SnapShot: Key Numbers in Biology

Uri Moran,[1] Rob Phillips,[2] and Ron Milo[1]
[1]Weizmann Institute of Science, Rehovot, Israel; [2]California Institute of Technology, Pasadena, CA, USA

Cell

Cell size

Bacteria *(E. coli):* ≈0.7-1.4 μm diameter, ≈2-4 μm length, ≈0.5-5 μm^3 in volume; 10^8-10^9 cell/ml for culture with OD_{600}≈1

Yeast *(S. cerevisiae):* ≈3-6 μm diameter ≈20-160 μm^3 in volume

Mammalian cell volume: 100-10,000 μm^3; HeLa cell: 500-5000 μm^3 (adhering to slide ≈15-30 μm diameter)

Length scales inside cells

Nucleus volume: ≈10% of cell volume

Cell membrane thickness: ≈4-10 nm

"Average" protein diameter: ≈3-6 nm

Base pair: 2 nm (D) x 0.34 nm (H)

Water molecule diameter: ≈0.3 nm

Energetics

Membrane potential ≈70-200 mV → 2-6 k_BT per electron ($k_BT \equiv$ thermal energy)

Free energy (ΔG) of ATP hydrolysis under physiological conditions
≈40-60 kJ/mol → ≈20 k_BT/molecule ATP; ATP molecules required during an *E. coli* cell cycle ≈$10-50 \times 10^9$

ΔG^0 resulting in order of magnitude ratio between product and reactant concentrations:
≈6 kJ/mol ≈60 meV ≈2 k_BT

Concentration

Concentration of 1 nM:
in *E. coli* ≈1 molecule/cell;
in HeLa cells ≈1000 molecules/cell

Characteristic concentration for a signaling protein: ≈10 nM-1 μM

Water content: ≈70% by mass; general elemental composition (dry weight) of *E. coli:* ≈$C_4H_7O_2N_1$; Yeast: ≈$C_6H_{10}O_3N_1$

Composition of *E. coli* (dry weight):
≈55% protein, 20% RNA, 10% lipid, 15% other

Protein concentration: ≈100 mg/ml = 3 mM. 10^6-10^7 per *E. coli* (depending on growth rate); Total metabolites (MW < 1 kDa) ≈300 mM

Division, replication, transcription, translation, and degradation rates

at 37°C with a temperature dependence (Q10) of ≈2-3

Cell cycle time (exponential growth in rich media): *E. coli* ≈20-40 min; budding yeast 70-140 min; HeLa human cell line: 15-30 hr

Rate of replication by DNA polymerase:
E. coli ≈200-1000 bases/s; human ≈40 bases/s. Transcription by RNA polymerase 10-100 bases/s

Translation rate by ribosome: 10-20 aa/s

Degradation rates (proliferating cells):
mRNA half life < cell cycle time; protein half life ≳ cell cycle time

Diffusion and catalysis rate

Diffusion coefficient for an "average" protein: in cytoplasm D≈5-15 $\mu m^2/s$ → ≈10 ms to traverse an *E. coli* → ≈10 s to traverse a mammalian HeLa cell; small metabolite in water D≈500 $\mu m^2/s$

Diffusion-limited on-rate for a protein:
≈10^8-10^9 $s^{-1}M^{-1}$ → for a protein substrate of concentration ≈1 μM the diffusion-limited on-rate is ≈100-1000 s^{-1} thus limiting the catalytic rate k_{cat}

