Chapter 7

Cell-Based Screening to Identify Cytoprotective Compounds

Domokos Gerő

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.73172

Abstract

Prevention of cellular injury and consequent cell death is expected to provide therapeutic benefit in various diseases, but with the complexity of cell damaging pathways involved, identification and validation of novel potential drug targets is not a trivial task. New drug targets are expected to take part in complex responses with wide-ranging effects on gene expression and cellular function and drug candidates rather modify these effects than act as simple agonists or antagonists to ultimately protect the cells from an injury. Phenotypic screening may help identify cytoprotective compounds in diseases, in which the lack of drug targets makes target-based approaches unfeasible. This chapter gives an overview of the strategy of cell-based assay development, primary screening, hit selection and confirmation. Considerations about the choice of small molecule compound libraries utilized in cell-based models are discussed as well as the use of clinical drugs for drug repurposing or repositioning. The choice of cell types and issues associated with cell culture techniques are overviewed and the most common assays and readouts are briefly described. Finally, the potential pitfalls of data analysis and hit selection are discussed.

Keywords: cell-based screening, high-throughput screening, cytoprotection, drug discovery, chemical genomics, drug target

1. Introduction

The development of small molecule therapeutics for the treatment of diseases has gone through various phases over the last few decades. While the number of approved drugs continuously increased from the 1970s till 1996, when a record number of drugs (53 drugs) was approved by the Center for Drug Evaluation and Research of the U.S. Food and Drug Administration (FDA), a decline has been observed since then with no more than 20–30 new drugs approved each year [1]. In the second half of the twentieth century, drug discovery saw a rise of synthetic organic

chemistry that made it economically feasible to produce large combinatorial libraries and biology progressed to provide many novel drug targets with structural details that promoted both high-throughput screening of small molecule libraries and computer-aided drug design as potential methodologies to identify novel therapeutic agents. In 2000, the number of known human drug targets were less than 500 but with the overestimated number of human genes and an expected number of 5–10 drug targets per disease genes, the number of potential drug targets were estimated to lie between 5000 and 10,000 [2–4]. The number of human genes shrank to 19,000 and the identified small molecule drug targets rose to mere 557 (549 for FDA-approved drugs and 8 further targets for drugs approved in the rest of the world) in 2017, while 146 additional human protein targets were curated for biological drugs [5, 6]. In fact, the number of biological drugs and respective targets increased considerably over the last 20 years: in 2000 only 59 biological drugs (recombinant proteins and monoclonal antibodies) were introduced, while 250 unique biological agents were available in 2016 [3, 5]. In 2002, Hopkins and Groom argued that only those disease-modifying genes can be used as drug targets that are druggable (contain domains that small molecules can bind to) and the overlap between the druggable genome and genes linked to diseases may be between 20 and 50% [7]. They predicted that no more than 600–1500 small molecule drug targets existed in humans if we had 30,000 genes. With fewer genes than anticipated, the number of potential drug targets also decreased, which partly explains the lower number of new molecular entities over the last decades. Examining the new first-in-class drugs (compounds that modulate a novel target or biological pathway that was not targeted before their introduction) approved by the FDA during the first decade after 1999, Swinney and Anthony found that 28 of the small molecule drugs were discovered using a phenotypic screening approach compared to 17 drugs discovered by a target-based approach suggesting that phenotypic screening may be more successful because of the unbiased identification of drugs with new molecular mechanisms of action [8–10]. However, the discovery of a first-in-class drug may not be simply attributed to a phenotypic screen or a target-based approach, the two approaches do not mutually exclude each other [11]. The launch of a new chemical entity often roots in a discovery that occurred one or two decades earlier and the development of a drug that reaches the market often involves both approaches [11]. For example, a phenotypic assay may be used for the discovery of a new drug target, while subsequent optimization of drug candidates may occur via a target-based approach, or the two strategies may be combined in multiple cycles during the process. The term of 'phenotypic screening' may be broadly used for all non-target-based approaches or more selectively for testing of compounds in a system-based approach using a target-agnostic assay that monitors phenotypic changes, thus the contribution of phenotypic assays may be differently interpreted in various contexts. Nonetheless, phenotypic screening is making a comeback in drug discovery with large pharmaceutical companies like Novartis AG and GlaxoSmithKline plc admittedly promoting it, but its overall share remains unclear in the efforts of pharmaceutical companies and academics [12].

2. Planning a phenotypic screen

A clear plan is needed for any early drug discovery projects and a phenotypic screen is not an exception [13, 14]. While the steps of the drug discovery project occur in a different order

than in a target-based project, it is necessary to establish a clear go/no-go plan for all the steps (**Figure 1**). Since the target is unknown in the beginning of the project and there might be rather wide expectations about the activity of a potential cytoprotective compound, it may not be necessary to define cut-off values in the beginning. However, it is necessary to make

