High-throughput Prediction of Nephrotoxicity in Humans

Lit-Hsin Loo[1,2] and Daniele Zink[3]

[1]Bioinformatics Institute (BII), Singapore; [2]Department of Pharmacology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore; [3]Institute of Bioengineering and Nanotechnology (IBN), Singapore

Summary — The Lush Science Prize 2016 was awarded to Daniele Zink and Lit-Hsin Loo for the interdisciplinary and collaborative work between their research groups in developing alternative methods for the prediction of nephrotoxicity in humans. The collaboration has led to the establishment of a series of pioneering alternative methods for nephrotoxicity prediction, which includes: predictive gene expression markers based on pro-inflammatory responses; predictive *in vitro* cellular models based on pluripotent stem cell-derived proximal tubular-like cells; and predictive cellular phenotypic markers based on chromatin and cytoskeletal changes. A high-throughput method was established for chemical testing, which is currently being used to predict the potential human nephrotoxicity of ToxCast compounds in collaboration with the US Environmental Protection Agency. Similar high-throughput imaging-based methodologies are currently being developed and adapted by the Zink and Loo groups, to include other human organs and cell types. The ultimate goal is to develop a portfolio of methods accepted for the accurate prediction of human organ-specific toxicity and the consequent replacement of animal experiments.

Key words: *high-content imaging, nephrotoxicity, phenotypic profiling, predictive alternative method, renal proximal tubular cell.*

Addresses for correspondence: *Lit-Hsin Loo, Bioinformatics Institute, 30 Biopolis Street, 07-01 The Matrix, Singapore 138671; Daniele Zink, Institute of Bioengineering and Nanotechnology, 31 Biopolis Way, 04-01 The Nanos, Singapore 138669.*
E-mail: *loolh@bii.a-star.edu.sg; dzink@ibn.a-star.edu.sg*

First Steps

The work carried out by the research groups of Daniele Zink and Lit-Hsin Loo, on the development of a high-throughput (HTP) method for the prediction of nephrotoxicity in humans, was awarded the Lush Science Prize in 2016. This work began at the end of 2011. At that time, although some toxicology-related work had been performed with *in vitro*-cultured renal cells, no *in vitro* or other alternative methods existed that could accurately predict compound-induced nephrotoxicity in humans or other species (1, 2).

Zink's laboratory had been working on kidney tissue engineering (3–6), and it was an interesting scientific challenge to apply knowledge gained from these prior studies to the development of predictive renal cell-based methods for safety screening. Hence, the beginning of the work was motivated by scientific curiosity, as well as by the desire to develop more useful and widely applicable human cell-based methods, which ultimately could replace animal experiments.

The focus of the project was strictly on the development of predictive methods, as the prediction of nephrotoxicity with alternative methods was an unresolved challenge. Part of the development of a predictive method requires the collection of data from large and statistically robust numbers of compounds, for use in pre-validation studies. Therefore, the initial work concentrated on assessing large numbers of compounds, without addressing the mechanisms of each one in detail. Although detailed mechanistic studies are informative, they are usually resource-intensive and can be performed on only a few compounds at one time. Due to the small numbers of compounds that can be analysed in this manner, it is not possible to assess the commonality of the identified mechanisms, or the general predictivity of the methods based on them. For example, until 2012, only three studies involving more than 10 compounds had been performed with *in vitro* cultured renal cells (7–9), and the predictive performance of these methods had either not been determined, or was found to be poorly predictive of nephrotoxicity.

As a first step, Zink's laboratory compiled a set of more than 40 reference compounds that were known to be nephrotoxic, or not nephrotoxic, in humans. The selection was primarily based on human clinical data, and not on animal data or renal cell-based *in vitro* results (10, 11). Detailed lists of these compounds, including information on their effects on human kidneys and the respective references, have been published in the original articles relating to the project (10, 11). Human data were primarily used for compound annotation because the relevance of most of the existing *in vitro*

results for nephrotoxicity prediction was unclear. Animal data were not used because compound-induced effects on animal kidneys can be different from those effects on human kidneys. An example is phenacetin-induced nephrotoxicity. Phenacetin was a popular over the counter painkiller, until it was banned in the 1980s due to its nephrotoxicity. In rats, phenacetin damages renal proximal tubular cells (PTC; 12), whereas in humans, the typical finding is interstitial nephritis, often combined with renal papillary necrosis and calcification (13, 14). In general, rodents are not a good model for analgesic nephropathy, which is a main cause of chemical-induced kidney disease in humans (13–15).

Distinguishing between direct toxicity to PTC and other ways of damaging the kidney was important, as the initial work was based on *in vitro*-cultured human PTC. This cell type is one of the main targets for compound-induced nephrotoxicity, due to its role in the transport and metabolism of xenobiotics (16–18), and most kidney-specific *in vitro* methods use PTC (1, 2). It is expected that compounds that are not known to directly damage PTC in human kidneys, such as phenacetin, should be predicted as negatives by an *in vitro* method based on PTC. Therefore, the selection of nephrotoxic compounds was further subdivided into compounds that were known to be directly toxic to PTC in humans, and compounds that were not directly toxic to this cell type (10, 11). Compounds belonging to the latter subdivision were used as part of the negative control group, for determination of the specificity of the methods developed.

Due to the researchers' eagerness to get started in 2011, 'everything they could get' was purchased for the initial compound library. The only criterion was the availability of human clinical data with respect to nephrotoxicity and PTC toxicity. The resulting compound library was a mix of very diverse xenobiotics, which comprised different types of drugs, industrial chemicals, herbicides and natural compounds. This turned out later to be useful, because the results with this compound library revealed that the methods could predict human PTC toxicity independently of the chemical structure or the nature of a compound.

