
DO YOU HAVE AN IDEA FOR YOUR PROJECT YET?
NO, I'M WAITING FOR INSPIRATION.
YOU CAN'T JUST TURN ON CREATIVITY LIKE A FAUCET. YOU HAVE TO BE IN THE RIGHT MOOD.
WHAT MOOD IS THAT?
LAST-MINUTE PANIC.

Week 8, Diffusion and Self Assembly

1/ Assembly instructions: Distributed pattern storage and self assembly

2/ Biology is 'small' - thermal energy matters! Behold - diffusion!

3/ Diffusion and self-assembly are paired at the hip

The fundamental mystery of living matter

If you mix proteins, DNA, and water, you get ....

The fundamental mystery of living matter

Ok, fine, if you mix water, DNA, proteins, and some fat, you get

But, that's nowhere near

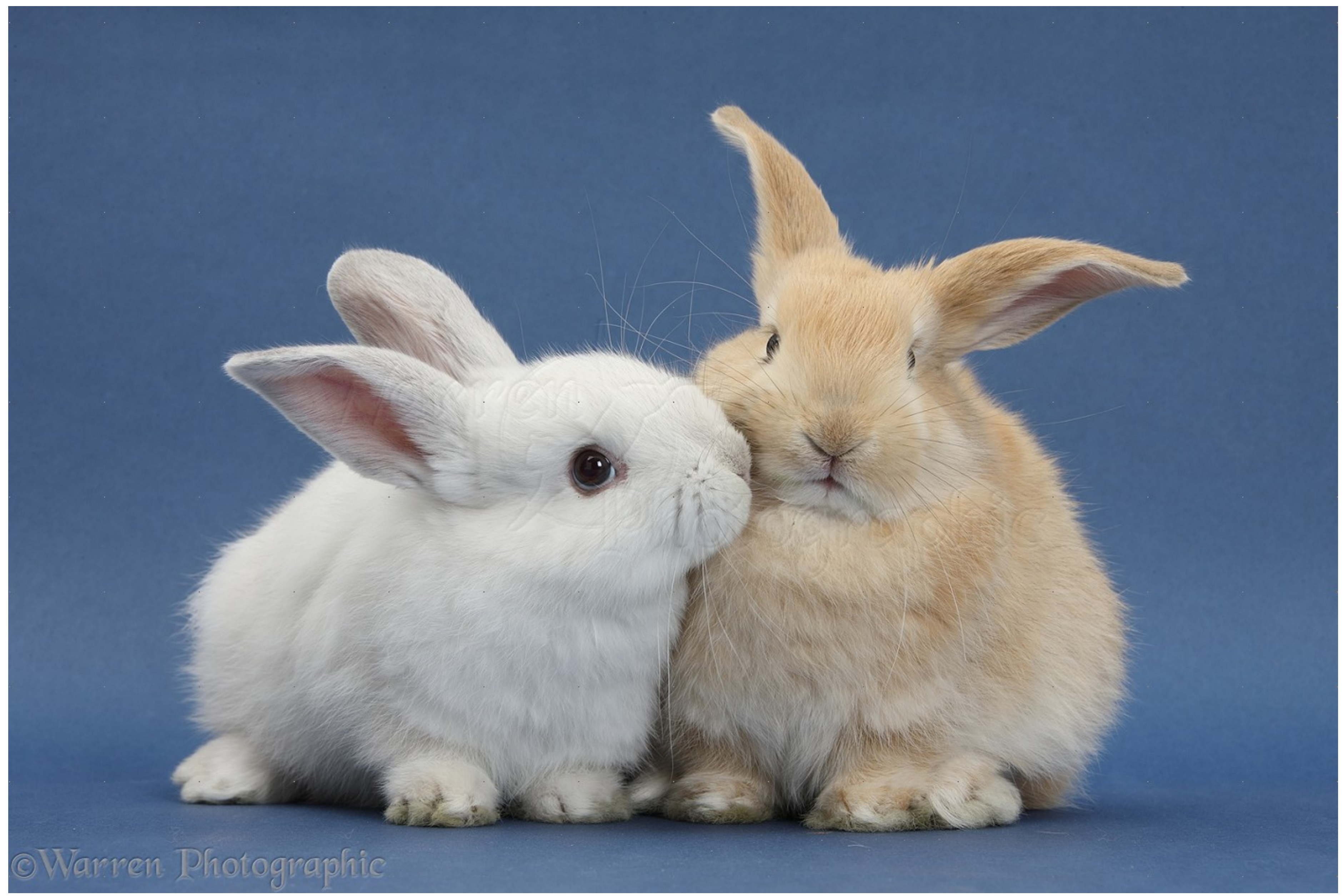
If we look around, living matter has myriad defined shapes, colors, mechanical characteristics...

How does that work?

"It's encoded in the DNA"

True but not insightful

How does one reliably pattern large living things from lots of teeny tiny little components?



How does one reliably pattern large living things from lots of teeny tiny little components?

You need to transduce a tiny molecular change (e.g. a single DNA mutation) into a larger macroscopic change,

and

the matter needs to assemble itself.

Today, we will look at <u>one</u> way to do that

There is really good news

You need to transduce a tiny molecular change (e.g. a single DNA mutation) into a larger macroscopic change,

and

the matter needs to assemble itself.

Both those things are closely connected…

#1

1/ Ways of specifying assembly instructions and assembling things

Idle Daydreams

It would be really really nice if the genome looked something like this:

```
void Initialize_Zebra( void )
{
   zebra_length          = 2.80; /*meters*/
   zebra_stripe_period   = 0.18; /*meters*/
   zebra_stripe_color    = _T("Black");
}
```

or like this:

```
….GCAATAAAzebralengthEQ2.80AAACGTACCA…
```

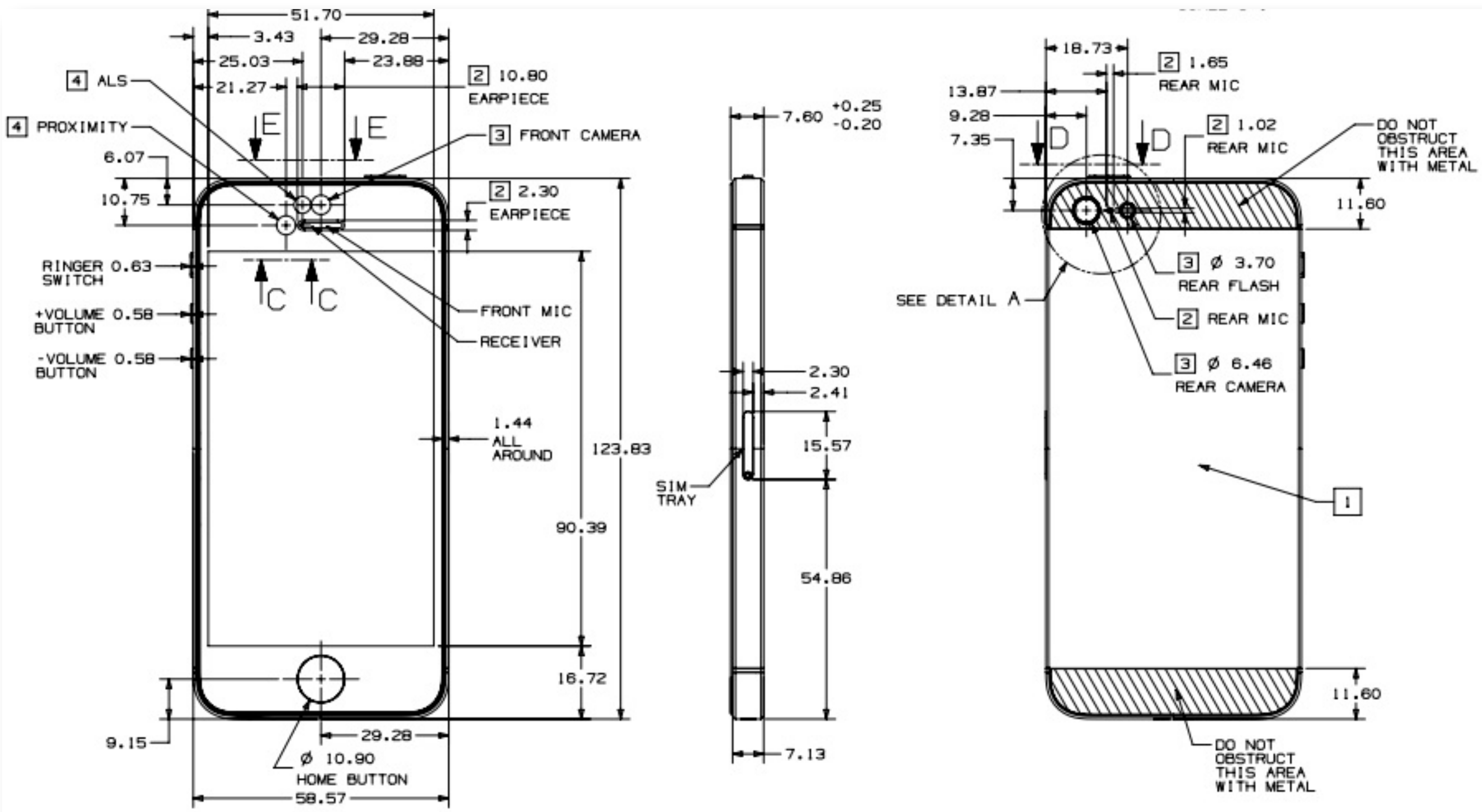
because then you could change stuff in seconds.

