Insights about collective decision-making at the genetic level

Guillermo Rodrigo¹

Received: 30 July 2019 / Accepted: 5 December 2019 / Published online: 16 December 2019

© International Union for Pure and Applied Biophysics (IUPAB) and Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract

By living in a collective, individuals can share and aggregate information to base their decisions on the many rather than on the one, thereby increasing accuracy. But a collective can also be defined at the molecular level. In the following, we reason that genes, by working collectively, share fundamental features with social organisms, which ends, without invoking cognition, in wiser responses. For that, we compile into a single picture the terms redundancy, stochastic resonance, intrinsic and extrinsic noise, and cross-regulation.

Keywords Gene regulation · Genetic architecture · Heterogeneity · Information theory · Molecular noise · Redundancy · Systems biology

About collective decision-making

Increasing evidence supports the fact that several organisms, not only humans, take advantage of their organization in collectives to make more accurate decisions based on consensus (Conradt and Roper 2003; Couzin et al. 2005; Sumpter et al. 2008; Strandburg-Peshkin et al. 2015). This is because biological systems perceive the signals coming from their immediate environments to respond accordingly (i.e., make decisions) in an imperfect and heterogeneous manner, a perception scheme that applies from genes to complex organisms. We use the terms imperfect as the underlying sensor-actuator machinery is implemented through molecular reactions that are subject to unpredictable fluctuations (Samoilov et al. 2006; Balázsi et al. 2011; Tsimring 2014), and heterogeneous because the activation/repression thresholds implementing this machinery (i.e., molecular switches) can differ between individuals due to dynamic internal states (e.g., due to cell cycle; Colman-Lerner et al. 2005), genetic variation (Zheng et al. 2010; Kasowski et al. 2010), and/or previous interactions recorded epigenetically (D’Urso and Brickner 2014). This way, two individuals may respond differently to the same stimulus (e.g., the case of the

movement of baboons; Strandburg-Peshkin et al. 2015), which may cause erroneous decisions in certain circumstances (even the same individual can respond differently to the same stimulus from time to time, especially when it has no memory). A general strategy to circumvent this issue consists in balancing a set of individual responses with the aim of minimizing the error associated with signal perception/transduction. This insight has already been exploited as a design principle in engineering (Moore and Shannon 1956a, b) and constitutes a slight implementation of the law of large numbers (Hoffmann-Jorgensen and Pisier 1976; Bahr and Passerini 1998). Besides, it permeates different levels of biological organization (see, e.g., Dussutour et al. 2009 in the case of insects; and, e.g., Padmanabhan and Urban 2010 in the case of neurons) and has been discussed for organisms including microbes (Shapiro 1998; Queller and Strassmann 2009; Reid et al. 2015). We here do so at a lower level by considering genes as individuals. Yet, we do not refer to selfish genetic elements (e.g., transposons; Hurst and Werren 2001), but in principle to any gene whose information-carrying dynamic expression can be shared to benefit the cell.

Box 1 List of concepts

Cross-regulation: Orthogonal regulation that occurs at any point between two signal transduction pathways.

Emergence: Ability of a system to have functional properties that are not anticipated from its fundamental constituents.

Extrinsic noise: Molecular fluctuations that occur outside the expression of a given gene and affect global elements in the cell. These fluctuations can be common for multiple genes.

✉ Guillermo Rodrigo
guillermo.rodrigo@csic.es

¹ Institute for Integrative Systems Biology (I2SysBio), CSIC – U. Valencia, 46980 Paterna, Spain

Intrinsic noise: Molecular fluctuations that occur in the different reactions involved in the expression of a given gene. These fluctuations are in principle independent from gene to gene.

Holistic approach: Study of a complex system from a global perspective without division.

Mutual information: Calculation of the ability to reproduce the input signal distribution from the output response distribution of a system. It quantifies the amount of information transmitted.

Receiver operating characteristic curve: Representation of the discrimination ability of a nonlinear system as its own threshold varies. It shows the trade-off between sensitivity and specificity.

Reductionist approach: Study of a complex system by dividing it into fundamental subsystems.

Redundancy: Replication of certain components of a system in order to enhance its functionality, either to increase reliability or performance.

