Supplementary Information for
Genetic architecture of dispersal and local adaptation drives accelerating range expansions

Jhelam N. Deshpande and Emanuel A. Fronhofer

Jhelam N. Deshpande.
E-mail: jhelam-nitin.deshpande@umontpeller.fr

This PDF file includes:
Supplementary text
Figs. S1 to S16
Table S1
SI References
We develop a discrete-time and discrete space individual-based metapopulation model of range expansion into an external environmental gradient. We model a sexually reproducing organism with non-overlapping generations. We assume that dispersal and local adaptation traits can evolve and are encoded by two non-interacting gene-regulatory networks (GRNs). We model a $500 \times 5$ landscape which means that there are $500$ patches in the $x$-direction and $5$ patches in the $y$-direction. The central $10 \times 5$ patches have spatio-temporally constant environment and represent the native habitat of the organism being modeled. On either side of the native habitat, there is a symmetric external environmental gradient only in the $x$-direction. Boundary conditions in the $y$-direction are periodic, thus, the landscape can be thought of as a hollow tube. We assume that individuals are initially confined to their native habitat for a duration of $20000$ generations, after which range expansions can begin into the remaining patches. Such a burn-in period ensures that the trait values before range expansions begin are constrained by the evolutionary history of the organism.

**Life cycle.** After birth, individuals first have the possibility to disperse, following which they mate and reproduce. At each generation, entire patches can go extinct with a probability $\epsilon$, representing temporally and spatially random catastrophic extinction events. Every patch in the native habitat (central $10 \times 5$ patches) is initialized with $N_0 = 100$ individuals which have equal chance of being male or female.

**Dispersal.** Dispersal is natal, and individuals can disperse to one of the eight surrounding patches (nearest-neighbour eight dispersal) with a probability $d$, determined by its genetic architecture (described below). Dispersing individuals may die with a probability $\mu$ while dispersing, which captures the multiple costs of dispersal (1).

**Reproduction and inheritance.** After dispersal, individuals mate in their destination patches. Each female mates with a male drawn randomly with replacement from the local population. The population dynamics follow logistic growth according to the Beverton-Holt model (2) with local density regulation:

$$N_{x,y,t+1} = N_{x,y,t} \frac{\lambda_0}{1 + \alpha N_{x,y,t}}$$

where $\lambda_0$ is the intrinsic growth rate, and $\alpha$ is the intraspecific competition coefficient. $K = \frac{\lambda_0 - 1}{\alpha}$ is the expected equilibrium density in the Beverton-Holt model. If it is completely adapted to local environmental conditions, a female produces a number of offspring drawn from a Poisson distribution with a mean given by $\lambda = \frac{2\lambda_0}{1 + \alpha N_{x,y,t}}$ which models demographic stochasticity. Otherwise, there is a reduction in its mean fecundity as described in the following section. The factor of two accounts for the fact that only half of the population (females) are capable of bearing offspring and makes $\lambda$ interpretable at the population level. The offspring have an equal probability of being male or female, therefore the expected sex ratio is $0.5$. After reproduction, the adult individuals die and are replaced in the metapopulation by their offspring. The modifications to this model to understand local adaptation will be explained in the following section.

Parameters (weights, thresholds and slopes that characterize a GRN; see Table S1 and the following section for a detailed explanation) of a dispersal or local adaptation genetic architecture are represented by freely recombining diploid loci and are realized by their mid-allelic value, that is, we do not assume dominance. Offspring inherit one allele from the mother and one from the father, randomly. The alleles may be modified with a mutation rate $m(t)$ by adding numbers drawn from a normal distribution with mean zero and variance $\sigma_m$. The mutation rate varies with time to generate variation at the beginning of the simulation while allowing the genotypes to stabilize towards the end. Hence, $m(t)$ decreases linearly till the $5000$th generation and then remains constant (see 3). We chose this approach to allow the genetic algorithm to explore the entire fitness landscape more easily without potentially getting stuck in local optima. Note that mutation rates only vary at the very beginning of the burn-in phase and are constant thereafter. This includes the range expansion phase, implying that any pattern observed during range expansion cannot be due to this mutation procedure choice.

$$m(t) = \begin{cases} m_{\text{max}} - \frac{(m_{\text{max}} - m_{\text{min}}) t}{5000} & t < 5000 \\ m_{\text{min}} & t \geq 5000 \end{cases}$$

We use these varying mutation rates only while optimizing GRN early in the burn-in period.

**Modeling local adaptation.**

**Landscape.** The equation describing the environmental gradient ($\tau_{\text{env}}(x)$) as a function of patch identity $x$ is specified by two parameters: $\tau_0$ is the constant environment in the burn-in region (between $x_{\text{left}}$ and $x_{\text{right}}$) and $b$ is the slope of the gradient.

$$\tau_{\text{env}}(x) = \begin{cases} b(x_{\text{left}} - x) + \tau_0 & x < x_{\text{left}} \\ \tau_0 & x_{\text{left}} \leq x \leq x_{\text{right}} \\ b(x - x_{\text{right}}) + \tau_0 & x > x_{\text{right}} \end{cases}$$

There is no gradient in $y$-direction.
Local adaptation. The effect of the environment on the fecundity of a female (or the survivorship of the offspring) is given by a Gaussian niche function. The local adaptation phenotype ($\tau$), defined as the external environmental value that maximizes the fecundity of a female, is genetically encoded (either by a GRN or a single quantitative locus, depending on the assumed genetic architecture), and the niche width ($\omega$) is assumed to be fixed. While the male mate does not directly influence $\tau$, $\tau$ is determined by both the female’s mother and father. Therefore, overall there is selection on male and female genotypes. If $1 - s$ is the reduction in fecundity of a female, then:

$$s = \exp \left(-\frac{(\tau_{env}(x) - \tau)^2}{\omega} \right) \quad [S4]$$

Therefore, the number of offspring produced by a female (based on Eq. S1) will now be drawn from a Poisson distribution with mean $\lambda = s \frac{2N_0}{1+nN_0}\frac{1}{\omega}$. Hence, a greater $s$ implies a higher level of adaptation to the external environment of a patch.