Sadly, no such luck.

Explicit and Concentrated Specification

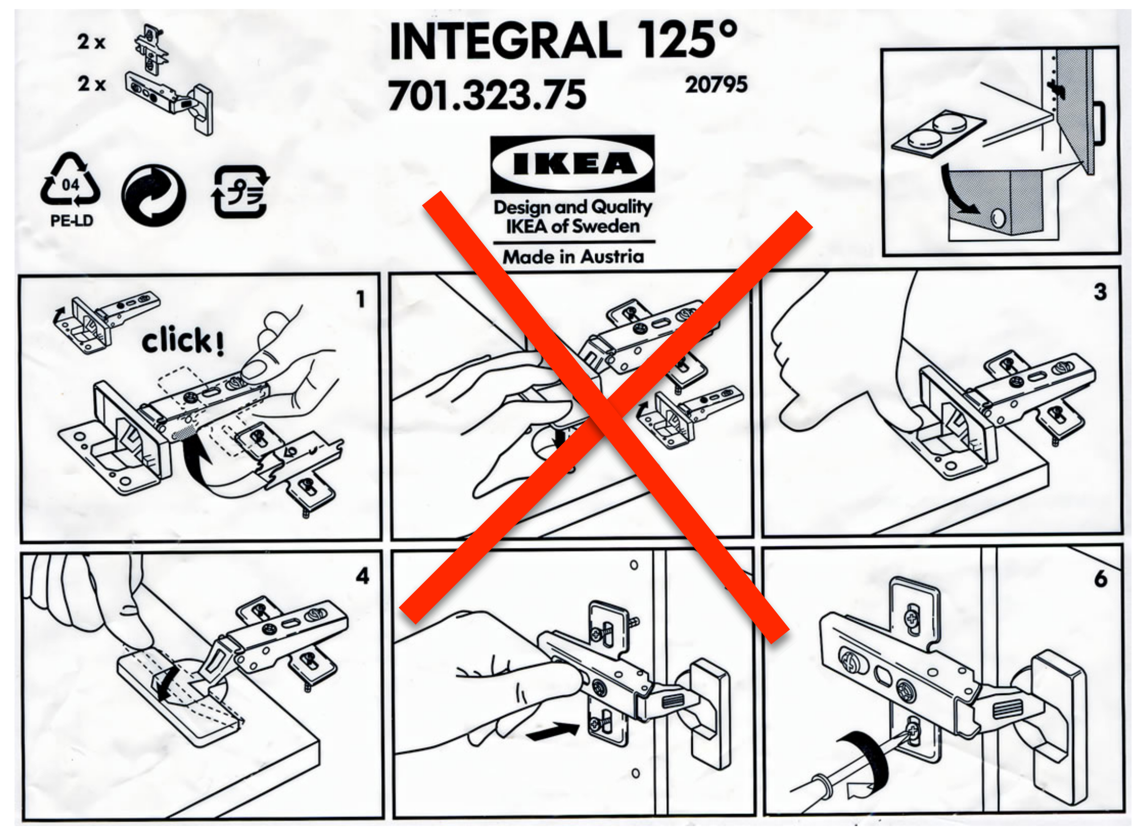
```
void Initialize_Zebra( void )
{
    zebra_length          = 2.80; /*meters*/
    zebra_stripe_period  = 0.18; /*meters*/
    zebra_stripe_color.Format(_T("Black"));
}
```

= Very hackable!



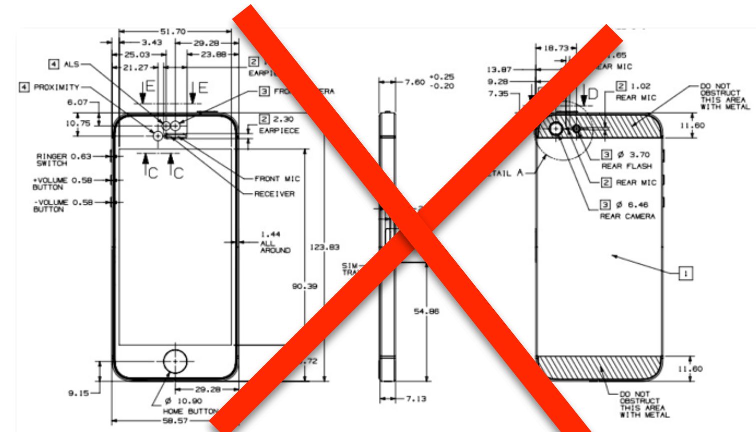
= Very hackable!

Forget macroscopic construction sites, orchestra conductors reading a score, etc.



Explicit and Concentrated Specification

```
void InitializeZebra( void )
{
   zebra_length        = 2.80; /*meters*/
   zebra_stripe_period = 0.18; /*meters*/
   zebra_stripe_color Format(_T("Black"));
}
```



Implicit and Distributed

Hard to hack

Key question:

For a given pattern, where and how is it stored?

Implicit and Distributed Storage of Information

Now that you have been deprogrammed:

A better way….

Self Assembly - what's your favorite mechanism?

Let's take a minute -
please discuss with your
neighbors

Forget:

macroscopic construction sites,
orchestra conductors reading a score, **hands**, rakes, molds…

Self Assembly - what's your favorite mechanism?

This is not so easy,
correct?

Forget:

macroscopic construction sites,
orchestra conductors reading a score, **hands**, rakes, molds…

A “magic” material - can you guess what it is?

Five clues:

