Trends in Biotechnology

CellPress

Opinion

Subpopulation networks: a scaling principle for engineering living systems

Frank Delvigne [1,*], Fabian Roeloffs [1], Laurie Josselin [1], Steve Pressé [2,3,4,5], Lisbeth Olsson [6,*], and Alexander Grünberger [7,*]

Cell populations underpin applications ranging from bioproduction and biotherapeutics to microbiome engineering and biocomputing. However, treating them as bulk quantities rather than as structured systems limits our understanding of population-level stability, productivity, and robustness. Cellular populations self-organize into coexisting subpopulations (SPs) representing phenotypically distinct states that emerge from the interplay of biological noise and selective pressures. We propose that networks of interacting SPs constitute a fundamental organizational layer for engineering stable cell population dynamics. Surveying published single-cell studies across bacterial, yeast, and mammalian systems, we show that populations consistently self-organize into a limited number of SPs. This dimensional reduction can be captured experimentally by clustering single-cell data and gives rise to scaling laws that link population structure to collective performance.

Highlights

Single-cell surveys show cell populations self-organize into few subpopulations.

Subpopulation networks compress high-dimensional dynamics and reveal emergent scaling laws.

Emergent subpopulation network properties define new engineering rules for living systems.

Cell population diversity is a resource to be exploited as a primary engineering variable

The ability to design, manipulate, and control cell populations has become a central ambition of modern biotechnology and synthetic biology. Engineered cellular systems now underpin applications ranging from the bioproduction of complex molecules and live biotherapeutics to microbiome-mediated health promotion, biosensing, active biomaterials, and biocomputing interfaces. Across these diverse applications, a common requirement emerges, that is, the need for reliable, long-term, population-level stability that remains robust against the inherent stochasticity of biological systems, driven by fluctuating real-world conditions [1,2]. A convergent body of evidence suggests that population-level stability rarely arises from homogeneous cellular systems [3]. Instead, cellular populations self-organize into coexisting **subpopulations (SPs)** (see Glossary) that represent phenotypically distinct states emerging from the interplay of biological noise and selective pressures. The coordinated dynamics of these SPs then generate emergent population-level behaviors.

Diauxic shift [4] and antibiotic persistence [5] in bacterial cultures, oscillations in gene expression responsible for circadian timekeeping [6], spatial patterning in tissues and biofilms [7], and the collective dynamics of gut microbiome communities [8] all reflect the action of identifiable SPs rather than average behavior. Advances in microbiome engineering and developmental biology have further shown that macroscopic collective behavior, including **community succession**, morphogenetic patterning, and tissue homeostasis, can be quantitatively predicted from the dynamics of a small number of interacting cell types or states, suggesting that population dynamics are intrinsically low-dimensional [9–11].

A fundamental challenge in deploying engineered cellular systems is not their design but their stability, that is, maintaining a desired population composition over time and across spatial

[1]Terra Research and Teaching Centre, Microbial Processes and Interactions (MiPI), Gembloux Agro-Bio Tech, University of Liège, Gembloux, Belgium
[2]Center for Single Molecule Biophysics, Biodesign Institute, Arizona State University, Tempe, AZ, USA
[3]Center for Biological Physics, Arizona State University, Tempe, AZ, USA
[4]Department of Physics, Arizona State University, Tempe, AZ, USA
[5]School of Molecular Sciences, Arizona State University, Tempe, AZ, USA
[6]Division of Industrial Biotechnology, Department of Life Sciences, Chalmers University of Technology, SE-411 96 Gothenburg, Sweden
[7]Microsystems in Bioprocess Engineering, Institute of Process Engineering in Life Sciences, Karlsruhe Institute of Technology, Karlsruhe, Germany

*Correspondence:
F.Delvigne@uliege.be (F. Delvigne), lisbeth.olsson@chalmers.se (L. Olsson), and alexander.gruenberger@kit.edu (A. Grünberger).

© 2026 Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

scales, or, more generally, the functional output it encodes, against perturbations, evolutionary drift, and environmental fluctuations [1]. Recent advances in the control of biological systems, including **cybergenetic** feedback architectures [12] and **optogenetic** actuators [13], have substantially improved our ability to steer cellular behavior in real time. Still, the population is mostly treated as a bulk quantity rather than as a structured, multistate system. A mechanistic and quantitative understanding of how SP structure shapes population-level stability, productivity, and robustness is crucially lacking. Therefore, principled design rules for engineering these systems urgently need to be developed.

We propose that networks of interacting SPs (SP networks) constitute a fundamental organizational layer to describe, predict, and engineer stable cell population dynamics. We further address two central questions in the development of SP networks: (i) can SP networks explain emergent properties of cellular populations that are difficult to explain at the single-cell level? (ii) Can SP networks be reliably inferred from experimental data to provide a principled explanation for the observed low dimensionality of population dynamics?

The origin and importance of cellular SPs

We begin with a simple, well-established mechanism for SP generation. Biological noise arising from stochastic fluctuations in transcription, translation, and downstream metabolic processes generates cell-to-cell variability in gene expression, even within a genetically homogeneous population (Figure 1A) [45]. This noise broadens the distribution of intracellular states, so that a single input can drive a range of phenotypic outputs. Beyond this intrinsic variability, noise propagates through gene regulatory networks, recruiting additional gene clusters and amplifying initial fluctuations into stable, coherent phenotypic programs. This propagation, combined with the competitive dynamics imposed by cultivation conditions, that is, differences in specific growth rate, substrate affinity, and stress tolerance, organizes the population into discrete, coexisting SPs, that is, clusters of cells that share a functional state and follow reproducible trajectories in gene expression, protein accumulation, and biological output (Figure 1A) [46].

Recent studies, however, indicate that SP formation can be more complex than this picture suggests, involving feedback from growth on gene expression as well as self-induced environmental feedback (Figure 1B) [47,48]. When gene expression carries a growth trade-off, cell populations can exhibit unexpected **bistability** [49] or oscillatory behavior [50,51], with different modes of switching between SPs. This link between growth-coupled gene expression and complex population behavior can be explained through resource allocation. Cells draw on a finite pool of resources to coordinate the activation and deactivation of gene circuits [52]. When a target gene circuit imposes a growth–expression trade-off, resource allocation accounts for the timing differences observed in gene expression across cells [53]. These timing differences, combined with self-induced environmental feedback (cells growing more slowly leave additional substrate available to actively proliferating cells), drive the emergence of SPs with pronounced phenotypic differences. This environmental feedback has recently been recognized as an important component in mitigating the trade-off between growth and gene expression and in promoting population stability [47,49,51,54]. Ultimately, ecological interactions may also arise between SPs exhibiting differential fitness [55].