Gene-regulatory network model.

Background on genotype-to-phenotype maps and gene-regulatory networks. One way in which the genetic architecture of traits has been conceptualized is the genotype-to-phenotype map (GP map). Alberch (4) proposed that the relationship between genotype and phenotype is non-linear and is mapped by dynamic developmental processes, determined by parameters (influenced by the genotype) associated with the developmental system. Stable phenotypes then correspond to steady states of a developmental mapping. According to this conceptualization, areas in parameter space correspond to phenotypes separated from each other by a boundary. Phenotypes are more robust if they are within a larger region of parameter space and away from the boundary of this area because perturbations to parameters would be neutral. This also implies that many genotypes correspond to the same phenotype. Therefore, these maps have the property of robustness but are also evolvable because they can neutralize explore phenotypic space as genetic variation can accumulate by mutation without changing the phenotype. This metaphor of the GP map has been used to study the evolution of RNA folding, protein structure, and gene-regulatory networks, for example (5). More recently, Nichol et al. (6) provide a more general definition for the GP map. These authors define a GP map as a relation that maps a genotype (along with a mutation operation) to a phenotype modulated by an environment. They further provide examples of such maps (phenotype landscapes, fitness landscapes, RNA secondary structures, gene-regulation, and neural networks for phenotypic plasticity) and discuss their general properties such as degeneracy, robustness, evolvability, and neutrality. By extending the definition such that the phenotype is modified by the environment, phenotypic plasticity can also be placed within the framework of GP maps. Additionally, this definition does not assume the GP map properties and includes more classical approaches as well. Therefore, we henceforward will use the definition given by Nichol et al. (6) and will refer to the conceptualization by Alberch (4) particularly if required.

A very well studied GP map is the gene-regulatory network (GRN) model, particularly the one introduced by Wagner (7) to study gene duplications. The GRN represents a developmental process. The GRN GP map consists of $n$ transcription factors which can regulate each other’s gene expression states (which are the phenotype), and the interaction between the genes is represented by a regulatory matrix (the genotype) containing weights. These weights are optimized by evolutionary simulations, where fitness is a function of gene expression levels (phenotype). Note that gene expression states in this model are usually discrete, that is, genes are either on or off. This is conceptually similar to the GP map described by Alberch (4) and has analogous properties. The parameters of the dynamical developmental system are the weights, and the gene expression levels are the steady states of this system. Various applications and variants of this model are reviewed in Spirov and Holloway (8). Particularly, the robustness to mutation, also known as genetic canalization, has been studied in GRNs extensively. Wagner (9) suggests that robustness to mutation in GRNs evolves due to stabilizing selection. Siegal and Bergman (10) however, suggest that GRNs can evolve to be robust in the absence of stabilizing selection when there is selection for developmental stability. Ciliberti et al. (11) visualize different GRN genotypes as a genotype network, in which genotypes with a single mutation are connected to each other. They find that the genotype network can mostly be explored without a change in the phenotype, and particular genotype network topologies can be simultaneously robust and innovative. Rünneburger and Rouzic (12) challenge the generalizability of the evolution of robustness in GRN models when gene expression is quantitative rather than discrete. They model a GRN with sigmoid, continuous gene expression. Stabilizing selection acts on few genes directly (selected genes), and the remaining genes are under indirect selection (only their regulatory function is selected). They find that, while the genes that are not directly selected always become canalized, selected genes are only canalized when there is selection for extreme levels of gene expression. This raises the further question of whether and how Wagner’s model can be used to understand phenotypes that are quantitative rather than discrete. Furthermore, can this model be extended in a meaningful way to understand complex traits, which do not necessarily result from transcription regulation alone but from the downstream effects of regulatory genes? Is robustness to mutation still a feature of such a model?

GRN models have also been used to study the evolution of evolvability under conditions of ecological change, including local adaptation. Draghi and Wagner (13) use the GRN model to study the evolution of evolvability. They conclude that evolvability can evolve, and networks with greater positive network excitation are more evolvable but less robust under fluctuating selection. Kimbrell and Holt (14) use Wagner’s model to understand local adaptation in a source-sink context. Optimal gene expression states are different in the source and sink. They find that adaptation to the sink happens in one or more steps and results from a breakdown in canalization of the local adaptation GRN, that is, an increase in sensitivity to mutation. Kimbrell (15) extends this model and finds that sensitivity to mutation increases serially along a gradient landscape with four patches. Malcom (16) models adaptation to an environmental shift governed by a quantitative trait encoded by the sum of expression states of a
boolean GRN. They find that smaller GRNs lead to a faster rate of adaptation. In Malcom (17), they extend this model to a metacommunity of two competing species of different network sizes and three patches. A local adaptation quantitative trait is represented by the sum of outputs of a Boolean GRN, and dispersal is a fixed rate. They find that for species with similar GRN sizes, the outcome of competition is determined by dispersal, and for comparable dispersal rates, the outcome of competition is determined by which species is able to adapt faster (those with networks of a smaller size). Finally, Melián et al. (18) suggest that including GRNs to model traits relevant to interspecific interactions can help us understand the stability and complexity of spatial communities like food webs.

**Gene-regulatory network model for the evolution of dispersal and local adaptation during range expansion.** The GRN assumed here is a modified version of the model by Wagner (7). However, as opposed to Wagner (7) and similar to Draghi and Whitlock (19), we assume that relevant traits (dispersal and local adaptation) are continuous and result from the downstream effects of a GRN. The GRN model by Wagner (7) assumes that a single developmental stage in a single cell type (tissue) is specified by equilibrium expression of n-transcription factors (either on or off). These n-transcription factors are represented by a discrete dynamical system and interact with each other cooperatively (switch-like response) during development. Such interactions are represented by regulatory matrices, which are the genotype of the individual. Thus, the GRN model maps a genotype (regulatory matrix) to a phenotype (gene expression levels) via a developmental process.