Self-organizing

Self-repairing

Responsive to environment

Different “species” live in different places

Implicit and distributed encoding of pattern

Self-organizing

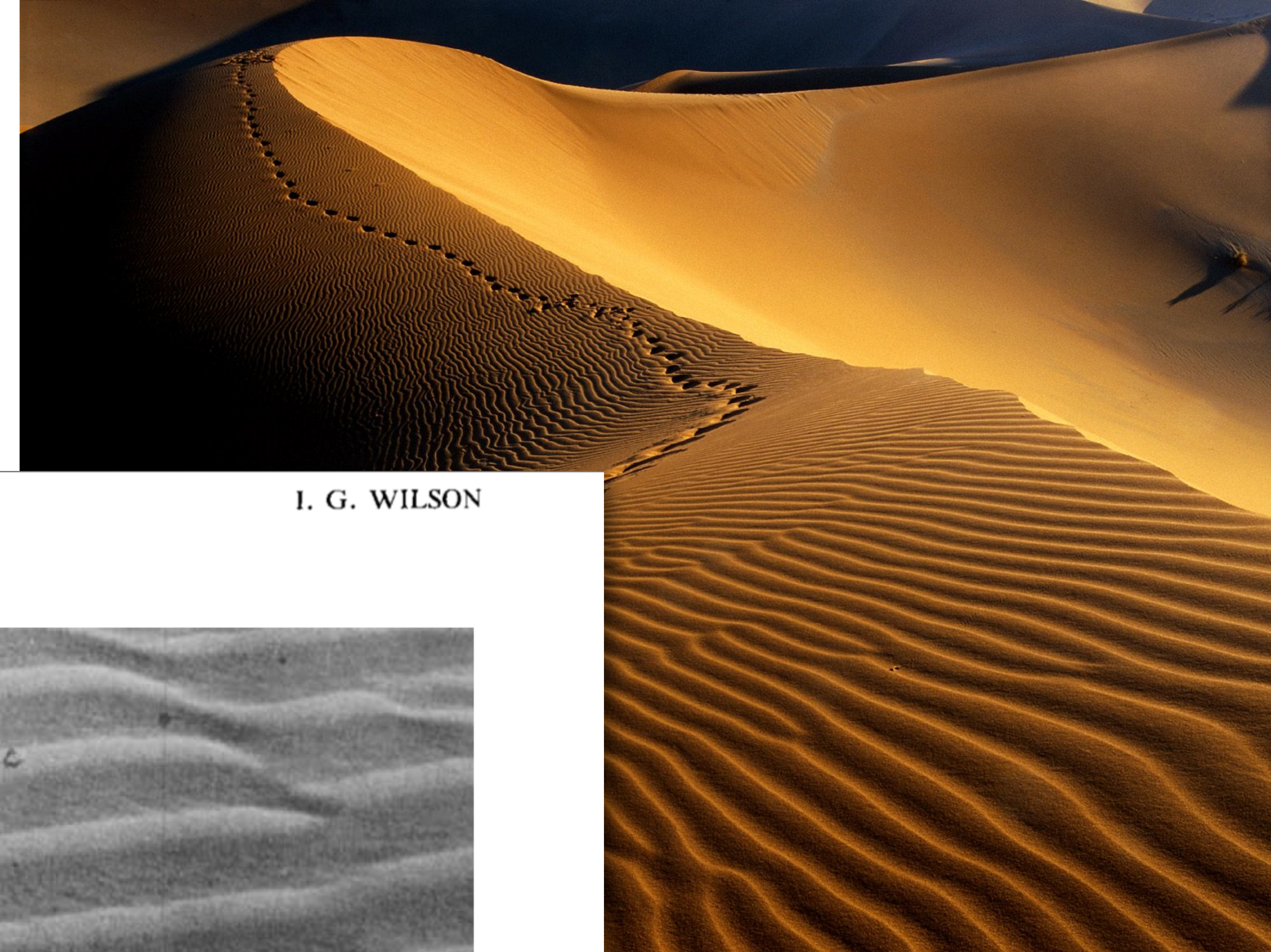
Self-repairing

Responsive to environment

Implicit and distributed encoding of pattern

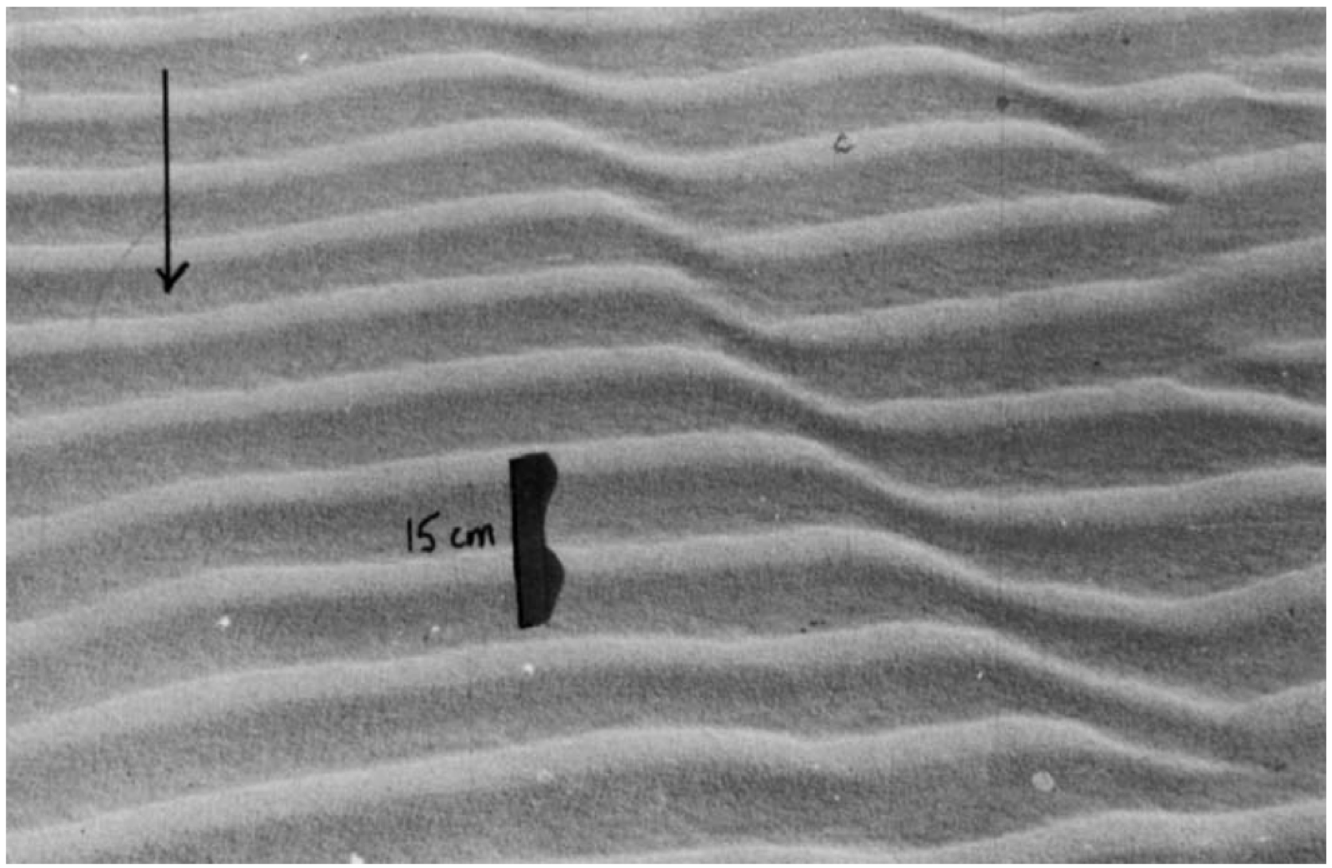
For a given pattern, where and how is it stored?

Period of sand ripples = 0.15 m



174 I. G. WILSON

PLATE I (pp.174–176)

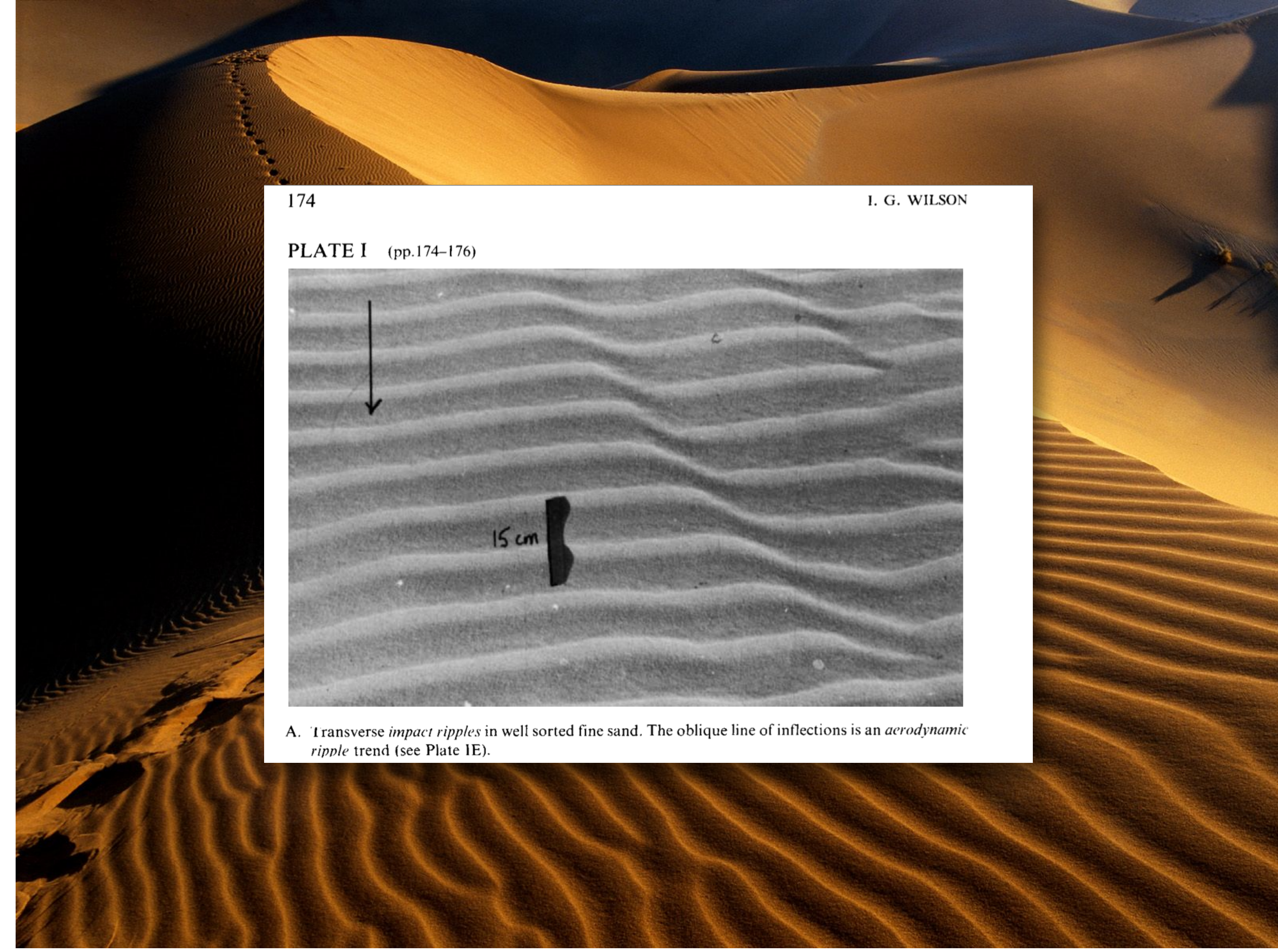


A. Transverse *impact ripples* in well sorted fine sand. The oblique line of inflections is an *aerodynamic ripple* trend (see Plate IE).

For a given pattern, where and how is it stored?

