Article

A hybrid approach unveils drug repurposing candidates targeting an Alzheimer pathophysiology mechanism

Highlights

- Drug-repurposing approach that combines *in silico* analyses and *in vitro* screenings

- A drug- and mechanism-oriented model, the Human Brain Pharmacome (HBP) was created

- The HBP was used to mine data related to drugs and targets to generate a hypothesis

- Experimental evidence validated predicted drug-target combinations

Authors

Vanessa Lage-Rupprecht, Bruce Schultz, Justus Dick, ..., Marc Jacobs, Carsten Claussen, Martin Hofmann-Apitius

Correspondence

vanessa.lage-rupprecht@ scai.fraunhofer.de

In brief

In our drug repurposing approach, we combined knowledge mined from a variety of sources with focused experimental screening data of selected drug candidates into our drug-target-mechanism-oriented data model, the Human Brain Pharmacome (HBP). We identified previously unreported drug-target combinations that show evidence as being viable therapeutic candidates for Alzheimer disease (AD). Data-driven approaches combining *in silico* and *in vitro* analyses are increasingly in the spotlight and represent a future path to knowledge discovery, especially in the context of complex diseases.

CellPress

Article

A hybrid approach unveils drug repurposing candidates targeting an Alzheimer pathophysiology mechanism

Vanessa Lage-Rupprecht,[1,5,6,*] Bruce Schultz,[1,5] Justus Dick,[3] Marcin Namysl,[2] Andrea Zaliani,[3,4] Stephan Gebel,[1] Ole Pless,[3,4] Jeanette Reinshagen,[3,4] Bernhard Ellinger,[3,4] Christian Ebeling,[1] Alexander Esser,[2] Marc Jacobs,[1] Carsten Claussen,[3,4] and Martin Hofmann-Apitius[1]

[1]Fraunhofer Institute for Algorithms and Scientific Computing SCAI, Department of Bioinformatics, Schloss Birlinghoven, 53757 Sankt Augustin, Germany
[2]Fraunhofer Institute for Intelligent Analysis and Information Systems IAIS, NetMedia Department, Schloss Birlinghoven, 53757 Sankt Augustin, Germany
[3]Fraunhofer Institute for Translational Medicine and Pharmacology ITMP, ScreeningPort, 22525 Hamburg, Germany
[4]Fraunhofer Cluster of Excellence Immune-Mediated Diseases CIMD, ScreeningPort, 22525 Hamburg, Germany
[5]These authors contributed equally
[6]Lead contact
*Correspondence: vanessa.lage-rupprecht@scai.fraunhofer.de
https://doi.org/10.1016/j.patter.2021.100433

THE BIGGER PICTURE Owing to current setbacks in the discovery and development of novel treatments tackling Alzheimer disease (AD), a re-evaluation of research and development (R&D) strategies is underway. Here, we present a holistic pharmacological approach that combines drug-target information with knowledge graphs that represent essential pathophysiology mechanisms. The resulting Human Brain Pharmacome (HBP) embeds hundreds of relevant drug-target interactions in the context of disease mechanisms governing AD. We demonstrate how such a tool can be used to aid AD research by identifying already-approved drugs that have the potential to treat the disease, thereby bypassing the expensive and time-consuming task of researching and developing a new drug. In our study, we identified new drug-target combinations and provided mechanistic explanations that help to improve our understanding of AD pathology and support future development of effective therapeutic strategies.

1 2 **3** 4 5   **Development/Pre-production:** Data science output has been rolled out/validated across multiple domains/problems

---

SUMMARY

The high number of failed pre-clinical and clinical studies for compounds targeting Alzheimer disease (AD) has demonstrated that there is a need to reassess existing strategies. Here, we pursue a holistic, mechanism-centric drug repurposing approach combining computational analytics and experimental screening data. Based on this integrative workflow, we identified 77 druggable modifiers of tau phosphorylation (pTau). One of the upstream modulators of pTau, HDAC6, was screened with 5,632 drugs in a tau-specific assay, resulting in the identification of 20 repurposing candidates. Four compounds and their known targets were found to have a link to AD-specific genes. Our approach can be applied to a variety of AD-associated pathophysiological mechanisms to identify more repurposing candidates.

INTRODUCTION

Aging societies face a rapidly growing number of people suffering from AD, resulting in an enormous socioeconomic burden.[1] Billions of dollars are invested to address this problem, but despite all efforts to find effective therapeutics, treatment options are sparse and have largely failed to achieve satisfying results in the fight against AD. More than 100 candidate compounds were abandoned in development or showed disappointing results in clinical trials (CTs), leaving only a limited number of remedies that are currently approved and used for symptomatic treatment of AD.[2]

Why do CTs in AD treatment continue to fail? An incomplete understanding of pathophysiological mechanisms, erroneous selection of target and drug candidates, shortcomings of pre-clinical animal models, wrong drug dosage regimens, and false recruiting of best responders are some of the reasons why these clinical studies have so far failed.[3,4] Novel therapeutic avenues must be established, and achieving this requires a re-evaluation of the strategies currently used.

Drug repurposing, the process of identifying new therapeutic uses for existing drugs, has grown in popularity in recent years and has been shown to be quite successful.[5,6] This strategy, also referred to as drug repositioning, offers an enormous advantage: substances that have already been well examined and approved with a wide spectrum of data, including safety and efficiency profiles, can be critically evaluated in a new disease context. Drug repurposing maximizes the therapeutic application of drugs while minimizing the risk of failure, leading to a reduced period of time by which they are ready for the market with a new indication context.[7]

Repositioning of drugs can be driven by structural *in silico* analyses (e.g., molecular docking, pharmacophore modeling, virtual screening, or cryoelectron microscopy (cryo-EM) or X-ray structural determination)[8] or by the experimental domain (*in vitro* biochemical and cellular assays, high-throughput screening, *in vivo* animal tests),[9,10] or approached through a combination of both fields.[11] Generally, drug repurposing strategies are drug, target, or disease centric depending on the research focus or pharmaceutical interest.[12] In a drug-centric approach, a drug is connected to a formerly unknown target (off-target) and its target-associated indications, whereas in a target-centric approach, the drug-target relationship is known (on-target) and linked to a new therapeutic context.

New avenues of drug repurposing are being explored, including signature-based approaches that focus on evidence-driven identification of new repurposing compounds,[13] which is in line with our mechanism-oriented strategy. In the repurposing approach described here, we are not restricted to a single approach, therefore allowing greater flexibility in terms of repositioning candidate identification. We combined knowledge mined from a variety of sources with focused experimental screening data of selected drug candidates into our drug-target-mechanism-oriented data model, the Human Brain Pharmacome (HBP). Information retrieval and extraction were applied in a reciprocal manner for the enrichment of targets, ligands, chemical similarity, or assays in the HBP. The enrichment of our computational model was complemented by information extracted from the patent literature as well as quantitative information extracted from tables in published articles. Dedicated mining strategies enabled us to use this information for drug/target prospective and retrospective validation as well as for assay refinement. The dynamic nature of literature and the continuous research efforts in AD prompted us to develop and implement processes for progressive updates of the knowledge graph with several internal checks to ensure that novel and even controversial findings were added back to the HBP with full provenance of the information source.

Here, we demonstrate our approach in an AD-biomarker-centric search strategy. The HBP model includes a highly granular representation of mechanistic underpinnings of tauopathies, a class of neurodegenerative disorders characterized by the depo-

sition of abnormal microtubule-associated tau protein (MAPT) in the brain in the form of neurofibrillary tangles (NFTs).[14,15] Phosphorylation of MAPT (pTau), and thus creation of insoluble aggregates, is considered one of the key regulatory factors that drives AD pathology, and its abundance is used as a predictive biomarker of AD.[16] Based on our work in the pTau context, we implemented a generalizable workflow (Figure 1; Video S1) for the identification of new druggable targets in relation to AD-specific pathophysiology processes. In our *in silico*-driven knowledge mining approach, we focused on mechanisms related to post-translational modifications. Starting with pTau as the root node (primary target, see methods), we sought to answer the following questions in the context of this AD pathological hallmark:

1) Are there druggable targets (secondary targets, see methods) that modulate pTau?
2) Which of the targets identified in (1) are putative repurposing targets based on qualitative and quantitative data analysis?
3) Which repurposing compounds are highly active in our pTau-focused experimental setting?
4) Can we validate unknown drug-target pairs in the context of AD therapeutics?

RESULTS

The HBP

We created a drug- and mechanism-oriented model, the HBP, that covers a broad spectrum of biological processes relevant in the pathophysiological context of AD (Figures 2A and 2B). By augmenting it with metadata, quantitative data, and contextual information, we have created a comprehensive framework for *in silico* identification of druggable mechanisms and a broad information base for evaluating potential therapeutic targets and agents.

Identification of druggable modulators of pTau

Using our secondary target selection pipeline (see methods), we identified 77 druggable targets (cited in 155 documents) that also have an upstream causal interaction with pTau (Figures 2A and 2C; Data S1). Our initial list contained proteins from several different gene families: MAPK1, MAPK3, MAPK8, MAPK9, MAPK10, MAPK11, MAPK12, MAPK13, and MAPK14 (mitogen-activated protein kinases); PPP1CA, PPP2CA, PPP2CB, PPP3CA, and PPP5C (protein phosphatase catalytic subunits); PPP2R1A and PPP2R2A (protein phosphatase 2 regulatory subunits); S100A1 and S100B (S100 calcium binding proteins); CDK1, CDK2, CDK5, and CDK5R1 (cyclin dependent kinases); CAMK2A and CAMK2B (calmodulin dependent protein kinases); and ABL1 and ABL2 (Abl family tyrosine kinases). Cross-validation of protein-protein-interaction data from several resources (IntAct, BioGRID, Pathway Commons, STRING, https://bikmi.pharmacome.scaiview.com/) confirmed 54 of these targets as direct interactors of pTau (Data S2).

