Title: Oscillating bacterial expression states generate herd immunity to viral infection

Authors: Christopher J.R. Turkington[1,3], Andrey Morozov[2], Martha R.J. Clokie[3], Christopher D. Bayliss[1]

Affiliations:

Department of Genetics and Genome Biology, University of Leicester, Leicester, LE1 7RH, UK.

[2]Department of Mathematics, University of Leicester, Leicester, LE1 7RH, UK.

[3]Department of Infection, Immunity and Inflammation, University of Leicester, Leicester, LE1 7RH, UK.

*Correspondence to: cdb12@le.ac.uk

Abstract:

Hypermutable loci are widespread in bacteria as mechanism for rapid generation of phenotypic diversity for survival of fluctuating, often antagonistic, selection pressures. However, it is unclear how hypermutation contributes to survival of multiple bacterial populations. Here we developed an 'oscillating prey assay' to examine phage spread through populations of *Haemophilus influenzae* wherein expression of the phage receptor gene (*lic2A*) is switched 'ON' and 'OFF' by mutations in a hypermutable tetranucleotide repeat tract. Phage extinction was frequently observed when the proportion of phage-resistant sub-populations exceeded 34%. *In silico* modelling indicated that phage extinction was interdependent on phage loss during transfer between populations and the frequency of resistant populations. We conclude that phase-variable hypermutable loci produce bacterial 'herd immunity' with resistant intermediary-populations acting as a barricade to reduce the viral load faced by phage-sensitive sub-populations. This paradigm of metapopulation protection is applicable to evolution of hypermutable loci in multiple bacteria-phage and host-pathogen interactions.

One Sentence Summary:

Hypermutable loci provide bacteria with a form of 'herd-immunity' against viral predation, protecting sensitive cells from attack by these infectious agents.

Main Text:

Hypermutable loci as mediators of survival of constantly fluctuating selection pressures are a predictable outcome of the evolution of evolvability as stated in the Red Queen hypothesis (*1*,*2*). Nonetheless we only have a limited understanding of the evolutionary pathways leading to hypermutable loci. During persistence in the human host, bacteria must navigate frequent exposure to both hostile host environments and predation by bacteriophages (phages). One class of hypermutable loci facilitate survival of these selective pressures by pre-emptive, frequent and reversible 'ON/OFF' generation of adaptive variants in a process known as 'phase variation' (*2-5*). Phage receptors are often subject to phase variation as these surface-located molecules confer advantages, such as immune resistance, in the 'ON' state but have to switch into an 'OFF' state to prevent killing by phages. Phase variable phage receptors have known benefits for individual populations, but could also confer significant advantages on geographically-adjacent sub-populations.

Phase variation can alter the resistance/susceptibility of bacteria to phage infection by producing 'ON/OFF' switches in expression of surface structures that are the receptors for phage attachment (*6-9*). This phenomenon is common to several bacterial species and mechanisms of phage defence (*6–11*). A major mechanism of phase variation involves increases and decreases in the numbers of identical DNA repeats (microsatellites) by slipped strand mispairing during DNA replication (*3,5,12*). The obligate human respiratory commensal and pathogen *H. influenzae* contains an expansive array of phase variable loci (*13–19*). Several of these phase variable loci are required for addition of sugar molecules onto the surface-exposed outer-core of the lipooligosaccharide (LOS) (*5,13*), allowing these bacteria to generate a high degree of surface variance. Phase variation of the UDP-galactose-LOS-galactosyltransferase encoding gene, *lic2A*, is mediated by repeat number changes in a tandem 5'CAAT repeat tract present in the open-reading frame (*20*). In *H. influenzae* strain Rd, the ON state of *lic2A* is associated with sensitivity to infection with phage HP1c1 while the OFF state is phage-resistant (*21*).

