



## ARTICLE

# Integrating yeast chemical genomics and mammalian cell pathway analysis

Fu-lai Zhou<sup>1,2</sup>, Sheena C Li<sup>3</sup>, Yue Zhu<sup>1,2</sup>, Wan-jing Guo<sup>1,2</sup>, Li-jun Shao<sup>1,2</sup>, Justin Nelson<sup>4</sup>, Scott Simpkins<sup>4</sup>, De-hua Yang<sup>1</sup>, Qing Liu<sup>1</sup>, Yoko Yashiroda<sup>3</sup>, Jin-biao Xu<sup>5</sup>, Yao-yue Fan<sup>5</sup>, Jian-min Yue<sup>5</sup>, Minoru Yoshida<sup>3,6,7</sup>, Tian Xia<sup>8</sup>, Chad L Myers<sup>4</sup>, Charles Boone<sup>3,9</sup> and Ming-wei Wang<sup>1,2</sup>

Chemical genomics has been applied extensively to evaluate small molecules that modulate biological processes in *Saccharomyces cerevisiae*. Here, we use yeast as a surrogate system for studying compounds that are active against metazoan targets. Large-scale chemical-genetic profiling of thousands of synthetic and natural compounds from the Chinese National Compound Library identified those with high-confidence bioprocess target predictions. To discover compounds that have the potential to function like therapeutic agents with known targets, we also analyzed a reference library of approved drugs. Previously uncharacterized compounds with chemical-genetic profiles resembling existing drugs that modulate autophagy and Wnt/ $\beta$ -catenin signal transduction were further examined in mammalian cells, and new modulators with specific modes of action were validated. This analysis exploits yeast as a general platform for predicting compound bioactivity in mammalian cells.

**Key words:** yeast; chemical genomics; tubulin cytoskeleton assembly; autophagy; Wnt/ $\beta$ -catenin signaling pathway

*Acta Pharmacologica Sinica* (2019) 40:1245–1255; <https://doi.org/10.1038/s41401-019-0231-y>

## INTRODUCTION

The budding yeast, *Saccharomyces cerevisiae*, is a powerful model organism for biological studies due to its genetic tractability and ease of propagation [1]. Although human and yeast are distant evolutionarily, essential genes and pathways are highly conserved between these organisms. Indeed, when a total of 414 essential yeast genes were replaced by their human orthologs, nearly half (47%) could be successfully humanized [2]. Because of homologous cell cycle, cell fate, and aging pathways, *S. cerevisiae* has been directly used to discover therapeutic leads against a wide variety of diseases, in both target-based and phenotypic-based screens [3–6].

Yeast is an ideal model for global genetic analysis. Numerous studies, including the analysis of haploid deletion mutants of non-essential genes, haploinsufficiency profiling of essential genes [7, 8], and the analysis of essential gene conditional alleles [9, 10], have been carried out to explore single gene function. More complex genetics can be examined through synthetic genetic array (SGA) analysis, an automated technique for studying genetic interactions on a large scale. A genetic interaction occurs when the combination of different gene perturbations leads to an unexpected phenotype, such as synthetic lethality, where two viable mutations combine to generate a lethal or severely sick

double mutant phenotype [11, 12]. Genome-wide SGA analysis generated a global genetic interaction map of budding yeast based on genetic interaction profile similarity [13, 14]. The set of genetic interactions for any given query gene comprise its genetic interaction profile. Genes that belong to the same biological pathway tend to share similar patterns of genetic interactions, thus genetic interaction profiles provide a quantitative description of every gene's biological function. This allows annotation of many genes to general bioprocesses, such as DNA replication, RNA splicing, protein degradation, and nuclear-cytoplasmic import, at an intermediate profile similarity threshold, or to specific protein complexes and signaling pathways such as the proteasome, the vacuolar H<sup>+</sup> ATPase, and the protein kinase C pathway, at higher profile similarity thresholds. Thus, the global genetic interaction map can be viewed as a comprehensive map of cellular function [14].

Chemical genomics combines genetic resources, such as large-scale mutant collections and chemical libraries to investigate the potential association between compounds and cellular targets [15–17]. Recently, we developed an unbiased, high-throughput, chemical genomic screening platform to functionally characterize bioactive compounds in a highly parallel assay [18]. To minimize the amount of compound required for screening, our pipeline

<sup>1</sup>The National Center for Drug Screening and the CAS Key Laboratory of Receptor Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences (CAS), Shanghai 201203, China; <sup>2</sup>University of Chinese Academy of Sciences, Beijing 100049, China; <sup>3</sup>RIKEN Center for Sustainable Resource Science, Wako, Saitama 3510198, Japan; <sup>4</sup>Bioinformatics and Computational Biology Program, University of Minnesota-Twin Cities, Minneapolis, Minnesota 55455, USA; <sup>5</sup>The State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China; <sup>6</sup>Department of Biology, The University of Tokyo, Bunkyo-ku, Tokyo 1138657, Japan; <sup>7</sup>Collaborative Research for Innovative Microbiology, The University of Tokyo, Bunkyo-ku, Tokyo 1138657, Japan; <sup>8</sup>Department of Electronics and Information Engineering, Huazhong University of Science and Technology, Wuhan 430074, China and <sup>9</sup>Donnelly Centre and Department of Molecular Genetics, University of Toronto, Ontario M5S 3E1, Canada

Correspondence: Chad L Myers (chadm@umn.edu) or Charles Boone (charlie.boone@utoronto.ca) or Ming-wei Wang (mwwang@simmm.ac.cn)  
These authors contributed equally: Fu-lai Zhou, Sheena C. Li

Received: 8 January 2019 Accepted: 14 March 2019

Published online: 28 May 2019

employs a diagnostic set of DNA-barcoded gene deletion mutant yeast strains, each constructed in a drug-hypersensitive genetic background. These mutants are pooled and grown competitively in the presence of bioactive compounds. Strain-specific DNA barcodes enable monitoring of strain abundance within a pool and identification of chemical-genetic interactions. A second set of barcodes unique to each compound treatment is introduced by polymerase chain reaction (PCR), allowing highly multiplexed experiments resolved using next-generation sequencing (NGS) [18].

The present study explores the utility of yeast chemical-genetic profiles for discovering small molecules with defined targets in mammalian systems. From a primary analysis of 333,200 compounds from the Chinese National Compound Library (CNCL) that contain small molecules from both synthetic and natural sources, we identified 5360 compounds that inhibited yeast cell growth and subjected this subset to a high-throughput screening (HTS) campaign using our diagnostic chemical genomic platform [18]. Ultimately, 1109 compounds with high-confidence target predictions, associated with 17 major bioprocesses, were identified, and their general cellular functions [18] annotated. To integrate these bioprocess level target predictions with mammalian cell pathways, the chemical-genetic profiles of these compounds were compared to those of a reference library consisting of 1898 approved drugs with known modes of action. Two bioprocesses, namely autophagy, which is highly conserved from yeast to humans [19] and the Wnt/ $\beta$ -catenin signaling pathway, which is found only in higher eukaryotes, were chosen to verify the predicted effects of CNCL compounds on mammalian cells. Our results suggest that yeast-based chemical genomic analysis provides a powerful and general approach for predicting cellular targets in mammalian cells.

## MATERIALS AND METHODS

### Reagents

Benomyl, *D*-sorbitol, furazolidone, galactose, glucose, methyl methanesulfonate (MMS),  $\beta$ -mercaptoethanol, and yeast extract were purchased from Sigma (St. Louis, MO, USA). Peptone and RNase A were bought from Sangon Biotech (Shanghai, China). QIAcube HT kits were procured from QIAGEN (Duesseldorf, Germany), GeneClean III kits and Zymolyase-100T from MP Biomedicals (Irvine, CA, USA), Universal Hot Start High-Fidelity PCR Master Mix from Biotool (Shanghai, China), Kapa Library Quantification kits from KAPA Biosystems (Wilmington, MA, USA), and Bortezomib from MedChem Express (Monmouth Junction, NJ, USA).

### Antibodies

The following antibodies were used: anti-LC3B (Sigma), anti- $\beta$ -tubulin (Proteintech, Rosemont, IL, USA), anti- $\beta$ -actin, anti-GAPDH, anti- $\beta$ -catenin and anti-rabbit IgG (Cell Signaling Technology, Danvers, MA, USA), as well as anti-VDAC1/porin (Abcam, Cambridge, MA, USA).

### Compound library

The 335 0988 synthetic and natural product-derived compounds used in this study are stored at the Chinese National Compound Library (CNCL), including 1898 commercially available medicines approved by the Food and Drug Administration (FDA, USA). Their structure diversity covers lactams, heterocycles, amides, secondary amides, sulfonates, sulphonamides, etc. The compounds were of high purity (>95%), checked randomly for quality control, and solubilized in 100% DMSO prior to application.

### High-throughput bioactivity screening

The *pdrl1 $\Delta$  pdr3 $\Delta$  snq2 $\Delta$*  drug-hypersensitive strain [18] (Y13206; Supplementary Table 1) and the pool of 310 non-essential gene

deletion strains in the same drug-hypersensitive genetic background [18] (diagnostic set) were cultured in YPGal medium (1% yeast extract, 2% peptone, 2% galactose) at 30 °C. All the compounds were dissolved in DMSO from a stock concentration of 2 mM. Positive controls included benomyl, bortezomib, and furazolidone (dissolved in DMSO from a stock solution of 10 mg/mL), as well as methyl methanesulfonate (MMS; dissolved in DMSO from a stock solution of 50 mg/mL).

