Deciphering epigenetic influence of curcumin on histone signatures in breast cancer using 3D network

Tina Tang[2], Mikhail Berezin[3], Shamim Mollah[1,2*]

[1] Department of Genetics, Washington University School of Medicine in St. Louis, MO 63110, USA.
[2] Institute for Informatics and Data Science, Washington University School of Medicine in St. Louis, MO 63110, USA.
[3] Mallinckrodt Institute of Radiology, Washington University School of Medicine in St. Louis, MO 63110, USA.

[*]Corresponding author: samollah1@gmail.com

Abstract

Aberrant epigenetic alterations are known to lead to cancer arising from normal non-cancerous cells. Unlike genetic mutations, these dysregulations can be reversible and are potential targets for anticancer drugs. Curcumin, a natural plant-derived compound, has been shown to have an anticancer effect via its influence on epigenetic regulation. However, the exact identity of those epigenetic changes, as well as their mechanisms of action, remain largely unknown. The goal of this study is to use a 3D network model to identify what histone codes are being modified by curcumin and hypothesize the enzymatic pathways through which those changes occur in breast cancer. We analyzed multi-omic data from the NIH LINCS program and identified two novel histone marks associated with curcumin that are influenced by twenty-three phosphoproteins involved in cell signaling which are profiled in this paper. These histone and phosphoprotein signatures can be used as potential biomarkers for future chromatin-based drug therapies in breast cancer.

GitHub URL Source code is available at https://github.com/mollahlab/ECHC

Introduction

Breast cancer is currently the most frequently diagnosed cancer according to the World Health Organization (WHO) with 2.3 million new cases every year. The occurrence and progression of breast cancer is determined by a complex interplay of genetic, epigenetic, and environmental factors. While traditionally genetic mutations, such as BRCA1 and BRCA2, have been the focus for determining breast cancer etiology, there is now more and more evidence of the significant role that epigenetic changes play in the cellular transition from a normal state to a cancerous one (1). These epigenetic aberrations can alter gene expression in tumor-related genes and disrupt key pathways such as cell differentiation, survival, migration, and invasion (2). In addition, these epigenetic abnormalities, unlike their genetic counterparts, are reversible and present novel targets for the development of

cancer therapies. However, knowledge about the identity of these epigenetic alterations as well as the mechanisms behind the cause and consequence of them remain largely unknown (3).

The epigenetic markers that this study focuses on are post-translational modifications (PTMs) on histone proteins, one of the structural building blocks that make up chromatin. These PTMs include four different types of covalent modifications on the N-terminal tails of histones: acetylation, methylation, phosphorylation, and ubiquitination (4). Depending on the type, number, and location of these PTMs, which constitute a histone code, they can have different effects on chromatin structure and gene regulation. For example, H3K4me3 promotes the open chromatin state and thus gene transcription; however, H3K27me3 is a repressive mark and indicative of a closed chromatin state (5). These histone PTMs, which are modulated by specialized enzymes termed "readers", "writers", and "erasers", are often the endpoints of protein signaling events following environmental stimuli, such as drug exposure (6). The phosphoproteins involved in these signaling networks can also serve as targets for epigenetic drug therapies (7).

Among the substances studied for their epigenetic activities, curcumin is a natural bioactive compound with pleiotropic biological effects. Derived from the spice turmeric, it has been used in traditional Asian cuisine and medicine for hundreds of years and reported to have an array of pharmacological properties including anti-inflammatory, antioxidant, neuroprotective, and anticancer (8, 9, 10). Curcumin has been shown to suppress breast cancer through the regulation of a multitude of cellular signaling pathways such as cell proliferation, apoptosis, cell-cycle arrest, reactive oxidative stress (ROS), and microRNAs with molecular targets ranging from NF-κB, cyclinD1, P13K, Raf-1, NRF2 and p53 (11, 12, 13, 14, 15, 16, 17, 18, 19). In addition to these mechanisms, curcumin has also been found to influence histone modifications, modulating histone acetylation and methylation in cancer cell lines. For example, in one study on triple-negative breast cancer (TNBC), curcumin was shown to inhibit oncogenic EZH2 by reducing H3K27me3 at the promoter for tumor suppressive DLC1, inducing apoptosis in tumor cells (2). However, there is still a need for more studies determining curcumin's effect on specific histone targets as well as the signaling mechanisms behind its epigenetic regulation.

The goal of this paper is to expand our understanding of curcumin's influence on histone signatures by using machine-learning techniques to integrate and analyze multi-omics data collected from various cell lines. This data was compiled as part of the NIH Library of Integrated Network-Based Cellular Signatures (LINCS) program, which aimed to systematically catalog cellular responses to a range of chemical, genetic, and disease perturbagens, including curcumin (20). Utilizing the 3D data connectivity model developed by Mollah et al. called iPhDNet (21) and it's generalized version Gen3DNet (22), we aim to profile the histone signatures that encapsulate the curcumin drug response in breast cancer as well as the relevant phospho-signaling proteins that underlie its mechanisms of action. To do this, we will 1) identify the functional modules of histone codes across different cell lines, 2) deduce which phosphoproteins contribute to these histone functional modules in breast cancer, and 3) hypothesize the biological pathways in which these phosphoproteins may operate.

Using this data-driven approach, we hope to provide deeper insight into curcumin's epigenetic mechanisms as well as its potential efficacy in breast cancer drug therapies.

Materials and Methods

Data Acquisition

Experimental data was obtained from the LINCS Proteomic Characterization Center for Signaling and Epigenetics (PCCSE) Panorama Repository (https://panoramaweb.org/LINCS). The data was collected from five different cancer cell lines including MCF7 (breast), PC3 (prostate), A375 (skin), A549 (lung), and YAPC (pancreatic), as well as two neural cell lineages: neural progenitor cells (NPC) and astrocytes. Each cell line was perturbed by 31 neuroactive drugs, curcumin being one of them. Afterwards, the cellular state was captured through three different types of assays that measured RNA transcripts (L1000), phosphoproteins (P100), and global chromatin profiles (GCP).

Both the P100 and GCP are mass-spectrometry based, proteomic assays. The P100 is targeted against 96 phosphorylated peptides represented in cellular signaling pathways while the GCP is targeted against 60 combinatorial histone PTMs (23). These assays were each acquired at Level 3, which is preprocessed to be log 2 transformed and normalized. The P100 was collected at 3 hours after treatment while the GCP was collected at 24 hours after treatment. The L1000 is a targeted gene expression assay against 978 landmark genes that are representative of the entire transcriptome (24). These were obtained for the MCF7 cell line at Level 5, which is preprocessed to be a differential expression signature called a CD-coefficient using the characteristic direction method developed by Ma'ayan et al (25). The L1000 assays were downloaded at 6 hours and 24 hours after treatment with curcumin from the SigCom LINCS web server (https://maayanlab.cloud/sigcom-lincs). A complete list of all data files used is found in **Supplementary Table 1**.

Data Preprocessing

For each of the 31 drugs in the proteomic assays (P100 and GCP), there were three biological replicates available. These replicates were averaged and used to impute any missing values. The differential fold change was then calculated by taking the fold change of each perturbed phosphoprotein or histone with respect to the control (DMSO). The final preprocessing result is two data matrices: a histone matrix [31 neuroactive drugs x 52 histone marks] and a phosphoprotein matrix [31 neuroactive drugs x 72 phosphopeptides]. For the identification of histone modules across different cell lines, another histone matrix was created that consisted of histone marks perturbed by curcumin in each cell line [7 cell lines x 52 histone marks].

