1 Model Description

1.1 The flask environment

Each flask contains a liquid matrix, the composition of which determines the environment of a microbial population suspended within it. Some of the chemicals present are 'nutrients' (in fact, different compounds of a particular nutrient element, e.g., nitrogen) that may be consumed as food by the microbial population and converted to biomass, while others are non-consumable and form part of the abiotic environment. It is assumed that the liquid medium has properties such as temperature, pH, salinity, etc., and that these both affect, and can be affected by, microbial activity. We refer to these non-consumable chemicals and physical properties of the flask environment collectively as 'abiotic factors'. (While chemical nutrients are also abiotic we feel that their role as the substrates of metabolism justifies this notational convenience.)

There is a flow of liquid medium through the flask which occurs continuously at a fixed rate. The inflow brings with it influxes of nutrients at fixed concentrations and steady inputs to abiotic factors, while the outflow removes a fixed proportion of stored nutrients and abiotic factors. The liquid medium in each flask is assumed to be well-mixed, so that in the absence of perturbation the composition of the medium in each flask will reach a homogeneous steady state. The actual composition of the abiotic environment results from the interactions of the input and output fluxes with the collective actions of the microbial population.

The state of the flask environment is given by a vector V:

$$V = (n_1, ..., n_N, a_1, ..., a_A) = (v_1, ..., v_{N+A})$$

where n_i is the concentration of nutrient i, a_i is the level of abiotic factor i, or equivalently, v_i is the level of the i^{th} environmental state variable. N is the number of chemical nutrients, and A is the number of abiotic factors. The change in V over time is given by Equation 1, which relates the rate of change of each v_i over time to the rates of influx and outflux of that variable and the net effect of microbial activity. I_i is the rate of influx of v_i, O_i is the rate of outflux, and E_i is the effect on v_i of current microbial activity. The form of Equation 1 is general to nutrients and abiotic factors, although E_i is calculated differently for nutrients and abiotic factors (see Sections 1.6 and 1.9). The term Ψ_i represents the flux of material due to diffusive transfer between flasks (see Equation 14 in Section 1.10).

$$\frac{dv_i}{dt} = I_i - O_i v_i + E_i + \Psi_i \tag{1}$$

1.2 Microbes

The model microbes consume and excrete nutrients in fixed proportions and affect the levels of abiotic factors in their environment as a side-effect of metabolism. The precise ratios in which nutrients are consumed and excreted, and the nature of the by-product effect on abiotic factors are genetically encoded for each individual, as are its preferred abiotic conditions (i.e., the state of the abiotic environment in which its growth rate will be maximised). Microbes grow by converting consumed nutrients to biomass and reproduce by splitting when their biomass reaches a fixed threshold. Biomass is reduced at a fixed rate to represent the inevitable thermodynamic inefficiency of metabolism and the cost of maintaining cellular machinery. Microbes die if their biomass drops below a fixed threshold, which can happen in sustained periods of nutrient limitation.

Each model microbe is represented by a vector M:

$$M = (B, \lambda, \mu, \alpha, \beta)$$

where B is the current biomass of the microbe, $\lambda = (\lambda_1, ..., \lambda_N)$ represents the ratio in which nutrients are consumed, $\mu = (\mu_1, ..., \mu_N)$ the ratio in which excreta is returned to the environment as nutrients,

$\alpha = (\alpha_1, ...\alpha_A)$ the microbe's effect on abiotic factors as a side-effect of metabolism, and $\beta = (\beta_1, ...\beta_A)$ the environmental levels of abiotic factors at which growth rate is maximised. Clearly $\sum_{i=1}^{N} \lambda_i = 1$ and $\sum_{i=1}^{N} \mu_i = 1$ hold since all materials consumed and excreted must be accounted for; there is no such constraint on α since the effect of the microbe on the abiotic environmental factors does not necessarily involve mass transfer and is thus treated generally. Of the quantities in the microbe state vector M, only B is a variable during the lifetime of an individual, since λ, μ, α and β are genetically encoded and thus fixed.

Most runs reported in this paper enforced a universal preference for the abiotic environment across all microbes. This was done by setting the β vector to some fixed value a_{opt} and not allowing it to mutate, thus overriding the standard genetic scheme.

1.3 Genotype

The genotype for a microbe is an array with $2N + 2A$ loci taking real values in the range $[-1.00, 1.00]$. The genotype is subdivided into two sets of N loci for consumption and excretion and two sets of A loci for influence on abiotic factors and preferred environmental conditions. The microbe phenotype is formed by mapping and transforming the values in its genotype according to fixed rules.