Establishing the first *in vitro* method for the accurate prediction of renal PTC toxicity in humans

In 2011, it was clear to the Zink group members that widely used *in vitro* toxicity endpoints — such as cell numbers, cell death, and adenosine triphosphate (ATP) or glutathione (GSH) depletion — could not be used for the development of a predictive renal assay, because a predictive kidney-specific *in vitro* method had not already been successfully developed, despite the widespread use of such endpoints. The limited usefulness of such widely used endpoints was later confirmed by the results obtained (10, 11). In addition, the use of potential biomarkers for acute kidney injury (AKI), such as hepatitis A virus cellular receptor 1 (HAVCR1/KIM1) or lipocalin 2 (LCN2/NGAL), in predictive *in vitro* assays, had not yielded convincing results. This was confirmed in the outcomes of the EU-funded Predict-IV project (3rd and 4th Project Periodic Reports), and by the initial results obtained in the Zink laboratory (11). The lack of predictivity of such potential novel *in vivo* biomarkers in kidney-specific *in vitro* assays was also confirmed in a recent study (19).

Therefore, other endpoints were considered. As the focus was strictly on developing a predictive method, the Zink group did not set out to follow the widely used strategy of determining which mechanisms and endpoints are specifically triggered by a particular compound, and then combining these compound-specific and mechanism-specific endpoints. Instead, the aim was to find common endpoints that would always be triggered if a compound were indeed toxic to PTC. Such endpoints would result in the best predictivity, as long as they were specifically induced by PTC-toxic compounds and not by other types of toxic compound.

Pro-inflammatory markers were good candidates, as damaged cells that do not die usually show a pro-inflammatory response (it was obvious from the literature that cell death is not consistently induced under *in vitro* conditions by PTC-toxic compounds, and this was confirmed by the results obtained by Zink's group). The group knew from previous work that *in vitro*-cultured PTC express interleukin-6 (IL-6; 5), and that this pro-inflammatory interleukin is associated with kidney damage and disease *in vivo* (20, 21). The same applies to IL-8 (now called CXCL8), which is also expressed *in vitro* (21–23). Furthermore, pro-inflammatory cytokines were known to play an important role in cisplatin-induced kidney damage (24), and IL-6 and IL-8 (CXCL8) were generally known to be induced in response to cell damage. Therefore, there was a good chance that IL-6 and IL-8 (CXCL8) would be induced in PTC, if these cells were damaged by a compound that is toxic to this cell type.

The library of more than 40 compounds was then screened with human primary PTC (HPTC) obtained from three different donors, and a widely used immortalised human PTC line (HK-2 cells). For comparison, the widely-used porcine PTC line, LLC-PK1, was also included. Expression levels of IL-6 and IL-8 (CXCL8) were measured by qPCR. The first results were obtained in November and December 2011. This phase of the work was extremely exciting, because the results showed very early on that IL-6 and IL-8 (CXCL8) were

indeed induced by compounds that were known to be toxic to PTC in humans, but not by other nephrotoxicants or compounds that were not nephrotoxic to humans. In addition, the results showed that the use of HPTC resulted in the highest predictivity (11).

A productive collaboration

The initial work in the Zink laboratory was based on data analysis with a relatively simple thresholding method (11), and the prediction performance was determined by using all the data samples, without separating them into training and test samples.

In 2012, the authors met for the first time and immediately decided to collaborate, in view of their complementary expertise [Zink is a cell biologist and Loo a computational biologist], as well as their shared determination to develop successful predictive methods and their similar philosophy of how one might do it. Loo then downloaded all of the published original data from the initial study (11) and established a machine learning method for the analysis of these data, which improved the predictivity of the method (25). Importantly, a cross-validation procedure that splits the data into training and test samples was used to properly estimate the predictivity of the IL-6/IL-8 (CXCL8) markers. Zink learned about this only when it had been confirmed by this second independent analysis that the IL-6/IL-8 (CXCL8)-based method would indeed predict PTC toxicity in humans.

The first stem cell-based methods for the accurate prediction of kidney toxicity

In the meantime, the groups of Jackie Y. Ying and D. Zink at IBN had developed the first method for the differentiation of stem cells into HPTC-like cells (26). The initial protocol was based on human embryonic stem cells, and this was then combined with the IL-6/IL-8 (CXCL8)-based method (Figure 1). This resulted in the first stem cell-based method for the prediction of PTC toxicity in humans (10).

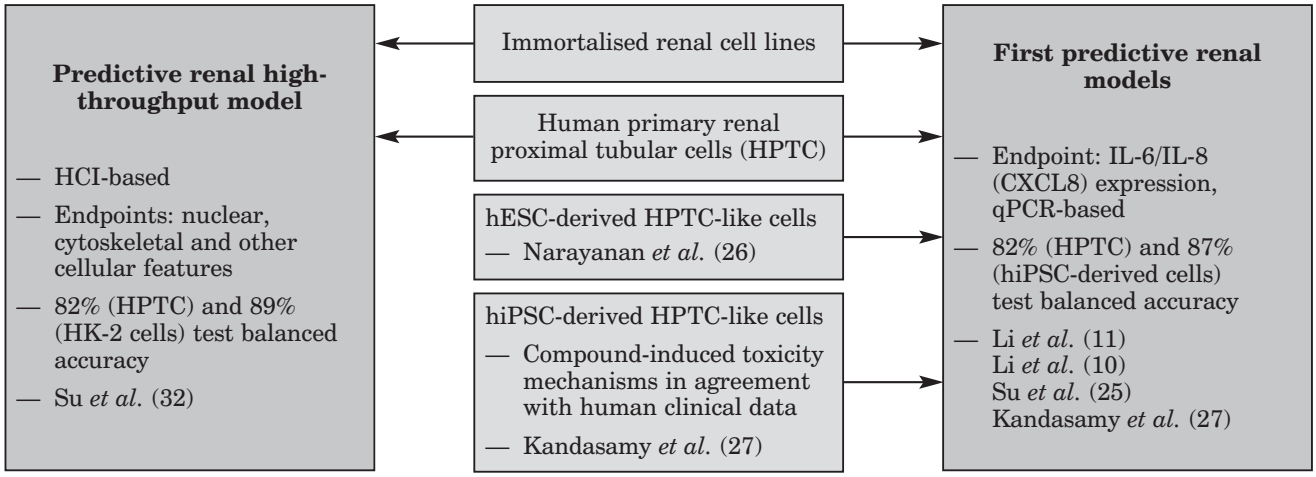
The stem cell-based protocol was further refined and adapted to human induced pluripotent stem cells (iPSC), which resulted in the currently most rapid and efficient method to generate iPSC-derived HPTC-like cells (27). HPTC-like cells with a purity of greater than 90% could be obtained with this one-step differentiation protocol in only eight days. Due to the high purity of the iPSC-derived HPTC-like cells, they could be applied immediately to compound screening. The iPSC-based protocol for generating HPTC-like cells was then combined with the IL-6/IL-8 (CXCL8)-based assay and machine learning methods, which resulted in the first iPSC-based method for the prediction of PTC toxicity in humans (27). This iPSC-based method and the hESC-based method are the only stem cell-based methods for the accurate prediction of renal toxicity in humans.