Thus, we describe here a GRN model extended from the Wagner (7) model for a general continuous trait $z$. In our model, we have two traits, dispersal probability ($z = d$) and the local adaptation phenotype ($z = \tau$) that can potentially evolve. We assume that each of these traits are encoded by the products of two non-interacting gene-regulatory networks. This means that we treat the GRNs underlying the two traits as units that do not influence the gene expression states of each other.

Particularly, we introduce changes to the model by Wagner (7), which are similar to Draghi and Whitlock (19) and van Gestel and Weissing (20). Like van Gestel and Weissing (20), we visualize the model as having three layers, an input layer, a regulatory layer, and an output layer (Fig. 1). The input layer $x_z = (x_{z,1}, ..., x_{z,m})$ is a vector of upstream signals which may result from a previous developmental step if one is modeling a constitutive trait or a set of environmental cues to model phenotypic plasticity. Therefore, since we assume that dispersal and local adaptation are constitutive, $m = 1$ and we provide a fixed input $x_1 = 0.5$ to both the dispersal and local adaptation GRNs. The regulatory layer represents the network of n-transcription factors as modeled in Wagner (7). In our model, we set $n = 3$ for both the dispersal and local adaptation traits. The normalized and transformed expression states of the regulatory genes after $I$ iterations for trait $z$ are $S_z(I) = (S_{z,1}(I), ..., S_{z,n}(I))$. The genes can take values in the interval $[-1, 1]$ and gene expression of a single gene is a continuous sigmoid function (10, 19) of the input.

The network is iterated to obtain equilibrium gene expression, which can be a fixed point or a limit cycle. Limit cycles are considered to represent developmental instability (7) and individuals with such GRNs die. While oscillatory GRNs are possible in biological scenarios (especially those that involve timekeeping, for example, circadian rhythms, (21)), the traits under consideration here, that is, dispersal and local adaptation, are characterized by a single value. However, one must be aware that this choice has consequences for evolved GRN properties, for example, the robustness of the GRN to mutations. Siegal and Bergman (10) find that selection for developmental stability leads to more robust GRNs. Espinosa-Soto (22) find that stabilizing selection for particular gene expression states, including limit cycles, leads to the evolution of mutational robustness.

We assume that genes in this GRN could be stochastically on or off before the beginning of the particular developmental stage being modeled. Therefore, the initial gene expression states are drawn from a uniform distribution in the interval $[-1, 1]$. It is possible that the equilibrium gene expression could change depending on this initialization. However, such networks are not highly frequent (23), and if the magnitude of the shift is high, then such a network might be eliminated from the metapopulation because of the fitness consequence of an incorrect dispersal decision or mismatch in trait optimum.

Similar to Draghi and Whitlock (19), we do not consider the vector of gene expression states as the phenotype under selection. The phenotype is continuous and results from downstream effects of regulatory genes (for example, expression levels of structural genes or metabolic pathways). Further, we assume that the phenotype varies linearly with gene expression states. Hence, this model applies to cases where the processes downstream of regulatory gene expression are continuous and linear. However, the downstream effects of regulatory gene expression can be switch-like or non-linear as well (for example, the decision of a bacterium to sporulate in van Gestel and Weissing (20)). This will not be considered in the present study, hence we do not assume a sigmoid activation function for the output layer. Therefore, the phenotypic value of a trait ($z$) is given by linear combination of gene expression levels at fixed point steady-state gene expression $S^\ast_z$. $V_z$ is a $1 \times n$ matrix containing weights connecting the regulatory layer to the output layer.
We assume that two distinct GRNs encode dispersal ($z = d$) and local adaptation ($z = \tau$). The gene expression states are iterated for 20 iterations $I$ (19) and then the continuous phenotype is calculated.

**Alternate models of genetic architecture.** We develop additional models, otherwise identical in their set-up and life-cycle to the above-described GRN model, but with differing genetic architectures. This allows us to understand how relaxing the assumption of additivity in the GRN model as opposed to standard additive quantitative genetics models modifies range dynamics. Since mutation effects are emergent in the GRN model, choosing a comparable additive quantitative genetics control is not straightforward. Thus, we choose to compare our GRN model to a model in which one locus each encodes dispersal and local adaptation (“simple additive genetic architecture”). While such a control is not informative in comparing absolute speeds of range expansion, it allows us to attribute temporal variation in range dynamics, including accelerating range expansions in the GRN model, to non-additivity. In order to understand if the comparison between the GRN and the simple additive architectures is sensitive to the number of loci and scaling of mutation effects, we develop additional models (“multi-locus additive genetic architecture” and “n locus control genetic architecture”) with varying numbers of loci. We also develop additional models to narrow down the relative contribution of dispersal versus local adaptation to range dynamics (“mixed genetic architecture” and “fixed dispersal genetic architecture” models).

**Simple additive genetic architecture.** We first compare the above-described GRN model to a model that assumes a simple additive architecture. In this model, the dispersal and local adaptation traits are represented by one diploid locus each. We initialize all individuals in this model with standing genetic variation for dispersal and local adaptation traits by drawing their allelic values from a uniform distribution going from 0 to 1. When individuals reproduce, the offspring inherit one allele from the female parent and another from the male. During inheritance, each allele may mutate with a probability $m_{\text{mut}} = 0.0001$. Mutation effects are drawn from a Gaussian distribution with mean 0 and standard deviation 0.1. Note that the mutation rates during range expansion are the same for the simple additive and GRN models. Similar to the GRN model, we assume that the individuals are initially adapted to a native habitat with a spatio-temporally constant environment before range expansions start, therefore, we again have a burn-in period of 20000 generations, after which range expansions can begin.

