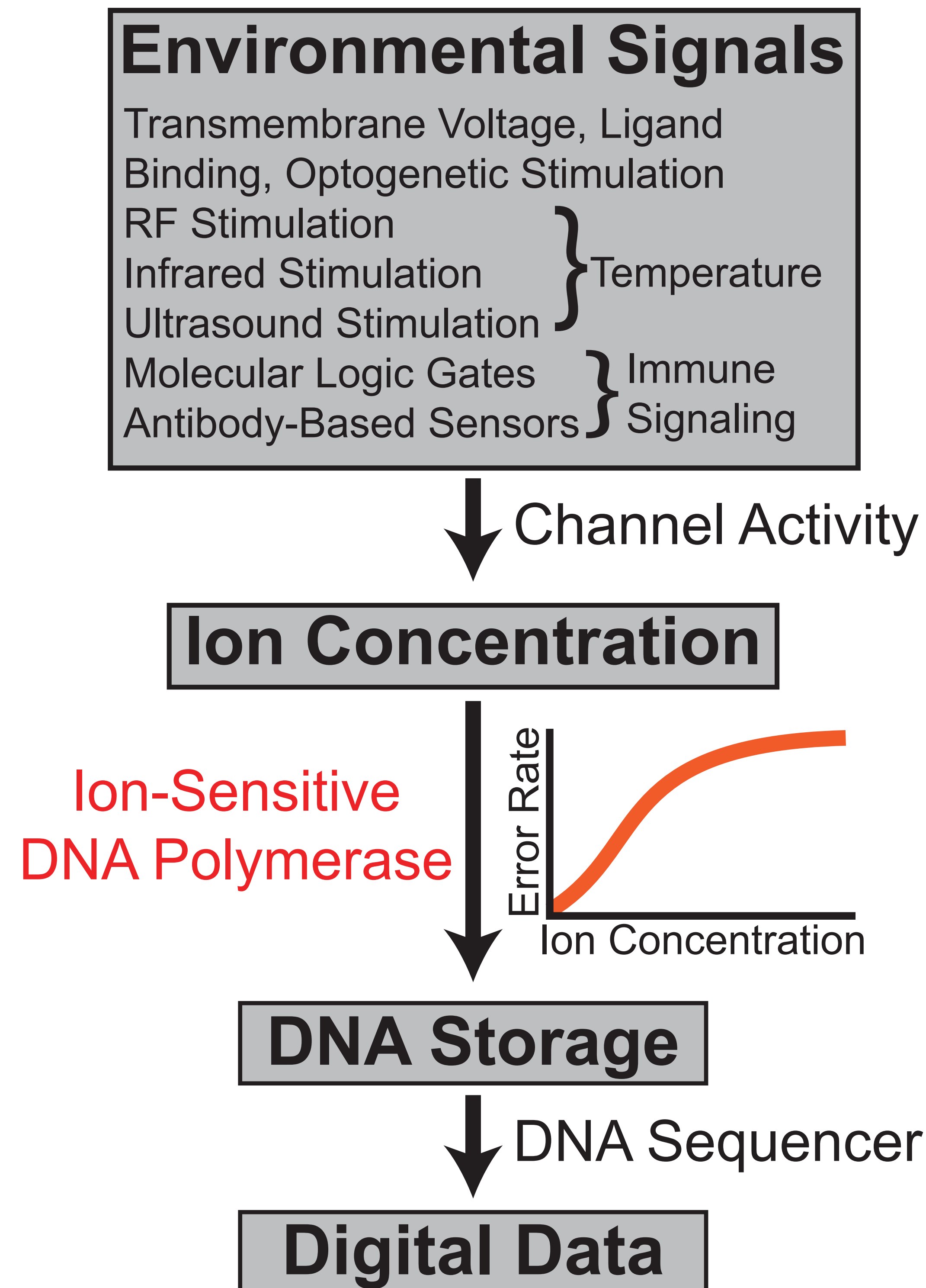
# molecular ticker-tapes

Adam Marblestone

*on behalf of the Church, Kording, Boyden and Tyo labs*

[original slides redacted to remove unpublished info]

# a (molecular) recording device inside each neuron



# ideal neuroscience experiment

*perfect for theorists*

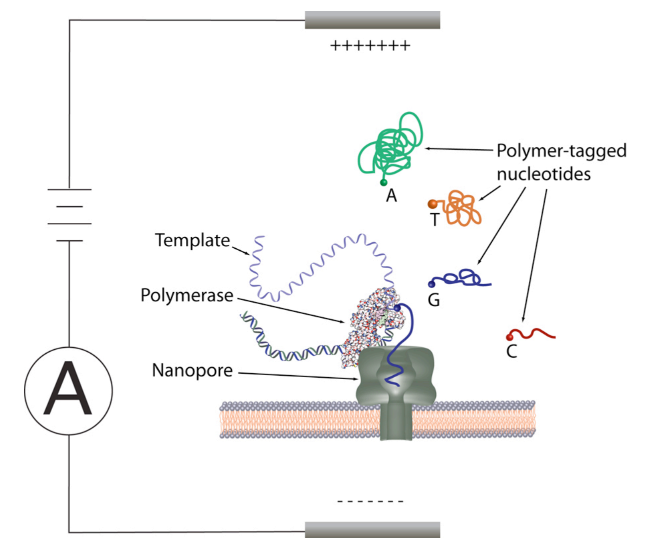
stimuli → behavior

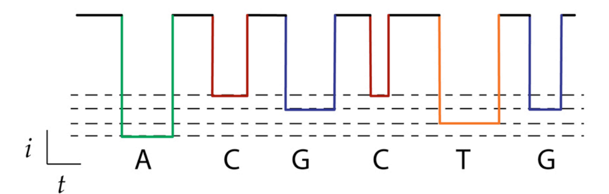
*mouse with no extra hardware*

kill →

extract DNA →

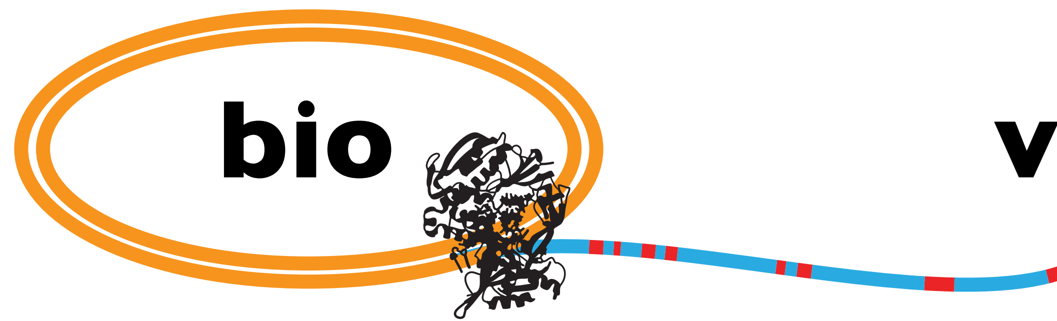
*nanopore sequencing*



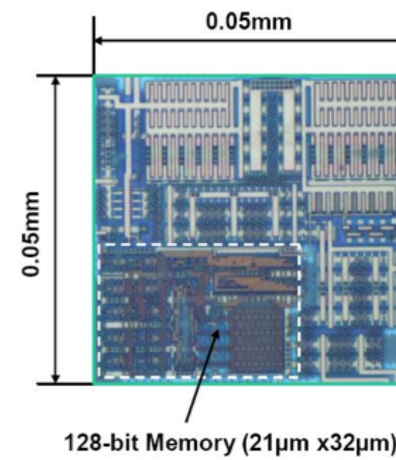




↓

**DATA**




vs.




