Robustness, Evolvability and Adaptive Architecture of Heterogeneous Cell Populations

Supplementary Figures and legends

Authors: Andrei Kucharavy\textsuperscript{a,b,c}, Boris Rubinstein\textsuperscript{d}, Jin Zhu\textsuperscript{a,b}, Rong Li\textsuperscript{a,b}

Affiliations:
\textsuperscript{a} Center for Cell Dynamics, Department of Cell Biology, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA.
\textsuperscript{b} Department of Chemical and Biomolecular Engineering, Johns Hopkins University, Baltimore, MD, 21218, USA
\textsuperscript{c} UMR 7238 CNRS - Université Pierre et Marie Curie, Paris, 75006, France.
\textsuperscript{d} Stowers Institute for Medical Research, Kansas City, MO, 64110, USA.

Correspondence:
Dr. Rong Li; e-mail: rong@jhu.edu
Supplementary Figure 1: Effect of increasing robustness on the likelihood of emergence of adaptive subpopulation

A-G: The likelihood of the emergence of adaptive subpopulation is represented both from the point of view of μ-σ correlation (right graphs), and values of H distribution (left graphs). In the right graphs, Points obtained by simulation are in black, theoretical results are in red, Y values are: 0.25 (A), 0.5 (B), 0.75 (C), 1 (D), 1.5 (E), 2 (F), 3 (G). In the left graph, red curve represents mean and pink curves - SD of H.
Supplementary Figure 2: Effect of robustness on fitness is independent from population dispersion in trait space and population size

A-C: The effect of increasing population heterogeneity on the emergence of the adaptive subpopulation. Population heterogeneity is represented as the spread in the trait space ($\sigma_A$), whereas the likelihood of emergence of adaptive subpopulation is represented both from the point of view of $\mu$-$\sigma$ correlation (right graphs), and value of $H$ (left graphs). $\sigma_A$ values are 0.25 (A), 0.5 (B), 0.75 (C).

D-E: The effect of increasing population size on the emergence of the adaptive subpopulations. No effect on the slope of mean-SD correlation (right), but highly adaptive subpopulations are more likely to emerge with larger population size (left, points above 0). Total number of subpopulations: 10 (D), 100 (E). Displays of right and left panels are as in Supplementary figure 1.
Supplementary Figure 3: A simple approximation can be used to accurately perform complexity and robustness regression.

A-D: Approximated values of the mean of H and standard deviation of H for different values of γ. The approximation is in black and the complete expression is in red. γ values: 0.75(A), 1(B), 1.5(C), 2(D). N is fixed at 40.

E-F: Approximated values of μ and σ for different values of trait space dimension N. Approximation is in black and complete expression is in red. N values: 10(E), 100(F)
Supplementary Figure 4: Fitness of Aneuploid yeast strains and the Breast Cancer cell lines under diverse growth conditions relative to no-stress conditions.

A: Fitness relative to the no-stress condition of different aneuploid strains of yeast. All the euploid strains (U1 - ControlHaploid, U2 - ControlDiploid, U3 - controlTriploid) are present in duplicates. Stresses with similar names differ by concentrations. Stresses ranked by the relative fitness decrease compared to the no stress conditions. Population generalist (Haploid) is in black.
B: Fitness relative to no drug conditions for breast cancer cell line collection. Treatments are ranked by the average relative fitness of all cell lines in a given treatment. Population generalist (BT483) is in black, WT_proxy is the average fitness between the reference non-cancerous cell lines (184A1 and 184B5). Drug concentrations are chosen for maximum differentiation between cell lines (see “Breast Cancer Cell lines collection response to drug treatments” section for more details).
Supplementary Figure 5: Effect of anisotropic population distribution in trait space.

In order to illustrate the potential effect of the anisotropy of the trait space with respect to adaptation to the environment, we simulated the aneuploid distribution while pooling SD along each axis according to log-normal distribution with parameters $\mu=0$, $\sigma=1$.

A: Regression plot for the isotropic population distribution (N=40, $\sigma_A=0.75$, black) matching well the anisotropic population distribution (N=60, $\sigma_A=0.59$, red).