Gen3DNet Data Analysis Pipeline

Generic 3D Network (Gen3DNet), first derived from Integrated Phosphoprotein Histone Drug Network (iPhDNet), is the data analysis pipeline used in this study that was developed by Mollah et

al (21, 22). Gen3DNet is the generalized version of iPhDNet that consists of two parts: a non-negative matrix factorization (NMF) and a partial least square regression (PLSR) component. It takes two data inputs: a histone matrix and phosphoprotein matrix. **Figure 1** summarizes this workflow below. The Gen3DNet R package is available on Github (https://github.com/mollahlab/gen3DNet).

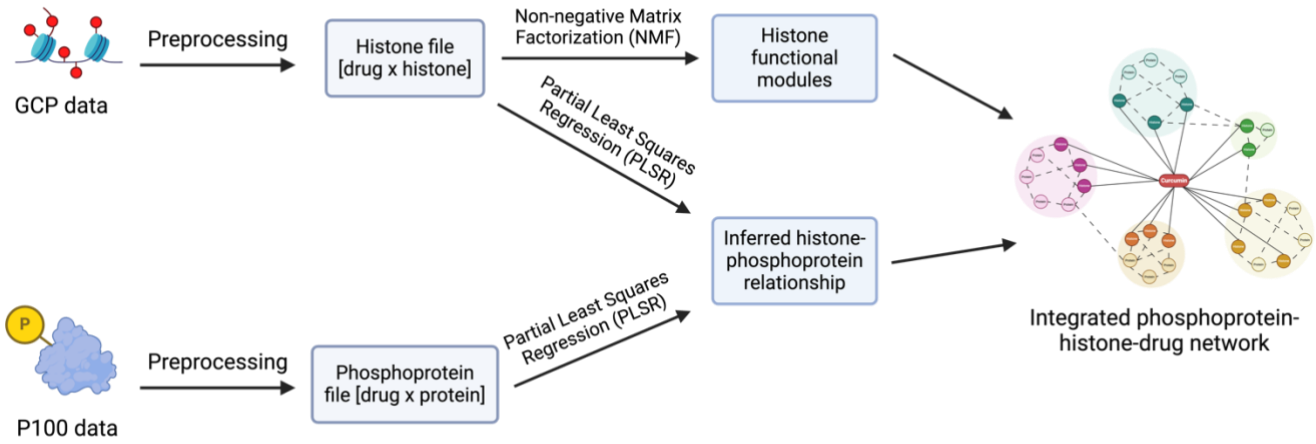


Figure 1: Gen3DNet Workflow. This workflow shows how an integrated three-dimensional network connecting drugs, phosphoproteins, and histone marks can be generated from two proteomic assays, GCP and P100, using Gen3DNet.

In the first component, NMF is used to identify the histone signatures. The goal of NMF is to extract a smaller number of latent factors (**k**) that can explain the data; these latent factors can be used to cluster the data into **k** modules. It does this by factoring the original matrix **A** (the histone matrix) into two matrices, **W** (basis matrix) and **H** (loading matrix), where the product of **W** and **H** approximate **A** (**A ≈ WH,** where **W, H ≥ 0**). The coefficients of the basis matrix signify the association of each histone code with a cluster, while the loading or mixture coefficients represent the connection of each drug (or cell line) to a cluster. This ultimately can be used to group a set of drugs with a collection of histone codes, which we refer to as a histone signature, to characterize the drug response.

In the second component, PLSR is used to create the histone-peptide interaction network. PLSR is a multivariate, linear regression model that finds the fundamental relations, known as the latent variables, between two matrices: **X** (predictor matrix) and **Y** (response matrix). In our application, the histone matrix is modeled as the response matrix (**Y**) and the phosphoproteins are the predictor variables (**X**). PLSR is useful in cases where there are many predictor variables with high collinearity, such as in this case where the phosphoproteins are highly correlated with each other. The resulting regression coefficients thus represent the contribution of each phosphoprotein to a specific histone code and the strength of their interaction. Finally, a t-test (**p < 0.01**) was used on these PLSR coefficients to determine the statistically significant phosphoproteins.

Integrated Histone-Phosphoprotein-Drug Network

To generate the three-dimensional network of interactions, the NMF histone-drug signatures for curcumin were integrated with the predicted PLSR histone-phosphoprotein relationships. This was visualized using Cytoscape where the histones, phosphoproteins, and drugs made up the individual nodes, and the NMF or PLSR coefficients were the edges connecting histones to curcumin or histones to phosphoproteins, respectively (26).

Results

Two functional histone modules identified across different cancer and neural cell lines

To better understand the context of curcumin's cellular response, we identified two functional modules in relation to the five cancer and two neural cell lines. Our results are displayed as heatmaps of basis and mixture components that correspond to each histone code and cell line's membership to a histone signature (**Figure 2A**). The basis components signify the strength of each histone code's membership to a functional module, while the mixture components reveal the contribution of each cell line towards a histone signature. These two histone modules characterize the cellular response state to curcumin. Thirty unique histone codes are associated with the first module, and nineteen are with the second module. The first module is made up of the cell lines MCF7 (breast cancer), PC3 (prostate cancer), NPC (neural progenitor cell), and astrocyte; while the second module consists of YAPC (pancreatic cancer), A375 (skin cancer), and A549 (lung cancer). Seeing how both MCF7 and PC3 are clustered together and are cancers affected by hormone dysregulation gave us good intuition that there were similarities in how curcumin affects the epigenetic landscape in both diseases. Likewise, since MCF7 was also grouped together with NPC and astrocyte, this indicates that curcumin may have related epigenetic mechanisms in breast cancer and neurological pathologies. This is further supported by the fact that MCF7 shares ectodermal origin with NPC and astrocyte cells (27, 28, 29). The detailed summary of histone modification and cell line clustering results can be found in **Supplementary Table 2**.

Phosphoprotein-histone associations mediated by curcumin in breast cancer

To further decipher curcumin's effect in breast cancer, we used network analysis to identify specific histone codes perturbed by curcumin and construct interactions between those histone codes and signaling phosphoproteins. First, NMF was used on the MCF7 GCP data to cluster the histone codes and neuroactive drugs into three distinct signatures, with curcumin being found in the second cluster along with two histone codes: H3K79me1 and H3K56me0 (**Figure 2B**). **Supplementary Table 3** recaps each histone signature with their respective drug groupings and associated drug pathways. Next, PLSR was utilized on the P100 phosphoproteins and GCP histones to predict phosphoprotein-histone relations that are then integrated with the histone signatures mentioned previously. After filtering for the statistically significant phosphoprotein interactions which resulted in 40 unique associations between 39 distinct phosphoproteins and the two histone codes, the outcomes are detailed in **Supplementary Table 4** and visualized in a three-dimensional network that connects

phosphoproteins to histones to curcumin (**Figure 2C**). The top enriched phosphoproteins in this network, which are circled on the Cytoscape-generated network, include: ZC3H14, NUP214, RBM17, TMSB4X, AP1GBP1, EIF4A3, PFKP, and TMPO.