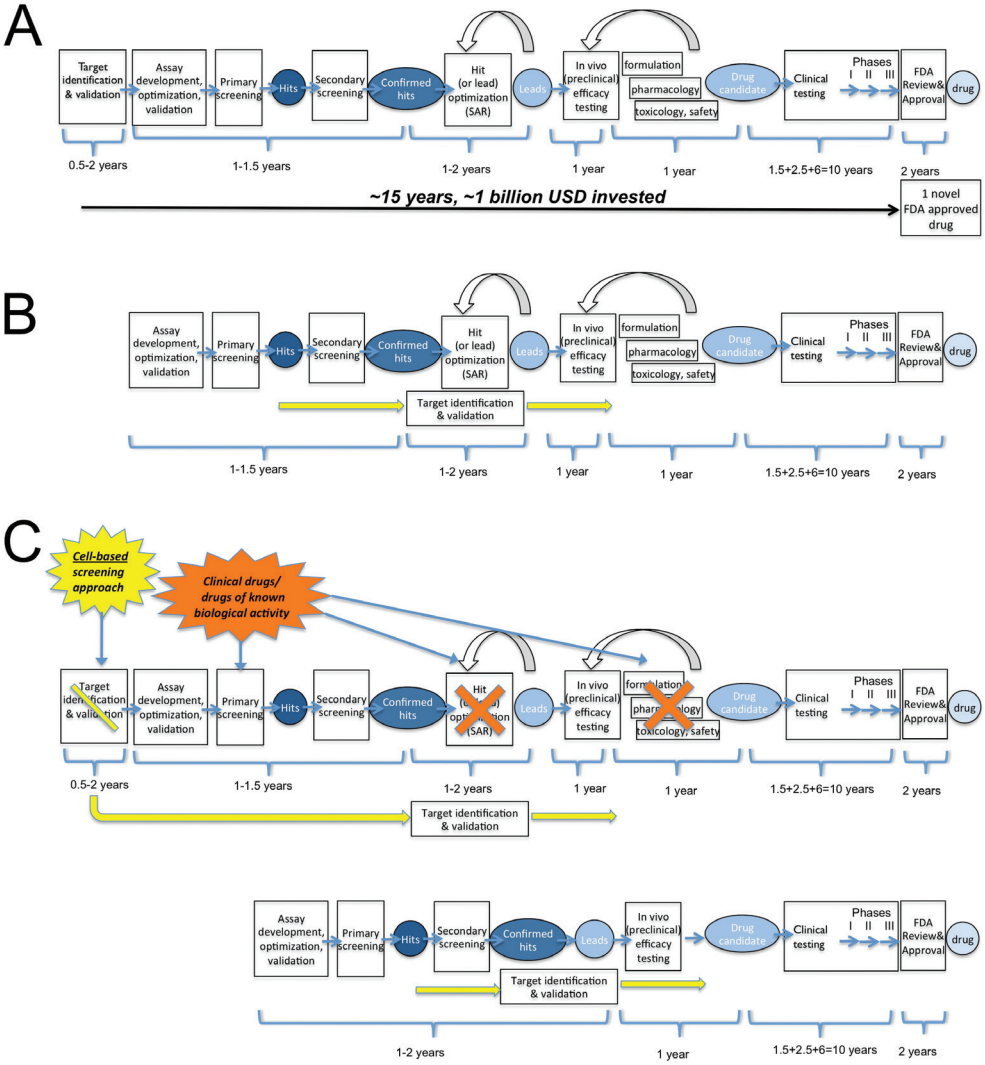

Figure 1. Basic plans of target-based and cell-based drug discovery approaches. (A) Target-based drug discovery represents the standard procedure in drug discovery. (B) Cell-based phenotypic screening skips the target identification step in the beginning of the project. This may lead to considerable time saving, since this step is usually performed simultaneously with the hit optimization. (C) Drug repurposing is possible in cell-based screening projects and it may accelerate the drug discovery process by skipping the hit optimization and toxicology steps. Please note the changes in the order of the respective steps and the expectable savings.

estimations about the market potential of a prospective drug candidate for the disease, so that we can decide on the necessary steps when the costs and potential investments arise. Fortunately, the earliest steps (optimization of the cell-based model and the primary screening) represent relatively low cost, foreseeable expenditures. Actually, the costs of the cell-based screening may not exceed the expenses associated with the identification of a novel target to start up a target-based project. Decisions can be made based on the results of the primary screen whether it is worth proceeding with a hit compound and what kind of secondary tests or models to be used based on the chemical nature of the compound. In general, higher level of flexibility is necessary during a cell-based drug discovery project than in a target-based project due to the unforeseeable nature of the drug target. Still, apart from the cell culture model of the disease it is worth having secondary *in vitro* models and a well-established animal model of the disease planned in the beginning of the project.

Phenotypic assays may be better suited for different disease areas than target-based approaches but there is no clear rule about its applicability. Santos et al. analyzed the therapeutic areas in which new drugs appeared and found that the majority of recent innovation occurred in the areas of cancer and immunology, while very little progress was seen in cardiovascular drugs [5]. On the other hand, the annual direct costs of cardiovascular disease and stroke were double the amounts of cancer-associated costs ($193.1 and $88.7 billion, respectively, in 2011) suggesting that higher progress is expected in this area [15]. When new molecular entities are considered, the lag behind other areas becomes even more apparent over recent years: cancer, infectious diseases and nervous system disorders are the leading areas of drug discovery and they all precede cardiovascular diseases [16–18]. One potential explanation for the disproportional representation of this area is the lack of new druggable targets in cardiovascular diseases, whereas the new cancer drugs and anti-infective agents are new molecular entities, they bind to novel protein targets. The other problem may be related to the nature of the injury: both anti-infective agents and cancer drugs are expected to kill the cells, while in cardiovascular diseases the expectation is to protect the cells from a harmful injury. In most instances, the mechanism of cell killing involves an inhibitory effect on the target, which is easier to attain than a stimulatory effect, just like it is simpler to produce an antagonist than an agonist for a given target since various compounds may block a binding site even if they do not fit perfectly in the active center but only a perfect molecular match can activate the target [19]. Phenotypic screening may provide a solution for the difficulties of cardiovascular drug discovery, since it is possible to find compounds that reduce the cellular injury even in the absence of a known drug target [20].

Orphan diseases represent another potential area of drug discovery using cell-based assays. Orphan diseases are rare disorders that affect small percentage of the population and thus possess a limited market potential, which led to the loss of interest by pharmaceutical companies and the lack of drug treatment in the majority of cases. The definition of orphan diseases are somewhat questionable, since there is no exact prevalence value associated with the term but in most cases if a disease affects less than 1 in 1000 or 2000 people, we call it a low prevalence or orphan disease [21–24]. While the incidence of rare diseases is low, the European Organization for Rare Diseases (EURORDIS) estimates that there are 5000–7000 distinct rare diseases and they affect approximately 6–8% of the population of the European Union [25].

The majority of these disorders are inherited diseases and drug therapy may be necessary throughout the lifetime that increases their market potential. Most governments have recognized the disproportionality of the potential profits and the necessary investments in case of rare diseases and have issued legislations to promote the development of new drug therapies for orphan diseases [26, 27]. As most of these diseases have a disease-linked gene and the mutations are easily reproduced in cellular models, they may represent the most important target diseases for cell-based phenotypic screening. Furthermore, if we accept the prediction about the number of druggable targets by Hopkins and Groom [7], the logical consequence is that many of these diseases will not be cured by an independent molecular entity but will share therapeutic drugs either with other orphan diseases or with more common illnesses. Thus, testing clinically used compounds with a drug repurposing approach may prove successful in many of these disorders, which reduces the overall costs of subsequent steps.

Cellular injury and cell death are the major challenges in today's drug discovery portfolio. While cancer drugs are on the rise, with cell killing as the principal mechanism of action of drugs, cellular injury remains the major theme of scientific research. In most diseases, researchers focus on cellular damage and cell death and investigate the underlying mechanisms that may help us understand how to interfere with the process. Unfortunately, this approach is not that successful in the discovery of new drug targets, but provides us with a multitude of cellular and animal models of various diseases [28]. These models usually allow only low throughput assays to be performed but may represent a good starting point for phenotypic screens.