electronic

**Bandwidth:** (1 bit / 1 ms / neuron) x (1e8 neurons) = **100 gigabit / s**

to read out a <u>mouse</u> brain at 1 bit per ms per neuron in real time

**readout of information <u>stored</u> in biomolecules can be <u>arbitrarily slow</u>**

**Power:** CMOS bit switching at 1000 Hz consumes more energy than a neuron

*molecular recorders are much more efficient*

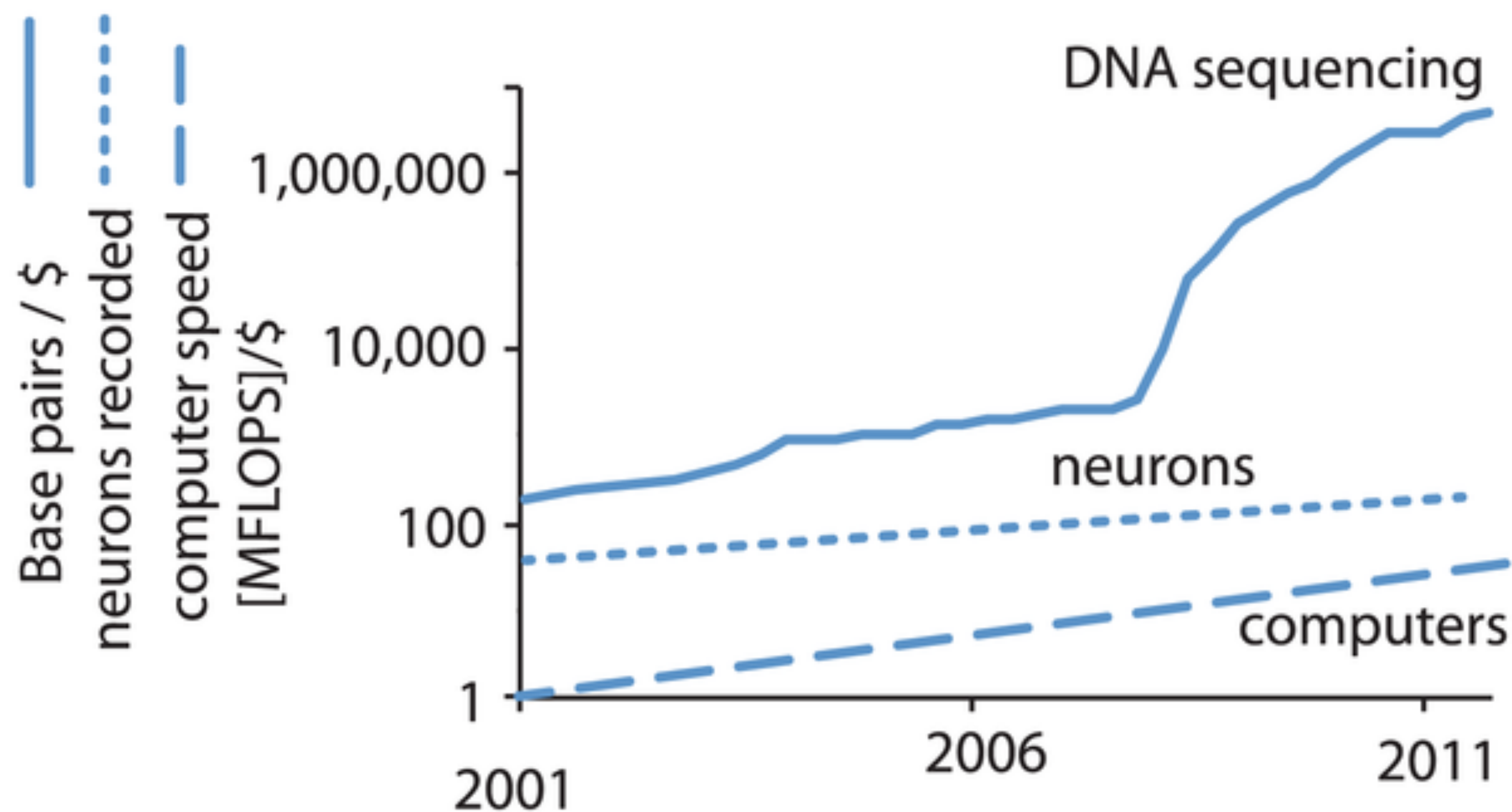
| | | |
|---|---|---|
| ultimate limit: | kT x log(2) per bit x 1000 bps | **3e-18 W** |
| practical limit for electronics: | 40 kT per bit x 1000 bps | **1e-16 W** |
| current CMOS: | ~1e6 worse than the 40 kT per bit limit | **~1e-10 W** |
| | 1fF x (1V)^2 x 1000 Hz | **1e-12 W** |
| 1 neuron: | (1e8 ATP / spike) x 100 Hz | **5e-10 W** |
| | (25W / human brain) x (1 human brain / 1e11 neurons) | **2.5e-10 W** |
| ATP consumption at 2000 Hz: | 2000 Hz x (5e-20J / ATP) | **1e-16 W** |

**(1000 nt/sec) x (1e8 neurons) x 60 sec / (3e9 nt per human genome)**

**~**

**2000 human genomes for <u>minimal</u> 1 minute whole mouse brain recording**

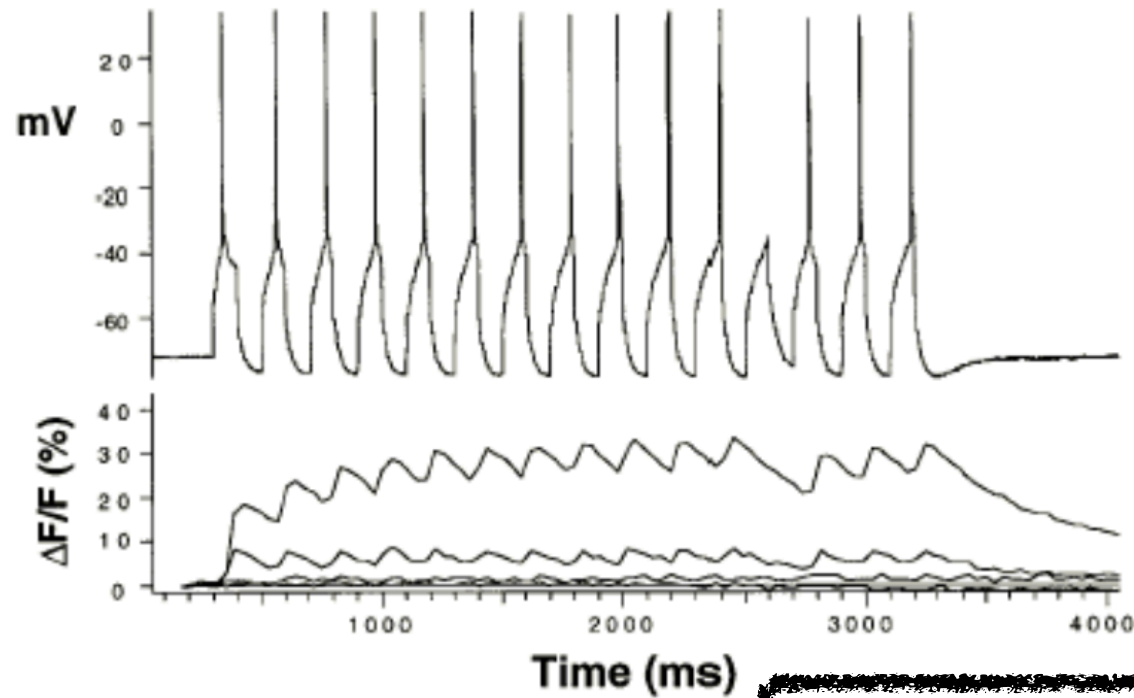
⟶ *need sequencing technology commensurate with zero-cost personal genomics*




DNA facts:
0.3 nm / base-pair
3e6 bp of dsDNA = 1 mm extended length
cell fits 3e9 bp into nucleus

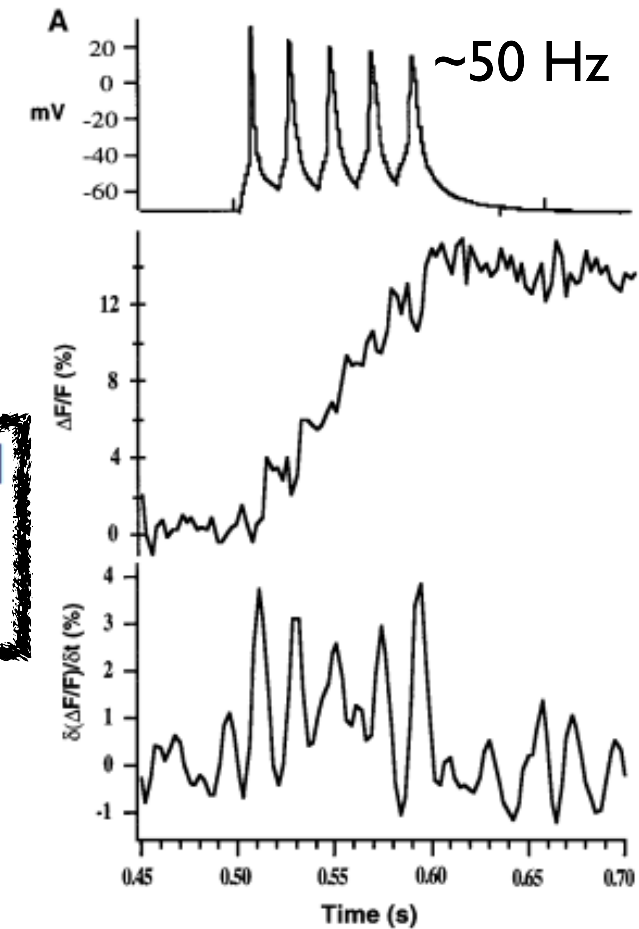
Kording 2011

# calcium as a proxy for neural firing




**Detecting Action Potentials in Neuronal Populations with Calcium Imaging**

Diana Smetters, Ania Majewska, and Rafael Yuste[1]



