HHS Public Access

Author manuscript

*Curr Opin Syst Biol*. Author manuscript; available in PMC 2019 April 01.

Published in final edited form as:

*Curr Opin Syst Biol*. 2018 April ; 8: 46–50. doi:10.1016/j.coisb.2017.11.009.

# Robustness, Accuracy, and Cell State Heterogeneity in Biological Systems

**Dr Roy Wollman**

## Abstract

The robustness of biological systems is often depicted as a key system-level emergent property that allows uniform phenotypes in fluctuating environments. Yet, analysis of single-cell signaling responses identified multiple examples of cellular responses with high degrees of heterogeneity. Here we discuss the implications of the observed lack of response accuracy in the context of new observations coming from single-cell approaches. Single-cell approaches provide a new way to measure the abundance of thousands of molecular species in a single-cell. Repeatedly, analysis of cell distributions identifies clusters within these distributions where cells can be grouped into specific cell states. If cells in a population occupy distinct cell states, the observed variable response could in fact be accurate for each cell conditioned on its own internal state. In this view, the observed lack of accuracy, i.e. response heterogeneity, could in fact be beneficial and a potentially regulated feature of cell state variability. Therefore, to truly determine whether the observed response heterogeneity is a desired property or a physical limitation, future analysis of signaling heterogeneity must take into account the internal states of cells in the population.

## Introduction – Biology is “Messy Yet Beautiful”

At the turn of the 21st century, biology was undergoing a technological and conceptual revolution. The human genome project was just completed, and technologies such as microarray and later next generation sequencing, were starting to move from their initial noisy start to being a robust experimental platform. These advances caused scientists from other disciplines to start pay attention to biology, with the hope that tools and approaches from physics, math, engineering, and computer science will help “crack” the enormous puzzles in biology. Indeed, the turn of the century was accompanied by a few “manifests” that tried to define new approaches to study biology [1–3].

As scientists trained in the various quantitative disciplines moved to biology, one key difference stood out: Biology is “messy”. Experimental physicists are trained that data collection should always be accompanied with a model of the measurement error such that each number has associated “significant digits” that are reliable in their measurements. While in principle these concepts are true in biology as well, in practice experimental biology is different. It is not uncommon for biologists to use various anthropomorphisms,

such as “my cells are not happy today”. What is the right number of significant digits relevant to cell happiness?

The observation that biology is “messy” causes one to ask: how does it even work? If biological systems are hard to control and manipulate, how do they tolerate natural environmental fluctuation? As all chemical reaction rates are affected by temperature, effectively all kinetic parameters inside a cell will change as temperature fluctuates. How can cells function with such variability in their internal parameters? How does “messy biology” even work? The beauty of biology, is that it does. Many biological systems are able to function at a wide range of conditions, and trying to understand this perceived tension acted as a call for action that powered the inception of systems biology.

## The Robustness of Signaling Networks

In a seminal paper in 1997, Barkai and Leibler coined the term “biological robustness”, arguing that: “the key properties of biochemical networks should be robust in order to ensure their proper functioning” [4]. They demonstrated that bacterial chemotaxis is robust and this robustness is accomplished through simple adaptation. This initial work was followed by many others that demonstrated low parameter sensitivity in a large number of systems [5–11]. At the time these studies were coming out, the call for a new “system” approach was at its peak, and many authors of these opinion pieces used robustness as an example of a system level property [1,2,12]. In an opinion piece titled: “Biological Robustness”, Kitano wrote that: “[Robustness] is one of the fundamental and ubiquitously observed systems-level phenomena that cannot be understood by looking at the individual components.” [2].

Of course, robustness is not a completely new concept, and it was not invented in 1997. In engineering disciplines, “control theory” emerged from mechanical engineering and became its own interdisciplinary subject area that combines engineering, math, and computer science. Control theory studies how to design, i.e. engineer, a system to produce a desired output despite potential fluctuations in inputs and system parameters. Naturally, there is a great deal of correspondence between biological robustness and control theory, and the similarities and differences were studied in detail [13,14].