In many diseases, the question arises whether reduction of the damage will be possible if currently there is no therapeutic drug for the disease in clinical practice. If, under experimental conditions, protection of the injured cells is accomplished by gene silencing or by gene therapy, there will be greater chance for establishing a pharmacological intervention in the future since it indicates the existence of disease-linked genes. It may not be necessary to induce orders-of-magnitude changes or to fully suppress the expression of a disease-linked gene to attain cell survival benefit, because the cell fate in an injury may be modified by small changes in the level of interacting proteins. Also, the existence of other experimental methods (as it was the case with ischemic pre- and post-conditioning) that induce cell protection may indicate the existence of potential targets prior to the identification of a disease-linked gene. Since phenotypic screening is a target-agnostic methodology, the outcome is of primary importance and not the underlying mechanism.

The mechanism of cytoprotection may not be identical with the blockage of known cell death pathways. Apoptosis and necrosis represent the two major cell death processes and they were long regarded as examples of "programmed" and "unprogrammed" cell death. Distinguishing a clear pathway or program in the mechanism of cell death allows us to interfere with specific components of the process and to block the execution of the program. In this respect, caspase inhibitors are the prototypical inhibitors of apoptosis [29]. However, necrosis can also be blocked by poly(ADP-ribose) polymerase (PARP) inhibitors and thus its unprogrammed classification is no longer valid, even if it took us longer to fully understand the process because of its rapid execution [30–32]. Many other regulated forms of cell death (autophagy, pyroptosis,

necroptosis, parthanathos and mitoptosis) have been identified and various drugs that block these cell-based processes played a key role in their discovery [33]. Since cell death may occur simultaneously via multiple pathways, our classification of the dominant cell death form may change over time in various diseases as we understand more details about an injury [34]. Understanding the key features of a disease or an injury and reproduction of these mechanistic details in a cellular model may be of higher value than close mimicking the cell death process, as the latter often represents the final steps in the damage that we have to prevent and not to interfere with. Interestingly, neither caspase inhibitors nor PARP inhibitors went through clinical trials in their originally conceived application, but paradoxically PARP inhibitors reached clinical practice in diseases, in which they had seemed to play a lesser role. Thus while PARP was discovered as the main contributor to necrotic cell death and the first PARP inhibitor started a trial in ST-Elevation Myocardial Infarction (STEMI) [35], PARP inhibitors reached the market later as cancer drugs: currently olaparib, rucaparib and niraparib are approved for ovarian cancer [36, 37]. Similarly, caspase inhibitors were initially suggested to play a role as potential drugs in a wide array of diseases including acquired immune deficiency syndrome (AIDS), ischemic diseases (myocardial infarction, stroke), neurodegenerative diseases, myelodysplatic syndrome and toxic liver injury [38], but clinical trials were only started in epilepsy, hepatitis C virus (HCV) infection and non-alcoholic steatohepatosis (NASH) and none of the caspase inhibitors have reached FDA approval [39–41].

3. Compounds libraries

The number of compounds used in a cell-based assay is often lower than what can be screened in a simple enzymatic reaction. The cell-based models are usually more complex and the assays typically require longer time to perform. Furthermore, the maintenance of a cell culture lab for high number of assays is more costly than what is needed for simple biochemical assays, thus it is worth considering whether the associated costs can be limited by testing fewer compounds. Also, the measurement results may show higher variance in cell-based assays than in biochemical assays and may require higher number of repeats that will considerably increase the expenses. Since there is limited information about the number of compounds that may protect the cells against an injury, we can start the screening with a validation set of compounds that may contain compounds that are known to reduce the cellular injury in that model and also include various other compounds to see the data variability. This can give us information about the expectable number of hits in larger sets of compounds and help us plan the screening strategy. A reference compound that protects against the injury may not be available when we study a new disease model, thus we may need to consider the use of controls in which the injury has not been induced and introduce the use of positive controls in the assay once we identified drugs that protect against the injury.

A two-step procedure may be preferred in the majority of phenotypic screens: starting with a smaller set of compounds with higher expectable hit ratio followed by a second screen of larger sets of compounds that may achieve a lower hit ratio. The number of hits is higher in sets of compounds that mostly contain biologically active compounds than in sets of drug-like

molecules. It is easily understandable that drugs that interfere with biological processes definitely possess binding sites, whereas those compounds that only show resemblance to other compounds do not necessarily have any targets in a cell. However, many of the biologically active compounds may show higher level of toxicity in the cells, since their known activity may not be related to cytoprotection and it may increase the data variability.

Hit selection may also include two steps in cell-based screens: (1) identification of hits in the primary screening and (2) a secondary confirmation assay of the hit molecules. Because of the higher variability of the assays and measurement values, the active compounds may show less cytoprotection in a single measurement than their average effect, thus it is better to use cut-off values that allow us to select a broader set of initial hits. These compounds will include many false positives, which will fail to show protection in repeated tests. If we expect a hit ratio around 1 in 100 molecules in the primary screen, and run confirmatory tests in repeats (e.g. in 3–6 repeats), it gives an the assay burden of ~50% for the hit confirmation step, which is substantially less than running the primary screen in duplicates. In many cases, the number of confirmed hits will be around or below 1 in 1000 compounds tested, so the set of test compounds should include a few thousand compounds in the first primary screen to produce a meaningful set of data. Fewer compounds may not contain protective molecules at all and the lack of confirmed hits often results in discontinuation of the project.

Clinical compounds can be used for initial screening efforts and a repurposing approach can speed up the drug discovery process. The number of drugs approved for human use is around 2400 in the USA and there are no more than 4000 molecular entities approved worldwide (including the US market) [24]. There are various sets available from a few vendors that contain a selection of clinically used drugs and may also include other compounds that went through toxicology studies but failed in the clinical phases (**Table 1**). If multiple of these sets are obtained from different companies, there is usually substantial overlap in the provided drugs but the vendors mostly use independent sources for the drugs and various salts of the same compound might be included in the different sets. Another option is to use a compound library of biologically active drugs. In this case, the majority of the drugs will have an annotated target in the cells but not all compounds will possess a binding site: for example, the drug target may not be expressed in the cell type in use or an anti-infective compound may not have a mammalian homolog. In general, these libraries mostly contain a similar number of potentially cytoprotective drugs as clinical libraries but there is a huge difference between them in the subsequent steps. Those compounds that have gone through formal toxicology studies, may be directly reused for other diseases, but those compounds whose toxicity have never been investigated will require more follow-up work.