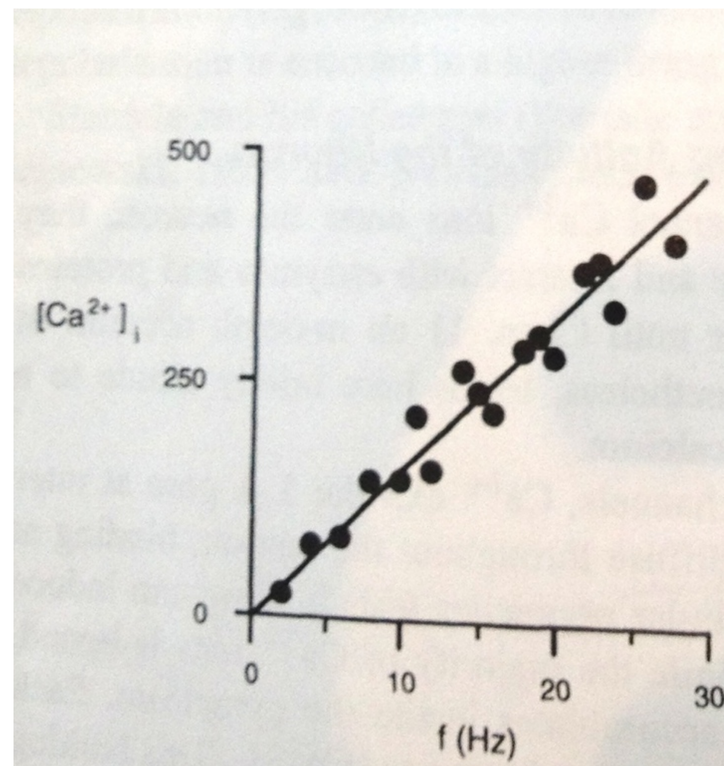
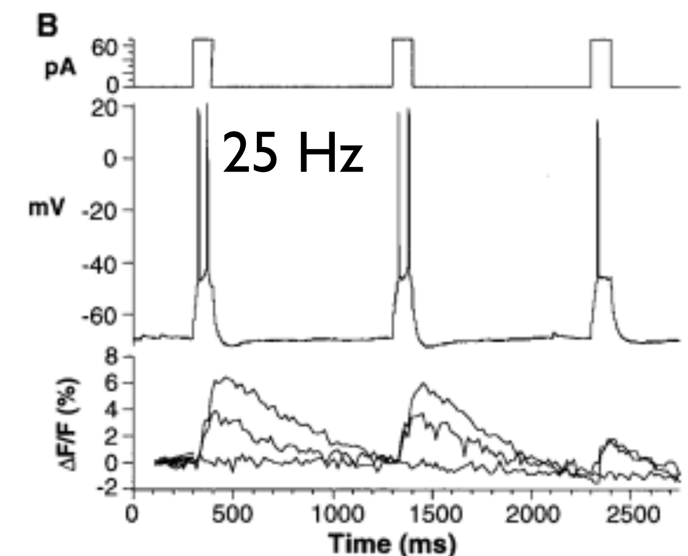
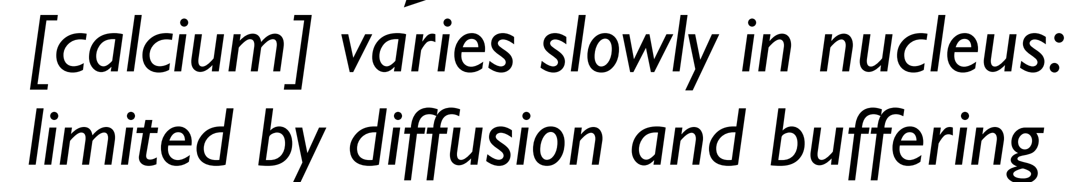




**Fig. 9.6 Calcium Buildup Proportional to Neuronal Activity** Following each action potential in a layer 5 pyramidal neuron, calcium rushes into the cell and accumulates, here recorded using a calcium-dependent fluorescent dye in the proximal apical dendrite (Helmchen, Imoto, and Sakmann, 1996). This increase in calcium reaches an equilibrium with a time constant of 200 msec. The final level (in nM) is linearly related to the firing frequency of the cell (evoked by current injections), with a slope of 15 n*M*/Hz (solid line). Reprinted in modified form by permission from Helmchen, Imoto, and Sakmann (1996).

[Koch, *Biophysics of Computation*]



# nucleic acid replication in the cytoplasm (not nucleus)?




## DNA

pox-viruses: cytoplasmic replication

**Vaccinia-like cytoplasmic replication of the giant Mimivirus**

Yael Mutsafi[a,1], Nathan Zauberman[a,1], Ilana Sabanay[b], and Abraham Minsky[a,2]

[a]Department of Organic Chemistry and [b]Electron Microscopy Center, The Weizmann Institute of Science, Rehovot 76100, Israel

## RNA

"...there is a special cytoplasmic membrane-associated transcription system in which DNA-dependent RNA polymerase II, which co-localizes with template cmDNA at the plasma membrane, can transcribe the membrane-associated 171-bp α-satellite repeat sequences into RNA..."

also lots of cytoplasmic RNA viruses!

# targeting proteins just to the axon or just to the synapse

## Axon

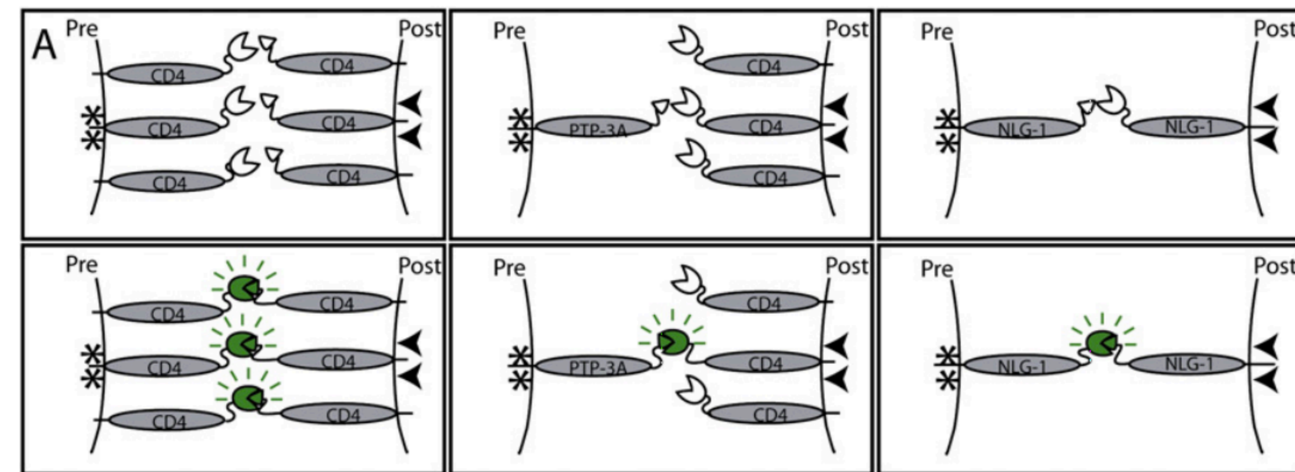
**$Na_v1.2$ targeting to axons**

The subcellular distribution of sodium channels from the $Na_v1$ family play a critical role in determining where action potentials initiate in neurons [18]. $Na_v1.2$ is localized to the axon hillock where it responds to depolarization by generating $Na^+$ currents necessary to trigger action potentials [36]. To identify signals that target this channel to the axon, a series of chimeric constructs consisting of cytoplasmic regions of the channel fused with heterologous proteins were tested in hippocampal neurons in dissociated culture [8]. These studies revealed that the C terminus contains a nine-amino acid, dileucine-containing motif that is sufficient to target heterologous proteins to the surface of the axon. Because these proteins are present intracellularly