The main concept suggested here builds on the self-organization of cellular populations into a limited number of SPs, opening new avenues for their engineering. However, the identification and classification of SPs critically depend on how single-cell data are analyzed and subsequently classified. Single-cell data collection ranges from low-dimensional readouts, such as single or multiple fluorescent reporters, to high-dimensional omics approaches, such as single-cell RNA-sequencing (scRNA-seq) (Figure 1C). Analysis of the number of SPs and clusters identified across

Glossary

Bet-hedging: a diversification strategy in which a genetically identical population produces phenotypically distinct SPs in advance of environmental change, trading some short-term fitness for reduced variance in reproductive success under fluctuating conditions.
Bistability: a property of a regulatory network in which two stable steady states (e.g., an 'ON' and an 'OFF' expression state) coexist under the same environmental conditions, typically arising from positive feedback and nonlinearity, allowing genetically identical cells to settle into distinct, stable phenotypes.
Community succession: the sequential, directional change in the composition and relative abundance of species or SPs within a microbial community over time, driven by resource depletion, metabolite exchange, and shifting selective pressures.
Cybergenetics: a discipline that applies control-theory principles (feedback, real-time sensing, and actuation) to steer gene expression or cellular behavior, often through computer-in-the-loop control or genetically encoded controllers.
Diauxic shift: the transition in a microbial culture from consuming a preferred carbon source to a secondary one, marked by a lag phase during which gene expression is reprogrammed (e.g., the classic glucose-to-lactose shift in *E. coli*).
Optogenetics: the use of light-sensitive proteins to control gene expression or cellular signaling with high spatial and temporal precision, enabling noninvasive, reversible perturbation of cellular behavior.
Subpopulation (SP): a cluster of cells that share the same genetic background and exhibit phenotypic diversification.

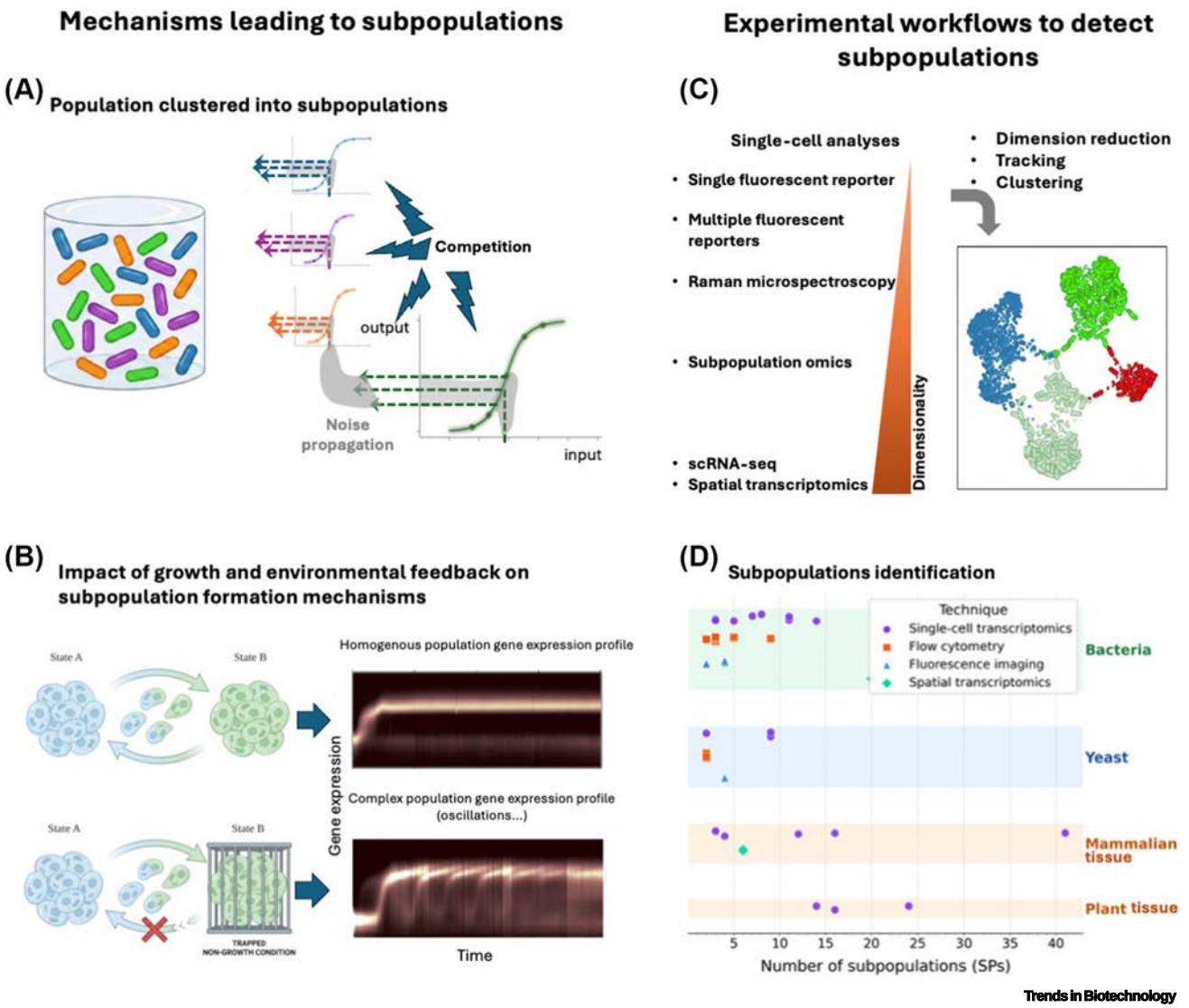


Figure 1. Mechanisms generating cellular subpopulations and experimental strategies for their detection. (A) Biological noise broadens the distribution of cellular responses to a given input, converting a homogeneous population into a phenotypically heterogeneous one. Noise propagation across regulatory networks and competitive interactions between cell types drive the organization of the population into discrete SPs, each following distinct trajectories of gene expression and biological function. (B) Growth feedback on gene expression adds another layer of complexity to cell population dynamics. (C) Experimental workflows for SP detection span increasing levels of dimensionality, from single fluorescent reporters and multiplexed biosensor panels to single-cell RNA-sequencing (scRNA-seq). Dimensionality reduction, temporal tracking, and clustering convert high-dimensional data into an interpretable SP map. (D) The application of various single-cell approaches to diverse biological systems has led to the discovery of a limited number of SPs (links to references and experimental details are provided in Table 1). SPs: subpopulations.

various biological systems and techniques (Figure 1D and Table 1) reveals that cellular populations tend to self-organize into a limited number of SPs. In our survey of published data, the number of SPs ranged from 2 to 41, with an average of five for yeast and bacterial systems. For plant and mammalian systems, the reported number of SPs was higher, averaging 15 and reflecting the higher resolution of the profiling techniques (predominantly scRNA-seq) used to analyze these more complex systems. Across case studies, the term ‘subpopulation’ appears to be broadly adopted by the scientific community, though it is sometimes used interchangeably with ‘cell clusters’ or, more rarely, replaced by ‘cellular states’ (Table 1).

Table 1. Detection of SPs or clusters of cells across biological systems following different single-cell analytical approaches[a,b]

