Cells find their identity using a mathematically optimal strategy

January 31 2019

Artistic depiction of gene regulatory machinery on the DNA that reads out gap gene concentration signals optimally to establish the precise seven-stripped patterns of pair rule gene expression in the fruit fly embryo. Credit: Mariela

Petkova

Organisms are made of many types of cells arranged in a precise and reproducible spatial pattern that gives rise to properly formed and well-functioning tissues and organs. But how do genetically identical cells in an organism become differentiated? A team of researchers, including Gašper Tkačik at the Institute of Science and Technology Austria (IST Austria), has now shown that in the developing fruit fly, expression levels of four genes, the so-called gap genes, can be jointly decoded into an optimal specification of position. This is the result of a study published today in *Cell*, by senior author Thomas Gregor, Eric Wieschaus, William Bialek, Mariela Petkova and Gašper Tkačik.

In the fruit fly embryo, the four gap genes are switched on in various domains along the long axis of a cigar-shaped embryo, forming an intricate spatial pattern. Individual cells in the embryo don't have a "global view" of their location within the embryo. Scientists have therefore hypothesized that a cell measures the concentrations of the gap genes, uses these concentrations as a chemical global positioning system to determine where it is, and makes a decision to become a certain cell type.

This decades-old paradigm, however, doesn't explain all observations, says Tkačik: "Several nagging questions persist, and most researchers so far have taken a mechanistic look at the molecules involved, how positions are read out, and what chemical reactions follow."

In the current study, however, the team took a different approach. "The signals that the cells receive about their position are noisy—the levels of gap gene expression fluctuate in time and between embryos. Having just four noisy concentration signals at which a cell looks only once limits the

precision with which this cell can compute its location, no matter what mechanisms or computations are involved. This is a fundamental physical limit. Given this noise in gap gene concentrations, we ask, how precisely could anyone—including the cell in the embryo—tell where they are?"

From a theoretical point of view, the approach to be taken is clear, Tkačik explains: "Theory tells us to use optimal decoding, an established statistical inference approach." To do so, the researchers measured gap gene expression levels with sufficient accuracy to characterize the biological noise in the system. Based on how the gap genes are switched on in wild-type embryos, they constructed an optimal decoder. To test the decoder, they asked what the decoder predicts would happen to patterning if any of the gap genes were genetically perturbed, and compared this prediction to what such mutant embryos really look like. The decoder correctly predicted how patterning is distorted in mutant embryos, with 1% accuracy and no free parameters which would have to be estimated from experiments. "This result is surprising. Without needing to know the mechanism of how cells determine their position, just assuming that this is done optimally, and simultaneously, using the absolute concentrations of all four gap genes, we can predict how positioning changes in mutants," Tkačik explains.

These results imply that in Drosophila embryos, the establishment of cell identities is close to optimal. "Evolution has been pushing this system strongly enough so that—whatever happens molecularly—the molecular hardware closely approximates the mathematically optimal computation of position," says Tkačik, "And while we didn't investigate which mechanism is at work, our results show that we don't understand the mechanism well enough. Models that fit data according to a presumed mechanism need many more parameters yet perform worse than simply assuming optimality with no free parameters to adjust, as we did here."

The results also put into question the textbook model of how positional information is conveyed in the Drosophila embryo. Eric Wieschaus and Christiane Nüsslein-Vollhard identified the genes required for patterning, a discovery for which they received the Nobel Prize in 1995. According to the prevailing view, very early signals from the mother get refined slowly across several layers of a patterning network. "Our results question this classic idea of a cascade that progressively refines noisy signals. Already in the earliest step of the cascade, at the level of gap genes, there is enough information to position all cells precisely," explains Tkačik.

Why the cascade exists remains an open question, but the additional layers could help cells process the available information, says Tkačik. "Just because the information is already present in the gap genes doesn't mean that this information is easy to read out for the fly. Maybe the additional layers transform the information into a format that is easier for cells to access and thereby commit reliably to a well-defined cell type."

Provided by Institute of Science and Technology Austria