Compressed CRISPR Guide Library for Genome-Wide Screening

Aidan Winters
Goodarzi Lab & Gilbert Lab

The Cancer Dependency Map (DepMap)

- Assayed genetic vulnerability for 800 cell lines ~30 cancer types), genome-wide
- Proven to be an extremely useful resource:
  - novel drug targets
  - specificity of vulnerability to cancer type
  - informs novel cancer biology

Genome-scale screens identify factors regulating tumor cell responses to natural killer cells

Michal Sheffer ✉, Emily Lowry, [...] Constantine S. Mitsiades ✉

Nature Genetics **53**, 1196–1206 (2021) | Cite this article

Selective Modulation of a Pan-Essential Protein as a Therapeutic Strategy in Cancer

Clare F. Malone, Neekesh V. Dharia, Guillaume Kugener, Alexandra B. Forman, Michael V. Rothberg, Mai Abdusamad, Alfredo Gonzalez, Miljan Kuljanin, Amanda L. Robichaud, Amy Saur Conway, Joshua M. Dempster, Brenton R. Paolella, Nancy Dumont, Volker Hovestadt, Joseph D. Mancias, Scott T. Younger, David E. Root, Todd R. Golub, Francisca Vazquez, and Kimberly Stegmaier

CANCER RESEARCH | GENOME AND EPIGENOME

Gene Fusions Create Partner and Collateral Dependencies Essential to Cancer Cell Survival

Riaz Gillani[1,2,3,4], Bo Kyung A. Seong[1,2], Jett Crowdis[2,5], Jake R. Conway[2,6], Neekesh V. Dharia[1,2,3,4], Saif Alimohamed[7], Brian J. Haas[2], Kyuho Han[8], Jihye Park[2,5], Felix Dietlein[2,5], Meng Xiao He[2,6], Alma Imamovic[2,5], Clement Ma[1], Michael C. Bassik[8,9,10], Jesse S. Boehm[2], Francisca Vazquez[2], Alexander Gusev[5,11], David Liu[2,5], Katherine A. Janeway[1,3,4], James M. McFarland[2], Kimberly Stegmaier[1,2,3,4], and Eliezer M. Van Allen[2,5,12]

CellPress

Cell

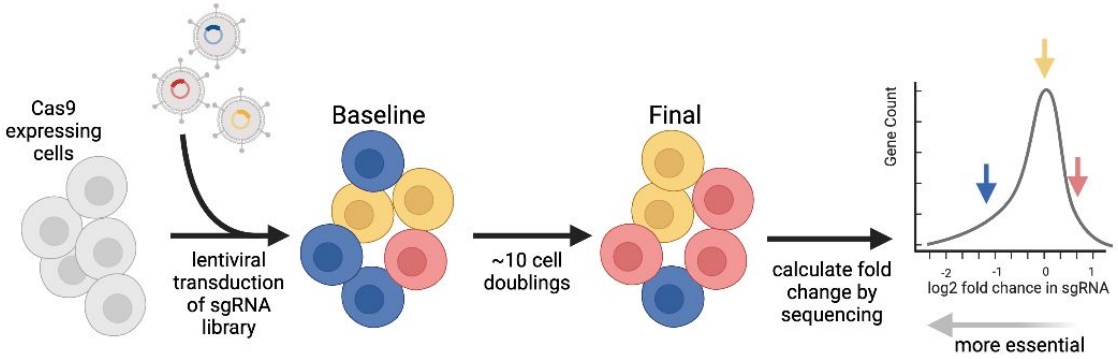
Leading Edge

Perspective

An expanded universe of cancer targets

William C. Hahn,[1,*] Joel S. Bader,[2] Theodore P. Braun,[4] Andrea Califano,[3] Paul A. Clemons,[10] Brian J. Druker,[4] Andrew J. Ewald,[16] Haian Fu,[5] Subhashini Jagu,[13] Christopher J. Kemp,[6] William Kim,[11] Calvin J. Kuo,[7] Michael T. McManus,[8] Gordon B. Mills,[9] Xiulei Mo,[5] Nidhi Sahni,[15] Stuart L. Schreiber,[10] Jessica A. Talamas,[1] Pablo Tamayo,[11] Jeffrey W. Tyner,[14] Bridget K. Wagner,[10] William A. Weiss,[12] Daniela S. Gerhard,[13] and the Cancer Target Discovery and Development Network