Biological system	Technique(s)	Terminology used	Number of SPs	Key observation	Refs
Bacterial systems					
Cocultures of *E. coli* and *Bacillus subtilis*	Microbial split-pool ligation transcriptomics (microSPLiT)	SPs/clusters	14 (+3 rare SPs i.e., exhibiting very low abundance)	Able to detect SP as low as 0.142%	[14]
Persistence in *E. coli*	Pooled scRNA-seq (PETRI-seq)	Clusters	7	Growth-dependent clusters resolved, with rare phage-lysis state identified	[15]
Clinical isolates of *Klebsiella pneumoniae*	BacDrop (bacterial scRNA-seq)	SPs	11 (including rare SPs)	Antibiotic treatment leads to the appearance of SPs	[16]
Plasmid conjugation in *Pseudomonas putida* KT2440	microSPLiT scRNA-seq	SPs/clusters	11	Occurrence of a nonconjugating SP	[17]
Pseudomonas aeruginosa biofilm	Parallel sequential fluorescence *in situ* hybridization	SPs/clusters	20	Mature biofilms contain functional SPs that are spatially localized	[18]
Clinical isolates of *E. coli* and *Staphylococcus aureus*	Microcolony-seq	SPs	*E. coli* EPEC 5 *E. coli* UPEC 8 *S. aureus* 3	Bacterial memory promotes phenotypic inheritance and clear SP trajectories	[5]
Culture of *B. subtilis* with metabolically different cells	Cell sorting and RNA-seq	SPs	3	*B. subtilis* planktonic culture self-organize into SPs with overflow metabolism and cross-feeding	[19]
Chemostat of *E. coli* and *P. putida*	Automated flow cytometry and propidium iodide staining	SPs	3	Microbial cultures in glucose-limited chemostat exhibit different SPs with different level of membrane permeability	[20]
Different *Pseudomonas* species cultivated in flask under different metabolic regimes	Flow cytometry with DAPI and NanoSIMS	SPs	5	Both NanoSIMS and flow cytometry pointed out the same number of SPs	[21]
Shake flask cultures of *E. coli*	Flow cytometry and Raman microspectroscopy	SPs/clusters	9	No SPs detected based on flow cytometry and 9 SPs detecting based on Raman microspectroscopy with PhenoGraph clustering method	[22]
Sinorhizobium meliloti in microfluidics	Flow cytometry and triple fluorescent reporter	SPs	3	Generation of SPs can occur in perfectly homogenous environment in microfluidics	[23]
Biocatalysis with P. putida	Flow cytometry and single fluorescent reporter	SPs	3	The relative abundance of SPs depends on the expression system, that is, either *alk* or *lac*-based	[24]
Xylanase production by Thermobacillus xylanilyticus	Flow cytometry and FSC (cell size) measurement	SPs	2	High FSC SP is associated with higher xylanase production	[25]
Recombinant protein production in E. coli strains with modified glucose transport capabilities	Flow cytometry and single fluorescent reporter	SPs	3	*lac*-based systems exhibit bimodal behavior and are affected by glucose transport alterations	[26]
Behavior of an arabinose-inducible system upon continuous culture of E. coli	Automated flow cytometry and single fluorescent reporter	SPs	2	Glucose and arabinose co-feeding in continuous culture triggers two SPs	[27]
Stress response in E. coli biofilm	Fluorescence imaging and triple fluorescent reporter	SPs	4	Stress response and corresponding SPs are spatially localized in biofilm	[7]
Cellular functions in B. subtilis biofilm	Fluorescence imaging and combinations of double fluorescent reporter (182 strains)	SPs	4 (+ overlapping states)	Many genes are coexpressed	[28]
Sugar and overflow metabolism of E. coli in microfluidics	Fluorescence imaging and double fluorescent reporter	SPs	2	The two SPs are only separated by 11 micrometers, suggesting that microbial metabolic interactions occur on short spatial scale	[29]

Table 1. (continued)

Biological system	Technique(s)	Terminology used	Number of SPs	Key observation	Refs
Yeast–bacteria coculture					
Dynamics of activation of *lacZ* and *stl1* in *E. coli* and *Saccharomyces cerevisiae*, respectively	Single-molecule RNA fluorescence *in situ* hybridization (smFISH)	Cell states	2 (*E. coli*) 4 (*S. cerevisiae*)	Based on snapshot smFISH data and Bayesian inference, the authors were able to infer the dynamics and state occupancy for two microbial systems	[30]
Yeast systems					
Bet-hedging (cell decision between growth and stress resistance) in *S. cerevisiae*	Flow cytometry and single fluorescent reporter	Cell states	2	Deep learning is used to capture the impact of gene expression and growth based on snapshot flow cytometry data	[31]
Osmostress response in *S. cerevisiae*	scRNA-seq	SPs/clusters	9	Under osmotic stress, cells explore various adaptive strategies through SP trajectories	[32]
Growth-synthesis coupling in *S. cerevisiae*	scRNA-seq	SPs/clusters	9	Contrary to expectations, growth–synthesis coupling did not alleviate SPs formation	[33]
Production stress in *S. cerevisiae* producing lycopene or violacein	Flow cytometry and fluorescent biosensor (pH sensor)	SPs	2	During bioprocesses, low- and high-intracellular pH SPs are generated and exhibit differential functionalities depending on the target product; that is, either lycopene or violacein	[34]
Amino acids cross-feeding in *S. cerevisiae*	Differential isotope labeling by amino acids	SPs	2	The producer consumer SPs are observed for several amino acids and are spatially located in agar colonies	[35]
Mammalian systems					
Cellular dynamics in human breast cancer	Spatial proteomics and one-shot tissue dynamics reconstruction	Cell types or states	6	Methods to predict cellular dynamics from snapshot data	[36]
Cellular states inside tumor from 77 studies	scRNA-seq	SPs/clusters	41 (metaprograms)	Intratumor heterogeneity can be clustered into 41 metaprograms	[37]
Analysis of 198 cancer cell lines from 22 cancer types	scRNA-seq	SPs/clusters	12	12 recurrent heterogeneous programs have been identified	[38]
Analysis of 15 cancer types	scRNA-seq	Clusters	16	16 consensus genes modules have been identified, but cancer cells can mobilize several modules	[39]
CHO cell cultures	Single nucleus sequencing	SPs/clusters	2–4	DNA methylation dynamics induce SPs	[40]
Induced pluripotent stem cells-derived endothelial cells	Droplet-based scRNA-seq	SPs/clusters	4	The four SPs are marked by enrichment of one specific gene	[41]
Plant systems					
Developmental program in *Arabidopsis thaliana*	scRNA-seq	SPs/clusters	24	The SPs are linked mainly to specific developmental programs. However, six of these SPs correspond to new cell types not previously identified	[42]
Developmental program during tissue formation in *A. thaliana*	scRNA-seq	Clusters/cell types	16	The authors have detected a finite number of cell clusters involved in the formation of tissue during the development of *A. thaliana*	[43]
Developmental program in root development of *A. thaliana*	scRNA-seq	Cell types	14	The authors have detected a finite number of cell clusters involved in the formation of root during the development of *A. thaliana*	[44]

[a]For all references across all systems, the number of SPs/clusters has been performed with the help of an AI agent (see the Supplementary Information online for details about the methodology).
[b]CHO: Chinese hamster ovary; DAPI: 4′,6-diamidino-2-phenylindole; FSC: forward scatter; NanoSIMS: nanoscale secondary ion mass spectrometry; PETRI-seq: Prokaryotic-Expression profiling by Tagging RNA In situ and sequencing.

Converting high-dimensional single-cell data into a tractable SP description requires dimensionality reduction techniques (e.g., t-distributed stochastic neighbor embedding or uniform manifold approximation and projection), which project data onto low-dimensional embeddings where clusters corresponding to SPs become computationally and visually accessible (Figure 1C). However, beyond the characterization of individual SPs at individual time points, understanding population dynamics requires describing how SPs interact and evolve over time, while most experimental techniques rely on the acquisition of snapshot biological data. The issue of recovering continuous dynamics from snapshot data is considered in the following section.

How can cellular SPs be tracked from snapshot data?

Reconstructing how each SP changes over time is essential for capturing their dynamic nature, but this is hard since most single-cell measurements only capture a snapshot of the population at one instant (Figure 2), and inferring how cells actually moved through time from a snapshot alone is a long-standing challenge in biophysics [30]. Here too, reducing the data to a small number of dimensions and constraining the search to a discrete set of SPs can make this trajectory reconstruction more robust.