**Polarized targeting of ion channels in neurons**

Don B. Arnold

## Synapse

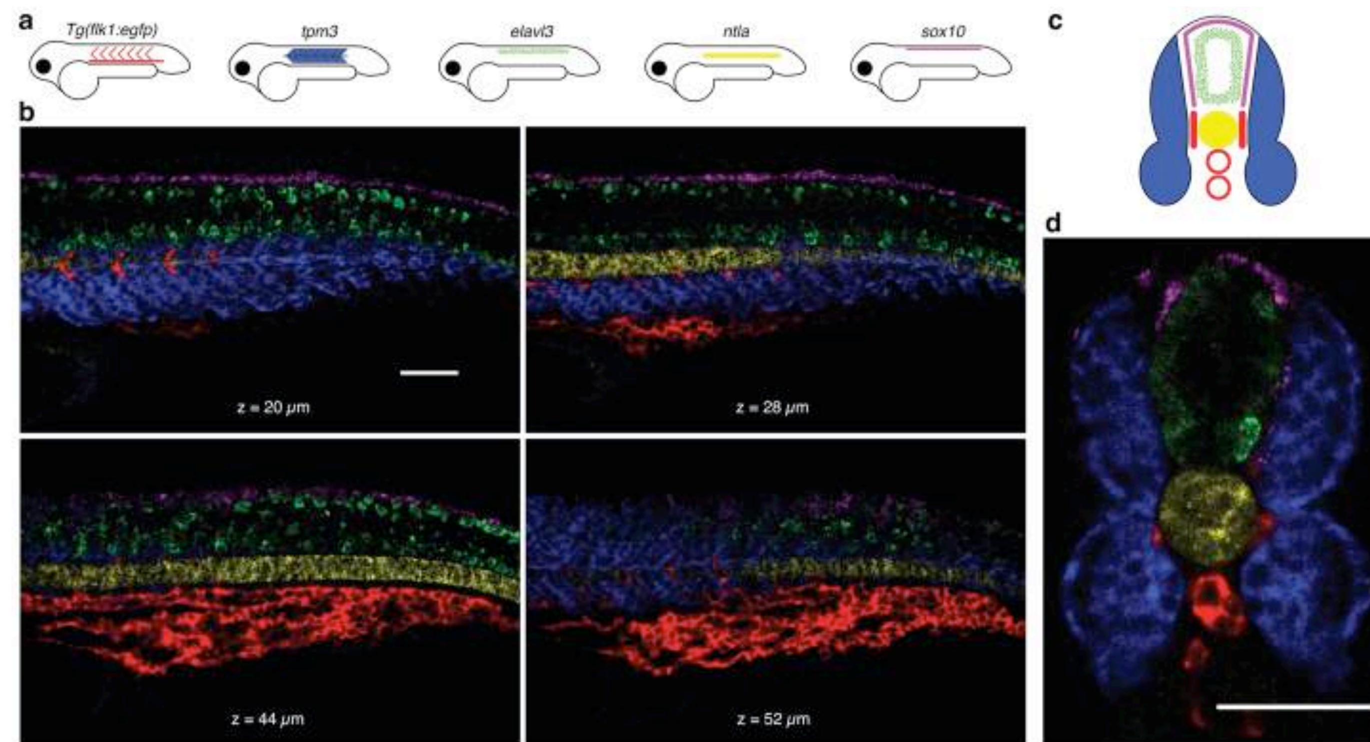



**GFP Reconstitution Across Synaptic Partners (GRASP) Defines Cell Contacts and Synapses in Living Nervous Systems**

Evan H. Feinberg,[1] Miri K. VanHoven,[2] Andres Bendesky,[1] George Wang,[2] Richard D. Fetter,[3] Kang Shen,[2,*] and Cornelia I. Bargmann[1,*]

# which DNA tape comes from which neuron?

**Fluorescent in-situ sequencing**: slice up the brain and sequence using a microscope




**Programmable in situ amplification for multiplexed imaging of mRNA expression**

Harry M.T. Choi,[1] Joann Y. Chang,[1] Le A. Trinh,[2] Jennifer E. Padilla,[1] Scott E. Fraser,[1,2] and Niles A. Pierce[1,3,*]

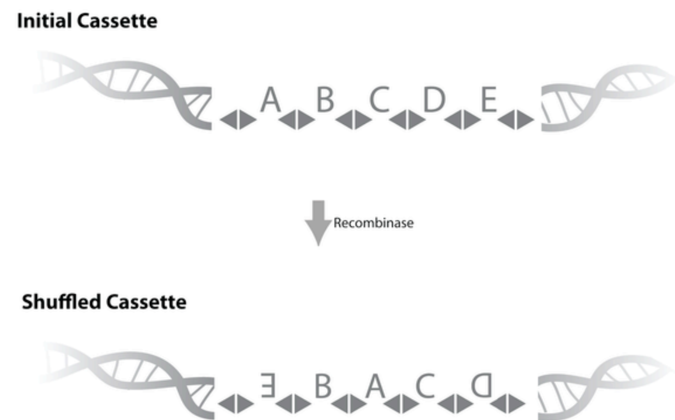
(e.g., could do 3D in-situ sequencing w/ CLARITY)

today: in-situ *hybridization* next step: in-situ *sequencing*

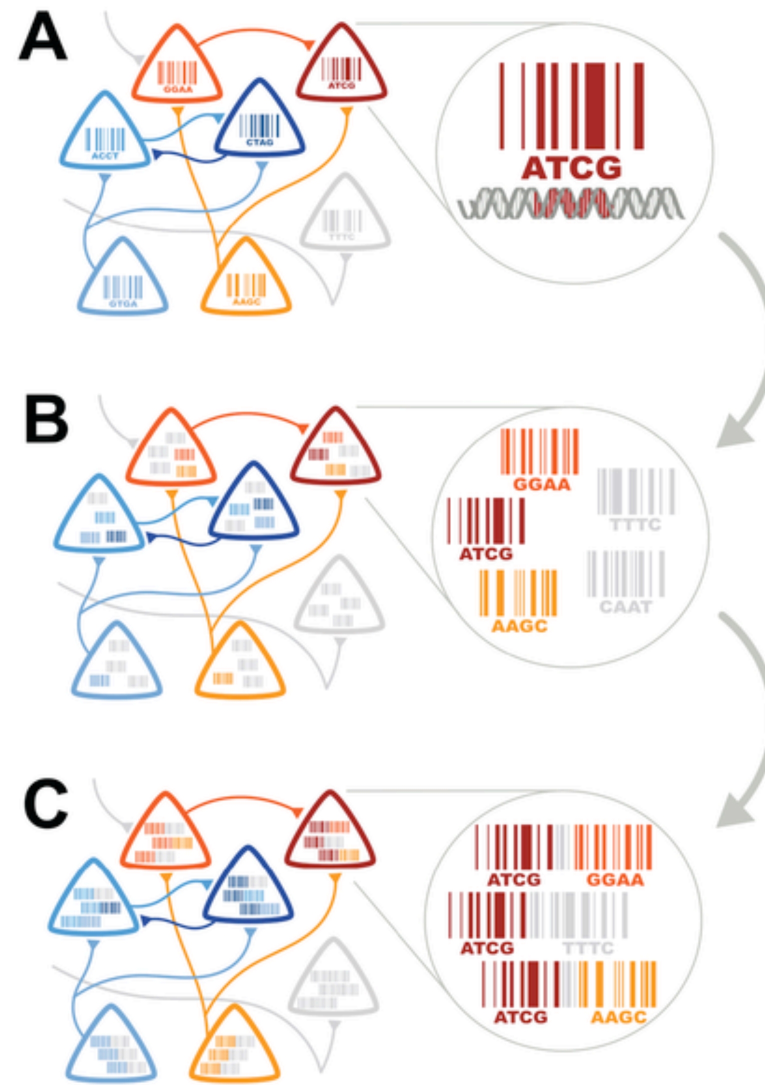
# which DNA tape comes from which neuron?

