

Creating a framework for mechanistic studies

Frank Pugh



ENCODE PI VISION: (Layer 3)

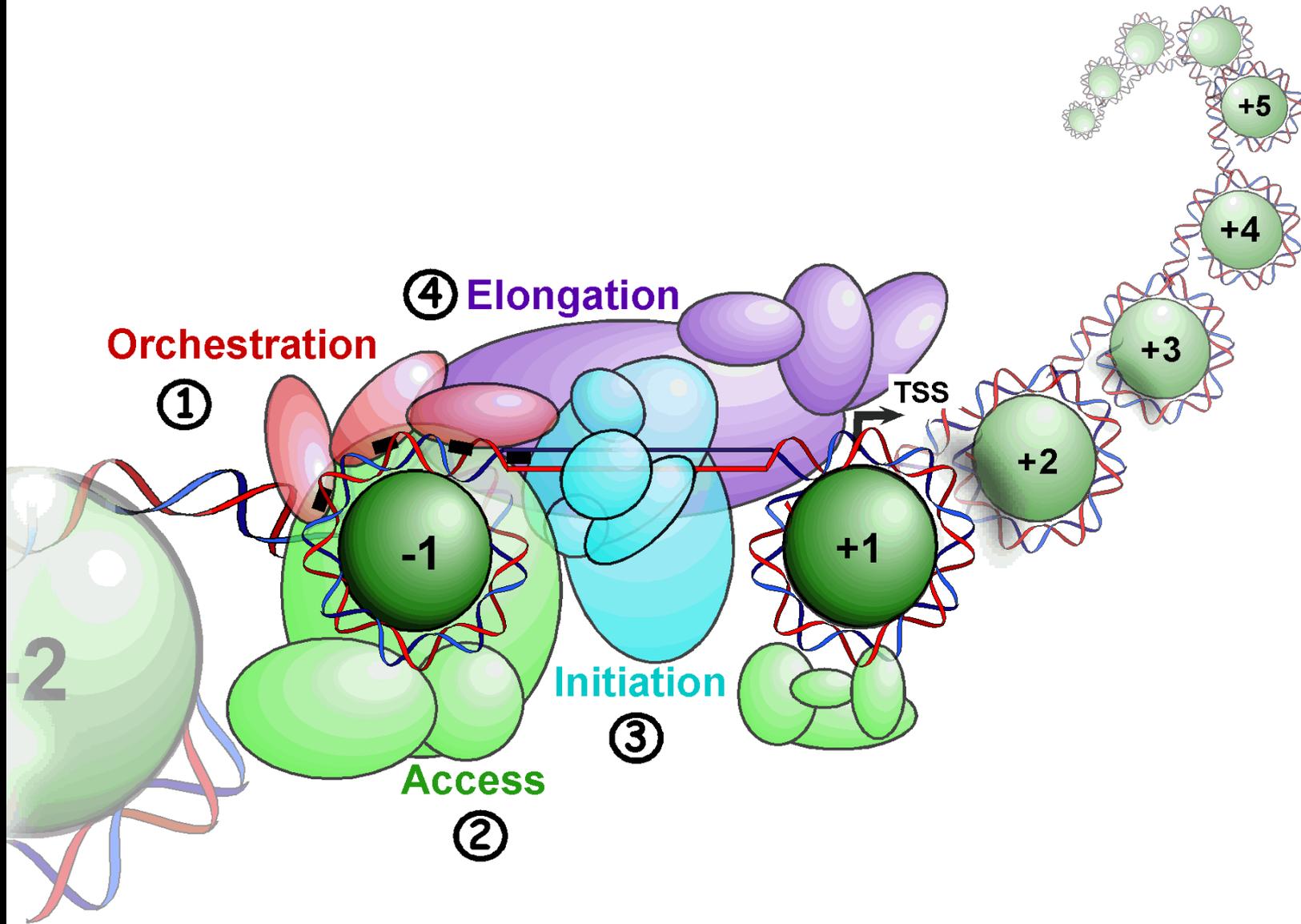
“an encyclopedia, wherein each entry describes ... what and how of each element”



“functional element”

“functional element”