Since biological systems are evolved and not designed, it is important to make a distinction between the two key timescales of fluctuations that the system needs to be robust against. The relevant timescale for comparison is the lifetime of the system, and fluctuation could be either faster or slower than the lifetime of a cell. Gene expression variability of a key enzyme in two cells of the same population will result in variable maximum rates of reaction (Vmax), which can potentially change during the lifetime of a cell. In contrast, mutations in the catalytic site of the enzyme change a reaction’s Vmax on an evolutionary timescale, far exceeding the lifetime of a cell. These two timescales can, for the most part, be analyzed separately, despite their many natural connections. Here we won’t discuss the evolutionary timescale, instead focusing on the relationship between robustness and variability on fast time scales.

The premise that biological systems are robust and therefore must be insensitive to parameter variability became so accepted that many researchers started using robustness as a criterion to probe biological systems. Both in specific mechanistic models [5,15,16] and in simple abstract 3-node networks [17–19], researchers used the criteria of robustness as a model selection tool [19,20]. This approach argues that if many different models, e.g. different wiring diagrams of a signaling network, all produce the same output, then the most likely one to be true is the one that is the most robust. In practice, for a given model many random sets of parameters are tested; the model that got the function “right” with the most parameter sets “wins” and is most likely the correct one.

But are all biological systems “robust”? The strength of robust systems is that they are insensitive to change. This is not always a benefit. Robustness could also mean rigidity and lack of ability to tune the output of the system to a changing environment. If biological systems are robust, e.g. the response of cells to a specific ligand was robust to intracellular changes in protein concentration, one would expect little response variability. However, simple observation of mammalian cells show that this is not the case. Genetically identical cells of the same cell type still show variation [21]. Even sister cells that are almost identical at the point of division diverge quickly, showing variable responses to identical environments [22]. But is the observed variability substantial? How can one determine if the observed variability is high or low? To do so necessitates new and better tools that can quantify cellular response variability.

## Quantifying the Accuracy of Signaling Networks

A useful and successful approach to quantify robustness is to analyze a population of cells and measure their phenotypic distribution. A simplistic approach is to look at the overall index of dispersion, or coefficient of variation, of the underlying distribution. This direct approach was utilized in the analysis of single protein variability in cells. However, this statistical measure has some limitations: it neither takes into account nor interprets the shape of the underlying distribution. For example, if cells vary 10% in their response to a specific ligand, is that high or low? If the variability between different ligands is 5%, then one would say that a 10% variability is high. On the other hand, if the difference between response to different ligands is 50%, then a 10% variability within each sample, seems small. Therefore, to quantify robustness of cellular response to changing environment, just measuring overall dispersion is insufficient and a more sophisticated approach is needed.

Information theory allows one to quantify the amount of mutual information between two distributions [23]. Furthermore, one could calculate mutual information directly from samples without any assumptions on the underlying distributions. Therefore, mutual information is a powerful statistical tool and as such was utilized multiple times in neuroscience [24,25], developmental biology [26], genetic networks [27–29],and signal transduction [21,30–34]. For signal transduction, the function of the network is to transmit information about the extracellular environment to various cellular response machineries. Estimation of the maximal possible mutual information between ligand and cellular response distributions can inform on the accuracy by which cells respond to a specific ligand concentration [35,36]. In pioneering work, Cheong et al [31] used such an approach to

directly measure the performance of these networks as communication channels. Based on the assumption that all cells are from an independent and identically distributed sample, one could simply quantify the degree of overlap in response distributions between different ligand concentrations and use that to estimate the ability of a signaling network to transmit information about that ligand [30,37]. Due to its historical roots in computer science, the information transmission capacity is measured in bits. Cheong et al applied these measures to many signaling networks, concluding that they are poor at their core task with less than 1 bit of information transmission capacity. One interpretation of this observation is that the system can only transmit 2 states without any ambiguity. This direct interpretation might be simplistic in biological systems, and many more states could be transmitted if the system can accept low level of error [38]. Regardless, information transmission of 1 bit does not allow for much wiggle room. More importantly than the measurement itself, is the claim that this is the upper limit of cellular information transmission capability. Therefore, they concluded that for the core function of information transmission, biological systems are non-robust and that biochemical “noise” limits the accuracy of these networks.