Several strategies exist for reconstructing trajectories from snapshots. One treats each snapshot as a sample drawn from an underlying process, with the same set of SPs assumed to persist throughout, with cells shifting between them in different proportions along an unobserved trajectory that happens to end at the moment the snapshot was taken. By writing down the probability of observing a given snapshot under this assumption and working backward from that probability (a Bayesian approach), researchers can recover the underlying kinetic parameters driving the system (e.g., how fast a protein is produced, how often a gene switches on and off, or how variable cell division is from one cell to the next). This strategy was recently used to reconstruct the trajectories yeast cells follow during a stress response, based only on flow cytometry snapshots[31], and has also been used to recover gene regulatory networks in yeast and *Escherichia coli* (Figure 2B) [30].

When cell numbers are large enough, the random cell-to-cell fluctuations described earlier average out, and the population can instead be described with simple deterministic equations tracking average abundances. This is essentially the same inference problem but in the large-population limit: rather than inferring the parameters of a random process from a spread of snapshot values, one instead infers a deterministic rule for how the population changes on average, directly from how heterogeneous it appears in a single snapshot. For instance, Somer *et al.* examined a single tissue biopsy and used the spatial pattern of a cell division marker (Ki67) as a proxy for local proliferation rates, asking how each cell type's chance of dividing depended on the makeup of its immediate neighborhood [36]. By fitting these neighborhood-dependent division rates to the biopsy's spatial heterogeneity, they built a dynamical model (a set of differential equations) describing how the abundance of each cell subtype (or SP) changes over time. This allowed them to reconstruct full temporal trajectories and phase portraits of the population's dynamics from a single static snapshot, without ever tracking cells over actual time.

A different way around the limits of snapshot-based inference is the molecular recorder, that is, a genetic system engineered to write a record of a cell's transcriptional history directly into its genomic DNA, to be read out later by sequencing [56,57]. These emerging tools offer a way to recover a cell's gene expression history at single-cell resolution, effectively giving each cell its own logged trajectory, pushing the temporal reach of SP analysis beyond what snapshot or time-lapse approaches can currently provide.

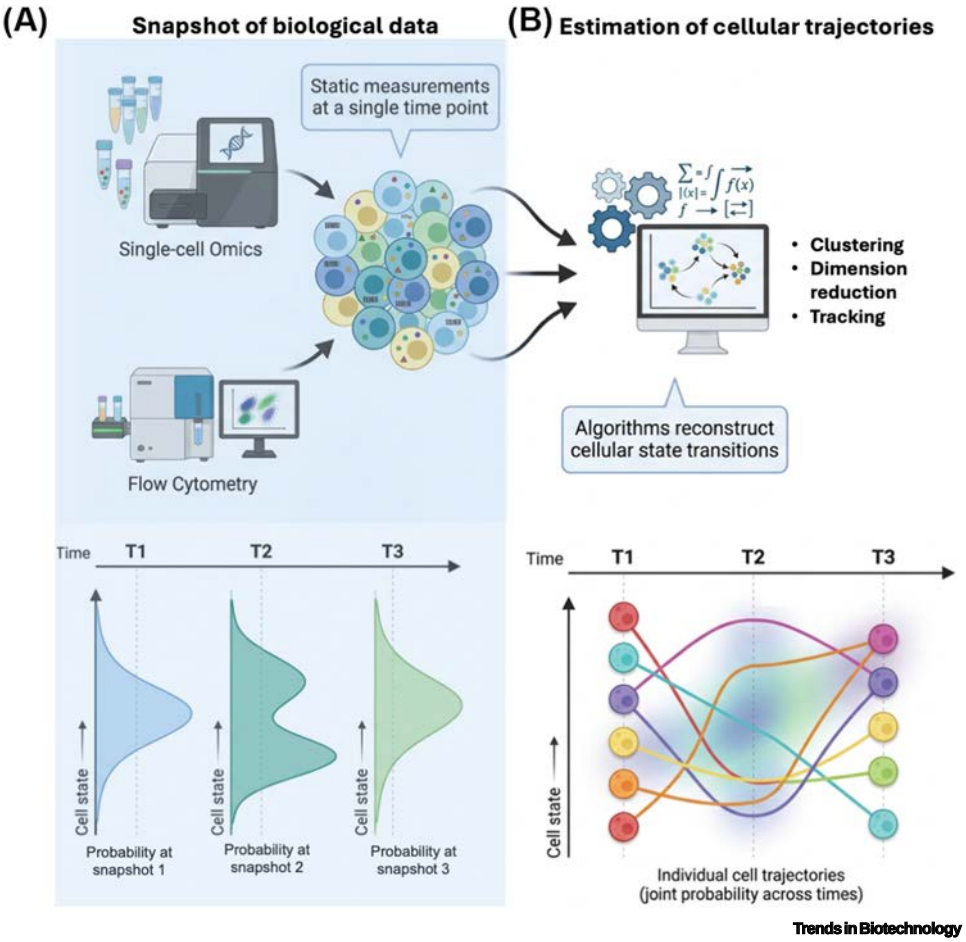


Figure 2. From single-cell snapshots to cellular trajectories: the challenge of reconstructing SP dynamics. (A) Snapshot measurements capture the marginal distribution of cell states at a single time point. Single-cell omics and flow cytometry generate measurements of heterogeneous cell populations, resolving distinct SPs (represented by differently colored and shaped cells). When collected at successive time points (T1, T2, and T3), each measurement provides only the snapshot of cell states at that instant, a series of disconnected cross-sections that does not directly reveal how individual cells or SPs transition between states over time. (B) Computational approaches reconstruct cellular trajectories from snapshot data. Applying dimensionality reduction, clustering, and trajectory inference algorithms to the snapshot data allows the reconstruction of cellular state transitions. As individual cells cannot be tracked continuously over time, the resulting trajectories represent slices of the probability distribution of cell states at various times, posing what is often an ill-posed inference problem whose solution depends on the constraints and assumptions built into the algorithm. Colored lines depict individual paths that cell SPs can take from T1 to T2 and then to T3, illustrating the diversity of possible transition paths between SPs. SPs: subpopulations.

Emerging engineering principles from SPs networks and applications

Phenotypic switching, competitive exclusion, cooperative cross-feeding, and differential responses to environmental perturbations collectively determine how population composition evolves over time (Figure 1A,B). We propose to capture these interactions in an SP network, that is, a single graph representation in which nodes represent distinct SPs and directed, weighted links encode rates of transition, competition, or cooperation between them (Figure 3A). This representation is formally analogous to microbial community networks, in which a network of interacting species framework has helped decode complex collective behavior from the properties of individual members [58–60]. An additional layer of complexity arises in single-species systems: cells can switch from one SP to another over time, and this dynamic