Phase variation of *lic2A* was shown to prevent propagation of phage HP1c1 in individual populations (*21*) but it is not known whether inhibitory effects would also occur as the phage spreads through multiple geographical-linked sub-populations. We developed 'the oscillating prey assay' to simulate phage spread through bacterial meta-populations. The assay comprised a known series of multiple resistant (*lic2A* phase OFF) and sensitive (*lic2A* phase ON) bacterial sub-populations (Figure S1). The HP1c1 phage was passaged through each sub-population for 50 minutes to allow for one cycle of phage replication. Loss of phages occurs during transmission between bacterial populations due to abiotic (e.g. pH change) and biotic factors (e.g. cilial clearance) but is also substantially impacted by stochastic dispersal over the surfaces between bacterial populations. The effect of dispersal will increase as a function of the distance between infected and non-infected bacterial populations such that low rates of loss occur during transmission between micro-colonies in the respiratory tract but high loss rates during inter-respiratory transfer between human hosts colonized with *H. influenzae*. Losses during transmission were arbitrarily simulated in our experimental model by imposing a 10-fold dilution of the phage population at each transfer.

Initial experiments were performed using defined patterns of linear oscillations between phage-resistant and phage-sensitive phase states of the *lic2A* gene (see Figure 1). Survival and propagation of phage in these assays was dependent on the proportion of phage-resistant sub-populations in each series of 20 passages. Survival of phage to the final passage was only observed when the proportion of phage-resistance populations was 34% or 0% (S66 and S100, respectively; Figure 1). Extinction events occurred within 5 to 16 cycles in all other heterogeneous and homogeneous populations at a rate that was dependent on the proportion of resistant sub-populations. Major decreases in phage density were also observed as exemplified for the S66 population series where densities were significantly decreased by the end of cycle 20 (Figure. 1; paired *t*-test: $t = 4.97$, $P < 0.05$; mean ± SEM PFU ml^{-1} values for phage densities at cycle 0 = $7.16 \pm 0.26 \times 10^5$ and cycle 20 = $2.82 \pm 0.9 \times 10^4$), indicating that further passages with a similar pattern would have resulted in extinction events. Contrastingly, phage density increased when all sub-populations were phage sensitive (S100). Thus, phage survival and density were both limited by the frequency of encounters with phage-resistant *lic2A* OFF phase variants during passage through a linear series of sub-populations.

Our exploration of the frequency of phage extinction events was extended to a wider range of linear and non-linear cycling patterns and combined with differing extents of phage loss

through a mathematical model. This model utilized key phage parameters for the adsorption, replication time, burst size and stability of phage HP1c1 (see the parameter setting experiments in Figure S2) and closely mimicked the experimental oscillation assay (Figure S3). In this model, the number of oscillations was extended to 105 viral replicative cycles, while cycling patterns were randomised for a specific overall proportion of *lic2A* ON/OFF sub-populations. Individual runs of the mathematical model produced comparable findings to the experimental setting with phage densities fluctuating from cycle-to-cycle and reducing as a function of the proportion of *lic2A* ON sub-populations in each series (Figure S4). Phage extinction events occurred in a stochastic manner within a similar but wider band of replication cycles demonstrating that random patterns of ON/OFF expression states for a phage receptor can limit phage spread.

Multiple simulations (n = 200) were performed for each combination of R (the percentage of phage-resistant *lic2A* OFF sub-populations) and C (the inverse of the dilution coefficient). Average phage densities were measured for all cycles and runs of each combination of R and C, and phage extinction was defined to occur whenever the density fell below 100 PFU ml^{-1}. Phage extinction was always observed when the number of resistant states exceeded 70% of the sub-populations even at a low dilution rate of 1 in 2 (C=0.5) (Figure 2). Similarly, when phage loss was >80% at each transfer (i.e. a dilution constant of 1:50, C=0.02), extinction of the phage population was observed for all population structures except those consisting of >95% sensitive bacterial populations (Figure 2). In between these extremes, there is an accelerating trade-off between R and C with respect to phage extinction and the average phage density; decreases in dilution rate are countered by a high prevalence of resistant sub-populations enabling bacterial populations to survive even when phage dispersal is low (Figure 2). This observation suggests that on-going evolution of localised hypermutability for a particular surface exposed bacterial epitope could be tuned to the prevalence and density of phages capable of using the phase-variable epitope as a receptor for surface attachment.