The consumption ratio λ (specifying the fixed microbe-specific proportions in which nutrients are consumed) and excretion ratio μ (specifying the fixed microbe-specific proportions in which excreta is returned to the environment as nutrients) are found by subtracting the absolute value of each excretion allele from the absolute value of the associated consumption allele, to give a vector of N values for net consumption. For each nutrient, positive net consumption values are then summed and normalised to give the nutrient consumption ratio λ. Negative net consumption values are taken to mean that a nutrient is excreted, and these values are summed and normalised to give the nutrient excretion vector μ. For example, if $N = 3$, the consumption loci of the genotype are $(-0.4, 0.7, 0.1)$, and the excretion loci are $(0.5, -0.2, 0.9)$, this would give a net consumption vector of $(-0.1, 0.5, -0.8)$, and thus phenotypic trait values of $\lambda = (0, 1, 0)$ and $\mu = (\frac{1}{9}, 0, \frac{8}{9})$. Microbes are constrained so that they may not consume their own waste, which is implemented by checking each phenotype for this occurrence; such phenotypes are declared inviable and have their consumption demands set to zero (e.g., $\lambda = (0, 0, 0)$), preventing growth or reproduction and ensuring their rapid removal from the population.

The vector α determining a microbe's effect on the abiotic factors in the environment is found by directly mapping the A alleles from the relevant part of the genotype to the values for the phenotypic trait without scaling or transformation, i.e., values remain in the range $[-1.00, 1.00]$. These values give the alteration caused in the level of each abiotic factor by the creation of a single unit of biomass during microbe metabolism (see Section 1.9 for more details). A microbe's preferred abiotic environment β is determined by linear mapping of the allele values from the range $[-1.00, 1.00]$ to the range $[\beta_{min}, \beta_{max}]$. The microbe's growth rate is maximised when the state of the abiotic environment matches this preference (see Section 1.8). Runs where abiotic preference is universally shared are achieved by setting $\beta_{min} = \beta_{max} = a^{opt}$.

The genotype-phenotype encoding determines the size of the phenotypic space for each microbe trait, and hence the number of different varieties of microbe that are possible in the model. For nutrient consumption and excretion traits, the value for each trait is a vector of N values in the range $[0, 1]$; even though each vector for a particular individual must sum to 1 and incorporate the constraint of not consuming its own waste, any value in the range is possible in the population. For abiotic preferences, trait values are vectors of A values in the range $[\beta_{min}, \beta_{max}]$. For abiotic effects, trait values are vectors of A values in the range $[-1, 1]$. The range of possible phenotypic variation in the model depends on the size of N and A; larger values give proportionately more variety.

1.4 Reproduction and mutation

If the genetic specification of the microbe causes it to be successful in its environment (i.e., if its nutrient demands and preferred abiotic conditions are suited to the current state of the liquid medium held in the flask), the microbe will consume nutrients and grow by increasing its biomass. If a microbe's biomass reaches the reproduction threshold T_R, it reproduces asexually by splitting. The parent microbe donates half of its biomass to the offspring microbe, which also receives a copy of the parent's genotype. Mutation of the offspring genotype may occur during reproduction, implemented as a potential for copying error at each locus of the genotype which causes the allele value for that locus to be randomly reassigned from the range $[-1.00, 1.00]$. Mutations occur at each locus with low probability P_{mut}. No mechanism for genetic recombination is implemented.

1.5 Maintenance cost and death

A fixed rate of biomass decrement is incorporated in the model as a proxy for the combined energy costs of maintaining cellular machinery and metabolic inefficiency. This 'cost of living' reduces biomass at a fixed rate γ per simulation timestep, with the decrement assumed to be lost from the flask environment as unrecoverable heat radiation. The inclusion of this cost ensures that nutrients cannot be infinitely recycled and thus preserves some thermodynamic integrity in the model; without this cost, system biomass could theoretically grow indefinitely. The loss of some biomass from the system via this 'cost of living' ultimately sets the carrying capacity of the system, which is reached when total heat dissipation from the system is equal to the energy supplied in the form of nutrient molecules [1].

If the biomass of a microbe falls below a threshold T_D the microbe is assumed to die from starvation. In addition to this, each living microbe may die 'from natural causes' with a low probability P_D at each timestep. This mechanism is intended to be a catch-all for death by predation, apoptosis, organismic senescence, etc., and serves to thin out the microbial population in an unbiased way and thus promote continuing selection and competition between microbes. Note that the value of P_D is related to the washout rate of living microbes in chemostat models. When a microbe dies it is assumed that its remaining biomass is lost from the system, as if the dead microbe were washed out of the flask.