**Multi-locus additive genetic architecture.** We develop a second additive model that exhibits the same number of loci as the GRN model. Given that $n$ is the number of regulatory genes and $m$ is the number of inputs each gene receives, the number of loci per trait in the GRN model can be written as the sum of the number of elements of the input matrix ($U_z$ with $mn$ elements), regulatory matrix ($W_z$ with $n^2$ elements), regulatory threshold vector ($\theta_z$, with $n$ elements), regulatory slope vector ($x_z$, with $n$ elements) and output matrix ($V_z$, with $n$ elements). Thus, the total number of loci in the GRN is given by $mn + n^2 + 3n$. Since the GRN model has $n = 3$ genes and $m = 1$ inputs to each gene for each trait, we assume 21 loci underlying each trait in our multi-locus additive model. In this model, each trait is represented as the mean value of these loci, thus mutation effects per locus are smaller in this model by a factor of 21, when compared to the simple additive model. Therefore, this model assumes that multiple loci of small effect contribute to each trait. This model allows us to control for the number of loci underlying the dispersal and local adaptation traits.

**n locus control genetic architecture.** We explore an additional $n$ locus model to understand how the scaling of mutation effects impacts range dynamics. Here, we assume the same number of loci of additive effect as there are regulatory genes in the GRN (here, $n = 3$ loci, corresponding to three regulatory genes). Dispersal and local adaptation traits are represented as the sum of these loci, thus mutation effects are comparable to the GRN model if all gene expression states are at an extreme ($-1$ or 1).

**Mixed genetic architectures.** We next seek to test the relative contribution of the genetic architecture of dispersal and local adaptation traits in modifying range expansion dynamics. We thus develop two additional models of range expansion with mixed genetic architectures: 1) GRN LA + simple additive dispersal model — in this model, local adaptation (abbreviated by LA) is encoded by a GRN and dispersal is encoded by one quantitative locus and 2) simple additive LA + GRN dispersal — here, one quantitative locus encodes local adaptation and a GRN encodes dispersal. Thus, we predict that if the range dynamics are primarily driven by the genetic architecture of local adaptation, then the qualitative behavior of the GRN model should be captured by the model in which local adaptation is encoded by a GRN as well (GRN LA + simple additive dispersal model). Conversely, if the genetic architecture of dispersal strongly impacts range dynamics, then we expect the simple additive LA + GRN dispersal model to lead similar qualitative dynamics as the GRN model.

**Fixed dispersal genetic architectures.** To understand how dispersal evolution impacts range expansions dynamics, we develop further models in which all individuals disperse with a fixed dispersal probability, that is, dispersal cannot evolve. The fixed probability is equal to the measured dispersal probability at the end of the burn-in period corresponding to each of the parameter combinations and scenarios, averaged over all 50 replicate simulations. We term these models: 1) GRN LA + fixed dispersal and 2) simple additive LA + fixed dispersal. By comparing these two models, we can determine how dispersal evolution modifies range expansion dynamics in the GRN and simple additive models relative to when dispersal does not evolve.
**Tracking range expansion dynamics.** For all of the genetic architectures described above, we track the position of the range front as a function of time every 50 time steps since the beginning of range expansion. The range front is defined as the farthest occupied patch from the range core. We do this for 50 replicate simulations, which amount to a total of 100 range expansions as the metapopulation in one replicate simulation can expand to both sides of the landscape. In order to make range expansion dynamics towards either side of the landscape comparable, we calculate the range front position relative to boundary of the range core on either side. We explore range dynamics for all the parameter combinations described in Table S1.

**Patch dynamics in the focal scenario.** In the focal scenario of the main text — GRN model and simple additive models ($b = 0.04$, $\mu = 0.1$, $\epsilon = 0$) — we track over space the level of adaptation ($s$), dispersal probability, sensitivity to mutation of the local adaptation trait (calculation described below) and mean absolute gene expression of the local adaptation trait as a function of time. We do this for 50 replicate simulations, that is, a total of 100 range expansions. The level of adaptation $s$ measures how adapted the expanding population is to a given patch according to Eq. S4 and is therefore a function of the level of mismatch between the local adaptation phenotype and the external environment. A larger $s$ implies that on an average the expanding population is better adapted to that cross section of the landscape. Since we model an external environmental gradient only in the $x$-direction, we calculate the level of adaptation averaged over a given patch cross-section. Dispersal probability is measured as the fraction of individuals leaving a patch and is also averaged over the patch cross-section. The sensitivity to mutation indicates how the phenotype changes with respect to a haploid mutation. The exact procedure for calculation is described below. Finally, we also calculate mean absolute gene expression ($\sum_{i=1}^{n}|S_{z,i}|$, where $n$ is the number of genes in the GRN), over all individuals in a given patch cross-section. If mean absolute gene expression is high, this indicates that gene expression is on an average extreme (closer to $-1$ or $1$).

**Calculation of the sensitivity to mutation in the GRN model.** In the gene-regulatory network model we assume that the loci interact non-additively to output a given trait. Therefore, how mutation changes a phenotype can evolve. This means that under conditions of strong stabilizing selection on the trait under consideration, the GRNs can become more robust (less sensitive) to mutation (for example 9). The sensitivity to mutation can also increase under directional selection for the optimum of a trait to change (for example, Kimbrell and Holt (14)). To quantify the sensitivity to mutation of a trait (here, local adaptation), we use an approach similar to Kimbrell and Holt (14), Kimbrell (15) and Draghi and Whitlock (19). The sensitivity to mutation of a trait is calculated by sampling individuals from a specified region in the landscape (range core, range front or a given patch cross-section) at a given time, and for each individual introducing 10 haploid perturbations drawn from $N(0,0.1)$ at one randomly selected locus of the GRN for that trait. The root mean squared difference between the perturbed, and the unperturbed phenotype over all the haploid perturbations and individuals is the sensitivity to mutation. To put it simply, it is a metric of how much the trait changes at the population level when a haploid mutation is introduced to one of the loci encoding it. Note that sensitivity can change only in the case of the GRN model but is fixed in the simple additive and multi-locus models because the latter genetic architectures are additive.
Table S1. Model parameters and variables, their meaning and tested values.