## **Zador**: DNA barcodes for *structural connectome*

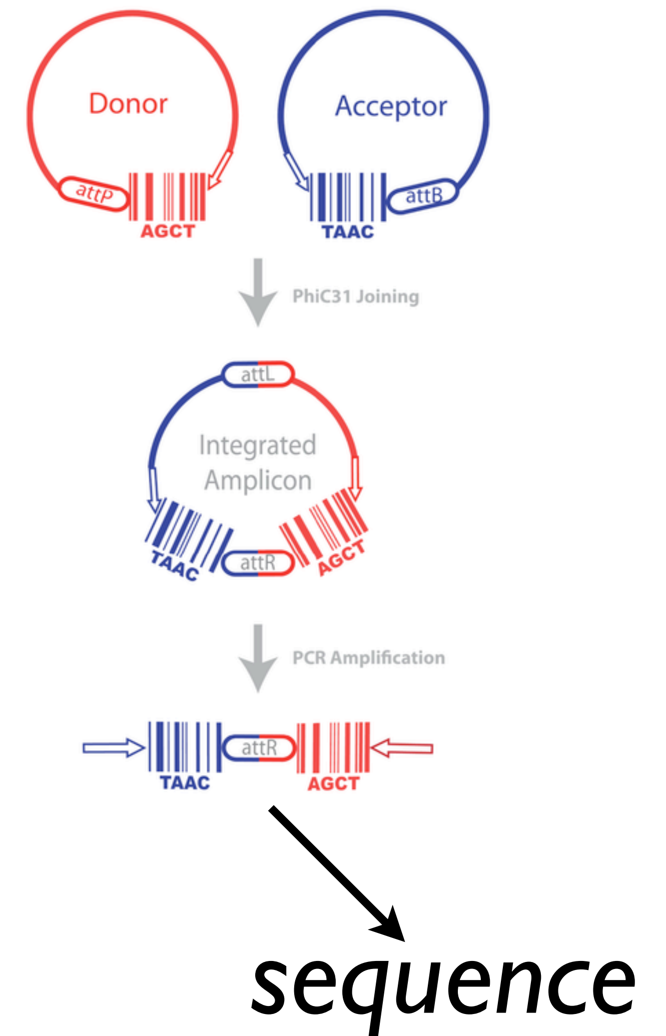
*make barcode*




*hop*




*stitch*

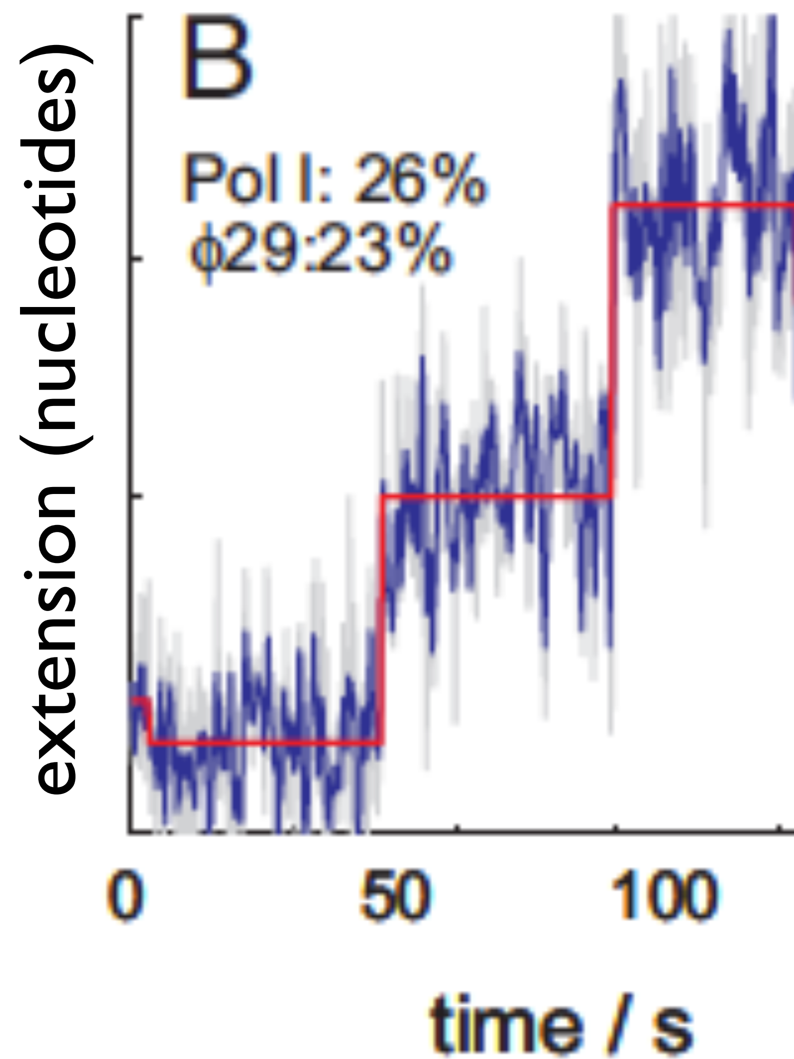



*sequence*

Sequencing the Connectome

Anthony M. Zador, Joshua Dubnau, Hassana K. Oyibo, Huiqing Zhan, Gang Cao, Ian D. Peikon

# is timing info preserved by a polymerase?



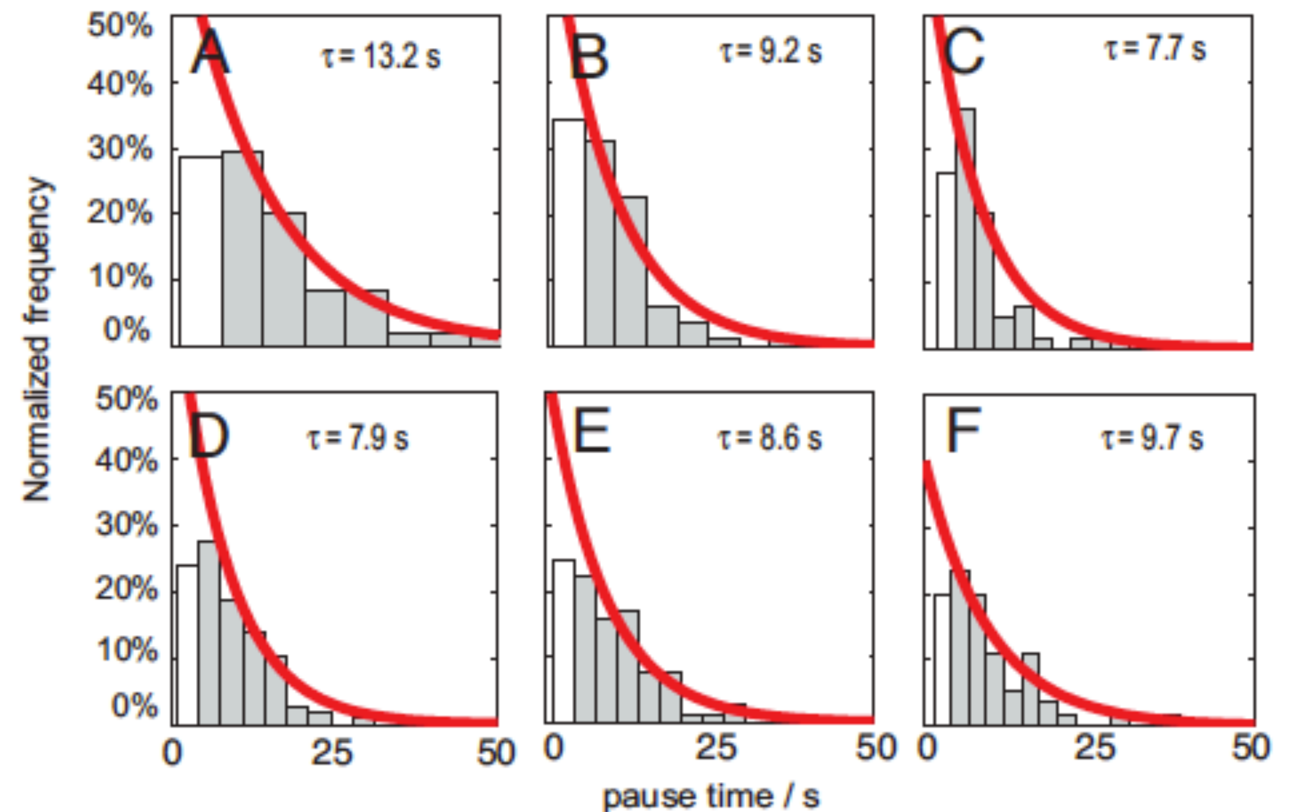
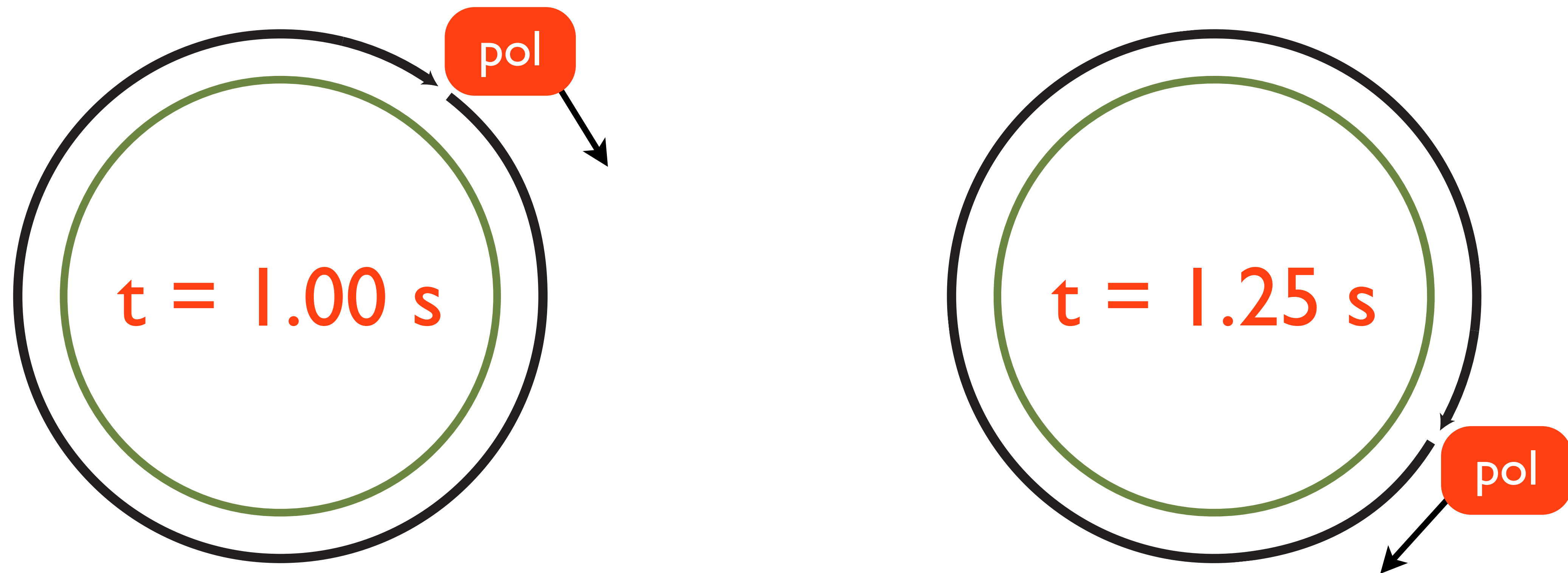




**Fig. 4.** Pause lifetimes for Pol I(KF) and φ29 were measured under different conditions for the sample and control sequences. The red curves are normalized single exponential fits given by $f = \tau^{-1} \exp(-\tau/\tau)$, where $\tau$ is the mean pause lifetime. Bins excluded from the fit due to undersampling are shown in white. (*A*) Pol I(KF) at 23 °C with the sample template; (*B*) Pol I(KF) at 23 °C with 1 M betaine with the sample template; (*C*) Pol I(KF) at 23 °C with the control template, (*D*) φ29 at 23 °C with the sample template, (*E*) φ29 at 23 °C with 1 M betaine with the sample template, and (*F*) φ29 at 23 °C with the control template.