Larger chemical libraries comprising 10,000–100,000 drugs may be screened as a second step following the screen of small libraries. The number of possible drug-like molecules is not indefinite, but it is certainly large enough to forget about testing all possible compounds. Virshup et al. estimated that the set of all synthetically feasible organic molecules of 500 Da molecular weight or less contained over 10^{60} structures ("the small molecule universe") [42]. Depending on the disease and target tissue, it might be possible to exclude certain chemistries and by using chemoinformatics filtering methods, the composition of the library might be limited to a set that is easier to handle [43]. A key concept is druglikeness, prediction of

Compound library	Number of drugs	Short description	Link
NIH Clinical Collection	450	Compounds that have already been in clinical use or in clinical trials	http://www.nihclinicalcollection.com/index.php?cPath=21
LOPAC (Library of pharmacologically active compounds)/ TocriScreen collection	1280/1120	Known receptor agonists, antagonists, modulators of cellular responses and signal transduction	http://www.sigmaaldrich.com/catalog/ProductDetail.do?lang=en&N4=LO1280|SIGMA&N5=SEARCH_CONCAT_PNO|BRAND_KEY&F=SPEC http://www.tocris.com/dispprod.php?ItemId=5381
US Drug Collection	1140	Compounds that have reached the stage of clinical trials in the USA	http://www.msdiscovery.com/usdrugs.html
International Drug collection	240	Compounds that have reached the stage of clinical trials in other countries	http://www.msdiscovery.com/
Killer Plates	160	Compounds with known effects on cellular viability	http://www.msdiscovery.com/killer.html
New Prestwick Chemical Library	1200	Drugs that are in clinical use	http://www.prestwickchemical.com
FDA Approved Drug Library	640	Compounds selected from another library of clinically used compounds	http://www.enzolifesciences.com/BML-2841/fda-approved-drug-library/
Apexscreen Library	5000	Diverse compound library, representative set of various chemotypes for screening	http://www.timtec.com/apexscreen.html
Chembridge Diversity Library	10,000	Drug-like diverse compound library for screening library	http://www.chembridge.com/screening_libraries/diversity_libraries/#DIVERSet
Myriascreen Library	10,000	Library of drug-like diverse compounds	http://www.timtec.com/myriascreen-diversity-collection.html
Actitarg-K library	6600	Library of kinase modulators and similar structures	http://www.timtec.com/kinase-modulators-actitarg-k-library.html
Natural Product Libtrary	640	Purified natural compounds, products	http://www.timtec.com/natural-compound-library.html
AMRI Diverse sample library	10,000	Diverse selection of 'lead-like' compounds that covers Albany Molecular Research Institutes' small molecular compound collections	http://www.amriglobal.com/products_and_services/products_detail_sub.cfm?prodID=1&subServID=4&subServID2=5

Table 1. Examples of compound libraries for cell-based screening projects.

the properties of a molecule based on the physicochemical properties of approved drugs and filtering the compound library according to these parameters [44]. The absorption, distribution, metabolism and excretion (ADME) of drug-like compounds may be predicted and subsets of drugs can be chosen for a specific organ or disease based on these data [45]. Lipinski introduced his "rule of five" concept, the filtering of molecules by solubility and permeability

prediction using the following parameters: the molecular weight is less than 500 daltons, CLog P is less than 5, the number of H-bond donors (the sum of OHs and NHs) is less than 5 and the number of H-bond acceptors (all nitrogen and oxygen atoms) is less than 10, that greatly reduces the number of potential drug-like molecules [46, 47]. This approach suggests that the number of drug-like molecules that we potentially use is closer to 10,000 than to a million drugs, since these compounds are sparsely distributed through "the small molecule universe" [47]. The use of a targeted library, which consists of drugs that are known to bind to certain types of targets and also contains highly similar molecules, is commonly used in chemical genomics and may prove useful in phenotypic screening, as well [43, 48].

The concentration of drugs used in cell-based screening is mostly determined by practicability and not by the effective or toxic concentrations of the individual compounds. The majority of compound libraries supplied compounds in solution at a fixed 1 or 2 mg/ml concentration in the past and nowadays, compounds are offered, mostly at 10 mM concentration in dimethyl sulfoxide (DMSO) [49]. The majority of the libraries are available in 96- or 384-well microplates, deep well plates or microtube racks. To simplify processing, dilutions of the compounds are best prepared at the same concentration for all drugs. While it would make more sense to use each clinical compound at a clinically relevant, effective concentration, equimolar concentrations are used most often to simplify and speed up the dilution steps. Compounds are usually screened at a concentration between 1 and 10 μmol/l, which might present dilution problems and may cause toxicity. Since the amount of DMSO must be limited as much as possible, dilutions of the drugs may be prepared in water-based solutions for cell-based assays and drug precipitation may occur as a result of poor water-solubility during dilution. Compound libraries are usually stored at −20°C in an upright position and compounds may settle down in the bottom of microtubes or wells during freeze-thaw cycles and some of the drugs may also precipitate. As a result, the cells may be treated with lower concentration of drugs than expected either due to incomplete mixing or dissolution of drugs. Also, compounds may lose activity during freezing and thawing, or due to an oxidative reaction with DMSO, which will also have various effects on the cells and may interfere with the assay [50]. Thus, hit compounds, if a fresh resupply is used in the confirmation studies, may show similar activity at a lower concentration than in the primary screen or even display higher activity.

Drug combinations may provide a further option to reduce the number of assays to be run [51, 52]. If you expect that the number of active compounds is low, you may consider pooling multiple compounds and testing them in combination. The number of assay runs may be reduced by an order of magnitude, if 8–12 compounds are pooled and only low number of "hits" is expected. Re-testing of the individual compounds will be necessary for each of the initial hits but since the expected hit ratio is low, it may not present excessive follow-up work. However, potential toxicity of the compounds needs to be considered: a toxic compound can mask the cytoprotective effect of an active compound if the cells are simultaneously treated with both drugs. Thus, the number or ratio of cytotoxic drugs might be the determining factor whether drug pooling is possible or not in a cellular model. On the other hand, this approach may allow us to search for potentiating compounds in a model if the test compounds are used in combination with a drug that provides limited protection [20, 53, 54].

4. Mammalian cell culture

Cell-based screening requires adaptation of cell culture techniques to higher throughput than what is typical in cellular assays and may present unexpected problems associated with the specific cell types or models [55]. The majority of these issues could have been sorted out by automation but the cost of robotic systems that handle cell cultivation remains extravagant. The first automated cell culture stations appeared in the market a decade ago, but these monstrous instruments remained outrageously expensive and very few units have been sold [56, 57]. Highly complex, cell-based models often include procedures that may not be automated and will require manual handling for specific steps. In this case, purchasing an expensive instrument that cannot fully replace the laboratory technician may increase the overall costs. Whereas training technicians specifically for higher throughput cell culture may be advantageous: provides flexibility, predictable costs and allows the introduction of newer models when needed.