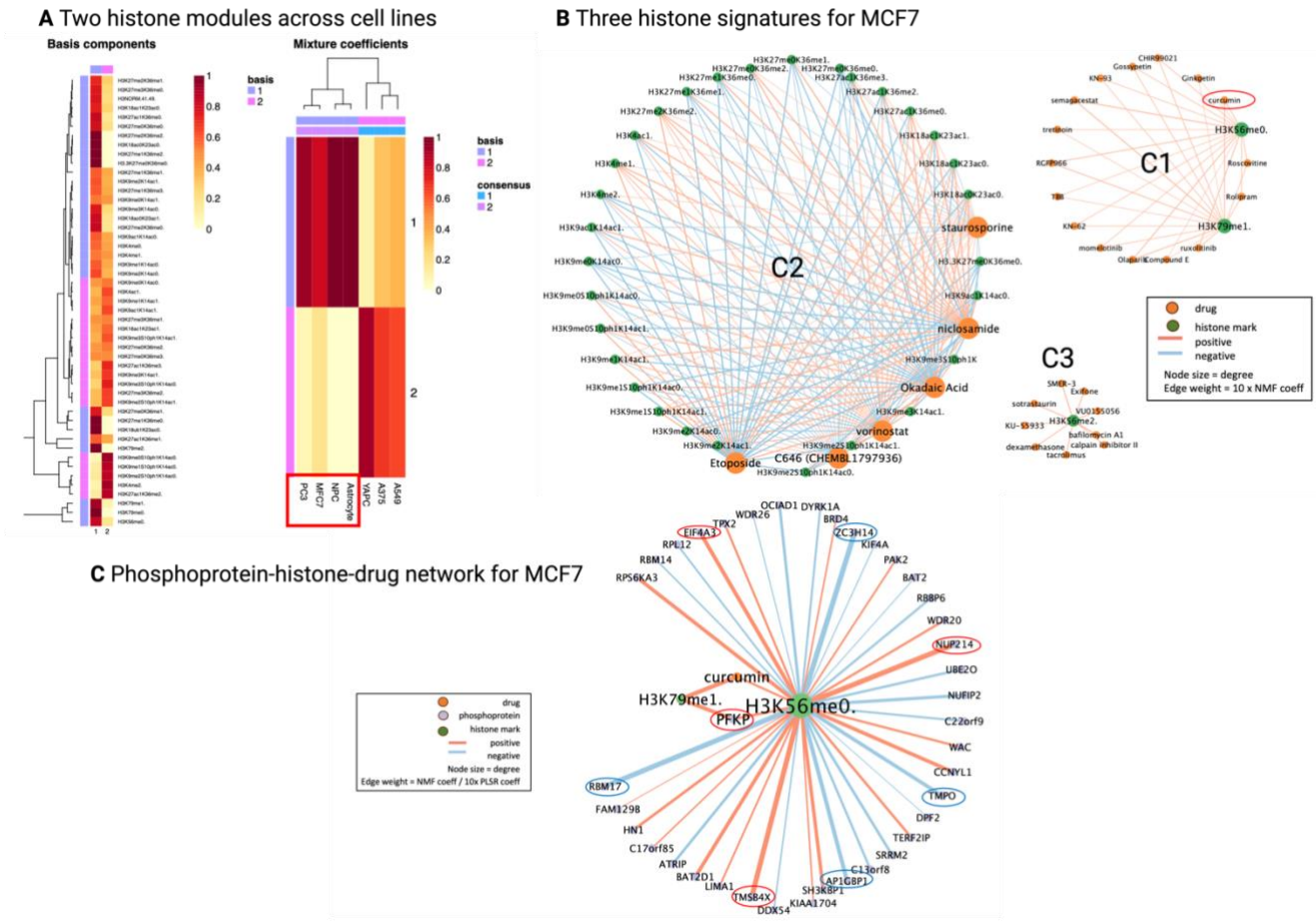


Figure 2: (A) Two histone modules characterized across seven cell lines. (B) Three histone signatures categorized for the MCF7 cell line. While the orange and green nodes represent neuroactive drugs or histone marks, respectively, the light red and blue edges represent elevated or reduced histone response levels. (C) Histone-phosphoprotein-drug network for the MCF7 cell line perturbed by curcumin. Curcumin is signified by the orange node, while phosphoproteins and histone marks are shown in purple and green, respectively. Elevated and reduced phosphoprotein/histone levels are represented as light red and blue edges. The top enriched phosphoproteins in this network are circled in red and blue standing for the top elevated and reduced peptides, respectively.

Two histone modifications connected to curcumin in breast cancer

A literature search on H3K56me0 and H3K79me1 reveals the histone PTMs found to be significantly connected to curcumin. The results are summarized in **Supplementary Table 5**. Both histone marks were found to be elevated in our study; to our best knowledge this is the first study of these histone codes in relation with curcumin. The first histone mark, H3K56me0, is a novel mark with limited data on its biological function. However, past studies have characterized H3K56me1 as an important

facilitator in DNA replication by serving as an important docking site for nuclear machinery (30), suggesting that the absence of methylation of K56 in H3K56me0 may lead to cell cycle disruption. This would help support the possible case of curcumin acting to induce cell cycle arrest in breast cancer cells to suppress tumor growth.

The second histone mark, H3K79me1, is associated with transcriptional activation along with its counterparts, H3K79me2/3 (31, 32). One study reports that the transition from mono-methylation to di- and tri-methylation at H3K79 is correlated with low- to high-level gene transcription activity (33). While there are no studies directly connecting H3K79me1 to breast cancer, there is evidence that H3K79me2 is positively correlated with breast cancer gene expression (34). Additionally, inhibition of H3K79 methylation by targeting DOT1L, the only known H3K79 methyltransferase, suppresses breast cancer proliferation (35, 36, 37). Though H3K79me1 is elevated in our study, it is possible that the role of mono-methylation in cellular regulation, and/or its location in the genome, is distinct from that of H3K79me2 and may still contribute to breast cancer repression.

In addition to their role in breast cancer, H3K79 methylation and DOT1L activity have also been found to be important for neurological function. For example, H3K79 methylation is necessary for neuronal development and regulation (38). DOT1L has also been discovered to help maintain proper cortical and cerebellar development (39). These findings suggest that H3K79 methylation may be a possible epigenetic target in the cellular mechanisms by which curcumin acts as a neuroactive drug.

Functional profiling of enriched phosphoproteins associated with histones PTMs

To augment our findings on the enriched P100 phosphoproteins obtained from our network analysis, we incorporated the L1000 gene expression data from the MCF7 cell line perturbed by curcumin. Of the 39 unique phosphoproteins identified as significant by PLSR, 23 of those matched to genes captured in the L1000 assay (**Figure 3A**). Given that the phosphoproteins were measured at 3 hours and RNA expression at 6 and 24 hours, we decided to compare the direction (positive or negative) of the L1000 CD-coefficient with the direction of the P100 fold change. For both CD-coefficient and fold change, a positive value signifies an increase in expression whereas a negative value indicates a decrease in expression. As expected, there were more L1000 genes sharing the same direction with the P100 phosphoproteins at 6 hours than at 24 hours, with 15 of the 23 genes matching at 6 hours but only 6 genes matching at 24 hours (four of these genes matched at every time point) (**Figure 3B**). Thereafter, we narrowed down our list of enriched phosphoproteins to these 23 that had corresponding gene expression values.

To determine which of the 23 phosphoproteins might be meaningful targets for drug interventions, we consulted several different resources and analysis tools. This included a literature search for previous studies connecting to breast cancer, functional enrichment analysis, gene annotation databases such as GeneCards, and additional neurological gene lists (40). The functional significance and results from the LINCS data of the top enriched phosphoproteins are profiled in **Supplementary Table 6**. Overall, the most enriched biological processes identified by Metascape for these 23 genes

included regulation of cytoskeleton organization, axon guidance, RNA splicing, and cell cycle (**Figure 4A**). Several of the top enriched phosphoproteins are associated with these biological functions, including EIF4A3 and RPS6KA3 in nervous system development, EIF4A3 and RBM17 in mRNA splicing, NUP214 and TMPO in cell cycle regulation (41). An extensive and detailed list of the top enriched processes and their correlated genes is provided in **Supplementary Table 7**.