| Model Parameter/ Variable | Description                                                                 | Values                                                                                           |
|---------------------------|-----------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------|
| $N_{x,y,t}$               | Population density in the patch $x, y$ at time $t$                          | dynamical need, encoded either by a GRN or simple additive                                       |
| $d$                       | Dispersal probability                                                       | evokes, encoded either by a GRN or simple additive                                              |
| $\tau$                    | Local adaptation phenotype                                                  | evokes, encoded either by a GRN or simple additive                                              |
| $z$                       | Any evolvable trait                                                         |                                                   |                                                   |
| $\tau_{\text{env}}(x)$    | Value of environmental variable at position $x$                            |                                                   |                                                   |
| $\omega$                  | Niche width                                                                |                                                   |                                                   |
| $\tau_0$                  | Value of environmental variable in range core                              | 1                                                  |                                                   |
| $\lambda_0$               | Intrinsic growth rate in Beverton-Holt model                               | 2                                                  |                                                   |
| $\alpha$                  | Intra-specific competition coefficient in Beverton-Holt model               | 0.01                                              |                                                   |
| $\mu$                     | Dispersal cost                                                             | 0.01, 0.1, 0.3                                     |                                                   |
| $\epsilon$                | Local patch extinction probability                                          | 0, 0.1                                            |                                                   |
| $b$                       | Slope of environmental gradient                                            | 0.02, 0.04                                        |                                                   |
| $m_{\text{min}}$          | Mutation rate at equilibrium                                               | 0.0001                                            |                                                   |
| $m_{\text{max}}$          | Mutation rate at the beginning                                             | 0.1 for GRN, $m_{\text{max}} = m_{\text{min}}$ for others                                    |
| $\sigma_m$                | Effect size (standard deviation) of mutations                              | 0.1                                               |                                                   |
| $x_z$                     | Vector of inputs to the GRN encoding trait $z$                            | in simulation                                      |                                                   |
| $S_z(I)$                  | Vector of expression states of regulatory genes for trait $z$ at iteration $I$ | in simulation                                      |                                                   |
| $U_z$                     | $n \times n$ matrix with each element $U_{z,ji}$ representing connection between the input $j$ and gene $i$ for trait $z$ | evolves, initialised from a normal distribution with sd = 1                                    |
| $W_z$                     | $n \times n$ matrix with each element $W_{z,ki}$ representing connection between the gene $k$ and gene $i$ for trait $z$ | evolves, initialised from a normal distribution with sd = 1                                    |
| $V_z$                     | $1 \times n$ matrix with each element $V_{z,ki}$ representing connection between the gene $k$ and the output | evolves, initialised from a normal distribution with sd = 1                                    |
| $\theta_z$                | Thresholds of regulatory genes                                             | evolves, initialised from a normal distribution with sd = 1                                    |
| $r_z$                     | Slopes of regulatory genes                                                 | evolves, initialised from a normal distribution with sd = 1                                    |
Fig. S1. Effect of dispersal costs ($\mu$) and slope of environmental gradient ($b$) on range front position as a function of time since the beginning of range expansion when there are no random patch extinctions ($\epsilon = 0$). The solid lines indicate the median position of the range front (defined as the occupied patch farthest from the range core) at every 50 time steps over 100 range expansions and the shaded regions represent the quartiles. From left to right, dispersal costs increase and from top to down, the slope of the environmental gradient increases. The general pattern of accelerating range expansions in the GRN model but not in the simple additive model is robust to changes in dispersal costs and the slope of the environmental gradient. Fixed parameters: female fecundity $\lambda_0 = 2$, intra-specific competition coefficient $\alpha = 0.01$, niche width $\omega = 1$, mutation rate during range expansion $m = 0.0001$, number of genes per GRN $n = 3$. 
Fig. S2. Effect of dispersal costs ($\mu$) and slope of environmental gradient ($b$) on range front position as a function of time since the beginning of range expansion when there are random patch extinctions with a probability $\epsilon = 0.1$. The solid lines indicate the median position of the range front (defined as the occupied patch farthest from the range core) at every 50 time steps over 100 range expansions and the shaded regions represent the quartiles. From left to right, dispersal costs increase and from top to down, the slope of the environmental gradient increases. This figure shows that the general pattern of accelerating range expansions in the GRN model but not in the simple additive model holds even if there are random patch extinctions. Fixed parameters: female fecundity $\lambda_0 = 2$, intra-specific competition coefficient $\alpha = 0.01$, niche width $\omega = 1$, mutation rate during range expansion $m_{min} = 0.0001$, number of genes per GRN $\kappa = 3$. 
Fig. S3. Effect of scaling of mutation effects on range dynamics without patch extinctions (ε = 0). We compare range dynamics in the GRN and simple additive models to a model in which 3 additive loci (per locus mutation effects equivalent to the GRN; “3 locus control”) or 21 loci of small effect (“multilocus additive model”) each encode dispersal and local adaptation traits. The solid lines indicate the median position of the range front (defined as the occupied patch farthest from the range core) at every 50 time steps over 100 range expansions and the shaded region represents the quartiles. From left to right, dispersal costs increase and from top to down, slope of the environmental gradient increases. We find that range expansions are overall faster in the 3 locus control model and slower in the multilocus additive model when compared to the GRN and simple additive models. However, like the simple additive model and as opposed the the GRN model, range dynamics are not accelerating in the multilocus additive and 3 locus control models. Range expansions speeds in the GRN model eventually converge to the 3 locus control due to a similar rate of adaptation (see Fig. S12) driven by the evolution of extreme gene expression states in the GRN model (see Fig. 4 of the main text for a detailed explanation). Fixed parameters: female fecundity λ₀ = 2, intra-specific competition coefficient α = 0.01, niche width ω = 1, mutation rate during range expansion m_{min} = 0.0001, number of genes per GRN n = 3.
Fig. S4. Effect of scaling of mutation effects on range dynamics with patch extinctions ($r = 0.1$). We compare range dynamics in the GRN and simple additive models to a model in which 3 additive loci (per locus mutation effects equivalent to the GRN; "3 locus control") or 21 loci of small effect ("multilocus additive model") each encode dispersal and local adaptation traits. The solid lines indicate the median position of the range front (defined as the occupied patch farthest from the range core) at every 50 time steps over 100 range expansions and the shaded region represents the quartiles. From left to right, dispersal costs increase and from top to down, slope of the environmental gradient increases. The general pattern of accelerating range expansions in the GRN model relative to additive architectures holds when there are patch extinctions. Fixed parameters: female fecundity $\lambda_0 = 2$, intra-specific competition coefficient $\alpha = 0.01$, niche width $\omega = 1$, mutation rate during range expansion $m_{\text{min}} = 0.0001$, number of genes per GRN $n = 3$. 