Choosing the right cell type for the primary screen might be the key for discovering novel drugs or drug actions. Primary cells are often propagated for cell-based screening projects, assuming that they more closely mimic the processes of the cells in whole organisms than cell lines. However, primary cells are usually more difficult to grow and their division potential is limited to a few passages, during which they may not maintain their original characteristics. Whereas cell lines are well-characterized, immortalized cells that are easy to maintain and can be resupplied at reasonable costs, if necessary. The reagent and cell maintenance costs are lower for cell lines, as they can be cultured in classic cell culture media and typically grow faster. Human cells may be better suited for drug discovery, since the potential drug targets may not have orthologs in other species or the dissimilarity between the representative proteins may change the binding of compounds to the target [5]. On the other hand, many of the cell-based models use non-human cells either because the preferred cell-line originates from rodents or because the primary cells cannot be freshly isolated from humans. Either way, choosing the best model is more important than the use of human cells, since it is better to loose a few hits because of the differences in orthologs than to identify hits in an irrelevant model. If there are multiple options to choose from, it is worth keeping the lower cost model for the primary screening and use the other cell type as secondary model to confirm the action of hit compounds [58, 59].

Assay miniaturization is necessary to reduce the costs of cell culture and the assay-associated expenses [60]. Although, the ultimate goal is to minimize reagent costs, the power of testing treatments in larger cell populations will remain high and may be preferential depending on the model or assay type [61, 62]. Thus, 96-well cell culture plates are often used in more complex assays, while 384- and 1536-well plates can be used for simple assays when the assay readout is expected to change considerably with the treatments. In a simple cell viability assay, if the cellular injury reduces the viability by 50% and we expect that our drug treatment may provide a partial protection against the injury (e.g. 20% survival benefit) then the pipetting error and imprecision of the measurement may not allow us to downsize the assay. It should be noted that in cell culture, miniaturization involves a more complex adaptation of the cell cultivation process than proportional scaling down of the number of seeded cells and all the reagents. Differences in aeration and temperature fluctuations occur between the wells based on their respective position on a plate: typically the edges are exposed to excessive

temperature changes, evaporation and better oxygenation, while the inner wells are more homogeneous in this respect but they do not have comparable gas exchange and the cells may grow slightly slower in them. As a result, the outer wells may be excluded from analysis, leaving only 60 wells on a 96-well plate or 308 wells on a 384-well plate, thus 5-times higher number of treatments might be tested on a 384-well plate.

Contamination is a major issue in cell-based screening, because of the higher throughput of assays and steps that may not be carried out in an aseptic environment [63–65]. Compound libraries are often provided in solutions that are not sterile, though the drugs are dissolved in DMSO and there is little chance for bacterial contamination. The dilution step of compounds may be associated with contamination risk, partly because the storage plates often use tube caps that require manual handling and cannot be removed by a robotic system or because the pipetting station used for the dilution step does not fit in a biosafety cabinet. Also, experimental devices that are necessary to induce cellular injury may be utilized in some models and these instruments may exceed the size limits of a laminar flow cabinet [59]. While contamination risk may be minimized by careful planning of experiments, it is often impossible to eliminate all sources of bacterial contaminants. If the cell cultivation procedure is longer because the cells require longer differentiation steps, or the assay investigates the long-term survival of cells, the risk of contamination will also increase. In general, assays terminated within 3–4 h following the injury or drug treatment, may allow that some of the steps are carried out on the benchtop and may not require to work in a biosafety cabinet, since the bacterial growth becomes exponential after a lag phase and within this time-frame the low level of contaminants impose little risk of interfering with the assay [66]. However, longer exposures or assay runs requisite that all the steps are performed in a biosafety cabinet using sterile technique. It is essential to plan regular decontamination and sterilization of the instruments to avoid generating false data because of contaminants [67].

5. Assay-related issues

The readout of the screening assay that may be a viability assay or another functional test plays an important role in the identification of hit compounds and it is a key contributor to the expenses of the project. A high sensitivity, low cost assay is the obvious choice for primary cell-based screening of compound libraries [68, 69]. High sensitivity does not mean that we need to use luminescence measurements, because it is not the detection sensitivity what is important but the sensitive detection of the changes in cellular function or viability. A simple, colorimetric lactate dehydrogenase (LDH) assay may be capable to measure minor changes in cell viability that costly methodologies cannot detect. Reagent cost makes a huge difference if the number of runs is high, thus a less expensive assay may be preferable. A custom-tailored assay with "home-made" reagents may be a better choice than a commercial assay, depending on the model, and may save on supplies.

Specificity of the assay and the instrumentation necessary to read the plates are of similar importance but reproducibility and higher sensitivity may allow more flexibility in these parameters. Since we expect that the number of cytoprotective compounds is low in most models, it is better to have fewer false negatives and more false positives among the initial hits, thus

sensitivity may overcome the limitations of specificity. Endpoint assays are typically preferred, since the measurement time can severely limit the number of runs in kinetic assays. If a plate is read for 30 min in a kinetic assay, and the endpoint measurement lasts 1 min, it will take more than 2 days to read 100 plates kinetically, while the endpoint measurement can be finished within 2 h. (And no more than 6000–8000 compounds are tested on 100 plates if 96-well plates are used and the outer wells are excluded.) Thus, if an endpoint assay can replace a kinetic read with no or little loss in sensitivity, it remains the preferred choice in cell-based screening.

Viability assays are obvious choices for identification of protective compounds in a cytotoxic injury and many simple assays are available to choose from [70–72]. The most commonly used assays depend on substrates that are converted to easily detectable products by metabolic enzymes in living cells and the results are linear with the number of viable cells within a limited range [73–75]. The least expensive options are colorimetric or fluorescent tests that requisite less costly plate readers. However, assay interference is a common problem with these methods: many of the test compounds are colored or fluorescent substances and they may produce false results. The metabolic pathways or enzymes may be up or down-regulated in the studied injury model, thus careful selection of the assay is necessary [53, 71, 72]. Homogeneous assays may be preferred, since fewer steps allow less processing error, and easier automation, robotic integration of the assay.

Simultaneous use of multiple viability assays can help compensate for the weaknesses of individual tests but makes the data analysis more complex [53, 58, 71]. For example, in a cellular injury model, when the methylthiazolyldiphenyl-tetrazolium bromide (MTT) assay is the primary viability assay, LDH release may be used as a secondary measurement from the simultaneously sampled cell culture supernatant [20, 53, 71, 72]. Thus, cytoprotection (increased viability) will be detectable as an increased signal in the MTT assay and also as a reduced LDH signal (diminished cell death) in the respective supernatant sample. Whereas the simultaneous increase or decrease of both the MTT and LDH signals indicates a false positivity in one of the assays and may not be of further interest.

