

Large scale compound selection guided by cell painting reveals activity cliffs and functional relationships

Received: 2 July 2025

Accepted: 29 December 2025

Cite this article as: Sanchez, M., Bourriez, N., Bendidi, I. *et al.* Large scale compound selection guided by cell painting reveals activity cliffs and functional relationships. *Commun Biol* (2026). <https://doi.org/10.1038/s42003-025-09500-y>

Maxime Sanchez, Nicolas Bourriez, Ihab Bendidi, Ethan Cohen, Ivan Svatko, Elaine Del Nery, Hamza Tajmouati, Guillaume Bollot, Laurence Calzone & Auguste Genovesio

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

Large Scale Compound Selection Guided by Cell Painting Reveals Activity Cliffs and Functional Relationships

Maxime Sanchez^{1,2,3,4,5}, Nicolas Bourriez¹, Ihab Bendid¹, Ethan Cohen¹, Ivan Svatko^{1,6}, Elaine Del Nery⁷, Hamza Tajmouati⁵, Guillaume Bollot⁵, Laurence Calzone^{2,3,4*}, Auguste Genovesio^{1*}

*co-correspondence: laurence.calzone@curie.fr, auguste.genovesio@ens.psl.eu

1 – IBENS, Ecole Normale Supérieure, Université PSL, F-75005, Paris, France

2 – Institut Curie, Université PSL, F-75005, Paris, France

3 – INSERM, U1331, F-75005, Paris, France

4 – Mines ParisTech, Université PSL, F-75005, Paris, France

5 – Iktos, 65 rue de Prony, 75017, Paris, France

6 – Université Paris Cité, IRD, Inserm, MERIT, F-75006 Paris, France

7 – Biophenics Laboratory, Institut Curie, PSL Research University, Department of Translational Research, Cell and Tissue Imaging Facility (PICT-IBISA), 26 rue d'Ulm, 75005 Paris, France

Abstract

Traditional structure-based pre-screen compound selection relies on the assumption that chemical similarity implies similar biological activity. This paradigm narrows the exploration of chemical space and often fails to account for functional convergence, where structurally diverse compounds act through distinct targets to produce similar phenotypic effects. As a result, compounds with therapeutic potential may be overlooked. To overcome this constraint, we introduce a training-free, transfer learning-based method for large scale compound preselection that leverages deep phenotypic profiling of human cells. Notably, this enables robust pairwise comparison of phenotypic signatures across any source of the entire JUMP-CP, the largest publicly available cell painting dataset (112,480 compounds), preserving biological signals while mitigating batch effects. Validated across 65 high-throughput assays—including in vitro and in cellulo systems—our method provides efficient pre-screen enrichment of biologically active compounds, bypassing the blind spots of structure-centric approaches. Interestingly, because it is large scale, it also allows for a comprehensive analysis of structure–phenotypic activity relationships, revealing potentially thousands of compound activity cliffs, where minimal chemical changes in structure may result in profound phenotypic shifts. We show that these cliffs capture subtle, atom-level determinants of bioactivity that cannot be accessed by structure-based models. Furthermore, we demonstrate that structurally diverse compounds targeting different genes in the same biological pathway can induce either convergent or opposite phenotypes—a phenomenon validated across 30 pathways, hundreds of genes, and thousands of compounds. Finally, to support the broader community, we propose Phenoseeker, a web-based tool enabling instant retrieval of JUMP-CP compounds with similar phenotypic profiles. Together, these findings position phenotypic profiling not merely as a complementary tool, but as a transformative and scalable framework for navigating chemical space through a biological lens. By capturing rich morphological signatures that reflect functional outcomes—regardless of structural similarity—this approach enables the discovery of bioactive compounds, novel mechanisms of action, and unexpected target–pathway relationships. Applied at the scale of the JUMP-CP dataset, phenotypic profiling emerges as a powerful strategy for prioritizing compounds, illuminating activity cliffs, and accelerating the identification of therapeutically relevant candidates across diverse biological contexts.

ARTICLE IN PRESS

Introduction

Over the past decades, phenotypic drug discovery (PDD) has re-emerged as a powerful strategy for identifying novel therapeutic agents^{1,2}. In contrast to target-based approaches that focus solely on predetermined molecular targets, such as specific proteins, PDD harnesses observable cellular or organismal phenotypes². Although both strategies have produced effective therapeutics, growing evidence indicates that PDD is more successful at delivering first-in-class drugs³. Its target-agnostic nature is especially advantageous for addressing polygenic diseases or conditions linked to undruggable targets. Moreover, PDD offers the simultaneous assessment of many potential targets⁴ and the filtering of compounds based on key criteria such as solubility, toxicity, and cellular accessibility⁵. It also entails drawbacks including the need for subsequent target deconvolution and the challenge of deciphering off-target interactions from direct effects. Importantly, the success of a drug discovery campaign highly depends on the relevance of the compound library, which, in the case of PDD, is hard to predict. In fact, an efficient pre-screening compound selection method for PDD should select compounds that are likely to reproduce the desired phenotype in the considered assay.

A relevant approach toward this goal is the Cell Painting assay⁶, a standardized, multiplexed imaging technique introduced in 2013⁷. This method produces high-content, high-resolution images that capture compound- or genetically- induced cellular changes by labeling non specific organelles (including nuclei, nucleoli, mitochondria, actin, and tubulin networks) chosen for their generic nature, unrelated to any specific phenotype. As a result, it provides an unbiased, holistic view of the cell state. Alone or in combination with other omics data, Cell Painting has helped to decipher cell death pathways⁸, to assess chemical exposures in non-tumorigenic breast cells⁹, to study gene roles in glaucoma¹⁰, or to predict drug-induced toxicity in muscle cells¹¹. Initially applied to U2OS cells, the assay has now been adapted for various cell lines¹², live imaging¹³, human induced pluripotent stem cells (hiPSC)¹⁴, and even 3D cell spheroids¹⁵, demonstrating its impact in both fundamental research and therapeutic discovery¹⁶.

In parallel, the resurgence of deep learning¹⁷ has revolutionized biotechnology, as highlighted by AlphaFold's breakthrough in protein structure prediction¹⁸. Researchers have integrated ideas and architectures from diverse deep learning areas, such as representation learning¹⁹, multimodal learning²⁰, and generative models²¹ to reveal more nuanced and distinct phenotypic patterns in cell imaging data. This progress has led to novel approaches, such as InfoAlign²², Cell Painting CNN²³, or CLOOME²⁴, which can predict molecular properties or drug mechanisms of action^{25,29,33,34}. It has even helped uncover biases in CRISPR-Cas9 genome editing²⁵ and has been applied in genetic perturbation studies^{26,27}. In short, deep learning

models can be used to convert microscopy images of cells to a vector called a *deep phenotypic profile* which can further be exploited to compute similarities between phenotypic responses to various perturbations.

If the hypothesis is made that two compound perturbations producing the same cell phenotype are likely to impact the same function, then phenotypic profiles computed from Cell Painting images could be used as a proxy to perform an efficient compound library pre-selection. We recently validated this hypothesis with a small proof of concept on 3 screens using a first Cell Painting screen of 10,000 compounds^{28,29}. Other works have used cell phenotypes to identify candidates most likely to exhibit a desired bioactivity (e.g., protein inhibition, cell death, or pathway activation)^{28,30–38}. Recent methods using supervised classifiers require hundreds of compounds with known bioactivities³⁸. Some works use handcrafted cell features (sometimes these features are combined with chemical structures³⁷ or genetic profiles), while other approaches use unimodal^{31,32,34} or multimodal²⁹ deep encoders. Altogether, these methods require a training procedure, (i.e., training data), and thus are not easily transferable to drug discovery projects. A transfer learning strategy such as the one we proposed in Cohen et al²⁸ does not necessitate any training step and is straightforward to use as it only requires a positive control reference. However, this preselection has been so far limited in scale due to the unavailability of large cell painting datasets.

Scaling our pre-selection method to a much larger library is now within reach, thanks to the introduction of the JUMP-CP dataset in 2023³⁹, which contains images of over 112,000 unique compound perturbations. However, for this massive dataset to be assembled, it was co-produced by 12 laboratories around the world, spread on various experimental batches and acquired with different microscopes. Consequently, it exhibits considerable non-biological variability, known as “*batch effects*”⁴⁰. Altogether scaling compound selection to a large library dispatched on multiple sites and multiple batches such as JUMP-CP is not straightforward and it is yet unclear what combination of deep encoder, normalization strategy and profile replicate aggregation would be optimal at this scale to produce phenotypic profiles robust enough for an efficient compound selection. For this reason, to our knowledge, there is no work to date that benefits from an exhaustive view on the dataset compound perturbations for this task.

In this work, we identified an effective encoding-normalization-aggregation strategy to make possible robust comparison of any couple of compounds through their phenotypic profiles across the entire JUMP-CP dataset. We then devised a training-free transfer learning approach to perform an efficient preselection of compounds prior to a screen. We thoroughly validated its efficiency across 65 diverse screens, from internal and public databases of high-throughput assays, for a total of 40k+ readouts, covering both *in vitro* and *in cellulo* bioactivities. Interestingly, we identified that our method performs large “jumps” in

the vast chemical space (10^{33} drug-like compounds⁴¹) using phenotypic similarity as a biological proxy. These results show a greater chemical diversity among the selected compounds confirming preliminary observations made using a small Cell Painting dataset from one of the twelve JUMP-CP laboratories³⁸. We further demonstrated its ability to uncover activity cliffs that are typically missed by traditional structure-based selection approaches⁴². Our findings also show that compounds targeting the same signaling pathway tend to induce related phenotypic profiles, a trend we validated on 30 pathways spanning hundreds of genes and thousands of compounds. In order to facilitate real-world applications and accelerate drug discovery, we created a website (<https://www.phenoseeker.bio.ens.psl.eu/>) where researchers can retrieve in one click a large preselection from the 112,480 JUMP-CP compounds that are most likely to exhibit a bioactivity that is similar to the positive control query.

Results

Building phenotypic profiles that enhances signal and reduces batch effect over the whole JUMP-CP

Being able to compare cell phenotypes induced by all the compound perturbations across the whole JUMP-CP dataset required several processing steps (encoding, aggregating and normalizing). However, no clear state-of-the-art method has yet emerged for the whole JUMP-CP. We have developed a flexible pipeline (Figure 1a) that includes various approaches for each of these steps, some of which can be found in previous works^{26,40}. Our goal was to identify the best combination that would minimize batch effects while preserving biologically relevant signals. The first step, extracting a feature vector from multiple 5-channel images acquired using various microscope-imaging configurations or plate formats (i.e., 384 or 1536 wells), can be achieved either using hand-crafted methods (e.g., CellProfiler⁴³) or various deep-learning encoders. These encoders may be pretrained on natural images (e.g., ResNet50⁴⁴, DINOv2⁴⁵) or on microscopy images (e.g., ChAda-ViT⁴⁶, OpenPhenom⁴⁷) (see [Methods](#)). Following this, the feature vectors of all fields of view from one well need to be aggregated in order to obtain one profile per experimental sample. Wells distributed over different plates in different batches from different laboratories then need to be aligned through a normalization step. For instance, one can use mock treatment control units in each batch to center and scale features or perform a Typical Variation Normalization (TVN)⁴⁸, which aligns both first-order statistics and covariance structures. Other methods from the field of transcriptomics, such as neural network-based scVI⁴⁹, mixture-model Harmony⁵⁰, or nearest neighbor Scanorama⁵¹, as well as dedicated training losses^{52,53} and deep-learning

architectures⁵² were considered, but these necessitate a training step which we aimed to avoid for robustness. Finally, replicate wells need to be aggregated into a single feature profile per condition referred to as “phenotypic profile”.

To identify which combination of these steps best preserves biological signal while mitigating batch effects, two metrics were used (Figure 1b). The first metric captures the preservation of biological signal, as measured by the mean average precision (mAP) for eight positive control compounds present on every experimental plate. We seek to obtain a high value since it indicates that profiles induced by the same compound treatment are close in phenotypic space. The second metric captures batch effect removal as measured by the mAP on experimental plate labels (see [Methods](#)). We seek to approach the low value of a random retrieval indicating that profiles from the same plate (therefore from the same batch and laboratory) are not particularly closer than those from different plates. By performing a thorough sweep of more than 5000 combinations (see [Methods](#)) of processing steps described above, we identified a relevant combination of steps that optimize our criteria. First, we determined that DINOv2-Giant provides a better encoding than other deep encoders (including those trained on microscopy images) or handcrafted solutions such as CellProfiler⁴³ (Figure 1b). Following this, we determined that the most effective normalization approach consisted in fitting a whitening matrix to the DMSO negative control wells of each plate, then applying it to all the sample wells of this plate⁵⁴ then finally applying an inverse normal transformation (INT) to ensure each feature follows a normal distribution across the plate (Figure 1). Before normalization, wells from the same laboratory tend to cluster together, reflecting strong batch effects and making phenotypic profile comparison across experimental batches irrelevant. After normalization, these clusters become more intermixed while, on the contrary, positive control replicates across laboratories cluster together, indicating a reduction in batch variation and an enhancement of biological signal over experimental noise (Figure 1d). This step enables the comparison of any pair of samples, regardless of their origin, thus making the JUMP-CP large compound library fully usable for further applications.

Large-scale Cell Painting-based selection boosts active compound yield

By leveraging on this representation that allows comparisons across a large range of perturbations, we propose a practical compound selection method using similarity of Cell Painting phenotypic profiles as a proxy. Indeed, we hypothesized that two different compound perturbations exhibiting the same phenotypic response shared similar bioactivity. Our pipeline then uses cosine similarity (see [Methods](#)) to compare the phenotypic profiles of 112,480 compounds in JUMP-CP to the positive control (Figure 2a), a

compound that is known to induce the desired bioactivity. Because our approach is limited to compounds contained in the JUMP-CP dataset, all validation analyses rely on the overlap between these 112,480 compounds and those tested in external cell-based assays. In the following, we name the cosine similarity between two phenotypic profiles of compounds the “phenotypic similarity” between those two compounds. This approach enables rapid and efficient selection of bioactive compounds without requiring any machine learning training or database of compounds with known bioactivities. Designed for simplicity and usability, our method allows any researcher to identify compounds with a desired bioactivity using only a single positive control. If the phenotypic profile of this positive control is not already available in the JUMP-CP dataset, it can be generated by imaging the control with the Cell Painting assay and computing a normalized phenotypic profile with our pipeline. We validated our method using 65 screens, some of which were previously performed at the BioPhoenix platform at Institut Curie (16) and others were collected from ChEMBL (49) open source data, as well as 5 targets from Lit-PCBA benchmark⁵⁵ (see [Methods](#)). To evaluate the quality of our compound selection for these screens, we employed the normalized enrichment factor (nEF), defined as the enrichment factor (EF) divided by the maximum possible EF (Figure 2a, [Methods](#)). As shown in Figure 2b, c, d, nEF values were calculated for the selection of the top 5% compounds with the highest phenotypic similarity to the positive control. We observed comparable results at other selection percentages (see **Supplementary Fig. 1**), and the raw EF values are provided in **Supplementary Fig. 2**. Because the compounds tested in the JUMP-CP dataset and in each evaluation screen only partially overlapped, we lacked perfect positive controls for each of them. To address this, we systematically treated each identified hit from Institut Curie, ChEMBL and Lit-PCBA screens (i.e., bioactive compound) as a potential positive control, calculating the nEF for each of them. We then considered two evaluation scenarios: a pessimistic scenario where the nEF is averaged over all hits, thus accounting for false positives or weaker hits (Figure 2b, c, d, green bars); and an optimistic scenario where the highest nEF obtained was considered, reflecting a more realistic configuration where the positive control is a tool compound with a strong phenotype and a well-characterized bioactivity (Figure 2b, c, d, blue bars). In most tested screens, selecting the top 5% compounds most similar to the positive control, under the pessimistic scenario, outperformed (or sometimes matched) the hit rate of the original selection (Figure 2b, c, d, purple bars). Importantly, our method almost never yields a worse hit ratio than the original screens, even when suboptimal or false positive compounds were taken into account. Moreover, when a highly effective positive control is considered (the optimistic scenario), our method always drastically improves hit enrichment, occasionally achieving the best possible selection of compounds and, at the very least, consistently doubling the proportion of active compounds compared to the initial selection. Notably, the method’s strong performance on 5 target proteins from the Lit-PCBA benchmark—which is specifically designed to assess target-dependent effects—demonstrates that it does not merely capture false positives or off-target signals, a common pitfall in phenotypic drug discovery. This target-specific validation confirms that our approach reliably enriches the

library for compounds with relevant bioactivity, thereby enhancing its practical utility in drug discovery.

Phenotypic-based compound selection increases the chemical diversity of biologically active compounds

We further investigated the structural diversity of the compounds selected by our approach. To encode the structural topology of a chemical compound, we use Morgan fingerprints, a type of circular molecular descriptor widely used in cheminformatics⁵⁶. Using the Tanimoto similarity⁵⁷, we measured the structural similarity between the positive control from each of the Institut Curie and ChEMBL screens and the compounds selected through our biological proxy (see [Methods](#)). This structural similarity remained consistently low (Figure 3a) and was statistically significantly lower than the structural similarity between the positive control and the top 5% compounds most similar in structure (see [Methods](#) and **Supplementary Table 1** for statistical details). Remarkably, the structural similarity between the positive control and the entire pool of screened compounds was comparable to that between the positive control and the top 5% of compounds with the highest phenotypic similarity (Figure 3a). We extended this finding for each compound of the entire JUMP-CP dataset by computing the Tanimoto similarity between its top 5% with highest phenotypic similarity and its top 5% most structurally similar (Figure 3a). To illustrate this, we generated UMAP⁵⁸ embeddings of both phenotypic profiles and Morgan fingerprints (Figure 3b) for compounds tested in a given cell-based fluorescence assay at the BioPhoenix platform. The positive control (in red) and the compounds selected with our method (in blue) appear in distinct regions of chemical space, further illustrating our approach's capacity to identify chemically diverse compounds that may share similar bioactivities. Our method identifies compounds with different chemical scaffolds that produce phenotypic profiles similar to the positive control (Figure 3c). This broadens the chemical diversity of selected compounds, and therefore future hits, and creates opportunities to discover compounds with potentially improved ADME-Toxicity profiles or fewer intellectual property constraints. Conversely, we observed instances where compounds with highly similar structures produced very different phenotypes (Figure 3d).

Clustering analogs per phenotype reveals activity cliffs

The last results motivated us to explore whether the JUMP-CP dataset could also reveal structural relationships between compounds explaining variation in their phenotypic profiles. Specifically, we aimed to examine whether compounds that share the same core scaffold, even with minor chemical variations, can produce distinct phenotypic profiles, denoting “phenotypic activity cliffs.” An activity cliff is a phenomenon in which two structurally similar compounds show a significant difference in biological activity.

First, compounds from the JUMP-CP dataset were grouped according to their Bemis–Murcko scaffold⁵⁹, resulting in 532 distinct chemical series with 6 compounds or more (Figure 4a, [Methods](#)). Then, in each of these series, compounds were split in two clusters of phenotypes (defined as the highest level of a hierarchical clustering of their phenotypic profiles) and the intra- and inter-cluster phenotypic and structural similarities were computed (Figure 4a). 81 series (2277 compounds in total, see **Supplementary Fig. 3** for examples of scaffolds) showed a substantially higher intra-cluster phenotypic similarity compared to their inter-cluster phenotypic similarity (Figure 4c, [Methods](#)) while showing minimal differences structure-wise by the same metrics (Figure 4d) thus defining potentially thousands of compounds activity cliffs. We repeated this experiment using hierarchical clustering with 3, 4, and 5 clusters, consistently obtaining similar results (see **Supplementary Fig. 4**). This finding suggests that phenotypic profiles reveal activity cliffs that are difficult to predict by quantitative structure–activity relationship (QSAR) methods⁴². This illustrates the advantage of PDD and the value of using Cell Painting phenotypes to complement the structural description of compounds, enhancing chemical space exploration in drug discovery projects.

In this analysis, compounds were initially grouped by their Bemis–Murcko scaffolds to reflect the practical context of medicinal chemistry, where optimization often proceeds within a chemical series defined by a common scaffold. However, Bemis–Murcko scaffolds capture only the molecular core, and compounds sharing a scaffold can still be chemically distinct. To validate and extend our observations, we adapted the Structure–Activity Landscape Index (SALI)⁶⁰ to phenotypic differences. The distribution of SALI scores across the dataset, as well as examples of phenotypic activity cliffs identified in this manner, are presented in **Supplementary Fig. 5**.

Clustering analogs per phenotype reveals key chemical functions

To expand on our previous finding, we cherry-picked examples of chemical series where compounds exhibit distinct phenotypic profiles (Figure 5, **Supplementary Fig. 6**, **Supplementary Fig. 7**, **Supplementary Fig. 8**). For instance, a cluster of such a series was made of three compounds (C1, C2, C3 Figure 5b) killing the cells (Figure 5a). Interestingly, they share three specific moieties: (1) a bromine atom, (2) a methyl group, and (3) a sp³ carbon chain linked to the nitrogen atom (Figure 5b, in green). These three compounds are also significantly different from the negative DMSO control (Figure 5a). In contrast, compounds C4 through C11 induce different phenotypes, both from C1–C3 and from one another (Figure 5c). These differences may stem from the protonation of the scaffold's nitrogen (green highlight), as well as other structural variations (purple highlights). For instance, compound C9, which is the Murcko scaffold reference of this series, induces a phenotype similar to DMSO (Figure 5a), indicating no significant phenotypic effect. Meanwhile, adding different chemical substitutions to the Murcko scaffold induces subtle changes in the cellular phenotype, yet these differences remain clearly distinguishable with our phenotypic profiles.