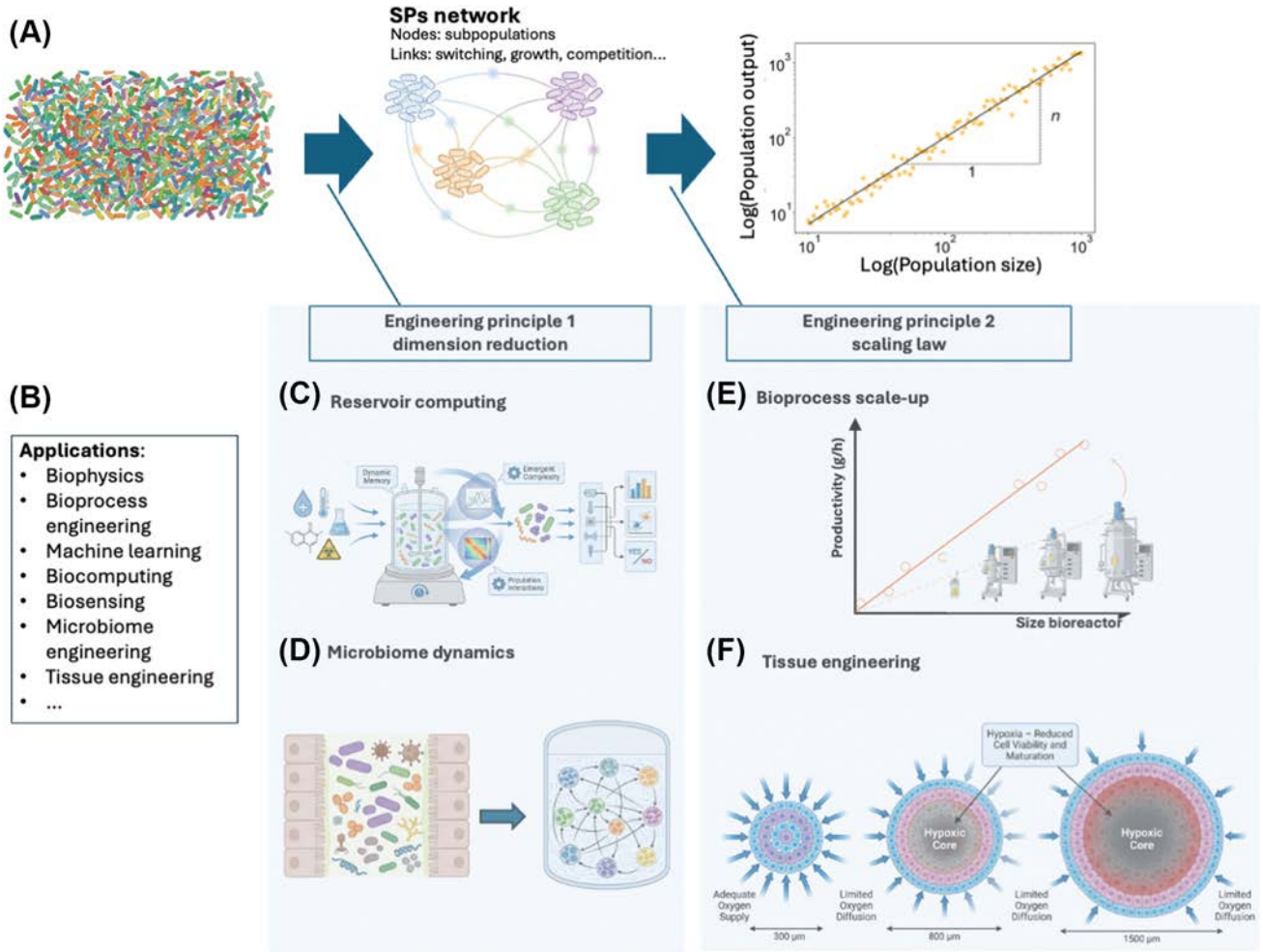


Figure 3. Emergence of new engineering principles from SP networks. (A) Left: a cell population that appears highly complex at the individual level, with each cell tracing its own trajectory through a vast phenotypic space driven by stochastic gene expression, metabolism, and cell cycle progression, self-organizes into a discrete set of functionally distinct SPs. Nodes represent SPs and directed, weighted edges encode rates of phenotypic switching and competitive or cooperative interactions between them, forming an SP network. (A) Center: this self-organization collapses the high-dimensional phenotypic space onto a low-dimensional manifold. Despite the thousands of molecular variables involved, the collective behavior of the population can be captured and controlled by a small number of emergent coordinates, represented here as a compressed low-dimensional state space. This dimensionality reduction is a general property of populations organized into SP networks, independent of the specific biological system considered. (A) Right: once collective dynamics operate on a low-dimensional manifold, scaling laws of the form $P \sim N^{\alpha}$ become accessible, where P is a measure of collective performance and N is population or system size. When α is greater than one, the system exhibits superlinear scaling, producing disproportionately greater output as it grows, a signature of collective efficiency arising from optimized interaction structure. The log–log linearity of this relationship, with slope greater than one, is the fingerprint of low-dimensional manifold dynamics. (B–F) Four applications illustrate the generality of these two principles across scientific fields: (C) reservoir computing, where microbial consortia provide naturally rich, heterogeneous, and memory-bearing dynamical responses that satisfy the computational requirements of a high-quality reservoir; (D) microbiome engineering, where SP-aware niche design and biosensor interpretation enable more accurate characterization and control of community function; (E) bioprocess scale-up, where SP network structure governs the trade-off between growth and production and may yield principled scale-up rules; and (F) tissue engineering, where oxygen transport scaling laws and dimensionality reduction of cell state dynamics together constrain and guide construct design across scales.

behavior complicates how edges in the SP network should be quantified and represented. Addressing this will require new computational tools, since existing approaches for constructing SP networks rely on cell-to-cell similarity inferred from sequencing profiles, a static measure that does not, on its own, capture the transition dynamics between SPs [61,62].

Two engineering principles emerge from this organization. First, representing a population as an SP network produces a dramatic reduction in the effective dimensionality needed to describe and control collective behavior [9,63]. Second, once a population operates on this low-dimensional manifold, scaling laws linking population size to collective performance become accessible (Figure 3A) [64–66]. At first glance, a cell population is highly complex. However, when the population self-organizes into a network of discrete SPs, each occupying a distinct functional state, the degrees of freedom collapse dramatically [67–70]. This feature was recently demonstrated by showing that the full temporal dynamics of microbial growth curves can be compressed into low-dimensional representations by autoencoder neural networks without losing predictive power, with these compressed embeddings performing as well as or better than raw data for identifying bacterial strains, predicting antibiotic resistance, and forecasting community dynamics [9]. Once a population operates on a low-dimensional manifold, a second property becomes accessible, that is, the emergence of scaling laws. A scaling law, expressed generally as $P \sim N^{\alpha}$ where P is a measure of collective performance and N is system size, reveals that output is governed not by the raw number of components but by the intrinsic structure of their interactions. When α is greater than one, the system is superlinear, doing more per unit size as it grows, a signature of genuine collective efficiency [71].

Superlinearity is widespread across biological and human systems. At the level of human populations, patent production scales superlinearly with city size, yielding roughly a 15% productivity gain in the largest cities, not because any individual becomes more creative, but because the population is organized as an optimized interaction network in which ideas, skills, and resources are exchanged with increasing efficiency as the network grows [72]. The log–log linearity of this relationship, with a slope greater than one, is the fingerprint of a system whose dynamics live on a low-dimensional manifold. We argue that SP networks in cell populations can exhibit the same property and that identifying and engineering the interaction structure of these networks is the key to unlocking superlinear collective performance.

Antibiotic tolerance provides a compelling example of how SP networks can explain emergent population-level behavior that is difficult to understand from single-cell observations alone [5]. In a seemingly homogeneous bacterial culture, cells often occupy distinct phenotypic states, including fast-growing, slow-growing, and persistent cells. Such SPs are connected through dynamic transitions driven by stochastic fluctuations and environmental cues [5,73]. While individual cells may transiently enter or leave tolerant states, population-level survival during antibiotic exposure emerges from the collective dynamics of the network rather than from the properties of any single cell [73]. For example, the presence of a small persister SP can ensure survival of the entire population despite the susceptibility of most cells. An SP network representation captures both the abundance of these SPs and the transition rates between them, thereby providing a low-dimensional and mechanistic description of antibiotic tolerance. In this framework, tolerance emerges as a network property arising from the flow of cells between phenotypic states rather than as an intrinsic characteristic of individual cells.