In cell-based assays, differences occur in the assay results depending on the respective position of individual wells, since the activity of metabolic enzymes is affected by the temperature fluctuations, inhomogeneity of gas exchange and potential pH variations. These effects may affect the results considerably in many injury models and corrections may be necessary to compensate for their impact. But even with the best compensations, the chances to find protective compounds in middle or outer wells may not be equal [53]. It is worth testing, that with the use of a positive control compound, our assay can similarly detect protection in all wells of the plate. Alternatively, we may confirm that mock-treated or vehicle treated (negative control) wells show similar results throughout the whole plate. If the measurements show larger variation, it may be necessary to run measurements in duplicates. In this case, duplicates need to be allocated to various positions and not to neighboring wells on a plate, so that the average protection will be similar for all test positions [53].

Independent repeats are expected in all cell culture experiments but running cell-based screens repeatedly several times would unnecessarily increase the expenses [76–79]. We can assume that only a handful of drugs may protect the cells and the majority of the compounds

will have little effect in the injury model. It is unnecessary to confirm the lack of effect for those drugs that exert no protection. The goal of the screen is to find potentially active drugs, thus screening a compound library needs a single measurement, even if we use a cell-based assay. The primary screen will not provide necessary data to state that the initial hits exert significant protection, we need to confirm the protective effect of potential hit compounds in repeated experiments to conclude that the effect is significant [77–79].

6. Data analysis

Hit identification is the main aim of data analysis in cell-based screening and this process may not follow the rules of statistics to the letter. In high-throughput screening, Z-factor has been used as the gold standard method to analyze the assay and data set quality [80]. This methodology assumes that the identification of hit compounds requisite good separation of positive and negative controls. However, the power of this methodology was questioned later, since higher background or noise values are typically present in several assays, and poor signal-to-background or signal-to-noise separation still allow the identification of "true hits" [81, 82]. Larger variation occurs in a cell-based assay than in biochemical assays, even so cell culture experiments have been successfully used to study the effects of drug treatments. Thus, the variable damage induced in many injury models is an acceptable feature of the experimental methodology but may require more flexibility during data analysis.

The ultimate goal is to identify compounds that provide protection against an injury, which means less damage or increased viability of the cells depending on the specific model. A Z-score, the number of standard deviations from the mean vehicle treated wells undergoing the injury, may be calculated using the whole dataset and it easily shows the outliers [83]. Based on the assumptions that (1) the wells treated with various ineffective drugs will show similar response to the vehicle-treated wells and (2) there are very few active drugs, the compound-treated wells can quasi replace the negative control wells in the analysis. However, if the dataset shows larger variation, the Z-scores will be smaller values, thus meaningful cut-off values should be chosen individually for the selection of hits in each experimental model [54, 58, 59]. Apart from the more complex procedures, like Z-score calculation, there are various other options to identify cytoprotective compounds. Selection of the wells, in which diminished cell death occurred, is often easier on a single plate than in a large dataset, thus it might be useful to establish cut-off values based on the results of each plate separately. Data may be categorized immediately at plate reading and the outliers can be labeled as potential hit compounds or toxic molecules depending on the changes in viability. In this way, you can start up the hit confirmation before the completion of the primary screen, which may be preferable if the primary screen of a larger compound library requisites longer time.

Hit selection may be overly complex if dual or multiple readouts are used to identify active compounds. In the simplest case, one of the parameters is used only at a fixed cut-off value and the second parameter is ranked according to the activity to find the protective compounds. In a cell injury model, viability values may be used for preliminary classification of drugs as toxic and non-toxic compounds, and the other parameter (e.g. ROS production or

an inflammatory signal) may be used to determine the protective effect of test compounds [54, 58]. Alternatively, the measurement values may be ranked separately for the individual parameters and cut-off Z-scores may be established for each of them. If a test compound performs above the cut-off values for each of the parameters, we may classify it as a hit that proceeds to the confirmation step. A further option is to streamline the decision making to the usual single parameter analysis: a new factor may be derived from the two primary measurement values, and the newly generated parameter can be ranked by the Z-scoring method.

Hit confirmation is the next step, following the identification of preliminary hits [83]. This procedure can take the usual steps of cell culture experiments. Independent repeats are needed and data analysis should be performed following the generally accepted rules of *in vitro* assays [79]. If more assays are available to test the activity of potential hits, it is worth including them at this step as secondary assays [83]. The cellular target of confirmed hits often represents new challenges, since the identification may not be a straightforward process in many of the disease models. However, the new targets are of high value for structure activity relationship (SAR) analysis and may shed new light on the mechanisms of disease development and progression.

7. Conclusion

Recent tendency in drug discovery suggests that target-based research will be complemented with target-agnostic approaches in the future. As opposed to classical target-based drug discovery approaches phenotypic assays may be necessary to identify novel compounds that show activity in orphan diseases or in common medical conditions that currently lack effective therapeutics. Cell-based models are often used to study various aspects of illnesses and many of these may be modified for powerful tools in drug discovery. The use of these models for cell-based screening may allow identification of potential drug candidates and chemical genomics approaches can promote reverse identification of novel drug targets [84].

Acknowledgements

D.G. received funding from the People Programme (Marie Curie actions) of the European Union's Seventh Framework Programme under the grant agreement number 62810.

Author details

Domokos Gerő[1,2]*

*Address all correspondence to: gerodomokos@yahoo.com

1 Institute of Pathophysiology, Semmelweis University Medical School, Budapest, Hungary

2 University of Exeter Medical School, Exeter, UK

References

[1] U. S. Food and Drug Administration. Summary of NDA Approvals & Receipts, 1938 to the present. Available from: https://www.fda.gov/aboutfda/whatwedo/history/product-regulation/summaryofndaapprovalsreceipts1938tothepresent/default.htm

[2] Pennisi E. Human Genome Project. And the gene number is…? Science. 2000;**288**(5469): 1146-1147

[3] Drews J. Drug discovery: A historical perspective. Science. 2000;**287**(5460):1960-1964

[4] Ewing B, Green P. Analysis of expressed sequence tags indicates 35,000 human genes. Nature Genetics. 2000;**25**(2):232-234

[5] Santos R et al. A comprehensive map of molecular drug targets. Nature Reviews. Drug Discovery. 2017;**16**(1):19-34

[6] Chi KR. The dark side of the human genome. Nature. 2016;**538**(7624):275-277

[7] Hopkins AL, Groom CR. The druggable genome. Nature Reviews. Drug Discovery. 2002; **1**(9):727-730

[8] Swinney DC. Phenotypic vs. target-based drug discovery for first-in-class medicines. Clinical Pharmacology and Therapeutics. 2013;**93**(4):299-301

[9] Swinney DC, Anthony J. How were new medicines discovered? Nature Reviews. Drug Discovery. 2011;**10**(7):507-519