Let’s take a minute - please discuss with your neighbors

Period of sand ripples = 0.15 m



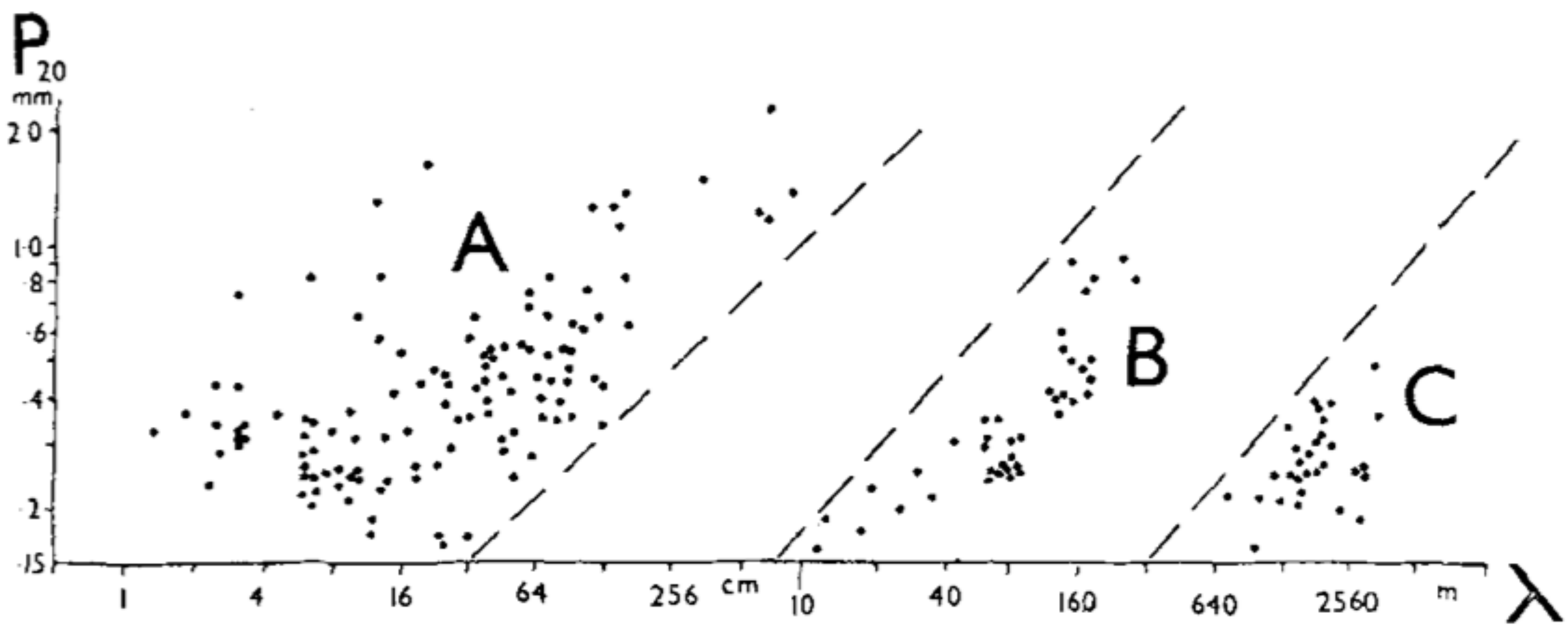


Fig.2. Scatter plot of grain-size (P_{20}—the coarse twenty-percentile) against wavelength for aeolian bedforms. There are no transitional forms between the three groups A, B, C, which correspond to ripples, dunes and draas.

What matters - consider extremes - blackboard - develop an expression

Gravitational force - if zero, sand blows away - no pattern

Gravitational force - if very large, sand does not move - no pattern

Size/shape of sand grains - if sand is too fine = sand storm/haze, no pattern

Wind - if zero, sand does not move - no pattern

Wind - if strong, sand blows away - no pattern

Guess - forces acting on sand need to be similar - gravity and wind

Lots of measurements

Lots of measurements

Sedimentology Elsevier Publishing Company, Amsterdam Printed in The Netherlands

AEOLIAN BEDFORMS—THEIR DEVELOPMENT AND ORIGINS

IAN G. WILSON†

Sedimentology Research Laboratory, Reading University, Reading (Great Britain)

(Accepted for publication November 5, 1971)

ABSTRACT

Wilson, I. G., 1972. Aeolian bedforms—their development and origins. *Sedimentology*, 19:173–210.

In wind, the sediment transport rate is given closely by the formula (after Bagnold, 1941, p.67):

$$q = 1.4 \sqrt{d} \cdot {u'_*}^{3} \cdot 10^{-5} \text{ g cm}^{-1} \text{ sec}^{-1}$$

where d (cm) is the grain diameter of the sand. Because it takes time for q to re-

$$u'_{*\mathrm{T}} = 0{\cdot}59\, \lambda^{1/3} \text{ cm/sec (draas)}$$

In short, other things being equal, the wavelengths of dunes and draas are approximately proportional to the cube of the wind velocity. The relationship based on the

Beginnings of an answer

Sand + Wind

Pattern is "stored" in

Local wind velocity

Grain diameter

Earth's Gravitational Field

The Navier-Stokes Equations

Typical solar flux = heating = wind

Conservation of Energy

Onset of Turbulence

AEOLIAN BEDFORMS—THEIR DEVELOPMENT AND ORIGINS 199

In wind, the sediment transport rate is given closely by the formula (after Bagnold, 1941, p.67):

$$q = 1.4 \sqrt{d} \cdot u_*'{}^3 \cdot 10^{-5} \text{ g cm}^{-1} \text{ sec}^{-1}$$

where d (cm) is the grain diameter of the sand. Because it takes time for q to re-

$$u'_{*T} = 0{\cdot}59 . \lambda^{1/3} \text{ cm/sec (draas)}$$

In short, other things being equal, the wavelengths of dunes and draas are approximately proportional to the cube of the wind velocity. The relationship based on the

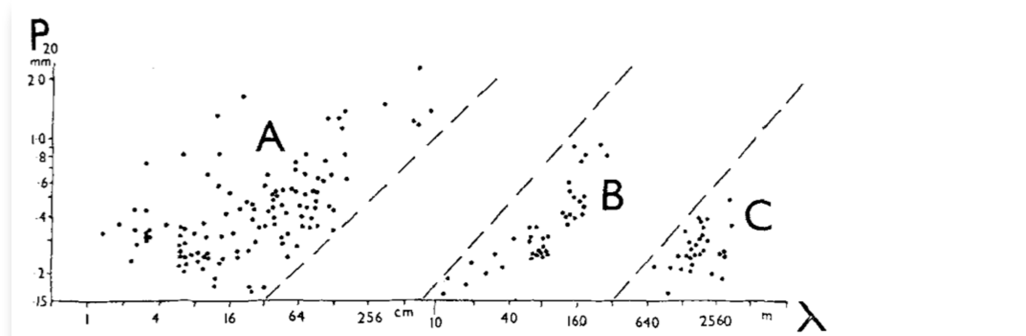


Fig.2. Scatter plot of grain-size (P_{20}—the coarse twenty-percentile) against wavelength for aeolian bedforms. There are no transitional forms between the three groups A, B, C, which correspond to ripples, dunes and draas.

Implicit and delocalized encoding/storage of pattern

complete absence of single, obvious, hack

Ok, so now please engineer for me a desert with user-specified ripple spacing.

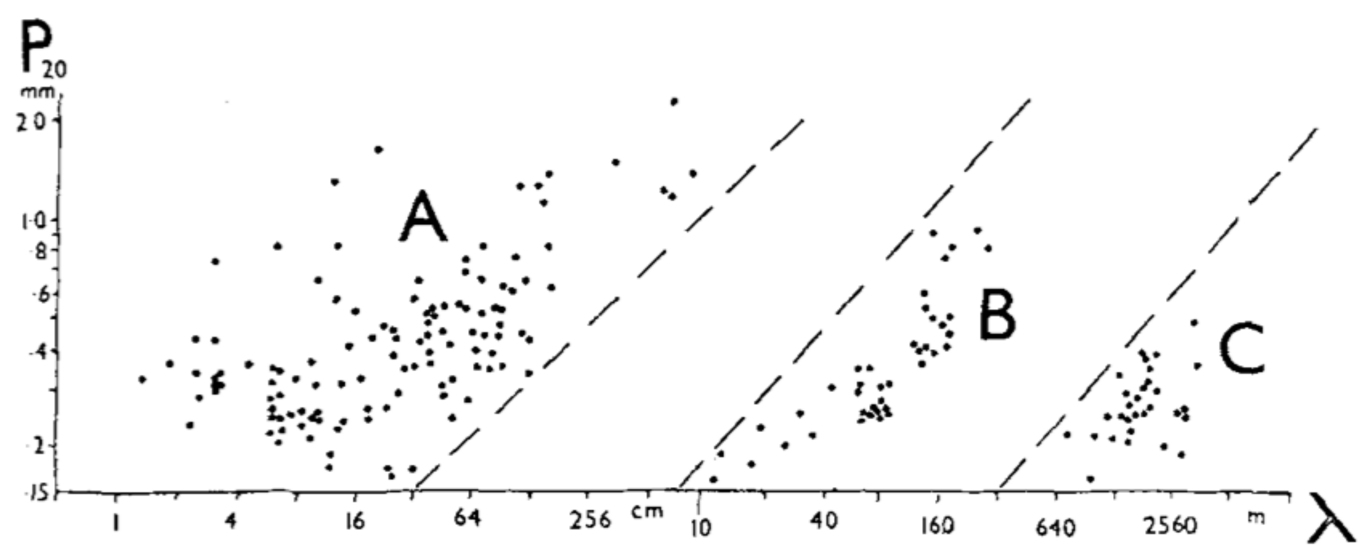


Fig.2. Scatter plot of grain-size (P_{20}—the coarse twenty-percentile) against wavelength for aeolian bedforms. There are no transitional forms between the three groups A, B, C, which correspond to ripples, dunes and draas.

