13 June 2024

A deep learning model of tumor cell architecture elucidates response and resistance to CDK4/6 inhibitors

Sungjoon Park et. al

Chanhee Lee

Interdisciplinary Program in Bioinformatics, Seoul National University, Seoul, Republic of Korea

nature cancer

Article

https://doi.org/10.1038/s43018-024-00740-1

A deep learning model of tumor cell architecture elucidates response and resistance to CDK4/6 inhibitors

Received: 16 May 2022

Accepted: 7 February 2024

Published online: 05 March 2024

Sungjoon Park[1,6], Erica Silva[2,6], Akshat Singhal[3,6], Marcus R. Kelly[1,4], Kate Licon[1], Isabella Panagiotou[1], Catalina Fogg[1], Samson Fong[5], John J. Y. Lee[1], Xiaoyu Zhao[1], Robin Bachelder[1], Barbara A. Parker[1,4], Kay T. Yeung[1,4] & Trey Ideker[1,3,4,5] ✉

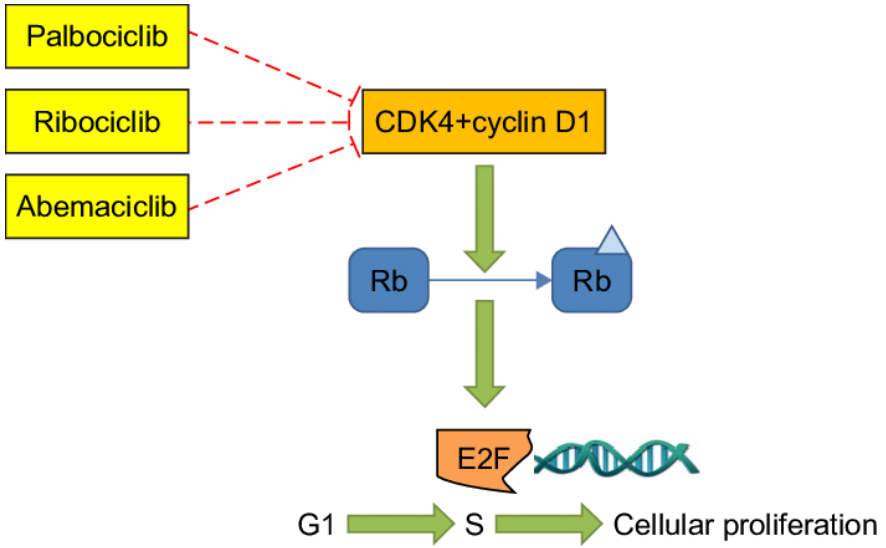
Outline

- **Introduction**
- **Results**
  - ✓ Implementation of a cancer-oriented VNN
  - ✓ Evaluation of prediction performance
  - ✓ Translation to patient-derived xenografts and patients
  - ✓ Interpreting the model to identify important protein assemblies
  - ✓ Systematic validation of core assemblies by loss-of-function screens
  - ✓ Exploration of gain-of-function alterations in a histone transcriptional assembly
- **Discussion**

Introduction

Understanding CDK4/6 Inhibitors and Resistance in Cancer Therapy

- Cell-cycle activation and sustained proliferation **"hallmarks of cancer"**
- Cyclin-dependent kinases 4 and 6 (CDK4/6) trigger cells to pass the G1/S cell-cycle restriction point
- Inhibiting these kinases has been of high interest in cancer drug development
  - palbociclib, ribociclib and abemaciclib
  - effective in metastatic breast cancer



Introduction

Understanding CDK4/6 Inhibitors and Resistance in Cancer Therapy

- **Limitations in Efficacy:** Less than 50% tumor response rate as first-line therapy, with prevalent resistance development.
- **Molecular Markers of Resistance**
  - loss-of-function in CDK pathway genes
  - gain-of-function in pro-growth genes.
  - RB1 mutation bears strongest burden of evidence.
- Yet molecular markers show inconsistent reliability in trials.

Aim : Understanding resistance mechanisms of CDK4/6 inhibitors.

Introduction

Enhancing Precision Medicine with ‘Visible’ neural network (VNN)

- **Visible Neural Networks (VNNs):** incorporates knowledge maps of cellular components and functions to create models that are both predictive and interpretable.
- Elmarakeby et al. suggests that MDM4 is linked to antiandrogen resistance in prostate cancer, predicted by genetic changes in the MDM-TP53 pathway.

Introduction

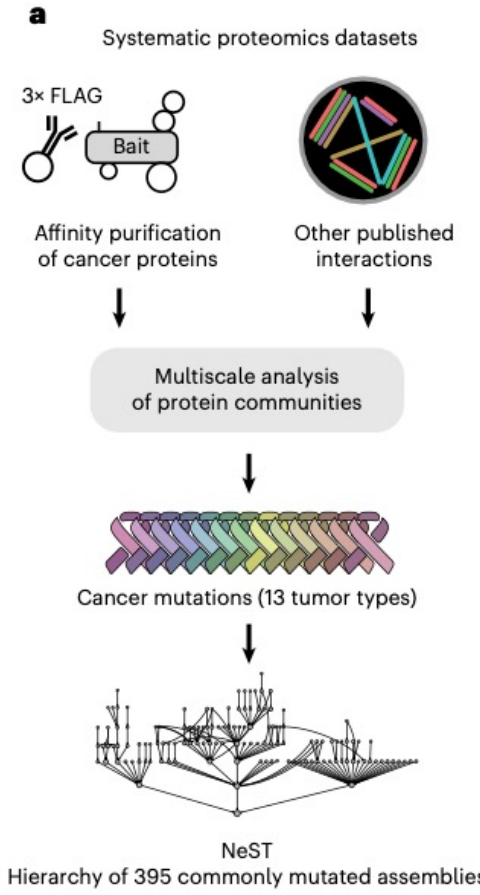
Enhancing Precision Medicine with 'Visible' neural network (VNN)

- Thus far, VNN models have been structured using Gene Ontology or Reactome
- However, these databases are not explicitly designed for capturing molecular pathways of cancer

Introduction

Enhancing Precision Medicine with ‘Visible’ neural network (VNN)

- To define and discover cancer mechanisms systematically, we recently developed a hierarchical map of multiprotein assemblies called **NeST**
- NeST was defined as the final hierarchy of 395 assemblies found to be under significant selection pressure for somatic mutations in one or more adult tumor types (pan cancer)



Introduction

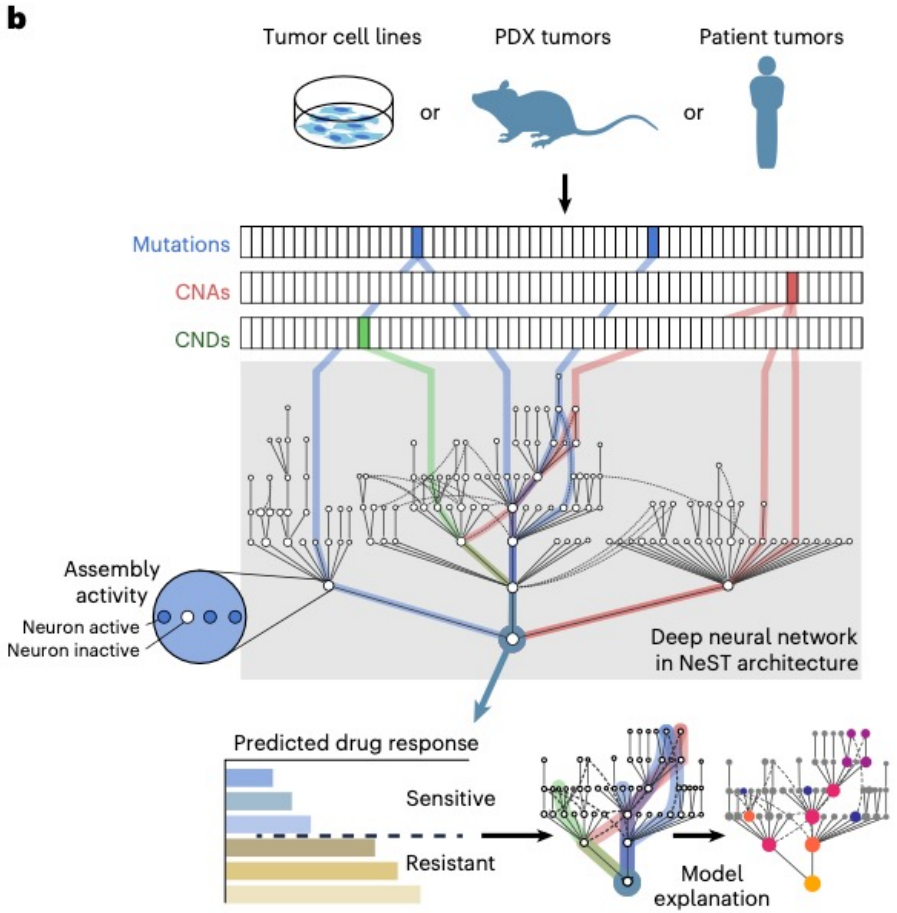
Enhancing Precision Medicine with 'Visible' neural network (VNN)