For bioactivity screening, hypersensitive yeast cells were seeded onto 96-well plates at a density of  $9.3 \times 10^4$  cells per well by Multidrop Combi (ThermoFisher Scientific, Rockford, IL, USA) and 1  $\mu$ L compound was added to each well using Freedom EVO platform (TECAN, Männedorf, Switzerland; 80 compounds were screened in each plate). After 24-h incubation at 30 °C,  $OD_{600}$  values were measured by an EnSpire Multimode Plate Reader (PerkinElmer, Boston, MA, USA) to evaluate cell viability. Compounds showing 45%–75% growth inhibition compared to DMSO were selected for chemical-genetic screening using the diagnostic set [18].

### High-throughput chemical-genetic screening

Yeast culture conditions were the same as described above. A total of 5520 initial bioactive hits were screened for chemical-genetic profiles against the diagnostic set of pooled yeast strains distributed in 96-well plates (each well containing  $4.65 \times 10^5$ /mL cells, 196  $\mu$ L YPGal medium, and 2  $\mu$ L compound).  $OD_{600}$  was measured after 24 h followed by additional 24-h incubation at 30 °C. The plates were then centrifuged at 1300 $\times g$  for 4 min before adding 125  $\mu$ L zymolyase dissolved in 1 M *D*-sorbitol (final concentration: 0.5 mg/mL) with 11.5  $\mu$ M  $\beta$ -mercaptoethanol. The samples were incubated for 1 h at 37 °C, centrifuged again at 2272 $\times g$  for 5 min and processed using the QIAamp 96 DNA kit. Genomic DNA extraction was performed with an automated high-throughput nucleic acid purification robot, QIAcube HT (QIAGEN).

Strain-specific DNA barcodes were amplified using multiplex primers and a communal U2 primer [18]. PCR conditions were set as follows: 3 min at 95 °C for initial denaturation, 30 cycles of 15 s at 95 °C, 15 s at 60 °C, 20 s at 72 °C, and a final extension time of 5 min at 72 °C. PCR products were purified from 2% agarose gels by the GeneClean III kit, quantified using Kapa qPCR kit and sequenced with Hiseq 2500 at WuXi AppTec (Shanghai, China).

### Data processing and target prediction

The relative abundances of barcoded mutants after compound treatment were quantified using amplicon sequencing. Chemical-genetic interaction z-scores for enrichment or depletion in the presence of the compound relative to the DMSO control were generated from sequencing data using the BEAN-counter software pipeline (<https://www.github.com/csbio/BEAN-counter>) [18, 20]. The screens were performed in four batches: each batch was processed independently using BEAN-counter. Prior to combining into the final dataset, the first singular value decomposition (SVD) component was removed from each of the four individual datasets to remove batch effects that were not reflective of true signal.

Genetic interaction information was obtained from the global *S. cerevisiae* genetic interaction dataset [14], which was derived from quantitative growth observations of non-essential, double mutant strains described previously [13, 14]. Perturbed biological processes (a subset of 1309 terms from the GO biological process ontology) [21] were predicted from chemical-genetic interaction profiles using the CG-TARGET software pipeline [22]. CG-TARGET leverages the similarity between chemical-genetic interaction profiles and the genetic interaction profiles of their target genes or proteins to predict the biological processes that are perturbed by compounds. It also uses experimental control and randomly re-sampled chemical-genetic interaction profiles to inform its statistical analyses and estimate the false discovery rate (FDR) of

each prediction. Predictions with a  $FDR \leq 25\%$  were considered to be of high-confidence, with the associated compounds considered high-confidence compounds. Biological process terms were mapped to the set of 17 functional neighborhoods identified by spatial analysis of functional enrichment (SAFE) of the global yeast genetic interaction network [14].

#### Chemical-genetic profile correlation analysis

To identify potential hits with pharmacological significance, we calculated the similarity between the chemical-genetic profiles of approved drugs (with known modes of action) and bioactive compounds (with high-confidence process predictions). Chemical-genetic profiles of compounds comprising the 1109 high-confidence set (HCS) were compared to that of 77 derived from the approved drug library. Similarity between any two profiles was calculated using Pearson correlation coefficient and ranked from highest similarity (1.0) to lowest similarity ( $-1.0$ ).

To assess the shared targets of FH535, a known inhibitor of the Wnt/ $\beta$  catenin signaling pathway [23], and 3305-2 G4 (also known as drupacine, an alkaloid isolated from *Cephalotaxus fortunei*), a pseudo-compound chemical-genetic profile was generated from the profiles of the two compounds. This new profile was generated by element-wise multiplication of the mean of the two profiles with a weight vector derived from the agreement of interactions across the two profiles. Specifically, the weight for each chemical-genetic interaction was calculated as the product of the corresponding interactions in the two compounds' profiles, with negative resulting values (i.e., sign disagreements) set to zero. This weight emphasizes chemical-genetic interactions that agree between the two compound profiles to produce a quantitative pseudo-compound profile that is composed of the consensus interactions. Direct targets were predicted by utilizing this profile to calculate a genetic interaction normalized cosine score between the pseudo-compound chemical-genetic interaction profile and potential genetic interaction profile targets. This methodology utilizes the same approach used for finding molecular targets with chemical-genetic interaction profiling [18, 22]. To assess the significance of the gene target predictions for the FH535/3305-2 G4 pseudo-compound profile, two sets of pseudo-compound chemical-genetic interaction profiles were generated and a genetic interaction normalized cosine score was calculated using the procedure described above: (1) FH535 with each of the CNCL compounds that had a high-confidence bioprocess target prediction, and (2) 3305-2 G4 with the same set of compounds. These cosine scores were used to assess the significance of the shared targets between FH535 and 3305-2 G4. We calculated a z-score for the FH535/3305-2 G4 pseudo-compound chemical-genetic interaction profile compared to a background distribution consisting of the two sets above. The minimum of the two z-scores generated from the two sets for each potential target was taken as a conservative estimate of the significance of the gene target prediction.

#### Cell culture

HEK293 cells were purchased from American Type Culture Collection (ATCC, Manassas, VA, USA). HeLa and HepG2 cells were generous gifts from Dr. Jian Ding (Shanghai Institute of Materia Medica, Chinese Academy of Sciences). HEK293 and HepG2 cells were cultured in Dulbecco's modified Eagle's medium (DMEM, Life Technologies, Carlsbad, CA, USA) supplemented with 10% fetal bovine serum (FBS; Gibco, Grand Island, NY, USA) and 1% penicillin/streptomycin (Invitrogen, Carlsbad, CA, USA). HeLa cells were cultured in RPMI-1640 medium (Life Technologies) with 10% FBS and 1% penicillin/streptomycin. They were incubated in a humidified chamber with 5%  $CO_2$  at 37 °C. To generate a cell line stably expressing GFP-LC3, the plasmid (Addgene, Cambridge, MA, USA) was transfected into HEK293 cells with Lipofectamine 2000 reagent (Invitrogen) according to the manufacturer's

introductions. Transfectants were selected by puromycin (Life Technologies) and GFP-positive cell lines were assessed by fluorescence microscopy.

#### Fluorescence microscopy

HEK293 cells stably expressing GFP-LC3 were seeded onto glass cover slips pretreated with poly-D-lysine (Sigma). After culturing for 24 h to reach 50%–70% confluence, the cells were treated with compounds in fresh medium for 24 h to enable efficient autophagosome (GFP-LC3 puncta) formation. This was followed by three-time wash of the slips with PBS, fixation of the cells with 4% paraformaldehyde for 15 min at room temperature and another three washes in PBS. Nuclei were counterstained with Hoechst 33258 (Sigma) for 5 min (three washes in PBS thereafter) and fluorescent images obtained using a TCS-SP8 STED confocal microscope (Leica, Wetzlar, Germany).

#### Western blotting

Cells were harvested after treatment and lysed in RIPA (Sigma) buffer supplemented with Protease Inhibitor Cocktail (Sigma). The whole cell lysate of each sample was separated by electrophoresis on Tricine-SDS-PAGE gels and transferred to a PVDF membrane (Millipore, MA, USA). The membrane was blocked by 5% non-fat milk in Tris buffered saline/Tween-20 (TBST, EpiZyme, Shanghai, China) and incubated with primary antibodies at 4 °C overnight. The secondary antibodies were used at a concentration of 1:10 000 and incubated at room temperature for 1 h. The protein bands were visualized using the SuperSignal West Dura Extended Duration Substrate (ThermoFisher Scientific). Quantification of the Western blot bands was achieved using the software Image J (National Institutes of Health, Bethesda, MA, USA) and the quantitative data presented were obtained from at least three independent experiments.

#### Transmission electron microscopy

HEK293 cells were scraped from dishes gently and immediately fixed with 2.5% glutaraldehyde in 0.1 M sodium cacodylate buffer, pH 7.4. After five washes with PBS for 15 min, the samples were post-fixed in 1% osmium tetroxide dehydrated in ethanol and acetone, and embedded in epon. They were then sliced and examined with an FEI Tecnai 12 transmission electron microscope (Hillsboro, OR, USA) operated at 120 kV.

#### Cell cycle

HeLa cells were seeded at a density of  $1 \times 10^5$  per well onto 24-well plates and treated with compounds for 18 h. They were digested by trypsin and collected by centrifugation. The pellets were washed with PBS and fixed with 70% (v/v) ethanol at 4 °C overnight. The cells were centrifuged at  $188 \times g$  for 5 min to discard the supernatant. The pellets were washed twice with PBS and resuspended in 100  $\mu$ L propidium iodide (PI) working buffer containing 0.2% Triton X-100, 100 mg/mL RNase A and 50  $\mu$ g/mL PI in PBS for 30 min at 4 °C in the dark. After incubation, fluorescence intensity was measured with a Novoocyte flow cytometer with the Novoexpress software (ACEA Biosciences, San Diego, CA, USA).