Since curcumin is a known pleiotropic drug with neuroactive properties, we further investigated which phosphoproteins overlap with genes that have been implicated in neural cells and neuropathy (42, 43). To do this, we first observed to see if shared phosphoproteins existed between the top enriched proteins from MCF7 and those identified in NPC and astrocyte cell lines for curcumin (**Figure 3C**). There were five genes that coincided with all three cell lines: LIMA1, RPL12, UBE20, SH3KBP1, and TMPO. We also cross-referenced with a list of dorsal root ganglion (DRG) genes that had been implicated in chemotherapy-induced neuropathy by oxaliplatin, a common anticancer drug, from a study done by a collaborating lab (Berezin et al; unpublished). Of the 271 significant DRG genes (FDR<0.05, fold-change > 1.5), the TMPO phosphoprotein was found to be enriched (**Figure 3D**).

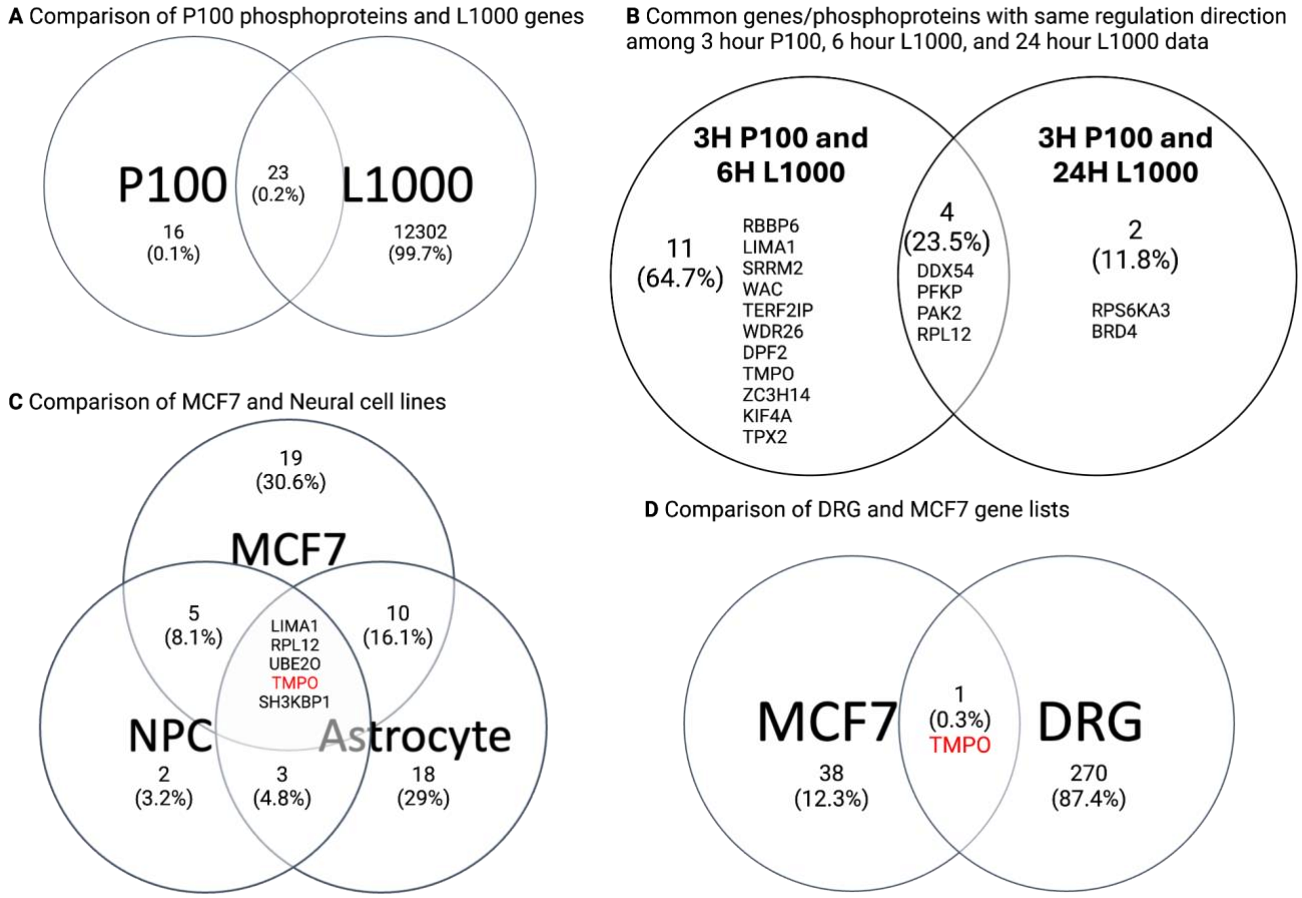


Figure 3: P100 (phosphoprotein), L1000 (transcriptomic), and neural cell gene list (NPC, Astrocyte, DRG) comparisons. Enriched P100 phosphoproteins of interest were compared (A) to the L1000 genes list, (B) across different time points by regulation direction (with 11 genes matching directions at 3 and 6 hours, 2 genes matching at 3 and 24 hours, and 4 genes matching across all time points), (C) to the phosphoproteins

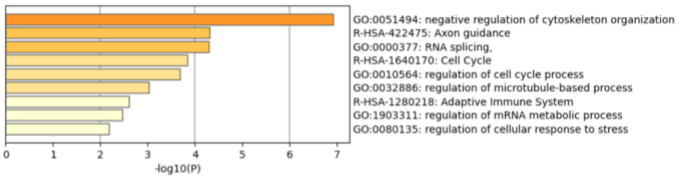
associated with curcumin in neural cell lines (NPC and astrocyte), and (D) to the list of significant DRG neuropathic genes.

Protein-protein interaction and transcription factor analysis exhibited potential signaling pathways

To identify possible gene regulatory networks, we inferred signaling pathways for curcumin by exploring connections between the phosphoproteins-of-interest and transcription factors (TFs). We used Enrichr, a transcription factor enrichment tool, to identify associated transcription factors from the gene list of 23 enriched genes (44). Based on probability scores and biological significance in breast cancer, we selected five transcription factors: YY1, ZBTB7A, EGR1, KAT2A, and STAT3 (**Figure 4B**). We then used the STRING database, a protein-protein interaction (PPI) tool, to predict interactions among those transcription factors and the 23 enriched phosphoproteins which function as possible TF targets (45). The resulting protein-protein interaction network is displayed in **Figure 4C**.

When proposing potential signaling pathways, we focused on relationships linking TFs to the top enriched phosphoproteins from the Gen3DNet network analysis. One example of a possible signaling pathway, involving the transcription factor ZBTB7A influencing the phosphoproteins PFKP and TMPO downstream, is outlined in **Figure 4C** and depicted in **Figure 4Di**. We observed each gene in this pathway across three time points to determine if the expression data, either increased ("up") or decreased ("down"), can match the role that each gene plays in the pathway and in breast cancer in order to determine the pathway's feasibility in breast cancer suppression. The expression of PFKP is increased at all available time points after curcumin exposure, while the expression of TMPO and ZBTB7A are initially decreased (note there is no P100 data available for ZBTB7A since it is not a phosphoprotein) (**Figure 4Dii**). However, both ZBTB7A and TMPO switch from being down-regulated to being up-regulated at 24 hours, implying that perhaps the influence of curcumin declines before 24 hours after first exposure. Additional potential pathways are characterized in **Supplementary Table 8**.