Both our experimental and mathematical oscillating prey assays demonstrated that variations in the linear pattern of phase-variant could influence phage spread. However, the distribution of phase variants across a surface (e.g. microcolonies on upper respiratory tract surfaces for *H. influenzae*) is anticipated to be random and hence to result in spatial effects on phage spread. Indeed, spatially structured environments are known to restrict spread of phages (*22*). The effects of spatial structuring were therefore explored by examining transmission of phage across a meta-population of fixed dimensions with varying levels of heterogeneity in *lic2A* expression (Figure S6). This fixed-area oscillating prey assay was initiated by inoculating phage into a single well of a 96-well plate and then expanding outward with each replicative cycle being seeded into neighbouring wells (Figure S5). Passage of the HP1c1 phage through the fixed-area oscillating prey assay resulted in uneven phage densities (Figure 3, and Figure S7). When 66% of populations were phage-sensitive (Figure 3b), densities ranged from 10^3 PFU ml^{-1} to 10^{10} PFU ml^{-1}, whereas densities always exceeded 10^8 PFU ml^{-1} if all population were sensitive (Figure 3a). Phage densities in specific regions were dependent on the direction of propagation through the randomly distributed sub-populations with passage through multiple resistant or sensitive sub-populations resulting in low or high phage densities, respectively. This model shows how spatial population heterogeneity created by a hypermutable locus can prevent equal dissemination of phage through host populations and allows survival of phage-sensitive sub-populations whose phenotypes may be beneficial for bacterial survival of other selective pressures. This irregular dissemination of phage may also have ramifications for evolution of hypermutable loci and other phenotypes in bacteria.

In environments where phage are present and the rate of loss is low, bacterial cells either lacking *lic2A* or with the gene in an OFF state would have a significant fitness advantage over cells with constitutive expression of this phage receptor. In human volunteer studies of colonisation with *H. influenzae*, expression of *lic2A* was not essential for human nasopharyngeal colonisation (*22*). However, the *lic2A* ON state has been associated with non-typeable *H. influenzae* pneumonia (*23*) while in *in vitro* studies, *lic2A*-dependent epitopes aid in survival against human immune responses (*24*). Analysis of 104 *H. influenzae* genome sequences indicated that *lic2A* is in an ON state in ~63% of all isolates and is predominantly in the ON state for multiple disease conditions, suggesting selection for expression of *lic2A* occurs across a wide range of niches for *H. influenzae* (Figure S8). These observations are consistent with a scenario where high densities of phage drive evolution of hypermutability as the *lic2A* gene oscillates between immune-resistant/phage-sensitive and immune-sensitive/phage-resistant states. As switching rates of phase-variable genes can evolve in response to alternating selection pressures (*25*-*27*), we developed a mathematical model to examine the impact of different switching rates and immune selection on the proportions of *lic2A* expression states (Supplementary text). Weak or no selection combined with medium to high switching rates resulted in phage resistant variants exceeding 1% of the population whereas low switching rates combined with weak selection drives resistant variants to <1%. We conclude that observed switching rates will generate spatially-structured populations where the probability of encountering phage resistant cells is between 0.01 and 0.3, which, in combination with high dilution rates (>1:10), are capable of inhibiting phage spread.

The generation of heterogeneity in *lic2A* expression states by localised hypermutation is a form of 'bacterial herd-immunity' wherein phage spread is retarded by resistant sub-populations creating barriers between the phage and phage-sensitive bacterial sub-populations. This model is analogous to population-based herd immunity produced by epidemic spread of a pathogen or vaccination programs in humans and animals. Theoretical models of heterogeneity in immunity/sensitivity to pathogens exhibit similar properties with reductions in the prevalence of human associated diseases being correlated with the level of vaccine coverage within a population (*28*). Our demonstration of a hypermutable locus retarding spread of an infectious agent within a prokaryotic meta-population suggests that the herd immunity phenomenon may be applicable to a wide variety of biological organisms and have deep evolutionary roots. An extension of these models would utilise our conceptual framework to explore how somatic hypermutation, a key example of localised hypermutation in eukaryotes, could have evolved through selection for sub-population heterogeneity linked to pathogen resistance.

References and Notes:

1. L. Van Valen, A new evolutionary law. *Evolutionary Theory* **1**, 1-30 (1973).

2. E. R. Moxon, P. B. Rainey, M. A. Nowak, R. E. Lenski. Adaptive evolution of highly mutable loci in pathogenic bacteria. *Current Biology* **4**, 24-33 (1994).