To find more target specific examples, we compared the list of pairs of compounds displaying activity cliffs extracted from ChEMBL by⁶¹ to the compounds we have in JUMP-CP dataset to find one example (out of 6 pairs of compounds present in both dataset) of specific activity change (inhibition of Carbonic anhydrase VII) clearly reflected in the phenotypic profiles. This example is shown in **Supplementary Fig. 9**

Compounds reproducing the positive control phenotype tend to hit targets in the same pathway

We further investigated if compound perturbations reproducing the phenotype of a positive control were modulating the same biological pathways. A biological or signaling pathway is a series of biomolecular interactions, such as biochemical reactions or signaling events, through which cells respond to internal or external stimuli. Pathways regulate critical cellular processes including metabolism, gene expression, and cell signaling, enabling coordinated cellular behavior. To illustrate this hypothesis, we considered data from a fluorescence cell-based assay screen to identify inhibitors of the

EGFR pathway (ChEMBL1613808). We chose Dezmapinod, a non-specific inhibitor of mitogen-activated protein kinases (MAPK), as our positive control (Figure 6a). Then, we considered the top 5% compounds (15 compounds out of 300 that appeared both in ChEMBL1613808 and JUMP-CP) with highest phenotypic similarity to Dezmapinod. Among these 15 compounds (Figure 6b), three were known to specifically inhibit proteins within the same pathway: EGFR for AG1478, BRD4 for Sb-202190, and MAPK8 for SP600125 (Figure 6c). We explored a prior-knowledge database to connect these genes into a signaling pathway. Among all the existing databases, we focused on SIGNOR⁶² (see [Methods](#)). Although BRD4 was not part of the initial SIGNOR signaling pathway, it is known to modulate MYC activity and conversely^{63,64} and therefore added to the network. These findings on a specific example suggest that compounds targeting the same pathway, even when they act on different proteins, can produce a similar phenotype.

Compounds targeting the same pathway tend to induce related phenotypes

We then aimed to evaluate the extent to which the previous observations, made on a single pathway and given a specific screen positive control, generalizes to any signaling pathways. To this end, we cross-referenced genes associated with the 30 pathways from MSigBD that had more than 10 known compound-gene interactions in BindingDB targeted by compounds found in JUMP-CP. For each pathway, we identified a few hundred compounds targeting up to a few tens of genes (see [Methods](#)). We then compared the phenotypic similarities between those compounds against the phenotypic similarity between random compounds in JUMP-CP for each pathway (Figure 7a). As a result, we observed that compounds targeting the same pathway exhibit a significantly higher frequency of extreme phenotypic similarities, both highly similar and almost opposite, compared to random pairs. This result is consistent across every one of the 30 studied pathways (see [Methods](#), **Supplementary Fig. 10** and Table 1 for all distributions and statistical details). On one hand, these findings confirmed the hypothesis that compounds targeting the same pathway could be found through their phenotypic similarity. On the other hand, interestingly, it showed that a high dissimilarity could also

be used as a criteria for selection. One explanation for the latter is that compounds targeting the same pathway may trigger opposite downstream effects, leading to opposite cellular responses because of the role of the target on the pathway. Illustrating it on the G2M pathway (Figure 7), the previously described procedure led to a significantly high amount of compounds inducing both highly similar or highly dissimilar phenotypes (Figure 7a). We then examined the effect of four well-characterized compounds known to interact with specific targets in this pathway and present in the JUMP-CP dataset (SCHEMBL1578316, Reversine, NSC-625987, and Gnf-5; Figure 7b) that are displaying both similar and opposite phenotypic profiles (in terms of cosine similarities, see Figure 7c). To this end, we constructed the network interaction with the 21 G2M pathway targeted genes using OmniPath interactions via the NeKo tool⁶¹ and considered the subnetwork with only genes linked to those compounds (Figure 7d). We found that the direction of their phenotypic similarities, either similar or opposite phenotypic profiles (Figure 7c), is coherent with their downstream effects on the interaction network. For instance, Reversine inhibits an inhibitor (AURKB) of cell proliferation (BIRK5), leading to a pro-proliferative outcome, whereas NSC-625987 inhibits CDK4, blocking proliferation. These two compounds target the same pathway with opposite effects. Even if these observations cannot, yet be generalised to any pair of compounds, the visualization of their effect on networks highlights which mechanisms each of these compounds targets and can explain the observed phenotypes. The presence of this bimodal effect (highly similar / highly opposite phenotypic profiles)²⁷ in our representation of compounds displays the power and the complexity of biological information captured by phenotypic profiles.

Discussion

QSAR-based library selection methods rely on the premise that chemically similar compounds often share biological activities. While valuable, this approach inherently limits the chemical diversity of selected compounds to those resembling the reference compound in structure. Through an extensive evaluation of deep image encoding and data processing methods, we identified an approach that enables robust pairwise comparison of phenotypic profiles of compounds across the entire JUMP-CP dataset while preserving biological signals and mitigating batch effects. Leveraging this approach, we were able, for the first time, to harness this vast dataset to develop a training-free transfer learning approach for the efficient selection of bioactive compounds for screening assay. Our method was extensively validated across 65 diverse high-throughput assays from both internal and public databases, covering *in cellulo* and *in vitro* bioactivities, underscoring the broad applicability and robustness of Cell Painting phenotypic profiling across biological systems. Importantly, because it is training-free, this approach is straightforward to use, as demonstrated by our web-based tool Phenoseeker, which allows researchers to simply submit the name of a positive control from their drug discovery project and retrieve, with a single click, a list of JUMP-CP compounds most likely to exhibit similar bioactivity.

Interestingly, we enabled, for the first time, a structure-activity relationship analysis spanning the largest publicly available dataset. On one hand, we systematically identified activity cliffs, where slight modifications in compound structure result in significant phenotypic variations, highlighting crucial compound properties that would remain undetected by structure-based approaches alone. By doing so, we demonstrated that deep phenotypic profiles can provide subtle yet critical information at the atomic level. On the other hand, we found that compounds with entirely different structures can target genes in the same pathway, inducing either significantly similar or distinct phenotypic profiles. This was validated across 30 pathways, hundreds of genes, and thousands of compounds. We illustrated this through examples leveraging known compound-to-gene and gene-to-gene interaction networks, reinforcing why phenotypic similarity and dissimilarity represent meaningful metrics for selecting bioactive compounds with a diverse range of molecular structures.

While we demonstrated the effectiveness of the proposed approach, some limitations should be acknowledged. First, although JUMP-CP is the largest publicly available dataset of its kind, it provides a phenotypic view based on a single cell line: U2OS. This constraint means that compounds targeting genes not expressed in U2OS may not produce the relevant phenotypic profiles, potentially leading to missed selections. An investigation of “what can be seen” in terms of specific pathways, gene functions or gene families was performed in a recent work²⁷. Another limitation is that compound selection is restricted to the 112,480 compounds in the JUMP-CP library. While substantial, this represents only a small subset of available compounds and an even smaller fraction of the theoretical or even the smaller

drug-like chemical space. A possible solution to this limitation would be to train a cross-modal model, which could predict a phenotypic profile for any given compound structure, allowing us to identify similar JUMP-CP phenotypes even for novel compounds. We recently trained such a model based on CLIP on the full JUMP-CP dataset⁵³, but it has not yet been extensively evaluated as a compound selection tool in existing screens. Moreover, learning the complex relationship between chemical structure and cell phenotype remains challenging due to the discontinuities introduced by activity cliffs and plateaus we observe in this work. While such a cross-modal model could estimate the phenotype of a given structure, it would not generate new structures likely to induce a desired phenotype—an ultimate goal in drug discovery. To achieve this, a generative model should be conditioned on target phenotypic profiles to design compounds with targeted bioactivity, and the approach we proposed here could provide a valuable foundation for such training.

Our findings underscore the power and scalability of the proposed compound pre-selection approach in both academic research and drug discovery. By harnessing phenotypic response as a biological proxy of compounds bioactivities, our method transcends the constraints of structural similarity, venturing into unexplored regions of chemical space to uncover novel bioactive compounds. This leap beyond traditional QSAR methods not only enhances structural diversity but also reveals drug-like candidates that might otherwise remain hidden. While structure-based approaches rely on chemical resemblance to predict bioactivity, inherently restricting exploration, our method captures diverse biological mechanisms and drug-target interactions, offering a powerful and complementary alternative to conventional selection strategies.

Methods

Cell Painting dataset

Cell Painting is a multiplexed high-content imaging assay designed to capture comprehensive phenotypic signatures of cellular responses. This method employs fluorescent dyes targeting multiple cellular compartments—including nuclei, cytoskeleton, mitochondria, endoplasmic reticulum, and Golgi apparatus.

We used the JUMP-CP dataset from the eponymous consortium ³⁹, which encompasses more than 136,000 chemical and genetic perturbations represented by over 5 million images acquired across 12 laboratories. These experiments were performed following the standardized Cell Painting assay protocol where each compound is tested at 10 μ M and U2OS cells are exposed during 48h to compound perturbation. For more details, see Bray et al ⁶. Each Cell Painting image consists of five grayscale channels, each highlighting one or several key cellular organelles. Each assay plate includes negative controls (DMSO mock treatment) with 32 replicates on 384-well plates and 128 replicates on 1536-well plates, and positive controls comprising eight selected compounds (**Supplementary Table 2**), each with 4 replicates on 384-well plates and 16 replicates on 1536-well plates.

Specifically, we use the compressed version from Watkinson et al ⁵³. It encompasses 112,480 unique chemical compounds.

Raw feature extraction from images

We utilized three main categories of models to extract features from the 5-channel Cell Painting images:

(1) Handcrafted feature extraction:

We used CellProfiler ⁴³, a software designed to extract handcrafted morphological features at the single-cell level. After segmenting individual cells, features were computed per cell and then aggregated at the well level. For this study, we relied on precomputed and normalized features made publicly available by the Broad Institute (https://broadinstitute.github.io/jump_hub/howto/11_retrieve_profiles.html).

(2) Deep learning models pretrained on natural images:

Two neural network architectures originally trained on natural image datasets were adapted to our Cell Painting dataset. To accommodate the five-channel format of Cell Painting, the first convolutional layer's weights initially trained on the RGB channels were duplicated: channels 4 and 5 used copies of the weights from the red (R) and green (G) channels, respectively. Note that the specific choice of the order of Cell Painting channels does not affect the model performance, as long as consistency is maintained during inference.

- ResNet50 ⁴⁴, a convolutional deep neural network architecture that leverages residual connections widely used to extract features with robust performance across diverse tasks.
- DINOv2 ⁴⁵, a state-of-the-art self-supervised Vision Transformer (ViT) that employs a teacher-student framework to learn robust, transferable representations from unlabeled data, thereby enabling efficient feature extraction across a range of tasks. We evaluated both the Small (DINOv2-S, 21M parameters) and the Giant (DINOv2-G,

1300M parameters) versions of DINOv2 to assess the impact of model depth and complexity on feature representation.

(3) Deep learning models pretrained on microscopy images:

We also explored two architectures specifically trained on microscopy data, offering potential advantages due to domain-specific pretraining:

- ChAda-ViT (Channel Adaptive Vision Transformer)⁴⁶, a ViT model tailored for microscopy images with heterogeneous input channels (ranging from 1 to 10). ChAda-ViT employs cross-channel attention mechanisms to adaptively learn joint representations, making it well-suited to multi-channel assays like Cell Painting.
- OpenPhenom⁴⁷, an open-source Masked Autoencoder (MAE)⁶⁶ architecture explicitly trained on Cell Painting images. Developed using millions of Cell Painting images, OpenPhenom learns robust and generalizable representations through a self-supervised reconstruction task.

A detailed summary of the compound selection results for all models tested is provided in the supplementary material (**Supplementary Fig. 11**).

Several notable observations emerged from our analysis. Although CellProfiler shows excellent performance for retrieving replicates of identical perturbations and effectively reducing batch effects, this may partly result from differences in normalization procedures: CellProfiler features provided by the Broad Institute were normalized using all available negative controls from the original JUMP-CP dataset, whereas our analysis relied on approximately 70% of the wells and 6 instead of 9 images per well, as we used the compressed JUMP-CP dataset version published by Watkinson et al. Interestingly, neural networks pre-trained on natural images show strong performance, despite never having encountered Cell Painting data during their training. Among these, DINOv2 stands out by producing highly discriminative phenotypic features. ChAda-ViT achieves good results, considering its relatively small training set of ~100k microscopy images (including around 55k fluorescence microscopy images). Overall, DINOv2-Giant appears to be the optimal model for our selection method, slightly outperforming OpenPhenom.

Features alignment on DMSO

We evaluated several feature-wise alignment methods to standardize phenotypic profiles, aiming to minimize batch effects while preserving biological signals. The tested methods included:

Rescaling (Res): Each feature x_i ($i \in [1, n_{features}]$, $n_{features}$ being 192, 384, 1515 or 2048 depending on the encoder) is linearly scaled to a defined range, either $[0, 1]$ or $[-1, 1]$, using:

For the range [0,1]: $x_{i,scaled} = \frac{x_i - x_{i,min}}{x_{i,max} - x_{i,min}}$

For the range [-1,1]: $x_{i,scaled} = 2 \frac{x_i - x_{i,min}}{x_{i,max} - x_{i,min}} - 1$

The minimum and maximum across all samples.

Removal (Rem): Features with low variability across samples were removed, based either on standard deviation (std) across samples, or mean absolute deviation (MAD) across samples, using thresholds.

With $MAD(x_i) = median(abs(x_i - median(x_i)))$

Z-score normalization (Z): Each feature x_i is standardized to zero mean and unit variance using its mean across samples (μ_i) and standard deviation across samples (σ_i):

$$x_{i,zscore} = \frac{x_i - \mu_i}{\sigma_i}$$

Robust Z-score normalization (rZ): A robust variant using median across samples and median absolute deviation across samples, less sensitive to outliers:

$$x_{i,Robust\ Zscore} = \frac{x_i - median(x_i)}{MAD(x_i)}$$

Inverse normal transform (Int): Features are mapped onto a standard normal distribution by applying the rank-based inverse normal transformation:

$$x_{i,INT} = \phi^{-1}\left(\frac{rank(x_i) - 0.5}{N}\right)$$

where ϕ^{-1} is the inverse cumulative normal distribution and N is the number of samples.

Sphering (Sph): We whiten each plate's feature matrix $X \in R^{n \times d}$ (with n samples, d features) by centering and then multiplying by the inverse square root of its covariance. Concretely:

$$X_{Centered} = (X - \mu)$$

where μ is the feature-wise mean vector.

$$\Sigma = \frac{1}{1-n} X_{Centered}^T X_{Centered} \quad X_{Centered} = V \Lambda V^T \text{ (Eigen-decomposition)}$$

where Σ is the covariance matrix and $\Lambda = diag(\lambda_1, \dots, \lambda_d)$

Two whitening methods were tried :

- PCA-whitening: $X_{Sph-PCA} = X_{Centered} V(\Lambda + \epsilon I)^{-\frac{1}{2}} \sqrt{n-1}$
- ZCA-whitening: $X_{Sph-ZCA} = X_{Centered} V(\Lambda + \epsilon I)^{-\frac{1}{2}} V^T \sqrt{n-1}$

where $\epsilon = 10^{-6}$ is a small “fudge factor” to avoid division by zero on near-rank-deficient plates.

ZCA preserves each sample’s orientation in the original feature space, while PCA rotates onto the principal-component axes.

Each method was evaluated individually as well as in various combinations. For Z-score and Sphering methods, transformations were fitted either solely on negative controls (DMSO wells) or across all wells on a given plate. All normalization procedures were performed on a plate-wise and feature-wise basis.

Metric used to compare phenotypic profiles

In this article, we used cosine similarity to compare phenotypic profiles due to its ability to identify biologically related samples based on the similarity in their patterns of change rather than on the magnitude of these changes⁶⁵. Additionally, cosine similarity provides directional information, capturing the relationship orientation between two phenotypic profiles.

The cosine similarity is defined as :

$$\text{Cosine Similarity}(a, b) = \frac{a \cdot b}{||a|| ||b||}$$

Where $a \cdot b$ is the scalar product between a and b , and $||.||$ is the Euclidean norm of the vector.

Metrics and data used to compare profiling methods

We employed specific metrics tailored to our downstream application—bioactive compound selection through ranking and retrieval. Although these metrics are not flawless, they effectively capture our primary objectives. In particular, we utilized the **mean average precision (mAP)**⁶⁷ metric, a standard evaluation measure widely employed in information retrieval and ranking tasks. Specifically, the average precision (AP) for a given query is calculated by averaging the precision values at each position in the ranked list where a relevant item appears. The mAP is then obtained by averaging AP scores across all queries, effectively capturing both the quality of retrieved items

and their ranking order. Higher mAP scores thus indicate superior retrieval performance and accuracy.

For a given query (a well phenotypic profiles), the Average Precision (AP) is defined as:

$$AP = \frac{1}{R} \sum_{k=1}^{n_{well\ profiles}} P(k) \cdot I\{label_k == label_{query}\}$$

Where :

- $P(k)$ is the precision at rank k (here, $\frac{\text{Number of true positives in the top } k \text{ wells}}{k}$)
- R the total number of relevant (ie, with the same label as the query one) well profiles
- $I\{.\}$ the indicator function (1 if condition holds, 0 otherwise)

To quantify batch effect reduction, mAP was computed using phenotypic profiles from both negative (DMSO) and positive control wells, with experimental batch identifiers serving as labels. Conversely, to assess the preservation of biological signals, mAP was computed exclusively on positive control wells, using compound identities as labels. The mAP was not computed with sample wells to avoid potential bias due to sample phenotypic profiles distributions.

We can observe that all 8 positive controls do not have the same phenotypic impact on cells (**Supplementary Fig. 12**, **Supplementary Fig. 13**). Indeed, some have strong effects (JCP2022_046054, JCP2022_037716 or JCP2022_035095), some have weaker effects (JCP2022_012818 or JCP2022_064022) and some are not distinguishable from DMSO negative control (JCP2022_085227, JCP2022_050797). This explains the huge difference in mAP value between those compounds and the average mAP across those 8 compounds being only around 0.5 for best models and normalisations. The strong batch effect observed on **Supplementary Fig. 13** also explains mAPs values.

For computational efficiency, we initially evaluated all possible combinations of normalization methods (over 5,000 in totals) using a subset of five plates per laboratory, each selected from distinct experimental batches. Subsequently, the 100 best-performing combinations were re-evaluated using approximately half of the JUMP-CP dataset (~600 plates), explicitly chosen to maximize stratification across experimental batches and laboratories. This evaluation process was conducted independently for each feature extractor: ChAda-ViT, ResNet50, DINOv2-G, DINOv2-S, and OpenPhenom.

Figure 1 shows the performance results for all tested combinations, while **Supplementary Fig. 14** presents detailed results for the 100 best-performing combinations evaluated on the larger subset (~600 plates). Importantly, results remained consistent from the initial smaller set of plates (45 plates) to the expanded validation (600 plates), supporting the robustness of our normalization approach.

Compound structural representation using Morgan fingerprints

To describe the structural topology of chemical compounds, we computed Morgan fingerprints⁵⁶ using the RDKit toolkit. Specifically, Morgan fingerprints were generated with a radius of 2 and a bit-vector length of 1024. This fingerprinting method encodes structural information by iteratively considering atom neighborhoods within the specified radius, capturing both local chemical environments and global topological properties of the compound. The similarity between compounds was quantified using the Tanimoto similarity⁵⁷, a widely adopted measure for comparing molecular fingerprints. The Tanimoto similarity assesses the overlap of structural features between pairs of compounds.

It is defined as:

$$T(A,B) = \frac{c}{a + b - c}$$

where:

- a and b are the number of bits set to 1 in the fingerprints of compounds A and B , respectively,
- c is the number of bits set to 1 in both fingerprints.

Chemical diversity of JUMP-CP

We compute a UMAP projection of the Morgan fingerprint of all compounds in JUMP-CP alongside those from ChEMBL and DrugBank to assess the extent of the chemical space covered by JUMP-CP compounds relative to known or previously studied compounds (see **Supplementary Fig. 15**). These UMAP projections show that JUMP-CP compounds broadly overlap with both ChEMBL and DrugBank while also occupying distinct regions in chemical space, underscoring the shared and novel structural diversity within the JUMP-CP dataset.

Assays description

Screen: A high-throughput experimental procedure to rapidly evaluate the activity of a library of compounds.

Assay: A standardized test that quantifies or qualifies the interaction, activity, or presence of a compound or biomolecule, used to validate screening results.

Institut Curie BioPhoenix screens (Supplementary Fig. 16)

We included assays from the Institut Curie's BioPhoenix screening platform based on specific criteria. Assays were selected if at least 100 compounds from the JUMP-CP dataset were screened, and if at least five active compounds (hits) were identified in the original assay. These screens were all conducted using cellular models, primarily employing imaging-based readouts, although some relied on fluorescence-based methods such as CellTiter-Glo. Many assays featured overlapping sets of compounds, typically originating from shared chemical libraries like the Prestwick library. Despite this overlap, hits identified varied significantly from one assay to another, even when the same compounds were tested. In total, 16 BioPhoenix screens met these criteria.

ChEMBL screens (<https://www.ebi.ac.uk/chembl/>) (Supplementary Fig. 17)

We also incorporated assays from the ChEMBL database, applying the same criteria as those used for BioPhoenix assays. Assays included had at least 5 active and 100 inactive compounds also present in the JUMP-CP dataset. To focus specifically on challenging screening tasks, we retained only assays with a hit rate lower than 35%. The selected ChEMBL assays represent diverse experimental methodologies and biological contexts, covering *in vitro* and *in cellulo* systems. Overall, this resulted in 44 selected ChEMBL screens.

Lit-PCBA benchmark ⁵⁵ (Supplementary Fig. 18)

Finally, we utilized the Lit-PCBA dataset for additional benchmarking. From Lit-PCBA, we retained only targets with at least 10 active compounds and a minimum of 100 total compounds present in the JUMP-CP dataset. This filtering resulted in five targets meeting our criteria: ALDH1, MAPK1, PKM2, CDR, and ESR1-antagonist.

Collectively, Lit-PCBA target, ChEMBL and BioPhoenix assays encompassed significant chemical diversity, covering more than 10,000 unique compounds from the JUMP-CP dataset—representing around 10% of the dataset in its entirety.

A supplementary table for each source (BioPhoenix, ChEMBL, Lit-PCBA) provides detailed information including the screen or target identifiers, assay type, hit rate, and total number of tested compounds.

Additionally, **Supplementary Fig. 19** illustrates the proportion of the JUMP-CP dataset evaluated across all selected assays.

Normalized enrichment factor

To evaluate the ability of our method to select bioactive compounds, we compute enrichment factors⁶⁸. We chose a threshold of 5% to select at least 5 compounds for all screens where we have 100 tested compounds (but results are consistent across thresholds, see **Supplementary Fig. 1**). Those values were normalized to help the comparison across screens with a very diverse number of compounds and hit rate. A normalized enrichment factor (nEF) of 100% means the best possible selection of compounds.