---

Jhelam N. Deshpande and Emanuel A. Fronhofer
Fig. S5. Range front position as a function of time in the absence of an external environmental gradient, with only dispersal evolving. The solid lines indicate median position of the range front (defined as the occupied patch farthest from the range core) at every 50 time steps over 100 range expansions and the shaded region represents the quartiles. From left to right, dispersal costs increase and from top to down, extinction probability increases. Both GRN and simple additive models lead to accelerating range expansion in the absence of an environmental gradient. Fixed parameters: female fecundity $\lambda_0 = 2$, intra-specific competition coefficient $\alpha = 0.01$, mutation rate during range expansion $m_{min} = 0.0001$, number of genes per GRN $n = 3$. 
Fig. S6. Effect of genetic architecture of dispersal on range front position as a function of time since the beginning of range expansion when there are no patch extinctions ($\epsilon = 0$). We compare range expansion dynamics resulting from the GRN, simple additive, GRN LA + simple additive dispersal and simple additive LA + GRN dispersal models. The solid lines indicate median position of the range front (defined as the occupied patch farthest from the range core) at every 50 time steps over 100 range expansions and the shaded region represents the quartiles. From left to right, dispersal costs increase and from top to down, the slope of the environmental gradient increases. Range dynamics are accelerating in the GRN and GRN LA + simple additive dispersal models but not in the simple additive and simple additive LA + GRN dispersal models. This indicates that accelerating range dynamics are observed as long as the local adaptation trait is encoded by a GRN. Fixed parameters: female fecundity $\lambda_0 = 2$, intra-specific competition coefficient $\alpha = 0.01$, niche width $\omega = 1$, mutation rate during range expansion $m_{\text{min}} = 0.0001$, number of genes per GRN $n = 3$. 
Fig. S7. Effect of genetic architecture of dispersal on range front position as a function of time since the beginning of range expansion when there are patch extinctions with probability $\epsilon = 0.1$. We compare range expansion dynamics in the GRN, simple additive, GRN LA + simple additive dispersal and simple additive LA + GRN dispersal models. The solid lines indicate the median position of the range front (defined as the occupied patch farthest from the range core) at every 50 time steps over 100 range expansions and the shaded region represents the quartiles. From left to right, dispersal costs increase and from top to down, the slope of the environmental gradient increases. The qualitative pattern that range expansions are accelerating if local adaptation is encoded by a GRN irrespective of the genetic architecture of dispersal holds, even when there are patch extinctions. Fixed parameters: female fecundity $\lambda_0 = 2$, intra-specific competition coefficient $\alpha = 0.01$, niche width $\omega = 1$, mutation rate during range expansion $m_{min} = 0.0001$, number of genes per GRN $n = 3$. 
Fig. S8. Effect of dispersal evolution on range front position as a function of time since the beginning of range expansion when there are no patch extinctions ($\epsilon = 0$). We compare range dynamics in the GRN, simple additive, GRN LA + fixed dispersal and simple additive LA + fixed dispersal models. The solid lines indicate the median position of the range front (defined as the occupied patch farthest from the range core) at every 50 time steps over 100 range expansions and the shaded region represents the quartiles. From left to right, dispersal costs increase and from top to down, the slope of the environmental gradient increases. The comparison between the GRN and GRN LA + fixed dispersal model range dynamics shows that if dispersal can evolve, range expansions are overall faster and show similar accelerating dynamics. Similarly, range dynamics are overall faster in the simple additive model relative to the simple additive LA + fixed dispersal models. Fixed parameters: female fecundity $\lambda_0 = 2$, intra-specific competition coefficient $\alpha = 0.01$, niche width $\omega = 1$, mutation rate during range expansion $m_{min} = 0.0001$, number of genes per GRN $n = 3$. 

\[ \mu = 0.01 \quad \mu = 0.1 \quad \mu = 0.3 \]
Fig. S9. Effect of dispersal evolution on range front position as a function of time since the beginning of range expansion when there are patch extinctions with probability $\epsilon = 0.1$. We compare range expansion dynamics in the GRN, simple additive, GRN LA + fixed dispersal and simple additive LA + fixed dispersal models. The solid lines indicate the median position of the range front (defined as the occupied patch farthest from the range core) at every 50 time steps over 100 range expansions and the shaded region represents the quartiles. From left to right, dispersal costs increase and from top to down, the slope of the environmental gradient increases. Unlike when there are no patch extinction (Fig. S8), dispersal evolution does not seem to speed up range expansion. However, the fixed dispersal scenarios (GRN LA + fixed dispersal and simple additive LA + fixed dispersal) lead to similar dynamics as the corresponding scenarios in which dispersal evolves (GRN and simple additive models respectively). Fixed parameters: female fecundity $\lambda_0 = 2$, intra-specific competition coefficient $\alpha = 0.01$, niche width $\omega = 1$, mutation rate during range expansion $m_{\text{min}} = 0.0001$, number of genes per GRN $n = 3$. 