The "other" factors



Creating a framework for mechanistic studies

1. Correlation ✓ “functional assays” & “states” ✗

2. Causation, Structure, & Dynamics

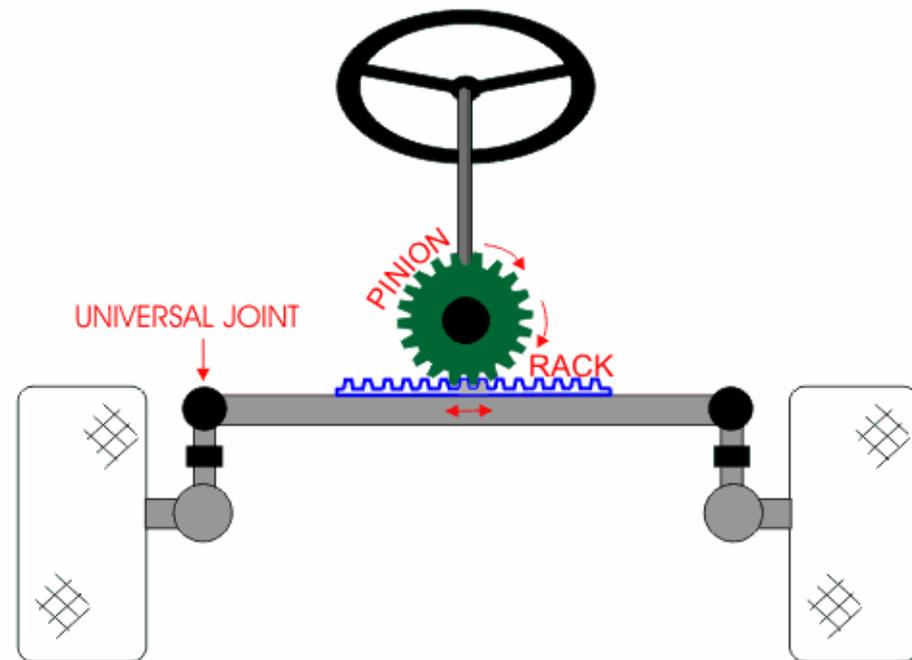
Mutation

Hi-Res assay

Timecourse

3. Mechanism

4. Drug Discovery



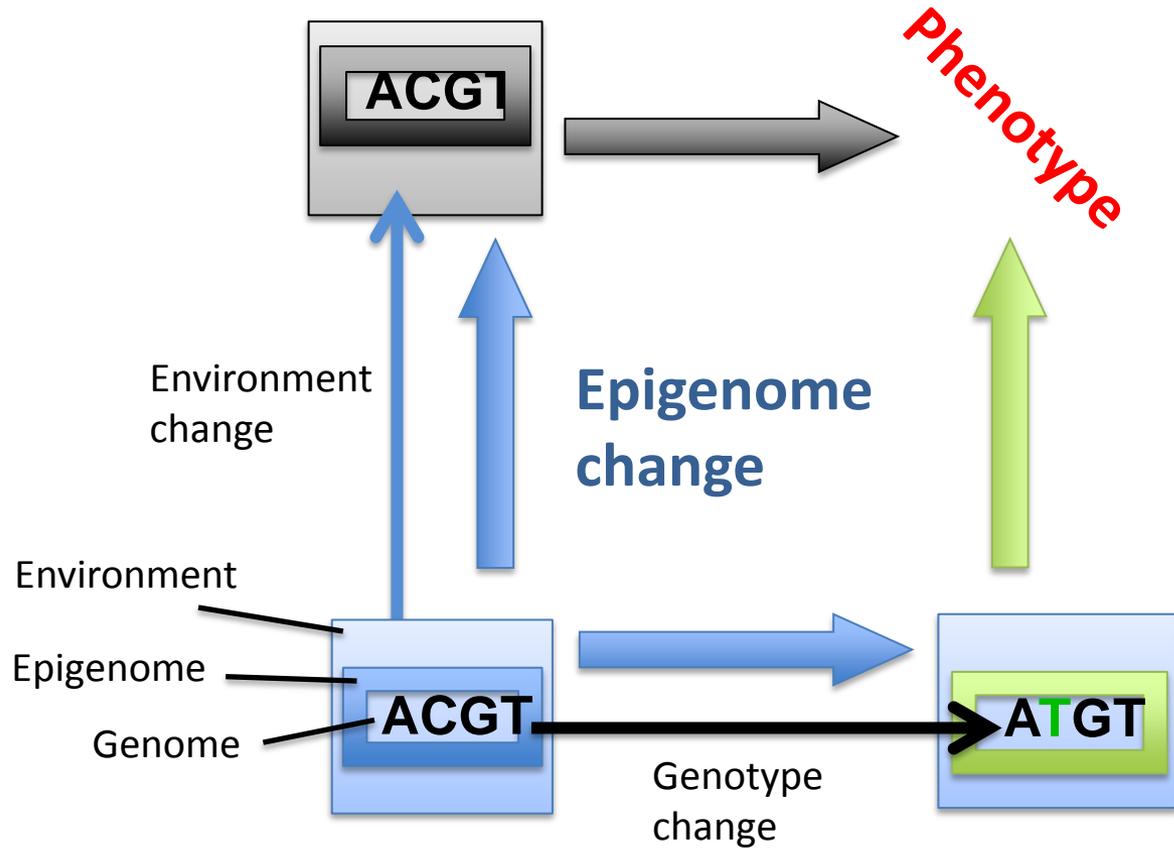
High resolution is good



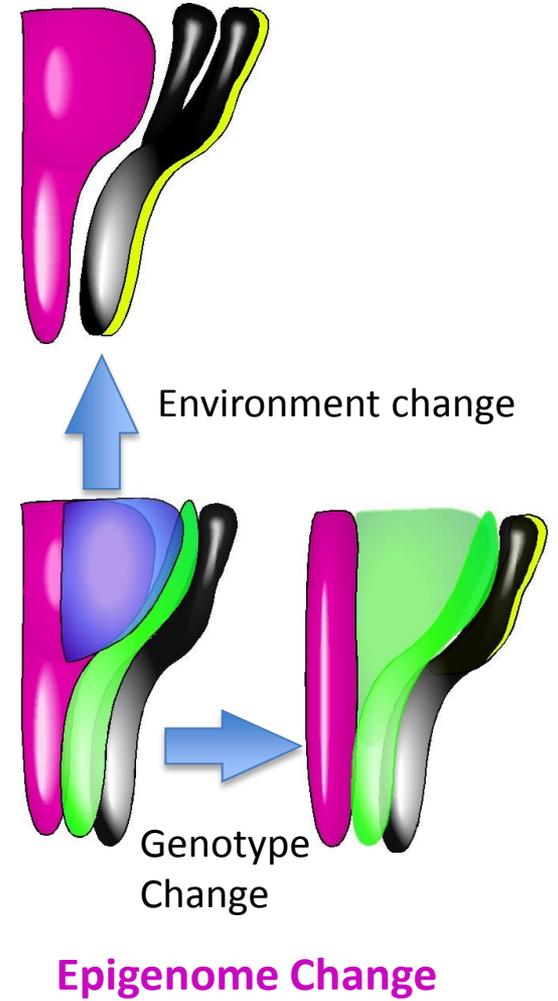
Creating a framework for
mechanistic studies

Proposal

Background



Preliminary results

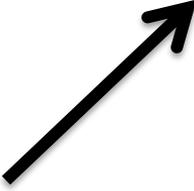
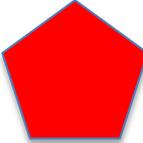
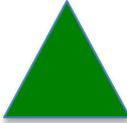


Proposed work

Cell System



Options
Timepoints (dynamics)
CRISPR/Cas9-based deletions (genotype)
-/+ Perturbation (environment)



Hi-throughput & Hi-resolution
Genome-wide assays and analyses

Quality Control
Antibody validation
Genotype validation
Full procedural documentation
Data standards

Feasibility and Budget

Assays	Cell types	Time points	Deletions	Factors	Replicates
9900	10	3	5	22	3
			↓	↘	
			SS TFs		

#	Factors	Assay
1	Chr Accessibility & FP	ATAC-seq
1	RNA	Net/Pro-seq
6	SS TF	ChIP-exo
1	Nucleosomes	Mnase-seq
5	Chromatin Regulators	ChIP-exo
5	Histone marks/variants	ChIP-exo
3	GTFs	ChIP-exo
22	Total	

Budget: \$3M (direct)

\$200/assay x 10,000 assays = \$2M

Fixed costs = \$1M

Community Enablement



Enabling requirements

Affinity capture	Antibody vs others; keeping up the validation
Noise minimization	
Spatial resolution	Exonuclease, MNase, DNase, Tn5, RNA 5' or 3' ends
Native context	
High throughput	True production vs variable compliance
Cell-type uniformity	
Standardization	"line-item" the DCC?
Quantify uncertainty	Objective vs subjective thresholds
Analysis accessibility	

Common questions:

- 1) What gap does your proposed project (s) fill and why is it a high priority?
Correlations → (gap) → **Mechanisms** → **Drugs**
- 2) Why is this project appropriate for NHGRI vs other ICs?
One-size-fits-all infrastructure for unbiased data collection
- 3) What new technological breakthroughs would be transformative?
Automation → **Ross's "800 million" datasets**
- 4) What additional unbiased data generation efforts would facilitate these studies?
High throughput genome alteration
- 5) Would this project benefit from a particular scale and/or organizational structure?
Yes!