Qualitative and quantitative data analysis reveal putative repurposing targets

One of our primary goals was to find new or repurposed compounds for treating pTau in the context of AD. To this end, we
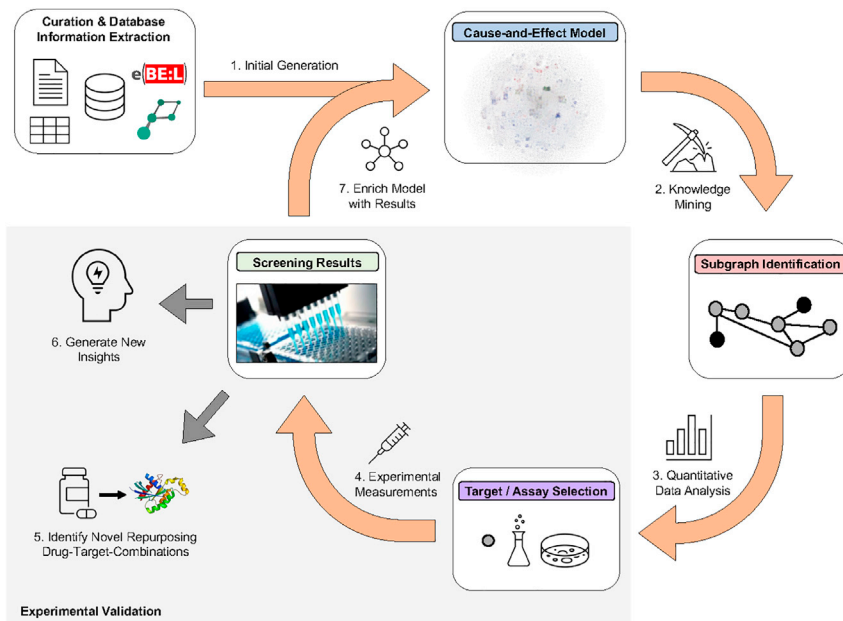
Figure 1. Generalized workflow for identifying novel therapeutic strategies
(1) A knowledge-based "cause-and-effect" disease model (here: Alzheimer disease) is initiated using manually curated information as well as standardized data from reputable repositories (see "The HBP" in results). (2) An entry point (primary target; here: pTau) associated with a pathophysiology mechanism of interest is selected, and upstream regulators (secondary interactors) modulating key elements of this mechanism are identified (see "Identification of druggable modulators of pTau" in results). (3) The secondary interactors are ranked according to chosen criteria, and the top hits are selected for further evaluation (see "Qualitative and quantitative data analysis reveal putative repurposing targets" and "Secondary-target candidate selection: HDAC6" in results). (4) Suitable biochemical and cellular assays are chosen for testing the activities of a library of compounds against top secondary interactors (see "Experimental validation of HDAC6 and candidate compounds" in results). (5) Experimental results are quantitatively analyzed to identify new repurposing drugs and compounds for each given secondary interactor (see "Mapping of identified HDAC6 inhibitors to the HBP" in results). (6) Novel target/compound combinations are designed and used to gain new insights for future experiments. (7) Earmarked combinations and screening data are mapped to the knowledge graph to enhance future queries (see "General applicability of the HBP" in results). From here, an entry point of interest can be selected to reinitiate the pipeline.

performed a statistical analysis of the secondary target metadata to identify the best candidate for additional screening (Data S1). First, we determined the number of BioAssays available for each secondary target to evaluate which of our candidate proteins had the most screening data available. These calculations are indicative of the number of known drug-target interactions as confirmed by several independent studies and serve as better values to evaluate protein/compound associations as opposed to looking purely at drugs that are approved and on the market. Our search resulted in a list of the top five ranked proteins: HDAC6, ABL1, CDK2, MAPK14, and SRC with 1,938, 1,929, 1,518, 1,490, and 1,363 BioAssay mentions, respectively. We then calculated the ratio of these counts to the number of available drugs for each protein, thereby determining which candidates had the greatest opportunity for drug repurposing exploration. Based on the information mined from DrugBank, HDAC6 had the highest BioAssay-to-known-drug ratio of the top five candidates with 1,938 available assays at the time of writing and only six approved drugs on the market, thus making it a prime candidate for testing potential new therapeutic compounds. Additionally, among the top five ranked proteins, HDAC6 had the fewest mentions within the parsed corpus of patents. Furthermore, research into the current clinical trial landscape revealed that there is only one trial in the recruiting phase in context of both AD and an approved HDAC6 selective compound (NCT03056495; Data S1). Therefore, assay availability, the lack of market-available or patented drugs, and the sparse clinical focus strongly argue in favor of HDAC6 as a suitable target candidate to demonstrate our repurposing selection workflow.

Secondary-target candidate selection: HDAC6
Research into HDAC6 revealed that it was previously the focus of cancer research[20] owing to its role in chromatin remodeling, cell motility, and gene regulation.[21] However, HDAC6 activity has also been shown to influence the structure of microtubules, thus making it an enticing candidate in the context of neurodegenerative diseases.[22,23] Interestingly, the HBP revealed a new potential intervention pathway: HDAC6 was shown to have a direct influence on acetylation of pTau.[19] According to Carlomagno et al., HDAC6 mediated deacetylation of tau at certain sites (such as K321; Figure 3A), which also allow for phosphorylation of the tau protein (pTau).[24] One such study reported that HDAC6 activity at K280 and K281 positively correlated with the pTau protein abundance,[25] while another group demonstrated that a direct acetylation of MAPT was shown to favor the generation of pathological pTau aggregates.[26] The study mentioned first suggests that the inhibition of HDAC6 and thereby the acetylation of K280 and K281, should lead to a decrease in harmful pTau aggregation by indirectly preventing phosphorylation from taking place at key residues. We have adapted this information in our hypothesis (Figure 2C). Though the dynamics of the post-translational modifications of MAPT, their mutual influence, and the downstream effects on the formation of pathological tau tangles are not fully understood, especially in a human context, HDAC6 does appear to have a central role in which disease-relevant signal cascades in AD are integrated.

Experimental validation of HDAC6 and candidate compounds
As classical HDAC6 assays often use consensus sequences derived from histone substrates, we sought to use a more specific readout for our experimental validation step. We therefore developed a substrate derived from amino acid (AA) sites 278–281 of the human tau protein, which was shown to be directly linked to pathological tau aggregation.[26] This peptide, Ile-Asn-(dimethyl) Lys-(ac)Lys was used for the development of our in-house
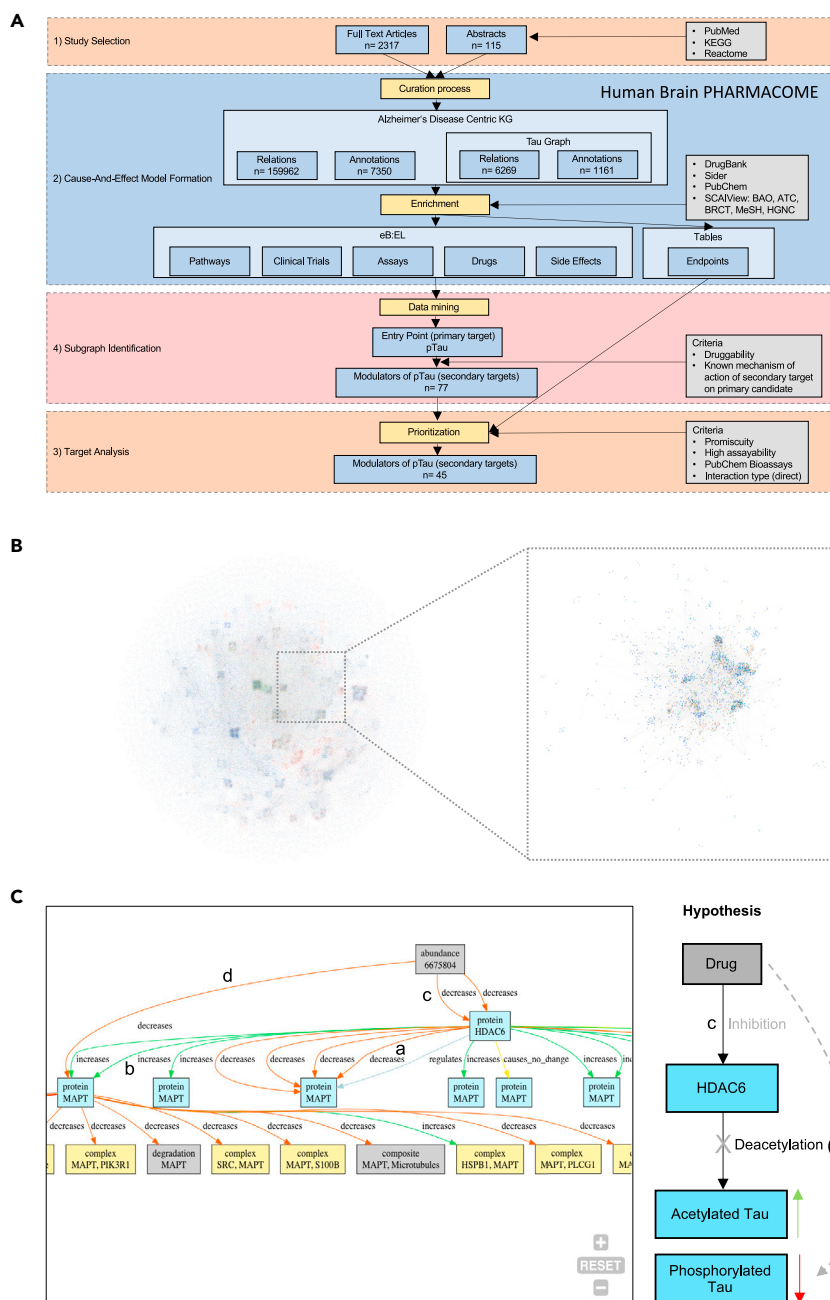
Figure 2. Cause-and-effect model generation, enrichment, and identification of druggable modulators of pTau

(A) Workflow illustrating the stepwise formation of a data-enriched cause-and-effect model, subsequent knowledge mining, and target analysis.