- Use NeST map as the foundation for a VNN to understand how patterns of genetic alterations govern the tumor response to CDK4/6 inhibition.
- This model is …
    1. Functionally predictive of palbociclib treatment outcomes
    2. Can be structurally interpreted : reveal focal set of protein assemblies on which common and rare cancer mutations converge to affect drug resistance or sensitivity

Results

Implementation of a cancer-oriented VNN

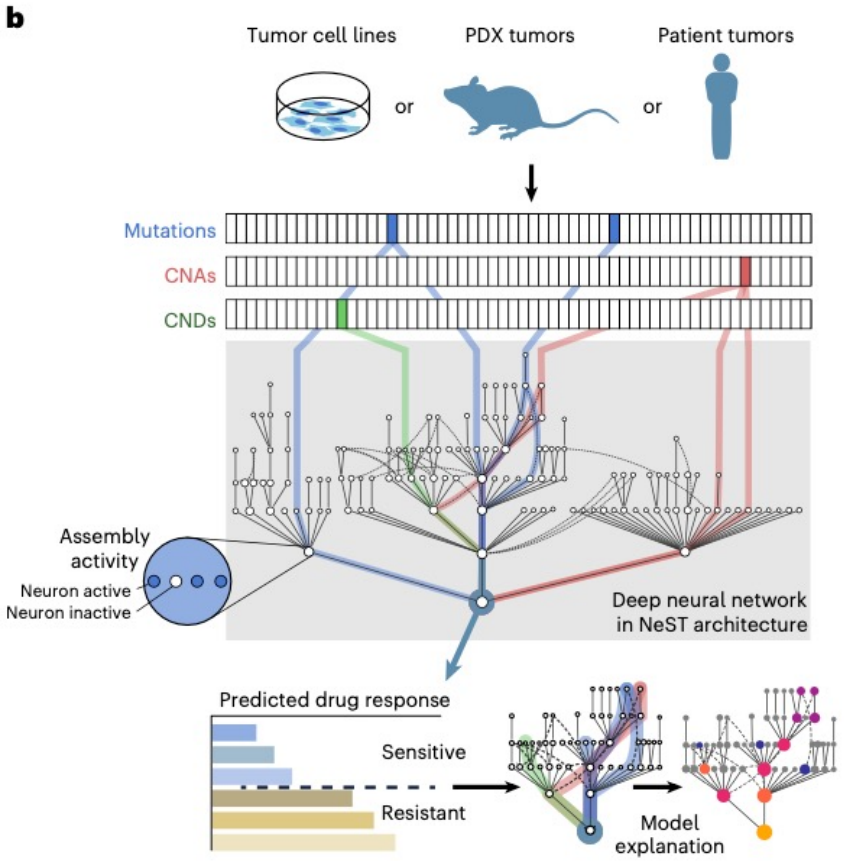
- **Feature selection** : 718 genes from clinical cancer panels
  - FoundationOne CDx, Tempus xT, and Project GENIE
- NeST-VNN uses a hierarchy of 131 protein assemblies from these genes to predict drug responses
- **Data** : Mutations, CNAs, and CNDs. (three binary input features per gene)



Results

Implementation of a cancer-oriented VNN

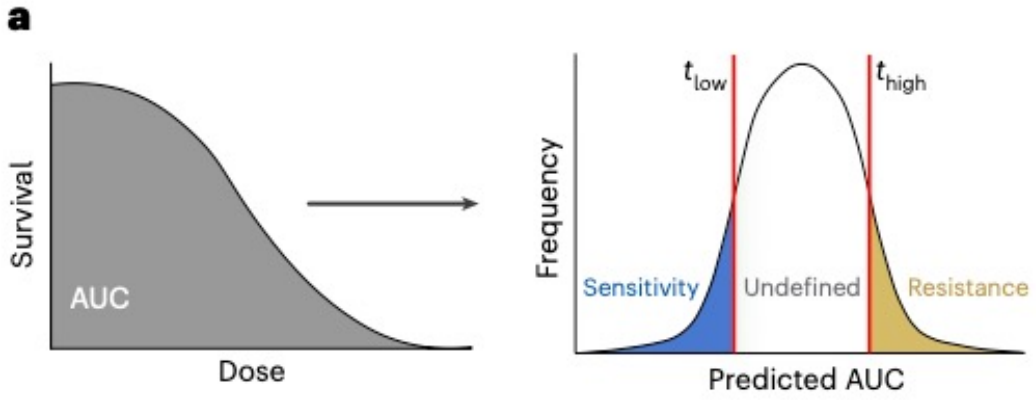
- Trained with drug response data from 1,244 tumor cell lines
  - harmonizing **CTRP** and **GDSC** databases, focusing on responses to palbociclib.
- Comparative benchmarking included 50 non-CDK drugs tested on over 200 cell lines
  - ensure the model's predictive accuracy for drug sensitivity and resistance.



Results

Evaluation of prediction performance

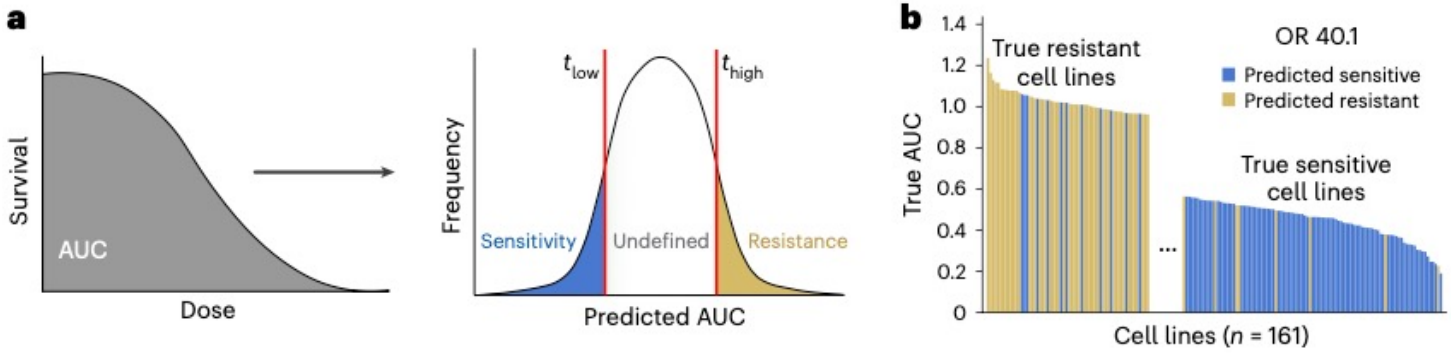
- Developed NeST-VNN models for palbociclib and 50 benchmark drugs using standard neural network for predicting AUC of dose-response curve.
- Employed nested five-fold cross-validation for training, ensuring no test data contamination and optimal utilization of data for testing.