Are signaling networks poor at their core task of information transmission? Follow-up work by Selimkhanov et al [21] further investigated this question. They showed that the existence of upper bounds depends on two assumptions: 1. That all cells are identical and hence any noise and variability in the system occurs stochastically during the process of signal transduction, and 2. That the cells only use a scalar for information transmission. If one quantifies cellular information transmission taking into account the multivariable dynamic profile of the signal, the empirical measurement of information transmission is almost double the previous estimates by Cheong et al. But perhaps more importantly, the reason for this increase is not from reduction of noise during the process of signal transmission itself. Rather, by transmitting a multivariate signal, cells can use some of that bandwidth to convey information about the cell state itself, reducing effective variability. The work by Selimkhanov et al effectively removed the upper bound on cell response accuracy for the case where the variability in cells is dominated by cell state variability. But is this really the case? Are genetically identical cells from a specific cell type homogenous or heterogeneous?

## Lack of Accuracy or Functional Variability?

Technological advances struck again. Improvements to high-throughput measurement tools, such as RNAseq, increased sensitivity sufficiently that they started to be used for demanding applications like single-cell measurements [39,40]. These systematic measurements began to show that there are specific patterns underneath the large degree of variability. Machine learning and dimensionality reduction techniques [41,42] constantly showed that for many different populations of identical cells, there is systematic variability that emerges from co-existence of specific cell states within the population.

The existence of specific patterns of single-cell gene expression suggests that kinetic parameters between cells will show specific differences. However, the relationship between kinetic parameters and gene expression is not easy to discern. In an attempt to understand the variability between cells in term of their kinetic parameters, Yao et al [43] performed parameter fitting to a differential equation model of cellular $Ca^{2+}$ response to ATP using

experimental dynamic single-cell data from hundreds of cells. They found that like gene expression patterns, kinetic parameters are separated into distinct subpopulations. In the NfkB network, two complementary approaches illuminate the relationship between cell state and cellular response. Sero et al [44] analyzed the relationship between cellular morphology and NfkB response across a panel of epithelial cell lines, finding strong correlation between morphology and signaling response. A more direct, but challenging, approach is to measure in the same single-cell both response dynamics and gene expression patterns; this was recently accomplished for NfkB signaling [45] and again supports the hypothesis that there are few key classes of cellular response.

What causes the existence of subpopulations in genetically identical cells from the same cell type? One hypothesis that needs further experimental support is that the specific patterns are a result of epigenetic regulation of gene expression, and that a hidden variable, such as a chromatin regulator, are different between these cell types. What is unclear is whether the chromatin regulators are simply variable between cells, and the decoding of this variability manifest itself as cell subpopulations? Or whether the chromatin regulators themselves are "part of the network" and the correct interpretation of each subpopulation is as a stable point in "cell state space" [46]. Answering this question will require a mechanistic understanding of the causal factors behind the different cell states, a formidable computational and experimental challenge.

The existence of distinct cellular states in the population does not necessarily mean that each of these cell states is functionally important. It is possible that these cell states are "minima" in some rugged epigenetic landscape, but that the states themselves are not functional units. Alternatively, it is likely that the distinct cell states are functionally important. Work in different systems has provided some indication that variability could have an important function in the form of bet-hedging and increasing response diversity. In addition, recent theoretical work showed that individual variation could increase population response accuracy in the case that the input itself is noisy [47]. Averaging sharp ultrasensitive responses results in a more graded curve and a reduced effect of noise in the levels on input. Therefore, cellular variability could potentially play an important functional role. If that is the case, it was likely evolved to serve that function and hence cellular variability itself should not be considered noise, but an important regulated property of biological systems.