Currently, other organoid-based approaches are being developed. These methods have been tested with only one or two nephrotoxic compounds, and no non-nephrotoxic compounds were used (28–30). Hence, the predictivity of such organoid-based approaches is unknown (for a comprehensive review on stem cell-derived renal cells and their applications see Chuah & Zink [31]).

In summary, at this stage, a battery of predictive methods based on various human PTC models (primary, immortalised and stem cell-derived) was developed, and the endpoint of these methods was compound-induced expression of IL-6 and IL-8 (CXCL8) (Figure 1, top). After the initial establishment of the *in vitro* methodology, the cell-based work was complemented by computational approaches, with machine learning methods being established for compound classification and predictive performance analysis. The rapid success of the work was due to several crucial factors, which included the strict focus on developing predictive methods, the consistent screening of a statistically robust set of compounds, careful compound annotation based on human clinical data, and a fruitful interdisciplinary collaboration driven by scientists with a similar philosophy but complementary expertise. Perhaps, a quantum of luck was also required for such work, as it is not self-evident that scientists from different disciplines think along the same lines.

Another crucial factor was the acceptance that toxicity mechanisms could not be addressed in depth during the initial phases of the project; and, as shown, understanding of the mechanisms was not necessarily required in order to develop a predictive method (Figure 2). However, it was important to furnish the methods subsequently with mechanistic understanding (Figure 2). Thus, the underlying mechanisms were investigated, in order to ensure that the cell models showed appropriate responses to nephrotoxicants, and that the results were concordant with human clinical data. In the course of the iPSC-based work, it was demonstrated that the PTC-damaging nephrotoxicants, cisplatin and aristolochic acid, induced reactive oxygen species (ROS) generation and lipid peroxidation in HPTC-like cells, as well as a DNA damage response and nuclear translocation of RELA (NF-κB p65), in concordance with human clinical data and other previous findings (27). Such responses were not induced by acarbose and ethylene glycol, which are also toxic in humans, but are not directly toxic to PTC (27). In addition, it was shown that the induction of IL-6 and IL-8 (CXCL8) by PTC-damaging compounds was

Figure 1: The methods developed for the accurate prediction of renal PTC toxicity in humans



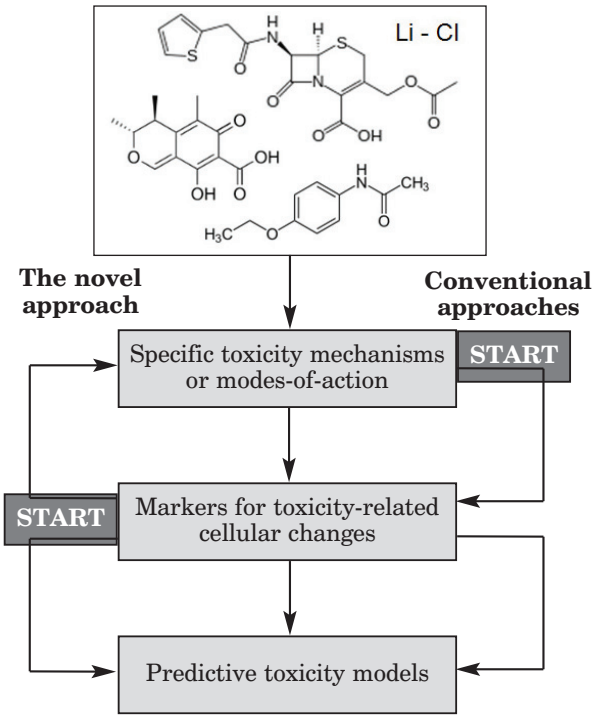
Several methods for the prediction of renal PTC toxicity in humans have been developed by the authors' research groups. The first methods for the accurate prediction of PTC toxicity in humans (First predictive renal models; right) were based on the detection of compound-induced expression of IL-6 and IL-8 (CXCL8). These methods employed different human PTC models (middle), such as immortalised and human primary PTC (HPTC), and pluripotent stem cell-derived HPTC-like cells. For each cell model, the predictivity of the method was determined, and the respective performance values and references were provided. The predictive renal HTP method (left) was established by using immortalised human PTC (HK-2 cells) and HPTC. This method combines high-content imaging (HCI) with phenotypic profiling and machine learning methods (see Figure 4).

dependent on transporters that mediate the uptake of these compounds (27). However, the priorities of this particular project were, from the outset, different compared to other approaches, which try to build predictive methods based on mechanistic understanding (Figure 2). As the achievement of a deep mechanistic understanding is a process that requires many years of work, and the development of a predictive method would occur only after that point, these conventional approaches are slow. However, there is a current urgent need for more and better alternative methods, in order to significantly reduce and replace animal experiments. Therefore, the strategy described here — to develop a predictive method first and elucidate the mechanisms or modes-of-action later — could greatly speed up the development process of alternative methods (Figure 2). Recent work from the authors on the development of a HTP method (described in the following section) has shown that this unbiased strategy might even lead to unexpected and novel mechanistic insights into cellular injury mechanisms or modes-of-action.