The enrichment factor at 5% is computed as:

$$EF_{5\%} = \frac{\frac{a}{n}}{\frac{A}{N}} = \frac{a \cdot N}{n \cdot A}$$

where:

- a is the number of active compounds in the top 5% selected compounds,
- n is the total number of compounds selected (5% of N),
- A is the total number of actives in the dataset,
- N is the total number of compounds tested.

The Normalized Enrichment Factor (nEF) is then defined as:

$$nEF_{5\%} = \frac{EF_{5\%}}{EF_{ideal,5\%}} \cdot 100$$

where: $EF_{ideal,5\%}$ is the theoretical maximum enrichment at 5% (i.e., when all selected compounds are active or all active compounds are selected).

Bemis–Murcko scaffolds and phenotypic subclusters

To systematically group compounds according to their structural similarity, we computed molecular scaffolds using RDKit's implementation of the Bemis–Murcko scaffold definition. Note that RDKit's implementation differs slightly from the original Bemis–Murcko definition⁵⁹:

it retains the first atom of exocyclic substituents attached via double bonds, distinguishing certain structural motifs (e.g., differentiating between C1CC1=O and C1CC1). Further details about these differences are documented in the RDKit community discussion (<https://github.com/rdkit/rdkit/discussions/6844>).

Once scaffolds were generated within each scaffold-defined group, we performed hierarchical clustering to further subdivide compounds into distinct phenotypic subgroups. We retained groups where each cluster contained at least three compounds to ensure sufficient representation for meaningful analysis. To explore the robustness of our approach, we systematically tested scenarios with two, three, four, or five clusters per scaffold, resulting in corresponding minimal scaffold group sizes of 6, 9, 12, or 15 compounds.

Detailed statistical results from these different clustering configurations are provided in **Supplementary Fig. 4**.

Scaffolds were selected as exposing subclusters with substantially different phenotypes when intra-cosine phenotypic similarity was higher than inter-cosine phenotypic similarity by at least 0.2 (in Figure 4). The value 0.2 corresponds to a cosine similarity greater than 3 standard deviations from the mean of the whole dataset (see **Supplementary Fig. 20** for distribution of similarity across all JUMP-CP compounds).

Pathways and networks construction

The EGFR pathway used in Figure 6 corresponds to the signaling network provided by The SIGNaling Network Open Resource (SIGNOR 3.0, <https://signor.uniroma2.it>) with a confidence superior to 0.2 for all interactions. Additionally, the Brd4 gene was manually included based on evidence from previous literature facts^{63,64}.

For Figure 7 and Table 1, gene sets were obtained from The Molecular Signatures Database (MSigDB, <https://www.gsea-msigdb.org/>), a curated collection of annotated gene sets derived from diverse biological sources, including published experimental data, expert-curated pathways, and computational predictions. We used the Hallmarks gene sets that focus on cancer but it can be generalized. Compound-target relationships were retrieved from BindingDB (<https://www.bindingdb.org/>), a comprehensive database of experimentally validated compound-gene interactions. We retained only compounds present in both BindingDB and the JUMP-CP dataset, along with genes targeted by these remaining compounds. Detailed statistics regarding genes, filtered genes, and remaining compounds are presented in Table 1.

The gene interaction network shown in Figure 7 was constructed using the NeKo tool⁶⁵, which aims at reconstructing connected networks from curated prior-knowledge databases (i.e.,

OmniPath (<https://omnipathdb.org/>), a comprehensive resource for intra- and intercellular signaling knowledge). We built the network specifically with the 21 genes from the G2M pathway targeted by JUMP-CP compounds and focused on the SIGNOR database only. This complete NeKo-generated network is available in **Supplementary Fig. 21** The subnetwork in Figure 7. highlights interactions relevant to the four selected compounds.

Pathways & Phenotypes table and statistics

For each gene set corresponding to a given biological pathway, we retained only those pathways in which at least 10 distinct genes were targeted by at least 50 compounds from the JUMP-CP dataset, resulting in 30 pathways for further analysis. For each selected pathway, we computed pairwise cosine similarities between phenotypic profiles of all compounds known to target genes within that pathway. As a comparative control, we also computed cosine similarities among a random set of non-targeting compounds, sampled at a ratio of 10 non-targeting compounds per active compound. The resulting distributions of cosine similarities for each pathway, including both active and random compound pairs, are shown in **Supplementary Fig. 10**.

Statistics and Reproducibility

All metrics for batch effect removal, signal conservation (mean Average Precision) and bioactive compounds retrieval (Enrichment Factors) are deterministic. Replicates are defined at the field of view, well level by field of view, or well that have been exposed to the same chemical compound perturbation (or no perturbation). FOV replicates are technical replicates. Well replicates are biological replicates. All metrics for batch effect removal and signal conservation are only evaluated with biological replicates.

To compare the distribution of Tanimoto similarities in Figure 3, we first checked normality of the paired differences using the Shapiro–Wilk test at a significance level of $\alpha=0.05$. Because most differences did not follow a normal distribution ($p<0.05$, see **Supplementary Table 1**), we employed the nonparametric Wilcoxon signed-rank test with a one-sided alternative hypothesis (e.g., “Top 5% Structures” > “All Tested Compounds”). This approach provides a robust evaluation of whether one distribution exhibits systematically higher (or lower) Tanimoto similarity compared to another.

The resulting p -values were then converted into significance thresholds ($p < 0.05$, < 0.01 , < 0.001 , etc.), which are visually depicted in box plots using the star notation (e.g. *, **, ***).

etc.). All test values and raw p-values are displayed in **Supplementary Table 1**, the number of samples in **Supplementary Fig. 19**

Activity cliffs detection is deterministic.

To determine whether active compounds exhibited more extreme similarity or dissimilarity, compared to random compound pairs, we first tested differences in variance between these two distributions using Levene's test. Additionally, to assess whether active compound pairs were significantly enriched in highly similar (cosine similarity $> \text{mean} + 3 \text{ standard deviations}$, see **Supplementary Fig. 20** for distribution of similarity across all JUMP-CP compounds) or highly dissimilar pairs (cosine similarity $< \text{mean} - 3 \text{ standard deviations}$, see **Supplementary Fig. 20** for distribution of similarity across all JUMP-CP compounds), we calculated the percentage of pairs meeting these thresholds in both active and random distributions, and assessed statistical significance using Fisher's exact test. Detailed statistical results, including test statistics and p-values, are provided in Table 1.

Data availability

JUMP-CP dataset is publicly available (see: https://broadinstitute.github.io/jump_hub/). Screens data from ChEMBL⁶⁹ is publicly available; we used version 35 of this database (see: <https://www.ebi.ac.uk/chembl/>). Screens data from BioPhoenix are not publicly available. Compounds-targets interactions are publicly available (see: <https://www.bindingdb.org/>). The gene sets are publicly available (see: <https://www.gsea-msigdb.org/>).

Code availability

We provide scripts at https://github.com/mxfly14/2025_sanchez_phenoseeker⁷⁰ for the automated preprocessing of the data, test and evaluation of normalization methods and reproduction of the figures and plots. The code to download the compressed version of JUMP-CP is available at https://github.com/gwatkinson/jump_download.

References

1. Vincent, F. *et al.* Phenotypic drug discovery: recent successes, lessons learned and new

- directions. *Nat. Rev. Drug Discov.* **21**, 899–914 (2022).
2. Swinney, D. C. Phenotypic vs. target-based drug discovery for first-in-class medicines. *Clin. Pharmacol. Ther.* **93**, 299–301 (2013).
 3. Swinney, D. C. & Anthony, J. How were new medicines discovered? *Nat. Rev. Drug Discov.* **10**, 507–519 (2011).
 4. Reisen, F. *et al.* Linking Phenotypes and Modes of Action Through High-Content Screen Fingerprints. *ASSAY Drug Dev. Technol.* **13**, 415–427 (2015).
 5. Brodin, P., DelNery, E. & Soleilhac, E. Criblage phénotypique à haut contenu pour la chémobiologie et ses enjeux. *médecine/sciences* **31**, 187–196 (2015).
 6. Bray, M.-A. *et al.* Cell Painting, a high-content image-based assay for morphological profiling using multiplexed fluorescent dyes. *Nat. Protoc.* **11**, 1757–1774 (2016).
 7. Gustafsdottir, S. M. *et al.* Multiplex Cytological Profiling Assay to Measure Diverse Cellular States. *PLoS ONE* **8**, e80999 (2013).
 8. Schorpp, K. *et al.* CellDeathPred: A Deep Learning framework for Ferroptosis and Apoptosis prediction based on cell painting. 2023.03.14.532633 Preprint at <https://doi.org/10.1101/2023.03.14.532633> (2023).
 9. Tapaswi, A., Cemalovic, N., Polemi, K. M., Sexton, J. Z. & Colacino, J. A. Applying Cell Painting in Non-Tumorigenic Breast Cells to Understand Impacts of Common Chemical Exposures. 2024.04.30.591893 Preprint at <https://doi.org/10.1101/2024.04.30.591893> (2024).
 10. Greatbatch, C. J. *et al.* High throughput functional profiling of genes at intraocular

- pressure loci reveals distinct networks for glaucoma. 2023.07.10.548340 Preprint at <https://doi.org/10.1101/2023.07.10.548340> (2023).
11. Lambert, R. *et al.* Drug-induced cytotoxicity prediction in muscle cells, an application of the Cell Painting assay. 2024.02.08.579439 Preprint at <https://doi.org/10.1101/2024.02.08.579439> (2024).
 12. Lejal, V., Rouquié, D. & Taboureau, O. Cell morphology and gene expression: tracking changes and complementarity across time and cell lines. 2024.08.30.610494 Preprint at <https://doi.org/10.1101/2024.08.30.610494> (2024).
 13. Garcia-Fossa, F. *et al.* Live Cell Painting: image-based profiling in live cells using Acridine Orange. 2024.08.28.610144 Preprint at <https://doi.org/10.1101/2024.08.28.610144> (2024).
 14. Platani, M. *et al.* Screening for variable drug responses using human iPSC cohorts. 2023.06.16.545161 Preprint at <https://doi.org/10.1101/2023.06.16.545161> (2023).
 15. Ringers, C. *et al.* High-content morphological profiling by Cell Painting in 3D spheroids. 2025.02.05.636642 Preprint at <https://doi.org/10.1101/2025.02.05.636642> (2025).
 16. Seal, S. *et al.* A Decade in a Systematic Review: The Evolution and Impact of Cell Painting. 2024.05.04.592531 Preprint at <https://doi.org/10.1101/2024.05.04.592531> (2024).
 17. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
 18. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
 19. Kim, V., Adaloglou, N., Osterland, M., Morelli, F. M. & Zapata, P. A. M. Self-supervision advances morphological profiling by unlocking powerful image representations.

- 2023.04.28.538691 Preprint at <https://doi.org/10.1101/2023.04.28.538691> (2023).
20. Zheng, S. *et al.* Cross-Modal Graph Contrastive Learning with Cellular Images. *Adv. Sci.* **11**, 2404845 (2024).
21. Lamiable, A. *et al.* Revealing invisible cell phenotypes with conditional generative modeling. *Nat. Commun.* **14**, 6386 (2023).
22. Liu, G. *et al.* Learning Molecular Representation in a Cell. Preprint at <https://doi.org/10.48550/arXiv.2406.12056> (2024).
23. Moshkov, N. *et al.* Learning representations for image-based profiling of perturbations. *Nat. Commun.* **15**, 1594 (2024).
24. Sanchez-Fernandez, A., Rumetshofer, E., Hochreiter, S. & Klambauer, G. CLOOME: contrastive learning unlocks bioimaging databases for queries with chemical structures. *Nat. Commun.* **14**, 7339 (2023).
25. Lazar, N. H. *et al.* High-resolution genome-wide mapping of chromosome-arm-scale truncations induced by CRISPR-Cas9 editing. 2023.04.15.537038 Preprint at <https://doi.org/10.1101/2023.04.15.537038> (2023).
26. Celik, S. *et al.* Building, Benchmarking, and Exploring Perturbative Maps of Transcriptional and Morphological Data. 2022.12.09.519400 Preprint at <https://doi.org/10.1101/2022.12.09.519400> (2024).
27. Chandrasekaran, S. N. *et al.* Morphological map of under- and overexpression of genes in human cells. *Nat. Methods* **22**, 1742–1752 (2025).
28. Cohen, E. *et al.* Cell painting transfer increases screening hit rate. *Biol. Imaging* **3**, e4

(2023).

29. Bray, M.-A. *et al.* A dataset of images and morphological profiles of 30 000 small-molecule treatments using the Cell Painting assay. *GigaScience* **6**, 1–5 (2017).
30. Wang, S. *et al.* PhenoScreen: A Dual-Space Contrastive Learning Framework-based Phenotypic Screening Method by Linking Chemical Perturbations to Cellular Morphology. 2024.10.23.619752 Preprint at <https://doi.org/10.1101/2024.10.23.619752> (2024).
31. Simm, J. *et al.* Repurposing High-Throughput Image Assays Enables Biological Activity Prediction for Drug Discovery. *Cell Chem. Biol.* **25**, 611–618.e3 (2018).
32. Seal, S. *et al.* Merging bioactivity predictions from cell morphology and chemical fingerprint models using similarity to training data. *J. Cheminformatics* **15**, 56 (2023).
33. Herman, D. *et al.* Leveraging Cell Painting Images to Expand the Applicability Domain and Actively Improve Deep Learning Quantitative Structure–Activity Relationship Models. *Chem. Res. Toxicol.* **36**, 1028–1036 (2023).
34. Hofmarcher, M., Rumetshofer, E., Clevert, D.-A., Hochreiter, S. & Klambauer, G. Accurate Prediction of Biological Assays with High-Throughput Microscopy Images and Convolutional Networks. *J. Chem. Inf. Model.* **59**, 1163–1171 (2019).
35. Trapotsi, M.-A. *et al.* Comparison of Chemical Structure and Cell Morphology Information for Multitask Bioactivity Predictions. *J. Chem. Inf. Model.* **61**, 1444–1456 (2021).
36. Ha, S. V. *et al.* Low concentration cell painting images enable the identification of highly potent compounds. *Sci. Rep.* **14**, 24403 (2024).

37. Moshkov, N. *et al.* Predicting compound activity from phenotypic profiles and chemical structures. *Nat. Commun.* **14**, 1967 (2023).
38. Haslum, J. F. *et al.* Cell Painting-based bioactivity prediction boosts high-throughput screening hit-rates and compound diversity. 2023.04.03.535328 Preprint at <https://doi.org/10.1101/2023.04.03.535328> (2023).
39. Chandrasekaran, S. N. *et al.* JUMP Cell Painting dataset: morphological impact of 136,000 chemical and genetic perturbations. 2023.03.23.534023 Preprint at <https://doi.org/10.1101/2023.03.23.534023> (2023).
40. Arevalo, J., Su, E., Dijk, R. van, Carpenter, A. E. & Singh, S. Evaluating batch correction methods for image-based cell profiling. 2023.09.15.558001 Preprint at <https://doi.org/10.1101/2023.09.15.558001> (2024).
41. Polishchuk, P. G., Madzhidov, T. I. & Varnek, A. Estimation of the size of drug-like chemical space based on GDB-17 data. *J. Comput. Aided Mol. Des.* **27**, 675–679 (2013).
42. van Tilborg, D., Alenicheva, A. & Grisoni, F. Exposing the Limitations of Molecular Machine Learning with Activity Cliffs. *J. Chem. Inf. Model.* **62**, 5938–5951 (2022).
43. Stirling, D. R. *et al.* CellProfiler 4: improvements in speed, utility and usability. *BMC Bioinformatics* **22**, 433 (2021).
44. He, K., Zhang, X., Ren, S. & Sun, J. Deep Residual Learning for Image Recognition. in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* 770–778 (2016). doi:10.1109/CVPR.2016.90.
45. Oquab, M. *et al.* DINOv2: Learning Robust Visual Features without Supervision.

Preprint at <https://doi.org/10.48550/arXiv.2304.07193> (2023).

46. Bourriez, N. *et al.* ChAda-ViT : Channel Adaptive Attention for Joint Representation Learning of Heterogeneous Microscopy Images. Preprint at <https://doi.org/10.48550/arXiv.2311.15264> (2023).
47. Kraus, O. *et al.* Masked Autoencoders for Microscopy are Scalable Learners of Cellular Biology. Preprint at <https://doi.org/10.48550/arXiv.2404.10242> (2024).
48. Ando, D. M., McLean, C. Y. & Berndl, M. Improving Phenotypic Measurements in High-Content Imaging Screens. 161422 Preprint at <https://doi.org/10.1101/161422> (2017).
49. Lopez, R., Regier, J., Cole, M. B., Jordan, M. I. & Yosef, N. Deep generative modeling for single-cell transcriptomics. *Nat. Methods* **15**, 1053–1058 (2018).
50. Korsunsky, I. *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
51. Hie, B., Bryson, B. & Berger, B. Efficient integration of heterogeneous single-cell transcriptomes using Scanorama. *Nat. Biotechnol.* **37**, 685–691 (2019).
52. Yan, C. *et al.* Triple-effect correction for Cell Painting data with contrastive and domain-adversarial learning. 2025.01.13.632874 Preprint at <https://doi.org/10.1101/2025.01.13.632874> (2025).
53. Gabriel, W. *et al.* Weakly supervised cross-modal learning in high-content screening. Preprint at <https://doi.org/10.48550/arXiv.2311.04678> (2023).
54. Kessy, A., Lewin, A. & Strimmer, K. Optimal whitening and decorrelation. *Am. Stat.* **72**, 309–314 (2018).

55. LIT-PCBA: An Unbiased Data Set for Machine Learning and Virtual Screening | Journal of Chemical Information and Modeling.
<https://pubs.acs.org/doi/10.1021/acs.jcim.0c00155>.
56. Morgan, H. L. The Generation of a Unique Machine Description for Chemical Structures—A Technique Developed at Chemical Abstracts Service. *J. Chem. Doc.* **5**, 107–113 (1965).
57. Bajusz, D., Rácz, A. & Héberger, K. Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations? *J. Cheminformatics* **7**, 20 (2015).
58. Healy, J. & McInnes, L. Uniform manifold approximation and projection. *Nat. Rev. Methods Primer* **4**, 1–15 (2024).
59. Bemis, G. W. & Murcko, M. A. The Properties of Known Drugs. 1. Molecular Frameworks. *J. Med. Chem.* **39**, 2887–2893 (1996).
60. Guha, R. Exploring Uncharted Territories – Predicting Activity Cliffs in Structure–Activity Landscapes. *J. Chem. Inf. Model.* **52**, 2181–2191 (2012).
61. Zhang, Z., Zhao, B., Xie, A., Bian, Y. & Zhou, S. Activity Cliff Prediction: Dataset and Benchmark. Preprint at <https://doi.org/10.48550/arXiv.2302.07541> (2023).
62. Lo Surdo, P. *et al.* SIGNOR 3.0, the SIGNaling network open resource 3.0: 2022 update. *Nucleic Acids Res.* **51**, D631–D637 (2023).
63. Devaiah, B. N. *et al.* MYC protein stability is negatively regulated by BRD4. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 13457–13467 (2020).
64. Kotekar, A., Singh, A. K. & Devaiah, B. N. BRD4 and MYC: Power couple in Transcription

- and Disease. *FEBS J.* **290**, 4820–4842 (2023).
65. Ruscone, M. *et al.* NeKo: a tool for automatic network construction from prior knowledg. Preprint at <https://doi.org/10.1101/2024.10.14.618311> (2024).
 66. He, K. *et al.* Masked Autoencoders Are Scalable Vision Learners. in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* 15979–15988 (2022). doi:10.1109/CVPR52688.2022.01553.
 67. Kalinin, A. A. *et al.* A versatile information retrieval framework for evaluating profile strength and similarity. 2024.04.01.587631 Preprint at <https://doi.org/10.1101/2024.04.01.587631> (2024).
 68. Lopes, J. C. D., dos Santos, F. M., Martins-José, A., Augustyns, K. & De Winter, H. The power metric: a new statistically robust enrichment-type metric for virtual screening applications with early recovery capability. *J. Cheminformatics* **9**, 7 (2017).
 69. Gaulton, A. *et al.* ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res.* **40**, D1100–D1107 (2012).
 70. mxfly14. mxfly14/2025_sanchez_phenoseeker: Publication DOI. Zenodo <https://doi.org/10.5281/zenodo.17914525> (2025).

Acknowledgements

This work has received support under the program « Investissements d’Avenir » launched by the French Government and implemented by the ANR, with the references: ANR-10-LABX-54 MEMO LIFE ANR-11-IDEX-0001-02 PSL* Research University, MS was funded by Iktos and ANRT under the CIFRE program. This work was granted access to the HPC resources of IDRIS under the allocation 2020-AD011011495 made by GENCI. We thank Gabriel Watkinson for

helping with the JUMP-CP dataset, the computing service of IBENS and Mary-Ann Letellier for editing the manuscript.

Author Contributions

MS designed the computational framework and performed all the analyses. NB, IB, EC and IS helped for the analyses. EdN provided screening data and expertise. HT and GB contributed to the discussion. AG and LC conceived the project and supervised the work. MS and AG wrote the manuscript. All authors revised the manuscript.

Competing interests

The authors declare the following competing interests: MS and HT are employees at Iktos. All other authors declare no competing interests.

Figure titles/legends

Figure 1 Generating robust Phenotypic Profiles from JUMP-CP Cell Painting images. (a) Typical pipeline steps from image acquisition to normalized well profiles. At each of these steps (feature extraction, aggregation and normalization), various methods can be chosen (b) We performed a sweep of possible pipelines using control well replicate samples from JUMP-CP. For each possible pipeline (represented on this plot by a single point), after processing all samples, we checked how well the pipeline retrieved sample replicates from the same perturbation (X axis, the higher the better) and could confound replicate samples of negative control whatever the experimental plate (Y axis, the lower the better) by computing two mean Average Precisions (mAP) values. The former evaluates the preservation of biological signals while the latter assesses effectiveness in correcting batch effects. (c) Scheme of the selected processing of Cell Painting images (d) UMAP⁵⁵ Visualization of Well Profiles Across Laboratories. UMAP⁵⁵ plots display well phenotypic profiles from 27 plates (only control wells) across 9 laboratories, before (left) and after (right) normalization, showing improved integration of phenotypic profiles.