Resonance: Improved response of a nonlinear system as a consequence of a definite input signal frequency.

Shannon entropy: Calculation of the uncertainty that contains a given source of information.

Stochastic resonance: Improved response of a nonlinear system as a consequence of a definite amount of noise. It can be subthreshold or suprathreshold according to the input signal amplitude

Aggregation of information at the genetic level to increase response accuracy

Genetic collectivity as an evolved genetic architecture

Genetic systems are paradigmatic examples in which unexpected behaviors can emerge as a consequence of precise regulatory architectures (see, e.g., Hart and Alon 2013). Such behaviors can refer to a new function but also to a way to implement precision in a noisy biochemical scenario, that is, to efficiently read the different signal levels that eventually trigger the mounting of a response. We argue that some regulatory architectures have evolved to effectively mount the intended response when needed (maximizing the rate of true positives) while avoiding unnecessary mountings (minimizing the rate of false positives). This can be quantified through the use of receiver operating characteristic curves (Turner 1978). We further argue that one of such architectures is collectivity (understood as redundancy or degeneracy; Edelman and Gally 2001), which serves to generate characteristic curves with greater detectability.

Intrinsic noise drives valuable heterogeneity

In principle, the lack of complete information about a given problem, together with its partial loss during the internal processing, leads to inaccurate decisions by the individuals. This also holds in the case of genes, with repercussion for the cell/organism. A collective decision allows, on the one hand, integrating more environmental information, as each individual in the collective has its

intrinsic ability to perceive stimuli (i.e., it has its own thresholds). In addition to genetic variability, intrinsic molecular noise uncouples dynamic responses from gene to gene, even when they have a common *cis*-regulatory region (Elowitz et al. 2002). This kind of noise is preponderant in scenarios of low number of molecules (proteins; mainly through weak transcription rates), burst-like activation of expression (Suter et al. 2011), reduced number of multiplexed regulations, and many nutrients in the medium (Swain et al. 2002; Becskei et al. 2005). In this regard, a set of intrinsically noisy genes might allow an increase in representation ability when the environmental information is scarce (subthreshold signaling regime), noting that only some genes would reach appropriate responses, what in essence is a phenomenon of stochastic resonance (Collins et al. 1996; Stocks and Mannella 2001). For instance, noise in ComK, the protein that regulates competence in *Bacillus* for DNA uptake, is mainly intrinsic (Maamar et al. 2007). A poor transcription rate together with a positive autoregulation can explain this behavior (Tao 2004). Interestingly, high noise in ComK is translated into high transient competence variability at the population level, which leads to more transformations of *Bacilli* (Maamar et al. 2007; Çağatay et al. 2009). We can imagine a similar scenario, although at a much lower scale, in which the different genes are located in the same cell (Fig. 1, top).

Heterogeneously expressed genes for better signal processing

On the other hand, a collective decision allows reducing the internal loss of information, as the deviation of those perceptions with respect to the optimal one is buffered by the collective (now, in a suprathreshold signaling regime). Functionally redundant genes (or genetic pathways), in addition to generate robustness to mutations (de Visser et al. 2003), can act as parallel communication channels. If the outputs of these channels, in principle somewhat heterogeneous due to independent processing, were aggregated to mount a unique response, we could see those genes (or pathways) as a collective in charge of making a decision (Fig. 1, top). Response accuracy, which can be quantified as the input-output information transfer (Shannon 1948), increases because the balance of different threshold crossings (decisions) buffers fortuitous errors in gene expression (typically induced by molecular noise). This is a way to ensure that important intracellular checkpoints are reached when appropriate (Kafri et al. 2006; Cheong et al. 2011; Rodrigo and Poyatos 2016). In physical terms as before, this corresponds to the phenomenon of suprathreshold stochastic resonance (Stocks 2000). Experimentally, it has already been shown how gene duplication results in an enhanced ability to transmit information (Hansen and O'Shea

Fig. 1 Graphical representation of decision-making at different levels of biological organization. There, environmental information is processed, either at the organismal (behavior), cellular (phenotype), or genetic (expression) level, to produce an adaptive response. A collective with

heterogeneous information processing will make a more accurate decision by balancing different individual responses, provided that factors coordinating these responses, such as uncontrollable externality or cross-talk between individuals, are inconsequential