High-throughput Prediction of PTC Toxicity in Humans

After establishing the first generation of predictive kidney-specific methods (Figure 1; top), the authors went on to establish the first high-throughput (HTP) platform for the prediction of PTC toxicity (32). As in the previous assays, this method was based on differentiated epithelia of PTC cultured in multi-well plates. The reason for choosing flat epithelia of PTC over more complex culture systems was the need to establish robust and relatively simple assays that could be easily carried out by many different users, including industry and environmental agencies. Methods that cannot be widely applied cannot substantially contribute to the replacement of animal studies. In addition, reproducibility is an important aspect, and high reproducibility cannot currently be achieved with more-complex culture systems. For instance, kidney organoids consist of 3-D structures with variable sizes and cell numbers, and complex and variable cellular compositions. In a previous study, about 80% of the cells in kidney organoids that were stained with *Lotus tetragonolobus* lectin (LTL; which specifically binds to PTC), also expressed podocyte (another renal cell type) and endothelial markers (28). The rest of the LTL-positive kidney organoids had different cellular compositions. Furthermore, only about two thirds of the generated 3-D structures represented kidney organoids; the remaining third had neuro-ectodermal characteristics (28). Renal organoid-based methods are also slow and costly, and currently they are not ready for testing large numbers of compounds. So far, kidney organoids were tested with only one or two nephro-

Figure 2: A novel strategy for building predictive human toxicity models, starting from phenotypic markers



Conventional approaches for building predictive human toxicity models start with the determination of toxicity mechanisms, or modes-of-action, that are specific for the compounds under consideration. Based on the results of this first step, specific cellular markers are then determined that are particular to these mechanisms or modes-of-action. Loo and Zink's novel approach starts directly with the determination of specific phenotypic markers by using image-based phenotypic profiling. Predictive models are then constructed, by using these identified markers. After the phenotypic markers are found, the underlying toxicity mechanisms, or modes-of-action, of the compounds are analysed. This approach can allow unbiased detection of novel, or unexpected, modes-of-action of the compounds, as the determination of phenotypic markers does not depend on expected mechanisms or modes-of-action.

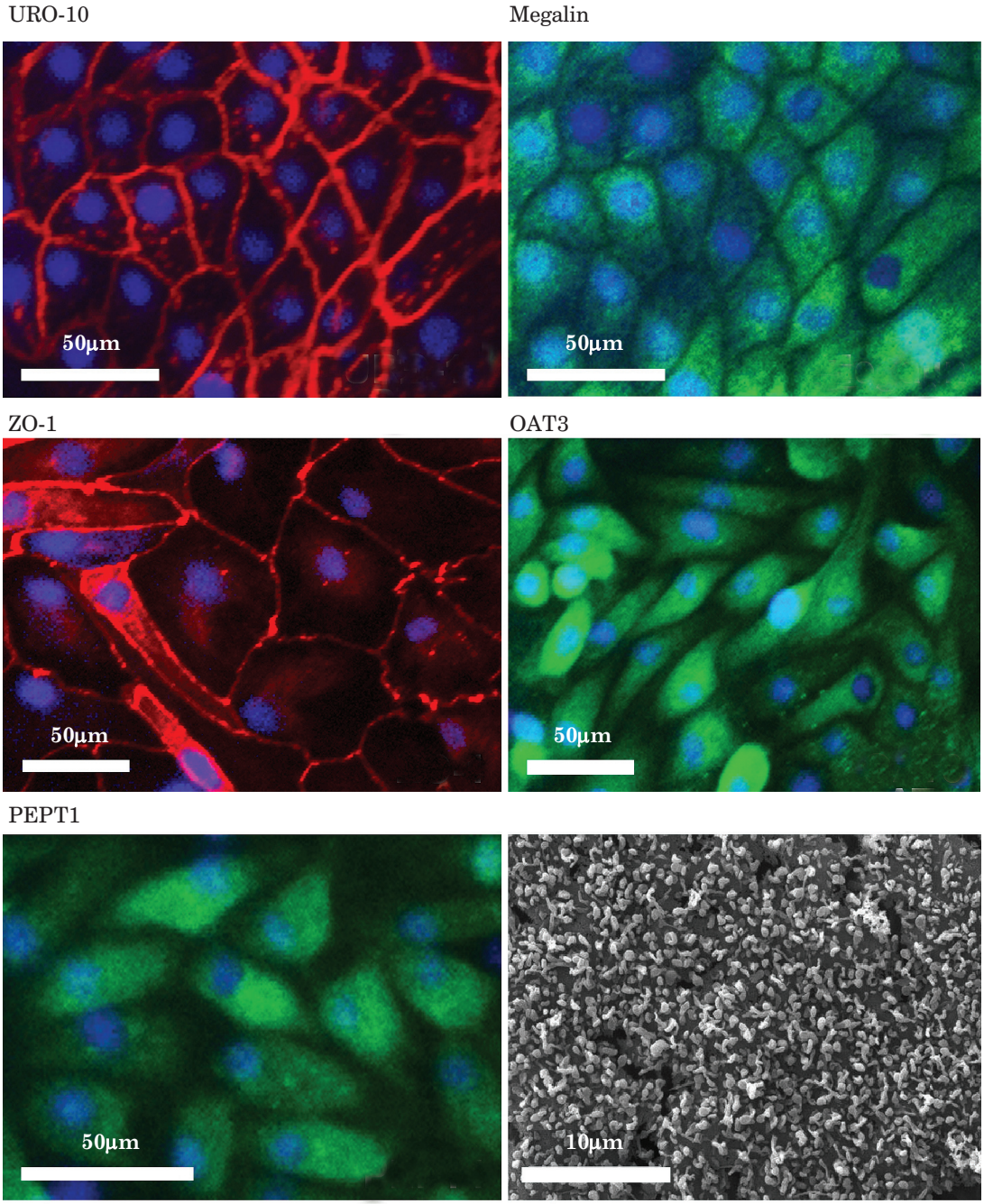
toxic compounds at a time, and no negative control compounds were used (28–30). Thus, apart from the reproducibility issues, the predictivity of organoid-based methods is also unknown, and the development of reproducible and predictive screening methods based on kidney organoids will likely remain difficult, at least in the near future.

