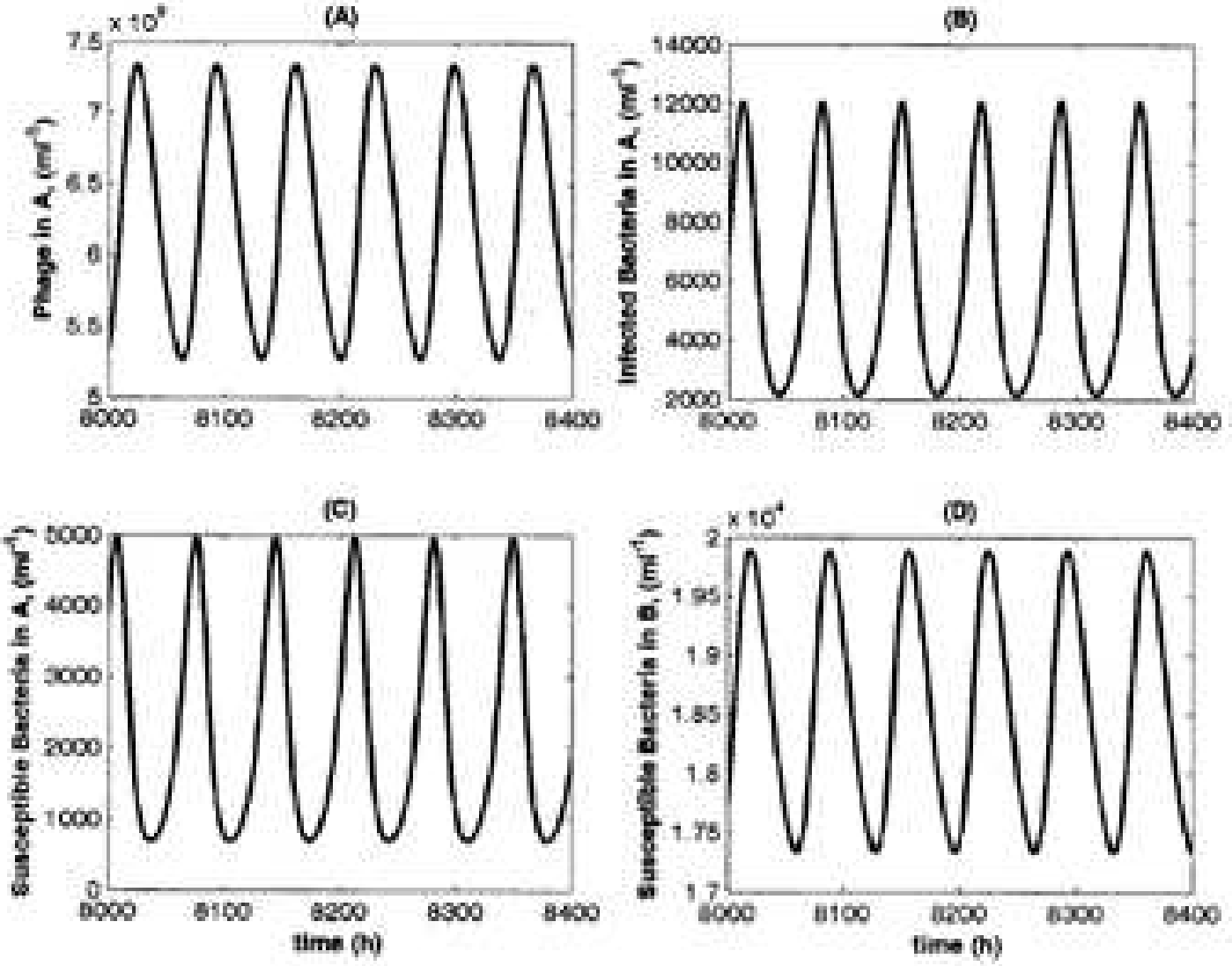
Link to lytic model

Lytic Results

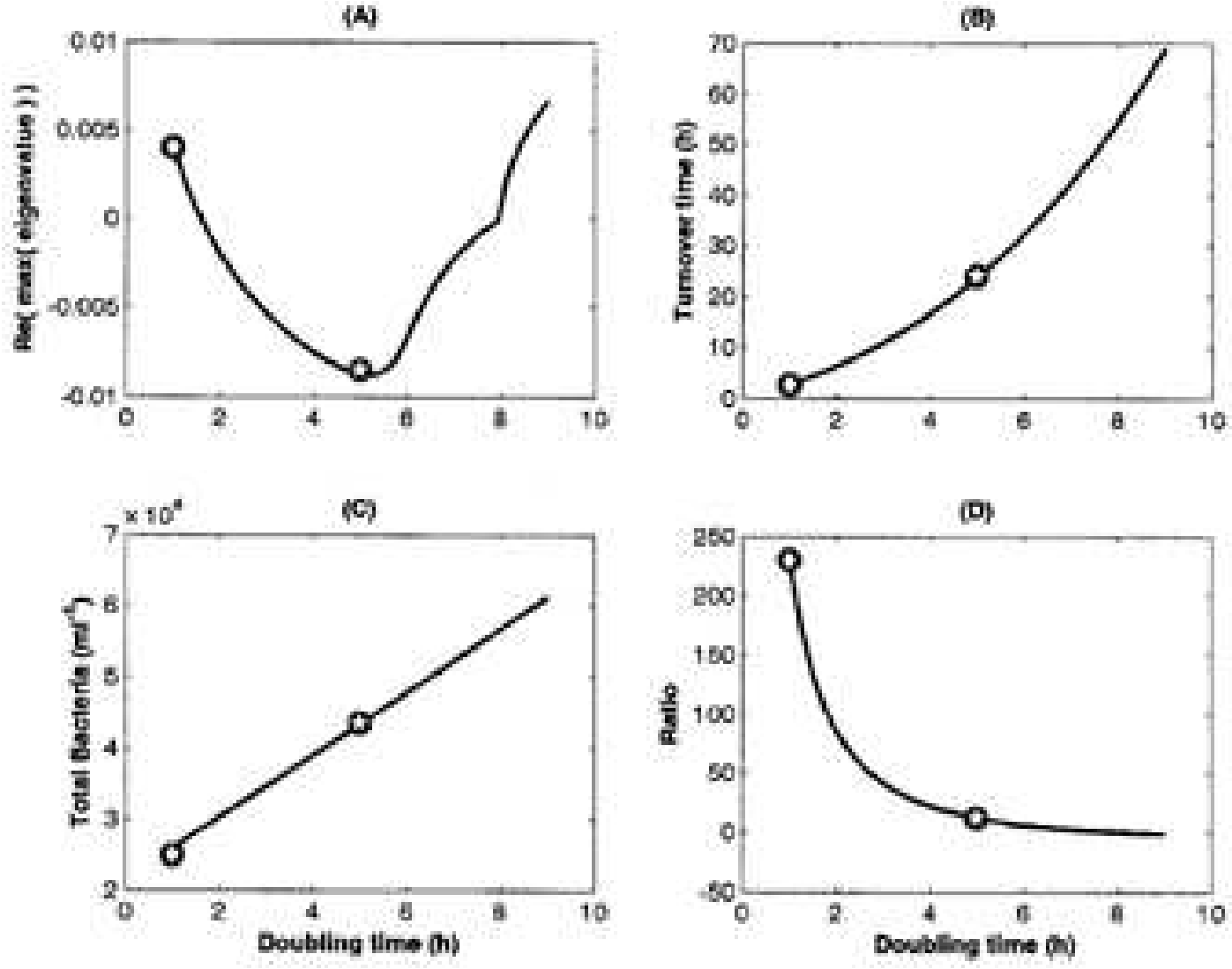


The equilibrium phage population is about 4.5×10^5, while the equilibrium bacteria population is about 3.8×10^4 in Compartment B (80%) and 0.7×10^4 in Compartment A.

Lytic Results



Lytic Results



Results of Lytic Model

- Two equilibria
- Found reasonable parameters
  - About 10:1 ratio of phage to bacteria
  - Approximately net 24 hour for bacterial half-life
  - Many parameters span a wide range, yet maintain biologically feasible solutions
- Stable equilibrium for marine conditions
- Oscillatory solutions for chemostat conditions

Results of Lysogenic Model

- Two equilibria
- Similar to the Lytic model except
  - Only stable behavior observed for non-trivial equilibrium
  - Parameters span a narrower range for biologically feasible solutions

Quorum Switching Model

- Assume phage become lytic when sensing sufficient active bacteria
- Combines lytic and lysogenic models with changes below
  - Lytic part of model includes infected inactive bacteria
  - Lysogenic part of model has no terms for lysis

Results of Quorum Switching Model

- Only some preliminary numerical results
- Mixing in active compartment leaves most bacteria infected