2015). Beyond information transfer and more in terms of the trade-off between sensitivity and specificity, decision accuracy can also be quantified if we analyze the probability of occurrence of false positive responses (i.e., false alarms; the system mounts a response that is not needed according to the actual signal level), as well as of false negative responses (i.e., misses; the system does not mount a response that is indeed required; Habibi et al. 2017). Interestingly, mammalian cells exploit two parallel signaling pathways (NF- κ B and JNK) to read the level of tumor necrosis factor (Oeckinghaus et al. 2011), noting that noise in the expression of the final transcription factors is both intrinsic and extrinsic (Tay et al. 2010). Thus, by combining information from both

pathways, cells can work with more precision (Cheong et al. 2011; Rhee et al. 2014).

Loss of genetic individuality to reduce the potential gain in response accuracy

Reduced gain due to extrinsic noise

There are, nevertheless, forces that can reduce the potential gain in accuracy by a shared decision. These forces can be external or internal to the collective and tend to homogenize the individual responses (Bahr and Passerini

1998). As a result, the required functional diversity to perform better as a collective than as a competent individual can be lost. Gene expression is sensitive to the intracellular state (proxied by the cell growth rate; Klumpp et al. 2009), in turn dependent on the extracellular environment. Irremediably, this dependence results in outward and erratic processes that commonly affect multiple genes. This is called extrinsic noise and tends to correlate gene expression profiles (Elowitz et al. 2002). More precisely, extrinsic noise integrates variability coming from the transcription-translation machinery and from different upstream molecular processes (other genes) that regulate expression (Volfson et al. 2006; Yang et al. 2014). We can argue that extrinsic noise is nothing different, at least in concept, from the action of climate change on animal behavior (Fig. 1, bottom), due to the change itself and also the influence in cascade exerted by other organisms in the ecosystem (Post et al. 1999). Moreover, even genes with significantly different thresholds can end with strongly correlated responses against multiple stimuli due to sequence-associated spatiotemporal effects within the cell beyond extrinsic noise (e.g., due to similar chromosomal positions, chromatin states, or *cis*-regulatory regions; Becskei et al. 2005; Karličić et al. 2010; Bussemaker et al. 2001). It has been already tested computationally that the resulting correlation in expression profiles reduces the gain in accuracy of a cellular response based on the aggregated expression of multiple genes exhibiting bistable or excitable responses (Rodrigo and Poyatos 2016).

Reduced gain due to cross-regulation

The homogenizing forces can also be internal to the genetic collective and depend on its particular structure, that is, how the genes are interconnected. Gene regulatory networks tend to be considered intricate due to the tinkering process occurred during evolution (Jacob 1977), despite several design principles can be recognized (Alon 2003). But interestingly, statistical mechanics analyses have revealed that some biological structures follow the principle of maximum entropy (i.e., maximal disorder), both in gene networks (Lezon et al. 2006), cell tissues (e.g., neuronal networks; Tkačik et al. 2014), and animal groups (e.g., bird flocks; Bialek et al. 2012). This means that the underlying structure (in those examples) is far from highly intricate and then that information is mostly processed in a parallel distributed manner. This scheme certainly favors the emergence of collective behaviors due to an increase in response accuracy. Moreover, even in absence of clearly defined communication channels, a decomposition into virtual subgroups with distinct information features and robust to noise can be produced (Prentice et al. 2016; based on data of neuronal activities in the retina). However, as long as the internal structure of the network (in the sense of collective) becomes more complex through the transcriptional or posttranscriptional cross-

regulation between genes (i.e., lower entropy), thereby leading to stronger correlations between individual responses (e.g., as it happens with *Hox* genes in vertebrates; Gould et al. 1997), the potential increase in accuracy through the collective response decreases. In particular, it has been quantified how information transfer declines with transcriptional cross-regulation within a set of independent genes that process a common signal in parallel (Rodrigo and Poyatos 2016). Hence, knowing that in human groups, for example, the performance in estimation tasks is reduced when some individuals, generally those presumably more informed, strongly influence over the autonomous decision ability of others (Lorenz et al. 2011; Madirolas and de Polavieja 2015), we can understand the genetic regulations occurring orthogonally to the main information flow in terms of social pressure (Fig. 1, bottom).