3. C. Bayliss, Determinants of phase variation rate and the fitness implications of differing rates for bacterial pathogens and commensals. *FEMS Microbiol. Rev.* **33,** 504–20 (2009).

4. P.B. Rainey, H. J. E. Beaumont, G. C. Ferguson, J. Gallie, C. Kost, E. Libby, X. X. Zhang, The evolutionary emergence of stochastic phenotype switching in bacteria. *Microb. Cell Fact.* **10 Suppl 1,** S14 (2011).

5. E. R. Moxon, C. Bayliss, D. W. Hood, Bacterial contingency loci: the role of simple sequence DNA repeats in bacterial adaptation. *Annu. Rev. Genet.* **40,** 307–33 (2006).

6. K. D. Seed, S. M. Faruque, J. J. Mekalanos, S. B. Calderwood, F. Qadri, A. Camilli, Phase variable O antigen biosynthetic genes control expression of the major protective antigen and bacteriophage receptor in *Vibrio cholerae* O1. *PLoS Pathog.* **8,** e1002917 (2012).

7. O. Sekulovic, M. Ospina Bedoya, A. S. Fivian-Hughes, N. F. Fairweather, L. C. Fortier, The *Clostridium difficile* cell wall protein CwpV confers phase-variable phage resistance. *Mol. Microbiol.* **98,** 329–342 (2015).

8. I. Cota, A. B. Blanc-Potard, J. Casadesús, STM2209-STM2208 (opvAB): a phase variation locus of *Salmonella enterica* involved in control of O-antigen chain length. *PLoS One* **7,** e36863 (2012).

9. M. C. H. Sørensen, L. B. van Alphen, A. Harboe, J. Li, B. B. Christensen, C. M. Szymanski, L. Brøndsted, Bacteriophage F336 recognizes the capsular phosphoramidate modification of *Campylobacter jejuni* NCTC11168. *J. Bacteriol.* **193,** 6742–6749 (2011).

10. L. Latino, C. Midoux, Y. Hauck, G. Vergnaud, C. Pourcel, Pseudolysogeny and sequential mutations build multiresistance to virulent bacteriophages in Pseudomonas aeruginosa. *Microbiology* **162,** 748–63 (2016).

11. A. Anjum, K. J. Brathwaite, J. Aidley, P. L. Connerton, N. J. Cummings, J. Parkhill, I. Connerton, C. D. Bayliss, Phase variation of a Type IIG restriction-modification enzyme alters site-specific methylation patterns and gene expression in *Campylobacter jejuni* strain NCTC11168. *Nucleic Acids Res.* gkw019 (2016). doi:10.1093/nar/gkw019

12. M. W. van der Woude, A. J. Bäumler, Phase and antigenic variation in bacteria. *Clin. Microbiol. Rev.* **17,** 581–611, table of contents (2004).

13. P. M. Power, W. A. Sweetman, N. J. Gallacher, M. R. Woodhall, G. A. Kumar, E. R. Moxon, D. W. Hood, Simple sequence repeats in *Haemophilus influenzae*. *Infect. Genet. Evol.* **9,** 216–28 (2009).

14. C. D. Bayliss, D. Field, E. R. Moxon. The simple sequence contingency loci of *Haemophilus influenzae* and *Neisseria meningitidis*. *J. Clin. Invest.* **107,** 657–62 (2001).

15. J. S. Kroll, E. R. Moxon, Capsulation and gene copy number at the cap locus of *Haemophilus influenzae* type b. *J. Bacteriol.* **170,** 859–64 (1988).

16. Z. Ren, H. Jin, P. W. Whitby, D. J. Morton, T. L. Stull, Role of CCAA nucleotide repeats in regulation of hemoglobin and hemoglobin-haptoglobin binding protein genes of *Haemophilus influenzae*. *J. Bacteriol.* **181,** 5865–70 (1999).

17. S. Dawid, S. J. Barenkamp, J. W. St Geme, Variation in expression of the *Haemophilus influenzae* HMW adhesins: a prokaryotic system reminiscent of eukaryotes. *Proc. Natl. Acad. Sci. U. S. A.* **96,** 1077–82 (1999).