## Outlook

The quest for identification of underlying principles is a driving force of scientific investigation. Initial work in systems biology focused on the question of robustness: how systems can become more uniform and respond in the same way despite existing biochemical variability. Both theoretical and experimental advances open a different view into the role of cellular variability. It suggests that variability is not simply something that has to be overcome, but instead an important functional property of the system. Does this means that we need to change how we think about robustness? Perhaps. A refined view of biological robustness needs to account for the fact that some time, it is the actual observed variability that has to be maintained. If indeed a specific distribution of cellular phenotypes in a population is optimal, how does the population of cells is maintained despite fluctuation

in the environment? Can a population of cells be robustly variable? Lastly, the existence of subpopulations does not mean that everything is variable. It is very possible that some systems properties need to be robust to variability in kinetic parameters across all cell subpopulations. Future work will be needed to further understand the interplay between population variability and response robustness.

## Acknowledgments

We thank Evan Maltz for critical reading of the manuscript. The work was supported by GM111404 and EY024960 from NIH.

## References

1. Kirschner MW. The meaning of systems biology. Cell. 2005; 121:503–504. [PubMed: 15907462]
2. Kitano H. Biological robustness. Nat. Rev. Genet. 2004; 5:826–837. [PubMed: 15520792]
3. Ideker T, Lauffenburger D. Building with a scaffold: emerging strategies for high- to low-level cellular modeling. Trends Biotechnol. 2003; 21:255–262. [PubMed: 12788545]
4. Barkai N, Leibler S. Robustness in simple biochemical networks. Nature. 1997; 387:913–917. [PubMed: 9202124]
5. Eldar A, Rosin D, Shilo B-Z, Barkai N. Self-enhanced ligand degradation underlies robustness of morphogen gradients. Dev. Cell. 2003; 5:635–646. [PubMed: 14536064]
6. von Dassow G, Meir E, Munro EM, Odell GM. The segment polarity network is a robust developmental module. Nature. 2000; 406:188–192. [PubMed: 10910359]
7. Sturm OE, Orton R, Grindlay J, Birtwistle M, Vyshemirsky V, Gilbert D, Calder M, Pitt A, Kholodenko B, Kolch W. The mammalian MAPK/ERK pathway exhibits properties of a negative feedback amplifier. Sci. Signal. 2010; 3:ra90. [PubMed: 21177493]
8. Tsai TY-C, Choi YS, Ma W, Pomerening JR, Tang C, Ferrell JE Jr. Robust, tunable biological oscillations from interlinked positive and negative feedback loops. Science. 2008; 321:126–129. [PubMed: 18599789]
9. Li F, Long T, Lu Y, Ouyang Q, Tang C. The yeast cell-cycle network is robustly designed. Proc. Natl. Acad. Sci. U. S. A. 2004; 101:4781–4786. [PubMed: 15037758]
10. Ma L, Iglesias PA. Quantifying robustness of biochemical network models. BMC Bioinformatics. 2002; 3:38. [PubMed: 12482327]
11. Bhalla US, Iyengar R. Robustness of the bistable behavior of a biological signaling feedback loop. Chaos. 2001; 11:221–226. [PubMed: 12779455]
12. Kitano H. Towards a theory of biological robustness. Mol. Syst. Biol. 2007; 3:137. [PubMed: 17882156]
13. Stelling J, Sauer U, Szallasi Z, Doyle FJ 3rd, Doyle J. Robustness of cellular functions. Cell. 2004; 118:675–685. [PubMed: 15369668]
14. Khammash M. An engineering viewpoint on biological robustness. BMC Biol. 2016; 14:22. [PubMed: 27007299]
15. Haney S, Bardwell L, Nie Q. Ultrasensitive responses and specificity in cell signaling. BMC Syst. Biol. 2010; 4:119. [PubMed: 20735856]
16. Yi TM, Huang Y, Simon MI, Doyle J. Robust perfect adaptation in bacterial chemotaxis through integral feedback control. Proc. Natl. Acad. Sci. U. S. A. 2000; 97:4649–4653. [PubMed: 10781070]
17. Ma W, Trusina A, El-Samad H, Lim WA, Tang C. Defining network topologies that can achieve biochemical adaptation. Cell. 2009; 138:760–773. [PubMed: 19703401]
18. Chau AH, Walter JM, Gerardin J, Tang C, Lim WA. Designing synthetic regulatory networks capable of self-organizing cell polarization. Cell. 2012; 151:320–332. [PubMed: 23039994]
19. Lim WA, Lee CM, Tang C. Design principles of regulatory networks: searching for the molecular algorithms of the cell. Mol. Cell. 2013; 49:202–212. [PubMed: 23352241]