Concluding remarks

Patterns of collective behavior appear even at the genetic level, shaping the decisions made upon the aggregation of different single-gene responses. Accordingly, it would be expected that some heterogeneous gene collectivities remained under selection (Lewontin 1970). Genetic redundancies and/or degeneracies would be favored as suitable molecular architectures to increase decision accuracy if the resulting benefit exceeded the cost of maintaining multiple genes (in terms of resources; Lynch and Marinov 2015). For example, duplicated genes maintaining absolute expression levels (Qian et al. 2010) can increase accuracy by reducing intrinsic noise fluctuations, which can be translated into cellular fitness increase on average in scenarios of trade-off without introducing any additional cost (Rodrigo and Fares 2018). But, of course, other molecular architectures could be favored to increase decision accuracy, such as negative feedback loops (Yu et al. 2008; Habibi et al. 2017), in detriment of or in addition to collectivity. Besides, as complexity goes parallel to redundancy and degeneracy in biology (Edelman and Gally 2001), the ideas here presented reveal a functional link and suggest that complex biological functions, at least in part, are founded on accurate information transfer. Certainly, more theoretical and experimental work is required, especially on the light of social evolution (West et al. 2015; Scott and West 2019), to derive a compelling picture that supports the emergence and prevalence of consensus and cooperation at any level of biological organization.

Acknowledgements G.R. wishes to express deep gratitude to Drs. M.C. Baquero and L.M. Floria (Hospital La Fe, Valencia) for their kind attention.

Funding information This work was supported by grants BFU2015-66894-P (MINECO/FEDER) and GV/2016/079 (GVA).