18. S. W. Glover, A. Piekarowicz, Host specificity of DNA in *Haemophilus influenzae*: restriction and modification in strain Rd. *Biochem. Biophys. Res. Commun.* **46,** 1610–7 (1972).

19. X. De Bolle, C. D. Bayliss, D. Field, T. van de Ven, N. J. Saunders, D. W. Hood, E. R. Moxon, The length of a tetranucleotide repeat tract in *Haemophilus influenzae* determines the phase variation rate of a gene with homology to type III DNA methyltransferases. *Mol. Microbiol.* **35,** 211–222 (2000).

20. N. J. High, M. E. Deadman, E. R. Moxon, The role of a repetitive DNA motif (5'-CAAT-3') in the variable expression of the *Haemophilus influenzae* lipopolysaccharide epitope alpha Gal(1-4)beta Gal. *Mol. Microbiol.* **9,** 1275–82 (1993).

21. P. Zaleski, M. Wojciechowski, A. Piekarowicz. The role of Dam methylation in phase variation of *Haemophilus influenzae* genes involved in defence against phage infection. *Microbiology* **151,** 3361–9 (2005).

22. J. Poole, E. Foster, K. Chaloner, J. Hunt, M. P. Jennings, T. Blair, K. Knudtson, E. Christensen, R. S. Munson Jr, P. L Winokur, M. A. Apicella, Analysis of nontypeable *Haemophilus influenzae* phase-variable genes during experimental human nasopharyngeal colonization. *J. Infect. Dis.* **208,** 720–7 (2013).

23. J. N. Weiser, N. Pan, Adaptation of *Haemophilus influenzae* to acquired and innate humoral immunity based on phase variation of lipopolysaccharide. *Molecular Microbiology*, **30**, 767–75 (1998).

24. S. E. Clark, K. R. Eichelberger, J. N. Weiser,Evasion of killing by human antibody and complement through multiple variations in the surface oligosaccharide of *Haemophilus influenzae*. *Mol. Microbiol.* **88,** 603–18 (2013).

25. M. Brockhurst, A. Buckling, P. B. Rainey. Spatial heterogeneity and the stability of host-parasite coexistence. *J. Evol. Biol.* **19,** 374–9 (2006).

26. M. Christen, C. Beusch, Y. Bösch, D. Cerletti, C. E. Flores-Tinoco, L. Del Medico, F. Tschan, B. Christen, Quantitative Selection Analysis of Bacteriophage φCbK Susceptibility in *Caulobacter crescentus*. *J. Mol. Biol.* **428,** 419–430 (2016).

27. M. E. Palmer, M. Lipsitch, E. R. Moxon, C. D. Bayliss, Broad conditions favor the evolution of phase-variable loci. *mBio*, **4**, e00430-12 (2013).

28. M. Brisson, É. Bénard, M. Drolet, J. A. Bogaards, I. Baussano, S. Vänskä, M. Jit, M. C. Boily, M. A. Smith, J. Berkhof, K. Canfell, H. W. Chesson, E. A. Burger, Y. H. Choi, B. F. De Blasio, S. J. De Vlas, G. Guzzetta, J. A. C. Hontelez, J. Horn, M. R. Jepsen, J. J. Kim, F. Lazzarato, S. M. Matthijsse, R. Mikolajczyk, A. Pavelyev, M. Pillsbury, L. A. Shafer, S. P. Tully, H. C. Turner, C. Usher, C. Walsh, Population-level impact, herd immunity, and elimination after human papillomavirus vaccination: a systematic review and meta-analysis of predictions from transmission-dynamic models. *Lancet Public Heal.* **1,** e8–e17 (2016).

29. A. Mazzocco, T. E. Waddell, E. Lingohr, R. P. Johnson, Enumeration of bacteriophages using the small drop plaque assay system. in Methods in molecular biology (Clifton, N.J.), vol. 501, pp. 81–85 (2009).

30. A. M. Kropinski, in Methods in molecular biology (Clifton, N.J.), vol. 501, pp. 151–155 (2009).

31. B. J. Cairns, A. R. Timms, V. A. A. Jansen, I. F. Connerton, R. J. H. Payne, Quantitative models of in vitro bacteriophage-host dynamics and their application to phage therapy. *PLoS Pathog.* **5**, e1000253 (2009).