We further challenge this expectation across four case studies spanning reservoir computing, microbiome engineering, bioprocess scale-up, and tissue engineering, showing in each case how the SP network perspective either reframes an existing challenge or points toward a solution that purely reductionist approaches have missed (Figure 3B–F).

Biocomputing and reservoir computing

Reservoir computing is a machine learning approach where a major part of the underlying computational neural network consists of a real physical, chemical, or biological reservoir. It

converts complex temporal inputs through a fixed, high-dimensional dynamical system whose rich responses are decoded by a simple, trained readout layer. A good reservoir requires two properties, that is, response richness, meaning distinguishable outputs for distinct inputs, and fading memory, meaning sensitivity to recent inputs that decays for more distant ones. SP networks naturally provide both; for example, Faigenbaum-Romm *et al.* showed that genetically identical bacteria spontaneously organize into stable SPs with distinct functional states, generating inherently heterogeneous temporal responses to the same input [5]. Duncker *et al.* demonstrated this computationally, engineering a microbial consortium in which distinct SPs sensed different chemical signals, producing rich multidimensional temporal outputs that a neural network compressed and a simple readout decoded accurately, even under strong crosstalk and in complex hospital wastewater samples [74]. This architecture is structurally identical to reservoir computing. In general, several studies have reported that SPs carry memory of past conditions through protein inheritance across cell divisions, effectively providing the reservoir with multiple tunable timescales of memory simultaneously, a property artificial reservoirs cannot easily achieve [5,31]. The low-dimensional structure of SP dynamics identified by Baig *et al.* further implies that even modest consortia may provide sufficient reservoir richness, reducing both biological complexity and data requirements for training [9].

Microbiome engineering

Microbial communities are heterogeneous, dynamic, and history-dependent in their composition and function. The SP network framework has direct implications for niche design: because a population contains functionally distinct SPs, local nutrient gradients, substrate composition, mechanical cues, and inducers can be tuned to selectively favor desired SPs [47,75]. A recent study further showed that microbial growth dynamics, though inherently high-dimensional time series data, can be compressed into low-dimensional representations using autoencoder neural networks and that these compressed representations retain, or even improve upon, the predictive power of the raw data [9]. Such embeddings perform as well as or better than raw growth curves for identifying bacterial strains, predicting antibiotic resistance, and forecasting community dynamics.

Bioprocess scale-up

Bioprocesses are typically scaled-up based on operational parameters from chemical process engineering, including volumetric power, mixing time, and oxygen transfer rate [76,77]. To date, no direct scale-up rule derived from cell population data has been established, and we argue that the SP network framework provides a more robust way to achieve scalability. Studies have identified coexisting SPs in bioreactors representing two functional extremes, that is, high-producing but slow-growing cells and fast-growing but low-producing cells [78,79]. These SPs participate in a network of interactions governing the trade-off between growth and production. Mapping this network and identifying its low-dimensional coordinates would enable a principled, population-aware scale-up rule, replacing the current empirical single-parameter approach with one grounded in the intrinsic structure of the cell community [79].

Tissue engineering

Cell populations in engineered tissues are similarly heterogeneous and history-dependent, making SP-aware scaffold design equally relevant here. Stem and progenitor cells seeded into a scaffold rarely differentiate synchronously: local variation in oxygen tension, substrate stiffness, and paracrine signaling drives individual cells along divergent trajectories, so that the resulting construct is itself a mosaic of SPs whose relative proportions and spatial arrangement shape functional outcome at least as much as any single average cell property. The field's canonical scaling constraint is the oxygen diffusion limit, which restricts construct thickness to roughly 100–200 micrometers in the absence of vascularization [80]. Beyond this scale, cells at the

construct core experience hypoxia and nutrient depletion that push them toward distinct, often quiescent, or necrotic SPs, so the transport limitation and the heterogeneity problem compound one another rather than acting independently. Combining SP-aware niche engineering, that is, designing scaffold architecture and spatially patterned signaling cues to steer the composition of the population itself, rather than only its mean phenotype, with this quantitative scaling constraint, alongside machine learning tools for real-time organoid monitoring and phenotype classification, offers a path toward constructs whose design is guided simultaneously by population structure and physical transport scaling [81].

Concluding remarks and future perspectives

Several open questions define the near-term research agenda for SP networks, mapping directly onto the two central questions raised in the first section, 'Cell population diversity is a resource to be exploited as a primary engineering variable', of this work. Addressing whether SP networks can be reliably inferred from data to explain the low dimensionality of the apparent population function requires resolving the origin of the limited SP count observed across systems. It remains unclear whether this low count is a genuine biological signature, reflecting selection that compresses a high-dimensional phenotypic space into a small number of stable, functionally distinct states, or whether it instead reflects the limited resolution of current experimental and clustering methods. Disentangling these possibilities will require systematically comparing SP counts obtained from techniques of increasing dimensionality, moving from single reporters to multiplexed biosensors, and scRNA-seq applied to the same biological system to test whether the inferred SP number converges or continues to grow with resolution. A count that converges across resolutions would support low dimensionality as an intrinsic, selectable property of the population rather than a measurement artifact.

The answer to the second question, that is, whether SP networks can explain emergent properties that are difficult to account for at the single-cell level, depends in turn on characterizing the mechanisms behind SP emergence itself, which remain incompletely understood. We have proposed a quantitative link between gene expression timing and resource allocation at the single-cell level: competition for shared resources such as RNA polymerase and ribosomes offers a plausible mechanistic basis for the differential expression timing that drives SPs to emerge and stabilize. Translating this into predictive, quantitative models, that is, by connecting circuit design, growth–expression trade-offs, and resulting SP composition, would move the field from qualitative explanation toward engineering design rules and would show that the existence, stability, and collective function of SPs are mechanistically traceable to resource-allocation dynamics that are not visible in a single-cell snapshot.

Finally, the intersection of SP networks with biocomputation and machine learning offers a particularly promising direction for advancing both questions at once. Treating SPs as the latent variables underlying compressed representations of population dynamics directly operationalizes the question. If population-scale behavior can be reconstructed from a small number of inferred SP states, that reconstruction is itself evidence that the observed low dimensionality is real and exploitable rather than an artifact of measurement. At the same time, using these inferred states to design control strategies that steer SP network structure directly leads to an important application, that is, a control strategy that reliably shapes emergent population behavior through SP-level intervention, rather than single-cell-level intervention. This would be strong evidence that the SP network is the right organizational layer at which to describe, predict, and engineer collective function. Addressing how SP organization connects to the broader scaling laws of biological systems will be essential to moving from describing SP networks to engineering them (see Outstanding questions).

Outstanding questions

Why do we typically observe a limited number of SPs? Is it an effect of selection (dimension reduction) or experimental artifact?

What is the link and quantitative biology between timing in gene expression and resource allocation at the single-cell level?

Can we select SPs based on ecological engineering, that is, providing specific metabolic niches such as in microbiome engineering? Can we find optimal SP network structure for targeted applications? From a biocomputation and machine learning perspective, could the SPs be assimilated to the latent variables describing the whole population dynamics?