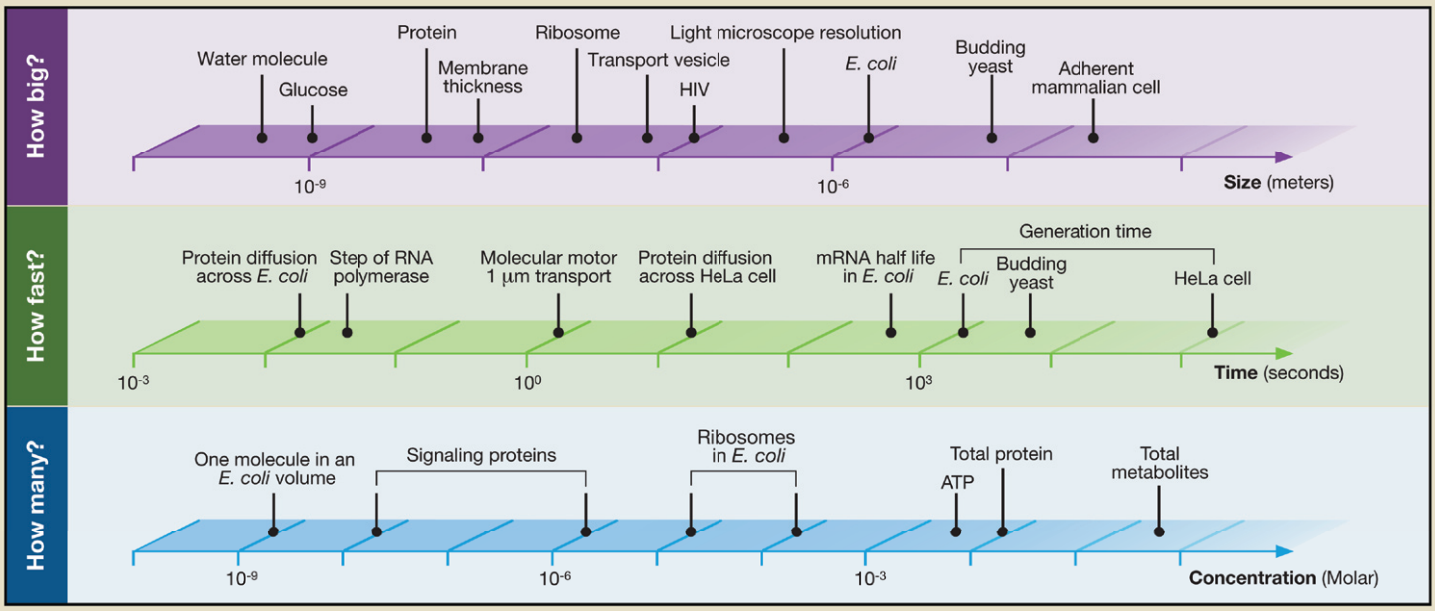
Genome sizes and error rates

Genome size:
E. coli (enterobacteria) ≈5 Mbp
S. cerevisiae (budding yeast) ≈12 Mbp
C. elegans (nematode) ≈100 Mbp
D. melanogaster (fruit fly) ≈120 Mbp
A. thaliana (plant) ≈120 Mbp
M. musculus (mouse) ≈2.5 Gbp
H. sapiens (human) ≈2.9 Gbp
T. aestivum (wheat) ≈16 Gbp

Number of protein-coding genes:
E. coli ≈ 4000; *S. cerevisiae* ≈ 6000; *C. elegans, A. thaliana, M. musculus, H. sapiens* ≈ 20,000

Mutation rate in DNA replication:
≈10^{-8}-10^{-10} per bp

Misincorporation rate:
transcription ≈10^{-4}-10^{-5} per nucleotide
translation ≈10^{-3}-10^{-4} per amino acid



Useful biological numbers extracted from the literature. Numbers and ranges should only serve as "rule of thumb" values. References are in the online annotated version at www.BioNumbers.org. See the website and original references to learn about the details of the system under study including growth conditions, method of measurement, etc.

See online version for legend and references.

SnapShot: Key Numbers in Biology

Uri Moran,[1] Rob Phillips,[2] and Ron Milo[1]
[1]Weizmann Institute of Science, Rehovot, Israel; [2]California Institute of Technology, Pasadena, CA, USA

Cell

Biology is becoming increasingly quantitative. Taking stock of key numbers in cell and molecular biology enables back-of-the-envelope calculations that test and sharpen our understanding of cellular processes. Further, such calculations provide a quantitative context for the torrent of data from new experimental techniques. However, such useful numbers are scattered in the vast biological literature in a way that often leads to a frustrating literature-mining ordeal. Here, we have collected a set of basic numbers in biology that we find extremely useful for obtaining an order of magnitude feel for the molecular processes in cells. Several examples (see below) show how to combine these numbers to think about biological questions. The values should be considered rules of thumb rather than definitive values as variety is the spice of life and variability is ever present in biology. This compilation is based on the BioNumbers wiki project (http://www.BioNumbers.org) where these and the values of several thousand other biological properties are provided together with their experimental context and references to the primary literature.

Is There Enough Time to Replicate the Genome?

The bacterium *Escherichia coli* has a genome of roughly 5 million base pairs (bp) and a replication rate in the range of 200–1000 bp/s. These numbers imply that it should take the two replisomes at least 2500 s to replicate the genome, a number that is much larger than the maximal division rate of ~20 min. How can this be? It turns out that, under ideal conditions, *E. coli* uses nested replication forks that begin to replicate the DNA for the granddaughter and great granddaughter cells before the daughter cells have even completed replication.

How Many Mutations in a 5 ml Culture of Bacteria?

Using the 10^{-9}/bp mutation rate of *E. coli* per replication and a genome size of $\sim 10^7$ (both strands), we predict $\sim 10^{-2}$ mutations per genome replication. In a 5 ml saturated culture (optical density ~2.0) of *E. coli*, there are about 10^9 to 10^{10} cells. The final doubling of this culture requires the replication of $\sim 10^9$ cells, thus even this last cell division event would be responsible for $\sim 10^7$ single base pair substitutions. If the culture started with a single bacterium, every single nonlethal base pair substitution in the *E. coli* genome is likely to be represented in the culture.

How Long to Reach Confluence?