It is also important to consider that cells might behave differently under *in vitro* conditions, and this applies also to cells in kidney organoids. This is not surprising, because organoids are still far from representing functional normal kidney tissue (33). For instance, in a previous study, where kidney organoids were treated with the PTC-damaging nephrotoxicant cisplatin, apoptosis was induced in cells that were positive for LTL and E-cadherin staining (30). Under *in vivo* conditions, PTC (which are stained by LTL) express predominantly N-cadherin, but distal tubules and other parts of the kidney express E-cadherin (34). Thus, in these kidney organoids, cells with features that do not exist *in vivo* (namely, being simultaneously positive for LTL and E-cadherin staining), were more sensitive to cisplatin. In contrast, cells with a staining pattern that is displayed by PTC *in vivo* (LTL^+ and $E\text{-cadherin}^-$) did not undergo apoptosis (30). How to properly interpret these results remains to be determined.

Another important consideration was the suitability of 2-D or 3-D cultures for PTC-based *in vitro* models. *In vivo*, PTC form a differentiated simple epithelium that consists of a single cell layer on a basal membrane. This form of organisation is rather similar to a differentiated epithelium that consists of a single cell layer on a cell culture substrate, as applied in these predictive models. It should be noted that not only 2-D monolayers were used, but confluent polarised epithelia — which have been shown to express a broad range of markers that are characteristic of differentiated PTC (Figure 3). Furthermore, they display tight junctions and apical brush borders (Figure 3) with typical activities of brush border enzymes, such as γ-glutamyl transferase (3, 11, 27). Similar results from such PTC epithelia generated *in vitro* were also shown in the authors' previous publications (3, 11, 27).

Based on the considerations outlined above, it was decided to proceed with the well-characterised and differentiated epithelia of PTC cultivated in multi-well plates, for the development of a HTP model for the binary (yes/no) prediction of PTC toxicity. A slightly modified library of more than 40 compounds was used for the pre-validation stage, and compound screening was performed by high-content imaging (HCI) at a dose range of up to 1mM. As mentioned, the goal was to develop a model for the binary prediction of PTC toxicity, and, in this respect, the dose range at which compounds would be tested could be freely chosen, as long as the specificity remained high and non-PTC-toxic compounds continued to produce negative results. It is important to note that organ-specific cell models cultured under *in vitro* conditions usually display relatively low expression levels of transporters and xenobiotic-metabolising enzymes (35–38). Therefore, their dose-responses are typically different from the dose-responses observed *in vivo* (9, 39–41). Indeed, it would be rather surprising if the cells still responded to the same dose ranges as they would in the body. It is likely that the dose range tested *in vitro* would need to be adapted to the altered cellular dose-response, in order to achieve a reasonable sensitivity. It is not

Figure 3: *In vitro*-generated differentiated PTC epithelia



Confluent simple epithelia of human primary PTC are shown after immunostaining of the indicated markers (red and green), and epifluorescence imaging of cell nuclei (blue). The following markers, which are typically expressed by PTC, were detected: urothelial surface glycoprotein-10 (URO-10); low density lipoprotein receptor-related protein 2 (megalin); zonula occludens-1 (ZO-1); solute carrier family 22 (organic anion transporter), member 8 (OAT3/SLC22A8); and solute carrier family 15 (oligopeptide transporter), member 1 (PEPT1/SLC15A1). The bottom-right panel displays the apical brush border of the polarised epithelium after imaging by scanning electron microscopy (SEM).

useful to test within a physiological dose range, if the *in vitro* cells do not respond at these concentrations. It is possible that the usage of inappropriately low *in vitro* test concentrations is often one of the major problems in the development of predictive *in vitro* methods.

It should also be noted that 3-D models, including PTC-based 3-D models, are known to not respond within the physiological dose range. For instance, in case of gentamicin, the IC50 value obtained with a gel-based 3-D model was 9mM (4298μg/ml; 42), whereas the human C_{max} value is in the range of 0.03–0.04mM (16–20μg/ml; 43), and gentamicin has substantial nephrotoxicity already at therapeutic doses. In another recently developed PTC-based 3-D model, significant responses to gentamicin were obtained only at concentrations of 2.5mM (mRNA level) and 5mM (protein level), respectively (19). The process of obtaining responses within the physiological dose range, and the subsequent *in vitro*–*in vivo* extrapolation, remain challenging, and improvements in this area are required in order to achieve full replacement of animal experiments.

Currently, the authors' *in vitro* methods for the prediction of compound toxicity give only binary predictions ('toxic' *versus* 'non-toxic'), but even with such a simplification of the toxicity prediction problem, the task is still very challenging and often not achieved. It is evident that models for the binary prediction of compound toxicity cannot be used for predicting quantitative features of the human dose-response, such as No Observed Adverse Effect Levels (NOAELs).

It might be expected that, in the end, different types of *in vitro* and *in silico* models will have to be combined in order to cover all steps of risk assessment, which would be required to fully replace animal experiments. Nevertheless, the rapid and efficient binary prediction of organ-specific toxicity already represents substantial progress, compared to having either no clue at all, or being wholly dependent on animal experiments. In many cases, a binary prediction would be sufficient to narrow down the numbers of compounds that would require a more comprehensive characterisation.

Another point of consideration was the selection of endpoints for the HTP method — this problem was addressed by phenotypic profiling, and the approach is outlined in the following section.