\[ \mu = 0.01 \] 
\[ \mu = 0.1 \] 
\[ \mu = 0.3 \] 

\[ b = 0.02 \] 
\[ b = 0.04 \]
Fig. S10. Example dynamics of dispersal, adaptation, sensitivity to mutation and gene expression during range expansion into an external environmental gradient for GRN and simple additive models. Every 15th patch cross-section from the boundary of the range core, we plot the quantities mentioned above as function of time for one example simulation of the GRN and simple additive models. A: Dispersal probability. Dispersal probability is measured as the fraction of individuals emigrating from a patch averaged over a patch cross-section and can change due to dispersal evolution. Some outliers (extremely low or high dispersal values) are visible when the expanding population initially colonizes a patch cross-section due to low population sizes. B: Level of adaptation. The level of adaptation indicates the mismatch between the environment and the genetically encoded local adaptation phenotype averaged over all individuals in the expanding population. We see that in the GRN model, adaptation in patches closer to the range core is slow but proceeds more quickly in patches farther away from the range core. However, in the simple additive model, the rate of adaptation remains the same. C: Sensitivity to mutation as a function of time. We calculate the sensitivity to mutation as describe in the Supplementary Methods. The sensitivity to mutation increases as the expanding population moves along the environmental gradient. D: Mean absolute gene expression. This quantity represents the mean of the absolute value of gene expression states averaged over the patch cross section. Gene expression evolves to extreme values farther from the range core. Focal scenario parameters: slope of gradient $b = 0.04$, fecundity of the females $\lambda_0 = 2$, intra-specific competition coefficient $\alpha = 0.01$, mutation rate during range expansions $m_{max} = 0.0001$, dispersal cost $\mu = 0.1$, extinction probability $\epsilon = 0$, niche width $\omega = 1$, number of genes per GRN $n = 3$. Jhelam N. Deshpande and Emanuel A. Fronhofer 17 of 24
Fig. S11. Relationship between the time to adapt of the expanding population and the sensitivity to mutation of the local adaptation trait, and between the sensitivity to mutation and mean absolute gene expression in the local adaptation GRN for a given environment assuming a GRN genetic architecture. A: Time to adapt as a function of the sensitivity to mutation of the local adaptation trait, for 50 replicates or 100 range expansions, color coded by the distance from the range core. B: Sensitivity to mutation of the local adaptation trait as a function of mean absolute gene expression of the local adaptation GRN, over 50 replicates or 100 range expansions, color coded by the distance from the range core. For a given patch cross section, expanding populations that are more sensitive to mutations adapt more quickly and are associated with extremes of gene expression. Focal scenario parameters: slope of gradient \( b = 0.04 \), fecundity of the females \( \lambda_0 = 2 \), intra-specific competition coefficient \( \alpha = 0.01 \), mutation rate during range expansions \( m_{\text{max}} = 0.0001 \), dispersal cost \( \mu = 0.1 \), extinction probability \( \epsilon = 0 \), niche width \( \omega = 1 \), number of genes per GRN \( n = 3 \).
Fig. S12. Comparison of evolved dispersal probability and time to adapt as a function of the distance from range core for the GRN, simple additive and 3 locus control model. All lines represent median and shaded regions interquartile ranges over 100 replicate range expansions. A: Dispersal probability as function of distance from range core. Dispersal probability increases farther away from the range core for the GRN, simple additive and 3 locus control models. B: The time to adapt as a function of the distance from the range core. Here, we find that time to adapt first increases in all the models explored, and then remains constant at a higher value for the simple additive and 3 locus control models, but decreases in the GRN model. Note that the initial increase is due to standing genetic variation present in the range core. For the GRN model, the time to adapt is initially greater than the simple additive and 3 locus control models but later decreases and eventually converges to the time to adapt in the 3 locus control model. This is because, based on our proposed mechanism for the evolution of increased sensitivity to mutation (see main text Fig. 4), gene expression states evolve to extremes in the GRN model as the expanding population moves farther along the environmental gradient, and mutation effects in this case are comparable to our 3 locus control model. The differences in evolved dispersal probability found in A can be explained by faster local adaptation as seen in B (i.e., reduced maladaptation) and therefore stronger spatial selection leading to increased dispersal. Fixed parameters: female fecundity $\lambda_0 = 2$, intra-specific competition coefficient $\alpha = 0.01$, mutation rate during range expansion $m_{min} = 0.0001$, number of genes per GRN $n = 3$. 
Fig. S13. Interquartile range (IQR) of the range front position as a function of the median when there are no patch extinctions ($e = 0$) for the GRN, GRN LA + simple additive dispersal, simple additive and simple additive LA + GRN dispersal models over 100 replicate range expansions. This representation allows us to compare range dynamics in the various models even though their overall range expansions speeds differ, as we are comparing variability in range dynamics at the same position in the external environmental gradient. From left to right, dispersal costs increase and from top to down, the slope of the environmental gradient increases. The interquartile range first increases with the median position for all the alternative genetic architectures displayed, and then decreases. The later decrease in interquartile range is an artifact because we assume a finite landscape, and the replicate expanding populations have reached the landscape boundary. In models where local adaptation is encoded by a GRN (GRN and GRN LA + simple additive dispersal models), the interquartile range associated with a range front position is greater at the same location in the external environmental gradient in comparison to the models in which local adaptation is encoded by a simple additive architecture (simple additive and simple additive LA + GRN dispersal models). Thus, range expansions are less predictable if a GRN genetic architecture underlies local adaptation. Fixed parameters: female fecundity $\lambda_0 = 2$, intra-specific competition coefficient $\alpha = 0.01$, mutation rate during range expansion $m_{min} = 0.0001$, number of genes per GRN $n = 3$. 
Fig. S14. Effect of random patch extinctions ($c = 0.1$) on the interquartile range (IQR) of range front position for GRN, GRN LA + simple additive dispersal, simple additive and simple additive LA + GRN dispersal models. We plot the interquartile range of the range front position over 100 replicate range expansions as a function of the median. From left to right, dispersal costs increase and from top to down, the slope of the environmental gradient increases. The qualitative result that range expansions are less predictable when a GRN genetic architecture underlies the local adaptation trait irrespective of the genetic architecture of dispersal is robust to the presence of random patch extinctions. Fixed parameters: female fecundity $\lambda_0 = 2$, intra-specific competition coefficient $\alpha = 0.01$, mutation rate during range expansion $m_{mut} = 0.0001$, number of genes per GRN $n = 3$. 