In a 96 multiwell plate, each well has a diameter of 5 mm (i.e., an area of $\approx 20\ mm^2 = 2 \times 10^7\ \mu m^2$). Given that the diameter of a HeLa cell is ≈25 μm (i.e., $\approx 500\ \mu m^2$ area), it takes roughly 40,000 cells to reach confluence. Starting with a single cell (obtained by cell sorting rather than cell splitting) with a generation time of about 1 day, the time to reach confluence is about 2 weeks.

How "Dense" Is a Saturated *E. coli* Culture?

A saturated *E. coli* culture has about 10^9 cells/ml. Given that each cell is about 10^{-12} grams, we get a cell concentration of about 1 mg/ml or about 1 part in a thousand of the mass (or volume). The mean spacing between the cells is roughly 10 μm (which is not as dense as the concentration of bacteria in the gut of the termite where densities are typically a factor of ten higher).

How Many Carbon Atoms Are in a Cell?

A cell with a volume of 1 μm^3 and a density of about 1 g/ml has a total mass of 10^{-12} grams. From the formula $C_4H_7O_2N_1$ and the weights of the elements, we derive a carbon content of about $12 \times 4/(12 \times 4 + 7 + 2 \times 16 + 14) = 48/101$ or about one half of the dry mass. With 30% dry mass (70% water), we obtain $\sim 10^{-13}$ gm of carbon. Next we transformed the number of molecules using Avogadro's constant: $6 \times 10^{23} \times 10^{-13}/12 = 5 \times 10^9$ carbon atoms per cell. To verify this, we have done the calculation in a different way: assuming there are about 3×10^6 proteins, each one consisting of about 300 amino acids, we get a total of $\sim 10^9$ amino acids. An amino acid has about five carbon atoms, so we arrive at a similar value. Both estimates depend linearly on the cell volume, which can vary significantly based on growth conditions.

How Far Can Macromolecules Move by Diffusion?

It takes about 10 s on average for a protein to traverse a HeLa cell. An axon 1 mm long is about 100 times longer than a HeLa cell, and as the diffusion time scales as the square of the distance it would take 10^5 seconds or ~2 days for a molecule to travel this distance by diffusion. This demonstrates the necessity of mechanisms other than diffusion for moving molecules long distances. A molecular motor moving at a rate of ~1 μm/s will take a "reasonable" time (~15 min) to traverse an axon 1 mm in length.

ACKNOWLEDGMENTS

We thank Uri Alon, Niv Antonovsky, Danny Ben-Zvi, David Cahen, Erez Dekel, Idan Efroni, Avigdor Eldar, Yuval Eshed, Nir Friedman, Hernan Garcia, Paul Jorgensen, Michal Kenan-Eichler, Jane Kondev, Marc Kirschner, Daniel Koster, Avi Levy, Michal Lieberman, Debbie Marks, Elliot Meyerowitz, Elad Noor, Chris Sander, Dave Savage, Moselio Schaechter, Maya Schuldiner, Eran Segal, Daniel Segre, Benny Shilo, Guy Shinar, Alex Sigal, Rotem Sorek, Mike Springer, Bodo Stern, Arbel Tadmor, Julie Theriot, Rebecca Ward, Detlef Weigel, Jon Widom, and Tali Wiesel for help in preparing this SnapShot.

REFERENCES

Alon, U. (2006). An Introduction to Systems Biology: Design Principles of Biological Circuits (New York: Chapman & Hall/CRC).

Altman, P.L., and Katz, D.S.D. (1978). Biology Data Book (New York: John Wiley & Sons Inc.).

Burton, R.F. (2000). Physiology by Numbers (Cambridge, UK: Cambridge University Press).

Goldreich, P., Mahajan, S., and Phinney, S. (1999). Order-of-magnitude physics: Understanding the world with dimensional analysis, educated guesswork, and white lies (http://www.inference.phy.cam.ac.uk/sanjoy/oom/book-a4.pdf).

Harte, J. (1988). Consider a Spherical Cow: A Course in Environmental Problem Solving (Mill Valley, CA: University Science Books).

Milo, R., Jorgensen, P., Moran, U., Weber, G., and Springer, M. (2010). BioNumbers – the database of key numbers in molecular and cell biology. Nucleic Acids Res. *38*, D750–D753.

Neidhardt, F.C., Ingraham, J.L., and Schaechter, M. (1990). Physiology of the Bacterial Cell: A Molecular Approach (Sunderland, MA: Sinauer Associates).

Phillips, R., and Milo, R. (2009). A feeling for the numbers in biology. Proc. Natl. Acad. Sci. USA *106*, 21465–21471.

Phillips, R., Kondev, J., and Theriot, J. (2008). Physical Biology of the Cell (London: Garland Science).

Weinstein, L., and Adam, J.A. (2008). Guesstimation: Solving the World's Problems on the Back of a Cocktail Napkin (Princeton, NJ: Princeton University Press).

 ©2010 Elsevier Inc. DOI 10.1016/j.cell.2010.06.019