Phenotypic profiling

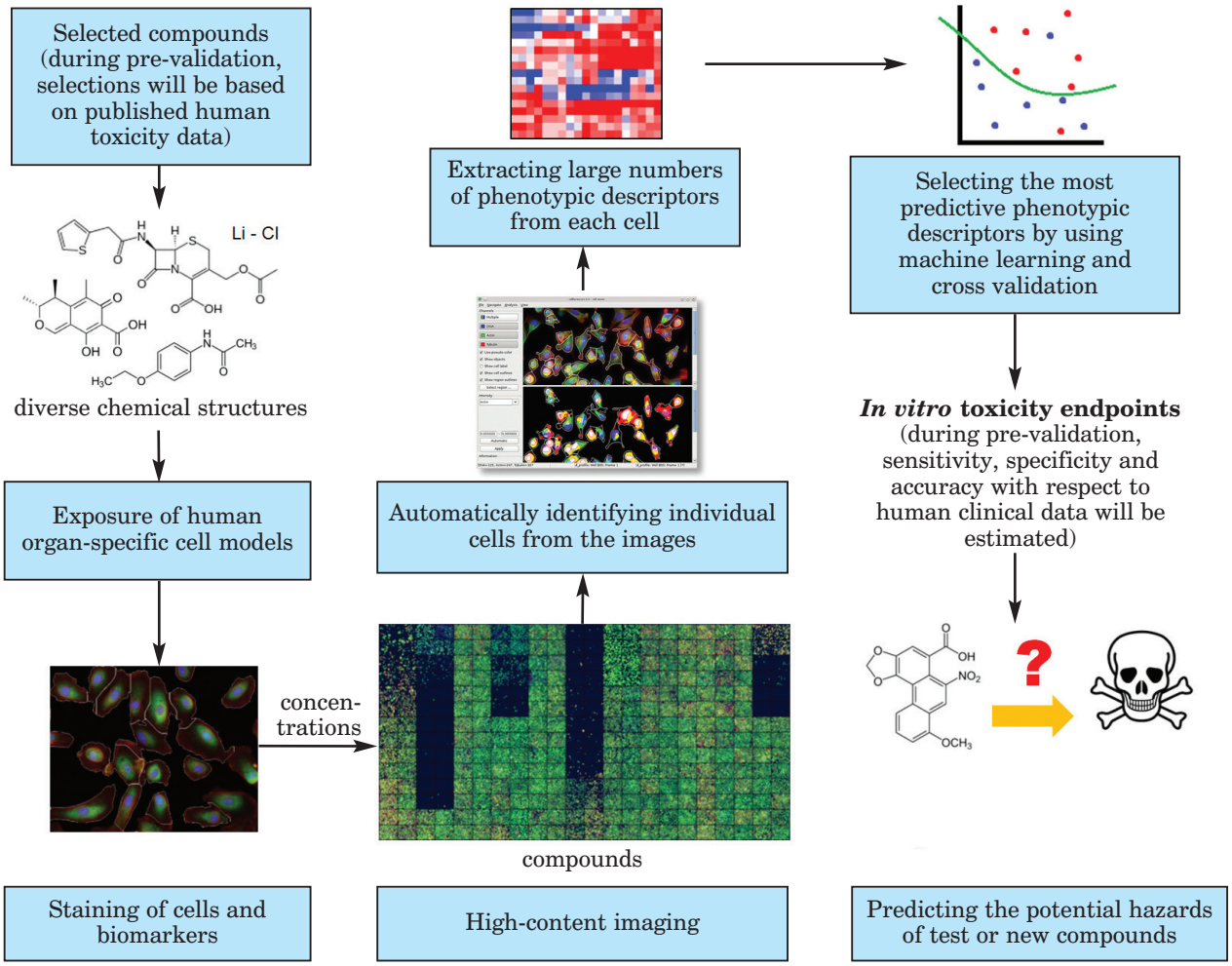
Toxic compounds may induce specific changes in cellular phenotypes, including induction, modification or translocation of intracellular biomolecules, reorganisation of cytoskeletal components, or changes in the structures or morphologies of cellular organelles. Therefore, quantitative descriptors of cellular phenotypes captured on microscopy images could potentially be used as *in vitro* toxicity endpoints. One of the challenges is how to quantify and search for the most predictive endpoints in an unbiased manner.

Phenotypic profiling is an approach used to construct quantitative representations (or 'profiles') of cellular phenotypes based on the microscopy images collected in HTP screening experiments (44). The approach usually consists of two main steps (Figure 4). The first step is to measure large numbers (in the order of hundreds to thousands) of phenotypic descriptors for every cell in the collected images. The second step is to use machine learning methods to automatically identify the most predictive descriptors. Conceptually, phenotypic profiling is very similar to genomic profiling, in which the expression levels of large numbers of genes are measured, and then the most predictive gene expression patterns are identified. Similarly, phenotypic profiling does not make *a priori* assumptions about the injury mechanisms or modes-of-action, and thus novel mechanisms or modes-of-action could be discovered (44). Therefore, phenotypic profiling is different from conventional high-content analysis, in which small numbers of readouts for pre-defined injury mechanisms are used (39, 41, 45).

Loo's laboratory has recently been working on the development of phenotypic profiling technologies for HTP imaging experiments (46). Previously, Loo and colleagues had shown that image-based phenotypic profiling is sufficient to automatically classify large numbers of small molecules with different targets and mechanisms (47), predict the biological functions of proteins based on their subcellular localisation patterns (48), and predict cellular sensitivity to cytotoxic agents and co-treatments (49). Therefore, in 2014, when steps were initiated to bring the methods for predicting PTC toxicity to the next level by establishing a HTP approach, it was immediately clear that these phenotypic profiling technologies could also be applied to the construction of a highly predictive model. At that time, high-content analysis based on standard endpoints was being used, but phenotypic profiling had not been applied to the building of *in vitro* toxicity models.

After screening the compound library by using HCI, phenotypic profiling was used to measure changes in 129 cellular features, and identify the features that were most predictive of the *in vivo* PTC toxicity of the screened compounds (Figure 5a). The mechanisms covered by the phenotypic endpoints identified were ultimately limited by the fluorescence markers used for the cellular staining in the imaging experiments. Initially, the cell nuclei (4′,6-diamidino-2-phenylindole [DAPI]), entire cells (whole cell stain), F-actin and RELA, were stained. F-actin and RELA were selected as markers for the compound-induced cytoskeletal

Figure 4: A HTP method for nephrotoxicity prediction, based on phenotypic profiling



The method starts with the selection of compounds for testing. During pre-validation of the model, reference compounds based on published human toxicity data are selected. Then, human organ-specific cell models are exposed to the selected compounds, and stained to detect specific biomarkers. The stained cells are imaged by using a high-content imaging system, and automatically identified with computational algorithms. During pre-validation, phenotypic profiling is used to measure large numbers of phenotypic descriptors from each cell, and to select the most predictive descriptors based on published human toxicity data. In its subsequent test application, only the final selected descriptors need to be measured in the cells, and these descriptors are used as toxicity endpoints to predict the potential hazards of new compounds.