[10] Zheng W, Thorne N, McKew JC. Phenotypic screens as a renewed approach for drug discovery. Drug Discovery Today. 2013;**18**(21-22):1067-1073

[11] Eder J, Sedrani R, Wiesmann C. The discovery of first-in-class drugs: Origins and evolution. Nature Reviews. Drug Discovery. 2014;**13**(8):577-587

[12] Kotz J. Phenotypic screening, take two. Science-Business eXchange. 2012;**5**(15):1-3

[13] Strovel JS, Sittampalam NP, Coussens M, Hughes J, Inglese A, Kurtz A, Andalibi L, Patton C, Austin M, Baltezor M, Beckloff M, Weingarten M, Weir S. Early drug discovery and development guidelines: For academic researchers, collaborators, and start-up companies. In: Sittampalam GS. et al., Editors. Assay Guidance Manual. Bethesda (MD): 2004. p. 3-44

[14] Hughes JP et al. Principles of early drug discovery. British Journal of Pharmacology. 2011;**162**(6):1239-1249

[15] Writing Group, M et al. Executive summary: Heart disease and stroke statistics-2016 update: A report from the American Heart Association. Circulation. 2016;**133**(4):447-454

[16] Kinch MS et al. Target selection for FDA-approved medicines. Drug Discovery Today. 2015;**20**(7):784-789

[17] Kinch MS, Merkel J, Umlauf S. Trends in pharmaceutical targeting of clinical indications: 1930-2013. Drug Discovery Today. 2014;**19**(11):1682-1685

[18] Munos B. A forensic analysis of drug targets from 2000 through 2012. Clinical Pharmacology and Therapeutics. 2013;**94**(3):407-411

[19] Rang HP. The receptor concept: Pharmacology's big idea. British Journal of Pharmacology. 2006;**147**(Suppl 1):S9-16

[20] Gero D et al. Oxidant-induced cardiomyocyte injury: Identification of the cytoprotective effect of a dopamine 1 receptor agonist using a cell-based high-throughput assay. International Journal of Molecular Medicine. 2007;**20**(5):749-761

[21] Baldovino S et al. Rare diseases in Europe: From a wide to a local perspective. The Israel Medical Association Journal. 2016;**18**(6):359-363

[22] Montserrat Moliner A, Waligora J. The European union policy in the field of rare diseases. Public Health Genomics. 2013;**16**(6):268-277

[23] Taruscio D et al. National registries of rare diseases in Europe: An overview of the current situation and experiences. Public Health Genomics. 2015;**18**(1):20-25

[24] Huang R et al. The NCGC pharmaceutical collection: A comprehensive resource of clinically approved drugs enabling repurposing and chemical genomics. Science Translational Medicine. 2011;**3**(80):80ps16

[25] EURORDIS: European Organisation of Rare Diseases. Rare Diseases: Understanding this Public Health Priority. 2005; Available from: https://www.eurordis.org/IMG/pdf/princeps_document-EN.pdf

[26] Dharssi S et al. Review of 11 national policies for rare diseases in the context of key patient needs. Orphanet Journal of Rare Diseases. 2017;**12**(1):63

[27] Gammie T, Lu CY, Babar ZU. Access to orphan drugs: A comprehensive review of legislations, regulations and policies in 35 countries. PLoS One. 2015;**10**(10):e0140002

[28] Horvath P et al. Screening out irrelevant cell-based models of disease. Nature Reviews. Drug Discovery. 2016;**15**(11):751-769

[29] Pronk GJ et al. Requirement of an ICE-like protease for induction of apoptosis and ceramide generation by REAPER. Science. 1996;**271**(5250):808-810

[30] Gero D, Szabo C. Poly(ADP-ribose) polymerase: A new therapeutic target? Current Opinion in Anaesthesiology. 2008;**21**(2):111-121

[31] Gero D et al. Modulation of poly(ADP-ribose) polymerase-1 (PARP-1)-mediated oxidative cell injury by ring finger protein 146 (RNF146) in cardiac myocytes. Molecular Medicine. 2014;**20**:313-328

[32] Jagtap P, Szabo C. Poly(ADP-ribose) polymerase and the therapeutic effects of its inhibitors. Nature Reviews. Drug Discovery. 2005;**4**(5):421-440

[33] Duprez L et al. Major cell death pathways at a glance. Microbes and Infection. 2009;**11**(13):1050-1062

[34] Pasparakis M, Vandenabeele P. Necroptosis and its role in inflammation. Nature. 2015;**517**(7534):311-320

[35] Morrow DA et al. A randomized, placebo-controlled trial to evaluate the tolerability, safety, pharmacokinetics, and pharmacodynamics of a potent inhibitor of poly(ADP-ribose) polymerase (INO-1001) in patients with ST-elevation myocardial infarction undergoing primary percutaneous coronary intervention: Results of the TIMI 37 trial. Journal of Thrombosis and Thrombolysis. 2009;**27**(4):359-364

[36] Davar D et al. Role of PARP inhibitors in cancer biology and therapy. Current Medicinal Chemistry. 2012;**19**(23):3907-3921

[37] Brown JS, Kaye SB, Yap TA. PARP inhibitors: The race is on. British Journal of Cancer. 2016;**114**(7):713-715

[38] Thompson CB. Apoptosis in the pathogenesis and treatment of disease. Science. 1995; **267**(5203):1456-1462

[39] O'Brien T, Linton SD. Design of Caspase Inhibitors as Potential Clinical Agents. CRC Enzyme Inhibitors Series. Boca Raton: CRC Press/Taylor & Francis Group; 2009. xii 299 p

[40] Kaur H, Kumar B, Medhi B. Antiepileptic drugs in development pipeline: A recent update. eNeurologicalSci. 2016;**4**(Supplement C):42-51

[41] Linton SD. Caspase inhibitors: A pharmaceutical industry perspective. Current Topics in Medicinal Chemistry. 2005;**5**(16):1697-1717

[42] Virshup AM et al. Stochastic voyages into uncharted chemical space produce a representative library of all possible drug-like compounds. Journal of the American Chemical Society. 2013;**135**(19):7296-7303

[43] Dandapani S et al. Selecting, acquiring, and using small molecule libraries for high-throughput screening. Current Protocols in Chemical Biology. 2012;**4**:177-191

[44] Bickerton GR et al. Quantifying the chemical beauty of drugs. Nature Chemistry. 2012; **4**(2):90-98

[45] Tao L et al. Recent progresses in the exploration of machine learning methods as in-silico ADME prediction tools. Advanced Drug Delivery Reviews. 2015;**86**:83-100

[46] Lipinski CA. Drug-like properties and the causes of poor solubility and poor permeability. Journal of Pharmacological and Toxicological Methods. 2000;**44**(1):235-249