Figure 2 Library selection through robust JUMP-CP profiles boost screening hit rates. (a) Schematic of our selection method where the top k% closest JUMP-CP neighbors (dotted sphere) of a screen positive control (red dot) are selected for screening. To evaluate the capability of this selection process to increase screening hit rate, we applied it to past screens and thus computed the fraction of hits that fall into this selection (orange dots). (b)(c)(d) Evaluation of our method across diverse screens and sources. Mean (green bars) and maximum (blue bars) normalized Enrichment Factor (nEF) are displayed, comparing our method with the original selection (purple bar). There are 49 screens from ChEMBL (b), 5 targets from Lit-PCBA (c), and 16 screens from the Curie Screening Platform (d). nEF is determined based on the top 5% compounds with the highest phenotypic similarity to the positive control.

Figure 3 Selection through similar Cell Painting profiles yield a variety of compound structures (a) For illustration, the Mean Tanimoto structure similarity between the positive control of a screen from Institut Curie and three groups of compounds: all compounds tested in the screen (blue), the top 5% phenotypically similar compounds (green), and the top 5% structurally similar compounds (red). The same computation is then performed for a screen from ChEMBL and the whole JUMP-CP using each compound as a positive control. Means (green triangle), medians (yellow line), quartiles are computed from : 16 samples (one per screen) for *Curie Institute*, 49 samples (one per screen) for *ChEMBL*, 10000 samples (one per compound) for *CP-JUMP* (b) On the left, UMAP⁵⁸ of the phenotypic profiles for compounds tested in the screen from Institut Curie. On the right, UMAP of the structural fingerprints for the same compounds. Positive control is shown in red; the 5% phenotypically closest compounds are shown in blue. Arrows point to illustration compounds from (c) and (d). (c) Examples of selected compounds (close to the positive control in phenotypic space but distant in structural space). (d) Examples of pairs of compounds that are close in structural space but distant in phenotypic space. For each compound, chemical structures and Cell Painting images are shown. Scale bars = 50 μ m.

Figure 4 **Systematic identification of activity cliffs. (a)** JUMP-CP compounds were grouped by Murcko scaffold. **(b)** For each group of compounds, a hierarchical clustering based on their phenotypic profiles was performed. Phenotypic (cosine, red) and Structural (Tanimoto, blue) similarity within and between these clusters were computed. **(c)** Intra- versus inter-cluster phenotypic similarity. 81 scaffold-based groups exposing a substantially higher intra- than inter-cluster phenotypic similarity are highlighted in red. **(d)** Intra- versus inter-cluster structural similarities. The 81 scaffolds highlighted in red on panel c are highlighted in blue here. The green squares represent the example of the chemical series illustrated in (b).

Figure 5 **(a) Example of Phenotypic Hierarchical Clustering of Compounds with the same Murcko Scaffold.** The distance matrix indicates the cosine similarities between phenotypes induced by these compounds and both the dendrogram and the matrix highlight two main phenotypic clusters: the first one is a group of very similar phenotypes while the second displays a variety of sub-phenotypes as highlighted by the matrix diagonal. The last row of the matrix indicates the distances to the negative control (DMSO). **(b)** Further investigation of the images and the structures shows that the first cluster includes 3 apparently toxic compounds each featuring an aliphatic chain highlighted in green. **(c)** The second cluster presents *compounds that induce different phenotypes*. Highlighted in green is the common protonated nitrogen atom. In purple the different substitution that could explain the different phenotypes induced. **(d)** Examples of negative control with DMSO, a mock treatment. Scale bars = 50 μm .

Figure 6 **Example of a screen where compounds reproducing the positive control phenotype hit different targets of the EGFR pathway (a)** UMAP⁵⁸ of phenotypic profiles of compounds from the ChEMBL1613808 screen, designed to identify *in cellulo* inhibitors of the EGFR pathway. Positive control (*Dezmapimod*) is shown in red, while compounds selected by our method are shown in blue. **(b)** Examples of selected compounds (SC) include known EGFR pathway inhibitors, acting through distinct mechanisms of action and targeting different genes involved in the pathway. **(c)** The EGFR signaling pathway extracted from the SIGNOR database. Targets of the selected compounds are highlighted in blue, illustrating their diverse role within the pathway and confirming why structural similarity between compounds is not always expected. Scale bars = 50 μm .

Figure 7 **Phenotypic similarities and dissimilarities of compounds targeting G2M pathway genes (a)** Violin plot of phenotypic similarity distributions: in green, similarities between all pairs within the 530 compounds active against the G2M pathway (280900 samples); in red, similarities between the same number of compound pairs not known to be active against the G2M pathway (280900 samples). **(b)** Cellular image (Phenotypes) and chemical structures of the four compounds. Scale bars = 50 μm . **(c)** Cosine similarity matrix showing the phenotypic similarity among the four compounds, each targeting different G2M pathway genes. Negative signs show opposite phenotypic profiles whereas positive signs show similar phenotypic profiles **(d)** Network of gene-gene interactions among G2M pathway genes targeted by JUMP-CP compounds. G2M pathway genes were obtained from MSigDB, interactions from OmniPath, and the network was constructed using NeKo. Red edges indicate inhibition, green edges indicate stimulation, and purple edges a bimodal effect.

Table 1: **Statistical Comparison of Phenotypic Similarities Between Pathway-Targeting and Random Compound Pairs**

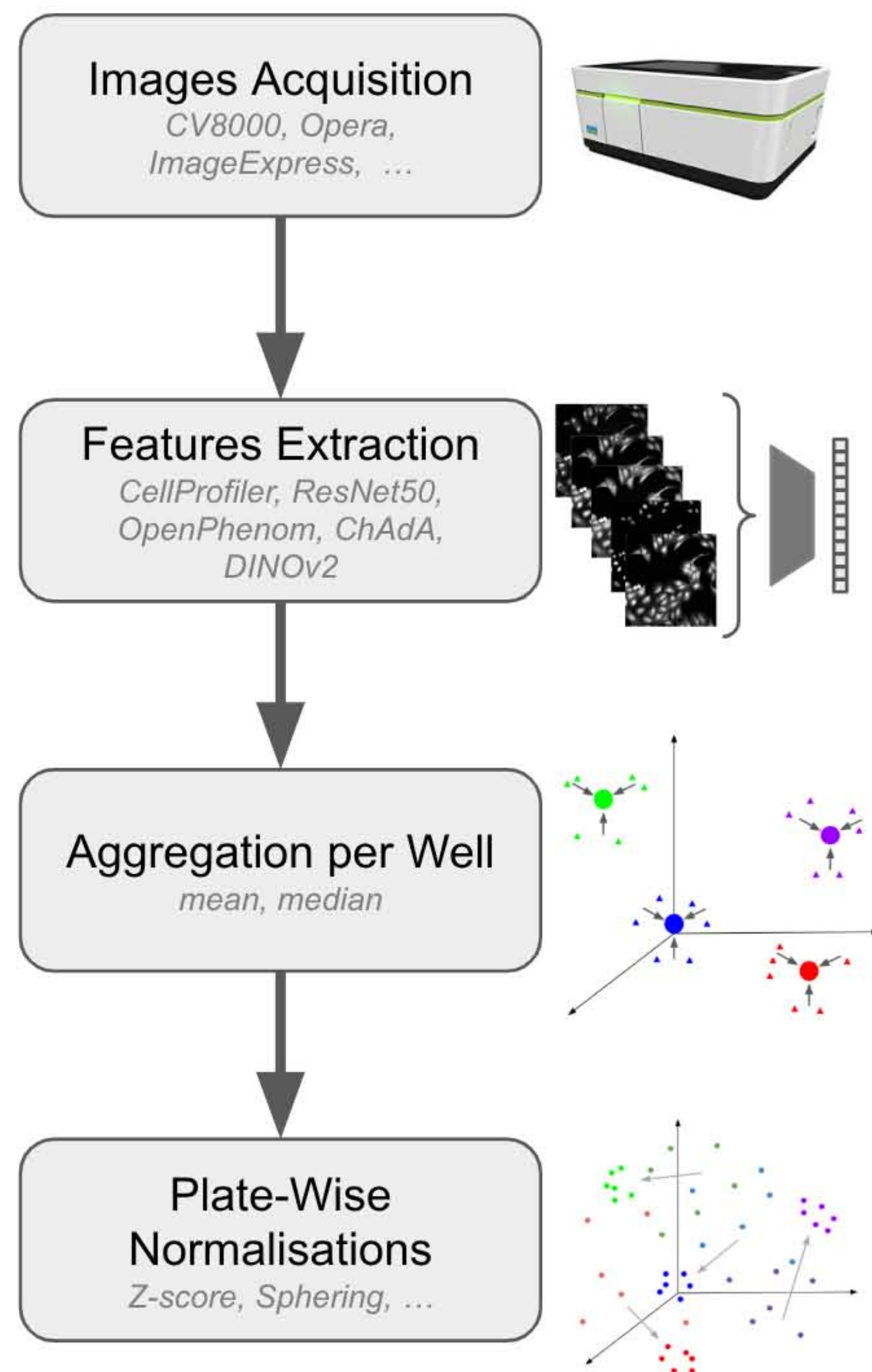
Editor Summary

Using DINOv2-derived batch-corrected phenotypic embeddings from 112k compounds, Phenoseeker enables structure-independent compound selection and systematic discovery of activity cliffs and pathway-consistent phenotypic convergence.

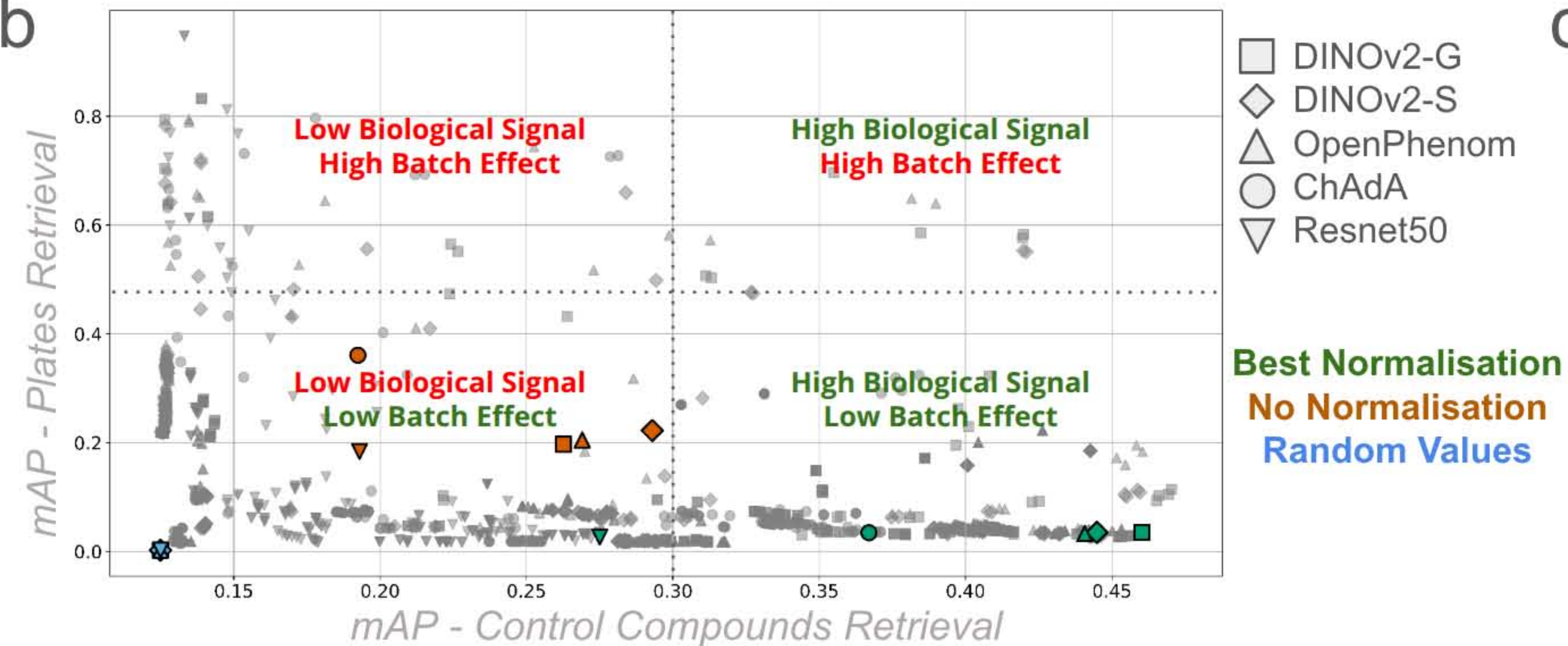
Peer Review Information

Communications Biology thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editors: Laura Rodríguez Pérez. A peer review file is available.

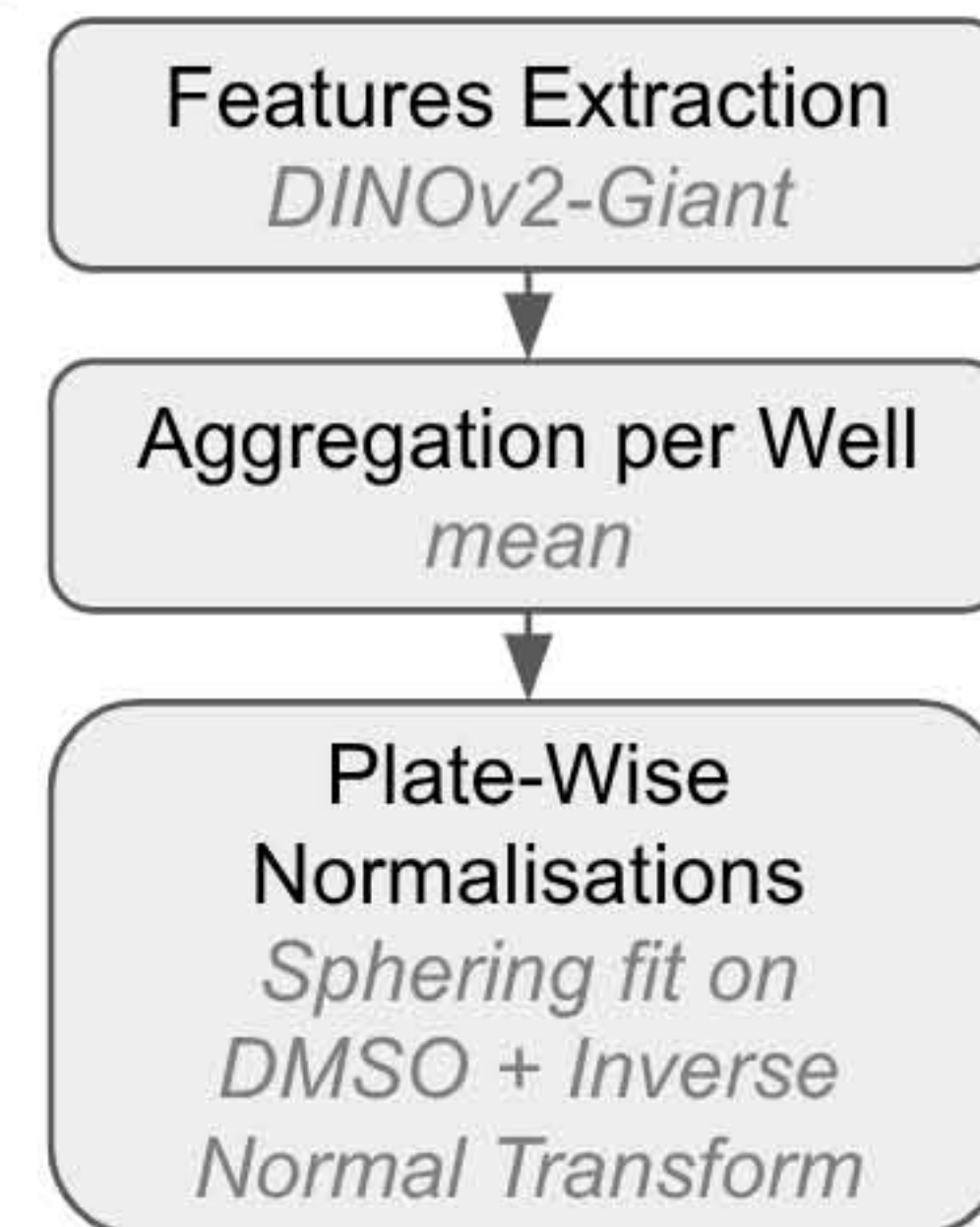
a



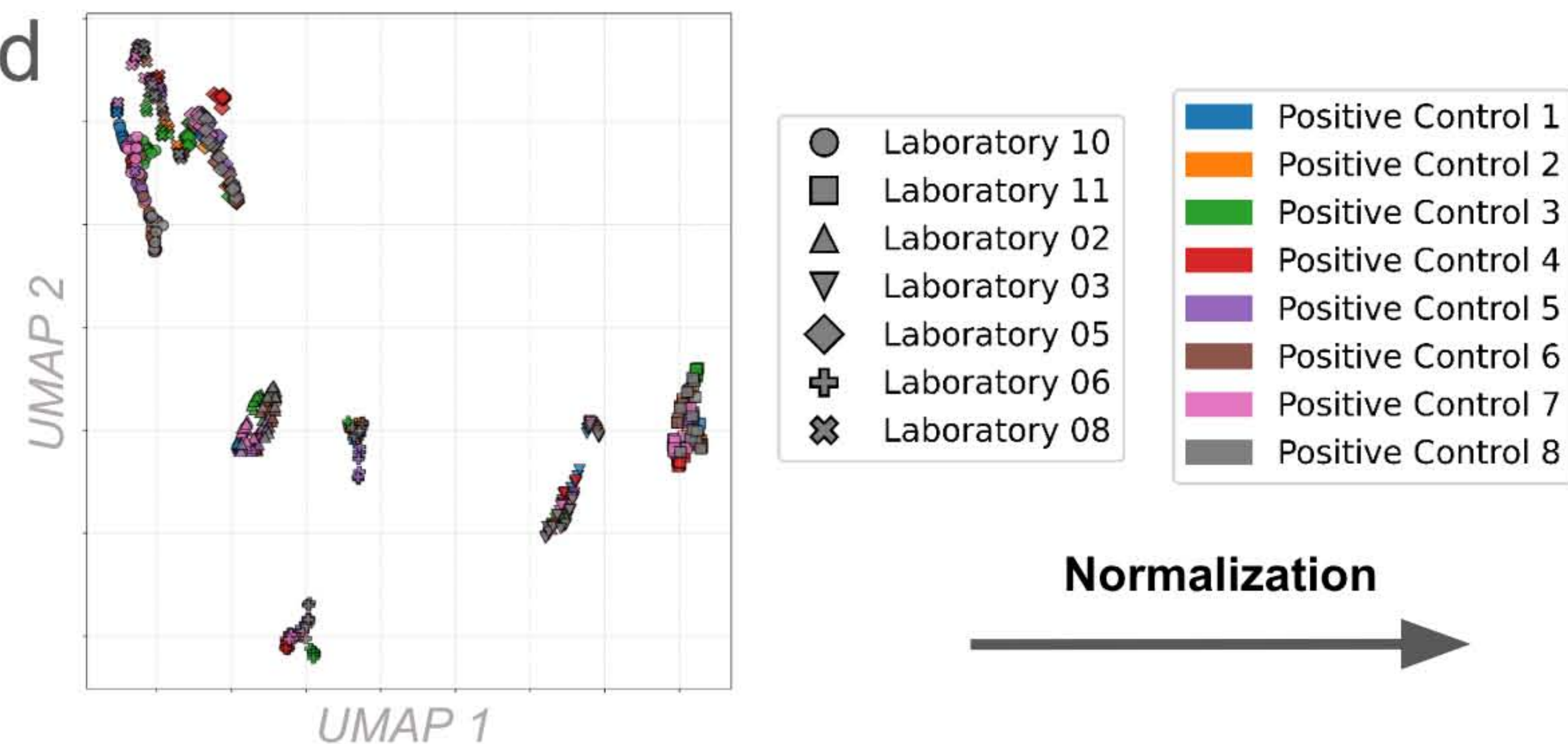
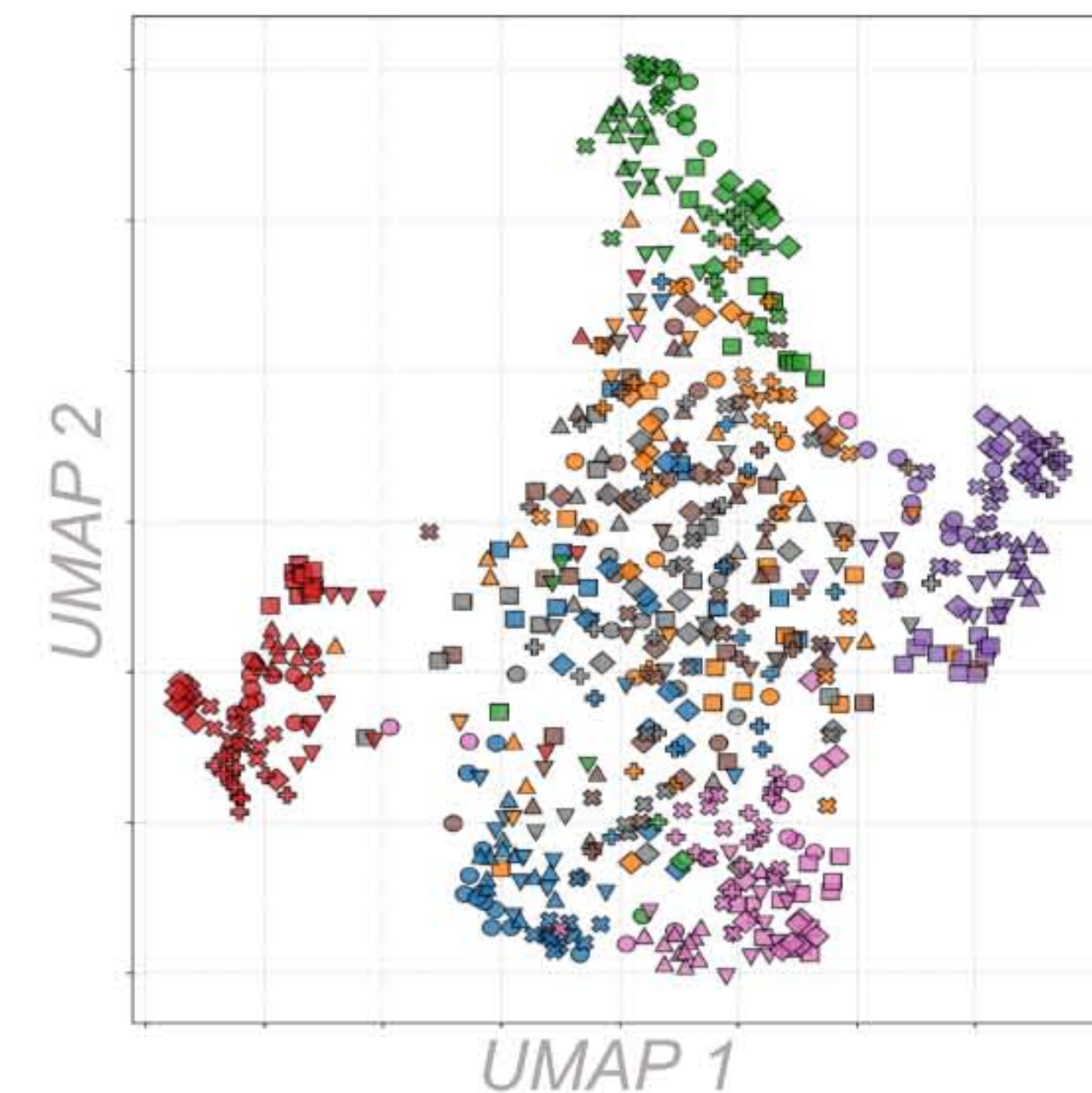
b