(B) Beginning with the entire Human Brain PHARMACOME, a primary target is selected (pTau) and its cause-and-effect context (tau subgraph, magnified square) is defined.
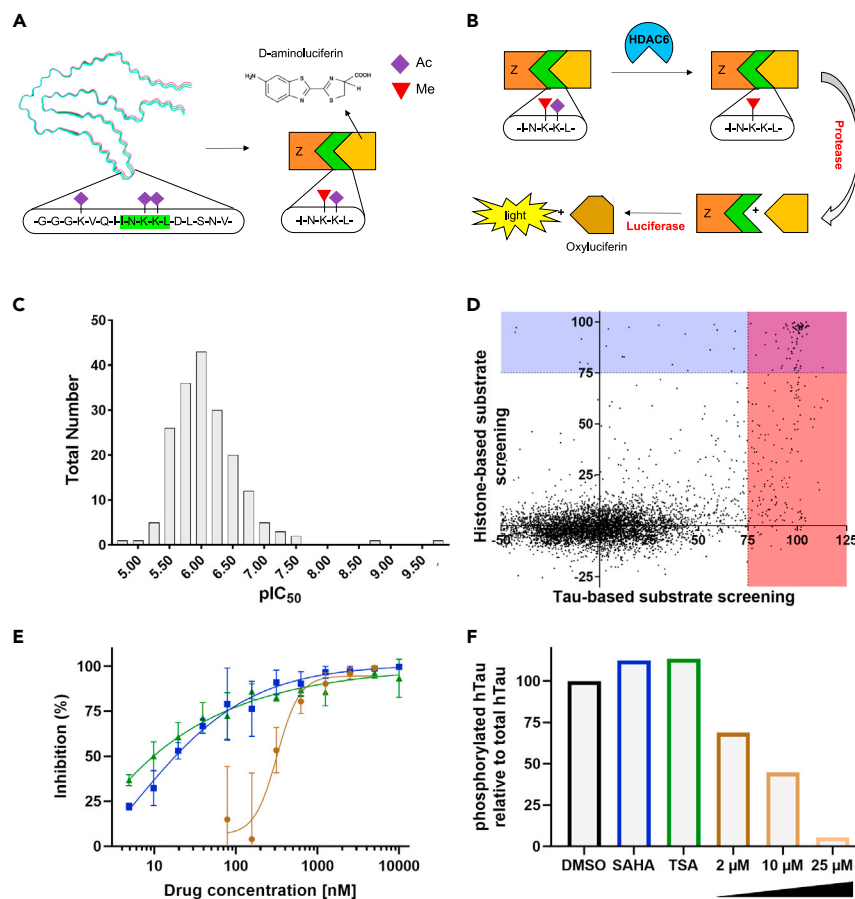
(C) Illustration of a drug-HDAC6-MAPT relational network[17,18] in BiKMi, available at https://bikmi.pharmacome.scaiview.com/ left). The graph can be generated by applying the path query: From: "6675804", To: "MAPT", Depth 2-2. The "Depth" limits the size of the network by specifying the number of edge steps searched for between the "From" and "To" entities. Each node (box) represents a specific subject/object (according to BEL). Each edge (line) is a unique source of evidence. The graphical representations from BiKMi are simplified representations of BEL subgraphs. Because it was listed as one of the druggable secondary targets, HDAC6 was further investigated using this tool. This network provided evidence (evidence boxes not shown) that HDAC6 is known to deacetylate tau[19] (edge highlighted with "a"), which correlates with tau phosphorylation[17] (edge highlighted with "b"). As the network (left) indicates, pharmacological HDAC6 inhibition (edge highlighted with "c") resulted in reduced tau phosphorylation[17] (edge highlighted with "d"). Based on these known relationships, we hypothesized (right) that inhibition of HDAC6 (indicated by "c") blocks (crossed out edge) the tau deacetylation process (indicated by "a"), which leads in turn to a decrease in pTau (effect indicated by "d," dotted arrow; reduction of pTau is indicated in red) by indirectly preventing phosphorylation at key residues through remaining acetylation (green). The edges (left and right) marked with the same lowercase letter correspond to each other.

In total, 5,632 compounds were screened on 16 plates with a z′ exceeding 0.5 on all plates. The average signal intensity of all data points was 1.4% inhibition, indicating no systematic errors were present within the screening, with a standard deviation of 0.35. Interestingly, 65% of the 279 active molecules from the primary screening were confirmed as hits, with a hit being defined as a compound showing at least 75% inhibition. Such a confirmation rate is reasonable for this type of complex and coupled enzymatic assay. To validate the 183 confirmed active compounds, a second HDAC6 screening was conducted in addition to dose response studies (Figure 3D). This assay employing a shorter, less specific peptide derived from histone H3 resulted in 104 compounds inhibiting to at least 75%. From these, 23 molecules were found to be specific for the histone H3 substrate, and 81 inhibited the conversion of both substrates. As a result, 102 compounds from the tau-based screening were specific for this

HDAC6 assay (Figures 3A and 3B) and validated against two unspecific standard inhibitors, vorinostat (SAHA) and trichostatin A. These inhibitors had half-maximal inhibitory concentration (IC_{50}) values of 3.1 nM and 2 nM respectively, concentrations that are similar to those found in a histone-based HDAC6 assay (6.2 nM and 3.8 nM, respectively).[28] This result not only reflects the overlapping specificity of the two catalytic HDAC6 domains,[29] but also demonstrates the feasibility of the approach and the direct involvement of HDAC6 at K281 deacetylation of human tau. To identify promising HDAC6 inhibitors that can modulate its activity on human tau, we screened a collection of drugs to identify those with the greatest effect on tau peptide deacetylation.

Figure 3. Schematic representation of the assay principle and screening results of HDAC6 inhibitors

(A) Schematic representation of the developed peptidic HDAC6 substrate. Cryo-EM structure was used to represent MAPT/hTau.[27] Acetylated lysines are highlighted. A D-aminoluciferin coupled peptide was developed, containing a dimethylated and an acetylated lysine. One letter abbreviation "Z" refers to the Cbz protecting group of the peptide.

(B) Workflow of the coupled enzymatic HDAC6 activity assay. HDAC6 (highlighted in blue) deacetylates the peptide, which becomes a substrate for a protease, thereby releasing the D-aminoluciferin from the peptide. Luciferase is catalyzing the last step of this cascade by converting D-aminoluciferin to oxyluciferin in a light-emitting reaction.

(C) All active compounds were subjected to dose response studies, and pIC_{50} values were determined.

(D) Validation of the hit population using two different substrates revealed distinct hit populations. Hits highlighted in blue are compounds acting only on histone-based substrate, hits in the red area are active against the peptidic substrate derived from MAPT/hTau sequence, while the purple corner shows compounds active against both substrate assays.

(E) Activity of the compound VU0361737 (brown), relative to vorinostat (SAHA; shown in blue) and trichostatin A (TSA; shown in green). Data are shown as mean with SD of three independent experiments.

(F) Densitometric analysis of a western blot (see also Figure S3) depicting SAHA, TSA, and increasing concentrations of VU0361737.

substrate and subjected to dose response studies (Figures 3C and 3D). In the following step, their activities were confirmed, and the 20 molecules that specifically were found to be the most active in inhibiting tau-based substrate conversion and had favorable physicochemical properties (Figure 3; Table S3) were selected. Suitable properties included drug likeliness, taking into account molecular weight (MW), solubility, and rule of five criteria,[30] as well as likeliness for central nervous system (CNS) penetration (topological polar surface area [TPSA] below 100) and having fewer than 8 rotatable bonds, although certain molecules were still included even if one of the rules was violated (Table S4). These compounds were further validated using western blots against pTau (Figures 3E and 3F). The reduced phosphorylation of tau further demonstrated the applicability of HDAC6 inhibitors in the context of tau hyperphosphorylation and may represent promising avenues for AD therapy.

Mapping identified HDAC6 inhibitors to the HBP
The compounds found to inhibit HDAC6 activity were queried against the ChEMBL database to determine any additional drug targets for these molecules (Figure 4A). During data collection, we made sure that two parameters were controlled: pCHEMBL >6 and confidence level >7 were used. A pCHEMBL value of more than 6 signifies that all measurementss of inhibition are less than 1 μM (pIC_{50} = negative log 6 [IC_{50} in μM]). The confidence level corresponds to the ChEMBL confidence score (see

https://chembl.gitbook.io/chembl-interface-documentation/frequently-asked-questions/chembl-data-questions).With these two parameters, we were able to control the quality of the data sampled as well as the potency of the interactions surveyed. Of the 20 compounds, 17 were found to interact with targets other than HDAC6. Using information collected from the GWAS Catalog,[31] a list of known, disease-associated genes was compiled based on identified single nucleotide polymorphisms (SNPs) ($p < 10^{-7}$) and compared against the compound targets. Only four of the drugs—quercetin (ChEMBL: CHEMBL50), GW441756X (ChEMBL:CHEMBL1516890), Debio 1347 (ChEMBL:CHEMBL3907479), and indirubin-3′-monoxime (ChEMBL:CHEMBL216543)—targeted proteins that were found to be associated with AD (Figure 4B; Data S3). It was not surprising that quercetin and GW441756X would have additional targets of interest, as these compounds were found to interact with the greatest number of molecules from those in our list (80 and 62, respectively). Interestingly, of the eight overlapping hits from the AD association analysis, three were found to be targeted by multiple drugs in our list. CDK1 and CSNK2A1 are both targets of quercetin and GW441756X, while GSK3B was also found to be modulated by these compounds as well as indirubin-3′-monoxime. Based on the information collected from the GWAS Catalog, we were able to calculate what percentage of each compound's target list was associated with AD (Figure 4C). We found that indirubin-3′-monoxime had the highest
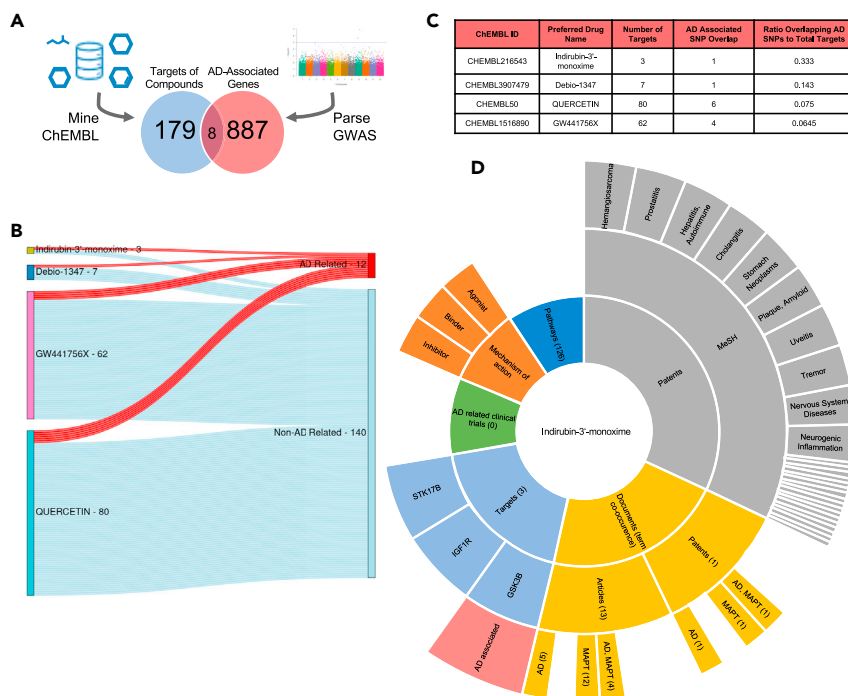
Figure 4. Computational analysis of drug repurposing candidates

(A) Targets of screened compounds were mined from the ChEMBL database and compared against AD-associated genes parsed from the GWAS Catalog. Sankey diagram depicting the ratio of AD-associated (red edges) to non-disease-related (blue edges) targets of the four compounds described in (C).