Let’s take a minute -
please discuss with your
neighbors

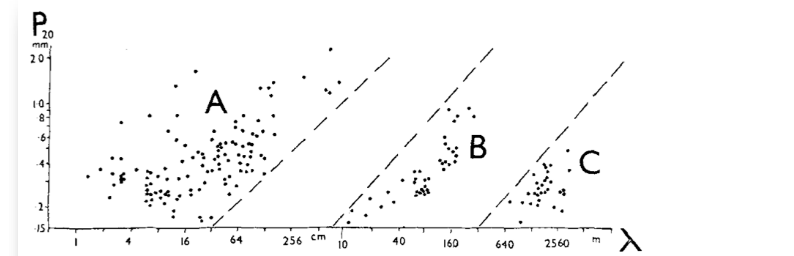


Fig.2. Scatter plot of grain-size (P_{20}—the coarse twenty-percentile) against wavelength for aeolian bedforms. There are no transitional forms between the three groups A, B, C, which correspond to ripples, dunes and draas.

Thought Experiment

Imagine you had a genome that encodes either small or large sand gains.

Genome A

Genome B

In the previous example, wind mattered…

What’s a biological equivalent of wind?

Thermal energy, which leads to Diffusion

Atomistic view:

All particles move due to thermal energy (1/2 $k_B T$ per DOF)

Movement direction is random

True for everything but 1/2 $k_B T$ is EXTREMELY SMALL so 'invisible' unless you are tiny!

Albert Einstein - Nobel Prize

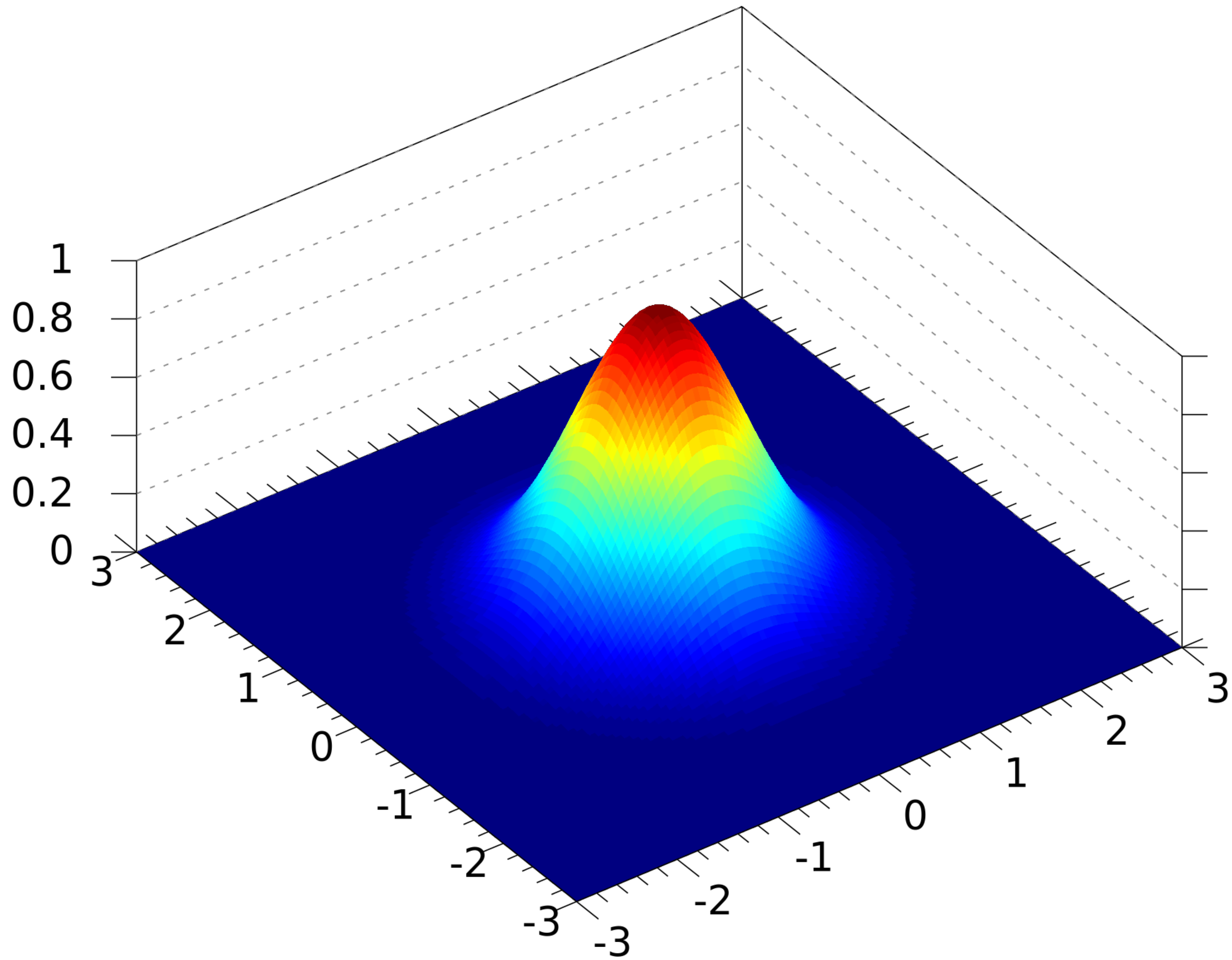
Phenomenological view:

The diffusion flux is proportional to the negative gradient of concentrations

"Driving force" due to concentration difference, stuff moves until the concentration difference is zero.

$$P(r,t) = \frac{e^{\left(\frac{-r^2}{4D \cdot t}\right)}}{4\pi D \cdot t}$$

$$D = \frac{k_B T}{6\pi\eta \cdot R}$$

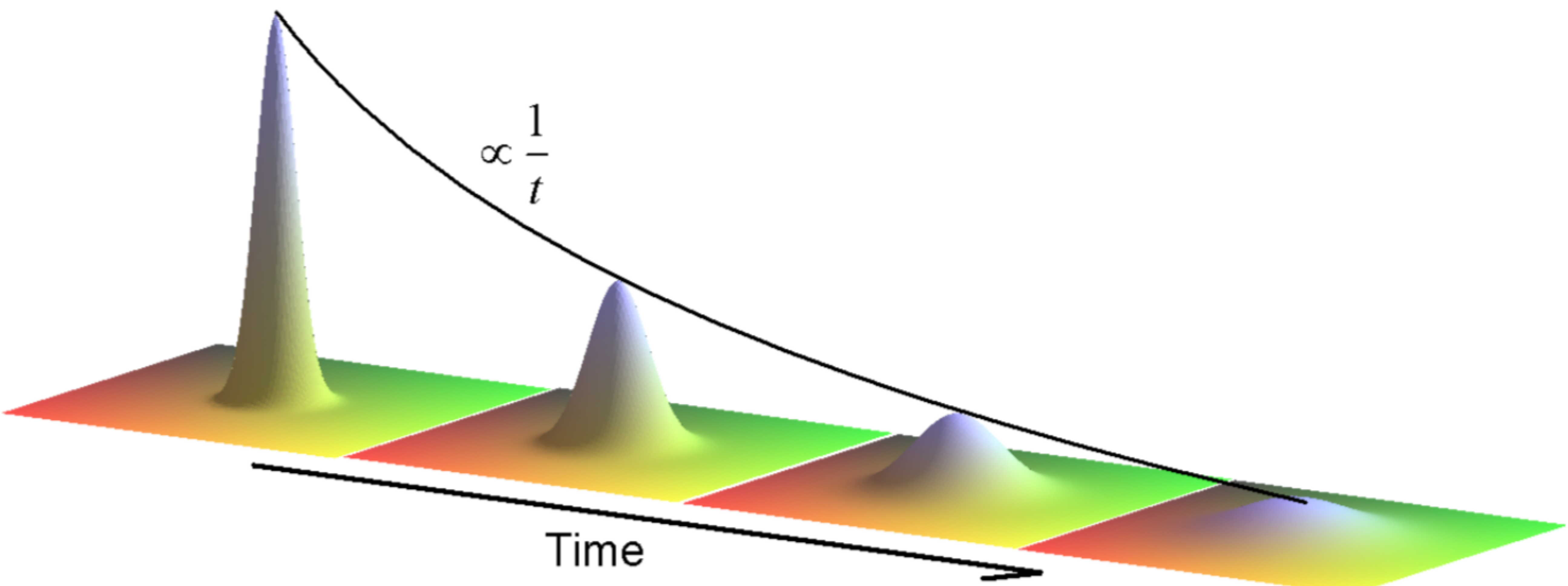


Further reading

http://www.rpgroup.caltech.edu/courses/aph162/2006/Protocols/diffusion.pdf

$$\boxed{P(r,t) = \frac{e^{\left(\frac{-r^2}{4D\cdot t}\right)}}{4\pi D\cdot t}}$$

$$\boxed{D = \frac{k_B T}{6\pi\eta\cdot R}}$$



Bigger proteins diffuse more slowly!