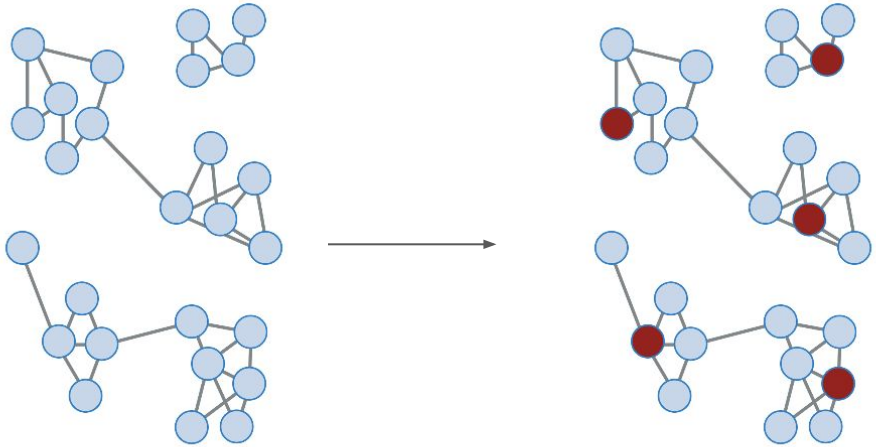
Limited by Scalability of CRISPR screening



- CRISPR (Cas9) Loss-of-function screens
  - Requires 20-200 million slides
- Severely limits utility for:
  - screens *in vivo*
  - screens across large panels of drugs

Goal: A Representative Gene Set for Genetic Dependencies

- High functional redundancy across genes
- “Representative” gene set that captures the genome-wide landscape
- Demonstrated with L1000 (LINCS) gene set
    - 1000 transcripts reconstruct 87% of non measured transcripts



My Approach: Concrete VAE (CAVE)

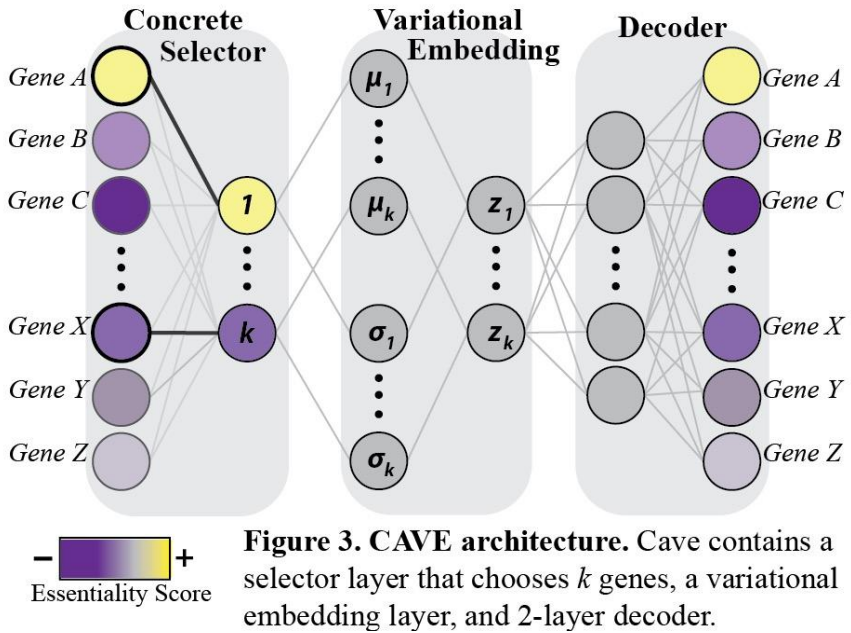
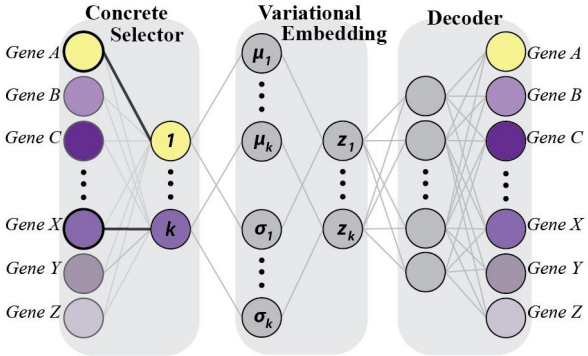


Figure 3. CAVE architecture. Cave contains a selector layer that chooses k genes, a variational embedding layer, and 2-layer decoder.