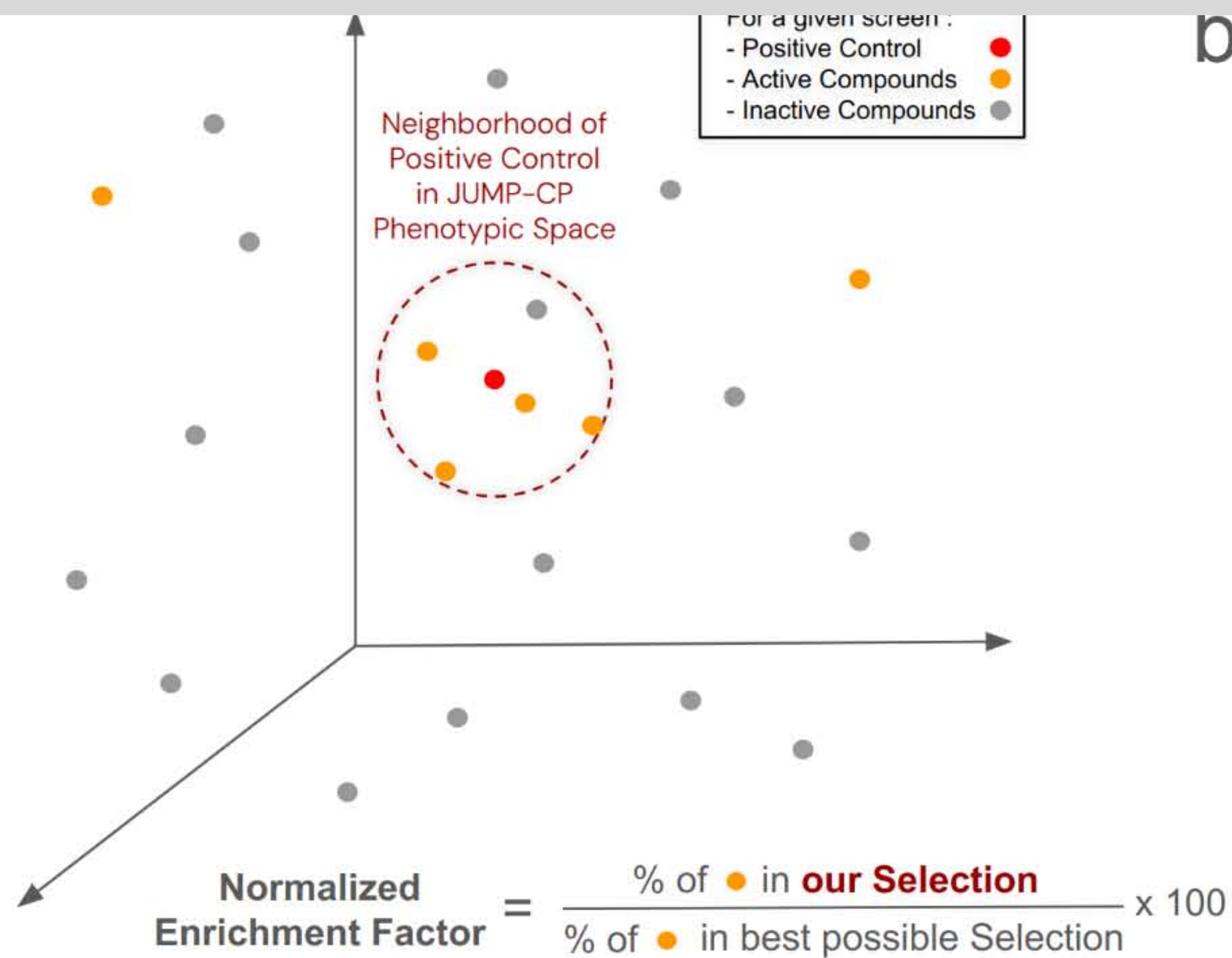
c



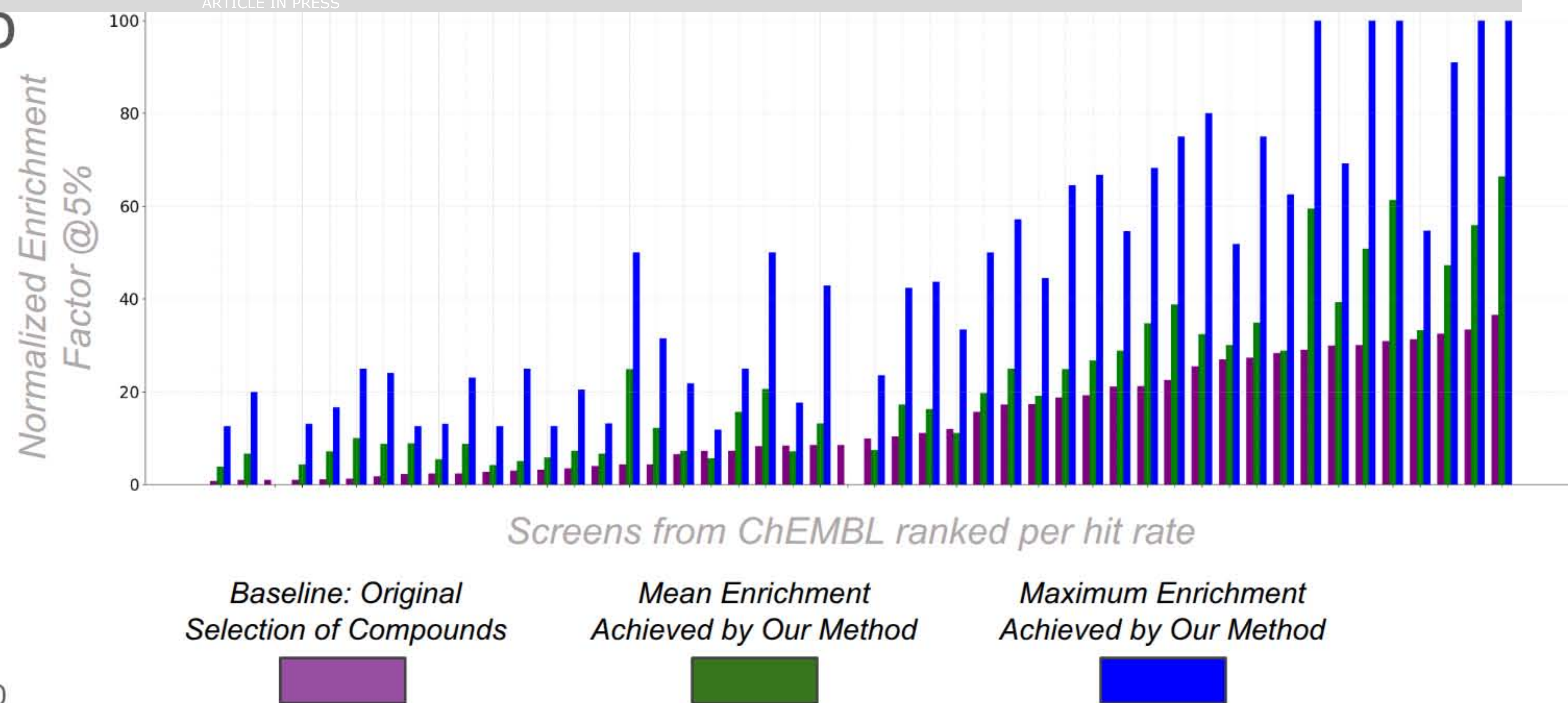
d



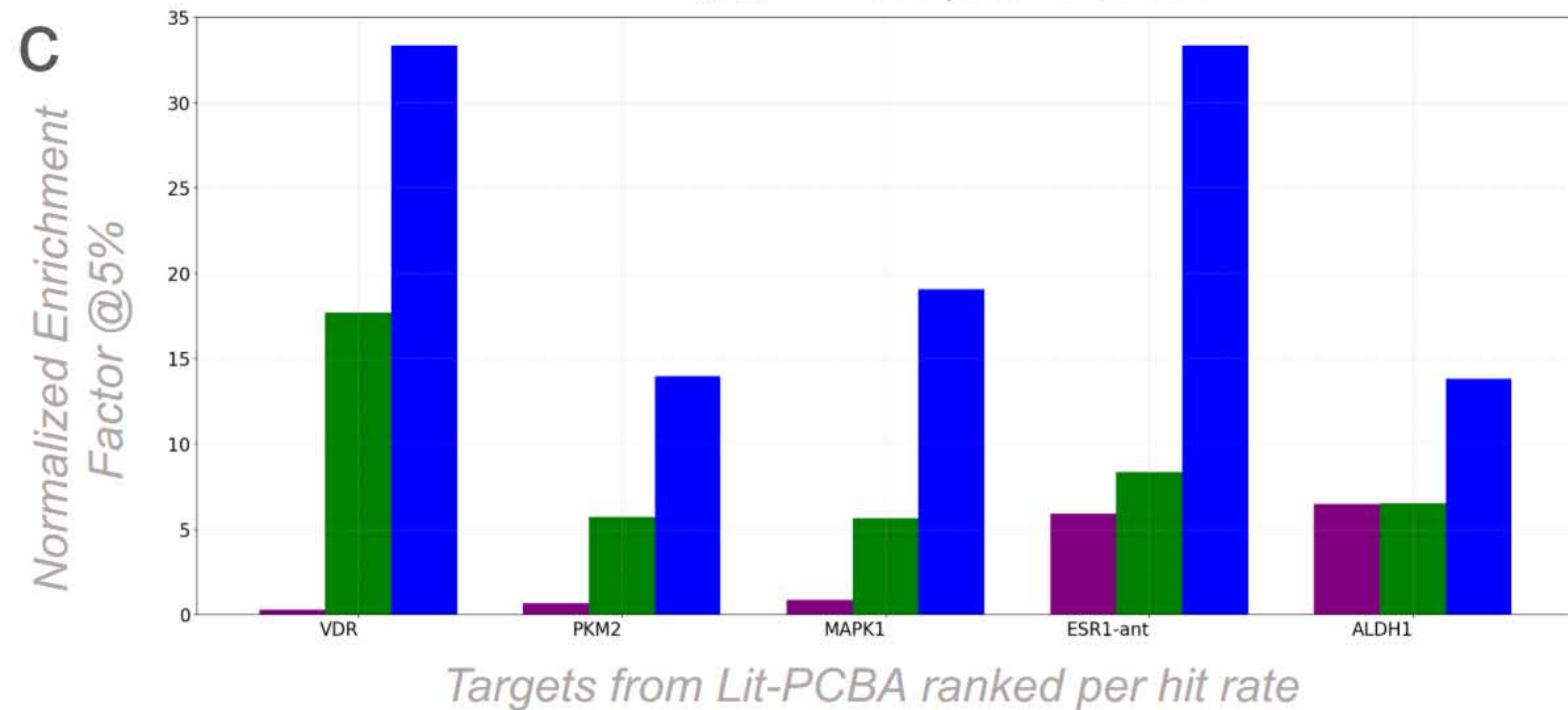
a



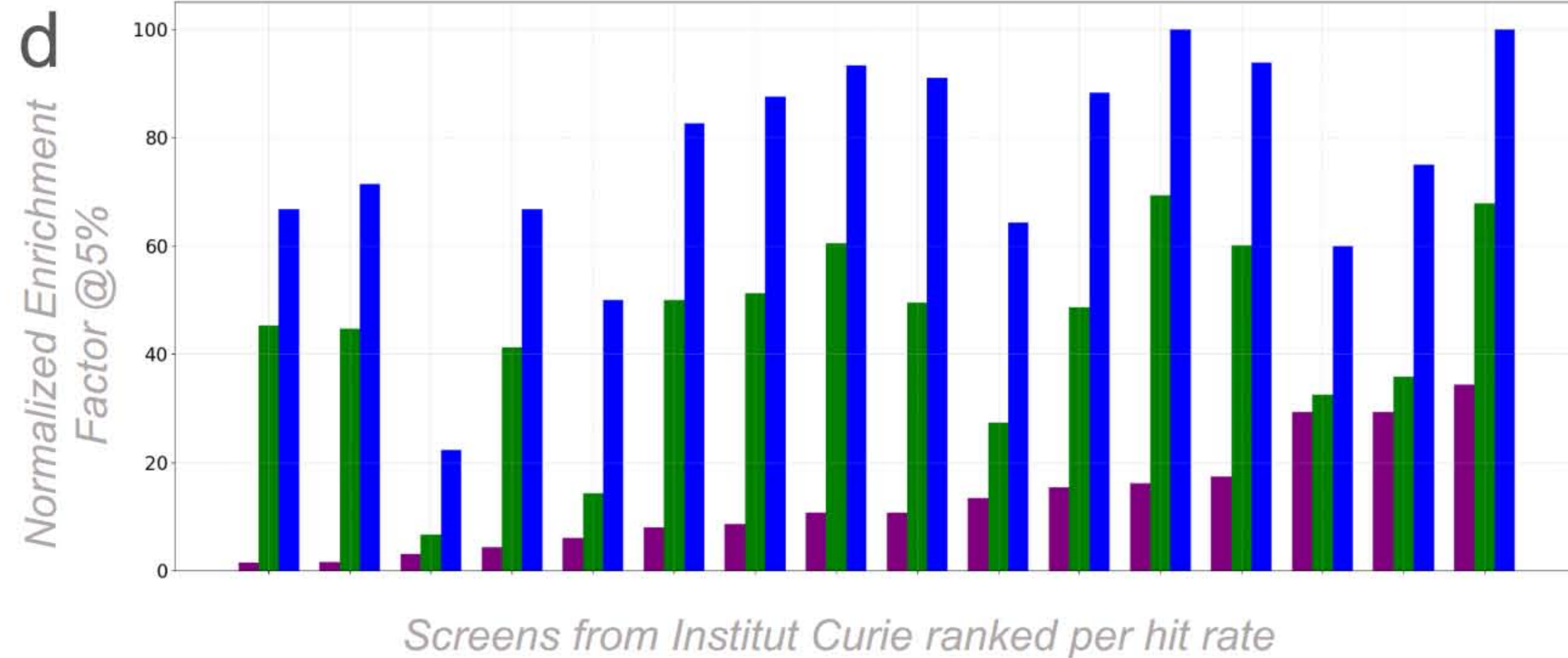
b

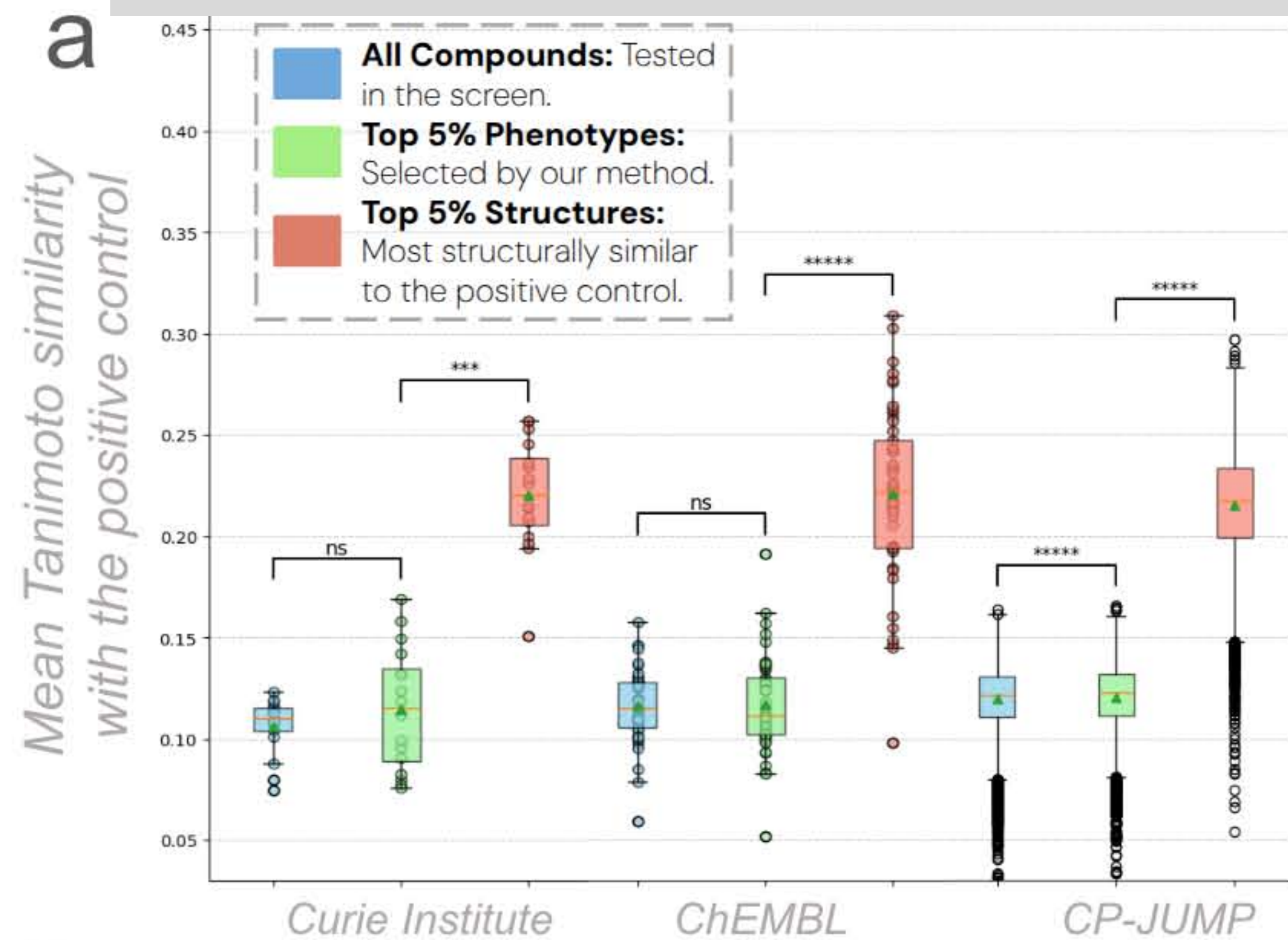
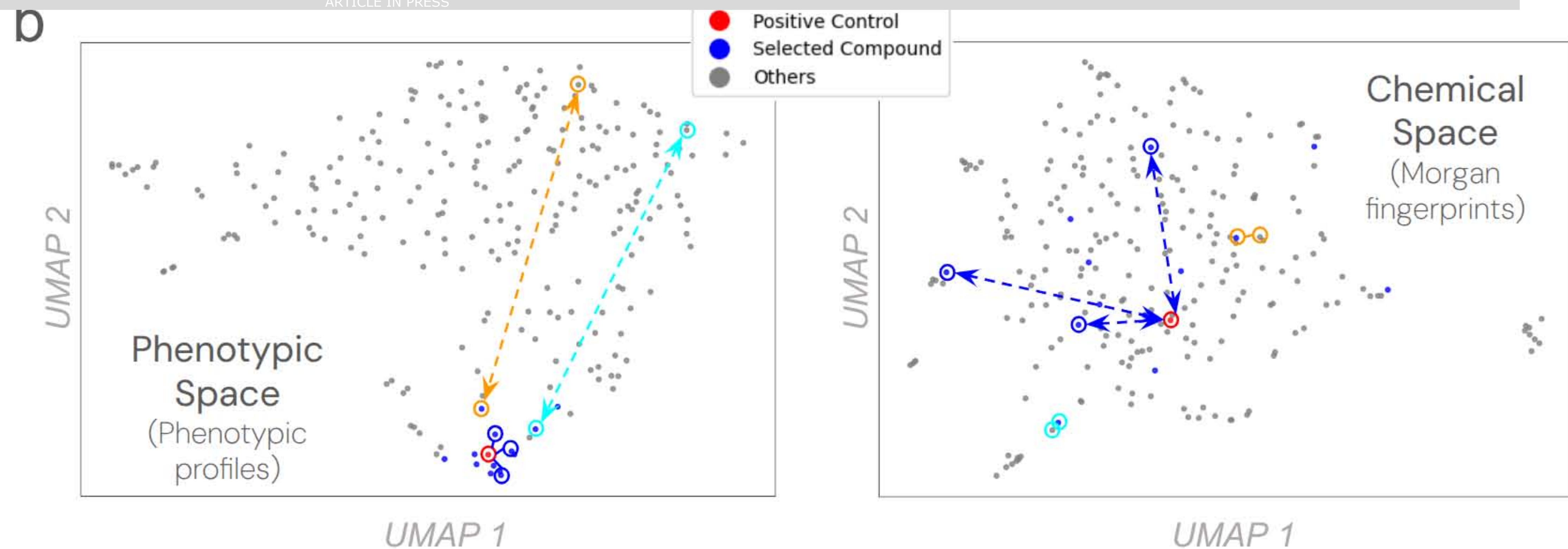
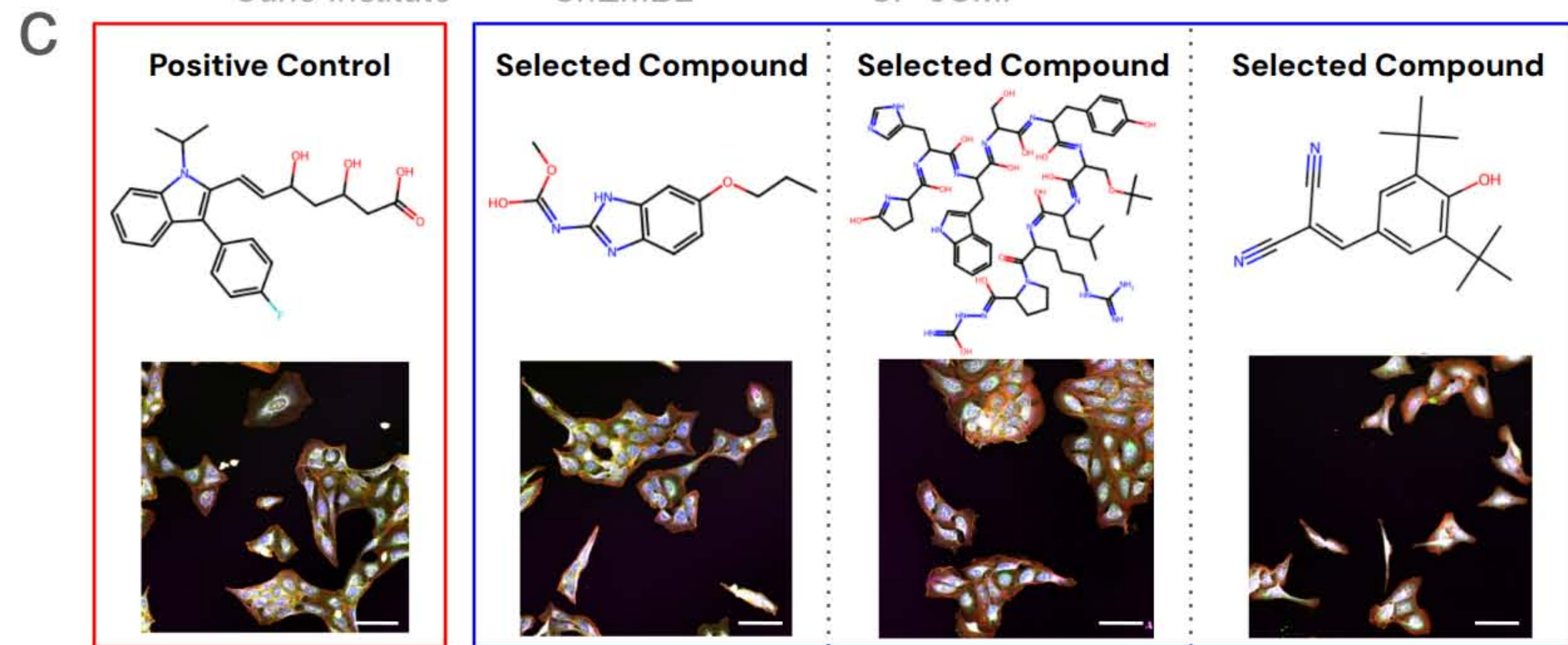
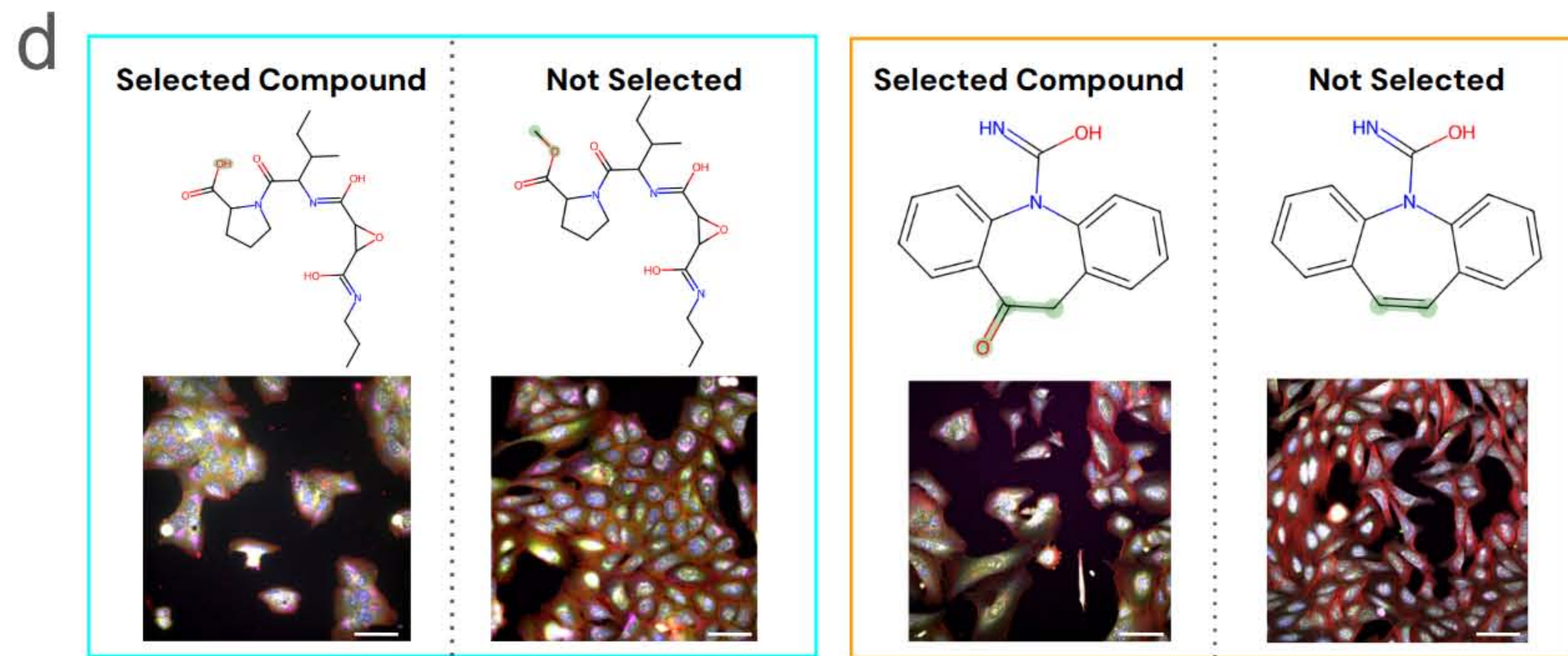