32. K. Krysiak-Baltyn, G. J. O. Martin, A. D. Stickland, P. J. Scales, S. L. Gras, Computational models of populations of bacteria and lytic phage. *Crit. Rev. Microbiol.* **42**, 942–68 (2016)

Acknowledgments:

C.J.R.T. was supported by the Biotechonology and Biological Sciences Research Council (BB/J014532/1) and the College of Life Sciences of the University of Leicester

Author contributions:

C.J.R.T. obtained the experimental data; A.M. constructed the mathematical models; C.J.R.T., A.M., M.R.J.C, and C.D.B. designed this study, analysed data and wrote the paper. All authors discussed the results and commented on the manuscript.

Fig. 1.

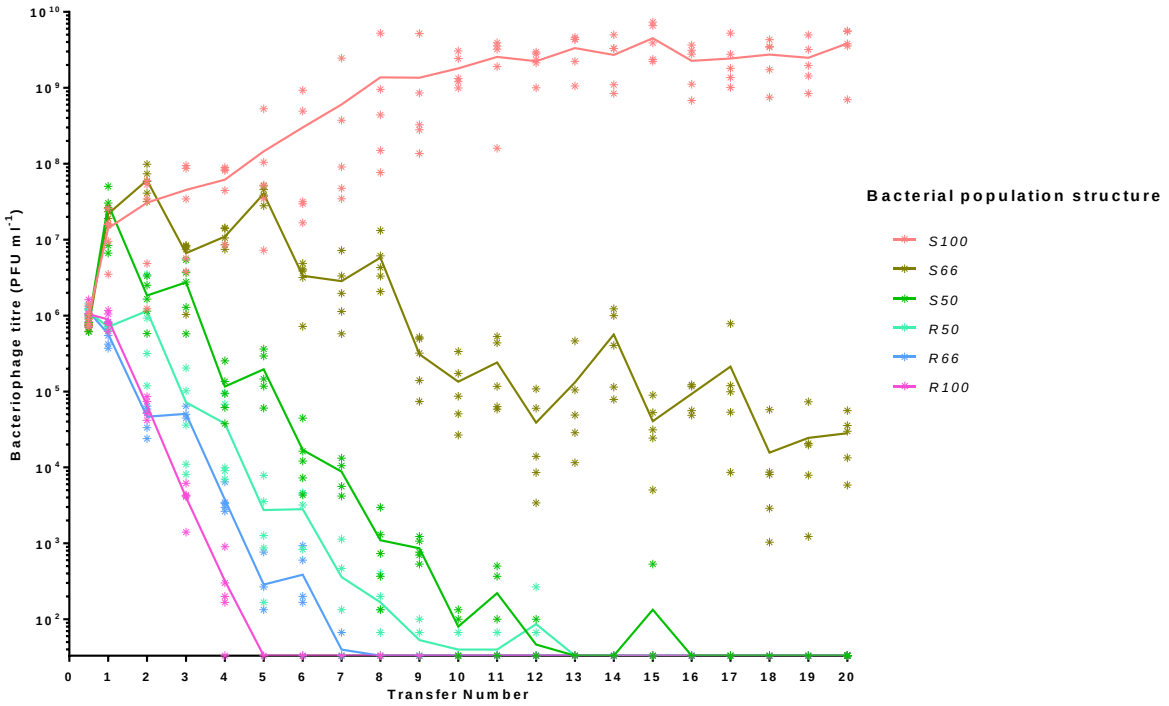

Fig. 1. Oscillating prey assay for phage HP1c1 infections of *H. influenzae* strain Rd populations with varying structures for *lic2A* expression. Mutations in a repetitive tract cause *lic2A* to switch between a phage sensitive (S) ON and a phage resistance (R) OFF state. Each line represents a defined series of cultures of two *H. influenzae* strain Rd phase variants namely Rd30S (S; *lic2A* ON) and Rd 30R (R; *lic2* OFF) as indicated in the legend (e.g. S100, all S; S66, S-S-R-S-S-R; S50, S-R-S-R; etc.). Each time point is the phage HP1c1 concentration produced following one phage replicative cycle (50 minutes). Points represent the phage titre observed from each of the five biological replicates, with the line showing the correspoinding mean phage titre of these replicates. After each time point the phage/host mix was filtered through a 0.22µm syringe filter to separate phage from host, and the resulting HP1c1 lysate was transferred to an uninfected culture of the selected *H. influenzae* phase variant for that simulated population series.