B: Distribution density for distance $d$ from the generalist subpopulation for the isotropic population distribution (N=40, $\sigma_A=0.75$, black) matching well that for anisotropic populations (N=60, $\sigma_A=0.59$, red)

C: Regression plot for the isotropic population distribution (N=60, $\sigma_A=0.59$, black) roughly matching the anisotropic population distribution (N=60, $\sigma_A=0.59$, red)
D: Distribution density for distance from the generalist subpopulation for the isotropic population distribution (N=60, σA=0.59, black) compared to that for the anisotropic population distribution (N=60, σA=0.59, red)
Robustness, Evolvability and Adaptive Architecture of Heterogeneous Cell Populations: Supplementary Information

Authors: Andrei Kucharavy a,b,c, Boris Rubinstein d, Jin Zhu a,b, Rong Li a,b

Affiliations:
a Center for Cell Dynamics, Department of Cell Biology, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA.
b Department of Chemical and Biomolecular Engineering, Johns Hopkins University, Baltimore, MD, 21218, USA
c UMR 7238 CNRS - Université Pierre et Marie Curie, Paris, 75006, France.
d Stowers Institute for Medical Research, Kansas City, MO, 64110, USA.

Correspondence:
Dr. Rong Li; rong@jhu.edu
Full derivation of the mathematical model:

Statement of the problem:

The general model is built of the following abstract components:

1. A heterogeneous set of subpopulations represented by points $\vec{r}_0$ distributed randomly in N-dimensional (ND) trait vector space according to a multinormal distribution;
2. A generalist system is at the origin $\vec{r}_0 = 0$;
3. An environment, acting similarly on all systems (including the generalist system), represented by a vector $\vec{S}$ in the ND trait space, pointing to the combination of traits that would enable optimal adaptation. Different environments are represented by vectors of different sizes and directions;
4. A fitness function $G$ maps distance $d$ between the environment optimum $\vec{S}$ and system position $\vec{r}_i$. It reaches unity at $d = 0$ and monotonically decreases on IR+, the positive real numbers.

We also introduce a transformation function $H$. It acts on the ratio of fitness $G_{int} = G(|\vec{r}_i + \vec{S}|)$ of the element $\vec{r}_i$ to the performance $G_{ref} = G(|\vec{S}|)$ of the generalist subpopulation, positioned at the origin, both being in the environment $\vec{S}$. We are interested in the statistical characteristics of the transformation function values as a function of the distance between the generalist subpopulation position and the point of optimal performance for the new conditions.

The choice of the functions $G$ and $H$ will be discussed below, but we will assume that the performance function $G$ is monotonous decreasing and positive, while the transformation function $H$ is monotonically increasing. The case where function $G$ belongs to the family $G_y(d) = \exp(-d^{2\gamma}/s^{2\gamma})$ indexed by the robustness parameter $\gamma$, where $s$ is the characteristic robustness distance in the trait space (scaling factor) and $H = \ln(G_{int}/G_{ref})$, is of particular interest to us as it captures a general, commonly encountered case.

General approach (symmetric multinormal distribution)

Consider computation of raw moments of the quantity that depends on the distance $d = |\vec{r}_i + \vec{S}|$ from the origin, where the shift vector $\vec{S}$ in ND space has a given length $l = |\vec{S}|$, and the point $\vec{r}_i$ follows a multinormal distribution with zero mean and standard deviation $\sigma_A$ for all its components. Without loss of generality, one can define the vector $\vec{S}$ having only a single nonzero component equal to $l$, i.e., $\vec{S} = \{l, 0, 0\}$.

Recall the definition of non-central $\chi^2(N, \lambda)$ distribution (Kendall and Stuart, 1961) of a random variable $q$ as a distribution of the sum

$$ q \equiv \sum_{i=1}^{N} \frac{X_i^2}{\sigma_i^2}, \quad \lambda = \sum_{i=1}^{N} \frac{\mu_i^2}{\sigma_i^2}, $$

(2.1)
where $X_i$ are independent normally distributed random variables with means $\mu = \{\mu_i\}$ and standard deviations $\sigma = \{\sigma_i\}$. When $\mu_i = \mu = 0$ and $\sigma_i = \sigma$, the probability distribution function (PDF) of $N$ variables $\bar{x} = \{x_i\}$ is found as a product of the independent normal distributions

$$P(\bar{x}) = \left(\frac{1}{2\pi\sigma^2}\right)^{N/2} \exp\left[-\frac{|\bar{x}|^2}{2\sigma^2}\right]$$