**Single molecule measurement of the ''speed limit'' of DNA polymerase**

Jerrod J. Schwartz and Stephen R. Quake[1]

# a possible experimental paradigm: *nick translation* + rapid stopping




*polymerase w/ 5'-3' exonuclease eats up all but its last few seconds of DNA recording*

rapidly halt DNA synthesis at a precisely-known time: e.g. w/ opto-genetics or flash freezing

working back from 3' end: **few sec ~ few kb of recording w/ precise time-stamp**

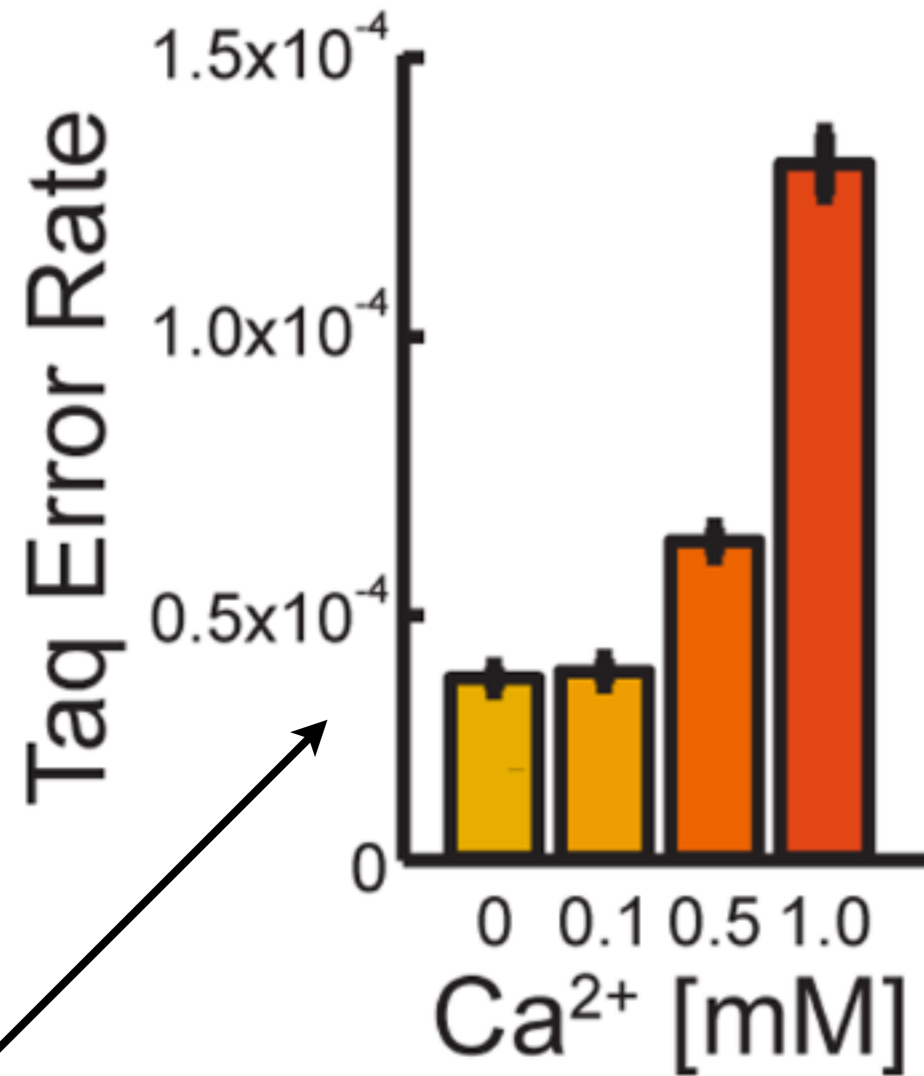
# Ions-to-Errors Baby Steps

**Q**: how can we rapidly screen ion-dependent effects on polymerase error rates?

***ideally*** *to identify a fast, highly processive polymerase with calcium-dependent error probabilities of up to several % per bp*

**Speed max**: 1000 nt/sec for E. coli pol III, 200 nt/sec for T7RNAP
**Error max**: 70% per base for Iota on template T
**Processivity max**: > 70,000 nt for phi29

# *calcium*

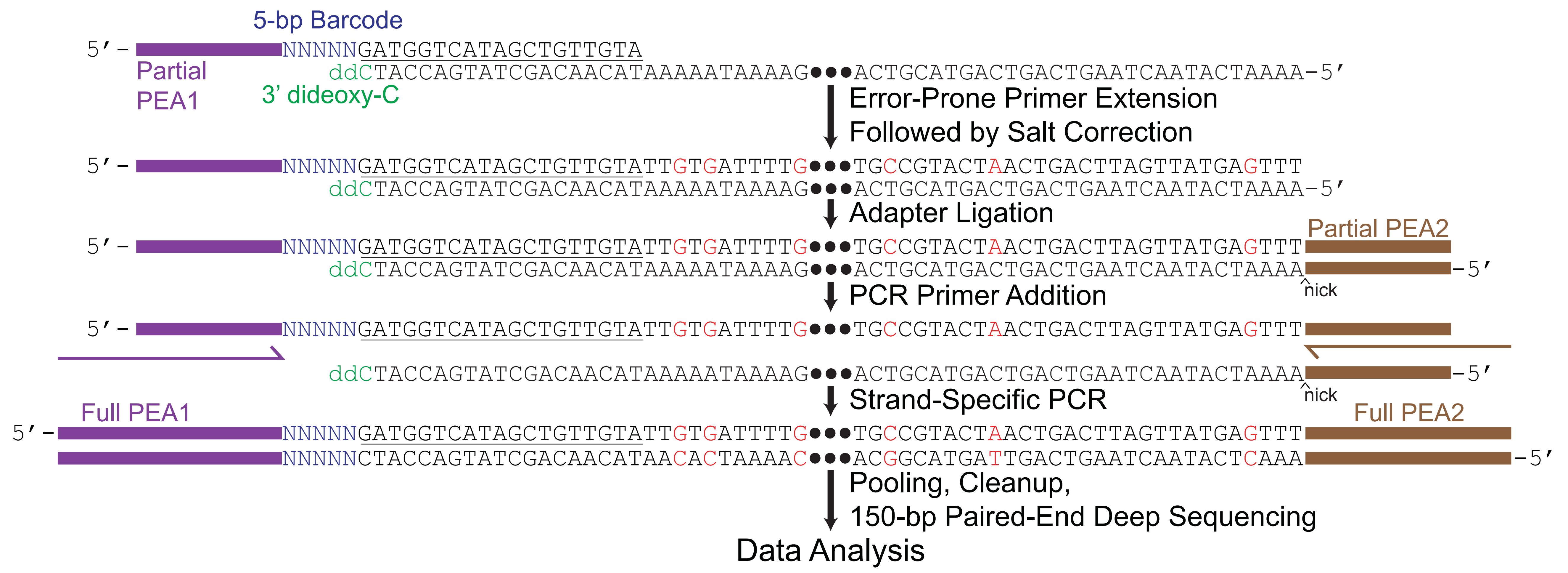



small error rates:

measured by colony counting not by sequencing

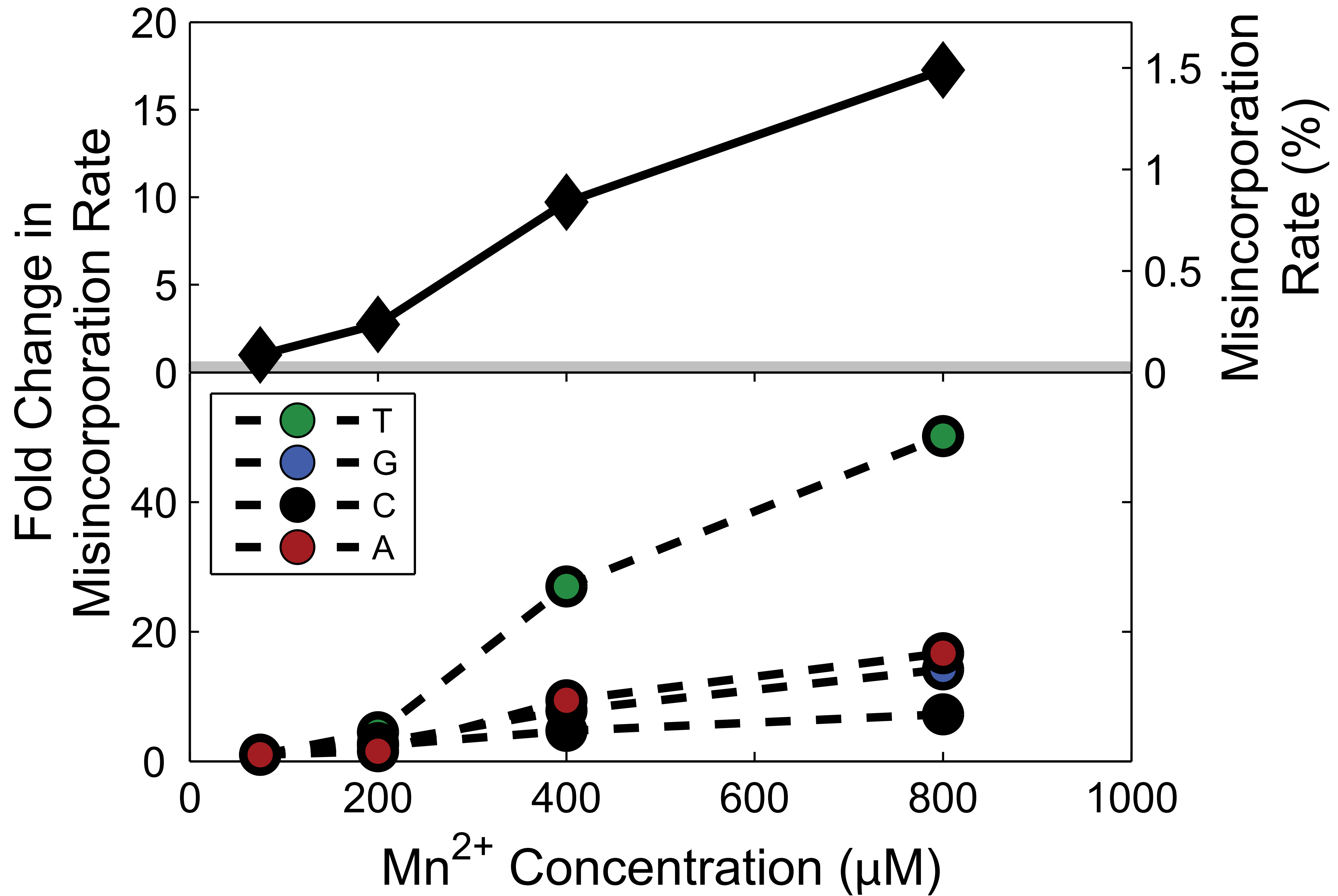
Daniel Schmidt

# measuring concentration-to-misincorporation transfer functions by deep sequencing



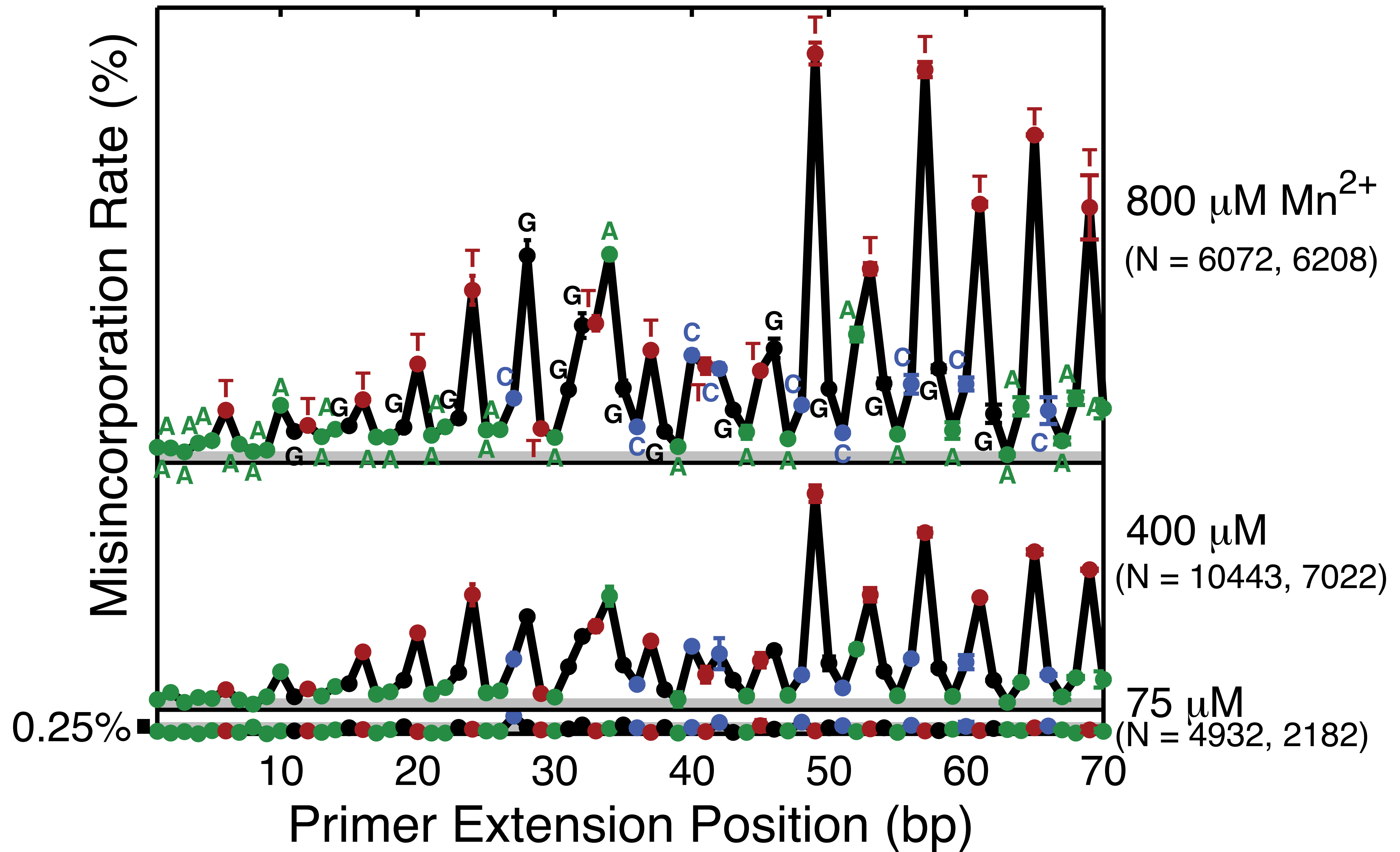

## with Brad Zamft

# measuring concentration-to-misincorporation transfer functions by deep sequencing




with Brad Zamft

# resolving details of polymerase mis-incorporation




with Brad Zamft

# big problems with Dpo4

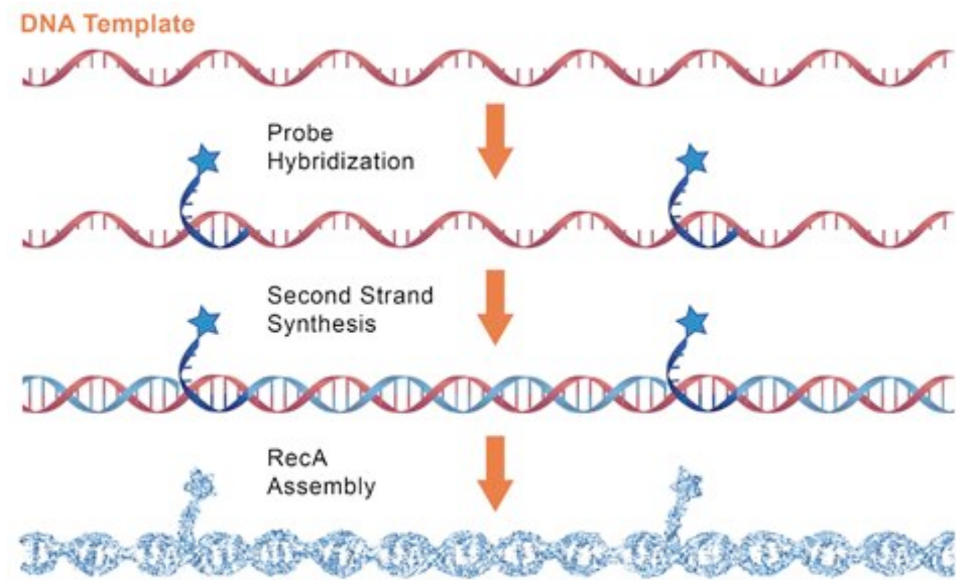
- Slow (~ 1 nucleotide per second)
- Non-processive
- Manganese (or pH) not Calcium