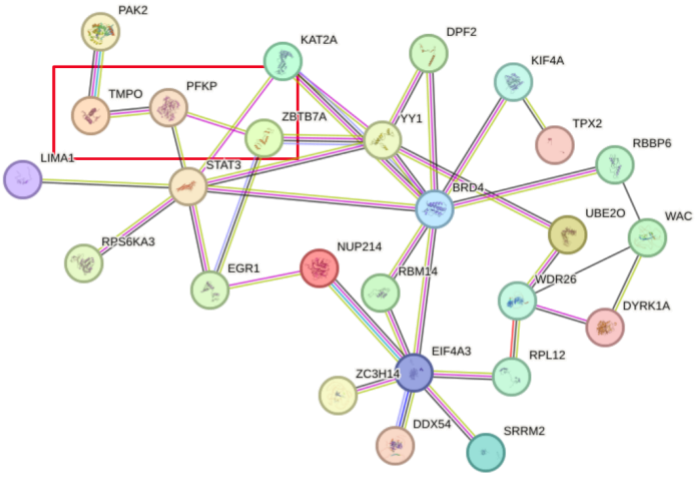
A Enriched biological processes from Metascape



B Transcription factor enrichment from Enrichr

Transcription Factor	Source	p-value
YY1	ENCODE and ChEA Consensus TFs from ChIP-X	2.04E-0.7
ZBTB7A	TRRUST Transcription Factors 2019	0.006881
EGR1	TRANSFAC and JASPAR PWMs	0.001719
KAT2A	TF ChIP-seq 2015	0.00001649
STAT3	TF-LOF Expression from GEO	0.0004224

C Protein-protein-interaction analysis from STRINGdb



D One possible signaling pathway

i.

ii.

Gene	3H P100	6H L1000	24H L1000
ZBTB7A	NA	Down	Up
PFKP	Up	Up	Up
TMPO	Down	Down	Up

Figure 4: Gene list enrichment analysis results. (A) Metascape pathway enrichment analysis. (B) Enrichr transcription factor enrichment analysis. (C) STRINGdb protein-protein-interaction analysis. (D) Example of potential drug signaling pathway analyzed across time points.

Discussion

This study serves as an example of how the application of bioinformatic tools on complex, multi-omic data allows for the extraction of a smaller number of meaningful features that can be further examined and constructed into data-driven hypotheses. This can allow us to begin to theorize potential epigenetic signaling pathways for curcumin in breast cancer. This study also brought to light several limitations that arise when answering these types of biological questions. The biggest limitation encountered during the analysis process was the lack of consistent time points across assays. The difference in collection time between the P100 and GCP assays left some ambiguity in the predicted phosphoprotein-histone relationships, as histone PTMs are very transient and can change quickly between 3 and 24 hours. While an attempt to supplement this with the L1000 assays at 6 and 24 hours was made, the reliability of using RNA expression to capture phosphoprotein levels is also inconsistent. Ideally, each type of assay would be collected at the same time point, across multiple time points. Without this data consistency, identifying a mechanism of action across multi-omic components and providing evidence for causation is challenging

One phosphoprotein of interest did emerge from our profiling of curcumin's epigenetic influence: TMPO. As one of the top enriched phosphoproteins in our histone-phosphoprotein network analysis, TMPO is a structural protein found in the nucleus and involved in several cellular processes including gene transcription, DNA replication, and cell cycle control (46). It has been implicated in cancer progression and found to be upregulated in several cancers, including breast cancer; its depletion is associated with cell cycle arrest and apoptosis in cancer cell lines (47). In our study, TMPO is downregulated, potentially corresponding to curcumin's anti-cancer activity. Furthermore, as neuropathy is a common side-effect in chemotherapy treatments for breast cancer, it is important to consider curcumin's dual role in the mediation of neuropathy in addition to breast cancer suppression (48, 49, 50). Based on the results of this study, TMPO may play a key role in both. When drawing comparisons with the neural cell lines, TMPO was among the shared enriched phosphoproteins between the breast cancer (MCF7), neural progenitor cell (NPC), and astrocyte cell lines. Moreover, TMPO was the sole gene that matched with the DRG genes linked to neuropathy study (Berezin et al; unpublished). Given its neurological importance, as well as its involvement in the cell cycle and cancer, we present TMPO as a possible key player in curcumin's epigenetic mechanism of action.

Among the possible pathways that involve TMPO, the most viable pathway involves the transcription factor ZBTB7A and phosphoprotein PFKP. ZBTB7A is a zinc finger protein that functions as a transcriptional repressor for a wide range of genes involved in cell proliferation and differentiation, including PFKP; it has been reported to accelerate breast cancer progression (51, 52). In our study, ZBTB7A is downregulated in the L1000 expression data captured at 6 hours after curcumin exposure which correlates with cancer suppression. PFKP, a key enzyme in glycolysis regulation, can play a role in metabolic reprogramming in some cancers, including bladder, breast, and lung; high levels of PFKP are associated with poor prognosis in cancer patients (53, 54). PFKP is upregulated in our study which correlates with ZBTB7A inhibition but does not, however, correspond with breast cancer repression according to the current literature. There is also no well-documented relationship between PFKP and TMPO at the present; though both proteins can be localized to the nucleus, further studies would be needed to validate this connection (55). In addition, to this date there are no studies on the relationships between PFKP and TMPO with the histone mark H3K56me0 (and similarly between PFKP and H3K79me1).

As mentioned earlier, much of the histone signatures profiled from this study are novel and have not been studied previously, especially in connection to curcumin. These signatures include the function and regulation of the histone PTMs identified as part of curcumin's histone-drug signature, the predicted phosphoprotein-histone relationships, as well as the signaling pathways for gene regulation. To be able to effectively evaluate and validate all these new inferred relationships, more sequence-driven assays, such as ATAC-seq, ChIP-seq, and CUT&RUN, would be required. This would allow us to identify the location of histone peaks for the histone PTMs of interest as well as perform sequence-driven TF-motif analysis. The next step in determining causality would be experimental validation in a biological model, such as performing a knockout of TMPO and measuring breast cancer progression.

In conclusion, in this study we provided a systematic approach to profile histone signatures and generate data-driven hypotheses on epigenetic signaling mechanisms induced by curcumin. We achieved this by using the network analysis pipeline of Gen3DNet to identify drug-histone signatures and predict histone-phosphoprotein interactions, with the final outcome being a connectivity map among curcumin, phosphoproteins, and histone codes. To our knowledge this is the first study to analyze and profile upwards of 50 histone codes in relation to curcumin in breast cancer and link novel histone PTMs with curcumin treatment. In this study we also found possible connections between curcumin's epigenetic targets in breast cancer and neurological disorders, suggesting potential common epigenetic pathways utilized by these diseases. Overall, our study elucidated possible new therapeutic targets for curcumin as an epigenetic modulator for breast cancer and other diseases.

Acknowledgements

This work was supported by Washington University in St. Louis School of Medicine Genetics departmental fund. We want to thank the NIH LINCS consortium for making these data available.