$$\boxed{J = -D\frac{dC}{dx}}$$

D ≡ diffusion coefficient

$D = (1/f)k_BT$

f ≡ frictional coefficient = $6 \pi \eta r$

Substance	Molecular Weight	$D(\text{cm}^2/\text{s})$
hydrogen	1	4.5×10^{-5}
oxygen	32	2.1×10^{-5}
carbon dioxide	48	1.92×10^{-5}
glucose	192	6.60×10^{-6}
insulin	5,734	2.10×10^{-6}
Cytochrome c	13,370	1.14×10^{-6}
Myoglobin	16,900	5.1×10^{-7}
Serum albumin	66,500	6.03×10^{-7}
hemoglobin	64,500	6.9×10^{-7}
Catalase	247,500	4.1×10^{-7}
Urease	482,700	3.46×10^{-7}
Fibrinogen	330,000	1.97×10^{-7}
Myosin	524,800	1.05×10^{-7}
Tobacco mosaic virus	40,590,000	5.3×10^{-8}

Journal of Bacteriology

Size Dependence of Protein Diffusion in the Cytoplasm of *Escherichia coli*

Anja Nenninger, Giulia Mastroianni and Conrad W. Mullineaux
***J. Bacteriol.* 2010, 192(18):4535. DOI: 10.1128/JB.00284-10.**
Published Ahead of Print 25 June 2010.

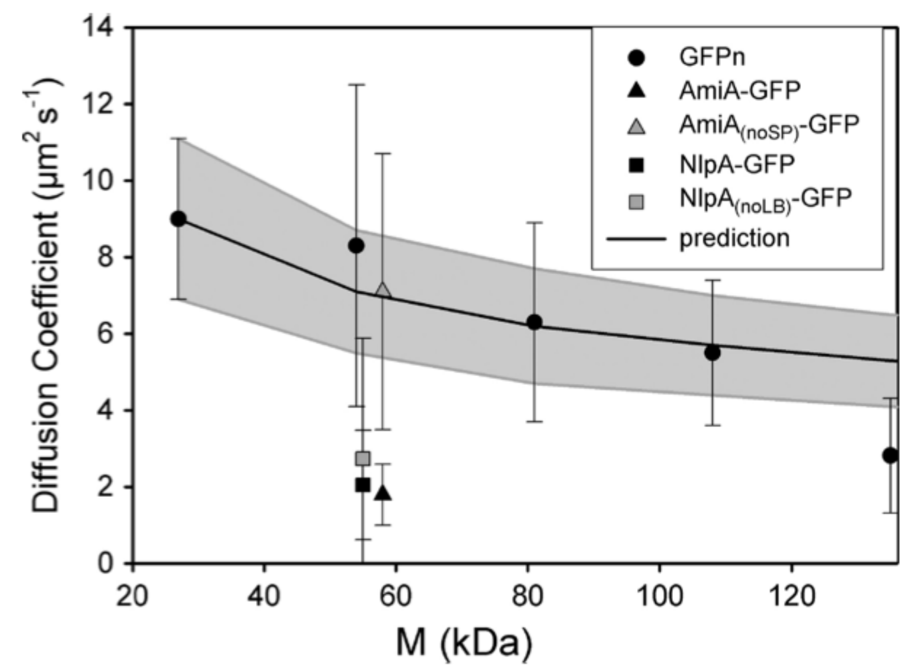


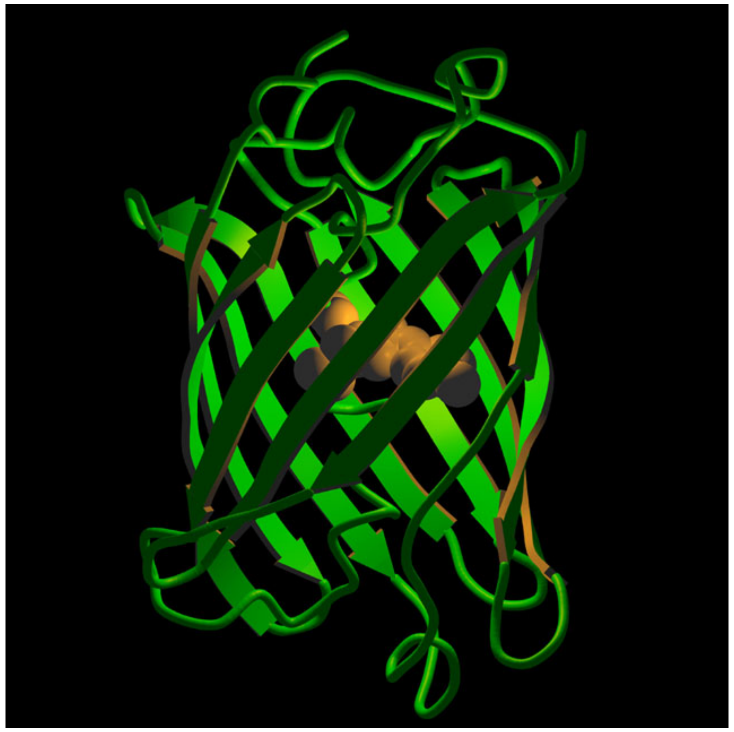
FIG. 2. Diffusion coefficients for GFP-tagged proteins in the *E. coli* DH5α cytoplasm. The mean diffusion coefficient ± SD is shown. "GFPn" represents multimers of torA-GFP from GFP1 to GFP5 (from left to right), as described in the text. The line shows the mean predicted *D* (± SD [gray-shaded area]) estimated by using the Einstein-Stokes equation to extrapolate from data for GFP1 (18) to larger proteins. Note that GFP multimers up to GFP4 show diffusion coefficients in line with this prediction. M, molecular mass.

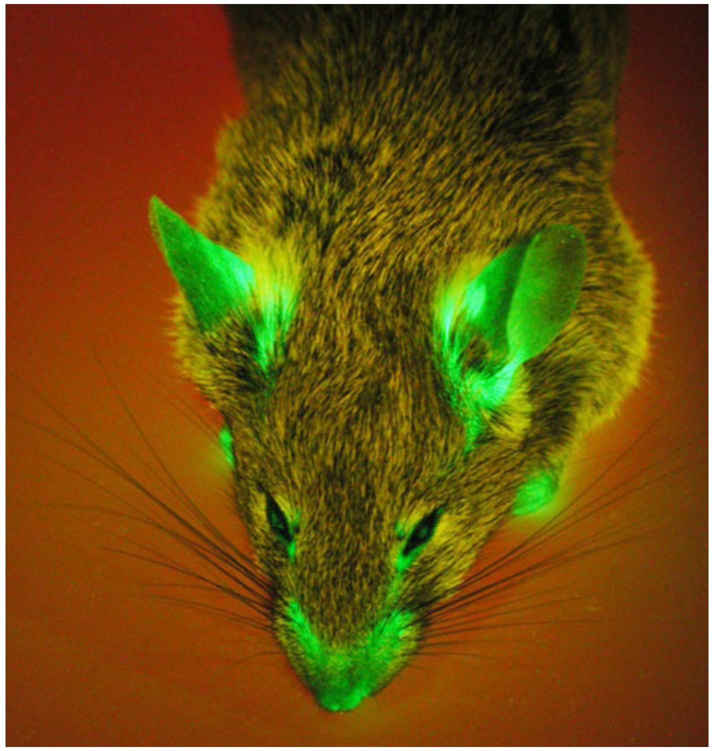
What is technically easy to control/alter in biology?

Gravity - nope.
Dynamic viscosity of water - nope.
Temperature - nope.
Boltzmann's Number - only in your dreams.

Protein size - YES
Relative protein levels - YES
Protein Protein interactions - YES
Enzyme kinetics - not so easy, but doable

Bigger proteins diffuse more slowly!