(B) Node labels indicate the number of targets associated with a given compound or disease state, and bar size is proportional to this measurement.

(C) Overlap analysis resulted in four compounds that targeted proteins associated with AD.

(D) Summary of computational analysis for a given compound (here: indirubin-3′-monoxime). Analyses of other compounds available in Figure S5 and Sunbursts in Data S5.

score, with 33.3% of its targets being associated with AD, while Debio-1347 had the second highest score at 14.7%. This suggests that indirubin-3′-monoxime would have the greatest chance of repurposing due to its strong focus on AD-related targets.

To extract additional information on the candidate compounds in relation to the primary (pTau)/secondary (HDAC6) targets, queries consisting of concepts and keywords were generated and executed using https://pharmacome.scaiview.com/ (Figure 4D; Data S4). Debio-1347 was not found in the context of HDAC6, MAPT, or AD, whereas quercetin has clear connections to these molecules (230 and 1,416 mentions, respectively) and AD (1,679 mentions). Meanwhile, both indirubin-3′-monoxime and GW441756X were found to be mentioned in publications in association with MAPT (13 and 4 mentions, respectively) and AD (6 and 7 mentions, respectively). Next, we searched both ClinicalTrials.gov and patent literature for recent information on the candidate compounds and any ongoing investigations into their therapeutic uses. Again, among these four compounds, only quercetin has been used in a clinical trial performed in the context of AD (a total of 4 trials were found). A patent search via dedicated concept queries and analysis of disease-related topics (Medical Subject Headings [MeSH]) extended the coverage of the elucidated contextual information. For Debio-1347 and GW441756X compounds, cancer-related terms dominated the results, with a relevance score (Kullback-Leibler divergence) of 6.12 for "small cell lung carcinoma" (Debio-1347) and 5.98 for "glioblastoma" (GW441756X), whereas indirubin-3′-monoxime appeared in an inflammatory context first ("prostatitis" relevance score: 3.67), followed by "hemangiosarcoma" (relevance score: 3.65). For quercetin, a broad range of disease context information could be identified (Figures S4 and S5; Data S4). All four compounds were mentioned together with the MeSH terms "arthritis," "asthma," "colitis," "fibrosis," "hemangiosarcoma," "multiple myeloma," "psoriasis," "stomach neoplasm," and "wounds and injuries." Only patents containing the compounds quercetin and indirubin-3′-monoxime specifically contained MeSH terms in the context of neurodegenerative disorders (quercetin: "Alzheimer disease," "Parkinson disease," "nervous system diseases"; both: "plaque, amyloid").

Because of their relevance to AD (Figure 4B), these four selected compounds and their identified targets were mapped to the HBP to improve the performance of subsequent iterations of the pipeline. A new node was created for any drug or protein that did not already exist within the network, and generated edges were annotated with the parsed interaction information.

In summary, our computational analysis identified repurposing candidate compounds based on their proximity to protein targets and biological pathways relevant to AD pathology. Among those, indirubin-3′-monoxime had the highest score based on the ratio of AD-relevant genes and pathways it targets, suggesting that it would serve as a suitable, focused therapeutic, while minimizing potential side effects due to off-target interactions. Furthermore, cellular characterization of AD-relevant cell types and evidence collected from literature that include this compound revealed the mechanism by which it can positively regulate AD, thereby strengthening it as a potential repurposing candidate and further validating our pipeline.

General applicability of the HBP

With the HBP, we have created a computational model that links drugs, their targets, and highly associated mechanisms. This complex framework of interactions spans multiple modalities and consists of biological processes complemented by highly granular representations of biological mechanisms. Though this model focuses primarily on AD, it can also be used to explore other diseases that share pathological aspects with AD. The HBP is enriched with qualitative and quantitative data, thus facilitating the ability to filter data using a variety of metadata, the

development of specific assays, and the assessment of experimental data.

Though our pipeline was initiated using a modified protein, it is not limited to this type of entity. A primary target can also be a biological process in the form of a single entity or an entire subgraph. A drug could also serve as the starting point for *in silico* processing and *in vitro* validation (or vice versa). The flexibility of this workflow enables far-reaching possibilities in terms of potential biological questions and hypothesis generation.

For example, assume that a biological process, such as autophagy, is the entry point (primary target) of our workflow. Pathological autophagic processes play an essential role across diseases[32] as in AD.[33] Therefore, this primary target would allow one to start from the *in vitro* perspective and answer the following question: Which repurposing compounds have a regulatory effect on autophagy? In a subsequent step, one could determine where the compound-specific targets (secondary targets) are located in the graph. These targets could be in a relationship with autophagy itself or affect a process or another protein target that are elements of the autophagic process. This information can be used to hypothesize potential mechanisms of action in the autophagy context, which in turn could be tested via appropriate *in vitro* assays. In theory, the HBP could provide possible explanations for how compounds work in the autophagy context.

In summary, our network and the associated hybrid workflow offer a wide range of opportunities to answer open questions in the search for appropriate approaches to new therapeutic treatments, extending beyond the context of AD.

DISCUSSION

AD research has suffered greatly from a lack of successful clinical trials and a scarcity of effective treatment options.[2] It has taken 18 years for a new drug, the antibody aducanumab which reduces Aβ plaques in AD patients,[34] to be approved, but its efficacy is highly controversial.[35] These shortcomings are due in part to our lack of understanding of the underlying mechanisms governing the disease, resulting in massive loss of time and money spent testing compounds that were not well understood. To avoid repeating such mistakes, we have developed the HBP, a comprehensive network of molecular interactions focused on the pathways and mechanisms that regulate the pathophysiology of AD.

Here, we demonstrated that the HBP can aid researchers in quickly screening and identifying novel therapeutics by applying the principles of drug repurposing to AD. In our example, we experimentally validated 5,632 compounds against HDAC6, a secondary interactor that was identified by our pipeline as being a druggable regulator of our biomarker of interest, pTau. Our results showed that of the compounds screened, 183 successfully increased the amount of acetylated tau, thereby decreasing the number of sites available for phosphorylation. Further characterization of these compounds revealed that four compounds already had links to AD (quercetin, GW441756X, indirubin-3-monoxime and Debio-1347), and as a result, metadata for these compounds were gathered and added to the HBP via text mining and table extraction approaches for use in subsequent iterations of the workflow.

To ascertain their potential in treating AD, we investigated our list of compounds and found evidence to support our selection. Quercetin is a flavonoid found in fruits and vegetables and has a wide range of biological associations, such as anti-carcinogenic, anti-inflammatory, and anti-viral activity.[36] GW441756X (synonym: GW-441756, see ChEMBL1516890) is a tropomyosin receptor kinase A (TrkA) inhibitor.[37] TrkA signaling has been recently described as a potential therapeutic target for pain[38] and, moreover, is suggested to contribute to degeneration of basal forebrain cholinergic neurons (BFCNs), which is associated with cognitive decline in AD.[39] Debio-1347 is a selective fibroblast growth factor receptor (FGFR) inhibitor and exhibits high antitumor activity. Fibroblast growth factors (FGFs) and their receptors have been shown to be involved in the pathogenesis of Parkinson disease and AD.[40,41] Indirubin-3′-monoxime is an inhibitor of cyclin-dependent protein kinases (CDKs), and recently inhibitors of these enzymes have been of interest for the treatment of cancer. CDK isoforms have also been shown to play a key role in cancer proliferation via abnormal cell-cycle regulation.[42] Interestingly, indirubin-3′-monoxime is described to play a role in neuronal apoptosis in AD,[43] and in Sf9 cells expressing human tau 23, the compound was found to suppress tau phosphorylation.[44]

Though these compounds all had links to AD and were of great interest to our drug repurposing efforts, one of them did not have enough metadata available for proper analysis because it is relatively novel. Indeed, VU0361737, a positive allosteric modulator (PAM) of metabotropic glutamate receptors,[45] did show promising results in our pTau specific assay. Interestingly, this compound is mentioned in a patent together with HDAC6 and MAPT (US patent 10500232); however, no focus on neurodegeneration is described.

The presented examples also highlight another problem of drug discovery approaches that target neurodegenerative diseases, which is the complexity of the underlying biology. Human tau can be phosphorylated at 85 different sites within the protein, and each site plays a different role in regard to microtubule binding and aggregation.[46] Two important post-translational modifications that have arguably the largest influence on activity of tau are phosphorylation and acetylation. Tau can be acetylated at more than 20 sites[47] by p300 or CREB binding protein or via autoacetylation[48] and deacetylated by SIRT1 and HDAC6.[48] Some of these acetylations are critical for tauopathies, for example, K280 and K281, on which the peptidic substrate of this work was based.[47] Unfortunately, the regulation of the acetylation patterns and their effect on phosphorylation and aggregation are complex. This complexity extends to the proteins responsible for the post-translational modifications, such as the protective versus detrimental effects of both SIRT1 and HDAC6 in terms of known tauopathy symptoms.[49] While some groups report a protective effect of HDAC6,[50] others suggest that inhibiting this protein can lead to improvement.[51,52] Some of these effects might be cell line specific or more relevant for certain tau isoforms, which adds another layer of complexity. The N2a cells used in this study represent a simple and robust model, which expresses many different tau isoforms.[53] More complex cell lines[54] as well as models based on induced pluripotent stem (iPS) cells can be applied in the future[55] to achieve a more profound biological validation of the compounds for progression

toward *in vivo* efficacy testing, but this is beyond the scope of the present study. Here, we emphasize the mechanism-centric approach complemented by targeted experimental validation, which should serve as an inspiration for the development of new drug repurposing approaches. In the future, the HDAC6 modulators identified in this study may help to dissect some of the contradicting effects observed by tau acetylation.