- Oscillating solution with Malthusian growth through threshold, then lysis decays to lower population

Future Directions

- Use studies for two NSF Biocomplexity grants
  - Help explain possible lytic/lysogenic switching behavior (Seasonal in Tampa Bay)
  - Explain varying diversity and concentrations (Solar Saltern study)
- Add nutrient or other limiting factor to 2-compartment model
- Include delays for lysis in model
- Examine additional refuge compartment or spatial component
- Perform detailed mathematical analysis

Conclusion

- Shotgun libraries of DNA from phage can be analyzed for species diversity
- Contig analysis often fits a power law giving estimates of species abundance, evenness, and diversity
- Automated program PHACCS for choosing rank-abundance model
- Heterogeneous environment suggests at least two compartments or some spatial component in model
- Dynamic models exhibit several behaviors
- Dynamic models aid parameter selection

Collaborators

- Beltran Rodriguez - Computational Sciences (SDSU-student)
- Anca Segall - Biology (SDSU)
- John Paul - Biology (Southern Florida University)
- Forest Rohwer - Biology (SDSU)
- Florent Angly - Biology (SDSU-student)
- Mya Breitbart - Biology (SDSU-student)
- Peter Salamon - Math (SDSU)
- Ben Felts
- Jim Nulton
- Numerous other students have contributed

Support from two NSF Biocomplexity grants