#### Reporter assay

The T-cell factor/lymphoid enhancer factor 1 (TCF/LEF1) response motif sequences were cloned into pGM-Lu plasmid ahead of the minimal TA promoter and luciferase reporter gene (Yeasen, Shanghai, China). Plasmids were transfected into HepG2 cells using Lipofectamine 2000 reagent (Invitrogen) following manufacturer's instructions. The cells were seeded onto 96-well plates 24 h after transfection and waited for 1 h to allow their adhesion to the plate bottom. After adding compounds (20  $\mu$ M), they were left at 37 °C for 24 h followed by incubation with Steady-Glo reagent (Promega, Irvine, CA, USA) for 10 min at room



temperature. Luminescence signals were measured by an EnVision plate reader (PerkinElmer).

Statistical analysis for biological experiments

All biological data were analyzed by GraphPad Prism 7 (San Diego, CA, USA) and the results were presented as means ± SEM. Generally, all biological experiments were carried out with  $n \geq 3$  replicates and the significance was determined using a two-tailed Student's *t*-test.  $P < 0.01$  was considered statistically significant.

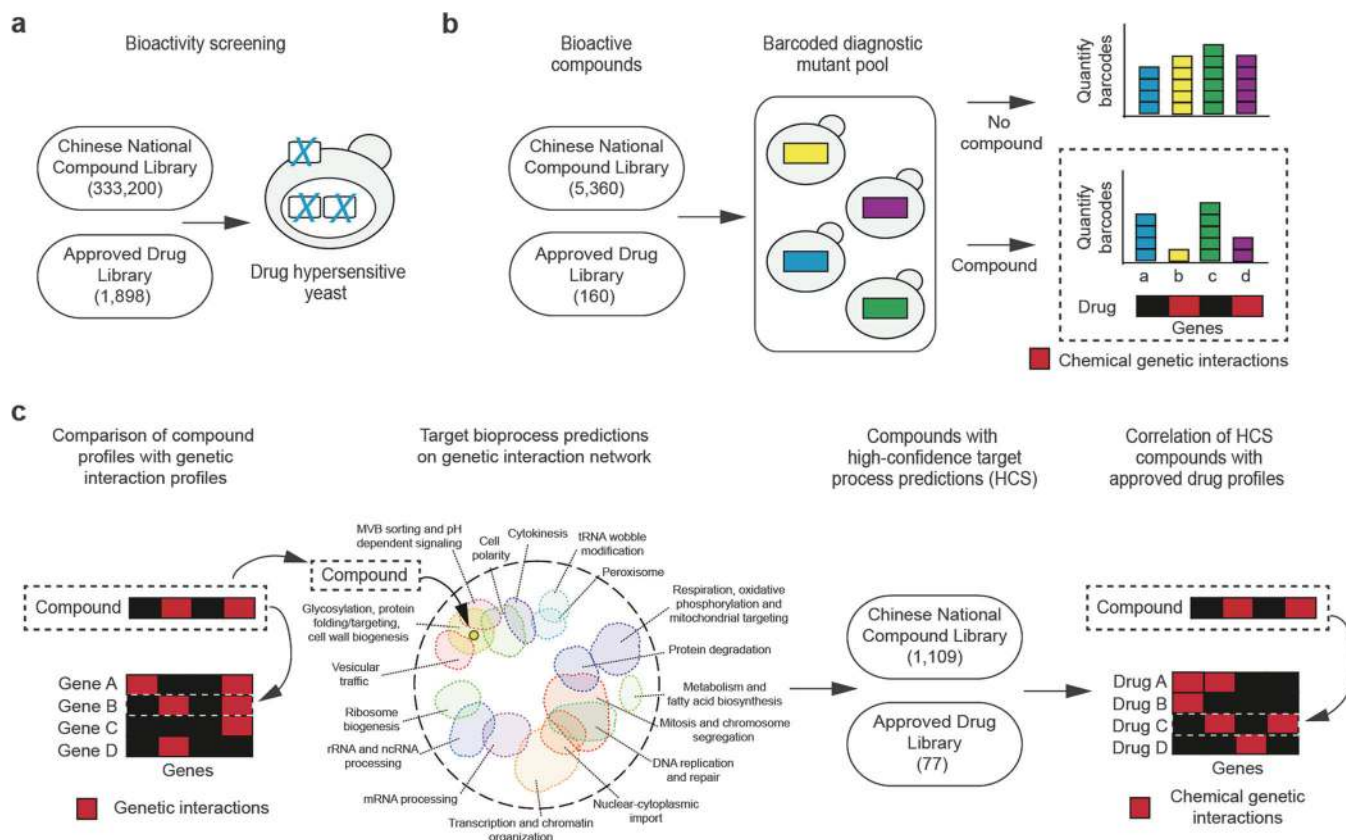
RESULTS

Functional annotation of bioactive compounds by correlation analysis of chemical-genetic profiles with the global yeast genetic interaction network

Of 333 200 synthetic and natural CNCL compounds screened for bioactivity against yeast, 5360 (1.6%) were selected for further chemical-genetic screening based on their ability to inhibit cell growth by more than 20% at a concentration of 20 μM (Fig. 1a). Bioactive CNCL compounds were screened against a diagnostic pool of 310 non-essential deletion mutants that were representative of all major biological processes in *S. cerevisiae* [18]. A compound-gene pair shows a chemical-genetic interaction if

compound treatment causes a relative depletion or accumulation of the corresponding barcoded mutant in pooled growth (Fig. 1b). The set of chemical-genetic interactions for a given compound is referred to as a chemical-genetic interaction profile [24]. Importantly, negative genetic interactions, in which two loss-of-function mutations combine to lead to a more extreme fitness defect than expected, are analogous to negative chemical-genetic interactions because a loss-of-function mutation in a specific gene models the effect of a compound which inhibits the gene's product (Fig. 1c). Hence, to annotate the major cellular function of a bioactive compound, we compare the chemical-genetic interaction profile of that compound to the compendium of genetic interaction profiles for all yeast genes [14]. Based on Gene Ontology (GO) annotations [21] of the genes that share highly similar profiles with a compound, high-confidence target bioprocess predictions can be inferred (Fig. 1c).

Among the screened bioactive compounds, 1109 (21%) were identified to have high-confidence bioprocess target predictions (Fig. 1c). A similar chemical-genetic analysis was carried out using an approved drug library (Selleck Chemicals, Houston, USA), in which all compounds were previously linked to characterized targets in human cells. The approved drug library is relatively small but had 77 (48%; out of 160 initial hits) high-confidence



**Fig. 1** Yeast chemical-genetic screening pipeline. **a** Bioactivity screening in a drug-hypersensitive yeast strain (*pdr1pdr3snq2*, Y13206) was conducted for the Chinese National Compound Library (333 200 compounds), and an approved drug library (1898 compounds). **b** Bioactive compounds that inhibited growth from 45% to 75% compared to the vehicle-only control were screened against a diagnostic set of 310 DNA-barcoded yeast gene deletion mutants. Mutants were grown in a pooled format, treated with each compound, and harvested. Genomic DNA was extracted from each pool and experiment-specific tags and mutant-specific barcodes were amplified by PCR. The relative abundance of each gene mutant after compound treatment was quantified by NGS of the PCR amplicons. From this, chemical-genetic profiles for each compound were constructed, with a compound and gene sharing a chemical-genetic interaction if the relative abundance of the gene mutant was significantly reduced after compound treatment. **c** The chemical-genetic profile of a compound (dashed-line box) was compared to the compendium of synthetic lethal genetic interaction profiles in yeast. A compound was mapped onto the genetic interaction network at the position of the gene with the most similar genetic interaction profile from the compound's top target bioprocess prediction (yellow circle). In this example, the compound's top target process prediction is glycosylation. The chemical-genetic interaction profiles of 1109 CNCL compounds with high-confidence target bioprocess predictions were compared to the profiles of 77 approved drugs with high-confidence target predictions. Unknown compounds that shared high profile similarity with approved drugs were identified for validation

bioprocess target predictions, a more than two-fold higher prediction rate than that of the bioactive CNCL compounds.

Compounds with high-confidence bioprocess target predictions from both the CNCL and approved drug libraries were mapped onto the global genetic interaction network (Fig. 2). This network contains 17 gene clusters corresponding to distinct bioprocesses (Fig. 2a). By placing each compound at the gene node with the most similar genetic interaction profile within the compound's top predicted biological process target, we constructed the chemical-genetic networks for the CNCL and the approved drug library (Fig. 2b, Supplementary Tables 2 and 3). The colored nodes shown on these networks correspond to the top predicted gene targets for each compound tested, many of which fall into a specific bioprocess region, as defined by the global genetic interaction network (Fig. 2, dashed outlines). Specific compounds from the CNCL library that were experimentally validated (see below) are highlighted in the network (Fig. 2b, Chinese National Compound Library). Supplementary Tables 2 and 3 list the high-confidence target GO process and direct gene target predictions for the approved drugs and CNCL compounds, respectively.