GFP

~238 aa

Sequences

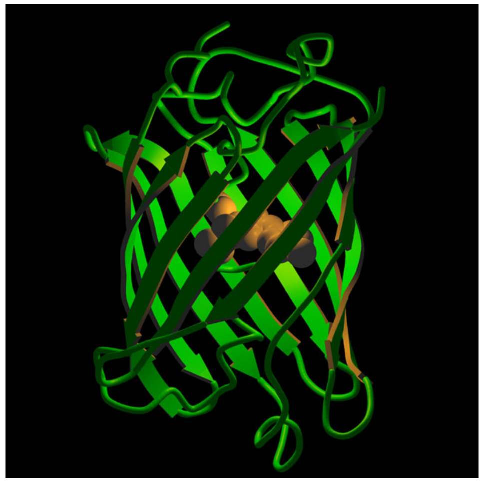
	Sequence		Length	Mass (Da)	Tools
☐	P42212 [UniParc].	FASTA	238	26,886	Blast go

Last modified November 1, 1995. Version 1.
Checksum: EA5A6F21FBFB6E05

```
        10         20         30         40         50         60
MSKGEELFTG VVPILVELDG DVNGHKFSVS GEGEGDATYG KLTLKFICTT GKLPVPWPTL

        70         80         90        100        110        120
VTTFSYGVQC FSRYPDHMKQ HDFFKSAMPE GYVQERTIFF KDDGNYKTRA EVKFEGDTLV

       130        140        150        160        170        180
NRIELKGIDF KEDGNILGHK LEYNYNSHNV YIMADKQKNG IKVNFKIRHN IEDGSVQLAD

       190        200        210        220        230
HYQQNTPIGD GPVLLPDNHY LSTQSALSKD PNEKRDHMVL LEFVTAAGIT HGMDELYK
```

« Hide

GFP sequence to color code



GFP

~238 aa

Sequences

Sequence		Length	Mass (Da)	Tools
P42212 [UniParc].	FASTA	238	26,886	Blast go

Last modified November 1, 1995. Version 1.
Checksum: EA5A6F21FBFB6E05

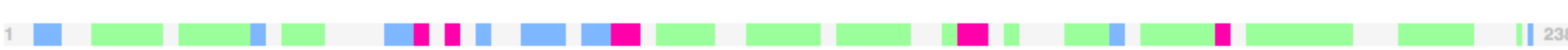
```
        10         20         30         40         50         60
MSKGEELFTG VVPILVELDG DVNGHKFSVS GEGEGDATYG KLTLKFICTT GKLPVPWPTL

        70         80         90        100        110        120
VTTFSYGVQC FSRYPDHMKQ HDFFKSAMPE GYVQERTIFF KDDGNYKTRA EVKFEGDTLV

       130        140        150        160        170        180
NRIELKGIDF KEDGNILGHK LEYNYNSHNV YIMADKQKNG IKVNFKIRHN IEDGSVQLAD

       190        200        210        220        230
HYQQNTPIGD GPVLLPDNHY LSTQSALSKD PNEKRDHMVL LEFVTAAGIT HGMDELYK
```

« Hide

Sequence annotation (Features)

Feature key	Position(s)	Length	Description	Graphical view	Feature identifier
Molecule processing					
Chain	1 – 238	238	Green fluorescent protein		PRO_0000192576
Amino acid modifications					
Modified residue	66	1	(Z)-2,3-didehydrotyrosine		
Cross-link	65 ↔ 67		5-imidazolinone (Ser-Gly)		
Natural variations					
Natural variant	100	1	F → Y. Ref.1		
Natural variant	108	1	T → S. Ref.1		
Natural variant	141	1	L → M. Ref.1		
Natural variant	219	1	V → I. Ref.1		
Experimental info					
Mutagenesis	65	1	S → G in mut3; highly fluorescent mutant when excited at 488 nm; when associated with A-72. Ref.4		
Mutagenesis	65	1	S → T: Increases fluoresence, photostability and shift the major exitation peak to 488 nm. Ref.4		
Mutagenesis	72	1	S → A in mut3; highly fluorescent mutant when excited at 488 nm; when associated with G-65. Ref.4		
Sequence conflict	2	1	S → G in CAA65278. Ref.3		
Sequence conflict	25	1	H → Q in CAA65278. Ref.3		
Sequence conflict	80	1	Q → R in CAA65278. Ref.3		
Sequence conflict	157	1	Q → P in AAA58246. Ref.2		
Sequence conflict	172	1	E → K in AAA58246. Ref.2		

Secondary structure

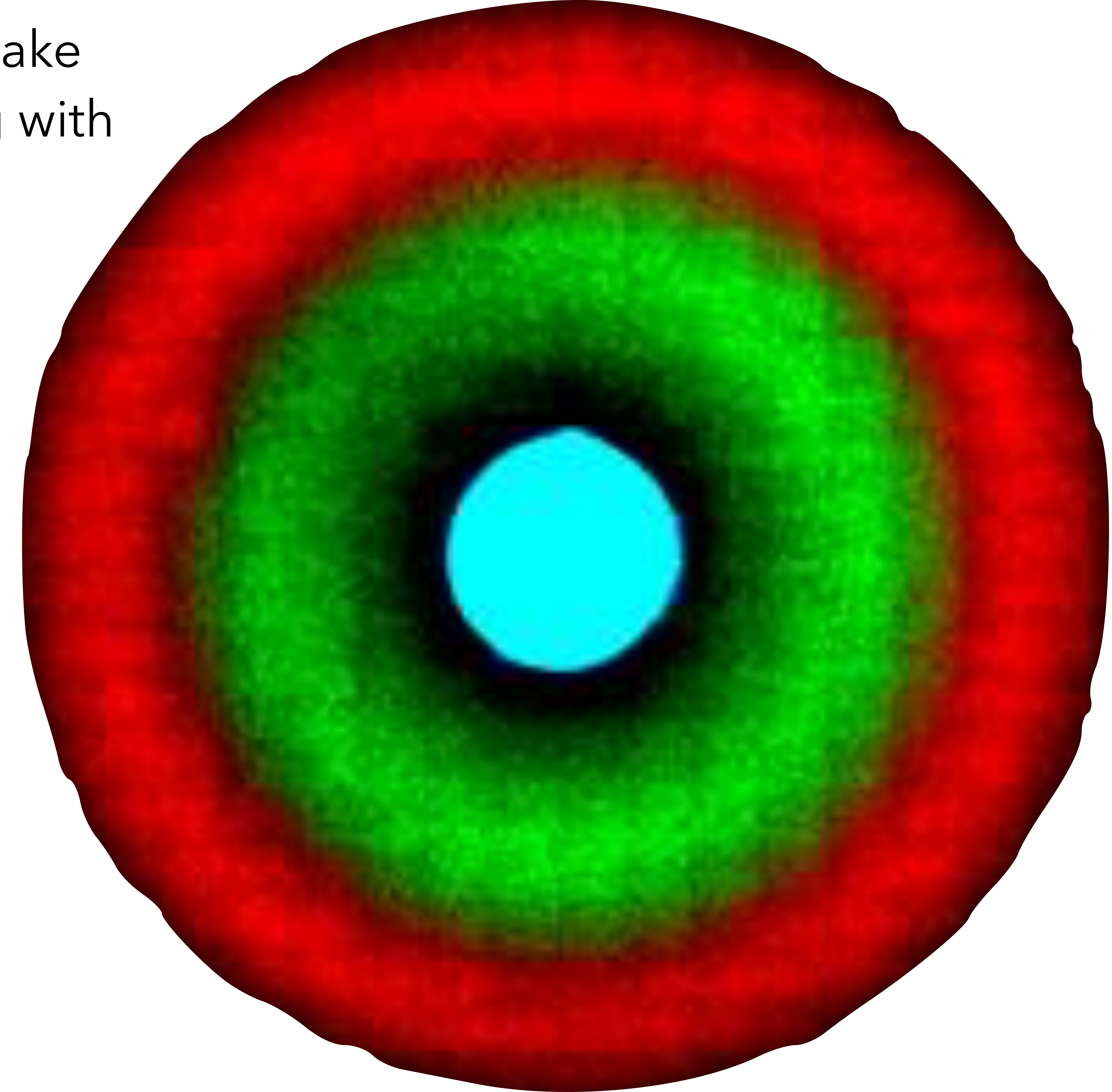


Helix Strand Turn

Details...

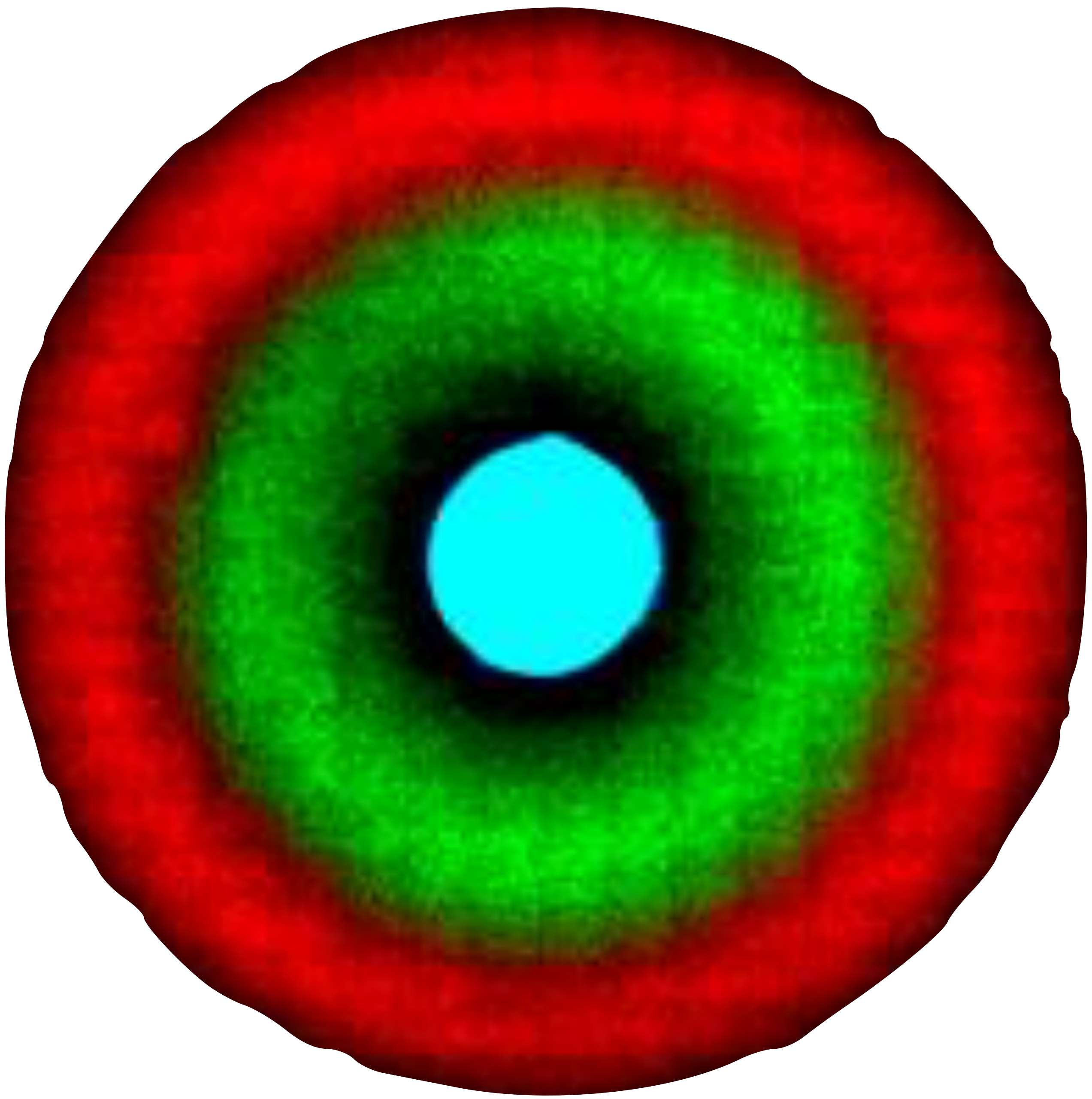
We know the sequence to color code!