\[ \mu = 0.01 \quad \mu = 0.1 \quad \mu = 0.3 \]
Fig. S15. Effect of scaling of mutation effects on the predictability of range expansion when there are no patch extinctions ($\epsilon = 0$). The interquartile range (IQR) of the range front position is plotted as a function the median for the GRN, simple additive, 3 locus control and multilocus genetic architectures over 100 replicate range expansions. This representation allows us to compare between replicate variation in range expansions across scenarios with differing median speeds. From left to right, dispersal costs increase and from top to down, the slope of the environmental gradient increases. We find that for a given median range border position the between replicate variation is greatest for the GRN model and comparable for the simple additive model and the 3 locus control, and lower for the multilocus additive model. Thus, irrespective of the scaling of mutation effects in the additive architectures (simple additive, 3 locus control and multilocus additive models), the GRN model yields overall more variable range expansions. Fixed parameters: female fecundity $\lambda_0 = 2$, intra-specific competition coefficient $\alpha = 0.01$, mutation rate during range expansion $m_{\text{min}} = 0.0001$, number of genes per GRN $n = 3$. 
Fig. S16. Effect of scaling of mutation effects on the predictability of range expansion when there are patch extinctions ($\epsilon = 0.1$). We plot the interquartile range of the range front position over 100 replicate range expansions as a function of median range front position. This representation allows us to compare between replicate variation in range expansion across scenarios with differing median speeds. From left to right, dispersal costs increase and from top to down, the slope of the environmental gradient increases. The qualitative result that the GRN model yields a greater between replicate variation in range expansion irrespective of the mutational scaling of the additive genetic architectures (simple additive, 3 locus control and multilocus additive models) holds when there are patch extinctions. Fixed parameters: female fecundity $\lambda_0 = 2$, intra-specific competition coefficient $\alpha = 0.01$, mutation rate during range expansion $m_{\min} = 0.0001$, number of genes per GRN $n = 3$. 
References

1. D Bonte, et al., Costs of dispersal. *Biol. Rev.* **87**, 290–312 (2012).
2. RJH Beverton, SJ Holt, *On the dynamics of exploited fish populations.* (Chapman & Hall, London), (1957).
3. HJ Poethke, T Hovestadt, Evolution of density–and patch–size–dependent dispersal rates. *Proc. Royal Soc. Ser. B: Biol. Sci.* **269**, 637–645 (2002).
4. P Alberch, From genes to phenotype: dynamical systems and evolvability. *Genetica* **84**, 5–11 (1991).
5. M Pigliucci, Genotype–phenotype mapping and the end of the ‘genes as blueprint’metaphor. *Philos. Trans. R. Soc. B-Biol. Sci.* **365**, 557–566 (2010).
6. D Nichol, M Robertson-Tessi, ARA Anderson, P Jeavons, Model genotype–phenotype mappings and the algorithmic structure of evolution. *J. R. Soc. Interface* **16**, 20190332 (2019).
7. A Wagner, Evolution of gene networks by gene duplications: a mathematical model and its implications on genome organization. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 4387–4391 (1994).
8. A Spirov, D Holloway, Using evolutionary computations to understand the design and evolution of gene and cell regulatory networks. *Methods* **62**, 39–55 (2013).
9. A Wagner, Does evolutionary plasticity evolve? *Evolution* **50**, 1008 (1996).
10. ML Siegal, A Bergman, Waddington’s canalization revisited: Developmental stability and evolution. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 10528–10532 (2002).
11. S Ciliberti, OC Martin, A Wagner, Innovation and robustness in complex regulatory gene networks. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 13591–13596 (2007).
12. E Rüneburger, AL Rouzic, Why and how genetic canalization evolves in gene regulatory networks. *BMC Evol. Biol.* **16** (2016).
13. J Draghi, GP Wagner, The evolutionary dynamics of evolvability in a gene network model. *J. Evol. Biol.* **22**, 599–611 (2009).
14. T Kimbrell, R Holt, Canalization breakdown and evolution in a source-sink system. *Am. Nat.* **169**, 370–382 (2007).
15. T Kimbrell, Canalization and adaptation in a landscape of sources and sinks. *Evol. Ecol.* **24**, 891–909 (2010).
16. JW Malcom, Smaller gene networks permit longer persistence in fast-changing environments. *PLoS ONE* **6**, e14747 (2011).
17. JW Malcom, Gene networks and metacommunities: Dispersal differences can override adaptive advantage. *PLoS ONE* **6**, e21541 (2011).
18. CJ Melián, et al., Deciphering the interdependence between ecological and evolutionary networks. *Trends Ecol. Evol.* **33**, 504–512 (2018).
19. JA Draghi, MC Whitlock, Phenotypic plasticity facilitates mutational variance, genetic variance, and evolvability along the major axis of environmental variation. *Evolution* **66**, 2891–2902 (2012).
20. J van Gestel, FJ Weissing, Regulatory mechanisms link phenotypic plasticity to evolvability. *Sci. Rep.* **6**, 24524 (2016).
21. B Novák, JJ Tyson, Design principles of biochemical oscillators. *Nat. Rev. Mol. Cell Biol.* **9**, 981–991 (2008).
22. C Espinosa-Soto, Selection for distinct gene expression properties favours the evolution of mutational robustness in gene regulatory networks. *J. Evol. Biol.* **29**, 2321–2333 (2016).
23. GS Nido, JM Williams, L Benuskova, Bistable properties of a memory-related gene regulatory network in *The 2012 International Joint Conference on Neural Networks (IJCNN).* pp. 1–6 (2012) ISSN: 2161-4407.