Results

Evaluation of prediction performance

- NeST-VNN generally matched or exceeded the performance of state-of-the-art models, being the top performer for over half of the drugs.
- For palbociclib, NeST-VNN outshone ElasticNet and ANN models and was on par with RF, indicating robust predictive capability.
- Implemented AUC thresholding to categorize tumor responses as 'sensitive,' 'resistant,' or 'undefined,' with high discriminative accuracy for stringent thresholds.

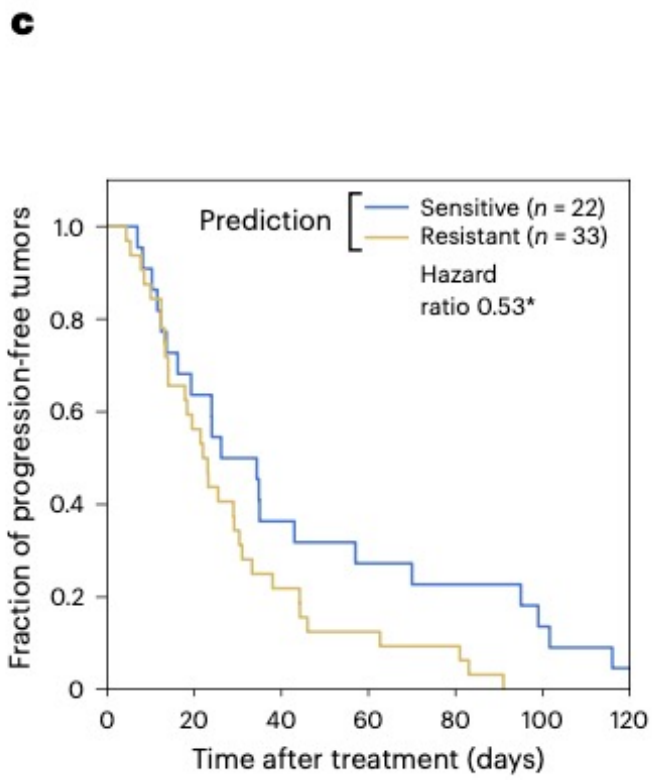


Results

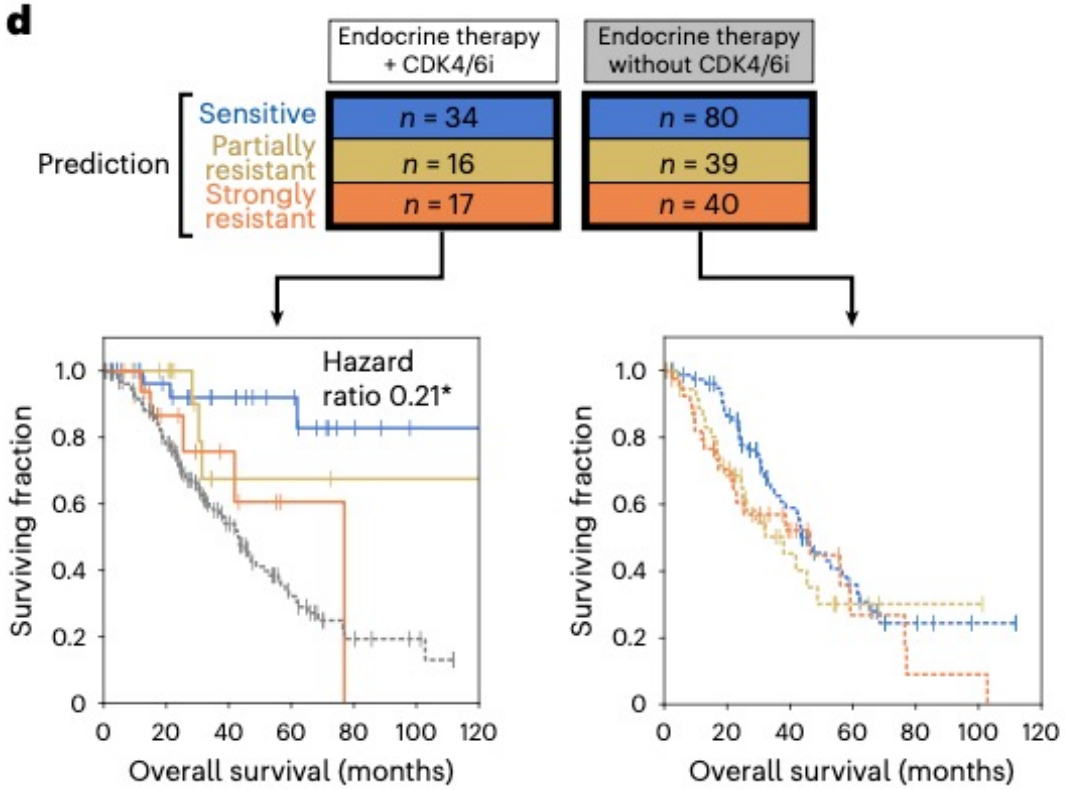
Translation to patient-derived xenografts and patients

- NeST-VNN tested on 172 **PDX samples** treated with **ribociclib** (CDK4/6i), predicting longer PFS in samples classified as sensitive. (Why?)
- In a cohort of 226 **breast cancer patients**, predictions of sensitivity to CDK4/6i aligned with significantly longer survival outcomes.
- The model distinguished survival outcomes better than single-gene markers for palbociclib resistance or sensitivity previously used in clinical trials.
- Predictions were specific to CDK4/6i response, not general survival prognostics, with no significant survival difference observed in patients untreated with CDK4/6i.

Results



Survival curve analysis for PDX samples.



Survival curve analysis for GENIE clinical trial patients

Results

NeST-VNN palbociclib model

1. translates to the population of patients with breast cancer
2. specifically predictive of response rather than generally prognostic of patient survival.

Results

Interpreting the model to identify important protein assemblies

- Computed a quantitative importance score ‘relative local improvement in predictive power’ for each assembly
- Of 33 assemblies that were of high importance for palbociclib response prediction in cell lines (importance ≥ 0.5), we focused on eight distinct minimally overlapping assemblies whose importance scores remained significant under multiple-hypothesis correction

Results

Interpreting the model to identify important protein assemblies

- Each assembly was modeled using linear regression, with the aim of evaluating how well its NeST-VNN neuron values capture the NeST-VNN overall drug response prediction.
- Each assembly k was assigned a $g \times N$ matrix P_k, where g is the number of samples and N is the number of neurons. P_k was then used in a linear ridge regression[73] model to predict the NeST-VNN drug response D, creating models M_1, M_2, ..., M_k.

Results

Interpreting the model to identify important protein assemblies

- The following function was minimized for each model:

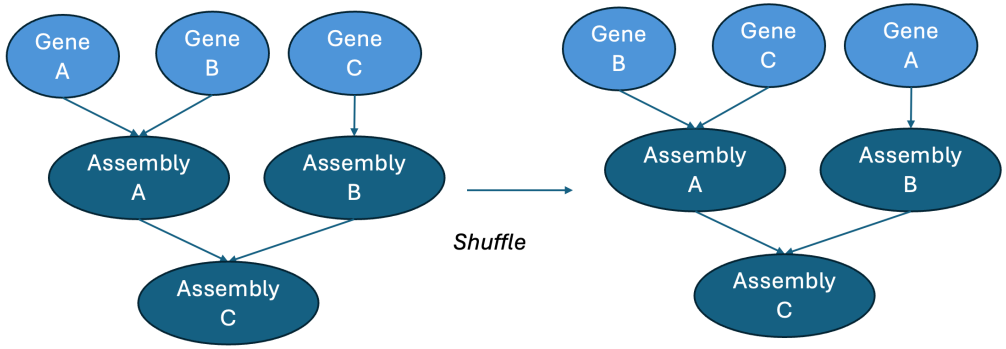
$$\min_{\mathbf{w}} ||P_k\mathbf{w} - D||_2^2 + \alpha||\mathbf{w}||_2^2$$

where $\mathbf{w}$ is a vector of the coefficients of length N and α imposes an L2 penalty on coefficient complexity. Assembly 'importance' (Fig. 3 and Extended Data Figs. 3 and 5) is the Spearman correlation (ρ) between M_k and D. The mean correlation of the five NeST-VNN models was reported.