Author Contributions

Original concept: S.M.; experimental design: S.M.; implementation: S.M., T.T; data interpretation: S.M., T.T.; DRG genes list generation: M.B.; manuscript writing: T.T., S.M.; research supervision: S.M.

References

1. Zhuang, J., Huo, Q., Yang, F., & Xie, N. (2020). Perspectives on the Role of Histone Modification in Breast Cancer Progression and the Advanced Technological Tools to Study Epigenetic Determinants of Metastasis. Frontiers in genetics, 11, 603552. https://doi.org/10.3389/fgene.2020.603552
2. Trager, M. H., Sah, B., Chen, Z., & Liu, L. (2021). Control of Breast Cancer Pathogenesis by Histone Methylation and the Hairless Histone Demethylase. Endocrinology, 162(8), bqab088. https://doi.org/10.1210/endocr/bqab088
3. Li, W., Wu, H., Sui, S., Wang, Q., Xu, S., & Pang, D. (2021). Targeting Histone Modifications in Breast Cancer: A Precise Weapon on the Way. Frontiers in cell and developmental biology, 9, 736935. https://doi.org/10.3389/fcell.2021.736935
4. Li, W., Wu, H., Sui, S., Wang, Q., Xu, S., & Pang, D. (2021). Targeting Histone Modifications in Breast Cancer: A Precise Weapon on the Way. Frontiers in cell and developmental biology, 9, 736935. https://doi.org/10.3389/fcell.2021.736935
5. Shanmugam, M. K., Arfuso, F., Sng, J. C., Bishayee, A., Kumar, A. P., & Sethi, G. (2019). Chapter 5—Epigenetic Effects of Curcumin in Cancer Prevention. In A. Bishayee & D. Bhatia (Eds.), *Epigenetics of Cancer Prevention* (Vol. 8, pp. 107–128). Academic Press. https://doi.org/10.1016/B978-0-12-812494-9.00005-6

6. Ardito, F., Giuliani, M., Perrone, D., Troiano, G., & Lo Muzio, L. (2017). The crucial role of protein phosphorylation in cell signaling and its use as targeted therapy (Review). International journal of molecular medicine, 40(2), 271–280. https://doi.org/10.3892/ijmm.2017.3036

7. Wang R, Yu H, Chen P, Yuan T, Zhang J. Integrated Transcriptome and Molecular Docking to Identify the Hub Superimposed Attenuation Targets of Curcumin in Breast Cancer Cells. International Journal of Molecular Sciences. 2023; 24(15):12479. https://doi.org/10.3390/ijms241512479

8. Wang, Y., Yu, J., Cui, R., Lin, J., & Ding, X. (2016). Curcumin in Treating Breast Cancer: A Review. *SLAS Technology*, *21*(6), 723–731. https://doi.org/10.1177/2211068216655524

9. Lee Y-M, Kim Y. Is Curcumin Intake Really Effective for Chronic Inflammatory Metabolic Disease? A Review of Meta-Analyses of Randomized Controlled Trials. Nutrients. 2024; 16(11):1728. https://doi.org/10.3390/nu16111728

10. Benameur T, Giacomucci G, Panaro MA, Ruggiero M, Trotta T, Monda V, Pizzolorusso I, Lofrumento DD, Porro C, Messina G. New Promising Therapeutic Avenues of Curcumin in Brain Diseases. Molecules. 2022; 27(1):236. https://doi.org/10.3390/molecules27010236

11. Jia, T., Zhang, L., Duan, Y. et al. The differential susceptibilities of MCF-7 and MDA-MB-231 cells to the cytotoxic effects of curcumin are associated with the PI3K/Akt-SKP2-Cip/Kips pathway. Cancer Cell Int 14, 126 (2014). https://doi.org/10.1186/s12935-014-0126-4

12. Fawzy, R. M., Abdel-Aziz, A. A., Bassiouny, K., & Fayed, A. M. (2024). Phytocompounds-based therapeutic approach: Investigating curcumin and green tea extracts on MCF-7 breast cancer cell line. *Journal of Genetic Engineering and Biotechnology*, *22*(1), 100339. https://doi.org/10.1016/j.jgeb.2023.100339

13. Li, R., Zhang, J., Zhou, Y., Gao, Q., Wang, R., Fu, Y., Zheng, L., & Yu, H. (2020). Transcriptome Investigation and In Vitro Verification of Curcumin-Induced HO-1 as a Feature of Ferroptosis in Breast Cancer Cells. *Oxidative Medicine and Cellular Longevity*, *2020*, 3469840. https://doi.org/10.1155/2020/3469840

14. Cao X, Li Y, Wang Y, Yu T, Zhu C, et al. (2022) Curcumin suppresses tumorigenesis by ferroptosis in breast cancer. PLOS ONE 17(1): e0261370. https://doi.org/10.1371/journal.pone.0261370

15. Quispe-Soto, E.T., & Quispe-Soto, E.T. (2016). Effect of curcumin and paclitaxel on breast carcinogenesis. International Journal of Oncology, 49, 2569-2577. https://doi.org/10.3892/ijo.2016.3741

16. Nirgude, S., Desai, S., & Choudhary, B. (2022). Curcumin alters distinct molecular pathways in breast cancer subtypes revealed by integrated miRNA/mRNA expression analysis. Cancer reports (Hoboken, N.J.), 5(10), e1596. https://doi.org/10.1002/cnr2.1596

17. Liu, J. L., Pan, Y. Y., Chen, O., Luan, Y., Xue, X., Zhao, J. J., Liu, L., & Jia, H. Y. (2017). Curcumin inhibits MCF-7 cells by modulating the NF-κB signaling pathway. Oncology letters, 14(5), 5581–5584. https://doi.org/10.3892/ol.2017.6860

18. Hosseini, S., Chamani, J., Hadipanah, M. R., Ebadpour, N., Hojjati, A. S., Mohammadzadeh, M. H., & Rahimi, H. R. (2019). Nano-curcumin's suppression of breast cancer cells (MCF7) through the inhibition of cyclinD1 expression. Breast cancer (Dove Medical Press), 11, 137–142. https://doi.org/10.2147/BCTT.S195800

19. Garufi, A., Baldari, S., Pettinari, R. et al. A ruthenium(II)-curcumin compound modulates NRF2 expression balancing the cancer cell death/survival outcome according to p53 status. J Exp Clin Cancer Res 39, 122 (2020). https://doi.org/10.1186/s13046-020-01628-5

20. Keenan, A. B., Jenkins, S. L., Jagodnik, K. M., Koplev, S., He, E., Torre, D., Wang, Z., Dohlman, A. B., Silverstein, M. C., Lachmann, A., Kuleshov, M. V., Ma'ayan, A., Stathias, V., Terryn, R., Cooper, D., Forlin, M., Koleti, A., Vidovic, D., Chung, C., … Pillai, A. (2018). The Library of Integrated Network-Based Cellular Signatures NIH Program: System-Level Cataloging of Human Cells Response to Perturbations. *Cell Systems*, *6*(1), 13–24. https://doi.org/10.1016/j.cels.2017.11.001

21. Mollah, S. A., & Subramaniam, S. (2020). Histone Signatures Predict Therapeutic Efficacy in Breast Cancer. IEEE open journal of engineering in medicine and biology, 1, 74–82. https://doi.org/10.1109/OJEMB.2020.2967105

22. Morrison, P., Tang, T., Lu, C., & Mollah, S. (2024). gen3DNet: An R Package for Generating 3D Network Models. bioRxiv, 2024.11.11.623060. https://doi.org/10.1101/2024.11.11.623060