1.6 Nutrient consumption/excretion

At each timestep of the simulation, each living microbe j will attempt to consume a total of C_j^{max} units of nutrient, with the contributions to this total made up of each different nutrient type fixed in the relative proportions defined by the microbe's genetically specified consumption ratio. No microbe may consume a nutrient that it also excretes, with this constraint enforced as described in Section 1.3 above. The size of C_j^{max} is limited by a global maximum level C^{max} and is calculated on an individual basis for each microbe j. This calculation takes into account the match between the current state of the abiotic environment and the genetically specified preferences of microbe j and will be covered in Section 1.8 below. The actual amount consumed C_j^{act} is less than or equal to C_j^{max} and depends on nutrient availability.

In order to ensure that the microbe population doesn't consume more nutrients than currently exist in the flask environment, individual demands may need to be scaled, and to ensure that no artefacts are introduced into the model by this scaling, it must not favour any particular individual. At each timestep the total nutrient demand for the entire microbe population is calculated and compared to the amounts of nutrients available. It is assumed that all microbes are continuously and simultaneously feeding, so in the case that there is an insufficient amount of a nutrient available to meet the entire population demand, the demand of every individual microbe that requires that nutrient is scaled down equally, so that the total amount consumed by the population matches what is available.

Mathematically, we have constraints $C_j^{max} \leq C^{max}$ and $C_j^{act} \leq C_j^{max}$, and then:

$$C_j^{act} = C_j^{max} \prod_{i=1}^{N} w_{ij} \tag{2}$$

where C_j^{act} is the actual total quantity of nutrients consumed by microbe j after scaling for nutrient limitation has been applied and w_{ij} is the scaling factor for nutrient i for microbe j. Values for w_{ij} are calculated sequentially by noting that the population demand D_i for nutrient i (after all individual demands for nutrient $i-1$ have been scaled appropriately) and w_{ij} are related. Recalling that λ_{ij} is the proportion of consumption taken as nutrient i by microbe j and n_i is the total amount of nutrient i currently available in the flask environment, we can derive the full set of all w_{ij} and D_i for every nutrient i and living microbe j by solving iteratively for each value of D and w, starting with an assumed value of $w_{0j} = 1$ (valid since nutrient 0 does not exist).

$$\begin{aligned} w_{0j} &= 1 && \forall j \\ w_{ij} &= \begin{cases} min(1, \frac{n_i}{D_i}) & \lambda_{ij} > 0 \\ 1 & \lambda_{ij} = 0 \end{cases} \end{aligned}$$

$$D_i = \sum_{j}^{living} \left(\lambda_{ij} C_j^{max} \prod_{k=0}^{i-1} w_{kj} \right)$$

Then having established the value of C_j^{act} we can go on to derive:

$$C_j = C_j^{act}(\lambda_{1j}, ..., \lambda_{Nj}) \tag{3}$$

$$C_i^{pop} = \sum_{j}^{living} \lambda_{ij} C_j^{act} \tag{4}$$

where C_j is the actual consumption vector for microbe j and details how much of each nutrient is consumed by microbe j at a particular timestep. C_i^{pop} is the total amount of nutrient i consumed by all living microbes.

This scheme means that each microbe always consumes nutrients in the relative proportions specified by its genetically determined consumption ratio. If nutrient limitation means that the amount of a particular nutrient consumed by a microbe is scaled down, the amounts of the other nutrients it consumes are also scaled down by an equivalent factor to maintain the fixed relative proportions of consumption.

Consumed nutrients are converted to biomass with a standard efficiency of θ, with the waste being excreted as nutrients (i.e., $C_j^{act} = 10$ units of food consumed with efficiency of $\theta = 0.6$ makes 6 units of biomass and 4 units of excreta). Excreta is returned to the environment as nutrients in the fixed proportions specified by the microbe's genetically determined excretion ratio. We can thus define the excretion vector X_j for every microbe j, and an expression for the total amount X_i^{pop} of nutrient i collectively excreted and returned to the environment by the population:

$$X_j = (1-\theta) C_j^{act}(\mu_{1j}, ..., \mu_{Nj}) \tag{5}$$

$$X_i^{pop} = \sum_{j}^{living} (1-\theta) C_j^{act} \mu_{ij} \tag{6}$$

1.7 Growth

At each timestep of the simulation a microbe j will consume C_j^{act} units of nutrient, which are converted to biomass with a fixed efficiency of θ. Taking into account the previously defined maintenance cost γ, we can now state the growth rate (rate of change of biomass) of a microbe j as:

$$\frac{dB_j}{dt} = \theta C_j^{act} - \gamma \tag{7}$$

Note that Equation 7 specifies the growth of an individual. Growth of a population occurs only as a result of individual growth and reproduction, and is not specified *a priori* as in more traditional population ecology models such as Lotka-Volterra systems.