Fig. 2.

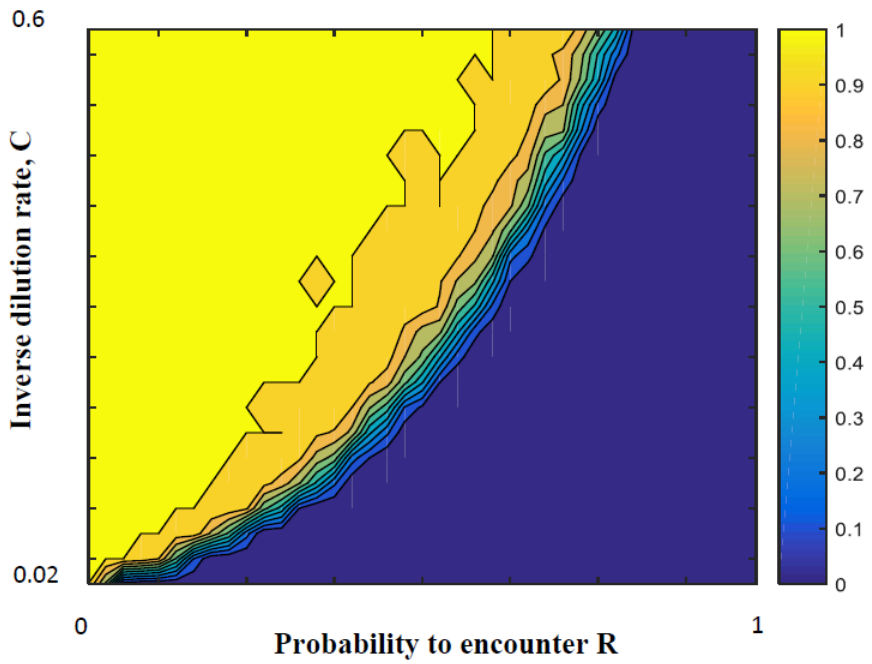

Fig. 2. Mathematical model of the impact of sub-population heterogeneity for phase-variable expression of a phage receptor and dilution rate on bacteriophage extinction events. This model simulates transmission of bacteriophage HP1c1 through 105 sub-populations of *H. influenzae* strain Rd with the phase-variable phage receptor in either a resistant (R) or sensitive state (see Fig. 1). Each series of sub-populations varied in the probability to encounter a resistant population during transfer of phage between sub-populations. Loss of phage during transfer between sub-populations was also variable and was simulated by having variable dilution coefficients (C). Multiple iterations (n = 200) were performed for each combination of R and C. The graph shows the mean proportion of lineages where bacteriophage extinction events occurred. Coloured areas represent the proportions of populations where bacteriophage were observed at the end of the experiment. Colour correlations are shown on the bar to the right of the graph with 1 representing no extinction events while 0 is extinction in all lineages. Proportions were calculated based on the number of lineages where bacteriophage densities had not fallen below the extinction threshold of 50 pfu/ml after 200 iterations.

Fig. 3.