In silico methods are becoming increasingly important in the pharmacological setting and have found their place alongside *in vitro* and *in vivo* methods.[56,57] Combining *in silico* with *in vitro* techniques is a relatively new approach, but it shows enormous potential in drug repurposing and drug discovery.[58] In particular, the coronavirus disease 2019 (COVID-19) pandemic has helped to rapidly advance these approaches, but it has also highlighted its limitations.[59] Models are currently limited to what information is present in databases and their underlying network, while limited capabilities restrict what can be validated in *in vitro* experiments. Another essential factor to consider is the lack of understanding of key biological mechanisms. With all these factors in mind, we sought to address as many of them as possible using the high-level HBP concept. The HBP benefits from a unique constellation of expertise, both *in silico* and *in vitro* based, and through mutual interaction, enrichment through the integration of experimental results, and periodic reviewing, the model continues to grow and evolve resulting in a dynamic computational ecosystem.

We are aware that our HBP concept is not free of limitations and biases. Directed data toward AD and tau pathology are complemented by a largely unbiased collection of data to balance the model. Other biases arise from our target-based filtering criteria. Selecting targets based on available screening data might introduce historical bias and give more weight to unsuccessful targets previously hyped, mostly in areas outside neurodegenerative research. Moreover, the non-exhaustive number of patents chosen as a filtering criterion to favor targets might create a tendency to exclude potentially robust, but underutilized targets associated with AD and other neurodegenerative diseases. We deliberately chose these criteria to specifically highlight more novel targets in the AD context. The use of protein-protein-interaction (PPI) data to narrow down target candidates is a tool that should also be considered with caution, as PPI data may not have been generated from CNS cells and may be limited in the types of target interactions listed in the specific use case.

Our strategy is generalizable and applicable to a wide variety of conditions as we have shown in Schultz et al.,[60] where we used a similar drug target mechanism graph (COVID-19 PHARMACOME) to predict the synergistic effect of specific drug pairs for combination therapy of SARS-CoV-2 infections. Our method is also flexible in terms of the entry point: from an *in silico* perspective, the context-enriched, qualitative biomedical knowledge graph is used to mine data related to drugs or targets to generate mechanistic hypotheses that can be validated through experimental testing. In turn, an *in vitro*-centric strategy would generate experimental output about putative drug-target combinations or actions of drugs on biological processes, which is then underpinned by computational insights about associated biological mechanisms as well as referential drug, target, genetic, and disease data.

With our drug repurposing strategy, we have found previously unreported drug-target combinations that represent very interesting candidates for AD therapeutic purposes. Furthermore,

we have gained knowledge about potential new targets through our dynamic HBP concept. Our methodology and the resulting findings could steer research in a new direction and pave the way toward successful therapeutic intervention.

Data-driven approaches are increasingly in the spotlight as the amount of available data and metadata grows and artificial intelligence (AI) approaches offer a way to keep track and derive insights. In particular, cross-validation over everything known, mediated by an interplay of *in silico* and *in vitro* analysis, represents a future path to knowledge discovery, especially for complex diseases.

EXPERIMENTAL PROCEDURES

Resource availability

Lead contact
Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Vanessa Lage-Rupprecht (vanessa.lage-rupprecht@scai.fraunhofer.de).

Materials availability
This study did not generate new unique reagents.

Data and code availability
The data and code for extracting and ranking druggable interactors are available at https://doi.org/10.5281/zenodo.5779103. A Jupyter Notebook is included in the code repository demonstrating how the repurposing target list for this article was generated. The HBP itself (i.e., the knowledge graph used for this ranking analysis) is available at https://graphstore.scai.fraunhofer.de and labeled as the "pharmacome" database. The HDAC6 screening dataset using tau-based substrate in an enzymatic assay is available at https://doi.org/10.5281/zenodo.5578613. Any additional information required to reanalyze the data reported in this article is available from the lead contact upon request.

Methods

Generation of a compound screening library
The collection of 5,632 compounds was assembled by an external partner (SPECS) in a manner aligned to the recommendations from the Broad Institute (Cambridge, MA, USA).[61] In assembling this "mirror" collection, compounds were purchased from the same set of more than 70 high-quality suppliers identified by the Broad Institute and quality controlled by liquid chromatography/mass spectrometry (LC/MS) for purity and identity (minimum purity >90%). Compounds at 10 mM were stored in 100% DMSO at $-20°C$. A curated database is available containing the compound, indication, primary target (where known), and mechanism of action as well as analysis tools that can assist in mechanism of action determination and target elucidation.[61] In total, the library contains 2,350 launched drugs, 95 withdrawn drugs, 1,600 drugs in clinical phases, and 1,587 drugs in preclinical development.

Construction of an AD-centered OpenBEL model
We re-curated our existing AD model[62,63] and adapted all relations to the latest version of OpenBEL (BEL v2.1; https://bel.bio/). BEL allows one to encode causal and correlative relations between entities, including information such as origin of the relation (PMID) and annotations of associated metadata (e.g., species, tissue, assay, or disease). The revision was carried out using defined processes, implemented in the software systems e(BE:L) and Py-BEL.[64] We expanded our model by further curation and enrichment procedures to increase the granularity of the existing data and incorporated additional biological information to sharpen the AD pathological picture (Table S1). In total, we curated 115 abstracts and 2,317 full-text articles, which yielded 159,962 relations (Table S1) with more than 7,350 annotations. Further, 13,580 drugs from DrugBank were mapped directly onto the HBP itself and were supplemented with compounds extracted from ChEMBL. The tau subgraph was identified and labeled within the HBP and comprised 6,269 relations (Table S1) with 1,161 unique annotations.

Enrichment of the OpenBEL model
To supplement our model with interactions that were not directly cited in the curated publications, we compiled data from several repositories and

converted them into the OpenBEL format so that we could enrich our AD model with established biological relationships and metadata (Table S2). Pathway data from KEGG,[65,66] Pathway Commons,[67,68] and Reactome[69] were transformed and normalized[70] to complement our AD knowledge graph as well as PPIs from IntAct,[71] BioGRID,[72] and StringDB[73] As our primary goal was to create a tool for investigating drug repurposing candidates, we also included data form both DrugBank[74] and Clinical Trials[75] to identify known drug interactions and the disease contexts in which they were studied, respectively. After this enrichment step, the total number of edges in our graph increased to 728,809, while the tau subgraph grew to 7,100 relations.

Integration of quantitative information
To extract quantitative information, we employed our general-purpose table extraction system that was developed to extract information from unstructured scientific documents, in which relevant content is arranged in tabular format (Figure S1). The details of this system are presented in Namysl et al.[76] Here, we will give a brief overview of the table recognition method and highlight its experimental evaluation.

We were able to parse two popular input formats: the born-digital PDF files and the digitized document images. Our table recognition system employs two ad hoc heuristics for table structure recognition. The first heuristic recognizes tables that are typeset with a LaTeX package *booktabs*. The second algorithm handles the bordered table format. Moreover, the system includes a regular expression- and graph-based table interpretation method for attribute-value pairs extraction. This complementary table interpretation phase additionally ensures that the tabular structure was correctly recognized. The system was validated using two well-known table recognition benchmarks.[77,78] It achieved results that are competitive with the best academic table recognition approaches on the ICDAR 2013 benchmark,[77] which consists of digital-born government and business documents. The flexibility of the system was additionally validated using the ICDAR 2019 dataset,[78] which comprises low-quality scanned document images. In particular, our table recognition system reached highest precision among the examined commercial and academic table recognition methods on both datasets. For more details about the evaluation using the general-purpose benchmarks, please refer to the article describing our table recognition system.[76]

Moreover, we evaluated our system using an in-house biomedical dataset consisting of documents collected in the context of tau aggregation and retrieved from PubMed Central.[79] The ground-truth annotations were generated semi-automatically from the XML article content and then manually curated. In total, our benchmark consists of more than 1,100 PDF documents containing over 1,600 tables. We compared our table recognition approach with a leading commercial document analysis system and two popular open-source table recognition toolkits. We outperformed all baselines in terms of precision in this scenario. The detailed description of our biomedical benchmarking dataset, the evaluation procedure, the experimental setup, and the results are presented in Adams et al.[80]

Furthermore, in Figure S2, we present an example of a table recognized using our system.

Integration of contextual information
For retrieval of contextual information, we used our semantic search engine https://pharmacome.scaiview.com/. This repository contains more than 37 million articles and 2.7 million reviews as well as more than 10 million patents.

More than 1.15 million documents were annotated with ontologies and terminologies, including the Anatomical Therapeutic Chemical (ATC) Classification System, Bioassay Ontology (BAO), the Brain Region and Cell Type terminology (BRCT), the DrugBank Database, HUGO Gene Nomenclature Committee, and Medical Subject Headings (MeSH). We performed queries to identify literature containing drug, target, and disease relations and used ontologies such as BAO to filter selectively for assay information. MeSH-specific analytics in patents for selected candidate compounds was used to identify disease-related terms, which were investigated in the context of the corresponding compound (Figure S4). Terms were ranked according to Kullback-Leibler divergence, an information-based measure of disparity among probability distributions.[81]

Selection of disease-related primary target/candidate compound
The HBP represents a high-level concept with various entry points, such as genes, proteins, drugs, or even cellular processes. The primary target defines the center node from which we start our mining routine.

Selection of the primary target in the context of AD was done based on one of the following criteria:

- Biomarker qualified by the European Medicines Agency (EMA) (pTau)
- Suggested biomarker identified in relevant literature
- Drug action indicated to be significantly relevant for AD therapeutic intervention

Secondary target selection pipeline
As many therapeutic strategies focus on regulating a key biomarker directly through compounds and medication, we chose to expand our range of targets by first establishing a workflow to identify interactors (secondary targets) that work upstream of a given node (primary target). In short, the pipeline is initiated with the gene symbol of a given protein of interest as well as with an optional post-translational modification. The node matching these parameters is identified within the graph, and a list comprised of neighboring nodes found to have a "causal" relationship to the matched node is generated. Individual proteins are extracted from these identified neighbors and are considered a possible secondary protein-target interactor.

For a secondary target to be valid for further consideration, the following conditions were defined:

- Druggability: A drug or compound targets and modulates the activity of the protein. Proteins in the model have links to drugs and compounds extracted from DrugBank as well as those added by experimental evidence.
- Known mechanism of action of secondary target on primary candidate: The effect of the secondary target on the primary candidate (either activation or inhibition) must be known to discern the outcome of a drug on the activity/abundance of the primary target.