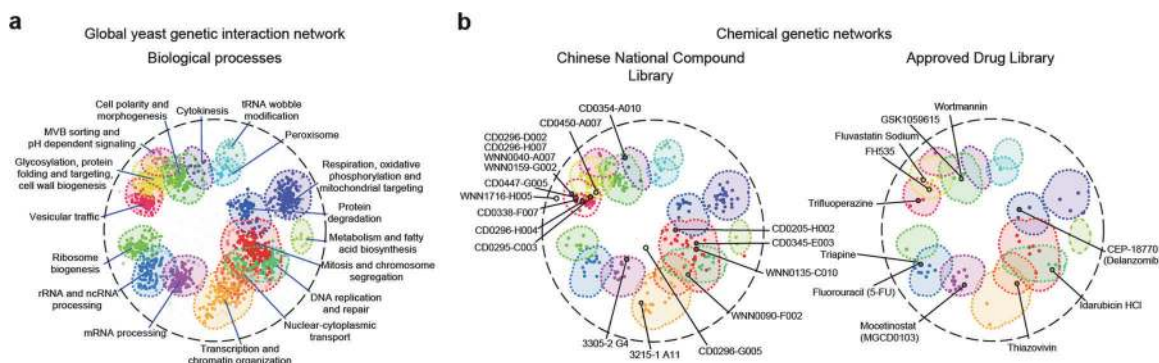
Selected drugs from the approved drug library map to different parts of the network (Fig. 2b, Approved Drug Library). In particular, drugs that target highly conserved complexes and pathways map to the expected bioprocesses. For instance, idarubicin, a chemotherapeutic agent that binds to DNA and interferes with DNA unwinding by topoisomerase II during replication [25], is positioned within the "DNA replication and repair" bioprocess. CEP-18770 (delanzomib), an inhibitor of the chymotrypsin-like activity of the proteasome [26], is located in the "Protein degradation" bioprocess and it has *RPT6*, which encodes a proteasome subunit, as its top predicted gene target. Other drugs with similar known modes of action map to the same yeast bioprocess and have the same predicted direct target gene. For example, triapine inhibits ribonucleotide reductase, while 5-fluorouracil (5-FU) suppresses thymidylate synthase. Both effectively deplete dNTP levels required for DNA and RNA synthesis. The genes encoding these targets have pleiotropic genetic interaction profiles and thus, do not localize to a distinct bioprocess cluster on the global genetic network. Nonetheless, triapine and 5-FU are mapped to the same predicted gene target

*NOP2*, which is required for processing ribosomal RNA, a finding that presumably reflects the similar role of these compounds in RNA-based cytotoxicity [27].

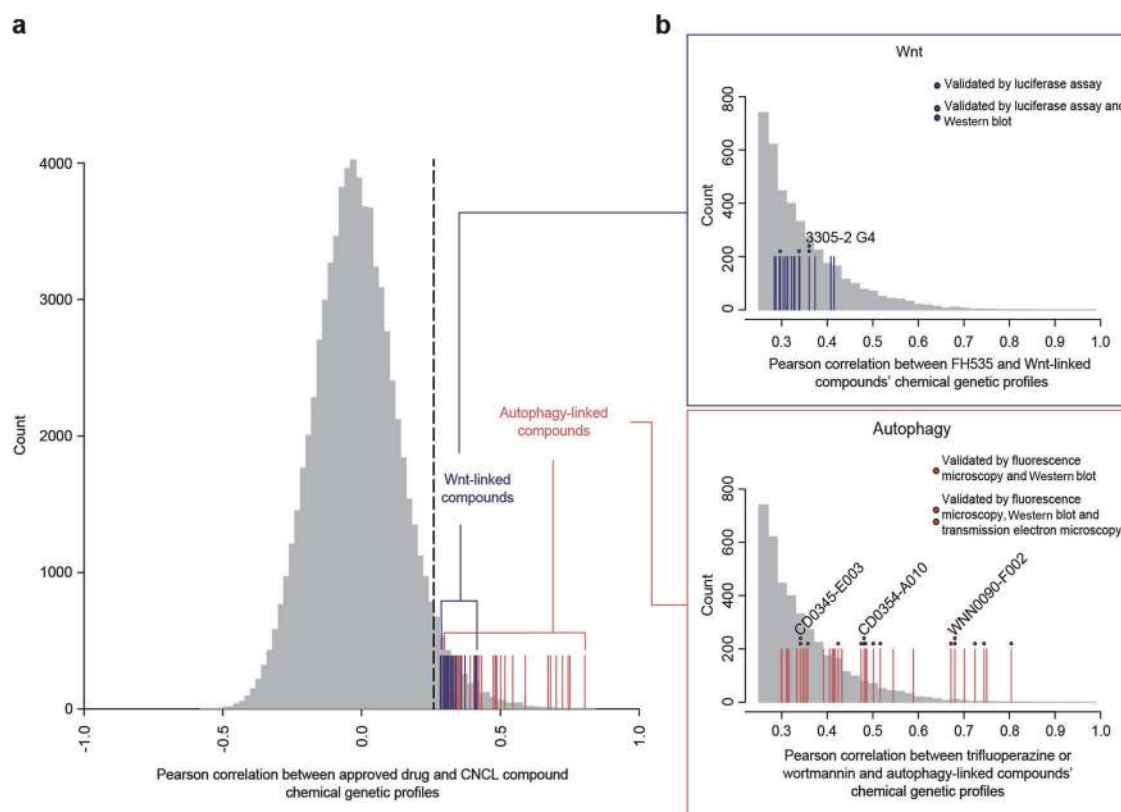
The functional signatures of each compound library can be visualized by plotting the fraction of compounds in a library that are annotated to a particular biological process, compared to the fraction of gene profiles in the genetic interaction background set that are annotated to the same process. For the CNCL high-confidence predictions, bioprocesses related to vesicle trafficking, glycosylation, and the metabolism were overrepresented, while bioprocesses related to mitosis, chromatin organization, mRNA processing, and DNA repair were strikingly underrepresented compared to a background distribution based on yeast gene annotations (Supplementary Fig. 1). The distribution is similar to those of other large libraries we have previously screened using this method, such as the RIKEN Natural Product Depository and the NIH Structural Diversity Set [18]. In contrast, the approved drug library has a more significant representation in a number of these bioprocesses, which, in part, reflects an enrichment for drugs such as idarubicin and mocetinostat (MGCD0103), which are anti-cancer agents that target highly conserved bioprocesses, such as DNA repair, mitosis, and epigenetic mechanisms of gene regulation [25, 28] (Supplementary Fig. 1).

Correlating chemical-genetic profiles to predict compound mode-of-action

Chemical-genetic profiles can also be compared directly to one another in order to identify molecules with similar modes of action. This has been shown to work for known drugs with direct yeast targets, such as concanamycin A and cerulenin [29]. While this approach is powerful, it has not been explored extensively for drugs whose primary target is absent in yeast. Comparing the chemical-genetic interaction profiles of CNCL compounds that satisfied our high-confidence cut-off to those of the approved drug library (Fig. 1c), we found that 861 (78%) CNCL compounds in the high-confidence prediction set (1109 hits) shared a similar chemical-genetic interaction profile (Pearson correlation coefficient (PCC) > 0.2588, corresponding to the top 5% of all pairs of chemical-genetic profiles tested, Supplementary Table 4) with at least one of 66 (86%) of the approved drugs from the high-



**Fig. 2** Chemical-genetic networks for the Chinese National Compound Library and a library of approved drugs. **a** The global yeast genetic interaction network. Colored nodes represent genes. Genes with highly similar genetic interaction profiles are in close proximity, while genes with dissimilar profiles are distant. The network is annotated using Spatial Analysis of Functional Enrichment (SAFE), resulting in 17 general bioprocess annotations for different regions of the network (colored outlines). Genes that were not annotated to a SAFE category are not shown. **b** The chemical-genetic networks for compounds from the Chinese National Compound Library (middle network) and approved drugs (right network) with high-confidence target predictions. Colored nodes represent compounds whose top predicted gene targets can be mapped to any general bioprocess defined by SAFE. The colored outlines are the regions of the genetic interaction network that correspond to the 17 bioprocesses. Each node/compound is placed on the map at the position of the gene within the compound's top predicted target process whose genetic interaction profile is the most similar to the chemical-genetic interaction profile of the compound. Compounds with positive validation results are highlighted in the CNCL network (big circles with black outlines). CD0296-G005 and WNN1716-H005 (white nodes) were validated to modulate autophagy, but do not belong to any SAFE bioprocess, even though they have high-confidence target GO process predictions. Other compounds with high-confidence target GO process predictions that do not map to any of the 17 SAFE bioprocesses are not shown



**Fig. 3** Distribution of profile correlation values between 1109 HCS compounds and 77 approved drugs. **a** Distribution of all correlation values between 1109 high-confidence compounds and 77 high-confidence approved drugs based on chemical-genetic profile similarity. The dashed line delineates the top 5% cut-off for correlation values (PCC > 0.2588) in the distribution. The Pearson correlation values of compounds that were selected for the autophagy validation are marked as red bars, while blue bars mark the correlation values of compounds selected for Wnt pathway validation. **b** Top: The Pearson correlation values between Wnt pathway-linked compounds and FH535, a known Wnt pathway inhibitor. Compounds with positive validation results, either from a luciferase assay to assess inhibition of gene expression from the TCF/LEF1 promoter, or from a Western blot assay to assess  $\beta$ -catenin expression levels, are marked. Treatment with 3305-G4, a novel natural product, inhibited TCF/LEF1 promoter activity and also led to reduced  $\beta$ -catenin expression. Bottom: The Pearson correlation values between autophagy pathway-linked compounds and either trifluoperazine or wortmannin, two compounds which are known to activate or inhibit autophagy, respectively. Compounds with positive validation results, either by microscopy-based assays to measure GFP-LC3 puncta and observe autophagic vacuoles, or a Western blot analysis to determine LC3-II levels, are marked. CD0345-E003, CD0354-A010, and WNN0090-F002 were bioactive hits in all assays