Acknowledgments

L.J. is supported by a postdoctoral research grant provided by the Service Public de Wallonie (SPW; Wagralim Proactif project) and by the Walloon Region and the Fonds européen de développement régional (FEDER) (Grant portfolio PHENIX n°244, project PHENIX_Technologique_ULiège n°397 and Grant portfolio PHENIX n°244, project PHENIX_FoodBooster_ULiège n°428). F.R. is supported by a Fond pour la Recherche en Industrie et Agriculture (FRIA) PhD grant from the Fonds de la Recherche Scientifique (F.R.S.-FNRS). A.G. and F.D. gratefully acknowledge support and funding from the Deutsche Forschungsgemeinschaft (project no. 516641394) and the F.R.S.-FNRS (project no. 40013556), respectively (Weave collaboration network). L.O. acknowledges support from the Swedish Research Council (Dnr 2024-0650).

Declaration of interests

The authors declare no competing interests.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used Claude Sonnet 4.6 in order to assist in collecting the data described in the supplementary information. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Supplemental information

Supplementary information associated with this article can be found online at https://doi.org/10.1016/j.tibtech.2026.09.018.

References

1. Liao, M.J. *et al.* (2019) Rock-paper-scissors: engineered population dynamics increase genetic stability. *Science* 365, 1045–1049
2. Rugbjerg, P. and Sommer, M.O.A. (2019) Overcoming genetic heterogeneity in industrial fermentations. *Nat. Biotechnol.* 37, 869–876
3. Martins, B.M.C. and Locke, J.C.W. (2015) Microbial individuality: how single-cell heterogeneity enables population level strategies. *Curr. Opin. Microbiol.* 24, 104–112
4. Solopova, A. *et al.* (2014) Bet-hedging during bacterial diauxic shift. *Proc. Natl. Acad. Sci. U. S. A.* 111, 7427–7432
5. Faigenbaum-Romm, R. *et al.* (2025) Uncovering phenotypic inheritance from single cells with Microcolony-seq. *Cell* 188, 5313–5331.e18
6. Isomura, A. and Kageyama, R. (2014) Ultradian oscillations and pulses: coordinating cellular responses and cell fate decisions. *Development* 141, 3627–3636
7. Zoheir, A.E. *et al.* (2023) A three-colour stress biosensor reveals multimodal response in single cells and spatiotemporal dynamics of biofilms. *NPJ Biofilms Microbiomes* 9, 57
8. Voigt, R.M. *et al.* (2016) Circadian rhythm and the gut microbiome. *Int. Rev. Neurobiol.* 131, 193–205
9. Baig, Y. *et al.* (2023) Autoencoder neural networks enable low dimensional structure analyses of microbial growth dynamics. *Nat. Commun. Engl.* 14, 7937
10. Garza, D.R. *et al.* (2025) Emergence of alternative states in a synthetic human gut microbial community. *Nat. Commun.* 17, 326
11. Karkaria, B.D. *et al.* (2021) Automated design of synthetic microbial communities. *Nat. Commun.* 12, 672
12. Aoki, S.K. *et al.* (2019) A universal biomolecular integral feedback controller for robust perfect adaptation. *Nature* 570, 533–537
13. Benisch, M. *et al.* (2023) Optogenetic closed-loop feedback control of the unfolded protein response optimizes protein production. *Metab. Eng.* 77, 32–40
14. Kuchina, A. *et al.* (2021) Microbial single-cell RNA sequencing by split-pool barcoding. *Science* 371, eaba5257
15. Blattman, S.B. *et al.* (2024) Identification and genetic dissection of convergent persister cell states. *Nature* 636, 438–446
16. Ma, P. *et al.* (2023) Bacterial droplet-based single-cell RNA-seq reveals antibiotic-associated heterogeneous cellular states. *Cell* 186, 877–891.e14
17. Cyriaque, V. *et al.* (2024) Single-cell RNA sequencing reveals plasmid constrains bacterial population heterogeneity and identifies a non-conjugating subpopulation. *Nat. Commun.* 15, 5853
18. Dar, D. *et al.* (2021) Spatial transcriptomics of planktonic and sessile bacterial populations at single-cell resolution. *Science* 373, eabi4882
19. Rosenthal, A.Z. *et al.* (2018) Metabolic interactions between dynamic bacterial subpopulations. *Elife* 7, e33099
20. Sassi, H. *et al.* (2019) Segregostat: a novel concept to control phenotypic diversification dynamics on the example of Gram-negative bacteria. *Microb. Biotechnol.* 12, 1064–1075
21. Calabrese, F. *et al.* (2019) Quantitation and comparison of phenotypic heterogeneity among single cells of monoclonal microbial populations. *Front. Microbiol.* 10, 2814
22. Garcia-Timermans, C. *et al.* (2020) Raman spectroscopy-based measurements of single-cell phenotypic diversity in microbial populations. *mSphere* 5 e00806-20
23. Schlüter, J.-P. *et al.* (2015) Classification of phenotypic subpopulations in isogenic bacterial cultures by triple promoter probing at single cell level. *J. Biotechnol.* 198, 3–14
24. Lindmeyer, M. *et al.* (2015) Variability in subpopulation formation propagates into biocatalytic variability of engineered *Pseudomonas putida* strains. *Front. Microbiol.* 6, 1042
25. Bouchat, R. *et al.* (2022) Xylanase production by *Thermobacillus xylanilyticus* is impaired by population diversification but can be mitigated based on the management of cheating behavior. *Microb. Cell Factories* 21, 39
26. Fragoso-Jiménez, J.C. *et al.* (2019) Growth-dependent recombinant product formation kinetics can be reproduced through engineering of glucose transport and is prone to phenotypic heterogeneity. *Microb. Cell Factories* 18, 26
27. Nguyen, T.M. *et al.* (2021) Reducing phenotypic instabilities of a microbial population during continuous cultivation based on cell switching dynamics. *Biotechnol. Bioeng.* 118, 3847–3859
28. Yannarell, S.M. *et al.* (2023) Extensive cellular multi-tasking within *Bacillus subtilis* biofilms. *mSystems* 8, e0089122
29. Dal Co, A. *et al.* (2019) Metabolic activity affects the response of single cells to a nutrient switch in structured populations. *J. R. Soc. Interface* 16, 20190182
30. Kilic, Z. *et al.* (2023) Gene expression model inference from snapshot RNA data using Bayesian non-parametrics. *Nat. Comput. Sci.* 3, 174–183