- Recently, **efficient feature selection** by concrete relaxation (gumbel softmax) for VAE
- Built for task of compression
- Learns probability distribution of latent layer
- **Multitasking**

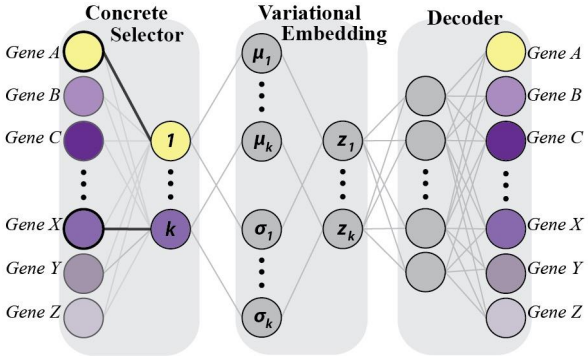
How To: Hybrid *in silico* CRISPR Screens

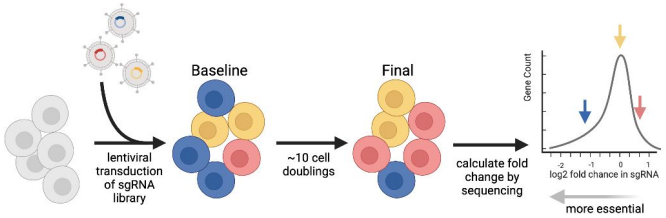
1. Selection of gene set for ‘compressed’ screening from genome-wide training data



How To: Hybrid *in silico* CRISPR Screens

1. Selection of gene set for ‘compressed’ screening from genome-wide training data

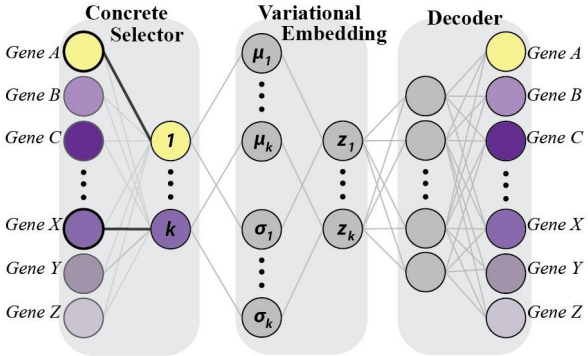


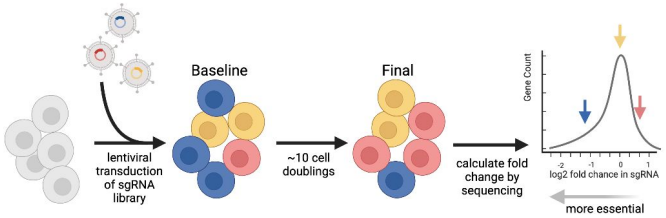


2. Perform compressed screen with guide library targeting selected genes only

How To: Hybrid *in silico* CRISPR Screens

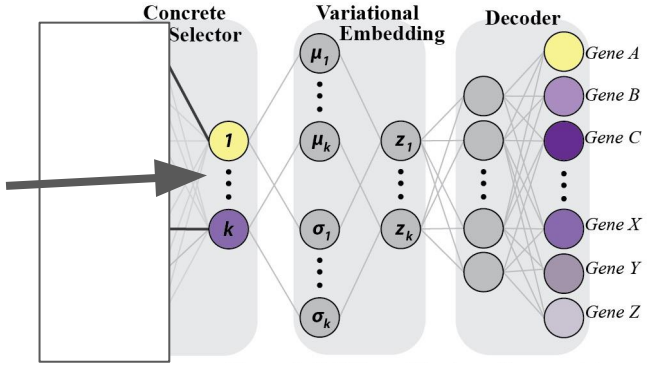
1. Selection of gene set for ‘compressed’ screening from genome-wide training data