References

- Alon U (2003) Biological networks: the tinkerer as an engineer. *Science* 301:1866–1867
- Bahr DB, Passerini E (1998) Statistical mechanics of opinion formation and collective behavior: micro-sociology. *J Math Sociol* 23:1–27
- Balázsi G, van Oudenaarden A, Collins JJ (2011) Cellular decision making and biological noise: from microbes to mammals. *Cell* 144:910–925
- Becskei A, Kaufmann BB, van Oudenaarden A (2005) Contributions of low molecule number and chromosomal positioning to stochastic gene expression. *Nat Genet* 37:937–944
- Bialek W, Cavagna A, Giardina I, Mora T, Silvestri E, Viale M, Walczak AM (2012) Statistical mechanics for natural flocks of birds. *Proc Natl Acad Sci U S A* 109:4786–4791
- Bussemaker HJ, Li H, Siggia ED (2001) Regulatory element detection using correlation with expression. *Nat Genet* 27:167–174
- Çağatay T, Turcotte M, Elowitz MB, Garcia-Ojalvo J, Süel GM (2009) Architecture-dependent noise discriminates functionally analogous differentiation circuits. *Cell* 139:512–522
- Cheong R, Rhee A, Wang CJ, Nemenman I, Levchenko A (2011) Information transduction capacity of noisy biochemical signaling networks. *Science* 334:354–358
- Collins JJ, Chow CC, Capela AC, Imhoff TT (1996) Aperiodic stochastic resonance. *Phys Rev E* 54:5575
- Colman-Lerner A, Gordon A, Serra E, Chin T, Resnekov O, Endy D, Pesce CG, Brent R (2005) Regulated cell-to-cell variation in a cell-fate decision system. *Nature* 437:699–706
- Conradt L, Roper TJ (2003) Group decision-making in animals. *Nature* 421:155–158
- Couzin ID, Krause J, Franks NR, Levin SA (2005) Effective leadership and decision-making in animal groups on the move. *Nature* 433:513–516
- D'Urso A, Brickner JH (2014) Mechanisms of epigenetic memory. *Trends Genet* 30:230–236
- de Visser JAGM, Hermisson J, Wagner GP, Meyers LA, Bagheri-Chaichian H et al (2003) Evolution and detection of genetic robustness. *Evolution* 57:1959–1972
- Dussutour A, Beekman M, Nicolis SC, Meyer B (2009) Noise improves collective decision-making by ants in dynamic environments. *Proc R Soc B* 276:4353–4361
- Edelman GM, Gally JA (2001) Degeneracy and complexity in biological systems. *Proc Natl Acad Sci U S A* 98:13763–13768
- Elowitz MB, Levine AJ, Siggia ED, Swain PS (2002) Stochastic gene expression in a single cell. *Science* 297:1183–1186
- Gould A, Morrison A, Sproat G, White RA, Krumlauf R (1997) Positive cross-regulation and enhancer sharing: two mechanisms for specifying overlapping Hox expression patterns. *Genes Dev* 11:900–913
- Habibi I, Cheong R, Lipniacki T, Levchenko A, Emamian ES, Abdi A (2017) Computation and measurement of cell decision making errors using single cell data. *PLoS Comput Biol* 13:e1005436
- Hansen AS, O'Shea EK (2015) Limits on information transduction through amplitude and frequency regulation of transcription factor activity. *eLife* 4:e06559
- Hart Y, Alon U (2013) The utility of paradoxical components in biological circuits. *Mol Cell* 49:213–221
- Hoffmann-Jorgensen J, Pisier G (1976) The law of large numbers and the central limit theorem in Banach spaces. *Ann Probab* 4:587–599
- Hurst GDD, Werren JH (2001) The role of selfish genetic elements in eukaryotic evolution. *Nat Rev Genet* 2:597–606
- Jacob F (1977) Evolution and tinkering. *Science* 196:1161–1166
- Kafri R, Levy M, Pilpel Y (2006) The regulatory utilization of genetic redundancy through responsive backup circuits. *Proc Natl Acad Sci U S A* 103:11653–11658
- Karlič R, Chung HR, Lasserre J, Vlahoviček K, Vingron M (2010) Histone modification levels are predictive for gene expression. *Proc Natl Acad Sci U S A* 107:2926–2931
- Kasowski M, Grubert F, Heffelfinger C, Hariharan M, Asabere A et al (2010) Variation in transcription factor binding among humans. *Science* 328:232–235
- Klumpp S, Zhang Z, Hwa T (2009) Growth rate-dependent global effects on gene expression in bacteria. *Cell* 139:1366–1375
- Lewontin RC (1970) The units of selection. *Annu Rev Ecol Syst* 1:1–18
- Lezon TR, Banavar JR, Cieplak M, Maritan A, Federoff NV (2006) Using the principle of entropy maximization to infer genetic interaction networks from gene expression patterns. *Proc Natl Acad Sci U S A* 103:19033–19038
- Lorenz J, Rauhut H, Schweitzer F, Helbing D (2011) How social influence can undermine the wisdom of crowd effect. *Proc Natl Acad Sci U S A* 108:9020–9025
- Lynch M, Marinov GK (2015) The bioenergetic costs of a gene. *Proc Natl Acad Sci U S A* 112:15690–15695
- Maamar H, Raj A, Dubnau D (2007) Noise in gene expression determines cell fate in *Bacillus subtilis*. *Science* 317:526–529
- Madirolas G, de Polavieja GG (2015) Improving collective estimations using resistance to social influence. *PLoS Comput Biol* 11:e1004594
- Moore EF, Shannon CE (1956a) Reliable circuits using less reliable relays I. *J Franklin I* 262:191–208
- Moore EF, Shannon CE (1956b) Reliable circuits using less reliable relays II. *J Franklin I* 262:281–297
- Oeckinghaus A, Hayden MS, Ghosh S (2011) Crosstalk in NF- κ B signaling pathways. *Nat Immunol* 12:695–708
- Padmanabhan K, Urban NN (2010) Intrinsic biophysical diversity decorrelates neuronal firing while increasing information content. *Nat Neurosci* 13:1276–1282
- Post E, Peterson RO, Stenseth NC, McLaren BE (1999) Ecosystem consequences of wolf behavioural response to climate. *Nature* 401:905–907
- Prentice JS, Marre O, Ioffe ML, Loback AR, Tkačik G, Berry MJ II (2016) Error-robust modes of the retinal population code. *PLoS Comput Biol* 12:e1005148
- Qian W, Liao BY, Chang AY, Zhang J (2010) Maintenance of duplicate genes and their functional redundancy by reduced expression. *Trends Genet* 26:425–430
- Queller DC, Strassmann JE (2009) Beyond society: the evolution of organismality. *Philos Trans R Soc B* 364:3143–3155
- Reid CR, Garnier S, Beekman M, Latty T (2015) Information integration and multiattribute decision making in non-neuronal organisms. *Anim Behav* 100:44–50
- Rhee A, Cheong R, Levchenko A (2014) Noise decomposition of intracellular biochemical signaling networks using nonequivalent reporters. *Proc Natl Acad Sci U S A* 111:17330–17335
- Rodrigo G, Fares MA (2018) Intrinsic adaptive value and early fate of gene duplication revealed by a bottom-up approach. *eLife* 7:e29739
- Rodrigo G, Poyatos JF (2016) Genetic redundancies enhance information transfer in noisy regulatory circuits. *PLoS Comput Biol* 12:e1005156
- Samoilov MS, Price G, Arkin AP (2006) From fluctuations to phenotypes: the physiology of noise. *Sci STKE* 2006:re17
- Scott TW, West SA (2019) Adaptation is maintained by the parliament of genes. *Nat Commun* 10:5163
- Shannon CE (1948) A mathematical theory of communication. *Bell Syst Tech J* 27:379–423
- Shapiro JA (1998) Thinking about bacterial populations as multicellular organisms. *Annu Rev Microbiol* 52:81–104
- Stocks NG (2000) Suprathreshold stochastic resonance in multilevel threshold systems. *Phys Rev Lett* 84:2310
- Stocks NG, Mannella R (2001) Generic noise-enhanced coding in neuronal arrays. *Phys Rev E* 64:030902