Results

Interpreting the model to identify important protein assemblies

- To assess statistical significance, we generated a null distribution of assembly importance scores, as follows.
- We randomly rearranged gene assembly memberships in the NeST-VNN while preserving the assembly size and parent–child relationships.



Results

Interpreting the model to identify important protein assemblies

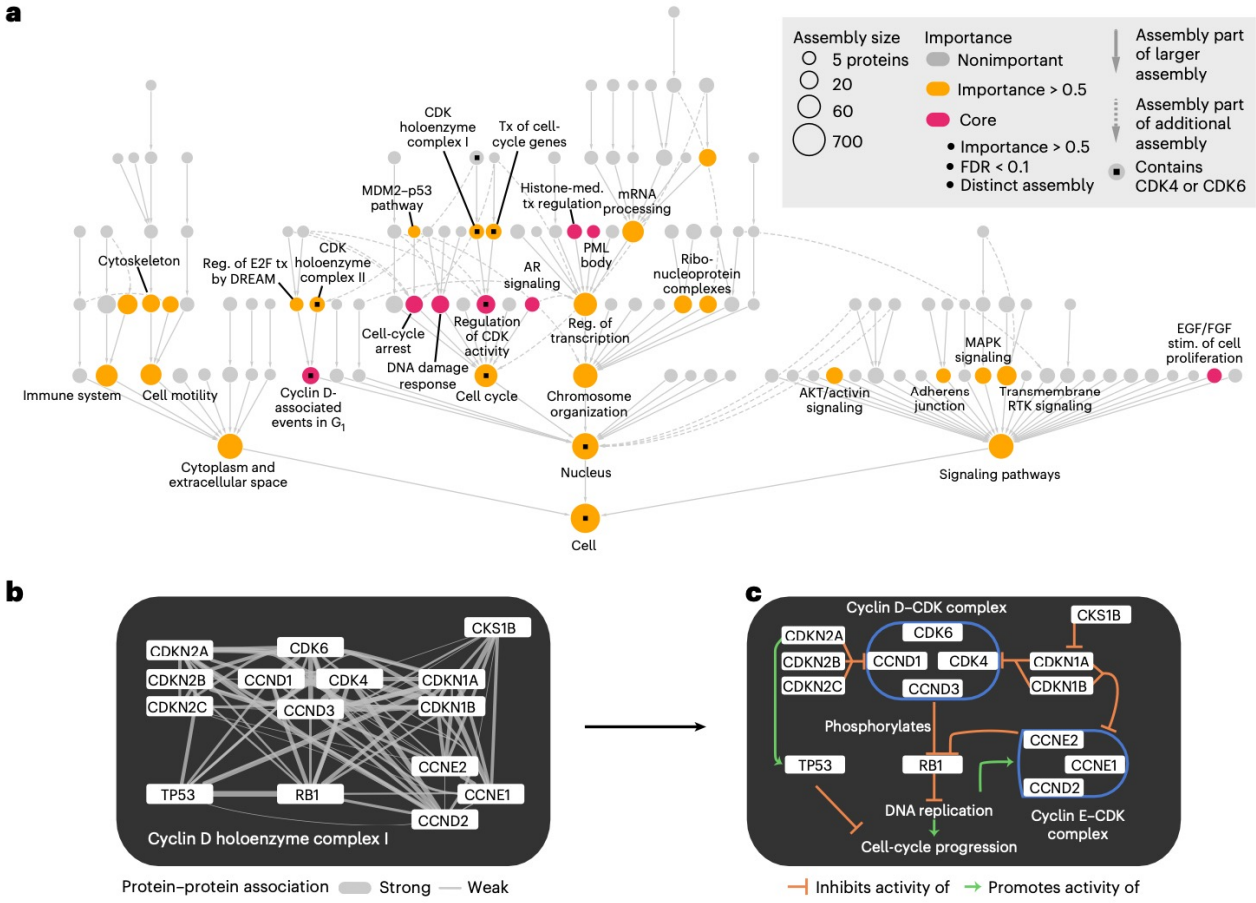
- We trained 500 null models with these random rearrangements and calculated assembly importance for each null.
- One-tailed t tests were used to evaluate whether the assembly importance scores from the five NeST-VNN models were greater than the assembly importance scores from the nulls, with a Benjamini–Hochberg control for false discovery rate

Results

Interpreting the model to identify important protein assemblies

- One-tailed t tests were used to evaluate whether the assembly importance scores from the five NeST-VNN models were greater than the assembly importance scores from the nulls, with a Benjamini–Hochberg control for false discovery rate
- Finally, we defined ‘core assemblies’ as those with an importance score of ≥0.5 and an FDR of ≤0.1, while excluding less important redundant assemblies (Jaccard similarity > 0.5).

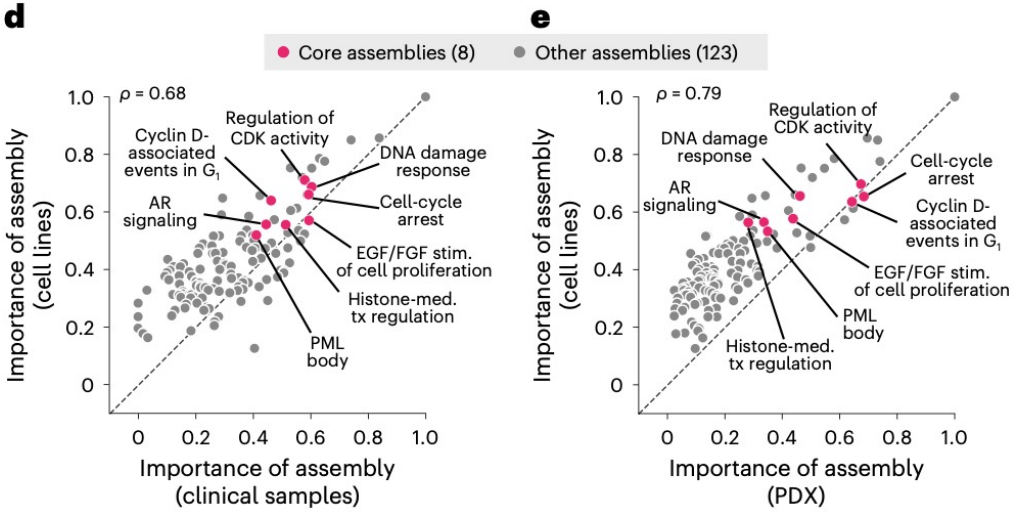
Results


a
Assembly size
5 proteins
20
60
700
Importance
Nonimportant
Importance > 0.5
Core
• Importance > 0.5
• FDR < 0.1
• Distinct assembly
Assembly part of larger assembly
Assembly part of additional assembly
Contains CDK4 or CDK6
CDK holoenzyme complex I
Tx of cell-cycle genes
MDM2–p53 pathway
Histone-med. tx regulation
mRNA processing
PML body
Ribo-nucleoprotein complexes
Cytoskeleton
Reg. of E2F tx by DREAM
CDK holoenzyme complex II
AR signaling
Cell-cycle arrest
Regulation of CDK activity
Reg. of transcription
DNA damage response
Cell cycle
Chromosome organization
MAPK signaling
EGF/FGF stim. of cell proliferation
Immune system
Cell motility
Cyclin D-associated events in G1
AKT/activin signaling
Adherens junction
Transmembrane RTK signaling
Cytoplasm and extracellular space
Nucleus
Signaling pathways
Cell
b
CKS1B
CDKN2A
CDK6
CDKN2B
CCND1
CDK4
CDKN1A
CDKN2C
CCND3
CDKN1B
CCNE2
TP53
RB1
CCNE1
CCND2
Cyclin D holoenzyme complex I
Protein–protein association
Strong
Weak
c
Cyclin D–CDK complex
CKS1B
CDKN2A
CDK6
CDKN2B
CCND1
CDK4
CDKN1A
CDKN2C
CCND3
CDKN1B
Phosphorylates
CCNE2
TP53
RB1
CCNE1
CCND2
DNA replication
Cyclin E–CDK complex
Cell-cycle progression
Inhibits activity of
Promotes activity of