31. Pessoa, P. *et al.* (2026) Simulation-based inference captures non-Markovian effects as exemplified in protein production kinetics through cell division. *Proc. Natl. Acad. Sci. U. S. A.* 123, e2517309123
32. Nadal-Ribelles, M. *et al.* (2025) A single-cell resolved genotype-phenotype map using genome-wide genetic and environmental perturbations. *Nat. Commun.* 16, 2645
33. Qin, L. *et al.* (2026) Design of the biosensor-dependent coupling system stabilizes the high-synthesis phenotype of cell factory. *Nat. Commun.* 17, 5081
34. Savigny, J. *et al.* (2025) Single cell profiling framework reveals metabolic subpopulations as drivers of bioproduction heterogeneity. *Nat. Commun.* 17, 645
35. Kamrad, S. *et al.* (2023) Metabolic heterogeneity and cross-feeding within isogenic yeast populations captured by DILAC. *Nat. Microbiol.* 8, 441–454
36. Somer, J. *et al.* (2026) Temporal tissue dynamics from a spatial snapshot. *Nature* 650, 490–499
37. Gavish, A. *et al.* (2023) Hallmarks of transcriptional intratumour heterogeneity across a thousand tumours. *Nature* 618, 598–606
38. Kinker, G.S. *et al.* (2020) Pan-cancer single-cell RNA-seq identifies recurring programs of cellular heterogeneity. *Nat. Genet.* 52, 1208–1218
39. Barkley, D. *et al.* (2022) Cancer cell states recur across tumor types and form specific interactions with the tumor microenvironment. *Nat. Genet.* 54, 1192–1201
40. Böhl, E. *et al.* (2024) Analysis of population heterogeneity in CHO cells by genome-wide DNA methylation analysis and by multimodal single-cell sequencing. *J. Biotechnol.* 396, 72–79
41. Paik, D.T. *et al.* (2018) Large-scale single-cell RNA-seq reveals molecular signatures of heterogeneous populations of human induced pluripotent stem cell-derived endothelial cells. *Circ. Res.* 123, 443–450
42. Zhang, T.-Q. *et al.* (2021) A single-cell analysis of the *Arabidopsis* vegetative shoot apex. *Dev. Cell* 56, 1056–1074.e8
43. Yang, B. *et al.* (2021) Non-cell autonomous and spatiotemporal signalling from a tissue organizer orchestrates root vascular development. *Nat. Plants* 7, 1485–1494
44. Shahan, R. *et al.* (2022) A single-cell *Arabidopsis* root atlas reveals developmental trajectories in wild-type and cell identity mutants. *Dev. Cell* 57, 543–560.e9
45. Eldar, A. and Elowitz, M.B. (2010) Functional roles for noise in genetic circuits. *Nature* 467, 167–173
46. Ackermann, M. (2015) A functional perspective on phenotypic heterogeneity in microorganisms. *Nat. Rev. Microbiol.* 13, 497–508
47. Henrion, L. *et al.* (2023) Fitness cost associated with cell phenotypic switching drives population diversification dynamics and controllability. *Nat. Commun.* 14, 6128
48. Youssef, A. *et al.* (2026) Resource competition and growth dilution modulate synthetic gene cascade dynamics. *ACS Synth. Biol.* 15, 181–190
49. Tan, C. *et al.* (2009) Emergent bistability by a growth-modulating positive feedback circuit. *Nat. Chem. Biol.* 5, 842–848
50. Namboothiri, H.R. *et al.* (2026) Resolving emergent transient oscillations in gene circuits with a growth-coupled model. *Sci. Adv.* 12, eadz2310
51. Vandenbroucke, V. *et al.* (2024) Biological oscillations without genetic oscillator or external forcing. *bioRxiv* https://doi.org/10.1101/2024.09.20.614027
52. Thiermann, R. *et al.* (2026) Decoupling of global metabolic flux and proteome partitioning in bacteria. *Science* 392, eaeb6410
53. Kim, J. *et al.* (2025) Gene expression depends on the interplay among growth, resource biogenesis, and nutrient quality. *ACS Synth. Biol.* 14, 2012–2029
54. Vandenbroucke, V. *et al.* (2026) Synthetic niches enable coculture bioprocessing but are prone to mutational escape. *ACS Synth. Biol.* 15, 1888–1901
55. Zagieboylo, A.P. *et al.* (2025) A commensal bacterium secretes effector proteins to establish population heterogeneity in the gut. *Cell Host Microbe* 33, 2115–2129.e10
56. Askary, A. *et al.* (2025) The lives of cells, recorded. *Nat. Rev. Genet.* 26, 203–222
57. Schmidt, F. *et al.* (2022) Noninvasive assessment of gut function using transcriptional recording sentinel cells. *Science* 376, eabm6038
58. Oña, L. *et al.* (2025) Disentangling microbial interaction networks. *Trends Microbiol.* 33, 619–634
59. Mattingly, H.H. and Emonet, T. (2022) Collective behavior and nongenetic inheritance allow bacterial populations to adapt to changing environments. *Proc. Natl. Acad. Sci. U. S. A.* 119, e2117377119
60. Haas, P.A. *et al.* (2020) Subpopulations and stability in microbial communities. *Phys. Rev. Res.* 2, 022036
61. Luo, S. *et al.* (2025) On metrics for subpopulation detection in single-cell and spatial omics data. *Nucleic Acids Res.* 53, gkaf921
62. Wang, Y. *et al.* (2019) Subpopulation detection and their comparative analysis across single-cell experiments with scPopCorn. *Cell Syst.* 8, 506–513.e5
63. Ding, Q. *et al.* (2024) Dimension reduction, cell clustering, and cell–cell communication inference for single-cell transcriptomics with DcjComm. *Genome Biol.* 25, 241
64. De, A. and Chaudhuri, R. (2023) Common population codes produce extremely nonlinear neural manifolds. *Proc. Natl. Acad. Sci. U. S. A.* 120, e2305853120
65. Pacheco, A.R. and Segrè, D. (2019) A multidimensional perspective on microbial interactions. *FEMS Microbiol. Lett.* 366, fnz125
66. Goldford, J.E. *et al.* (2018) Emergent simplicity in microbial community assembly. *Science* 361, 469–474
67. Chen, Q. and Zernicka-Goetz, M. (2026) Decoding the origins of cellular self-organization for engineered biology. *Nat. Biotechnol.* 44, 1108–1120
68. Jiang, J. *et al.* (2026) Metabolic adaptation drives self-organization during skin organoid morphogenesis. *Nat. Commun.* 17, 5065
69. Saetzler, K. *et al.* (2011) Systems biology beyond networks: generating order from disorder through self-organization. *Semin. Cancer Biol.* 21, 165–174
70. Hari, K. *et al.* (2025) Low dimensionality of phenotypic space as an emergent property of coordinated teams in biological regulatory networks. *iScience* 28, 111730
71. Watts, D.J. and Strogatz, S.H. (1998) Collective dynamics of 'small-world' networks. *Nature* 393, 440–442
72. Bettencourt, L.M.A. *et al.* (2007) Invention in the city: increasing returns to patenting as a scaling function of metropolitan size. *Res. Policy* 36, 107–120
73. Kussell, E. *et al.* (2005) Bacterial persistence: a model of survival in changing environments. *Genetics* 169, 1807–1814
74. Duncker, K.E. *et al.* (2026) Engineering microbial consortia for distributed signal processing. *Nat. Commun.* 17, 8791
75. Martinez, J.A. *et al.* (2025) Automated adjustment of metabolic niches enables the control of natural and engineered microbial co-cultures. *Trends Biotechnol. Engl.* 43, 1116–1139
76. Delvigne, F. *et al.* (2017) Bioprocess scale-up/down as integrative enabling technology: from fluid mechanics to systems biology and beyond. *Microb. Biotechnol.* 10, 1267–1274
77. Wang, G. *et al.* (2020) Developing a computational framework to advance bioprocess scale-up. *Trends Biotechnol. Engl.* 38, 846–856
78. Trivellin, C. *et al.* (2022) Quantification of microbial robustness in yeast. *ACS Synth. Biol.* 11, 1686–1691
79. Delvenne, M. *et al.* (2026) Overriding bioprocess perturbations with a cell–machine interface for reliable microbial stress-response control. *Microb. Biotechnol.* 19, e70329
80. Zhuang, R.Z. *et al.* (2025) Integrating hierarchical vasculature into engineered cardiac tissues. *Adv. Mater. Ger.* 37, e09242
81. Ramesan, S.P. *et al.* (2026) Next-generation discovery: empowering organoid research with machine learning, artificial intelligence, and mathematical modeling. *Trends Biotechnol.* https://doi.org/10.1016/j.tibtech.2026.01.009