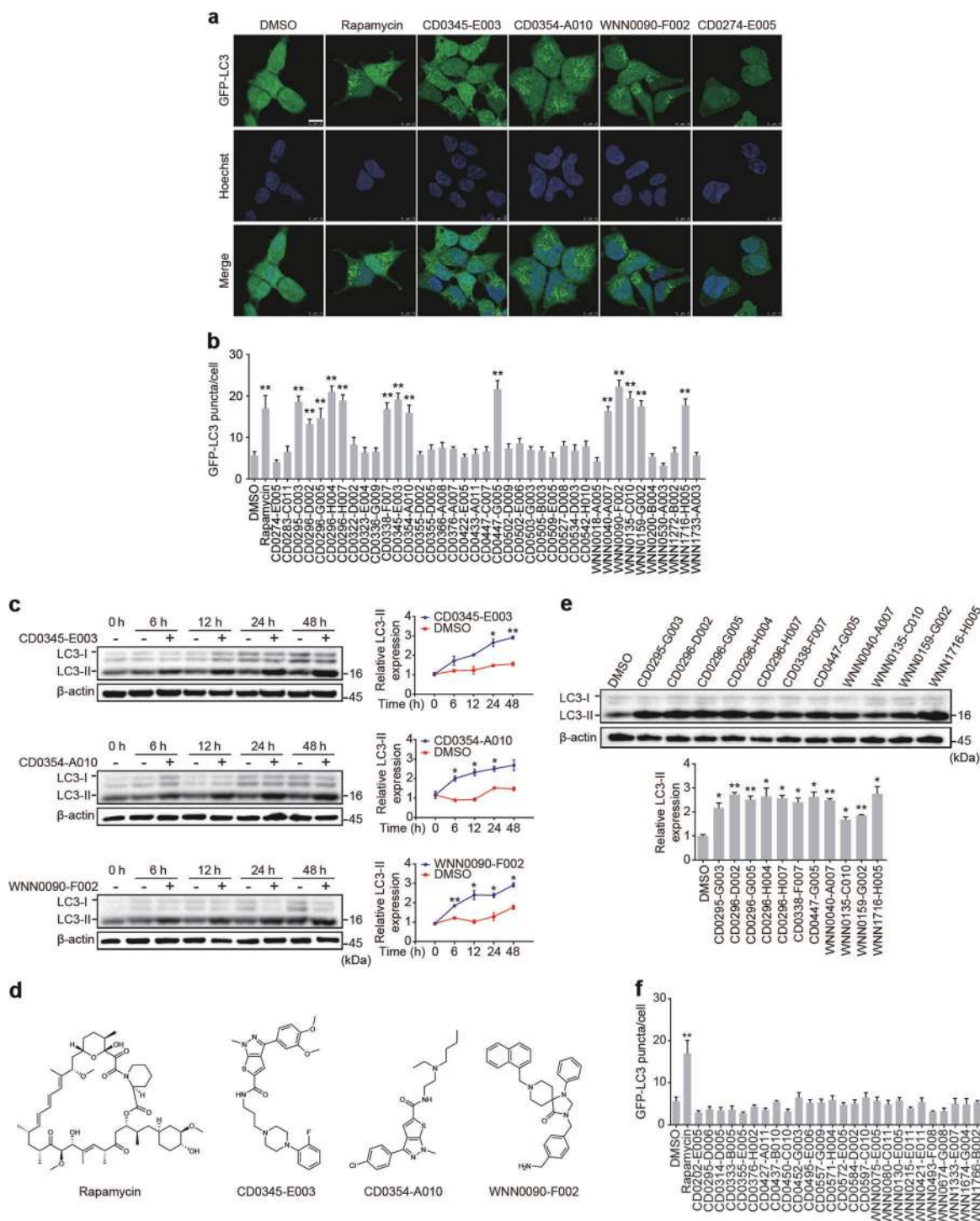
confidence prediction set (77 hits, Fig. 3). For more detailed analysis, we selected CNCL compounds that showed highly correlated chemical-genetic profiles with trifluoperazine [30] or wortmannin [31], which affect autophagy. In fact, 11 out of the 38 CNCL chemical-genetic profiles that passed the PCC threshold (>0.2588) were among the most similar to the autophagy compound profiles relative to possible tested pairs of chemical-genetic profiles (top ~0.4%, PCC > 0.5, Fig. 3). In addition, we also examined FH535, a compound that targets the Wnt/ $\beta$ -catenin signaling pathway [23], which is not found in yeast, and selected 18 CNCL compounds with chemical-genetic profiles that exhibited strong overlap with the Wnt compound profile (top ~4% of overall distribution, PCC ranging between 0.28 and 0.42, Fig. 3).

#### Pathway validation

**Autophagy.** Autophagy generally refers to macroautophagy, an evolutionarily conserved cellular process with a key role in maintaining intracellular homeostasis, and is induced by stress conditions such as nutrient starvation or drug treatment [32, 33]. The initiation of autophagy includes phagophore formation, followed by expansion into an autophagosome and fusion between autophagosomes with lysosomes, causing degradation of intracellular components [33]. In total, 38 CNCL compounds with profiles resembling that of either trifluoperazine or

wortmannin were selected for further validation (Fig. 4 and Supplementary Fig. 3). LC3 (microtubule-associated protein 1 light chain 3) is a widely used autophagy marker; the number of GFP-LC3 puncta increases either when autophagy is induced or when autolysosome fusion is blocked [33]. We found that 37% (14 of 38) of compounds predicted to target autophagy exhibited increased GFP-LC3 puncta at similar levels to rapamycin, the canonical inhibitor of mTORC and an activator of autophagy (Fig. 4a, b) [33, 34]. Since autophagy is a dynamic process, observing autophagic flux is required to confirm the action exerted by the compounds in question. During autophagy induction, cytoplasmic LC3-I is converted into LC3-II by conjugation with phosphatidylserine, and this initiates the formation and lengthening of autophagosomes [32]. As a result, LC3-II levels increase with more autophagosome formation [32, 33]. We examined LC3-II protein levels, focusing on three compounds (CD0345-E003, CD0354-A010, and WNN0090-F002, Fig. 4c, d) as representatives, and found that all three increased LC3-II in a time-dependent manner. We then tested the remaining 11 compounds with a 24-h treatment scheme, and the results showed that they all significantly increased the LC3-II protein levels (Fig. 4e), indicating that all 14 compounds were capable of modulating autophagy. Furthermore, we observed autophagic vacuoles [34], including autophagosomes (or initial autophagic vacuoles, AVi) and





**Fig. 4** Functional validation of high-confidence predicted compounds in autophagy pathway. **a** HEK293 cells stably expressing GFP-LC3 were treated with DMSO, rapamycin (200 nM) and compounds (10  $\mu$ M; 5  $\mu$ M for WNN1716-H005 due to cytotoxicity) from the predicted set, respectively. GFP-LC3 puncta formation was monitored with confocal microscopy and representative fluorescent images of cells counterstained with Hoechst (nuclei) are shown. Scale bar, 10  $\mu$ m. In the high-confidence predicted set, CD0345-E003, CD0354-A010, and WNN0090-F002 are representative active compounds which induced GFP-LC3 puncta formation, while CD0274-E005 is an inert compound. **b** Comparison of the autophagy-related phenotype induced by 38 predicted compounds based on the quantification of GFP-LC3 puncta. Green fluorescent spots represent autophagosomes ( $n > 50$  cells per sample). Data shown are means  $\pm$  SEM of three independent experiments. **c** HEK293 cells were treated with CD0345-E003, CD0354-A010 and WNN0090-F002 (10  $\mu$ M) for 6, 12, 24, and 48 h. LC3-II and  $\beta$ -actin levels were assessed by Western blotting. Densitometric values were quantified and normalized to the control which was set to 1.0. Data shown are means  $\pm$  SEM of three independent experiments. **d** Chemical structures of rapamycin, CD0345-E003, CD0354-A010, and WNN0090-F002. **e** HEK293 cells were treated with 11 other predicted compounds (10  $\mu$ M; 5  $\mu$ M for WNN1716-H005 due to cytotoxicity) for 24 h, LC3-II and  $\beta$ -actin levels were assessed by Western blotting. Densitometric values were quantified and normalized to the control which was set to 1.0. Data shown are means  $\pm$  SEM of three independent experiments. **f** Comparison of the autophagy-related phenotype induced by 26 random compounds based on the quantification of GFP-LC3 puncta ( $n > 50$  cells per sample). Data shown are means  $\pm$  SEM of three independent experiments. Statistical significance was determined with a two-tailed Student's *t*-test; \* $P < 0.01$  and \*\* $P < 0.001$

degradative autophagic vacuoles (Avd), in human HEK293 cells treated with CD0345-E003, CD0354-A010 and WNN0090-F002 using transmission electron microscopy (Supplementary Fig. 4).

As a negative control, we selected compounds with high-confidence bioprocess target predictions that did not overlap with autophagy-related pathways, excluding compounds with structures that resemble those predicted to modulate autophagy. None of 26 control compounds induced GFP-LC3 puncta formation (Fig. 4f and Supplementary Fig. 5). Furthermore, the mean PCC value for the chemical-genetic profiles of the 14 compounds with autophagic activity was 0.568, which was higher than that of complete set of 38 predicted compounds (0.456), suggesting that chemical-genetic profile similarity has a positive impact on experimental validation: for the autophagy bioprocess, a compound with a higher correlation score is more likely to be validated biologically.

Mapping the validated compounds linked to autophagy onto the CNCL chemical-genetic interaction network, we observed compounds with distinct validation phenotypes segregating into two groups. Two compounds (WNN0090-F002 and CD0345-E003) that had positive LC3 fluorescence microscopy, Western blot and transmission electron microscopy results mapped to the DNA replication/repair and mitosis/chromosome segregation bioprocesses in the network (Fig. 2b). Recently, it has been shown that agents causing different types of DNA damage can lead to targeted autophagy by activating the yeast Mec1/ATR, Tel1/ATM, and Rad53/CHEK2 kinase cascade [35]. However, the majority of the compounds (CD0295-C003, CD0296-D002, CD0296-H004, CD0296-H007, CD0338-F007, CD0447-G005, WNN0040-A007, and WNN0159-G002), validated by fluorescence microscopy and Western blotting, were linked to genes in the vesicular traffic bioprocess (Fig. 2b), which plays a major role in numerous aspects of autophagy [36]. The bioprocess predictions for these compounds included "ATP hydrolysis coupled proton transport" and "vacuolar acidification", relating to the lysosomal H<sup>+</sup> ATPase, whose activity is required for proper autophagic flux [36]. Taken together, these data show that our yeast chemical-genetic profiling platform provides a powerful approach for predicting compounds that modulate autophagy in human cells.