Results

Interpreting the model to identify important protein assemblies

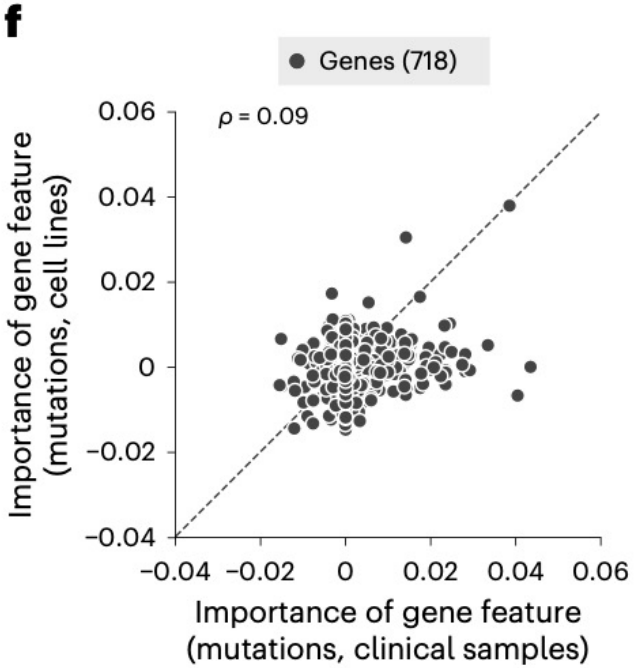
- For all drug models, assembly importance tended to increase with size and depth in the hierarchy, reflecting the progressive integration of genetic information.
- Assembly importance was similar between cell lines and patient tumors (Fig. 3d) or PDX samples (Fig. 3e).



Results

Interpreting the model to identify important protein assemblies

- In contrast, little correlation was observed between cell lines and clinical samples on individual gene mutations

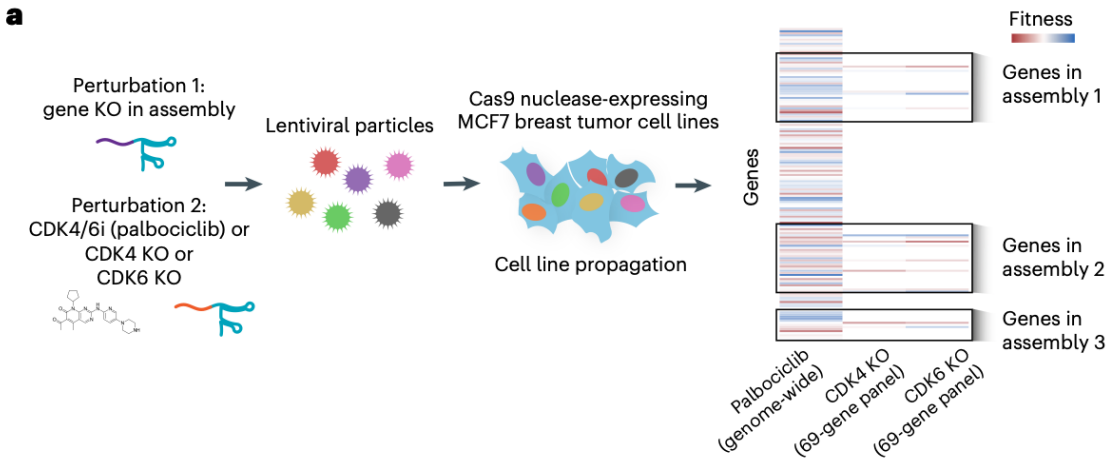


Most individual genetic alterations are rare, with variable incidence across contexts, but effects on protein assemblies can be substantially more stable.

Results

Systematic validation of core assemblies by loss-of-function screens

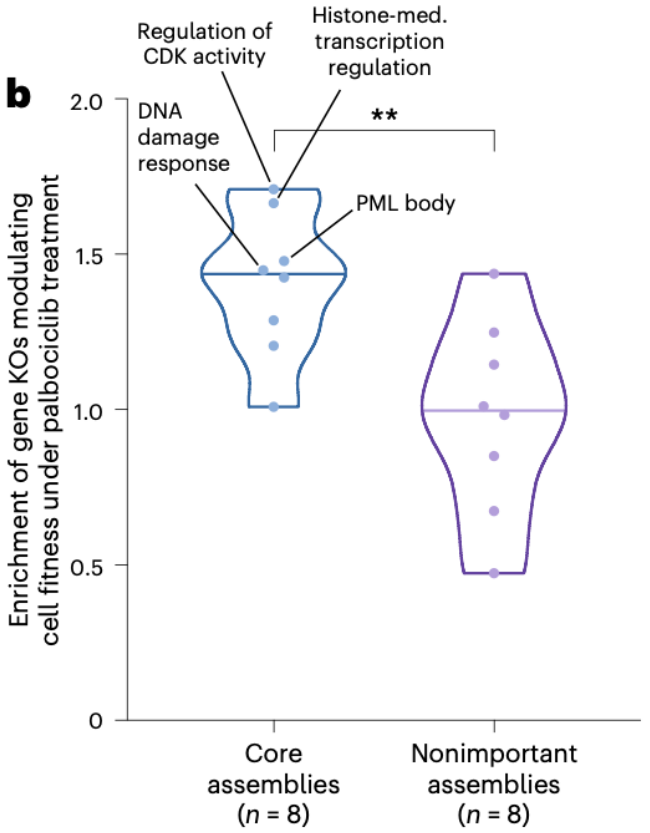
- Validated palbociclib response predictions using two CRISPR loss-of-function screens: a chemogenetic genome-wide KO screen and a dual CRISPR screen targeting CDK4/6.
    1. **Chemogenetic Screen**: This involved genome-wide knockout (KO) of single genes in combination with palbociclib treatment.
    2. **De Novo Dual CRISPR Screen**: This involved KO of genes in selected assemblies paired with a second KO targeting CDK4 or CDK6.