Identified secondary targets were further filtered based on following criteria:

- Promiscuity: Proteins within our model that are considered "hubs" (edge count >1,000) were excluded from our list. Targeting such proteins would likely result in unforeseen downstream effects in associated pathways and processes.
- High assayability: To verify our findings, secondary targets for which an assay exists that allows one to measure their effect on the activity or abundance of the primary target were given preference.
- PubChem BioAssays: The number of assays available according to PubChem[82] were considered during evaluation. Secondary targets were excluded and labeled with "suboptimal experimental accessibility" when no confirmatory assay was given, while the highest numbers were taken as a measure for "degree of focus of interest."
- Interaction type: secondary targets that could not be clearly confirmed as direct interactors modulating pTau were excluded.

PPI analysis
A PPI analysis was performed to limit the list of druggable pTau modulators to possible direct interactors. We extracted our druggable targets list from BiKMi and compared these PPI data with four other sources: IntAct, BioGRID, Pathway Commons, and StringDB. Selection criteria were applied only when extracting data from String DB. It was specified that for a valid interaction an experiment had to be present. Targets were sorted by frequency of occurrence. Targets that occurred repeatedly in three or four sources were recognized as putative MAPT interactors with direct influence on pTau. Targets with fewer than two mentions were manually reviewed and rejected if biological data indicated an indirect influence on pTau (e.g., signaling cascades between primary and secondary target) and/or if a direct influence on phosphorylation could not be supported by a study. For each putative direct modulator of pTau, a reference was given.

Drug prioritization pipeline
Once secondary targets were identified and filtered, a list of drugs and compounds was generated for each protein. Each compound was ranked using the available metadata:

- Primary target outcome: The effect of the drug on its target results in the desired downstream change in activity or abundance of the primary

target. Drugs producing unknown or undesirable outcomes were excluded from further analysis.

- Patent: Drugs that are currently patented were not prioritized.
- Clinical trials: Drugs that have not been tested in a clinical stage II trial in context of AD were preferred.
- Generics: Drugs with generic, cost-effective versions available were preferred over drugs that are on patent.
- Promiscuity: Compounds with fewer targets (thus being more specific) were ranked higher than those that bound non-specifically to a wide range of proteins.

HDAC activity assay

The assay analyzing the tau-based substrate was designed around a coupled enzymatic reaction (Figures 3A and 3B). Deacetylation of a luciferi- labeled substrate peptide through HDAC6 allows for trypsin, a serine protease, to cleave an N-terminal aminoluciferin, thereby enabling a quantifiable biolumi-nescent reaction. Assays were performed in 384-well microtiter plates (Greiner 784904) with 15 μL total volume. Before the assay, 20 nL of compound or con-trol substance was added to designated wells using an Echo550 liquid handler (Beckman Coulter, USA). Initially, 10 μL enzyme buffer containing 2 nM HDAC6, 50 μM aminoluciferin peptide, and 0.05% BSA in HEPES buffer solu-tion (25 mM, with added 137 mM NaCl, 2.7 mM KCl, and 1 mM MgCl$_2$, pH 7) was added to each well and incubated for 30 min at room temperature. After incubation, 5 μL of detection reagent (0.5 μg luciferase [Biomol, Germany]) (final concentration 0.067 μg/μL), 2 μg trypsin (Sigma-Aldrich, USA) (0.133 μg/μL), and 400 μM ATP (Jena Bioscience, Germany) (final concentra-tion 133 μM) in HEPES (25 mM, with added 137 mM NaCl, 2.7 mM KCl, and 1 mM MgCl$_2$, pH 7) were added and measured directly using an EnVision multi-mode reader (PerkinElmer, USA) at 200 ms integration time.

To analyze a histone-based peptide substrate, HDAC-Glo I/II Assay and Screening System (Promega, USA) was used according to the instructions of the manufacturer. Briefly, reagents were prepared as described in the manual, and plates and compounds were handled as described previously. HDAC6 was added to all wells in 5 μL assay buffer (0.2 nM, final concentration 0.1 nM) and incubated for 10 min. Afterward, 5 μL of substrate reaction was added and incubated for 10 min followed by luciferase measurement using an EnVision multimode reader as described previously. Assay quality was assured by Z′ calculation with all plates above 0.5 regarded as valid.

Western blot

Neuro-2a cells were cultivated in 75 cm^2 flasks using DMEM with penicillin (100 U/mL), streptomycin (0.1 mg/mL), L-glutamine (2 mM), and FBS (10%). Medium was changed every third day, and cells were incubated at 37°C and 5% CO$_2$. Passaging was done in 1:5 dilution, as soon as confluency was around 90%, usually once a week. For cell seeding, the monolayer was washed using 10 mL PBS, followed by detachment and dissociation of cells using 1.5 mL of 0.05% trypsin/0.02% EDTA solution (Capricorn Scientific, Germany) for 3 min at 37°C. Cells were resuspended in culture medium, centrifuged at 300 × g for 5 min, resuspended in fresh medium, and counted.

Cells were seeded at 500,000 cells per well in non-coated 6-well plates (Greiner BioOne, Germany). After 12 h, compounds and controls were added at the desired concentrations in 4 mL total volume and incubated for an addi-tional 12 h. After treatment, the wells were washed with PBS for 3 min, and 150 μL lysis buffer (RIPA buffer [Thermo Fisher Scientific, USA] supplemented with protease and phosphatase inhibitor cocktails [both Sigma-Aldrich, USA]) was applied, followed by a shaking incubation for 10 min on ice. Cells were scraped off, collected in Eppendorf tubes, and centrifuged (21 × g at 4°C) for 10 min. The supernatant was collected in a separate tube, and protein con-centration was measured using Pierce BCA Assay kit (Thermo Fisher Scienti-fic, USA) according to the manufacturer's recommendations.

For western blot experiments, protein from cell lysates was prepared in loading buffer and reducing agent. In each sample, 24 μg protein, 3 μL 10× SDS, and 7.5 μL 4× LDS were filled with deionized H$_2$O to a final volume of 30 μL. Contents of each tube were boiled at 95°C for 10 min. Gradient gels (1.0 mm, 4%–12% [Thermo Fisher Scientific, USA]) were used and run using MOPS buffer (Thermo Fisher Scientific, USA) for 45–60 min at 200 V and 4°C. Protein transfer was done as a "wet" transfer using activated polyvinyli-dene fluoride (PVDF) membrane for 1 h at 400 mA at 4°C.

After protein transfer, membranes were washed in TBS-T twice for 5 min and blocked in 5% BSA in TBS-T for 1 h. Primary antibodies were applied in different dilutions (monoclonal mouse anti-α-tubulin [DM1A] 1:2,000, Cell Signaling, Cat. No. 3873; monoclonal rabbit anti-acetyl-α-tubulin [Lys40] 1:800, Cell Signaling, Cat. No. 5335; polyclonal rabbit anti-phospho-Tau [Ser396] 1:500, Invitrogen, Cat. No. 44-752G; monoclonal mouse anti-Tau antibody [HT7] 1:100, Invitro-gen, Cat. No. MN1000) in blocking buffer overnight on a spinner at 4°C. Tubulin and tau levels, either total protein levels or post-translational modifications, were detected simultaneously. After washing for 5 min, 3 times with TBS-T, the secondary antibody (WS HRP polyclonal goat anti-mouse, Li-Cor, Cat. No. A11029; WS HRP polyclonal goat anti-rabbit, Li-Cor, Cat. No. A31576) was applied at 1:15,000 dilution for 1 h in the dark on a spinner at 4°C. After-ward, the membrane was washed 3 times for 5 min each with PBS and analyzed using chemiluminescent detection reagent (Li-Cor, USA).

After detection, the membranes were stripped to apply a different set of an-tibodies. Therefore, the membrane was washed with PBS, then a western blot stripping buffer (Thermo Fisher Scientific, USA) was applied for 7 min. After washing (3 × 5 min, PBS), the membranes could be blocked again and sub-jected to the next antibody. Rabbit antibodies were generally applied first, fol-lowed by the antibodies raised in mouse.

Validated compound target comparison against disease-specific proteins

Experimentally validated compounds were initially selected and tested based on their downstream effect on the primary target via a specific secondary in-teractor. However, drugs often interact with several different components, thereby requiring one to assess how each compound fits into the disease-spe-cific interactome. To this end, we compiled a list of all known targets for each compound that passed the experimental screening and compared them against those known to be associated with AD. Drug targets were collected from the ChEMBL database, while genes of SNPs that were found to have a significant link (p < 5 × 10^{-7}) with AD were collected from the GWAS Catalog dataset stored within our knowledge graph. The resulting overlapping set of symbols and their found drug interactions were subsequently mapped to the knowledge graph to be used as new starting points for the workflow described above.

AUTHOR CONTRIBUTIONS

Conceptualization, M.H.-A. and C.C.; methodology, M.H.-A., C.C., M.J., V.L.-R., B.S., S.G., B.E., O.P., A.Z., M.N., and A.E.; investigation, M.J., V.L.-R., B.S., S.G., B.E., O.P., A.Z., J.R., J.D., C.E., M.N., and A.E.; formal analysis, V.L.-R., B.S., C.E., M.N., A.E., M.J., A.Z., B.E., J.R., and J.D.; software, B.S., C.E., and M.J.; visualization, V.L.-R., B.S., M.N., and B.E.; data curation, B.S. and S.G.; writing-original draft, M.H.-A., C.C., M.J., V.L.-R., B.S., S.G., B.E., O.P., A.Z., and M.N.; writing-review & editing, M.H.-A., C.C., M.J., V.L.-R., B.S., S.G., B.E., O.P., A.Z., C.E., and M.N.; supervision, M.H.-A., M.J., and V.L.-R.; resources, M.H.-A. and C.C.; project administration, M.H.-A., M.J., and C.C.; funding acquisition, M.H.-A. and C.C.

REFERENCES

1. Castro, D.M., Dillon, C., Machnicki, G., and Allegri, R.F. (2010). The economic cost of Alzheimer's disease: family or public-health burden? Dement. Neuropsychol. *4*, 262–267.

2. Mehta, D., Jackson, R., Paul, G., Shi, J., and Sabbagh, M. (2017). Why do trials for Alzheimer's disease drugs keep failing? A discontinued drug perspective for 2010–2015. Expert Opin. Invest. Drugs *26*, 735–739.

3. Cummings, J. (2018). Lessons learned from Alzheimer disease: clinical trials with negative outcomes. Clin Transl Sci *11*, 147–152.