(2.2)

where $\sigma = \sigma_A / \sqrt{N}$. Comparing this definition with (2.1) one can see that the value $d^2 = |\bar{r}_i + \tilde{S}|^2$ has a distribution expressed as

$$d^2 = \sum_{i=1}^{N} X_i^2 = \sigma^2 \sum_{i=1}^{N} \frac{X_i^2}{\sigma_i^2} \equiv \frac{\sigma_i^2}{N} \chi^2(N, \lambda), \quad \lambda = \frac{N\sigma^2}{\sigma_A^2},$$

(2.3)

where the non-central $\chi^2(N, \lambda)$ has $\mu_i = 1$ and $\mu_i = 0$ ($2 \leq i \leq N$). From (2.3) it follows that the random variable $q$ can be expressed as $q = N\sigma^2 / \sigma_A^2 > 0$, and it satisfies non-central $\chi^2(N, \lambda)$ distribution with a density

$$P_N(q, \lambda) = \frac{e^{-(1+\lambda)/2} q^{(N-1)/2} \sqrt{\lambda}}{2(\lambda q)^{N/4}} I_{N/2-1}(\sqrt{\lambda q})$$

$$= 2^{-N/2} q^{N/2-1} \exp[-(\lambda + q) / 2] \bar{F}_1(\lambda / 2; N / 2; \lambda q / 4)$$

(2.4)

where $I_{\nu}$ is the modified Bessel function and $\bar{F}_1$ denotes regularized confluent hypergeometric function.

We use the analytic transformation function

$$H(r, l) = H\left(\frac{G(r)}{G(l)}\right),$$

where $G$ denotes a performance function. In the simplest case, $H$ can be chosen as an identity function $H(x) = x$, other choices are $H(x) = \ln x$, or $H(x) = \log_2 x$. Expressing $r = \sigma_A \sqrt{q / N}$, and $\alpha = \sigma_A / \sqrt{N}$ we have for the k-th raw moment of the function $H(r, l)$

$$m_k(\lambda) = \int_0^\infty P_N(q, \lambda) H^k(\alpha \sqrt{q}, \alpha \sqrt{\lambda}) dq,$$

(2.5)
Below we denote by $G(q) = G(\sqrt{\lambda}) = \sum_{n=0}^{\infty} C_{kn}(\lambda, \alpha)q^n$. As the function $H_k(r, l)$ can always be expanded in the Taylor series

$$H^k(\alpha \sqrt{q}, \alpha \sqrt{\lambda}) = \sum_{n=0}^{\infty} C_{kn}(\lambda, \alpha)q^n,$$

one has to find a general form for the raw moment of $q$ which is computed as

$$\mu'_v = \int_0^\infty P_N(q, \lambda)q^vdq = 2^\nu e^{-\lambda/2}(\nu + N/2)_{\nu + N/2}(\nu + N/2 + 1/2),$$

which is also valid for both integer ($v = n$) and noninteger values of $v$. The moments for small integer $v$ read

$$\mu'_0 = 1$$
$$\mu'_1 = N + \lambda$$
$$\mu'_2 = N^2 + 2N(\lambda + 1) + \lambda(\lambda + 4)$$
$$\mu'_3 = N^3 + 3N^2(\lambda + 2) + N(3\lambda^2 + 18\lambda + 8) + \lambda(\lambda^2 + 12\lambda + 24)$$
$$\mu'_4 = N^4 + 4N^2(\lambda + 3) + 2N^2(3\lambda^2 + 24\lambda + 22) + 4N(\lambda^3 + 15\lambda^2 + 44\lambda + 12) + \lambda(\lambda^3 + 24\lambda^2 + 144\lambda + 192)$$

which at $\lambda = 0$ simplify to $\mu'_v = 2^\nu \Gamma(\nu + N/2)/\Gamma(N/2)$. In the evaluation of these expressions, we used the formula (Prudnikov et al., 1992)

$$\int e^{-pq^2}q^{a-1}I_v(cq)dq = 2^{-(v+1)}c^vp^{-(a+v)/2}\left(\frac{\alpha + v}{2}\right)_{\nu + l + 1}.$$
The k-th moment $m_k$ reads

$$m_k = \int_0^\infty P_N(q, \lambda) \ln^k (\exp(-\alpha^2 \gamma (q^m - \lambda^m)) dq$$

$$= (-1)^k (\alpha / s)^{2k} \int_0^\infty P_N(q, \lambda)(q^m - \lambda^m)^k dq$$

$$= (-1)^k (\alpha / s)^{2k} \sum_{n=0}^{\infty} C^n_k (-1)^{(k-n)} \lambda^{\gamma(k-n)} \int_0^\infty P_N(q, \lambda) q^m dq$$

$$= (\alpha^2 \lambda / s^2)^k \sum_{n=0}^{\infty} C^n_k (-1)^n \frac{\mu'_n}{\lambda^m}$$