Results

Systematic validation of core assemblies by loss-of-function screens

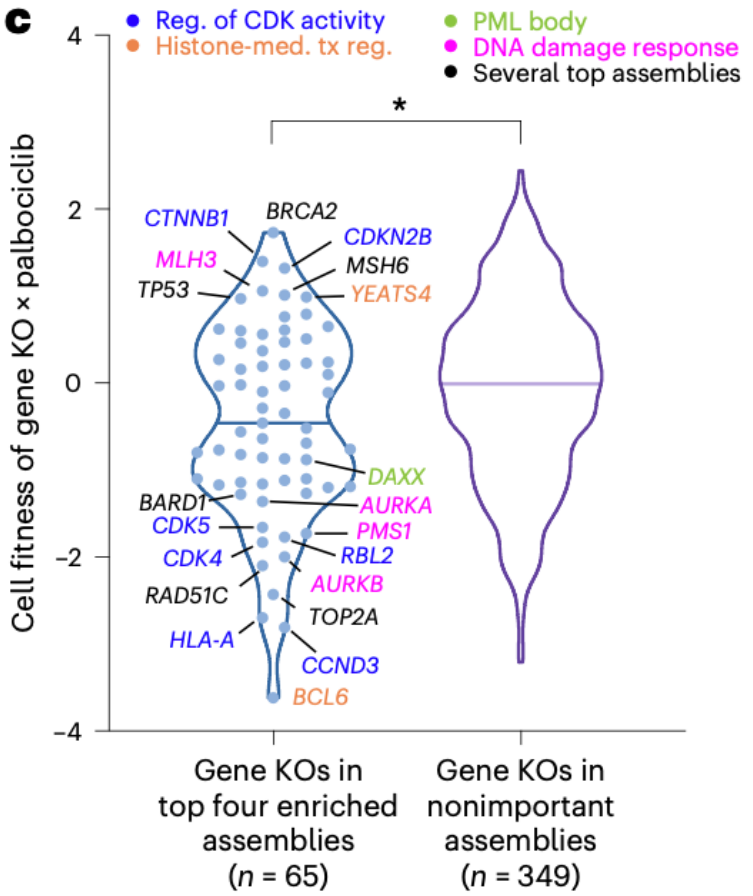
- The enrichments of the eight core assemblies tended to be significantly higher than those of nonimportant controls(**Fig 4b**)



Results

Systematic validation of core assemblies by loss-of-function screens

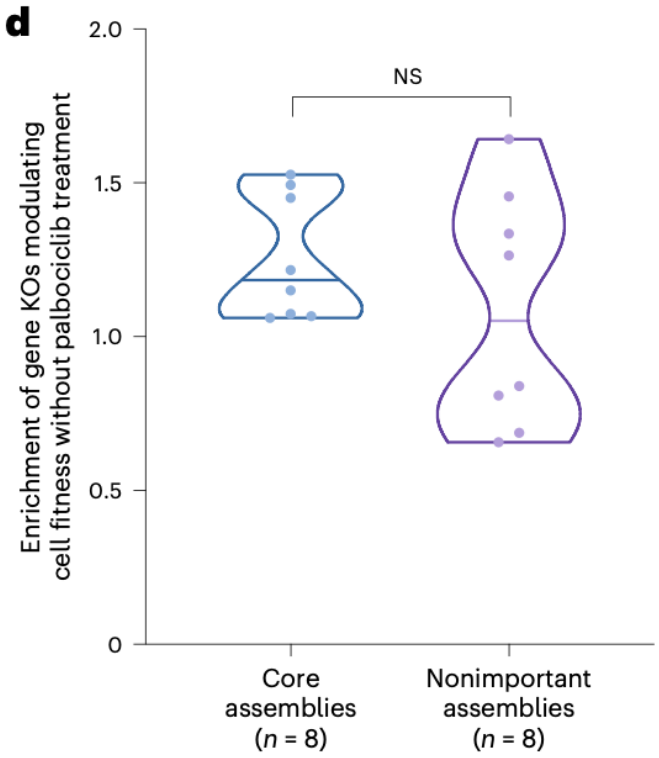
- Gene KOs within these assemblies included those causing extreme loss-of-fitness (e.g., BCL6, CCND3, CDK4) Some resulted in gain of fitness (e.g., BRCA2, CDKN2B).(**Fig 4c**)



Results

Systematic validation of core assemblies by loss-of-function screens

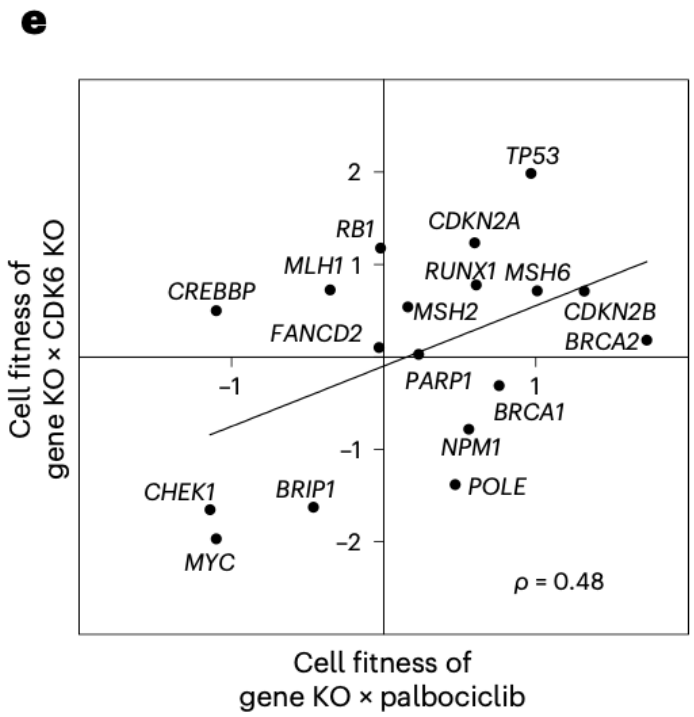
- No enrichment for gene KOs in the absence of palbociclib treatment, highlighting gene–drug interaction effects. (**Fig 4d**)



Results

Systematic validation of core assemblies by loss-of-function screens

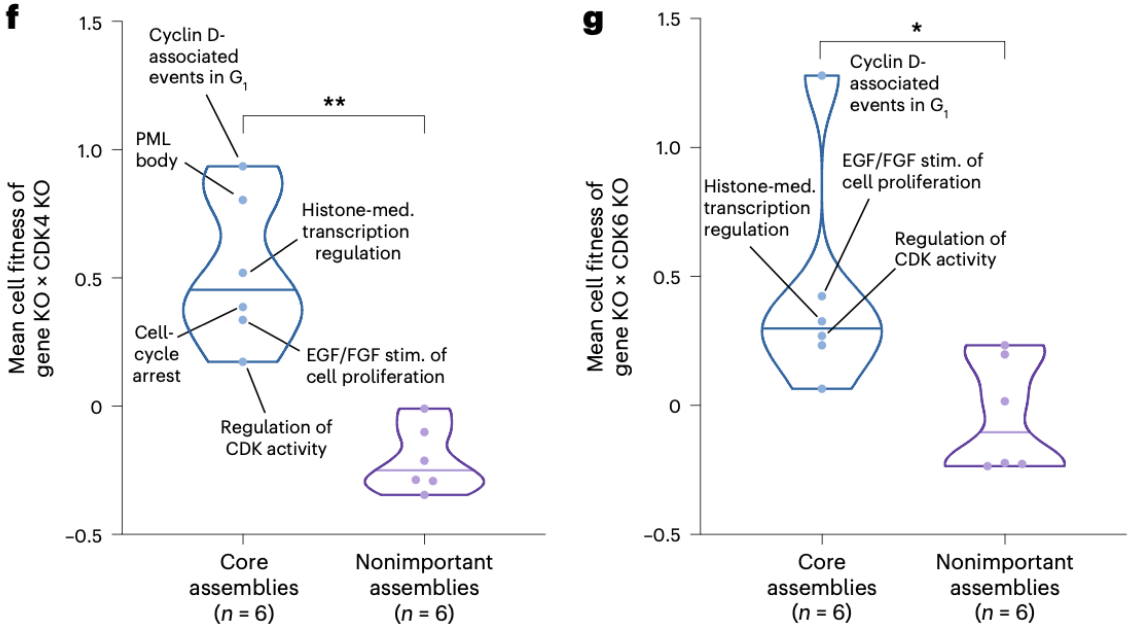
- The dual CRISPR KO screen confirmed the fitness effects from the chemogenetic screen, underscoring the role of core assemblies in cell growth during CDK4/6 inhibition. (**Fig 4e**)



Results

Systematic validation of core assemblies by loss-of-function screens

- Disruptions in all six of the core assemblies with sufficient coverage in our gene KO panel displayed a trend toward increased cell fitness (**Fig 4f,g**)