4. Yiannopoulou, K.G., Anastasiou, A.I., Zachariou, V., and Pelidou, S.H. (2019). Reasons for failed trials of disease-modifying treatments for Alzheimer disease and their contribution in recent research. Biomedicines *7*, 97.

5. Ghofrani, H.A., Osterloh, I.H., and Grimminger, F. (2006). Sildenafil: from angina to erectile dysfunction to pulmonary hypertension and beyond. Nat. Rev. Drug Discov. *5*, 689–702.

6. Fukushiro-Lopes, D., Hegel, A.D., Russo, A., Senyuk, V., Liotta, M., Beeson, G.C., Beeson, C.C., Burdette, J., Potkul, R.K., and Gentile, S. (2020). Repurposing Kir6/SUR2 channel activator minoxidil to arrests growth of gynecologic cancers. Front Pharmacol. *11*, 577.

7. Ashburn, T.T., and Thor, K.B. (2004). Drug repositioning: identifying and developing new uses for existing drugs. Nat. Rev. Drug Discov. *3*, 673–683.

8. Hodos, R.A., Kidd, B.A., Shameer, K., Readhead, B.P., and Dudley, J.T. (2016). In silico methods for drug repurposing and pharmacology. Wiley Interdiscip. Rev. Syst. Biol. Med. *8*, 186–210.

9. Pushpakom, S., Iorio, F., Eyers, P.A., Escott, K.J., Hopper, S., Wells, A., Doig, A., Guilliams, T., Latimer, J., McNamee, C., et al. (2018). Drug repurposing: progress, challenges and recommendations. Nat. Rev. Drug Discov. *18*, 41–58.

10. Zhang, Z., Zhou, L., Xie, N., Nice, E.C., Zhang, T., Cui, Y., and Huang, C. (2020). Overcoming cancer therapeutic bottleneck by drug repurposing. Signal Transduct. Target. Ther. *5*, 1–25.

11. Milani, M., Donalisio, M., Bonotto, R.M., Schneider, E., Arduino, I., Boni, F., Lembo, D., Marcello, A., and Mastrangelo, E. (2021). Combined in silico and in vitro approaches identified the antipsychotic drug lurasidone and the antiviral drug elbasvir as SARS-CoV2 and HCoV-OC43 inhibitors. Antiviral Res. *189*, 105055.

12. Parisi, D., Adasme, M.F., Sveshnikova, A., Bolz, S.N., Moreau, Y., and Schroeder, M. (2020). Drug repositioning or target repositioning: a structural perspective of drug-target-indication relationship for available repurposed drugs. Comput. Struct. Biotechnol. J. *18*, 1043–1055.

13. Shukla, R., Henkel, N.D., Alganem, K., Hamoud, A.R., Reigle, J., Alnafisah, R.S., Eby, H.M., Imami, A.S., Creeden, J.F., Miruzzi, S.A., et al. (2021). Signature-based approaches for informed drug repurposing: targeting CNS disorders. Neuropsychopharmacology *46*, 116–130.

14. Morris, M., Maeda, S., Vossel, K., and Mucke, L. (2011). The many faces of tau. Neuron *70*, 410–426.

15. Wang, Y., and Mandelkow, E. (2016). Tau in physiology and pathology. Nat. Rev. Neurosci. *17*, 5–21.

16. Schöll, M., Maass, A., Mattsson, N., Ashton, N.J., Blennow, K., Zetterberg, H., and Jagust, W. (2019). Biomarkers for tau pathology. Mol. Cell. Neurosci. *97*, 18–33.

17. Simões-Pires, C., Zwick, V., Nurisso, A., Schenker, E., Carrupt, P.A., and Cuendet, M. (2013). HDAC6 as a target for neurodegenerative diseases: what makes it different from the other HDACs? Mol. Neurodegener. *8*, 7.

18. Ding, H., Dolan, P.J., and Johnson, G.V.W. (2008). Histone deacetylase 6 interacts with the microtubule-associated protein tau. J. Neurochem. *106*, 2119–2130.

19. Carlomagno, Y., Chung, D., Eun, C., Yue, M., Castanedes-Casey, M., Madden, B.J., Dunmore, J., Tong, J., DeTure, M., Dickson, D.W., Petrucelli, L., et al. (2017). An acetylation–phosphorylation switch that regulates tau aggregation propensity and function. J. Biol. Chem. *292*, 15277–15286.

20. Seidel, C., Schnekenburger, M., Dicato, M., and Diederich, M. (2015). Histone deacetylase 6 in health and disease. Epigenomics *7*, 103–118.

21. Sakamoto, K.M., and Aldana-Masangkay, G.I. (2011). The role of HDAC6 in cancer. J. Biomed. Biotechnol. *2011*, 875824.

22. d'Ydewalle, C., Bogaert, E., and Van Den Bosch, L. (2012). HDAC6 at the intersection of neuroprotection and neurodegeneration. Traffic *13*, 771–779.

23. Sabnis, R.W. (2021). Novel histone deacetylase 6 inhibitors for treating Alzheimer's disease and cancer. ACS Med. Chem. Lett. *12*, 1202–1203.

24. Gorsky, M.K., Burnouf, S., Dols, J., Mandelkow, E., and Partridge, L. (2016). Acetylation mimic of lysine 280 exacerbates human Tau neurotoxicity in vivo. Sci. Rep. *6*, 22685.

25. Cook, C., Gendron, T.F., Scheffel, K., Carlomagno, Y., Dunmore, J., DeTure, M., and Petrucelli, L. (2012). Loss of HDAC6, a novel CHIP substrate, alleviates abnormal tau accumulation. Hum. Mol. Genet. *21*, 2936–2945.

26. Cohen, T.J., Guo, J.L., Hurtado, D.E., Kwong, L.K., Mills, I.P., Trojanowski, J.Q., and Lee, V.M.Y. (2011). The acetylation of tau inhibits its function and promotes pathological tau aggregation. Nat. Commun. *2*, 252.

27. Zhang, W., Tarutani, A., Newell, K.L., Murzin, A.G., Matsubara, T., Falcon, B., Vidal, R., Garringer, H.J., Shi, Y., Ikeuchi, T., et al. (2020). Novel tau filament fold in corticobasal degeneration. Nature *580*, 283–287.

28. Halley, F., Reinshagen, J., Ellinger, B., Wolf, M., Niles, A.L., Evans, N.J., Kirkland, T.A., Wagner, J.M., Jung, M., Gribbon, P., et al. (2011). A bioluminogenic HDAC activity assay: validation and screening. J. Biomol. Screen. *16*, 1227–1235.

29. Miyake, Y., Keusch, J.J., Wang, L., Saito, M., Hess, D., Wang, X., Melancon, B.J., Helquist, P., Gut, H., and Matthias, P. (2016). Structural insights into HDAC6 tubulin deacetylation and its selective inhibition. Nat. Chem. Biol. *12*, 748–754.

30. Lipinski, C.A. (2016). Rule of five in 2015 and beyond: target and ligand structural limitations, ligand chemistry structure and drug discovery project decisions. Adv Drug Deliv. Rev. *101*, 34–41.

31. Buniello, A., MacArthur, J.A.L., Cerezo, M., Harris, L.W., Hayhurst, J., Malangone, C., et al. (2019). The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. Nucleic Acid Res. *47*, D1005–D1012. https://doi.org/10.1093/nar/gky1120.

32. Levine, B., and Kroemer, G. (2008). Autophagy in the pathogenesis of disease. Cell *132*, 27–42.

33. Uddin, M.S., Stachowiak, A., Al Mamun, A., Tzvetkov, N.T., Takeda, S., Atanasov, A.G., Bergantin, L.B., Abdel-Daim, M.M., and Stankiewicz, A.M. (2018). Autophagy and Alzheimer's disease: from molecular mechanisms to therapeutic implications. Front. Aging Neurosci. *10*, 04.

34. Sevigny, J., Chiao, P., Bussière, T., Weinreb, P.H., Williams, L., Maier, M., Dunstan, R., Salloway, S., Chen, T., Ling, Y., et al. (2016). The antibody aducanumab reduces Aβ plaques in Alzheimer's disease. Nature *537*, 50–56.

35. Mullard, A. (2021). Landmark Alzheimer's drug approval confounds research community. Nature *594*, 309–310.

36. Li, Y., Yao, J., Han, C., Yang, J., Chaudhry, M.T., Wang, S., Liu, H., and Yin, Y. (2016). Quercetin, inflammation and immunity. Nutrients *8*, 167.

37. Wang, J., Hancock, M.K., Dudek, J.M., and Bi, K. (2008). Cellular assays for high-throughput screening for modulators of Trk receptor tyrosine kinases. Curr. Chem. Genomics *1*, 27–33.

38. Hirose, M., Kuroda, Y., and Murata, E. (2016). NGF/TrkA signaling as a therapeutic target for pain. Pain Pract. *16*, 175–182.

39. Canu, N., Amadoro, G., Triaca, V., Latina, V., Sposato, V., Corsetti, V., Severini, C., Ciotti, M.T., and Calissano, P. (2017). The intersection of NGF/TrkA signaling and amyloid precursor protein processing in Alzheimer's disease neuropathology. Int. J. Mol. Sci. *18*, 1319.

40. Turner, C.A., Eren-Koçak, E., Inui, E.G., Watson, S.J., and Akil, H. (2016). Dysregulated fibroblast growth factor (FGF) signaling in neurological and psychiatric disorders. Semin. Cell Dev. Biol. *53*, 136–143.

41. Li, X., Wang, C., Xiao, J., McKeehan, W.L., and Wang, F. (2016). Fibroblast growth factors, old kids on the new block. Semin. Cell Dev. Biol. *53*, 155–167.

42. Whittaker, S.R., Mallinger, A., Workman, P., and Clarke, P.A. (2017). Inhibitors of cyclin-dependent kinases as cancer therapeutics. Pharmacol. Ther. *173*, 83–105.

43. Zhang, S.G., Wang, X.S., Zhang, Y.D., Di, Q., Shi, J.P., Qian, M., Xu, L.G., Lin, X.J., and Lu, J. (2016). Indirubin-3′-monoxime suppresses amyloid-betainduced apoptosis by inhibiting tau hyperphosphorylation. Neural Regen. Res. *11*, 988–993.

44. Leclerc, S., Garnier, M., Hoessel, R., Marko, D., Bibb, J.A., Snyder, G.L., Greengard, P., Biernat, J., Wu, Y.Z., Mandelkow, E.M., et al. (2001). Indirubins inhibit glycogen synthase kinase-3β and CDK5/P25, two protein kinases involved in abnormal tau phosphorylation in Alzheimer's disease. A property common to most cyclin-dependent kinase inhibitors? J. Biol. Chem. *276*, 251–260.