# phi29 DNAP

*rolling-circle amplification*

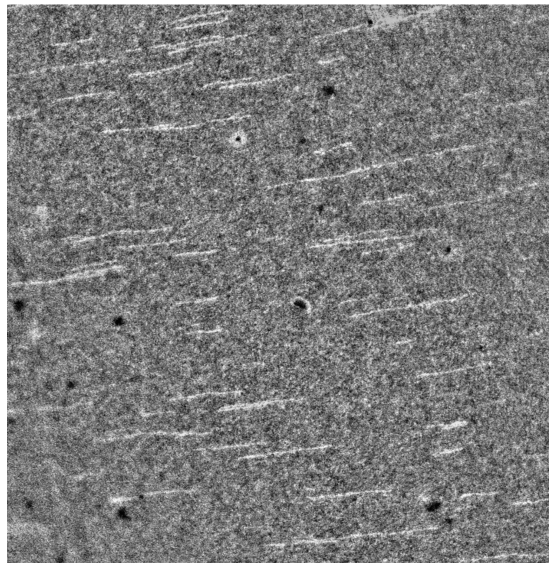
- ~ 50 nt/sec
- processive (~ 70,000 nt)
- must engineer calcium sensitivity

# crude long-read "sequencing" methods for ticker-tapes




NabSys: "hybridization-assisted" nano-pore sequencing

---



*meniscus-combed*
*48 kbp lambda dsDNA*

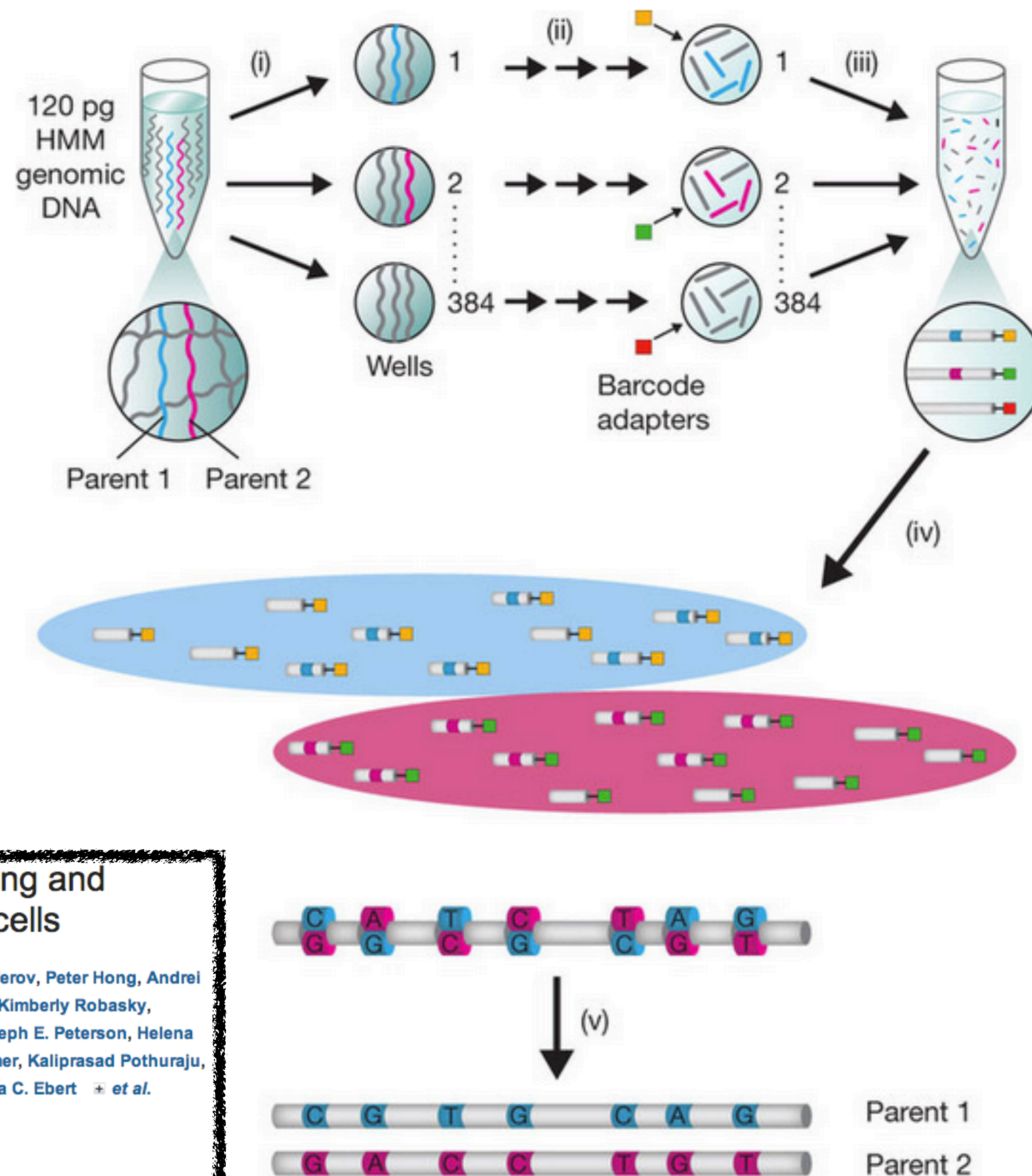
DNA combing & fluorescence microscopy

---

# simulating long-read sequencing with short-read sequencing

key ideas:

1) dilution
2) barcoding




Accurate whole-genome sequencing and haplotyping from 10 to 20 human cells

Brock A. Peters, Bahram G. Kermani, Andrew B. Sparks, Oleg Alferov, Peter Hong, Andrei Alexeev, Yuan Jiang, Fredrik Dahl, Y. Tom Tang, Juergen Haas, Kimberly Robasky, Alexander Wait Zaranek, Je-Hyuk Lee, Madeleine Price Ball, Joseph E. Peterson, Helena Perazich, George Yeung, Jia Liu, Linsu Chen, Michael I. Kennemer, Kaliprasad Pothuraju, Karel Konvicka, Mike Tsoupko-Sitnikov, Krishna P. Pant, Jessica C. Ebert + *et al.*

Affiliations | Contributions | Corresponding authors

*Nature* **487**, 190–195 (12 July 2012) | doi:10.1038/nature11236

# sequencing below the intrinsic error rate of the sequencer

*uses redundant reads of same molecule for error-correction*

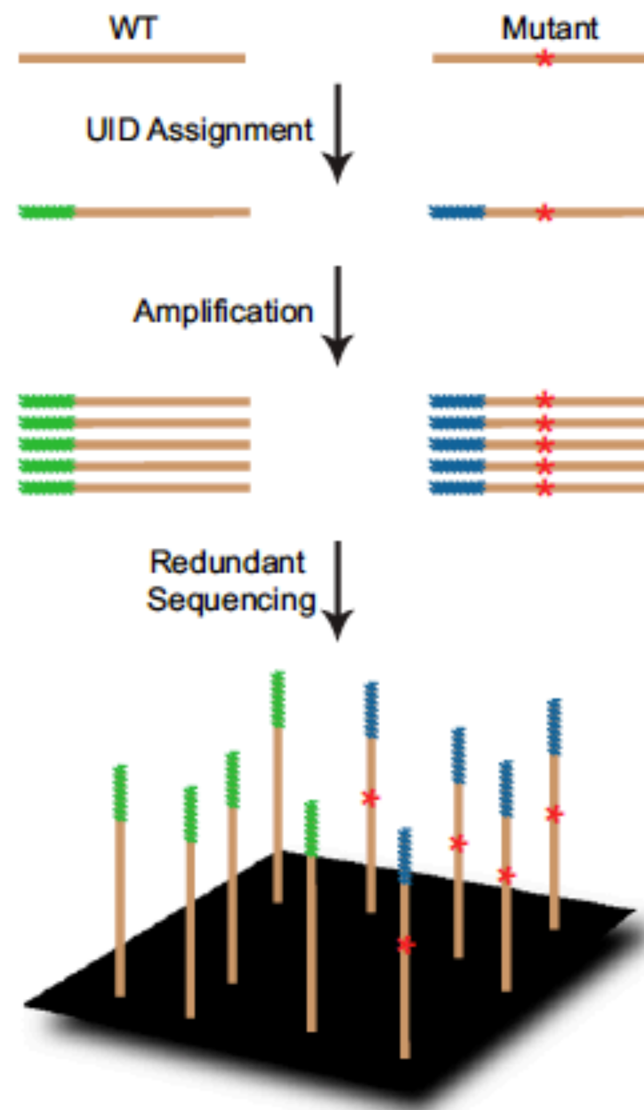



**Fig. 1.** Essential elements of Safe-SeqS. In the first step, each fragment to be analyzed is assigned a unique identification (UID) DNA sequence (green or blue bars). In the second step, the uniquely tagged fragments are amplified, producing UID families, each member of which has the same UID. A supermutant is defined as a UID family in which ≥95% of family members have the same mutation.

**Detection and quantification of rare mutations with massively parallel sequencing**

Isaac Kinde, Jian Wu, Nick Papadopoulos, Kenneth W. Kinzler[1], and Bert Vogelstein[1]