[47] Lipinski CA et al. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Advanced Drug Delivery Reviews. 2001;**46**(1-3):3-26

[48] Downey W, Liu C, Hartigan J. Compound profiling: Size impact on primary screening libraries. Drug Discovery World. 2010;(Spring)

[49] Chan JA, Hueso-Rodríguez JA. Compound library management. In: Janzen WP. editor. High Throughput Screening: Methods and Protocols. Humana Press: Totowa, NJ; 2002 pp. 117-127

[50] Jacob SW, Herschler R. Pharmacology of DMSO. Cryobiology. 1986;**23**(1):14-27

[51] Kainkaryam RM, Woolf PJ. Pooling in high-throughput drug screening. Current Opinion in Drug Discovery & Development. 2009;**12**(3):339-350

[52] Elkin LL et al. Just-in-time compound pooling increases primary screening capacity without compromising screening quality. Journal of Biomolecular Screening. 2015;**20**(5):577-587

[53] Szoleczky P et al. Identification of agents that reduce renal hypoxia-reoxygenation injury using cell-based screening: Purine nucleosides are alternative energy sources in LLC-PK1 cells during hypoxia. Archives of Biochemistry and Biophysics. 2012;**517**(1):53-70

[54] Gero D et al. Identification of pharmacological modulators of HMGB1-induced inflammatory response by cell-based screening. PLoS One. 2013;**8**(6):e65994

[55] Maddox CB, Rasmussen L, White EL. Adapting cell-based assays to the high throughput screening platform: Problems encountered and lessons learned. JALA Charlottesv Va. 2008;**13**(3):168-173

[56] Lehmann R et al. Biomek cell workstation: A variable system for automated cell cultivation. Journal of Laboratory Automation. 2016;**21**(3):439-450

[57] Ozturk SS. Equipment for large-scale mammalian cell culture. Advances in Biochemical Engineering/Biotechnology. 2014;**139**:69-92

[58] Gero D et al. Cell-based screening identifies paroxetine as an inhibitor of diabetic endothelial dysfunction. Diabetes. 2013;**62**(3):953-964

[59] Lopez-Garcia I, Gero D, Szczesny B, Szoleczky P, Olah G, Modis K, Zhang K, Gao J, Wu P, Sowers LC, DeWitt D, Prough DS, Szabo C. Development of a stretch-induced neurotrauma model for medium-throughput screening in vitro: Identification of rifampicin as a neuroprotectant. Br J Pharmacol. 2018;1**75**(2):284-300

[60] Sundberg SA. High-throughput and ultra-high-throughput screening: Solution- and cell-based approaches. Current Opinion in Biotechnology. 2000;**11**(1):47-53

[61] Maffia AM 3rd, Kariv II, Oldenburg KR. Miniaturization of a mammalian cell-based assay: Luciferase reporter gene readout in a 3 microliter 1536-well plate. Journal of Biomolecular Screening. 1999;**4**(3):137-142

[62] Mayr LM, Bojanic D. Novel trends in high-throughput screening. Current Opinion in Pharmacology. 2009;**9**(5):580-588

[63] Cote RJ. Aseptic technique for cell culture. Current Protocols in Cell Biology. 2001;Chapter 1: Unit 1. 3

[64] Stacey GN. Cell culture contamination. Methods in Molecular Biology. 2011;**731**:79-91

[65] Phelan K, May KM. Mammalian cell tissue culture techniques. Current Protocols in Pharmacology. 2016;**73**:12.1.1-12.1.23

[66] Naveen V, Hsiao CD. NrdR transcription regulation: Global proteome analysis and its role in *Escherichia coli* viability and virulence. PLoS One. 2016;**11**(6):e0157165

[67] Callaway E. Contamination hits cell work. Nature. 2014;**511**(7511):518

[68] Fernandes PB. Technological advances in high-throughput screening. Current Opinion in Chemical Biology. 1998;**2**(5):597-603

[69] Kelkar M, De A. Bioluminescence based in vivo screening technologies. Current Opinion in Pharmacology. 2012;**12**(5):592-600

[70] Chan GK et al. A simple high-content cell cycle assay reveals frequent discrepancies between cell number and ATP and MTS proliferation assays. PLOS One. 2013;**8**(5):e63583

[71] Modis K et al. Cytoprotective effects of adenosine and inosine in an in vitro model of acute tubular necrosis. British Journal of Pharmacology. 2009;**158**(6):1565-1578

[72] Modis K et al. Adenosine and inosine exert cytoprotective effects in an in vitro model of liver ischemia-reperfusion injury. International Journal of Molecular Medicine. 2013; **31**(2):437-446

[73] Prabst K et al. Basic colorimetric proliferation assays: MTT, WST, and Resazurin. Methods in Molecular Biology. 2017;**1601**:1-17

[74] Stoddart MJ. Cell viability assays: Introduction. Methods in Molecular Biology. 2011; **740**:1-6

[75] Riss TL, Moravec RA, Niles AL, Duellman S, Benink HA, Worzella TJ, Minor L. Cell viability assays. In: Sittampalam GS. et al., Editors. Assay Guidance Manual. Bethesda (MD): 2004. p. 355-385

[76] How robust are your data? Nature Cell Biology. 2009;**11**(6):667

[77] Curtis MJ, Abernethy DR. Advice on statistical analysis, and new journal guidance for experimental design and analysis. Pharmacology Research & Perspectives. 2015;**3**(1):e00095

[78] Curtis MJ et al. Experimental design and analysis and their reporting: New guidance for publication in BJP. British Journal of Pharmacology. 2015;**172**(14):3461-3471

[79] Lazic SE. Experimental Design for Laboratory Biologists : Maximising Information and Improving Reproducibility. Cambridge, United Kingdom: Cambridge University Press; 2016. xv 412 p

[80] Zhang JH, Chung TD, Oldenburg KR. A simple statistical parameter for use in evaluation and validation of high throughput screening assays. Journal of Biomolecular Screening. 1999;**4**(2):67-73

[81] Shun TY et al. Identifying actives from HTS data sets: Practical approaches for the selection of an appropriate HTS data-processing method and quality control review. Journal of Biomolecular Screening. 2011;**16**(1):1-14

[82] Sui Y, Wu Z. Alternative statistical parameter for high-throughput screening assay quality assessment. Journal of Biomolecular Screening. 2007;**12**(2):229-234

[83] Malo N et al. Statistical practice in high-throughput screening data analysis. Nature Biotechnology. 2006;**24**(2):167-175

[84] Jones LH, Bunnage ME. Applications of chemogenomic library screening in drug discovery. Nature Reviews. Drug Discovery. 2017;**16**(4):285-296