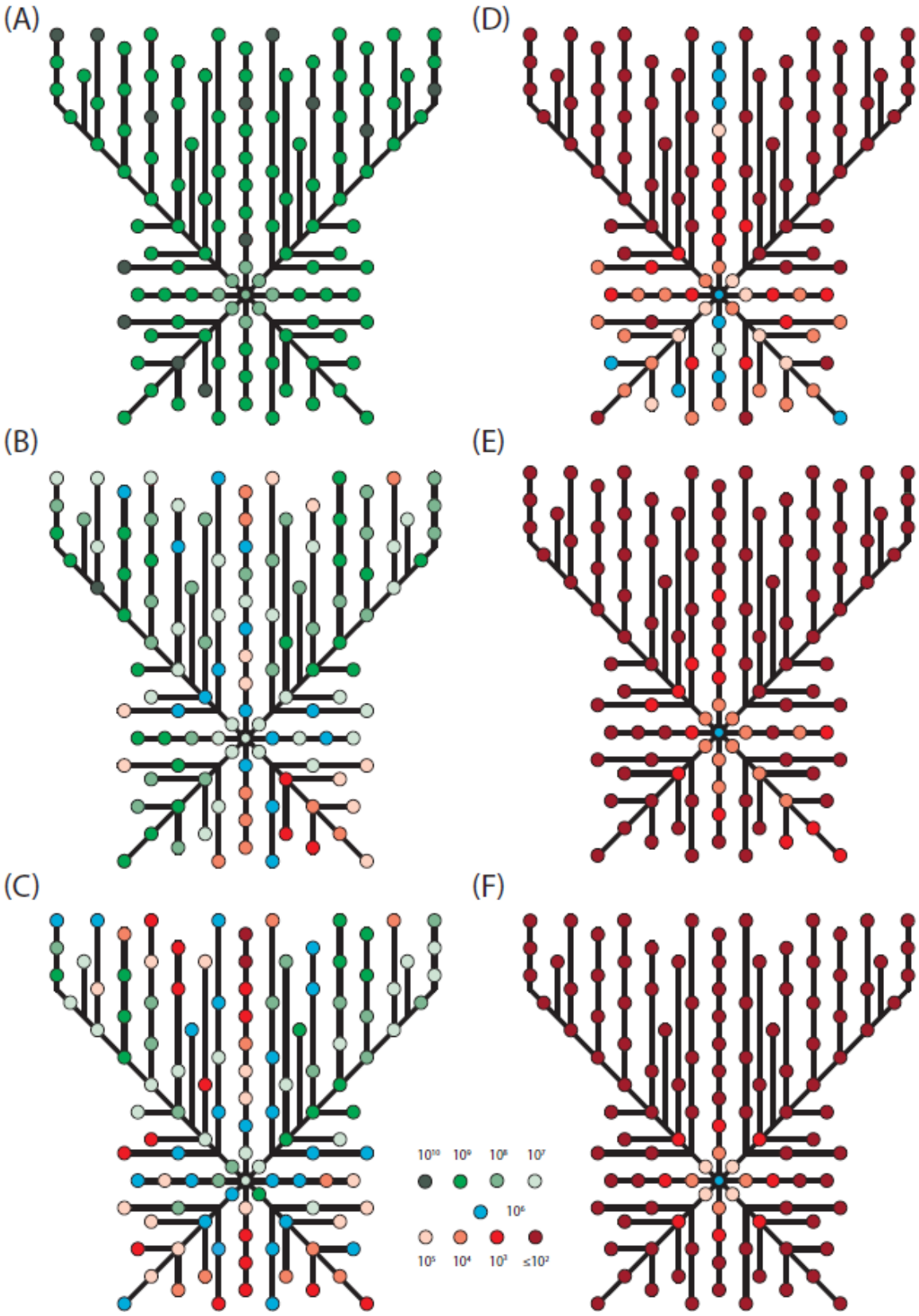

Fig. 3. Examination of bacteriophage survival in spatially-structured sub-populations of *lic2A* phase variants of *H. influenzae*. The fixed area oscillating prey assay was developed to examine spread of bacteriophage HP1c1 through *H. influenzae* sub-populations with differing spatial structures for expression of the phage receptor due to ON/OFF switching of the *lic2A* gene. The proportion of phage sensitive sub-populations (*lic2A* ON) in each fixed area structure are as follows:- (A) 100 % *lic2A* ON, (B) 66 % *lic2A* ON, (C) 50 % *lic2A* ON, (D) 50 % *lic2A* ON, (E) 33 % *lic2A* ON, and (F) none (i.e. all *lic2A* OFF, phage resistant). Each node represents a well in which bacteriophage density was measured. The colour of each node indicates the concentration of bacteriophage detected in a specific well. Lines indicate the route taken from the initially inoculated location to the observed position, with line length being proportional to the number of cycles between each node. The central well indicates the bacteriophage titre obtained at the end of the first infective cycle (all populations were inoculated with 10^6 PFU ml^{-1}).