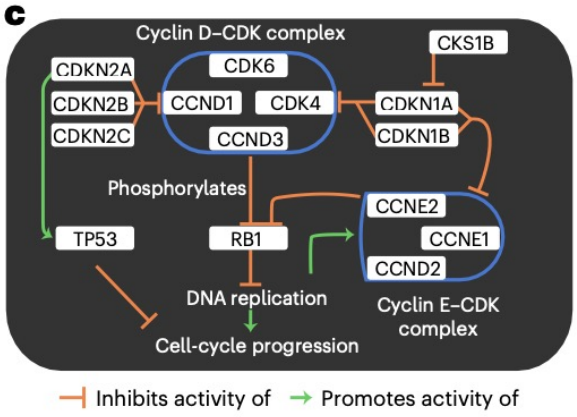
Results

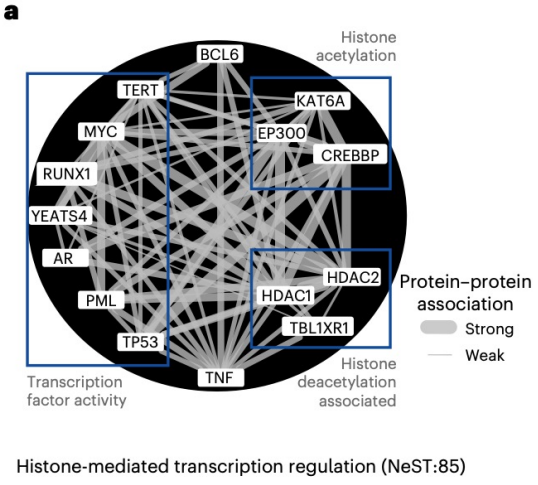
Taken together, these results indicate that engineered genetic disruptions in protein assemblies identified by NeST-VNN can influence tumor cell growth in the setting of CDK4/6 inhibition, whether such inhibition is induced by a drug (Fig. 4b) or CDK4/6 KO (Fig. 4f,g).

Results

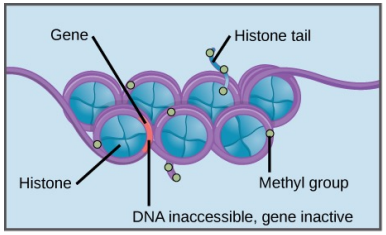
Exploration of gain-of-function alterations in a histone transcriptional assembly

- How does CDK4/6 and the G1/S transcriptional program interact with other cell functions?
- **Key Role of NeST:85 Assembly**: Central to histone modification, affecting the CDK4/6 inhibitor response.

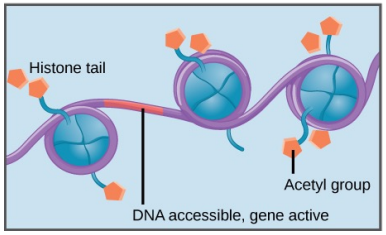




Histone-mediated transcription regulation (NeST:85)



Methylation of DNA and histones causes nucleosomes to pack tightly together. Transcription factors cannot bind the DNA, and genes are not expressed.

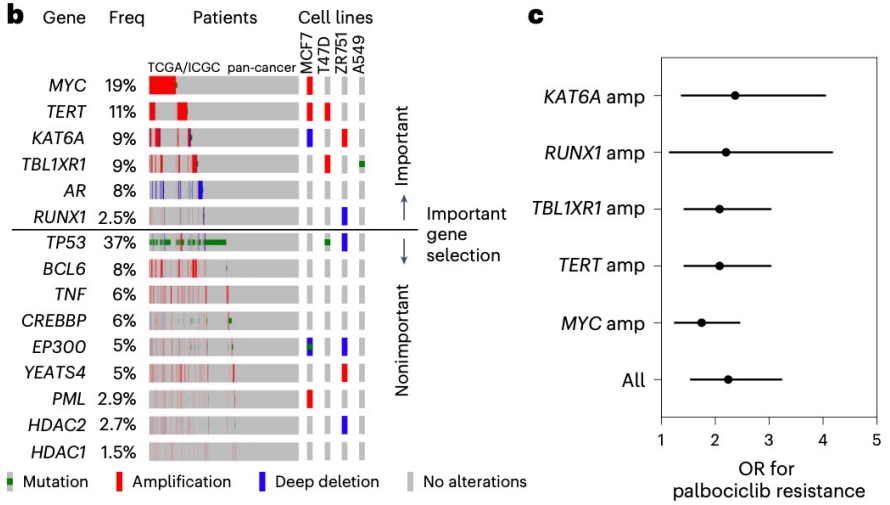


Histone acetylation results in loose packing of nucleosomes. Transcription factors can bind the DNA and genes are expressed.

Results

Exploration of gain-of-function alterations in a histone transcriptional assembly

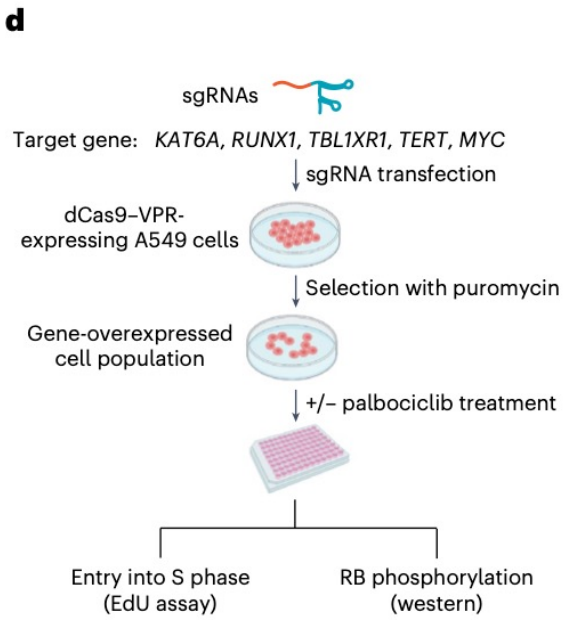
- **CNA-Driven Resistance**: Frequent gene amplifications in NeST:85 linked to resistance, notably in lung and gynecologic cancers.
- CNAs also accounted for the top five genetic alterations in this assembly that were most predictive of palbociclib resistance (OR of approximately 2.0)

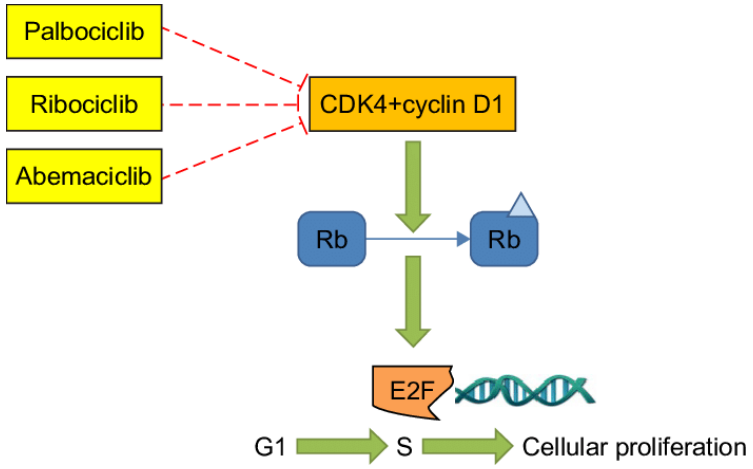


Results

Exploration of gain-of-function alterations in a histone transcriptional assembly

- **CRISPRa Technique**: Utilized to upregulate NeST:85 genes, revealing their effect on cell cycle progression (A549 lung carcinoma epithelial cells)