23. Dele-Oni, D. O., Christianson, K. E., Egri, S. B., Vaca Jacome, A. S., DeRuff, K. C., Mullahoo, J., Sharma, V., Davison, D., Ko, T., Bula, M., Blanchard, J., Young, J. Z., Litichevskiy, L., Lu, X., Lam, D., Asiedu, J. K., Toder, C., Officer, A., Peckner, R., MacCoss, M. J., … Jaffe, J. D. (2021). Proteomic profiling dataset of chemical perturbations in multiple biological backgrounds. Scientific data, 8(1), 226. https://doi.org/10.1038/s41597-021-01008-4

24. Subramanian, A., Narayan, R., Corsello, S. M., Peck, D. D., Natoli, T. E., Lu, X., Gould, J., Davis, J. F., Tubelli, A. A., Asiedu, J. K., Lahr, D. L., Hirschman, J. E., Liu, Z., Donahue, M., Julian, B., Khan, M., Wadden, D., Smith, I. C., Lam, D., Liberzon, A., … Golub, T. R. (2017). A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. Cell, 171(6), 1437–1452.e17. https://doi.org/10.1016/j.cell.2017.10.049

25. Clark, N.R., Hu, K.S., Feldmann, A.S. et al. The characteristic direction: a geometrical approach to identify differentially expressed genes. BMC Bioinformatics 15, 79 (2014). https://doi.org/10.1186/1471-2105-15-79

26. Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., & Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome research, 13(11), 2498–2504. https://doi.org/10.1101/gr.1239303

27. Zhang, Q., Fan, H., Shen, J., Hoffman, R. M., & Xing, H. R. (2010). Human breast cancer cell lines co-express neuronal, epithelial, and melanocytic differentiation markers in vitro and in vivo. PloS one, 5(3), e9712. https://doi.org/10.1371/journal.pone.0009712

28. Martínez-Cerdeño, V., & Noctor, S. C. (2018). Neural Progenitor Cell Terminology. Frontiers in neuroanatomy, 12, 104. https://doi.org/10.3389/fnana.2018.00104

29. Marín-Teva, J. L., Cuadros, M. A., Martín-Oliva, D., & Navascués, J. (2011). Microglia and neuronal cell death. Neuron Glia Biology, 7(1), 25–40. doi:10.1017/S1740925X12000014

30. Yu, Y., Song, C., Zhang, Q., DiMaggio, P. A., Garcia, B. A., York, A., Carey, M. F., & Grunstein, M. (2012). Histone H3 lysine 56 methylation regulates DNA replication through its interaction with PCNA. *Molecular cell*, *46*(1), 7–17. https://doi.org/10.1016/j.molcel.2012.01.019

31. Farooq, Z., Banday, S., Pandita, T. K., & Altaf, M. (2016). The many faces of histone H3K79 methylation. *Mutation research. Reviews in mutation research*, *768*, 46–52. https://doi.org/10.1016/j.mrrev.2016.03.005

32. Wood, K., Tellier, M., & Murphy, S. (2018). DOT1L and H3K79 Methylation in Transcription and Genomic Stability. Biomolecules, 8(1), 11. https://doi.org/10.3390/biom8010011

33. Steger, D. J., Lefterova, M. I., Ying, L., Stonestrom, A. J., Schupp, M., Zhuo, D., Vakoc, A. L., Kim, J. E., Chen, J., Lazar, M. A., Blobel, G. A., & Vakoc, C. R. (2008). DOT1L/KMT4 recruitment and H3K79 methylation are ubiquitously coupled with gene transcription in mammalian cells. Molecular and cellular biology, 28(8), 2825–2839. https://doi.org/10.1128/MCB.02076-07

34. Jin, W., Li, Q.-Z., Liu, Y., & Zuo, Y.-C. (2020). Effect of the key histone modifications on the expression of genes related to breast cancer. *Genomics*, *112*(1), 853–858. https://doi.org/10.1016/j.ygeno.2019.05.026

35. Nassa, G., Salvati, A., Tarallo, R., Gigantino, V., Alexandrova, E., Memoli, D., Sellitto, A., Rizzo, F., Malanga, D., Mirante, T., Morelli, E., Nees, M., Åkerfelt, M., Kangaspeska, S., Nyman, T. A., Milanesi, L., Giurato, G., & Weisz, A. (2019). Inhibition of histone methyltransferase DOT1L silences ERα gene and blocks proliferation of antiestrogen-resistant breast cancer cells. Science advances, 5(2), eaav5590. https://doi.org/10.1126/sciadv.aav5590

36. Byun, W. S., Kim, W. K., Han, H. J., Chung, H. J., Jang, K., Kim, H. S., Kim, S., Kim, D., Bae, E. S., Park, S., Lee, J., Park, H. G., & Lee, S. K. (2019). Targeting Histone Methyltransferase DOT1L by a Novel Psammaplin A Analog Inhibits Growth and Metastasis of Triple-Negative Breast Cancer. Molecular therapy oncolytics, 15, 140–152. https://doi.org/10.1016/j.omto.2019.09.005

37. Zhang, L., Deng, L., Chen, F., Yao, Y., Wu, B., Wei, L., Mo, Q., & Song, Y. (2014). Inhibition of histone H3K79 methylation selectively inhibits proliferation, self-renewal and metastatic potential of breast cancer. Oncotarget, 5(21), 10665–10677. https://doi.org/10.18632/oncotarget.2496

38. Bovio, P. P., Franz, H., Heidrich, S., Rauleac, T., Kilpert, F., Manke, T., & Vogel, T. (2019). Differential Methylation of H3K79 Reveals DOT1L Target Genes and Function in the Cerebellum In Vivo. Molecular neurobiology, 56(6), 4273–4287. https://doi.org/10.1007/s12035-018-1377-1

39. Park, J., Lee, K., Kim, K. et al. The role of histone modifications: from neurodevelopment to neurodiseases. Sig Transduct Target Ther 7, 217 (2022). https://doi.org/10.1038/s41392-022-01078-9

40. Stelzer, G., Rosen, N., Plaschkes, I., Zimmerman, S., Twik, M., Fishilevich, S., Stein, T. I., Nudel, R., Lieder, I., Mazor, Y., Kaplan, S., Dahary, D., Warshawsky, D., Guan-Golan, Y., Kohn, A., Rappaport, N., Safran, M., & Lancet, D. (2016). The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. Current protocols in bioinformatics, 54, 1.30.1–1.30.33. https://doi.org/10.1002/cpbi.5

41. Zhou, Y., Zhou, B., Pache, L. et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat Commun 10, 1523 (2019). https://doi.org/10.1038/s41467-019-09234-6

42. Zhu, X., Li, Q., Chang, R., Yang, D., Song, Z., Guo, Q., & Huang, C. (2014). Curcumin alleviates neuropathic pain by inhibiting p300/CBP histone acetyltransferase activity-regulated expression

of BDNF and cox-2 in a rat model. PloS one, 9(3), e91303. https://doi.org/10.1371/journal.pone.0091303