(3.1)

The asymptotics of the $\mu'_\nu$ for large values of $\lambda$ reads

$$\mu'_\nu \approx \lambda \nu \sum_{k=0}^{\infty} \frac{2^\nu \Gamma(\nu + N/2 - 1)\Gamma(\nu)}{k! \Gamma(\nu + N/2 - k - 1)\Gamma(\nu - k)\lambda^2} + O(1/\lambda^N)$$

(3.2)

Thus we have

$$m_1 = \mu$$

$$= (\alpha / s)^{2\gamma} (\lambda^2 - \mu'_\gamma)$$

$$\approx -\gamma(2\gamma + N - 2) \frac{\lambda}{\sigma^2} (\alpha^2 \lambda / s^2)^\gamma$$

(3.3)

$$\approx -\gamma(2\gamma + N - 2) \frac{\sigma^2}{N} (l / s)^{2\gamma}$$

The last expression can be rewritten as

$$\mu \approx -\frac{\gamma(2\gamma + N - 2)}{N} \frac{(l / s)^{2\gamma}}{(l / \sigma_\lambda)^2},$$

(3.4)

and we observe that the mean value of the transformation function is always non-positive and depends on the new environment deviation from the original generalist state as $l^{2(\gamma-1)}$, so that it monotonically decreases for $\gamma > 1$, monotonically increases for $\gamma < 1$, and is constant for $\gamma = 1$, as it was shown in the previous subsection. The asymptotic at small $\lambda \ll 1$ reads

$$\mu'_\nu \approx 2^\nu \frac{\Gamma(\nu + N/2)}{\Gamma(N/2)} \left[ 1 + \frac{\lambda \nu}{N} + \frac{\lambda^2 \nu(\nu - 1)}{2N(N + 2)} + O(\lambda^3) \right].$$

(3.5)

Using it we obtain for $\lambda = 0$
\[ m_l(0) = -(2\alpha^2/s^2)^\gamma \frac{\Gamma(\gamma+N/2)}{\Gamma(N/2)} = -(2\sigma_A^2/N s^2)^\gamma \frac{\Gamma(\gamma+N/2)}{\Gamma(N/2)}, \]

The second moment reads
\[ m_2 = (\alpha^2/s^2)^{2\gamma} (\lambda^{2\gamma} - 2\mu_{2\gamma} \lambda^{\gamma} + \mu_{2\gamma}^2), \]
and we obtain
\[ \sigma^2 = (\alpha^2/s^2)^{2\gamma} (\mu_{2\gamma} - \mu_{2\gamma}^2) \implies \sigma = (\alpha^2/s^2)^\gamma \sqrt{\mu_{2\gamma} - \mu_{2\gamma}^2}. \quad (3.6) \]

Using the asymptotic expansion (3.2) we find
\[
\begin{align*}
\mu_{2\gamma} - \mu_{2\gamma}^2 &\approx \lambda^{2\gamma} \left[ 1 + \frac{2\gamma(4\gamma+N-2)}{\lambda} + \frac{\gamma(2\gamma-1)(4\gamma+N-2)(4\gamma+N-4)}{\lambda^2} \right] \\
&\quad - \lambda^{2\gamma} \left[ 1 + \frac{2\gamma(2\gamma+N+2)}{\lambda} + \frac{\gamma^2(2\gamma+N-2)^2}{\lambda^2} \right] \\
&\quad + \lambda^{2\gamma} \left[ \frac{4\gamma^2}{\lambda} + \frac{\gamma^2}{\lambda^2} (\gamma-1)(N^2 + 28\gamma(\gamma-1) + 8) + 2N(1+2(3\gamma-1)(\gamma-1)) \right]
\end{align*}
\]

so that in the lowest order we obtain
\[ \sigma \approx (\alpha^2 \lambda / s^2)^{\gamma} \frac{2\gamma}{\sqrt{\lambda}} = \frac{2\gamma}{\sqrt{N}} \frac{(l/s)^{2\gamma}}{(l/s_A)}. \quad (3.7) \]