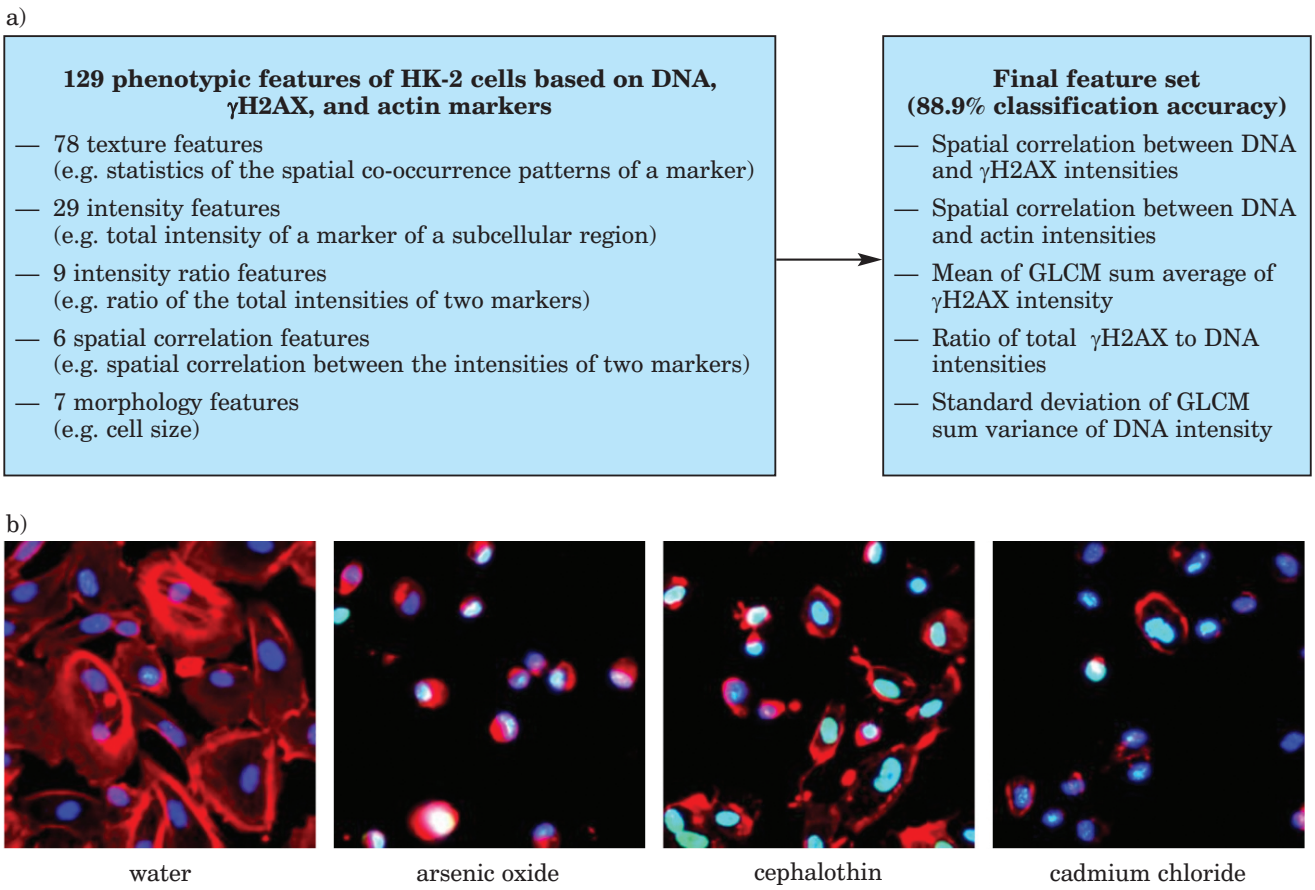
changes and pro-inflammatory responses that were consistently observed in the authors' previous studies. However, after performing phenotypic profiling, it turned out that changes in DNA arrangements were highly predictive for PTC toxicity in humans. Based on these results, it was postulated that a DNA damage response could be the underlying mechanism, and thus another round of screening was performed with three batches of HPTC and HK-2 cells, replacing RELA with a DNA damage marker, γH2AX. It turned out that γH2AX expression was consistently and specifically increased by compounds that are toxic to PTC in humans, and it was apparent that the inclusion of γH2AX as a marker resulted in higher accuracy prediction. The final sets of predictive features selected consisted of a number of chromatin and cytoskeleton features (e.g. spatial correlation between DNA and γH2AX intensities; see Figure 5a). Five features were selected for HK-2 cells, and four for HPTC. Random Forest classifiers based on these features could achieve test sensitivities of 98.8% (HK-2 cells) and 83.7% (HPTC), test specificities of 79.0% (HK-2) and 79.5% (HPTC), and test balanced accuracies of 88.9% (HK-2) and 81.6% (HPTC).

All of the performance values were estimated by using a 10-fold cross-validation procedure, which

uses independent samples to train or test the classifiers (32). The procedure is recommended by the Organisation for Economic Co-operation and Development (OECD) for validating predictive models for chemical safety assessment (OECD Environment Directorate, Guidance Document on the Validation of [Quantitative] Structure–Activity Relationship [(Q)SAR] Models, Environment Health and Safety Publications Series on Testing and Assessment, No. 69; 50). In many previous studies, the classifiers (or decision thresholds) used were determined based on all the samples, and the sensitivity and specificity of the classifiers were either not reported (45) or determined based on the same set of samples (8, 39). Therefore, these reported values were not obtained from independent test samples, and thus were not directly comparable to the performance values obtained for the model described here.