**Wnt/ $\beta$ -catenin pathway.** We also examined the effect of compounds that had profile similarities with FH535, a known inhibitor of the Wnt/ $\beta$ -catenin pathway, in human cells [23]. FH535 suppresses  $\beta$ -catenin/TCF-mediated transcription and exhibits anti-proliferative effects in various cancer cell lines. Although this pathway is not present in yeast [37], we found that chemical-genetic profiles associated with 18 CNCL compounds were significantly correlated with the chemical-genetic profile of FH535 (Fig. 3). Wnt/ $\beta$ -catenin signaling is an important pathway that influences cell differentiation, proliferation and fate [37, 38].  $\beta$ -catenin activity in this pathway can be grouped into two processes: accumulation and nuclear transposition. The former is linked with phosphorylation, while the latter results in the formation of the  $\beta$ -catenin-TCF/LEF1 complex, which activates transcription of downstream genes. Based on this, we conducted a reporter gene assay to investigate the effects of predicted compounds on  $\beta$ -catenin downstream transcription controlled by  $\beta$ -catenin-TCF/LEF1 complex [39], and Western blot analysis to determine compound effects on  $\beta$ -catenin accumulation and phosphorylation mediated by glycogen synthase kinase (GSK) [40].

For the first assay, we used cells with the TCF/LEF1 promoter element linked to the luciferase reporter gene, and measured luciferase activity after compound treatment. Like FH535, which decreased luciferase activity by 72.08% after 24-h treatment at 20  $\mu$ M, luminescence intensity was reduced by CD0450-A007 (39.61%), 3215-1 A11 (61.64%), and 3305-2 G4 (53.08%), respectively (Fig. 5a), suggesting that these compounds

inhibited transcription from Wnt/ $\beta$ -catenin signaling responsive elements (TCF/LEF1). We also studied the effects of these compounds on the expression level of  $\beta$ -catenin. While there was no significant difference between CD0450-A007 and 3215-1 A11 and the DMSO control, 3305-2 G4 treatment clearly decreased  $\beta$ -catenin expression, consistent with reduced luciferase activity in the TCF/LEF1 reporter assay (Fig. 5b). Dose-response experiments for luciferase activity demonstrated that 3305-2 G4 and FH535 had a similar IC<sub>50</sub> (7.44  $\mu$ M vs. 7.43  $\mu$ M, Fig. 5c), a finding that was further supported by Western blot analysis (Fig. 5d). Thus, both FH535 and 3305-2 G4 altered the expression of active  $\beta$ -catenin. For negative controls, the effects of 24 random compounds were examined similarly, and while two compounds displayed FH535-like inhibition in the luciferase assay, namely CD0203-G006 (68.69%) and CD0251-G005 (71.75%, Fig. 5e), none of the random compounds exhibited any impact on  $\beta$ -catenin expression (Fig. 5f).

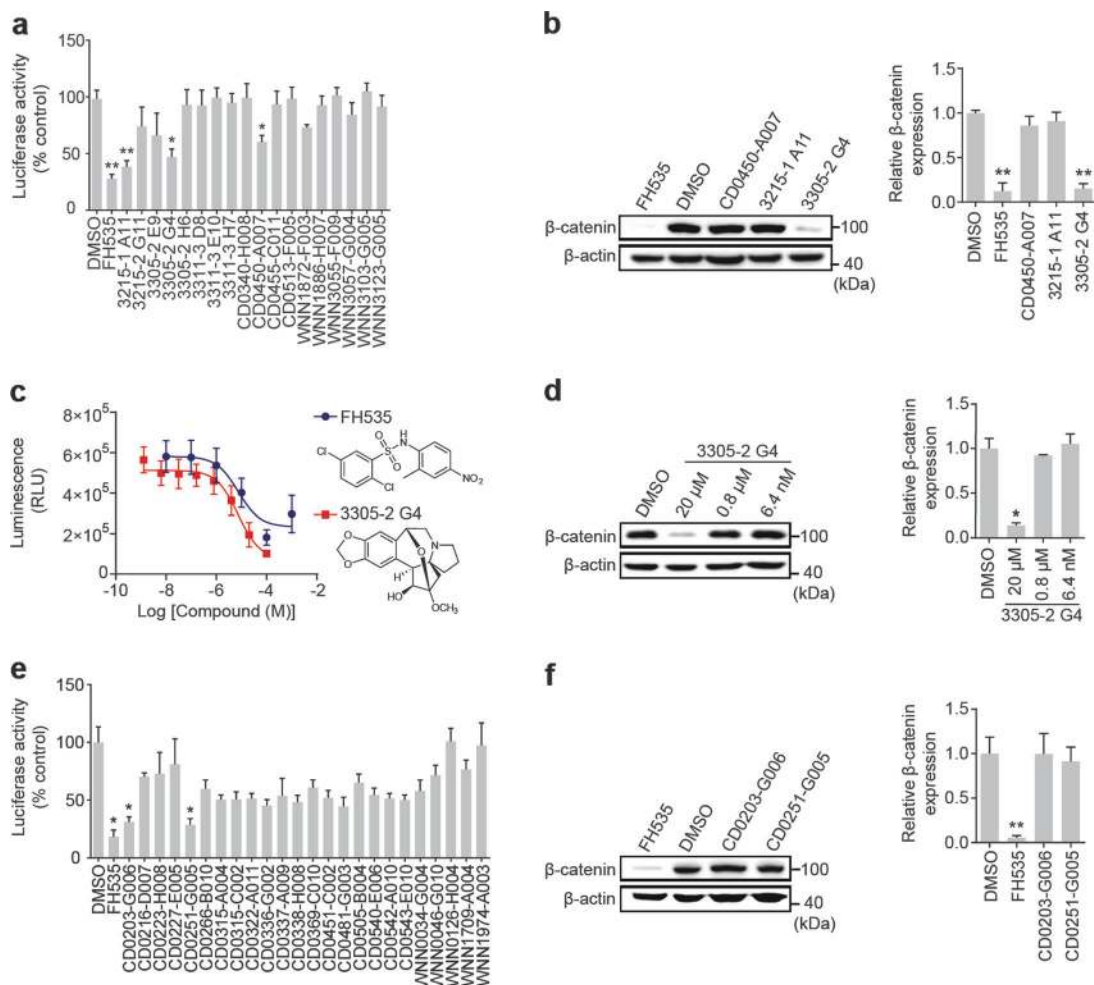
3305-2 G4 is a known alkaloid called drupacine [41] (Fig. 5c), isolated from *Cephalotaxus fortunei*, whose structure was identified by spectroscopic data analysis [42]. 3305-2 G4 mapped to "mRNA processing" (Fig. 2b), while FH535 mapped to "glycosylation, protein folding and targeting, cell wall biogenesis" bioprocesses (Fig. 2b). However, a closer examination of high-confidence target process predictions for these compounds showed that FH535 and 3305-2 G4 share a common process prediction, "Response to acidic pH" (Supplementary Table 5). The top driver genes for this target prediction are *BCK1* and *SLT2*, which encode MAPKKK and MAPK of the protein kinase C signaling pathway in yeast that controls cell wall integrity [43]. Though the Wnt pathway, the human target for FH535, is not present in yeast, it is possible that FH535 and 3305-2 G4 both act on a conserved pathway that is relevant to Wnt signaling mechanism.

To further assess the shared targets of FH535 and 3305-2 G4, a "pseudo-compound" chemical-genetic profile was calculated as a weighted mean profile of the two compounds, which accentuates large chemical-genetic interaction scores that are common between FH535 and 3305-2 G4. The mean chemical-genetic profile of FH535 and 3305-2 G4 was weighted by an element-wise product between the two compounds. Direct targets were predicted for this pseudo-compound profile, using the same process applied to calculate direct targets for real compounds. Among the top gene targets calculated for the merged FH535/3305-2 G4 chemical-genetic profile were *CDC73* (top 5th out of 4625 ranked genes, Supplementary Table 6) and *MRK1* [18] (top 6th out of 4625 ranked genes, Supplementary Table 6). *CDC73* is a component of the PAF1 complex which binds to and modulates the activity of RNA polymerases I and II, and is required for gene expression, histone modification, and telomere maintenance [44]. Its human homolog, parafibromin, is a tumor suppressor linked to parathyroid, renal, and uterine cancers [45]. Parafibromin and other components of the PAF1 complex are required for the nuclear transduction of the Wnt signal, and bind directly to the C-terminal portion of  $\beta$ -catenin to control the activity of Wnt target genes [46]. *MRK1* is one of four GSK-3 homologs in yeast, and activates Msn2-dependent transcription of stress responsive genes that function in protein degradation [47]. GSK-3 is a key enzyme in Wnt signaling, and phosphorylates  $\beta$ -catenin leading to its subsequent degradation [48]. In spite of the Wnt pathway itself being absent in yeast, the merged chemical-genetic profiles of FH535 and 3305-2 G4 were capable of predicting direct gene targets which are highly conserved and are important players in Wnt signal transduction.

## DISCUSSION

Mode-of-action discovery for novel compounds is a well-recognized rate-limiting factor for rational drug development.