Your task - make me this thing with living matter



Thought Experiment

Imagine you had a genome that encodes a **large green protein** and a **small red protein**, and the cell secretes these proteins. Which pattern will you get?



Recap

Approach the problem differently!

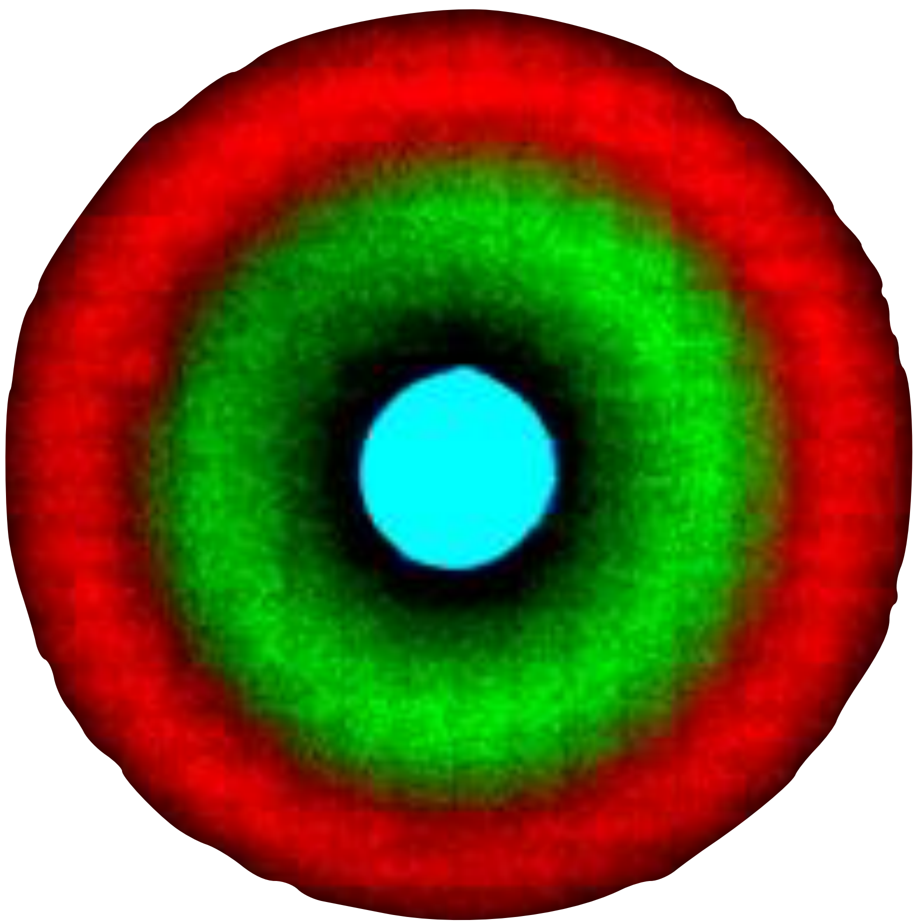
Implicit and Distributed Encoding of Pattern

Store the pattern in biology (genome) in a highly obfuscated form; implicitly make many assumptions about local physics (velocities, viscosities, adhesion, gravity…)

Pattern emerged from **interplay** of sand grain size and wind velocity

interplay of two (or 1000 things)

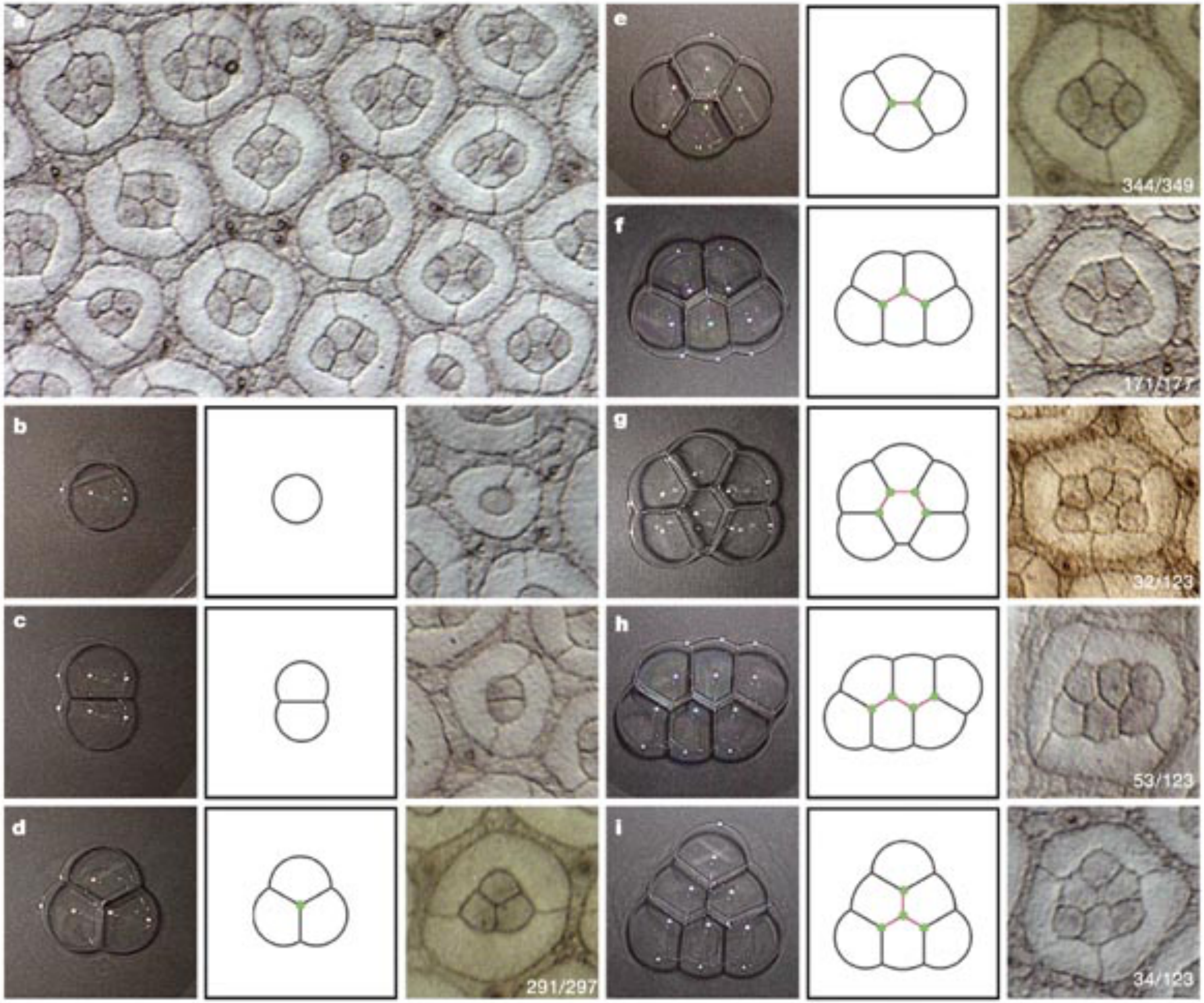
That was conceptually useful, but, there is no way you can tell me that diffusion, surface tension, + genome etc. can give more interesting patterns…



Nevermind…

Surface mechanics mediate pattern formation in the developing retina

Takashi Hayashi & Richard W. Carthew



https://www.nature.com/nature/journal/v431/n7009/full/nature02952.html

Fail

Your task - make me this thing with living matter