thus the standard deviation is inversely proportional to square root of the trait space dimension \( N \). It depends on the new environment deviation from the original generalist state magnitude as \( l^{2(\gamma-1)} \), so that it monotonically decreases for \( \gamma < 1/2 \), monotonically increases for \( \gamma > 1/2 \), and is constant for \( \gamma = 1/2 \). The expansion (3.5) in the lowest order produces
\[ \sigma(0) = \left( \frac{2\sigma_A^2/N s^2}{\Gamma(N/2)} \right)^\gamma \sqrt{\frac{\Gamma(2\gamma+N/2)\Gamma(N/2)}{\Gamma(N/2)^2}}. \]

From (3.3, 3.6) a simple relation between the mean and the standard deviation for large \( l \) follows
\[ \sigma = -\frac{2\mu l \sqrt{N}}{(2\gamma+N-2)\sigma_A}. \]

Expressing here \( l \) from (3.7) we find the asymptotic relation
\[
\sigma = (2^\gamma)^{-\frac{1}{2(\gamma-1)}} \left( \frac{s\sqrt{N}}{\sigma_A} \right)^{\left(\frac{\gamma}{(\gamma-1)}\right)} \left( \frac{2}{2^\gamma + N - 2} \right)^{\frac{2\gamma-1}{2\gamma(\gamma-1)}} (-\mu)^{\frac{2\gamma-1}{2\gamma(\gamma-1)}}.
\] (3.7)

**Further improvements:**

The approach presented in the manuscript produces general conclusions but also has several limitations to its applicability. One of the major model assumptions which drastically simplifies the analysis consists in the isotropy of multinormal distribution of the points \( \bar{p}_i \). The anisotropic version of this distribution described in (Kotz et al., 1967) would lead to much cumbersome computation of the distribution moments. Calculating it is an enterprise of its own and is outside the scope of this paper.

Another assumption is that the external environment shifts the element distribution without distorting it; in the general case one can expect an anisotropic effect of the environment change \( \vec{S}(\vec{r}_i) \), shifting the points \( \vec{r}_i \) with respect to each other. In the first approximation one can set \( \vec{S}(\vec{r}_i) = \vec{S}_0 + \delta \vec{S}(\vec{r}_i) \), where \( \vec{S}_0 \) represents the vector applied to the generalist system \( \vec{r}_0 \), \( |\delta \vec{S}(\vec{r}_i)| \ll |\vec{S}_0| \), and develop a perturbation theory. Once again, this is outside the scope of this paper.

The usage of \( \chi^2 \) distribution is valid only for Euclidean norm \( L_2 \) on the environment, which is consistent with independent action of shifts along each axis of the trait space affecting fitness independently in case \( \gamma = 1 \). Different norm definitions are possible a priori, but are unlikely to be analytically tractable.

**Supplementary Table 1:**

Chemicals and conditions used in aneuploid yeast phenotypic profiling

| Perturbation           | Supplier                  | Concentration       |
|------------------------|---------------------------|---------------------|
| Cycloheximide          | Sigma-Aldrich C4859       | 0.25 µg/mL          |
| Hydroxyurea            | MP BIOMEDICALS 02102023   | 100mM               |
| 4-Nitroquinoline N-oxide | Sigma-Aldrich N8141     | 0.4 µg/mL          |
| Benomyl                | Sigma-Aldrich 381586     | 30 µg/mL            |
| Dimethyl sulfoxide (DMSO) | Sigma-Aldrich D2650   | 1%                  |
| Thiolutin              | Sigma-Aldrich T3450       | 8.3 µg/mL           |
| Bleomycin              | Sigma-Aldrich B5507       | 0.01U/mL            |
| High pH                |                           | pH 8                |
| 16°                    |                           |                     |
| 23°                    |                           |                     |
**References:**

Kendall, M. G., and Stuart, A. (1961). The advanced theory of Statistics - Inference and Relationship, New York, NY: Hafner.

Kotz, S., Johnson, N., and Boyd, D. W. (1967). Series representations of distributions of quadratic forms in normal variables. I. Central case. Ann. Math. Stat. 38, 823–837.

Prudnikov, A., Brychkov, Y., and Marichev, O. (1992). Integrals and Series: Special Functions, Philadelphia, PA: Gordon.