**Fig. 5** Functional validation of high-confidence predicted compounds in Wnt/ $\beta$ -catenin signaling pathway. **a** Three out of eighteen compounds induced luminescence signal changes in HepG2 cells transiently transfected with a luciferase reporter gene under the control of TCF/LEF1 response element. Luciferase activity was determined after compound exposure for 24 h. Data shown are means  $\pm$  SEM of four independent experiments. **b** 3305-2 G4 and FH535 decreased  $\beta$ -catenin expression in HepG2 cells. Quantitative densitometric analysis of Western blotting from three independent experiments and the data shown are means  $\pm$  SEM. **c** Dose-response curves of luminescence signal induced by 3305-2 G4 and FH535. Data shown are means  $\pm$  SEM of three independent experiments. **d** Compound 3305-2 G4 decreased  $\beta$ -catenin expression in a dose-dependent manner in HepG2 cells. Quantitative densitometric analysis of Western blotting from three independent experiments and data shown are means  $\pm$  SEM. **e** Two out of twenty-four random set compounds induced luminescence signal changes in HepG2 cells. Data shown are means  $\pm$  SEM of three independent experiments. **f** None of the compounds decreased  $\beta$ -catenin expression in HepG2 cells. Quantitative densitometric analysis of Western blotting from five independent experiments and data shown are means  $\pm$  SEM. Statistical significance was determined with a two-tailed Student's *t*-test; \**P* < 0.01 and \*\**P* < 0.001

Chemical-genetic profiling in yeast identifies target pathways in an unbiased manner and has proven to be amenable to HTS campaigns, though the utility of this method in finding compounds that impinge on metazoan pathways, especially those that are absent from yeast but found in higher organisms, remains relatively unexplored. The present analysis annotates a complex library of small molecules employing a diagnostic chemical genomic pipeline [18]. We used a comprehensive database of yeast genetics interactions [14] as a key for interpreting the chemical-genetic profiles and linking the predicted compounds to general target bioprocesses. Indeed, when 16 compounds predicted to target the tubulin cytoskeleton assembly pathway were evaluated using techniques such as cell cycle analysis, two of them were successfully validated to disrupt tubulin function in HeLa cells (Supplementary Fig. 2). Importantly, we also developed a dataset of chemical-genetic profiles for approved drugs with known modes of action, and applied it to identify new hits whose chemical-genetic profiles are highly similar to those of known

drugs with well-established pharmacology in metazoans. Notably, we have shown that matching the chemical-genetic profiles of hits from HTS campaigns to those of known drugs allows the identification of compounds with targeted activity in mammalian cell pathways.

As proof of concept, several drugs with highly conserved cellular targets have chemical-genetic target gene and process predictions that are consistent with their known modes of action. Idarubicin, delanzomib, and mocetinostat act on diverse cellular targets (topoisomerase II, proteasome, and histone deacetylase, respectively), and these compounds were accurately linked to their target pathways and bioprocesses in yeast, namely, DNA replication and repair, proteasome/protein degradation, and mRNA processing (Fig. 2b).

Correlation analysis of chemical-genetic profiles, which compared approved drugs and CNCL compounds, was effective in linking novel compounds to autophagy, a highly conserved pathway from yeast to human, and the Wnt/ $\beta$ -catenin signaling

pathway, which is not present in yeast. The autophagy-linked compounds generally had higher PCC values (~0.8) than that of the Wnt-linked compounds (~0.35). The difference in distribution of these values likely reflects the high degree of conservation between autophagy mediators in yeast and metazoans. Thus, this guilt-by-association analysis may be particularly powerful for finding compounds that work in essential, highly conserved biological pathways.

One distinct advantage of this approach is the ability to group compounds into functional categories that may reflect their mode-of-action in cells. A striking observation was the segregation of the autophagy-linked compounds into two regions of the chemical-genetic interaction network: the “vesicular traffic” bioprocess for a majority of the compounds that were only validated by puncta formation as viewed by fluorescence microscopy and a general measurement of LC3-II expression by Western blotting, and the “DNA replication and repair” and “nuclear-cytoplasmic transport” parts of the network for compounds that had positive validation results for the more rigorous assays to assess autophagic activity (transmission electron microscopy and a LC3-II expression time course). The functional segregation of these validated hits may reflect commonalities in mode-of-action that would not have been apparent from an assay that focuses on a single phenotypic outcome such as autophagic vesicle formation.

Interestingly, the autophagy-modulating compounds can all be described as cationic amphiphilic drugs (CADs) with common physiochemical properties. CADs have a hydrophobic ring structure and a hydrophilic side chain with a charged cationic amine group, and comprise a diverse array of various substance classes such as antiarrhythmics and antipsychotics [49]. It has been recently shown that high concentrations of CADs after long exposure times tend to localize at autophagosomal compartments [50]. Accumulation of the tested CADs into the autophagosome may, at least in part, explain the observed modulation of autophagy by these compounds.

We also identified compounds that act on the Wnt/ $\beta$ -catenin signaling pathway, which is completely absent in yeast, pointing to the possibility of using yeast chemical-genetic profile similarity as a way to link compounds to drugs that may not have conserved yeast targets. Notably, FH535 and 3305-2 G4 map to separate parts of the chemical-genetic interaction network: “Glycosylation, protein folding, and cell wall biogenesis” for the former and “mRNA processing” for the latter (Fig. 2). This suggests that these compounds have distinct functional consequences in yeast, though there is significant overlap in their chemical-genetic profiles, leading to a PCC value of ~0.35 ( $P = 1.05 \times 10^{-8}$ ). Interestingly, by merging the chemical-genetic profiles of FH535 and 3305-2 G4 and weighting the interaction scores to emphasize the overlap between their profiles, we predicted direct target yeast genes (*CDC73* and *MRK1*) whose human homologs are key players in Wnt signaling pathway which either bind to  $\beta$ -catenin directly to activate expression of target genes upon transduction of the signal into the nucleus (parafibromin, homolog of *CDC73*), or stimulate cytoplasmic degradation of  $\beta$ -catenin by phosphorylation (GSK-3, homolog of *MRK1*). Furthermore, 3305-2 G4 is a known natural compound (drupacine) that is capable of suppressing cancer cell growth in vitro [51]. Whether this effect is related to Wnt/ $\beta$ -catenin signaling pathway remains to be explored. Orthologous phenotypes have been previously used for systematic discovery of non-obvious human disease models [52], and our study provides a framework for identifying orthologous chemical-genetic gene targets in yeast for compounds that act on therapeutically relevant non-conserved pathways in humans.

In conclusion, our findings demonstrate that compounds with target bioprocess predictions identified in an unbiased yeast-based chemical-genetic system can be validated in mammalian cells. These target predictions can be derived directly for highly

conserved cell functions, or they can be inferred from chemical-genetic profiles of therapeutics with known modes of action. This general approach shows great promise for utilizing yeast chemical-genetic profiling for identifying pathway-specific drug leads towards a wide variety of human diseases.

## ACKNOWLEDGEMENTS

We are indebted to Jia-mi Han and Jin-yan Tao for technical assistance, and Dian-wen Ju, Kuninori Suzuki and Michael Costanzo for valuable discussions. MWW is supported by a Shanghai Municipality Science and Technology Development Fund (Grant number 18430711500). CB and YY are supported by a JSPS Grant-in-Aid for Scientific Research on Innovative Areas (17H06411). SCL was supported by a RIKEN Foreign Postdoctoral Research Fellowship. TX is supported by National Natural Science Foundation of China (Grant number 61571202).

## AUTHOR CONTRIBUTIONS

SCL, YY, MY, CLM, and CB designed chemical genomics studies; DHY and MWW designed HTS campaign and pharmacological experiments; JN and SS performed computational analysis (processing chemical-genetic profiles from raw sequencing data, data normalization, correlation analyses and plot generation); SCL, YY, MY, CLM, and CB interpreted computational results; FLZ, YZ, WJG, LJS, QL, JBX, TX, and YYF performed HTS and validation experiments; FLZ, SCL, YZ, DHY, JMY, YY, MY, CLM, CB, and MWW analyzed the data; SCL, FLZ, YZ, DHY, CB, and MWW wrote the paper.

## ADDITIONAL INFORMATION

The online version of this article (<https://doi.org/10.1038/s41401-019-0231-y>) contains supplementary material, which is available to authorized users.

**Competing interests:** The authors declare no competing interests.

## REFERENCES

1. Botstein D, Fink G. Yeast: an experimental organism for 21st Century biology. *Genetics*. 2011;189:695–704.
2. Kachroo AH, Laurent JM, Yellman CM, Meyer AG, Wilke CO, Marcotte EM. Systematic humanization of yeast genes reveals conserved functions and genetic modularity. *Science*. 2015;348:921–5.
3. Mager WH, Winderickx J. Yeast as a model for medical and medicinal research. *Trends Pharmacol Sci*. 2005;26:265–73.
4. Simon J, Bedalov A. Yeast as a model system for anticancer drug discovery. *Nat Rev Cancer*. 2004;4:481–92.
5. Verduyck M, Vignaud H, Bynens T, Van den Brande J, Franssens V, Cullin C, et al. Yeast as a model for Alzheimer's disease: latest studies and advanced strategies. *Methods Mol Biol*. 2016;1303:197–215.
6. Tenreiro S, Franssens V, Winderickx J, Outeiro TF. Yeast models of Parkinson's disease-associated molecular pathologies. *Curr Opin Genet Dev*. 2017;44:74–83.
7. Winzeler E, Shoemaker D, Astromoff A, Liang H, Anderson K, Andre B, et al. Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science*. 1999;285:901–6.
8. Giaever G, Chu A, Ni L, Connelly C, Riles L, Véronneau S, et al. Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature*. 2002;418:387–91.
9. Mnaimneh S, Davierwala A, Haynes J, Moffat J, Peng W, Zhang W, et al. Exploration of essential gene functions via titratable promoter alleles. *Cell*. 2004;118:31–44.
10. Li Z, Vizeacoumar FJ, Bahr S, Li J, Warringer J, Vizeacoumar FS, et al. Systematic exploration of essential yeast gene function with temperature-sensitive mutants. *Nat Biotechnol*. 2011;29:361–7.
11. Baryshnikova A, Costanzo M, Dixon S, Vizeacoumar FJ, Myers CL, Andrews B, et al. Synthetic genetic array (SGA) analysis in *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*. *Methods Enzymol*. 2010;470:145–79.
12. Kuzmin E, Costanzo M, Andrews B, Boone C. Synthetic genetic array analysis. *Cold Spring Harb Protoc*. 2016;2016:326–32.
13. Costanzo M, Baryshnikova A, Bellay J, Kim Y, Spear ED, Sevier CS, et al. The genetic landscape of a cell. *Science*. 2010;327:425–31.
14. Costanzo M, VanderSluis B, Koch EN, Baryshnikova A, Pons C, Tan GH, et al. A global genetic interaction network maps a wiring diagram of cellular function. *Science*. 2016;353:1381.
15. Stockwell B. Chemical genetics: ligand-based discovery of gene function. *Nat Rev Genet*. 2000;1:116–25.
16. Bredel M, Jacoby E. Chemogenomics: an emerging strategy for rapid target and drug discovery. *Nat Rev Genet*. 2004;5:262–75.