1.8 Effect of abiotic factors on metabolic rate

The model is designed so that the state of the abiotic environment affects the growth rate of microbes, and this is implemented as a feedback from the environmental state variables for abiotic factors onto the consumption demands of all microbes. As mentioned above, each microbe will attempt to consume a maximum amount C_j^{max} of nutrients at each timestep, with this demand being met depending on nutrient availability. The attempted consumption amount C_j^{max} is calculated for each microbe j as a function of the match between the current state of the abiotic environment and the microbe's genetically specified preferences. This function has Gaussian form and falls away smoothly from its maximum value as the distance between the current environment and the optimum increases. Mathematically, we can capture this as below:

$$C_j^{max} = \psi_j C^{max} \tag{8}$$

$$\psi_j = e^{-(\tau \rho_j)^2} \tag{9}$$

$$\rho_j = \sqrt{\sum_{i=1}^{A} (a_i - \beta_{ij})^2} \tag{10}$$

where C^{max} is a universal constant defining the maximum rate of consumption for any microbe, ψ_j is a microbe-specific measure of the microbe's satisfaction with the current abiotic environment given its preferences, τ is a universal constant parameter that sets the level of influence of the abiotic environment on growth rate (high τ means a stronger influence, $\tau = 0$ means no influence), ρ_j is a measure of the distance between the current environmental level of each abiotic factor a_i and the microbe's preferred level β_{ij}.

1.9 Effect of microbial activity on environment

Microbes can affect their abiotic environment as a side effect of their metabolism. We make the size of a microbe's effect on the environment proportional to its growth rate, so that fast-growing and more fecund microbial strains have a greater influence than dormant or slow-growing strains.

Equation 1 relates the rate of change of each environmental variable v_i to the combined effect E_i of the microbial population on that variable. The calculation of E_i is treated differently for nutrients and abiotic factors. First of all, let the vector E be the vector of effects of the population on all environmental variables, and note that this vector can be sub-divided into nutrient and abiotic components.

$$E = (E_n, E_a) = (E_{n1}, ..., E_{nN}, E_{a1}, ..., E_{aA})$$

We can use our previous definitions of nutrient consumption/excretion to work out the population effect on nutrient state variables. This gives us an expression for the effect E_{ni} of the population on nutrient i:

$$E_{ni} = -C_i^{pop} + X_i^{pop} \tag{11}$$

The effect of each microbe on each abiotic factor is determined by its genetically specified effect vector α and is applied for every unit of biomass created. The expressions for microbe growth combined with this vector can be used to work out the population effect E_{ai} on each abiotic state variable i:

$$E_{ai} = \sum_{j}^{living} \frac{dB_j}{dt} \alpha_{ij} \tag{12}$$

1.10 Spatial structure

Spatial structure is added by connecting F flasks in a ring topology to create a simple one-dimensional spatial environment. Connectivity is by adjacency only and no long-distance links are modelled. Thus each flask f_i is adjacent to flasks f_{i-1} and f_{i+1}. Flasks f_F and f_1 are connected to complete the ring.

The connection between neighbouring flasks is an approximation of diffusive mixing at a steady rate, implemented as the exchange of a fixed amount of liquid between neighbouring flasks at each timestep. Since each flask is well-mixed it is assumed to have an even distribution of nutrients, abiotic factors, and microbes. Thus if the total volume of a flask is V_T, then diffusion acting at a rate such that a volume V_D of liquid is exported from each flask at each timestep means that a fraction $\frac{V_D}{V_T}$ of all materials and microbes will leave each flask. No directionality is modelled in the flow of liquid, so this amount is divided equally between each of the flask's neighbours. The donor flask receives from its two neighbours a total amount of liquid equal to the amount it exports, so the amount of liquid in each flask remains constant. All liquid transferred carries with it material and microbes at the same concentrations as the donor flask. If environmental gradients exist between neighbours, these will be reduced over time by the diffusive transfer.