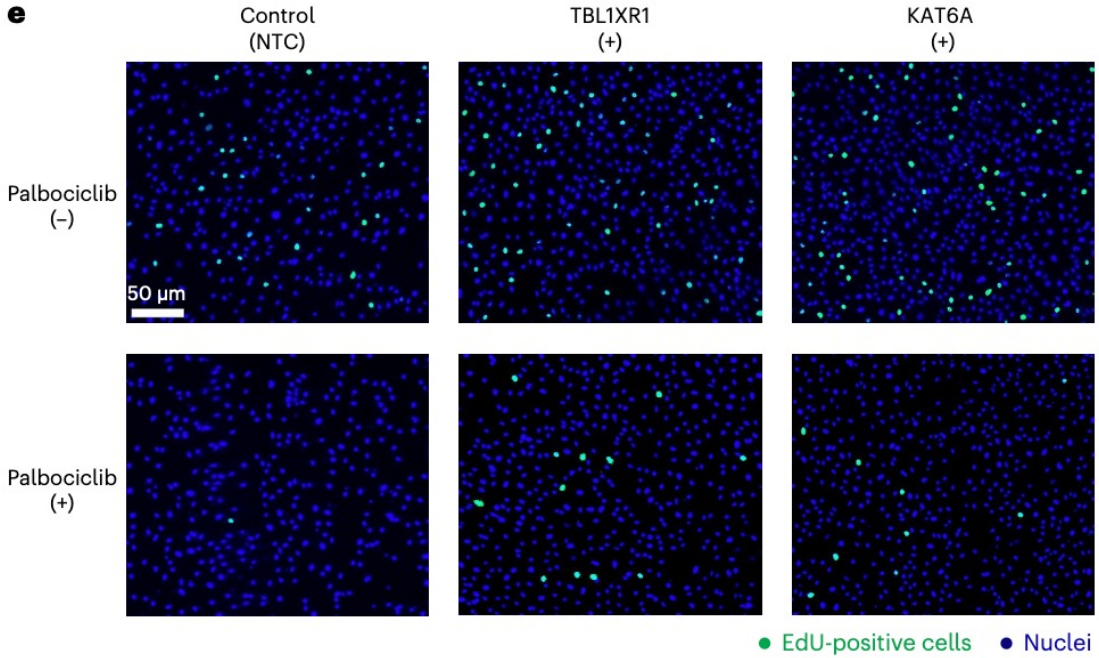


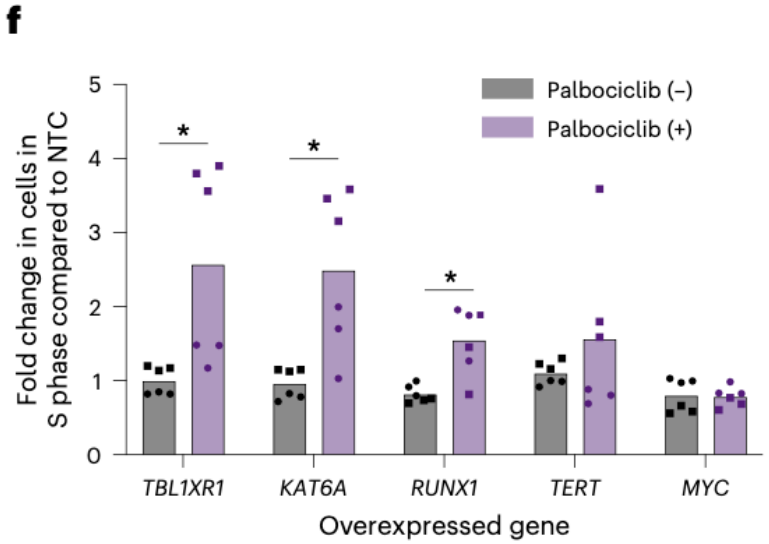


Results

Exploration of gain-of-function alterations in a histone transcriptional assembly

- **Increased S Phase Entry**: Overexpressing NeST:85 genes in A549 cells leads to more cells entering the S phase during palbociclib treatment.

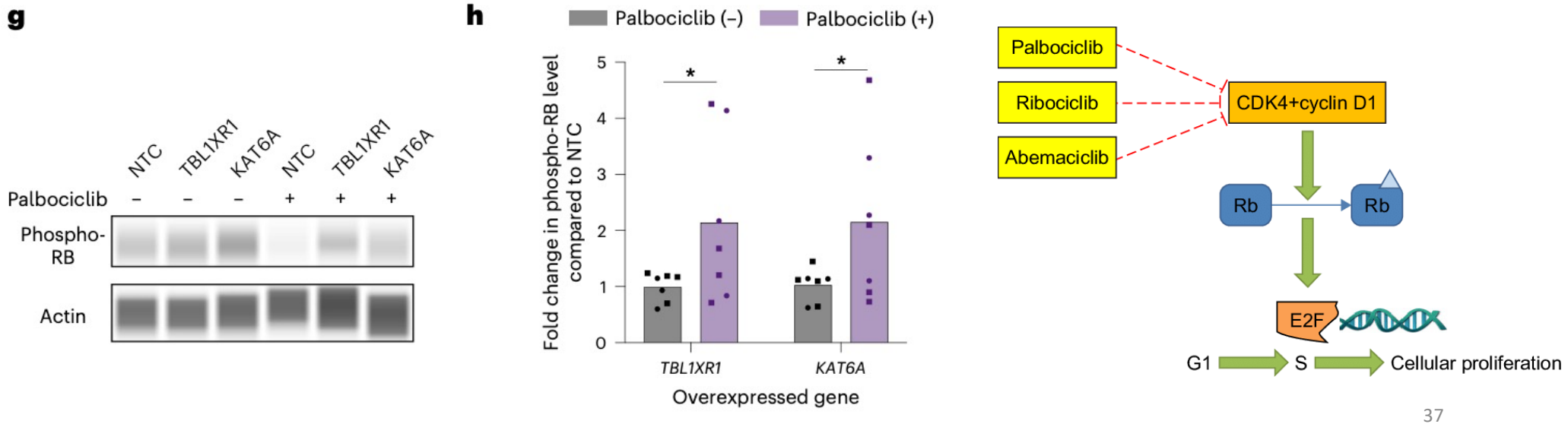




Results

Exploration of gain-of-function alterations in a histone transcriptional assembly

- **Phospho-RB Indicator**: Overexpression correlates with higher phospho-RB levels, **indicating potential CDK4/6 inhibitor resistance.**



Discussion

Why has the prediction of CDK4/6i responses remained challenging?

- Markers with promise in cell lines (for example, *CCND1* amplification) do not consistently translate to patient populations
- Individual genetic alterations that are clinically predictive may occur too rarely to have broad utility(Rb)
- A wider, more integrative analysis is needed to understand CDKi resistance fully

Discussion

Nest-VNN

- Synthesizes both rare and common genetic events across a repertoire of drug response pathways
- Quantitative, integrated assessment of drug response with a map of tumor cell components + DNN
- Key subcellular assemblies of models that accurately capture drug responses in vitro and that translate to in vivo (for example, PDX) and clinical settings
- Candidate biomarker selection + resistance score calculation

Discussion

Nest-VNN

- Previous approaches : use knowledge of cell structure or known pathway DB
- Based on NeST, a whole-cell map of cancer protein complexes derived from proteomics
  - can incorporate information from numerous rare mutations in predicting a drug response
  - almost certainly does not include all relevant protein assemblies (false negatives)
  - may be imperfect or irrelevant to a given tumor population (false positives)
- NeST knowledgebase is dynamic entity, which can be updated

Discussion

Summary

- Identified a set of eight core assemblies associated with anti-CDK4/6 response
- Seven of which were validated by one or more CRISPR screens
- In NeST-VNN, the EGF/FGF complex combines alterations, which have largely been reported separately, into a single integrated effect
- The model also highlights a notable role for NeST:85 (histone- mediated transcription regulation). underscores possible combination of HDAC inhibitor therapies with cell-cycle inhibitors

Discussion

Summary

- In such an integrated analysis, diverse effects converge on biological machinery at multiple levels to produce an overall treatment outcome.
- Solve difficulty in identifying individual genetic biomarkers + heterogeneity between patients

Questions

💡 Thank you for listening. Feel free to ask any questions now 🙂