20. Morohashi M, Winn AE, Borisuk MT, Bolouri H, Doyle J, Kitano H. Robustness as a measure of plausibility in models of biochemical networks. J. Theor. Biol. 2002; 216:19–30. [PubMed: 12076125]
21. Selimkhanov J, Taylor B, Yao J, Pilko A, Albeck J, Hoffmann A, Tsimring L, Wollman R. Accurate information transmission through dynamic biochemical signaling networks. Science. 2014; 346:1370–1373. ** Analysis of signaling response accuracy that took into account both the different sources of “noise” as well as the multivariate nature of cellular response. [PubMed: 25504722]
22. Spencer SL, Gaudet S, Albeck JG, Burke JM, Sorger PK. Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis. Nature. 2009; 459:428–432. * Sister-cell analysis of response variability. [PubMed: 19363473]
23. Tkačik G, Bialek W. Information Processing in Living Systems. Annu. Rev. Condens. Matter Phys. 2016; 7:89–117. * Comprehensive review on the interface between information theory and biology.
24. Schneidman E, Bialek W, Berry MJ 2nd. Synergy, redundancy, and independence in population codes. J. Neurosci. 2003; 23:11539–11553. [PubMed: 14684857]
25. Victor JD. Binless strategies for estimation of information from neural data. Phys. Rev. E Stat. Nonlin. Soft Matter Phys. 2002; 66:051903. [PubMed: 12513519]
26. Dubuis JO, Tkacik G, Wieschaus EF, Gregor T, Bialek W. Positional information, in bits. Proc. Natl. Acad. Sci. U. S. A. 2013; 110:16301–16308. [PubMed: 24089448]
27. Tkacik G, Callan CG Jr, Bialek W. Information flow and optimization in transcriptional regulation. Proc. Natl. Acad. Sci. U. S. A. 2008; 105:12265–12270. [PubMed: 18719112]
28. Walczak AM, Tkacik G, Bialek W. Optimizing information flow in small genetic networks. II. Feed-forward interactions. Phys. Rev. E Stat. Nonlin. Soft Matter Phys. 2010; 81:041905. [PubMed: 20481751]
29. Rieckh G, Tkačik G. Noise and information transmission in promoters with multiple internal States. Biophys. J. 2014; 106:1194–1204. [PubMed: 24606943]
30. Uda S, Saito TH, Kudo T, Kokaji T, Tsuchiya T, Kubota H, Komori Y, Ozaki Y-I, Kuroda S. Robustness and compensation of information transmission of signaling pathways. Science. 2013; 341:558–561. [PubMed: 23908238]
31. Cheong R, Rhee A, Wang CJ, Nemenman I, Levchenko A. Information transduction capacity of noisy biochemical signaling networks. Science. 2011; 334:354–358. ** Pioneering work that brought the tools of information theory into signal transduction reserach. [PubMed: 21921160]
32. Uda S, Kuroda S. Analysis of cellular signal transduction from an information theoretic approach. Semin. Cell Dev. Biol. 2016; 51:24–31. [PubMed: 26710655]
33. Mehta P, Goyal S, Long T, Bassler BL, Wingreen NS. Information processing and signal integration in bacterial quorum sensing. Mol. Syst. Biol. 2009; 5:325. [PubMed: 19920810]
34. Hansen AS, O’Shea EK. Limits on information transduction through amplitude and frequency regulation of transcription factor activity [Internet]. Elife. 2015; 4
35. Levchenko A, Nemenman I. Cellular noise and information transmission. Curr. Opin. Biotechnol. 2014; 28:156–164. [PubMed: 24922112]
36. Habibi I, Cheong R, Lipniacki T, Levchenko A, Emamian ES, Abdi A. Computation and measurement of cell decision making errors using single cell data. PLoS Comput. Biol. 2017; 13:e1005436. [PubMed: 28379950]
37. Brennan MD, Cheong R, Levchenko A. How Information Theory Handles Cell Signaling and Uncertainty. Science. 2012; 338:334. [PubMed: 23087235]
38. Bowsher CG, Swain PS. Environmental sensing, information transfer, and cellular decision-making. Curr. Opin. Biotechnol. 2014; 28:149–155. * Important discussion on the interpretation of mutual information statistics. [PubMed: 24846821]
39. Vera M, Biswas J, Senecal A, Singer RH, Park HY. Single-Cell and Single-Molecule Analysis of Gene Expression Regulation. Annu. Rev. Genet. 2016; 50:267–291. [PubMed: 27893965]
40. Wang Y, Navin NE. Advances and applications of single-cell sequencing technologies. Mol. Cell. 2015; 58:598–609. [PubMed: 26000845]