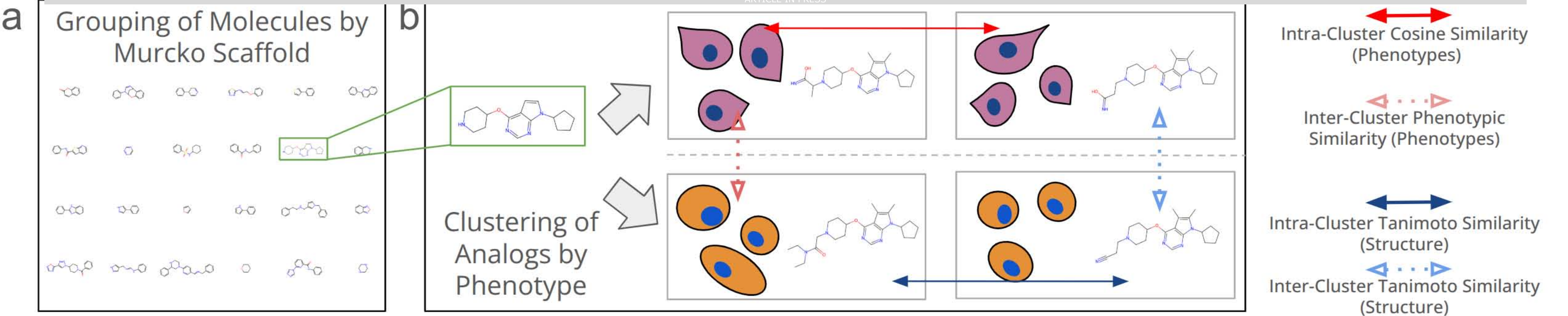
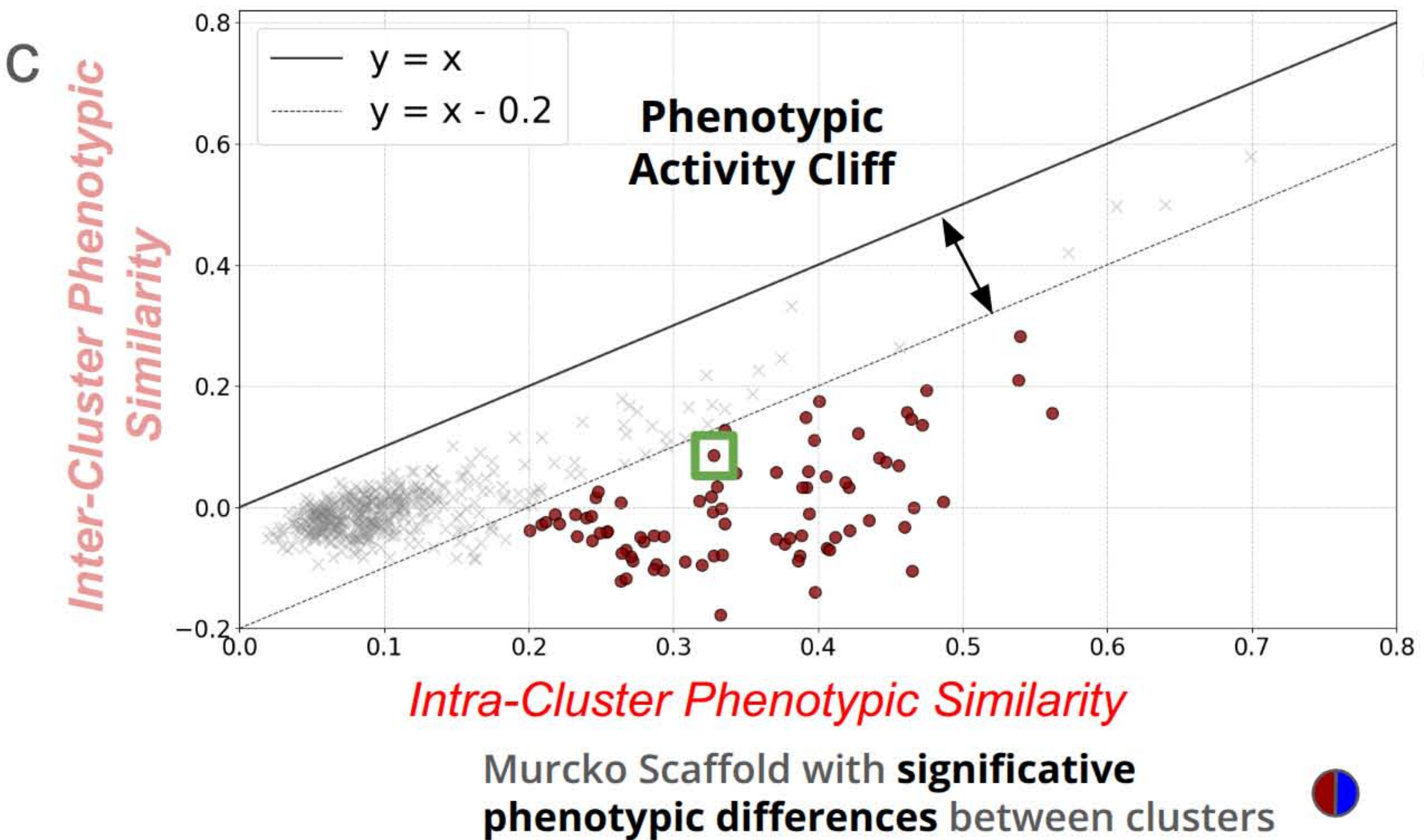
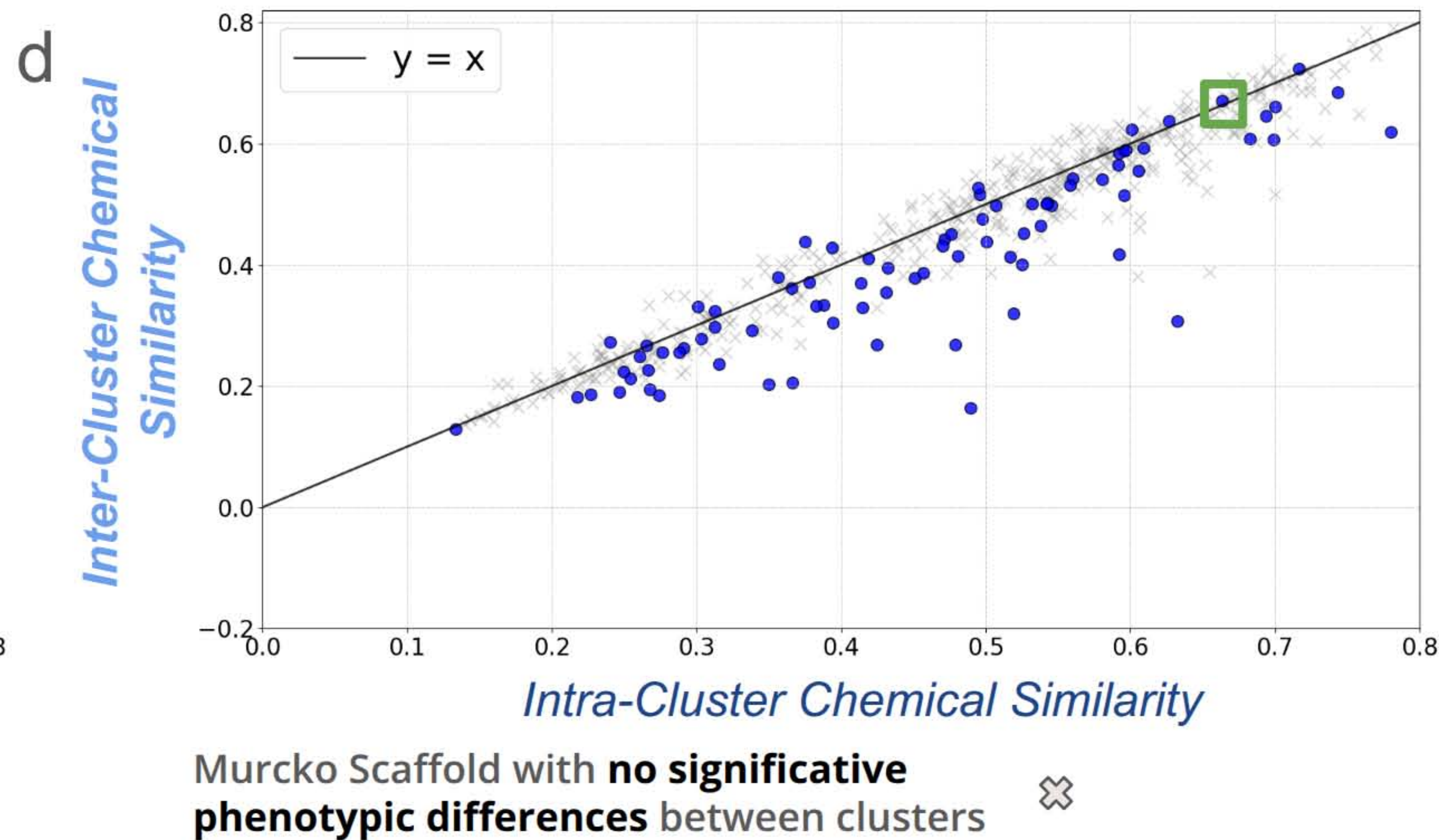
c