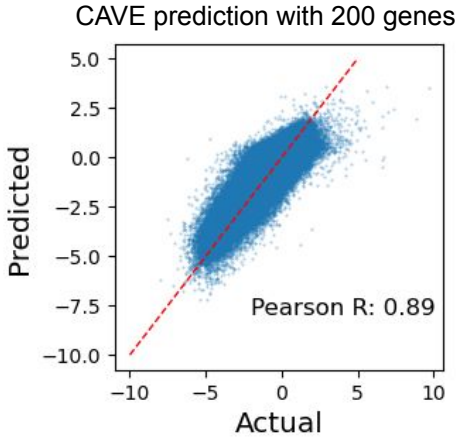
2. Perform compressed screen with guide library targeting selected genes only

3. Predict non-measured genes from results of compressed screen

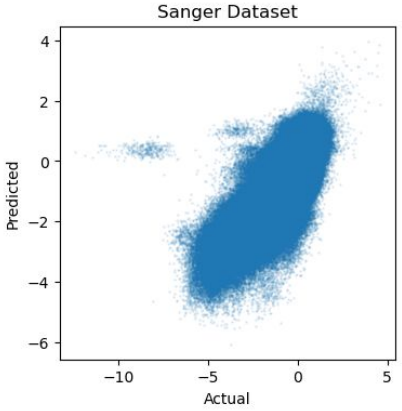


200 Genes Infer Genome-Wide Dependencies

- Predicts well with 200 genes (10x reduction in library size)
  - 0.89 R^2
  - 0.99 AUPRC, 0.97 AUPRC on selectively essential

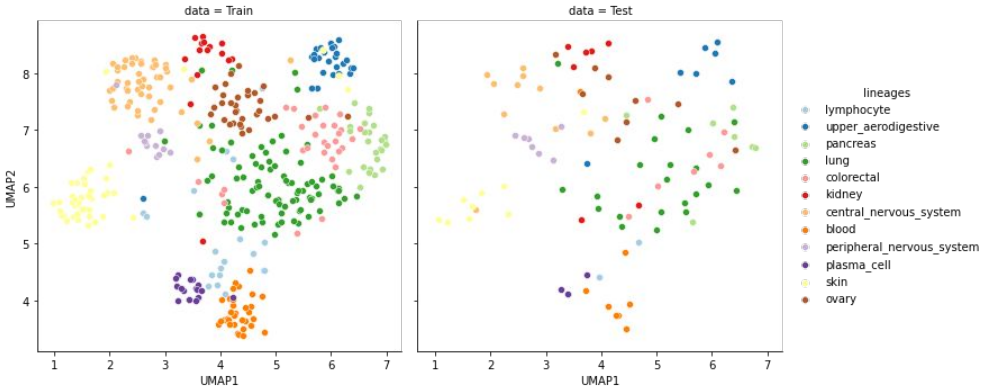


- Generalizes to new technical dataset
  - Project Score: ~350 cell lines, different experimental setup
  - 0.79 R^2

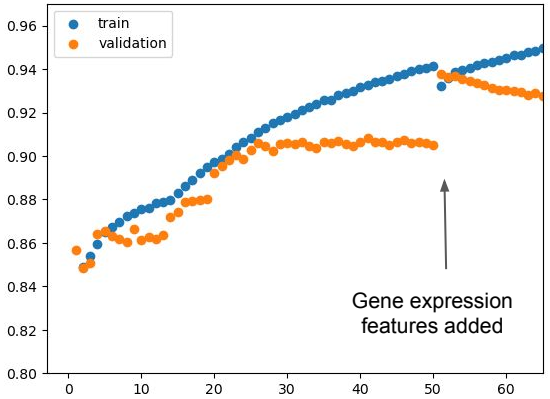


Improving CAVE with Lineage Specificity and Gene Expression

- Lineage prediction task increases prediction 0.05 - 0.1 R^2



- Adding Gene expression increases prediction ~0.05 R^2



Future Directions

- Test generalizability to CRISPRi LOF screens
- Experimental validation:
  - Run a compressed screen in tandem with genome-wide screen
- Test compressed screening approach across panel of drug conditions
- Test compressed screen in mouse xenograft

Thanks!

Goodarzi Lab

Mehran Karimzadeh

Gilbert Lab