17. Ho CH, Piotrowski J, Dixon SJ, Baryshnikova A, Costanzo M, Boone C. Combining functional genomics and chemical biology to identify targets of bioactive compounds. *Curr Opin Chem Biol.* 2011;15:66–78.
18. Piotrowski JS, Li SC, Deshpande R, Simpkins SW, Nelson J, Yashiroda Y, et al. Functional annotation of chemical libraries across diverse biological processes. *Nat Chem Biol.* 2017;13:982–93.
19. Meijer WH, van der Klei IJ, Veenhuis M, Kiel JA. ATG genes involved in non-selective autophagy are conserved from yeast to man, but the selective Cvt and pexophagy pathways also require organism-specific genes. *Autophagy.* 2007;3:106–16.
20. Simpkins SW, Deshpande R, Nelson J, Li SC, Piotrowski JS, Ward HN, et al. Using BEAN-counter to quantify genetic interactions from multiplexed barcode sequencing experiments. *Nat Protoc.* 2019;14:415–40.
21. Gene Ontology Consortium. Gene Ontology Consortium: going forward. *Nucleic Acids Res.* 2015;43:D1049–D1056.
22. Simpkins SW, Nelson J, Deshpande R, Li SC, Piotrowski JS, Wilson EH, et al. Predicting bioprocess targets of chemical compounds through integration of chemical-genetic and genetic interactions. *PLoS Comput Biol.* 2018;14:e1006532.
23. Gedaly R, Galuppo R, Daily MF, Shah M, Maynard E, Chen C, et al. Targeting the Wnt/beta-catenin signaling pathway in liver cancer stem cells and hepatocellular carcinoma cell lines with FH535. *PLoS One.* 2014;9:e99272.
24. Parsons A, Brost R, Ding H, Li Z, Zhang C, Sheikh B, et al. Integration of chemical-genetic and genetic interaction data links bioactive compounds to cellular target pathways. *Nat Biotechnol.* 2004;22:62–9.
25. Gewirtz DA, Randolph JK, Chawla J, Orr MS, Fornari FA. Induction of DNA damage, inhibition of DNA synthesis and suppression of c-myc expression by the anthracycline analog, idarubicin (4-demethoxy-daunorubicin) in the MCF-7 breast tumor cell line. *Cancer Chemother Pharmacol.* 1998;41:361–9.
26. Piva R, Ruggeri B, Williams M, Costa G, Tamagno I, Ferrero D, et al. CEP-18770: A novel, orally active proteasome inhibitor with a tumor-selective pharmacologic profile competitive with bortezomib. *Blood.* 2008;111:2765–75.
27. Hoskins J, Scott Butler J. Evidence for distinct DNA- and RNA-based mechanisms of 5-fluorouracil cytotoxicity in *Saccharomyces cerevisiae*. *Yeast.* 2007;24:861–70.
28. Fournel M, Bonfils C, Hou Y, Yan PT, Trachy-Bourget MC, Kalita A, et al. MGCD0103, a novel isotype-selective histone deacetylase inhibitor, has broad spectrum antitumor activity in vitro and in vivo. *Mol Cancer Ther.* 2008;7:759–68.
29. Hoepfner D, Helliwell SB, Sadlish H, Schuierer S, Filipuzzi I, Brachet S, et al. High-resolution chemical dissection of a model eukaryote reveals targets, pathways and gene functions. *Microbiol Res.* 2014;169:107–20.
30. Haddad D, Nakamura K. PINK1-based screen shines light on autophagy enhancers for Parkinson's disease. *Cell Chem Biol.* 2017;24:429–30.
31. Wu YT, Tan HL, Shui G, Bauvy C, Huang Q, Wenk MR, et al. Dual role of 3-methyladenine in modulation of autophagy via different temporal patterns of inhibition on class I and III phosphoinositide 3-kinase. *J Biol Chem.* 2010;285:10850–61.
32. Mizushima N, Yoshimori T. How to interpret LC3 immunoblotting. *Autophagy.* 2014;3:542–5.
33. Klionsky DJ, Abdelmohsen K, Abe A, Abedin MJ, Abeliovich H, Acedo Arozena A, et al. Guidelines for the use and interpretation of assays for monitoring autophagy (3rd edition). *Autophagy.* 2016;12:1–222.
34. Ylä-Anttila P, Vihinen H, Jokitalo E, Eskelinen E. Monitoring autophagy by electron microscopy in mammalian cells. *Methods Enzymol.* 2009;452:143–64.
35. Eapen VV, Waterman DP, Bernard A, Schiffmann N, Sayas E, Kamber R, et al. A pathway of targeted autophagy is induced by DNA damage in budding yeast. *Proc Natl Acad Sci U S A.* 2017;114:E1158–E1167.
36. Noda T. Autophagy in the context of the cellular membrane-trafficking system: the enigma of Atg9 vesicles. *Biochem Soc Trans.* 2017;45:1323–31.
37. Nusse R. Wnt signaling in disease and in development. *Cell Res.* 2005;15:28.
38. Nusse R, Clevers H. Wnt/β-Catenin signaling, disease, and emerging therapeutic modalities. *Cell.* 2017;169:985–99.
39. Shimizu N, Kawakami K, Ishitani T. Visualization and exploration of Tcf/Lef function using a highly responsive Wnt/β-catenin signaling-reporter transgenic zebrafish. *Dev Biol.* 2012;370:71–85.
40. Chen H, Hsu L, Shia Y, Lin M, Lin C. The β-catenin/TCF complex as a novel target of resveratrol in the Wnt/β-catenin signaling pathway. *Biochem Pharmacol.* 2012;84:1143–53.
41. Powell RG, Madrigal RV, Smith CR Jr, Mikolajczak KL. Alkaloids of *Cephalotaxus harringtonia* var. *drupacea*. 11-hydroxycephalotaxine and drupacine. *J Org Chem.* 1974;39:676–80.
42. Zhang H, Shyaula SL, Li JY, Li J, Yue JM. Hydroxylated daphniphyllum alkaloids from *Daphniphyllum himalense*. *J Nat Prod.* 2015;78:2761–7.
43. Errede B, Cade RM, Yashar BM, Kamada Y, Levin DE, Irie K, et al. Dynamics and organization of MAP kinase signal pathways. *Mol Reprod Dev.* 1995;42:477–85.
44. Shi X, Chang M, Wolf AJ, Chang CH, Frazer-Abel AA, Wade PA, et al. Cdc73p and Paf1p are found in a novel RNA polymerase II-containing complex distinct from the Srbp-containing holoenzyme. *Mol Cell Biol.* 1997;17:1160–9.
45. Newey PJ, Bowl MR, Thakker RV. Parafibromin—functional insights. *J Intern Med.* 2009;266:84–98.
46. Mosimann C, Hausmann G, Basler K. Parafibromin/hyrax activates Wnt/Wg target gene transcription by direct association with β-catenin/armadillo. *Cell.* 2006;125:327–41.
47. Hirata Y, Andoh T, Asahara T, Kikuchi A. Yeast glycogen synthase kinase-3 activates Msn2p-dependent transcription of stress responsive genes. *Mol Biol Cell.* 2003;14:302–12.
48. Ikeda S, Kishida S, Yamamoto H, Murai H, Koyama S, Kikuchi A. Axin, a negative regulator of the Wnt signaling pathway, forms a complex with GSK-3β and β-catenin and promotes GSK-3β-dependent phosphorylation of β-catenin. *EMBO J.* 1998;17:1371–84.
49. Halliwell W. Cationic amphiphilic drug-induced phospholipidosis. *Toxicol Pathol.* 1997;25:53–60.
50. Vater M, Möckl L, Gormanns V, Schultz Fademrecht C, Mallmann A, Ziegart-Sadowska K, et al. New insights into the intracellular distribution pattern of cationic amphiphilic drugs. *Sci Rep.* 2017;7:44277.
51. He YR, Shen YH, Li B, Li B, Lu L, Tian JM, et al. Alkaloids from *Cephalotaxus lanceolata* and their cytotoxicities. *Chem Biodivers.* 2013;10:584–95.
52. McGary KL, Park TJ, Woods JO, Cha HJ, Wallingford JB, Marcotte EM. Systematic discovery of nonobvious human disease models through orthologous phenotypes. *Proc Natl Acad Sci U S A.* 2010;107:6544–9.