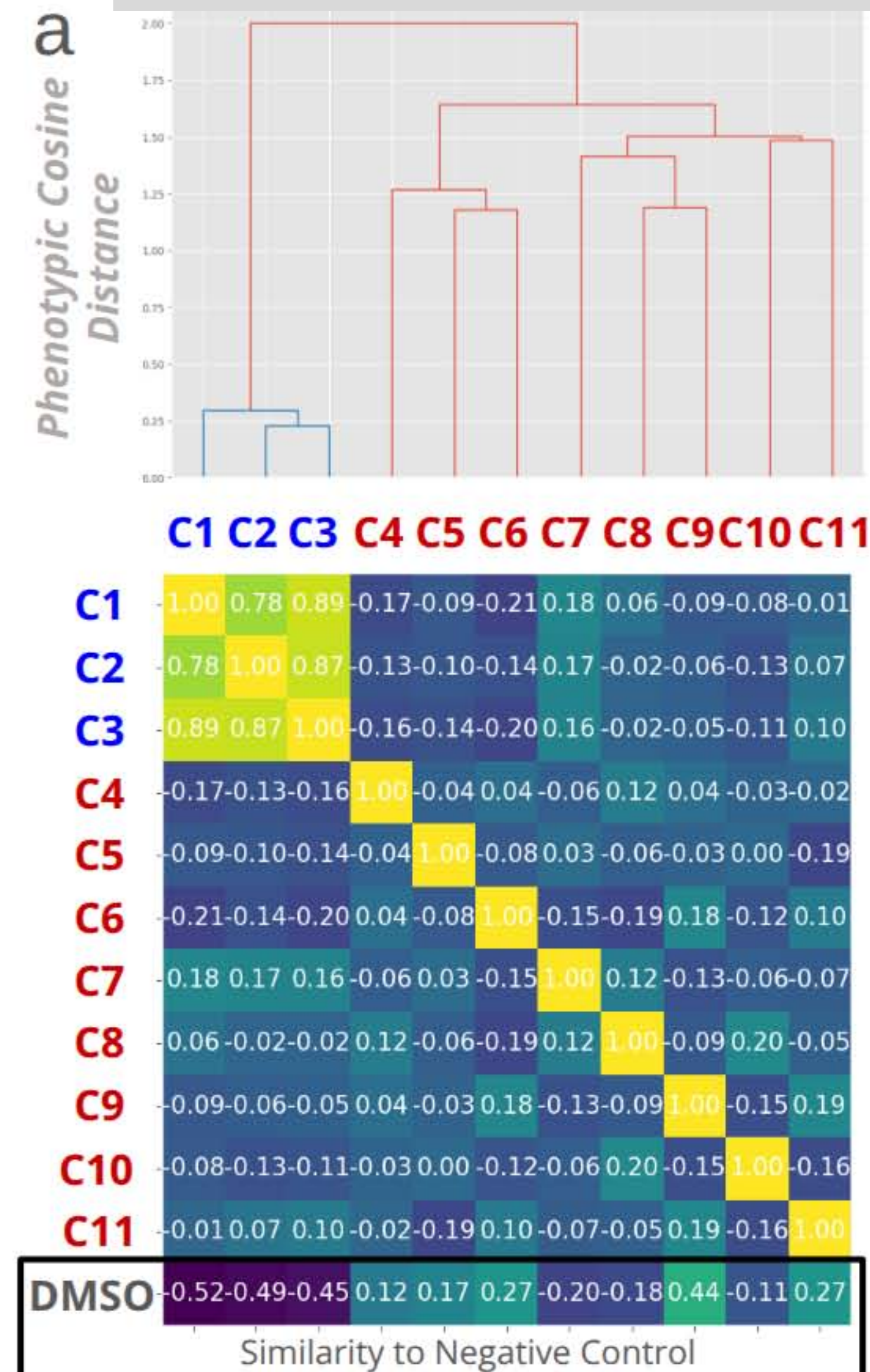
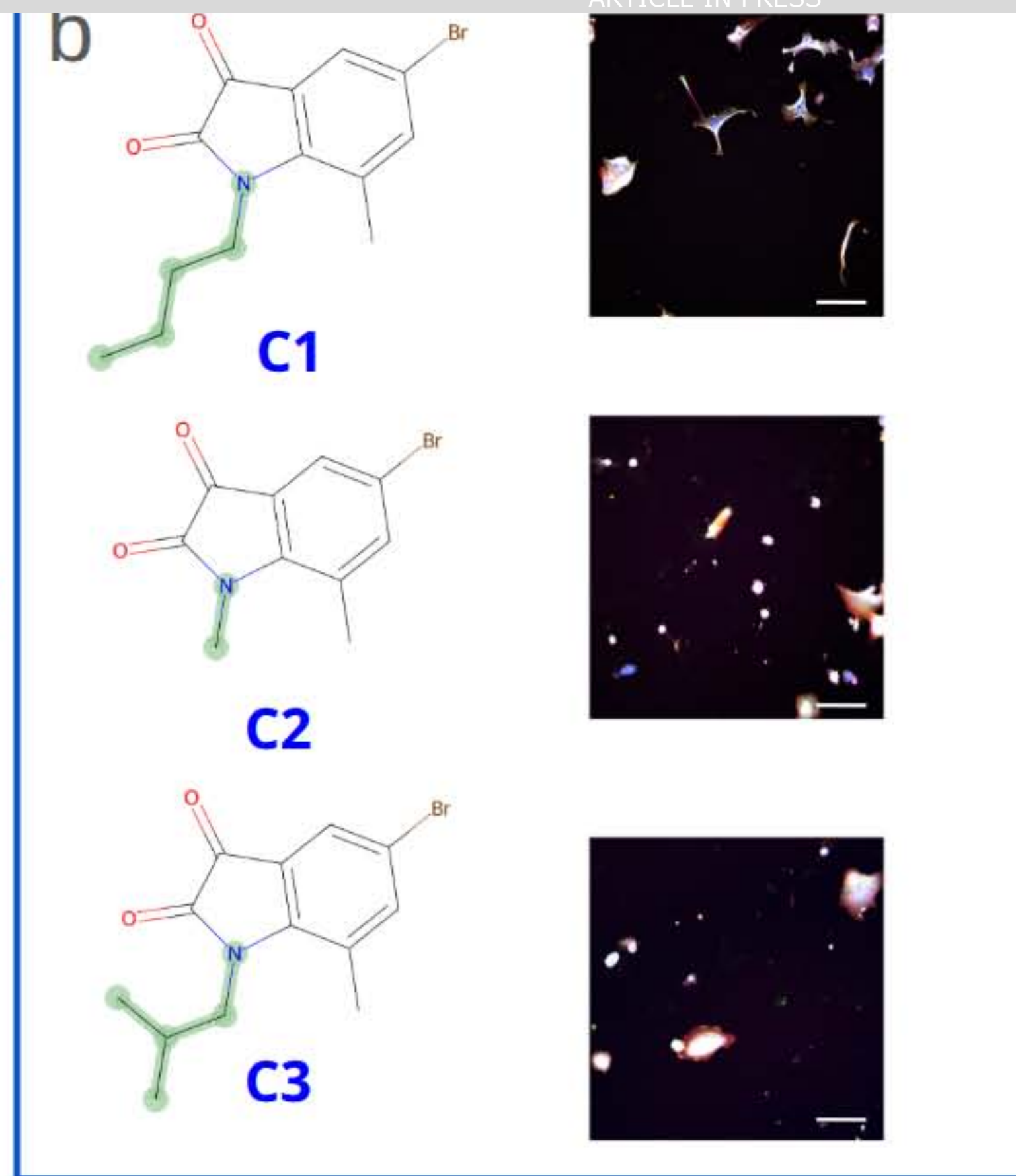
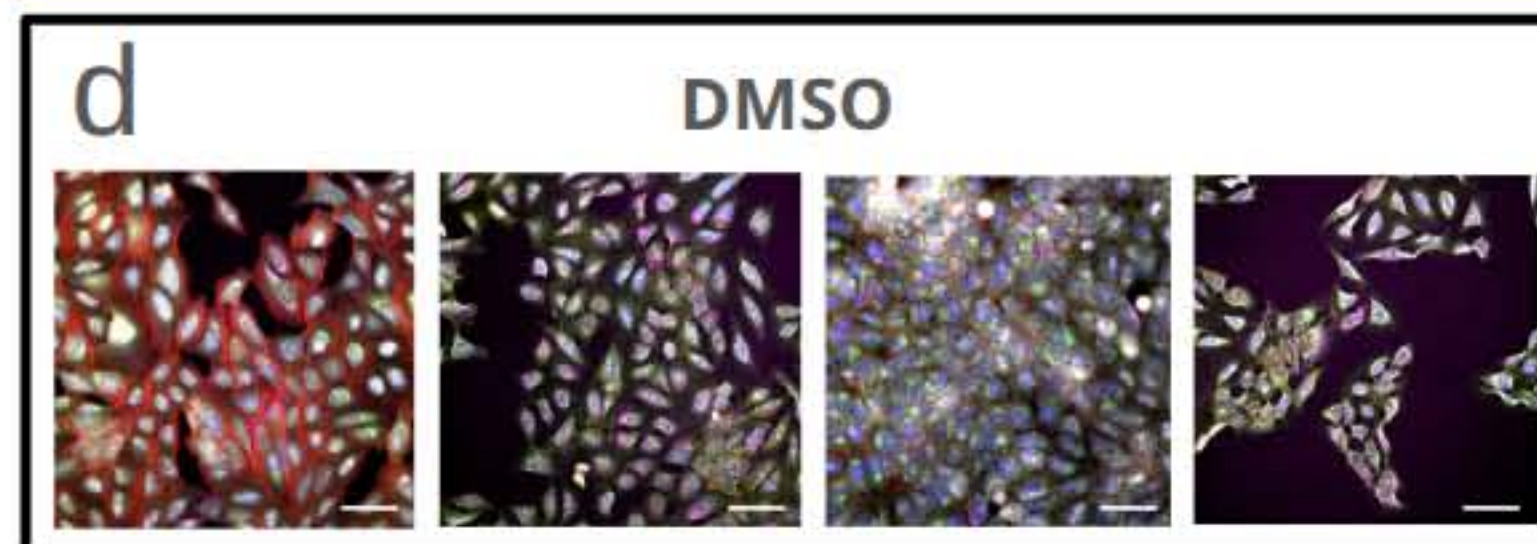
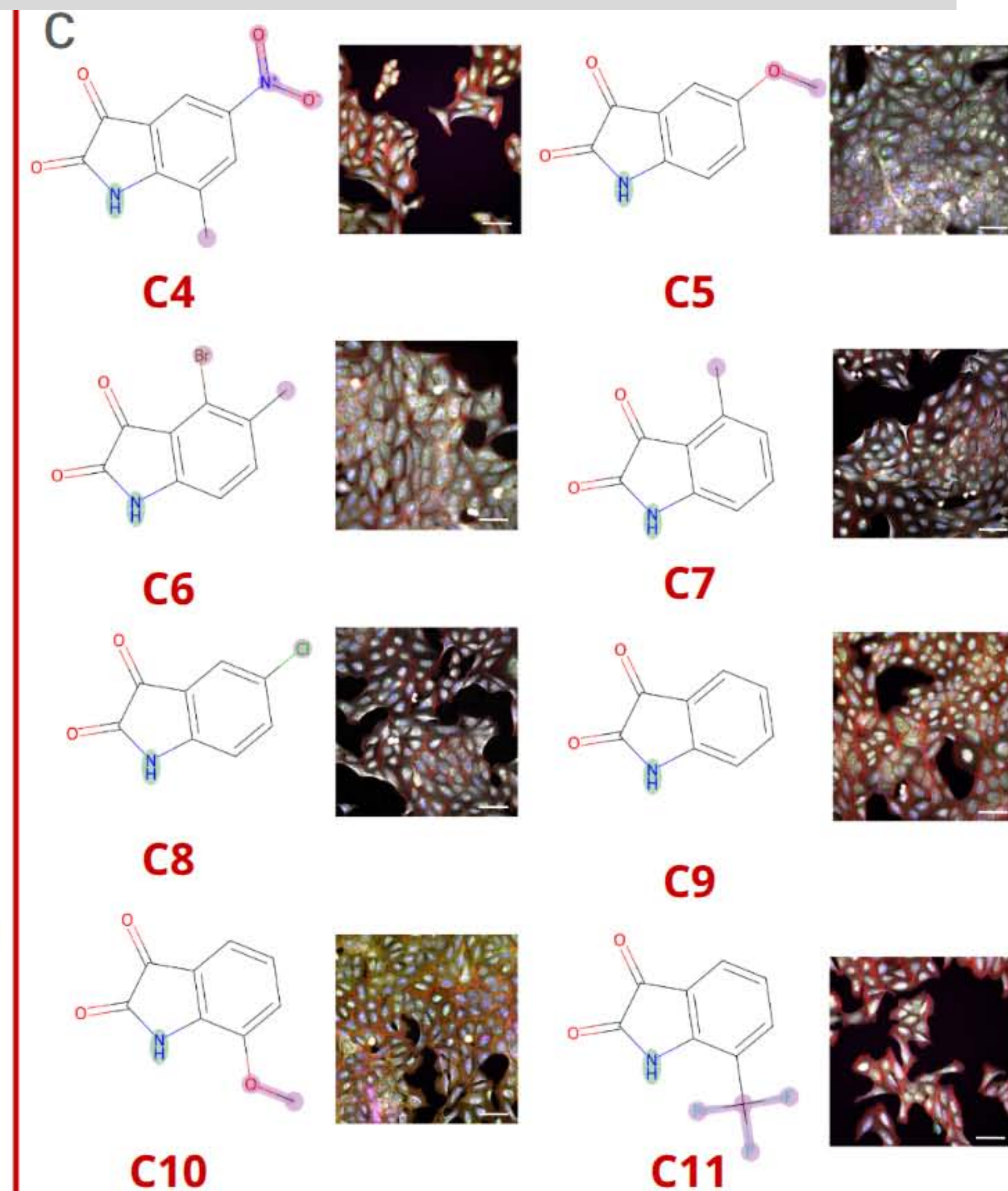


d



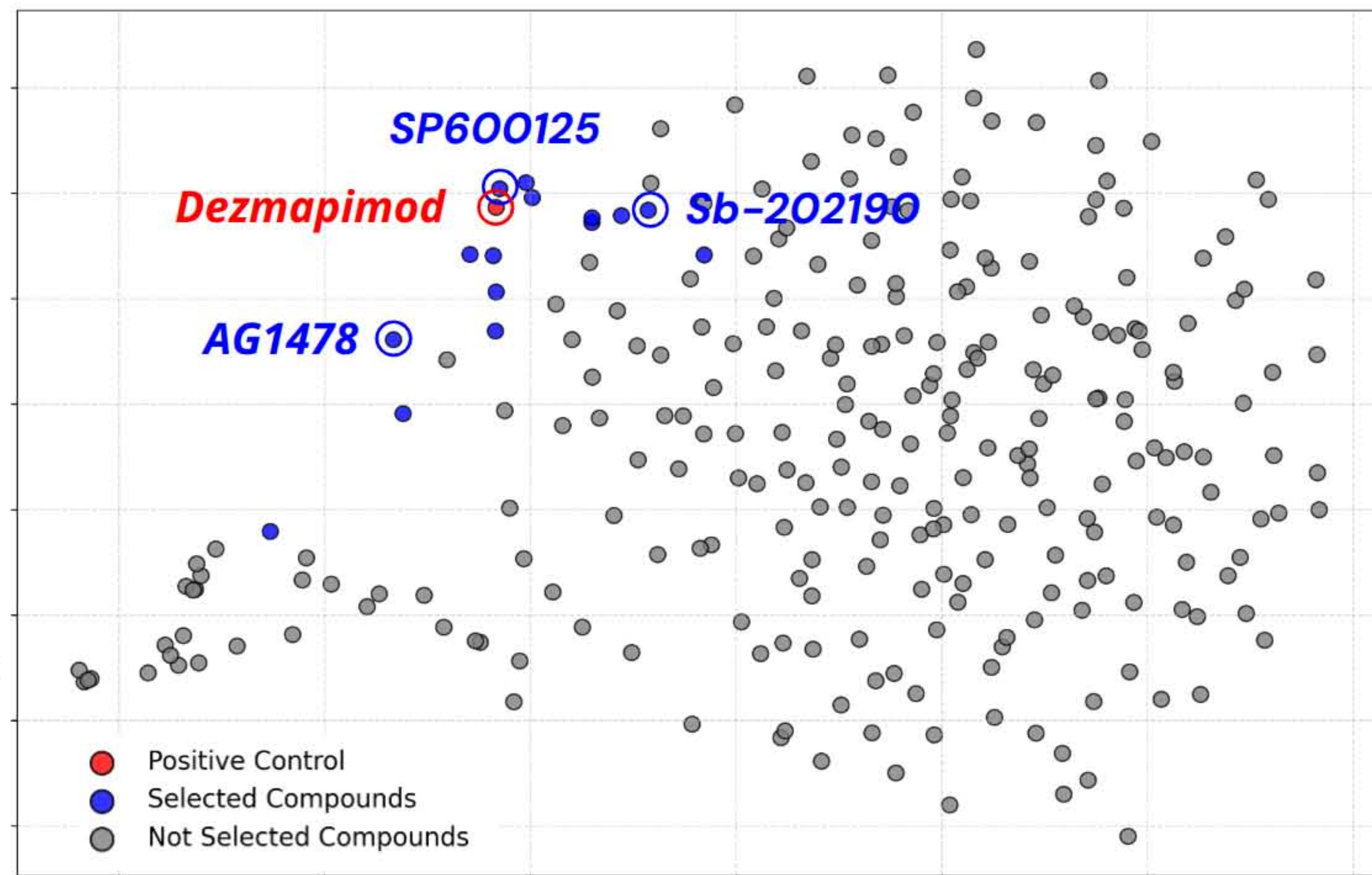
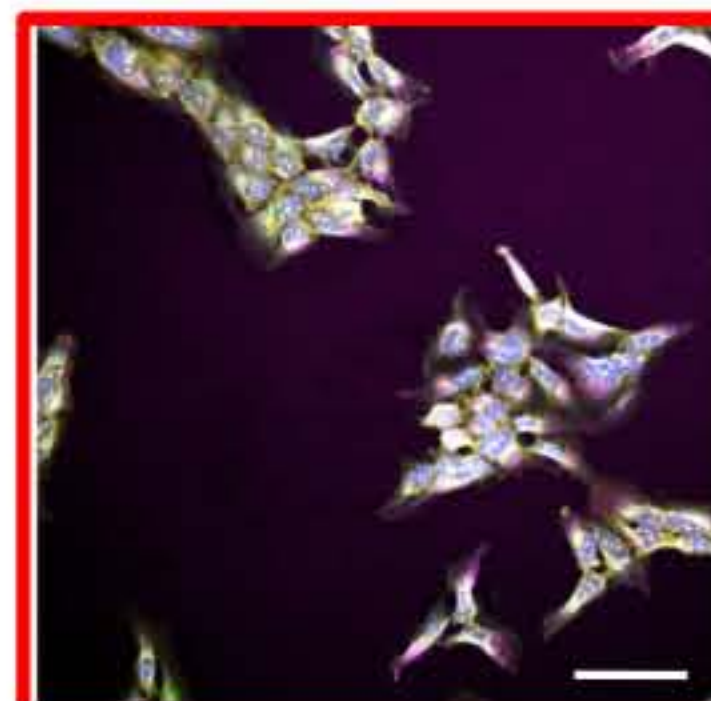
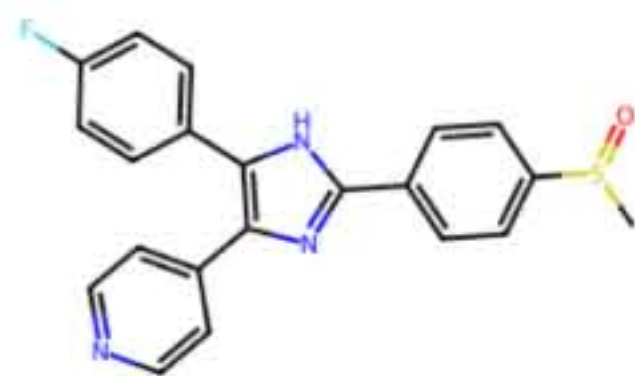






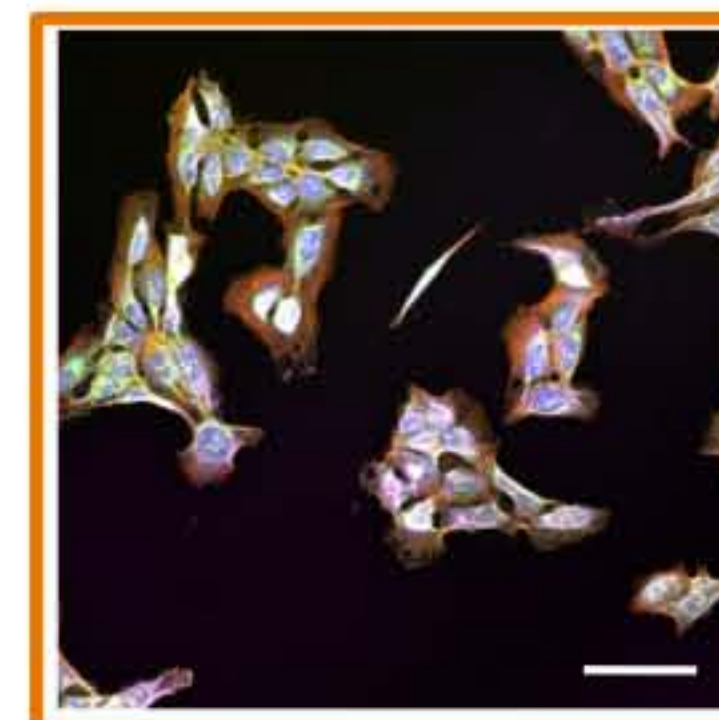
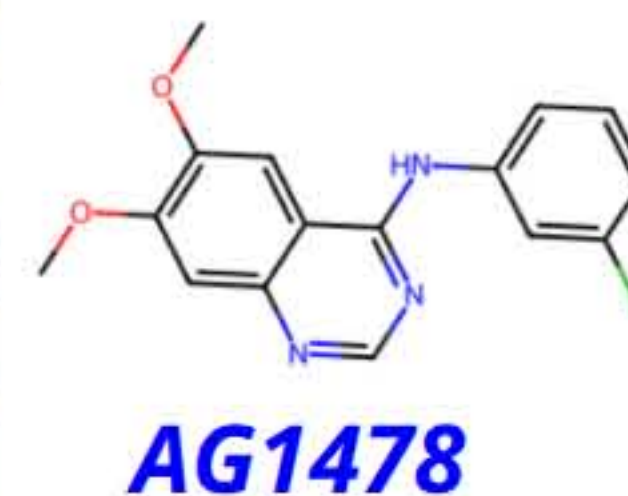
a

Phenotypic Profiles - UMAP 2

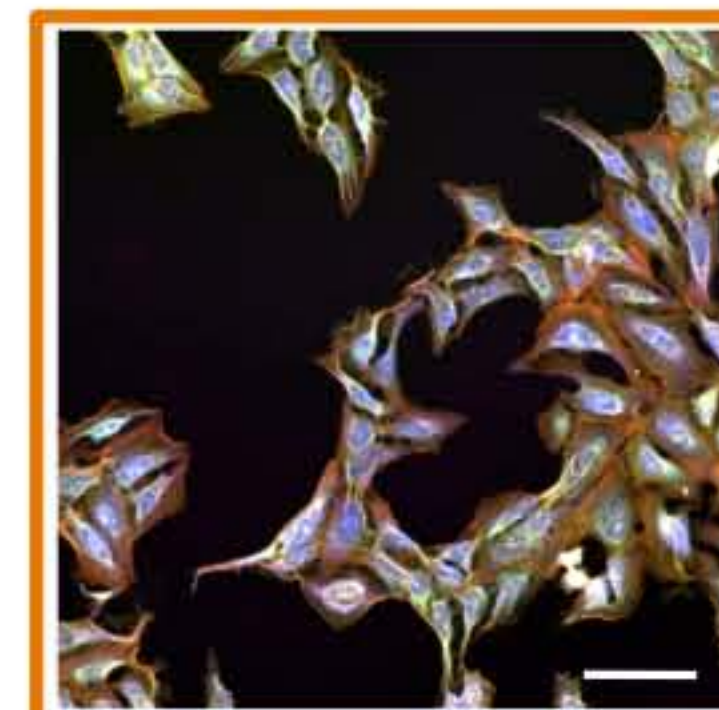
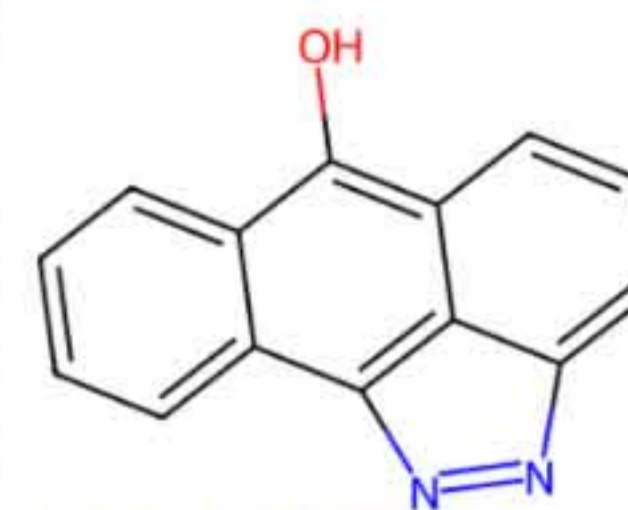