43. Liu, S., Li, Q., Zhang, MT. et al. Curcumin ameliorates neuropathic pain by down-regulating spinal IL-1β via suppressing astroglial NALP1 inflammasome and JAK2-STAT3 signalling. Sci Rep 6, 28956 (2016). https://doi.org/10.1038/srep28956
44. Chen, E. Y., Tan, C. M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G. V., Clark, N. R., & Ma'ayan, A. (2013). Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC bioinformatics, 14, 128. https://doi.org/10.1186/1471-2105-14-128
45. Szklarczyk, D., Kirsch, R., Koutrouli, M., Nastou, K., Mehryary, F., Hachilif, R., Gable, A. L., Fang, T., Doncheva, N. T., Pyysalo, S., Bork, P., Jensen, L. J., & von Mering, C. (2023). The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic acids research, 51(D1), D638–D646. https://doi.org/10.1093/nar/gkac1000
46. Sun, D. P., Liew, P. L., Lin, C. C., Hung, S. T., Chen, T. C., Fang, C. L., & Lin, K. Y. (2019). Clinicopathologic and Prognostic Significance of Thymopoietin-α Overexpression in Gastric Cancer. Journal of Cancer, 10(21), 5099–5107. https://doi.org/10.7150/jca.30738
47. Zhang, L., Wang, G., Chen, S. et al. Depletion of thymopoietin inhibits proliferation and induces cell cycle arrest/apoptosis in glioblastoma cells. World J Surg Onc 14, 267 (2016). https://doi.org/10.1186/s12957-016-1018-y
48. Zajączkowska, R., Kocot-Kępska, M., Leppert, W., Wrzosek, A., Mika, J., & Wordliczek, J. (2019). Mechanisms of Chemotherapy-Induced Peripheral Neuropathy. International journal of molecular sciences, 20(6), 1451. https://doi.org/10.3390/ijms20061451
49. Liang, L., Lutz, B. M., Bekker, A., & Tao, Y. X. (2015). Epigenetic regulation of chronic pain. Epigenomics, 7(2), 235–245. https://doi.org/10.2217/epi.14.75
50. Liu, S., Li, Q., Zhang, MT. et al. Curcumin ameliorates neuropathic pain by down-regulating spinal IL-1β via suppressing astroglial NALP1 inflammasome and JAK2-STAT3 signalling. Sci Rep 6, 28956 (2016). https://doi.org/10.1038/srep28956
51. Liu, X. S., Haines, J. E., Mehanna, E. K., Genet, M. D., Ben-Sahra, I., Asara, J. M., Manning, B. D., & Yuan, Z. M. (2014). ZBTB7A acts as a tumor suppressor through the transcriptional repression of glycolysis. Genes & development, 28(17), 1917–1928. https://doi.org/10.1101/gad.245910.114
52. Mao, A., Chen, M., Qin, Q., Liang, Z., Jiang, W., Yang, W., & Wei, C. (2019). ZBTB7A promotes migration, invasion and metastasis of human breast cancer cells through NF-κB-induced epithelial-mesenchymal transition in vitro and in vivo. Journal of biochemistry, 166(6), 485–493. https://doi.org/10.1093/jb/mvz062
53. Peng, J., Li, P., Li, Y. et al. PFKP is a prospective prognostic, diagnostic, immunological and drug sensitivity predictor across pan-cancer. Sci Rep 13, 17399 (2023). https://doi.org/10.1038/s41598-023-43982-2
54. Inaishi, T., Shibata, M., Ichikawa, T., Kanda, M., Hayashi, M., Soeda, I., Takeuchi, D., Takano, Y., Tsunoda, N., Kodera, Y., & Kikumori, T. (2022). Platelet isoform of phosphofructokinase accelerates malignant features in breast cancer. Oncology reports, 47(1), 9. https://doi.org/10.3892/or.2021.8220

55. Gao, X., Qin, S., Wu, Y., Chu, C., Jiang, B., Johnson, R. H., Kuang, D., Zhang, J., Wang, X., Mehta, A., Tew, K. D., Leone, G. W., Yu, X. Z., & Wang, H. (2021). Nuclear PFKP promotes CXCR4-dependent infiltration by T cell acute lymphoblastic leukemia. The Journal of clinical investigation, 131(16), e143119. https://doi.org/10.1172/JCI143119

56. Wigington, C. P., Morris, K. J., Newman, L. E., & Corbett, A. H. (2016). The Polyadenosine RNA-binding Protein, Zinc Finger Cys3His Protein 14 (ZC3H14), Regulates the Pre-mRNA Processing of a Key ATP Synthase Subunit mRNA. The Journal of biological chemistry, 291(43), 22442–22459. https://doi.org/10.1074/jbc.M116.754069

57. Liu, L., Vujovic, A., Deshpande, N. P., Sathe, S., Anande, G., Chen, H. T. T., Xu, J., Minden, M. D., Yeo, G. W., Unnikrishnan, A., Hope, K. J., & Lu, Y. (2022). The splicing factor RBM17 drives leukemic stem cell maintenance by evading nonsense-mediated decay of pro-leukemic factors. Nature communications, 13(1), 3833. https://doi.org/10.1038/s41467-022-31155-0

58. Koedoot, E., Wolters, L., van de Water, B., & Dévédec, S. E. L. (2019). Splicing regulatory factors in breast cancer hallmarks and disease progression. Oncotarget, 10(57), 6021–6037. https://doi.org/10.18632/oncotarget.27215

59. Huang, X., Tan, W., Liu, Z., Fu, X., Li, Z., Lai, S., Li, Q., Zhong, X., Qu, F., Zhang, H., & Li, H. (2023). EIF4A3-induced circZFAND6 promotes breast cancer proliferation and metastasis through the miR-647/FASN axis. Life sciences, 324, 121745. https://doi.org/10.1016/j.lfs.2023.121745

60. Ju, C., Zhou, M., Du, D. et al. EIF4A3-mediated circ_0042881 activates the RAS pathway via miR-217/SOS1 axis to facilitate breast cancer progression. Cell Death Dis 14, 559 (2023). https://doi.org/10.1038/s41419-023-06085-4

61. Meng, X., Li, W., Meng, Z., & Li, Y. (2022). EIF4A3-induced circBRWD3 promotes tumorigenesis of breast cancer through miR-142-3p_miR-142-5p/RAC1/PAK1 signaling. BMC cancer, 22(1), 1225. https://doi.org/10.1186/s12885-022-10200-7

62. Zheng, K., Yao, S., Yao, W., Li, Q., Wang, Y., Zhang, L., Chen, X., Xiong, H., Yuan, X., Wang, Y., Zou, Y., & Xiong, H. (2021). Association Between RSK2 and Clinical Indexes of Primary Breast Cancer: A Meta-Analysis Based on mRNA Microarray Data. Frontiers in genetics, 12, 770134. https://doi.org/10.3389/fgene.2021.770134

63. Ma, J. H., Qin, L., & Li, X. (2020). Role of STAT3 signaling pathway in breast cancer. Cell communication and signaling : CCS, 18(1), 33. https://doi.org/10.1186/s12964-020-0527-z

64. Cao, L., Wang, R., Liu, G., Zhang, Y., Thorne, R. F., Zhang, X. D., Li, J., Xia, Y., Guo, L., Shao, F., Gu, H., & Wu, M. (2023). Glycolytic Pfkp acts as a Lin41 protein kinase to promote endodermal differentiation of embryonic stem cells. EMBO reports, 24(3), e55683. https://doi.org/10.15252/embr.202255683

65. Saha, S. K., Islam, S. M. R., Saha, T., Nishat, A., Biswas, P. K., Gil, M., Nkenyereye, L., El-Sappagh, S., Islam, M. S., & Cho, S. G. (2021). Prognostic role of EGR1 in breast cancer: a systematic review. BMB reports, 54(10), 497–504. https://doi.org/10.5483/BMBRep.2021.54.10.087