- Strandburg-Peshkin A, Farine DR, Couzin ID, Crofoot MC (2015) Shared decision-making drives collective movement in wild baboons. *Science* 348:1358–1361
- Sumpter DJ, Krause J, James R, Couzin ID, Ward AJ (2008) Consensus decision making by fish. *Curr Biol* 18:1773–1777
- Suter DM, Molina N, Gattfield D, Schneider K, Schibler U, Naef F (2011) Mammalian genes are transcribed with widely different bursting kinetics. *Science* 332:472–474
- Swain PS, Elowitz MB, Siggia ED (2002) Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc Natl Acad Sci U S A* 99:12795–12800
- Tao Y (2004) Intrinsic and external noise in an auto-regulatory genetic network. *J Theor Biol* 229:147–156
- Tay S, Hughey JJ, Lee TK, Lipniacki T, Quake SR, Covert MW (2010) Single-cell NF- κ B dynamics reveal digital activation and analogue information processing. *Nature* 466:267–271
- Tkačik G, Marre O, Amodei D, Schneidman E, Bialek W, Berry MJ II (2014) Searching for collective behavior in a large network of sensory neurons. *PLoS Comput Biol* 10:e1003408
- Tsimring LS (2014) Noise in biology. *Rep Prog Phys* 77:–026601
- Turner D (1978) An intuitive approach to receiver operating characteristic curve analysis. *J Nucl Med* 19:213–220
- Volfson D, Marciniak J, Blake WJ, Ostroff N, Tsimring LS, Hasty J (2006) Origins of extrinsic variability in eukaryotic gene expression. *Nature* 439:861–864
- West SA, Fisher RM, Gardner A, Kiers ET (2015) Major evolutionary transitions in individuality. *Proc Natl Acad Sci U S A* 112:10112–10119
- Yang S, Kim S, Lim YR, Kim C, An HJ, Kim JH, Sung J, Lee NK (2014) Contribution of RNA polymerase concentration variation to protein expression noise. *Nat Commun* 5:4761
- Yu RY, Pesce CG, Colman-Lerner A, Lok L, Pincus D, Serra E, Holl M, Benjamin K, Gordon A, Brent R (2008) Negative feedback that improves information transmission in yeast signalling. *Nature* 456:755–761
- Zheng W, Zhao H, Mancera E, Steinmetz LM, Snyder M (2010) Genetic analysis of variation in transcription factor binding in yeast. *Nature* 464:1187–1191

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.