**Positive Control -
Dezmapimod** : a non
specific inhibitor of
mitogen-activated protein
kinases (MAPK).

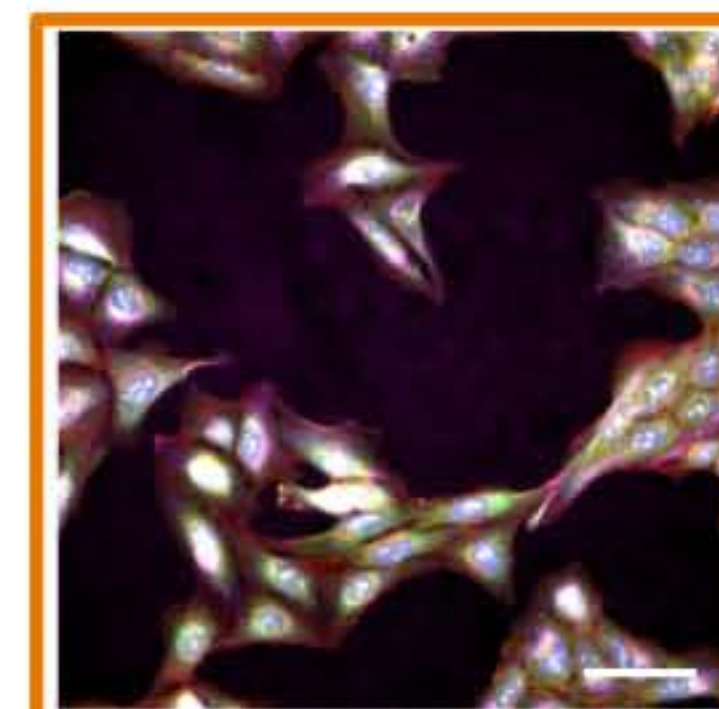
b



AG1478

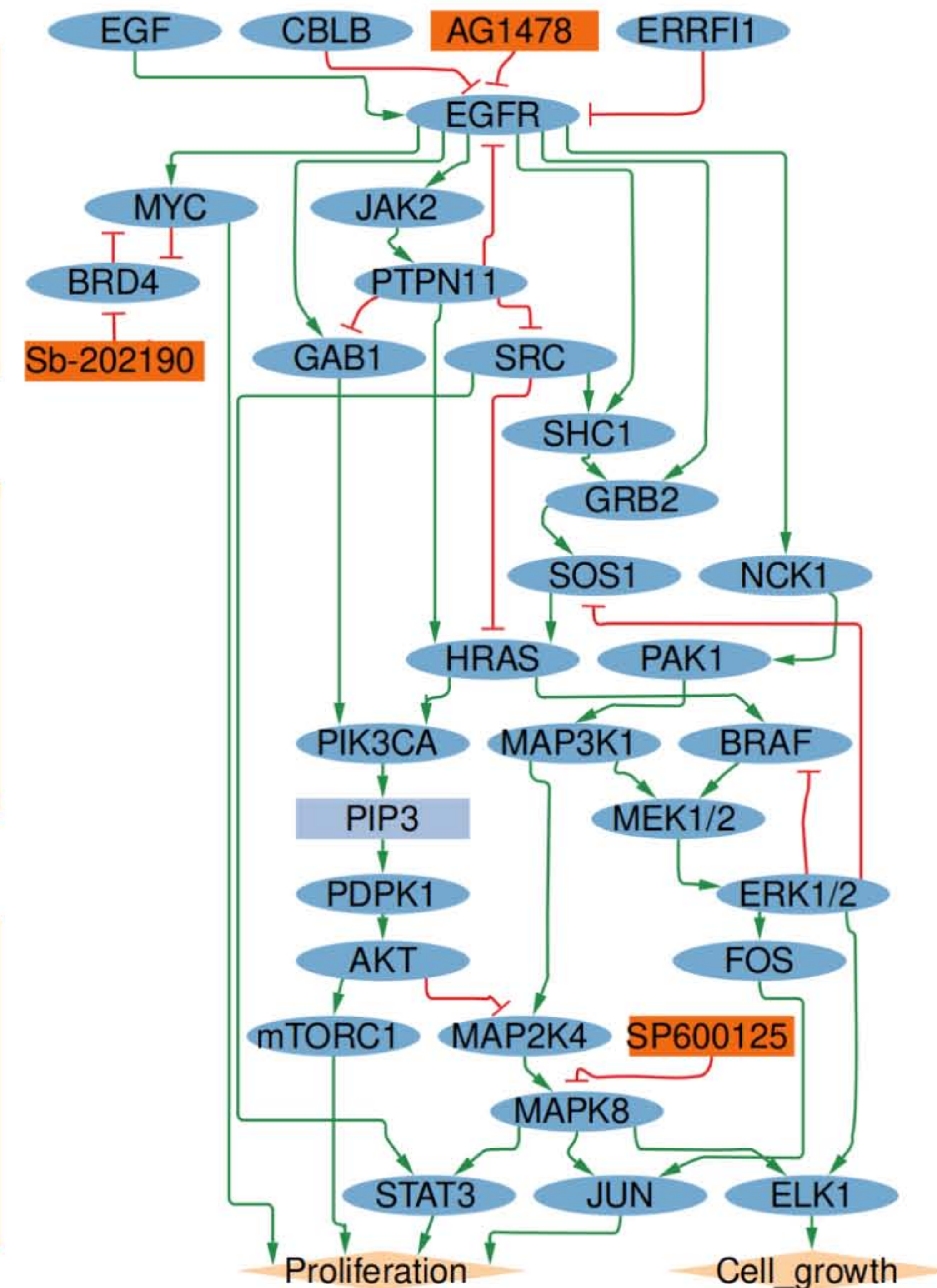


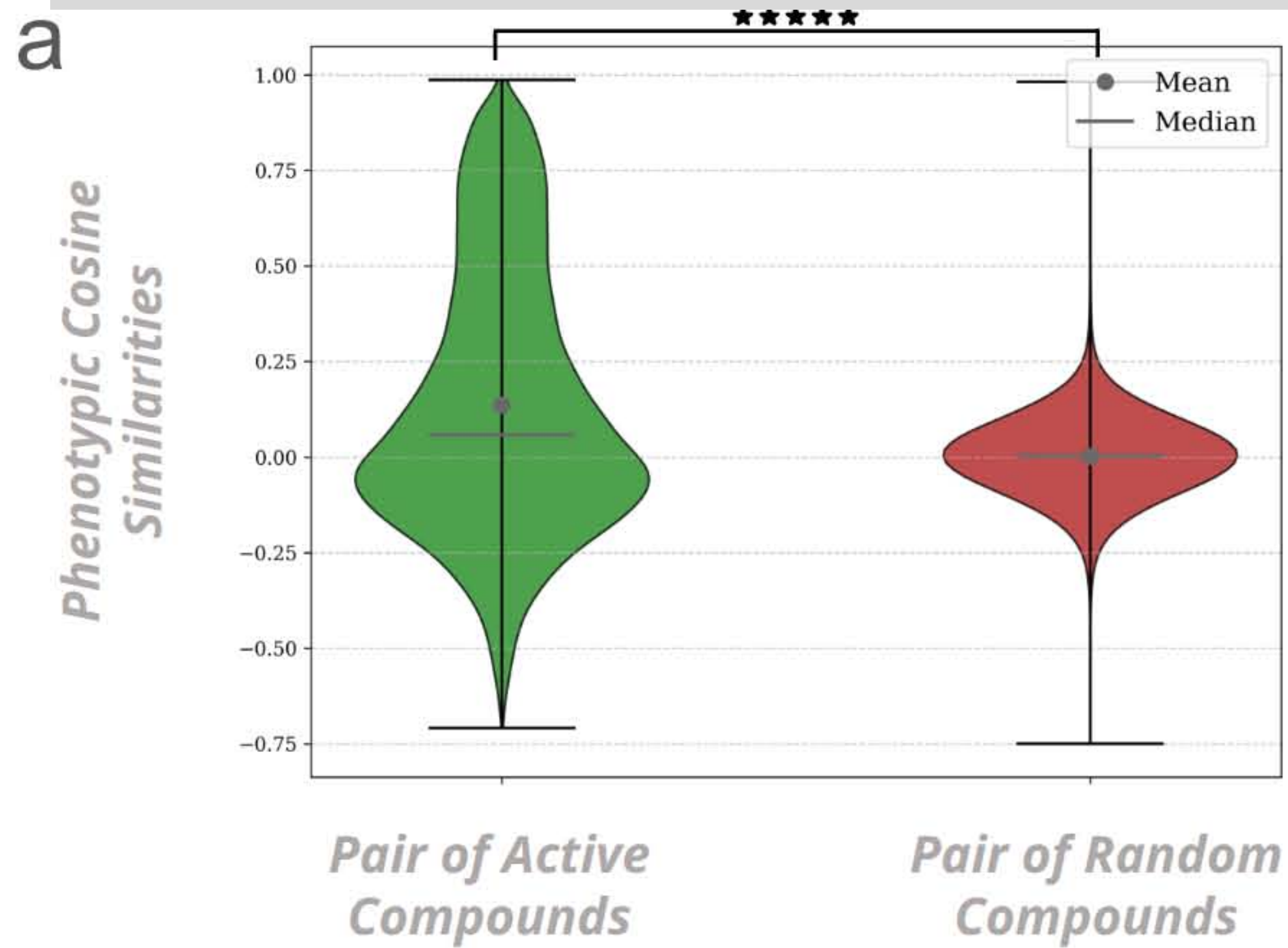
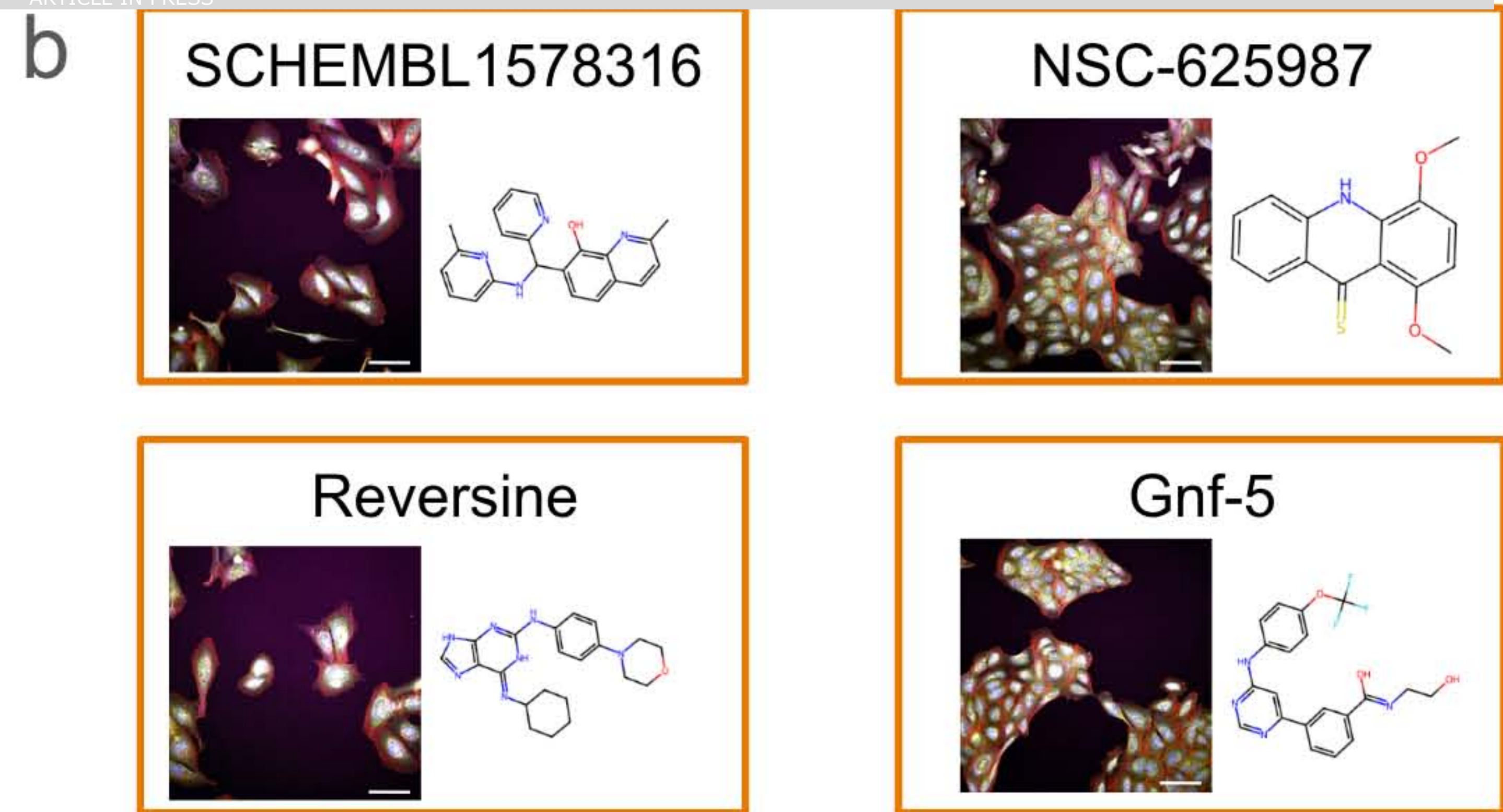
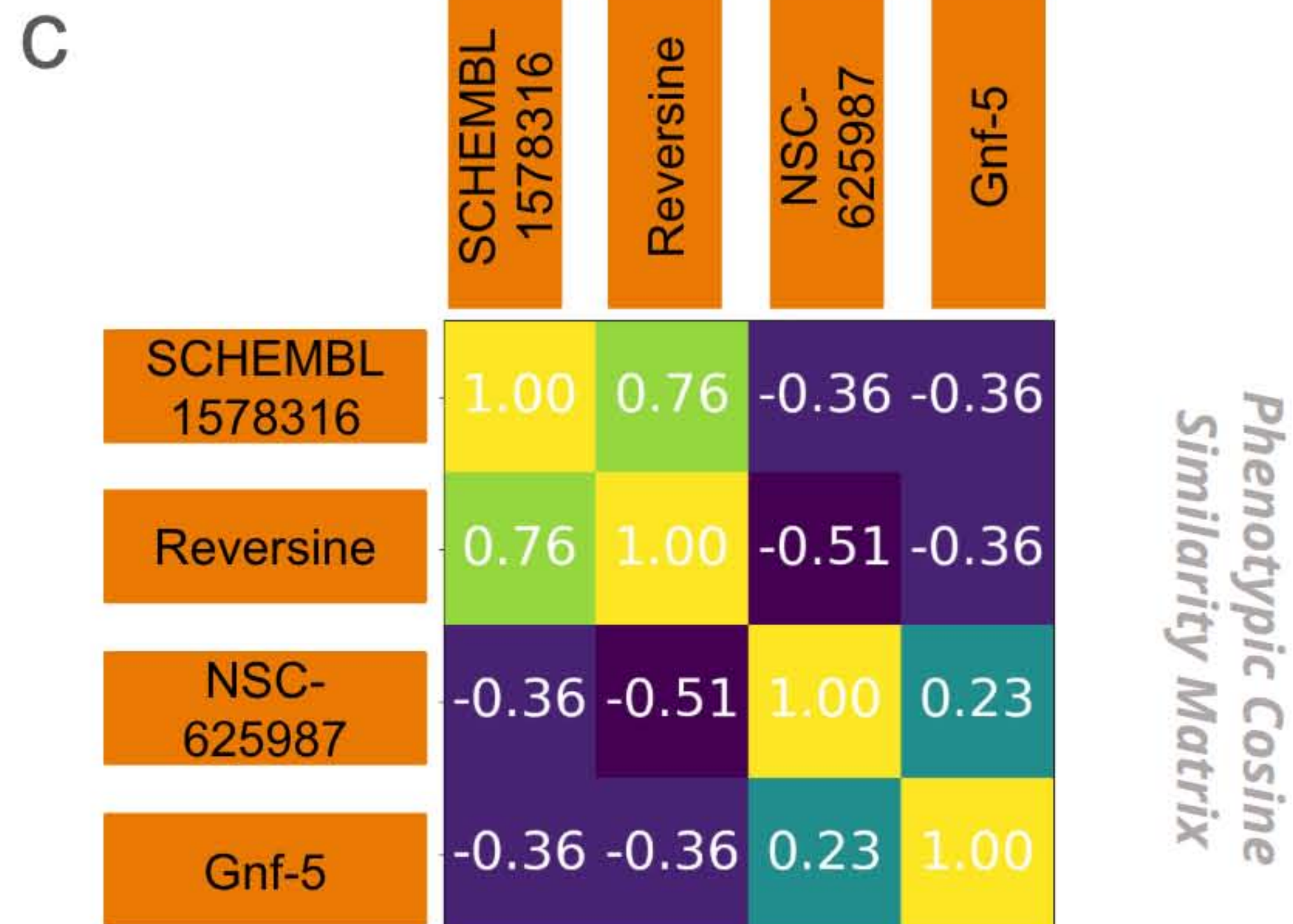
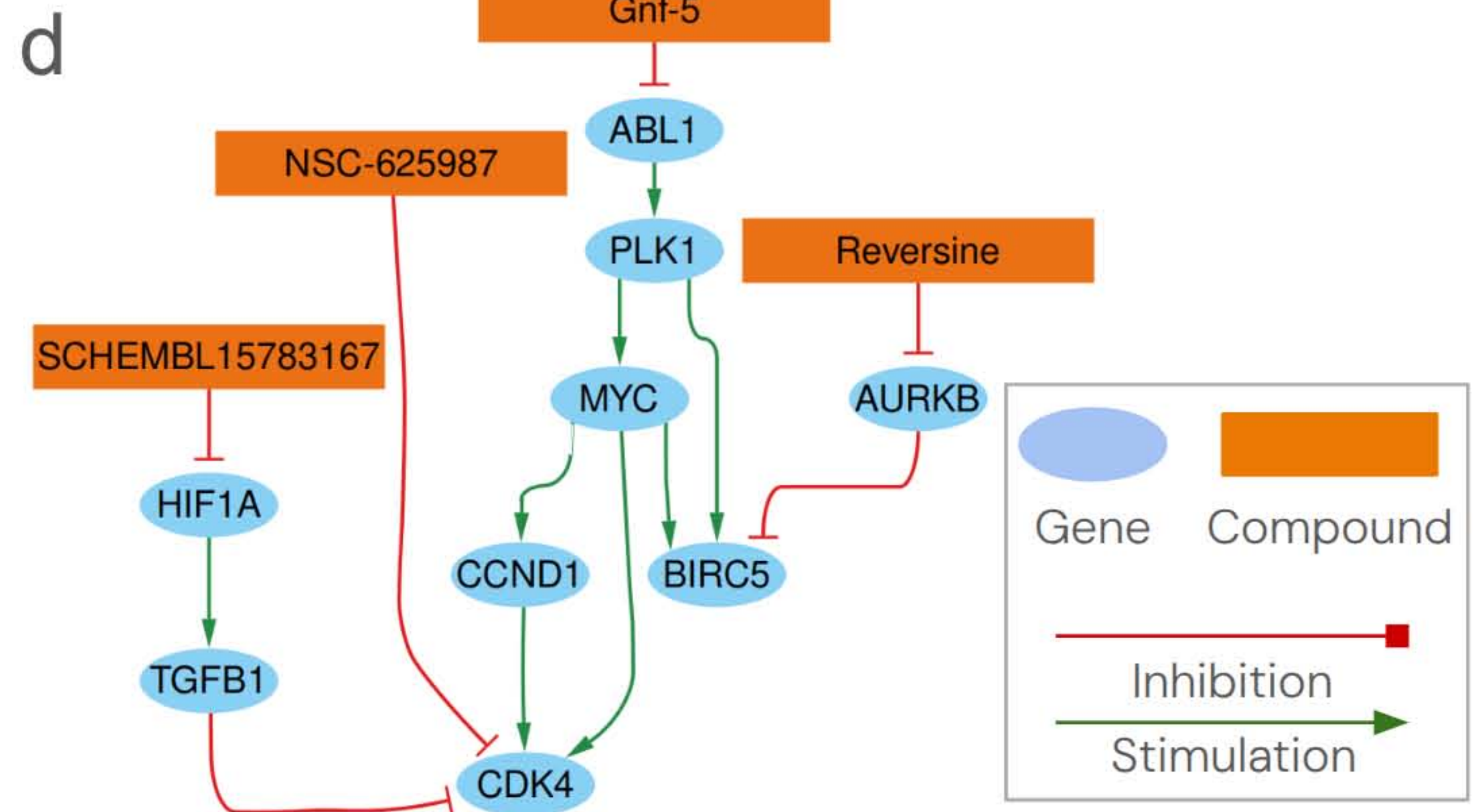
SP600125



Sb-202190

c





Pathway	Number of Genes	Number of Impacted Genes	Number of Compounds Active on the Pathway	Pair-wise Similarity Variance of Active Compounds	Pair-wise Similarity Variance of Other Compounds	Statistical Significance of Differences Between Variances	Fraction of Active Pairs with Cosine Similarity Below -0.2	Fraction of Other Pairs with Cosine Similarity Below -0.2	Statistical Significance of Differences Between Fractions	Fraction of Active Pairs with Cosine Similarity Above 0.2	Fraction of Other Pairs with Cosine Similarity Above 0.2	Statistical Significance of Differences Between Fractions
APOTHSIS	161	24	488	0.053000	0.012000	****	11.79%	3.19%	****	14.57%	3.27%	****
ALLOGRAFT REJECTION	200	21	588	0.075700	0.012100	****	13.16%	3.26%	****	22.84%	3.27%	****
ANDROGEN RESPONSE	101	10	120	0.109300	0.011900	****	14.46%	3.15%	****	36.6%	3.22%	****
APICAL JUNCTION	200	12	305	0.091700	0.011900	****	21.08%	3.11%	****	26.89%	3.17%	****
COAGULATION	138	13	144	0.033100	0.012200	****	9.25%	3.26%	****	7.84%	3.21%	****
COMPLEMENT	200	22	486	0.066000	0.012000	****	11.12%	3.15%	****	15.28%	3.19%	****
E2F TARGETS	200	17	424	0.123200	0.012000	****	16.26%	3.19%	****	37.21%	3.22%	****
ESTROGEN RESPONSE EARLY	200	15	398	0.071400	0.012100	****	10.51%	3.26%	****	22.19%	3.31%	****
ESTROGEN RESPONSE LATE	200	15	475	0.079600	0.011700	****	9.73%	3.05%	****	23.45%	3.1%	****
GSM CHECKPOINT	200	21	530	0.118300	0.011800	****	15.47%	3.18%	****	35.96%	3.2%	****
GLYCOLYSIS	200	12	420	0.106000	0.012300	****	13.55%	3.16%	****	21.26%	3.24%	****
HYPOXIA	200	11	267	0.025200	0.011900	****	7.67%	3.18%	****	23.26%	3.16%	****
IL2 STAT5 SIGNALING	169	20	484	0.068800	0.013000	****	13.08%	3.18%	****	16.75%	3.29%	****
IL6 IAK STAT3 SIGNALING	87	10	148	0.048000	0.021100	****	13.18%	3.23%	****	14.27%	3.25%	****
INFLAMMATORY RESPONSE	200	21	495	0.037700	0.011800	****	8.57%	3.33%	****	21.16%	3.37%	****
INTERFERON GAMMA RESPONSE	200	13	180	0.083000	0.011600	****	15.9%	3.14%	****	16.86%	3.15%	****
KRAS SIGNALING DN	200	13	153	0.060800	0.012200	****	14.7%	3.07%	****	16.3%	3.09%	****
KRAS SIGNALING UP	199	18	568	0.109000	0.012300	****	13.84%	3.27%	****	36.85%	3.28%	****
MITOTIC SPINDLE	200	14	474	0.045200	0.011900	****	8.32%	3.1%	****	12.73%	3.16%	****
MITOCL SIGNALING	200	11	237	0.129000	0.021100	****	15.32%	3.23%	****	35.19%	3.20%	****
MYC TARGETS V1	200	12	387	0.075700	0.011900	****	19.88%	3.16%	****	23.03%	3.17%	****
MYOGENESIS	200	19	366	0.091900	0.012500	****	16.25%	3.31%	****	24.37%	3.31%	****
PI3K PATHWAY	200	12	387	0.075700	0.011700	****	9.11%	3.25%	****	14.31%	3.33%	****
PI3K AKT MTOR SIGNALING	105	18	966	0.045500	0.012000	****	14.9%	3.06%	****	42.48%	3.3%	****
SPERMATOGENESIS	135	11	342	0.126600	0.011700	****	13.22%	3.17%	****	32.08%	3.27%	****
TGF BETA SIGNALING	54	10	255	0.115100	0.011800	****	12.11%	3.13%	****	13.05%	3.18%	****
TNFA SIGNALING VIA NFKB	200	16	248	0.046100	0.011600	****	18.75%	3.0%	****	18.98%	3.02%	****
UV RESPONSE DN	144	17	323	0.069600	0.012300	****	13.96%	3.33%	****	16.43%	3.33%	****
UV RESPONSE UP	158	16	654	0.051900	0.012300	****	14.97%	3.14%	****	16.43%	3.33%	****
XENOBIOTIC METABOLISM	200	11	91	0.075500	0.011900	****	14.97%	3.14%	****	21.2%	3.18%	****

Table 1: Tableau récapitulatif des données des pathways avec la colonne de signification pour la haute similarité opposée positionnée après la proportion d'extrêmes à gauche inactive.