41. Setty M, Tadmor MD, Reich-Zeliger S, Angel O, Salame TM, Kathail P, Choi K, Bendall S, Friedman N, Pe'er D. Wishbone identifies bifurcating developmental trajectories from single-cell data. Nat. Biotechnol. 2016; 34:637–645. [PubMed: 27136076]

42. Amir E-AD, Davis KL, Tadmor MD, Simonds EF, Levine JH, Bendall SC, Shenfeld DK, Krishnaswamy S, Nolan GP, Pe'er D. viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia. Nat. Biotechnol. 2013; 31:545–552. [PubMed: 23685480]

43. Yao J, Pilko A, Wollman R. Distinct cellular states determine calcium signaling response. Mol. Syst. Biol. 2016; 12:894. [PubMed: 27979909]

44. Sero JE, Sailem HZ, Ardy RC, Almuttaqi H, Zhang T, Bakal C. Cell shape and the microenvironment regulate nuclear translocation of NF-κB in breast epithelial and tumor cells. Mol. Syst. Biol. 2015; 11:790. [PubMed: 26148352]

45. Lane K, Van Valen D, DeFelice MM, Macklin DN, Kudo T, Jaimovich A, Carr A, Meyer T, Pe'er D, Boutet SC, et al. Measuring Signaling and RNA-Seq in the Same Cell Links Gene Expression to Dynamic Patterns of NF-κB Activation. Cell Syst. 2017; 4:458–469.e5. [PubMed: 28396000]

46. Trapnell C. Defining cell types and states with single-cell genomics. Genome Res. 2015; 25:1491–1498. [PubMed: 26430159]

47. Suderman R, Bachman JA, Smith A, Sorger PK, Deeds EJ. Fundamental trade-offs between information flow in single cells and cellular populations. Proc. Natl. Acad. Sci. U. S. A. 2017; 114:5755–5760. [PubMed: 28500273]

- Robustness is an important emergent property of biological systems
- Many cells populations show high response heterogeneity
- Single-cell technologies uncover complex structure of cell state distribution
- Response heterogeneity is possibly result of internal cell state and not lack of accuracy