45. Kil, K.E., Zhang, Z., Jokivarsi, K., Gong, C., Choi, J.K., Kura, S., and Brownell, A.L. (2013). Radiosynthesis of N-(4-chloro-3-[11C]methoxy-phenyl)-2-picolinamide ([11C]ML128) as a PET radiotracer for metabotropic glutamate receptor subtype 4 (mGlu4). Bioorg. Med. Chem. *21*, 5955–5962.

46. Martin, L., Latypova, X., and Terro, F. (2011). Post-translational modifications of tau protein: implications for Alzheimer's disease. Neurochem. Int. *58*, 458–471.

47. Tracy, T., Claiborn, K.C., and Gan, L. (2019). Regulation of tau homeostasis and toxicity by acetylation. In Advances in Experimental Medicine and Biology (Springer), pp. 47–55.

48. Kontaxi, C., Piccardo, P., and Gill, A.C. (2017). Lysine-directed post-translational modifications of tau protein in Alzheimer's disease and related tauopathies. Front. Mol. Biosci. *4*, 56.

49. Min, S.W., Sohn, P.D., Li, Y., Devidze, N., Johnson, J.R., Krogan, N.J., Masliah, E., Mok, S.A., Gestwicki, J.E., and Gan, L. (2018). SIRT1 deacetylates tau and reduces pathogenic tau spread in a mouse model of tauopathy. J Neurosci *38*, 3680–3688.

50. Trzeciakiewicz, H., Ajit, D., Tseng, J.H., Chen, Y., Ajit, A., Tabassum, Z., Lobrovich, R., Peterson, C., Riddick, N.V., Itano, M.S., et al. (2020). An HDAC6-dependent surveillance mechanism suppresses tau-mediated neurodegeneration and cognitive decline. Nat. Commun. *11*, 1–18.

51. Xia, Y., Bell, B.M., and Giasson, B.I. (2021). Tau K321/K353 pseudoacetylation within KXGS motifs regulates tau–microtubule interactions and inhibits aggregation. Sci. Rep. *11*, 1–9.

52. Choi, H., Kim, H.J., Yang, J., Chae, S., Lee, W., Chung, S., Kim, J., Choi, H., Song, H., Lee, C.K., et al. (2020). Acetylation changes tau interactome to degrade tau in Alzheimer's disease animal and organoid models. Aging Cell *19*, e13081.

53. de Garcini, E.M., Corrochano, L., Wischik, C.M., Nido, J.D., Correas, I., and Avila, J. (1992). Differentiation of neuroblastoma cells correlates with an altered splicing pattern of tau RNA. FEBS Lett. *299*, 10–14.

54. Dubey, S.K., Ram, M.S., Krishna, K.V., Saha, R.N., Singhvi, G., Agrawal, M., Ajazuddin Saraf, S., Saraf, S., and Alexander, A. (2019). Recent expansions on cellular models to uncover the scientific barriers towards drug development for Alzheimer's disease. Cell. Mol. Neurobiol. *39*, 181–209.

55. Mungenast, A.E., Siegert, S., and Tsai, L.H. (2016). Modeling Alzheimer's disease with human induced pluripotent stem (iPS) cells. Mol. Cell. Neurosci. *73*, 13–31.

56. Yan, G., Wang, X., Chen, Z., Wu, X., Pan, J., Huang, Y., Wan, G., and Yang, Z. (2017). In-silico ADME studies for new drug discovery: from chemical compounds to Chinese herbal medicines. Curr. Drug Metab. *18*, 535–539.

57. Park, K. (2019). A review of computational drug repurposing. Transl. Clin. Pharmacol. *27*, 59–63.

58. Li, H., Wang, X., Yu, H., Zhu, J., Jin, H., Wang, A., and Yang, Z. (2017). Combining in vitro and in silico approaches to find new candidate drugs targeting the pathological proteins related to the Alzheimer's disease. Curr. Neuropharmacol. *16*, 758–768.

59. Galindez, G., Matschinske, J., Rose, T.D., Sadegh, S., Salgado-Albarrán, M., Späth, J., Baumbach, J., and Pauling, J.K. (2021). Lessons from the COVID-19 pandemic for advancing computational drug repurposing strategies. Nat. Comput. Sci. *1*, 33–41.

60. Schultz, B., Zaliani, A., Ebeling, C., Reinshagen, J., Bojkova, D., Lage-Rupprecht, V., Karki, R., Lukassen, S., Gadiya, Y., Ravindra, N.G., et al. (2021). A method for the rational selection of drug repurposing candidates from multimodal knowledge harmonization. Sci. Rep. *11*, 1–10.

61. Corsello, S.M., Bittker, J.A., Liu, Z., Gould, J., McCarren, P., Hirschman, J.E., Johnston, S.E., Vrcic, A., Wong, B., Khan, M., et al. (2017). The drug repurposing hub: a next-generation drug library and information resource. Nat. Med. *23*, 405–408.

62. Kodamullil, A.T., Younesi, E., Naz, M., Bagewadi, S., and Hofmann-Apitius, M. (2015). Computable cause-and-effect models of healthy and Alzheimer's disease states and their mechanistic differential analysis. Alzheimer's Dement. *11*, 1329–1339.

63. Domingo-Fernández, D., Kodamullil, A.T., Iyappan, A., Naz, M., Emon, M.A., Raschka, T., Karki, R., Springstubbe, S., Ebeling, C., and Hofmann-Apitius, M. (2017). Multimodal mechanistic signatures for neurodegenerative diseases (NeuroMMSig): a web server for mechanism enrichment. Bioinformatics *33*, 3679–3681.

64. Hoyt, C.T., Konotopez, A., and Ebeling, C. (2018). PyBEL: a computational framework for biological expression language. Bioinformatics *34*, 703–704.

65. Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2016). KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res. *44*, D457–D462.

66. Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., and Morishima, K. (2017). KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res. *45*, D353–D361.

67. Cerami, E.G., Gross, B.E., Demir, E., Rodchenkov, I., Babur, Ö., Anwar, N., Schultz, N., Bader, G.D., and Sander, C. (2011). Pathway Commons, a web resource for biological pathway data. Nucleic Acids Res. *39*, D685–D690.

68. Rodchenkov, I., Babur, O., Luna, A., Aksoy, B.A., Wong, J.V., Fong, D., Franz, M., Siper, M.C., Cheung, M., Wrana, M., et al. (2020). Pathway Commons 2019 Update: integration, analysis and exploration of pathway data. Nucleic Acids Res. *48*, D489–D497.

69. Jassal, B., Matthews, L., Viteri, G., Gong, C., Lorente, P., Fabregat, A., Sidiropoulos, K., Cook, J., Gillespie, M., Haw, R., et al. (2020). The reactome pathway knowledgebase. Nucleic Acids Res *48*, D498–D503.

70. Domingo-Fernández, D., Hoyt, C.T., Bobis-Álvarez, C., Marín-Lláo, J., and Hofmann-Apitius, M. (2019). ComPath: an ecosystem for exploring, analyzing, and curating mappings across pathway databases. NPJ Syst. Biol. Appl. *5*, 1.

71. Hermjakob, H., Montecchi-Palazzi, L., Lewington, C., Mudali, S., Kerrien, S., Orchard, S., Vingron, M., Roechert, B., Roepstorff, P., Valencia, A., et al. (2004). IntAct: an open source molecular interaction database. Nucleic Acids Res. *32*, D452–D455.

72. Oughtred, R., Stark, C., Breitkreutz, B.J., Rust, J., Boucher, L., Chang, C., Kolas, N., O'Donnell, L., Leung, G., McAdam, R., et al. (2019). The BioGRID interaction database: 2019 update. Nucleic Acids Res. *47*, D529–D541.

73. Szklarczyk, D., Gable, A.L., Nastou, K.C., Lyon, D., Kirsch, R., Pyysalo, S., Doncheva, N.T., Legeay, M., Fang, T., Bork, P., et al. (2021). The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. Nucleic Acids Res. *49*, D605–D612.

74. Wishart, D.S., Feunang, Y.D., Guo, A.C., Lo, E.J., Marcu, A., Grant, J.R., Sajed, T., Johnson, D., Li, C., Sayeeda, Z., et al. (2018). DrugBank 5.0: a

major update to the DrugBank database for 2018. Nucleic Acids Res. *46*, D1074–D1082.

75. Zarin, D.A., Tse, T., Williams, R.J., Califf, R.M., and Ide, N.C. (2011). The ClinicalTrials.gov results database — update and key issues. N Engl J Med *364*, 852–860.

76. Namysl, M., Esser, A.M., Behnke, S., and Köhler, J. (2021). Flexible table recognition and semantic interpretation system. arXiv https://arxiv.org/abs/2105.11879.

77. Gobel, M., Hassan, T., Oro, E., and Orsi, G. (2013). ICDAR 2013 table competition. In Proceedings of the International Conference on Document Analysis and Recognition (IEEE), pp. 1449–1453.

78. Gao, L., Huang, Y., Dejean, H., Meunier, J.L., Yan, Q., Fang, Y., Kleber, F., and Lang, E. (2019). ICDAR 2019 competition on table detection and recognition (cTDaR). In Proceedings of the International Conference on Document Analysis and Recognition, ICDAR (IEEE Computer Society), pp. 1510–1515.

79. Adams, T., Namysl, M., Kodamullil, A.T., Behnke, S., and Jacobs, M. (2021). Benchmarking Table Recognition Performance on Biomedical Literature on Neurological Disorders. Bioinformatics. https://doi.org/10.1093/bioinformatics/btab843.

80. Adams, T., Namysl, M., Kodamullil, A.T., Behnke, S., and Jacobs, M. (2021). Benchmarking table recognition performance on biomedical literature on neurological disorders. Bioinformaticsbtab943. https://doi.org/10.1093/bioinformatics/btab843.

81. Kullback, S., and Leibler, R.A. (1951). On information and sufficiency. Ann. Math. Stat. *22*, 79–86.

82. Kim, S., Thiessen, P.A., Bolton, E.E., Chen, J., Fu, G., Gindulyte, A., Han, L., He, J., He, S., Shoemaker, B.A., et al. (2016). PubChem substance and compound databases. Nucleic Acids Res. *44*, D1202–D1213.