We begin to formalise the diffusive transfer scheme in Equation 13:

$$V_D = R_D V_T \tag{13}$$

where R_D is the rate of diffusion. Material exchange Ψ is by removal/adding of the correct proportion of material at each timestep:

$$\Psi_{i,j} = \Delta v_{i,j} = -R_D v_{i,j} + \frac{R_D}{2} v_{i,j-1} + \frac{R_D}{2} v_{i,j+1} \tag{14}$$

where $v_{i,j}$ is the level of the i^{th} environmental variable (nutrient or abiotic factor) in flask j. The transfer of microbes is handled stochastically. The probability of a microbe being located in the volume of liquid transferred from a donor flask is $R_D (= \frac{V_D}{V_T})$. Thus if M_j denotes the population size at flask j, the change in population size at each timestep is in the limit given by:

$$\Delta M_j = -R_D M_j + \frac{R_D}{2} M_{j-1} + \frac{R_D}{2} M_{j+1} \tag{15}$$

Microbes that are transferred are moved intact from the donor flask to the recipient flask. The recipient flask for each microbe is chosen from the donor's two neighbour flasks with equal probability. Microbe biomass and genotype do not affect the transfer process, so the export group represents a fair sample of the donor population.

The special case of perfect mixing is handled differently. Rather than use the above scheme with $R_D = 1$, which would be problematic, perfect mixing is implemented by re-distributing all microbes and materials evenly across the multi-flask system at every timestep. Perfect mixing thus instantaneously

Parameter	Value	Description
N	2	Number of nutrients
A	1	Number of abiotic factors
T_R	120	Reproduction threshold (biomass units)
T_D	50	Starvation threshold (biomass units)
P_{mut}	0.01	Probability of mutation at each genotype locus during reproduction
P_D	0.002	Probability of death by natural causes at each timestep
γ	1	Maintenance cost (biomass units/timestep)
θ	0.6	Nutrient conversion efficiency
C^{max}	10	Maximum nutrient consumption rate (units/timestep)
τ	0.02	Level of influence of abiotic environment on metabolism
I_N^{min}	0	Minimum rate of nutrient influx (units/timestep)
I_N^{max}	150	Maximum rate of nutrient influx (units/timestep)
O_N^{min}	0.01	Minimum rate of nutrient outflux (percentage/timestep)
O_N^{max}	0.25	Maximum rate of nutrient outflux (percentage/timestep)
I_A^{min}	10	Minimum rate of abiotic factor influx (units/timestep)
I_A^{max}	20	Maximum rate of abiotic factor influx (units/timestep)
O_A^{min}	0.1	Minimum rate of abiotic factor outflux (percentage/timestep)
O_A^{max}	0.25	Maximum rate of abiotic factor outflux (percentage/timestep)
K_m	100	Number of individuals in flask innoculum
T_{prep}	500	Flask equilibriation time prior to seeding (timesteps)
T_{run}	30000	Duration of run (timesteps)
F	10	Number of flasks used in spatial ring topology
R_D	0.01% - 100%	Rate of diffusion between flasks (percentage of volume per timestep)
β_{min}	a^{opt}	Lower bound for abiotic environmental preference
β_{max}	a^{opt}	Upper bound for abiotic environmental preference
a^{opt}	150	Target level for abiotic environment (universally applied)
V_T	1	Nominal volume of each flask

Table 1: Parameter values used in the simulation. Sensitivity analysis on these parameters shows no significant changes to qualitative behaviour of the model are caused by varying these values, except where described in the text.

removes any gradients in environmental variables or microbe distributions, so that the system is always homogeneous.

1.11 Numerical method

At the start of each simulated timestep, nutrient/abiotic influx is added to the environment. This is followed by simultaneous update of all microbes in the population for metabolism, death (by starvation or random selection), and reproduction, in that order. After the biota are updated, nutrient/abiotic outflux is removed from the environment. At this point diffusive transfer of material and microbes is performed and the system update is complete for that timestep.

The update equation for the flask environment (Equation 1) and the microbial growth equation (Equation 7) are continuous differential equations. At each timestep these differential equations are discretised by calculating their instantaneous value and adding it to the existing values of the quantities concerned. For environmental update we have $\Delta v = \frac{dv}{dt}$ and then $v_t = v_{t-1} + \Delta v$. For microbe growth we have $\Delta B = \frac{dB}{dt}$ and $B_t = B_{t-1} + \Delta B$. This is equivalent to numerical integration using Euler's forward method with an integration timestep equal to one simulated timestep in the flask ecosystem.

References

[1] Williams, H. T. P & Lenton, T. M. (2007) The Flask model: Emergence of nutrient-recycling microbial ecosystems and their disruption by environment-altering 'rebel' organisms. *Oikos* **116**, 1087–1105.