One of the most surprising findings from the HTP study was that DNA changes and the DNA damage marker were consistently induced, although not all PTC-toxic compounds that were included were known to have DNA-damaging properties. In the meantime, by using comet assays, it was confirmed that the observed changes were indeed true DNA damage, and that they seemed to be due to ROS generation in the case of compounds not known to damage DNA directly (K.G. Eng and D. Zink, unpublished results). This occurs already at low compound concentrations, and seems to be PTC-specific, as a similar response in other organ-

Figure 5: The predictive cellular phenotypic features for nephrotoxicity



a) A list of 129 phenotypic features of human HK-2 cells were measured, based on DNA, γH2AX and actin markers. Five categories of features, measuring different aspects of cellular phenotypes and marker subcellular localisation patterns, were used. An automated feature selection algorithm was then used to identify a final set of five features that gave 88.9% classification accuracy. b) Examples of immunofluorescence images showing HK-2 cells that were treated with three nephrotoxic compounds, namely: arsenic oxide; cephalothin; cadmium chloride; and the solvent control (i.e. water); blue = DNA, green = γH2AX, red = actin. The cells treated with the nephrotoxic compounds had increased nuclear activation of γH2AX, which was represented by: the 'spatial correlation between the DNA and γH2AX intensities' feature; disrupted actin cytoskeleton, which was represented by the 'spatial correlation between DNA and actin intensities' feature; and other phenotypic changes, represented by the other three selected features. GLCM = Grey-Level Co-occurrence Matrix.

specific cell types, such as hepatocytes, has not been observed.

These results illustrate how phenotypic profiling can reveal unexpected mechanistic insights, that might not have been detected if the method had been based on expected mechanisms and predetermined endpoints. In addition to providing mechanistic insights in an unbiased way, this method permits the development of predictive HTP methods without *a priori* knowledge of compound-induced mechanisms (Figure 2). In this way, the development of predictive HTP methods and the adaption of the existing technology to other organs and cell types can be rapidly achieved. HTP methods that do not depend on *a priori* knowledge of compound-induced mechanisms could be particularly useful for the classification of poorly characterised chemicals, which is still a challenging problem for regulators and governmental agencies. They could also be useful for early screening of novel compound libraries during drug discovery.

Conclusions and Outlook

This renal HTP method (32) is the first and only method that allows the prediction of kidney-specific toxicity in a HTP manner. The method is currently undergoing further pre-validation in collaboration with the US Environmental Protection Agency; the goal of this research collaboration is to predict the human PTC toxicity of hundreds of ToxCast compounds. Due to the lack of efficient predictive screening technologies, the nephrotoxicity of most ToxCast compounds is unknown. Collaborations with other environmental agencies are also planned.

One of the authors' current goals is to develop a portfolio of similar HTP methods specific for other organs. This would allow a comprehensive prediction of organ-specific toxicity in humans. As phenotypic profiling does not depend on fixed endpoints, the developed method can be rapidly adapted to other cell types and organs. Similar predictive HTP methods that are specific for liver, lung and vasculature cells have recently been established, and respective publications are in preparation.

The most favourable outcome from the work described here would be for the methods to be adopted by academia, industry and governmental agencies. The desire to develop high-quality and robust predictive methods that can be applied by many different users to refine, reduce and ultimately replace animal experiments has been a major driving force behind this work. The authors were delighted and most grateful that the renal HTP method was awarded the Lush Science Prize 2016, not least because it helps to raise awareness of this novel technology. The Zink and Loo groups are currently working toward regulatory acceptance of the methods, and this is the next crucial challenge.

Acknowledgements

The authors thank all members of the Zink and Loo groups who contributed to the work and are cited in the references. We are particularly grateful to Dr Su Ran (BII) and Dr Sijing Xiong (IBN), who contributed to the development of the HTP method. We thank Dr Karthikeyan Kandasamy (IBN), Dr Farah Tasnim (IBN), Dr Rensheng Deng (IBN) and Dr Yao Li (IBN) for providing images. The work was supported by a grant from the Joint Council Office (Agency for Science, Technology and Research [A*STAR]) Development Program, the Institute of Bioengineering and Nanotechnology, and the Bioinformatics Institute (Biomedical Research Council, A*STAR, Singapore).

References

1. Davies, J. (2014). Engineered renal tissue as a potential platform for pharmacokinetic and nephrotoxicity testing. *Drug Discovery Today* **19**, 725–729.
2. Tiong, H.Y., Huang, P., Xiong, S., Li, Y., Vathsala, A. & Zink, D. (2014). Drug-induced nephrotoxicity: Clinical impact and preclinical *in vitro* models. *Molecular Pharmaceutics* **11**, 1933–1948.
3. Ni, M., Teo, J.C., Ibrahim, M.S., Zhang, K., Tasnim, F., Chow, P.Y., Zink, D. & Ying, J.Y. (2011). Characterization of membrane materials and membrane coatings for bioreactor units of bioartificial kidneys. *Biomaterials* **32**, 1465–1476.
4. Oo, Z.Y., Deng, R., Hu, M., Ni, M., Kandasamy, K., Bin Ibrahim, M.S., Ying, J.Y. & Zink, D. (2011). The performance of primary human renal cells in hollow fiber bioreactors for bioartificial kidneys. *Biomaterials* **32**, 8806–8815.
5. Oo, Z.Y., Kandasamy, K., Tasnim, F. & Zink, D. (2013). A novel design of bioartificial kidneys with improved cell performance and haemocompatibility. *Journal of Cellular & Molecular Medicine* **17**, 497–507.
6. Zhang, H., Tasnim, F., Ying, J.Y. & Zink, D. (2009). The impact of extracellular matrix coatings on the performance of human renal cells applied in bioartificial kidneys. *Biomaterials* **30**, 2899–2911.
7. Duff, T., Carter, S., Feldman, G., McEwan, G., Pfaller, W., Rhodes, P., Ryan, M. & Hawksworth, G. (2002). Transepithelial resistance and inulin permeability as endpoints in *in vitro* nephrotoxicity testing. *ATLA* **30**, Suppl. 2, 53–59.
8. Lin, Z. & Will, Y. (2012). Evaluation of drugs with specific organ toxicities in organ-specific cell lines. *Toxicological Sciences* **126**, 114–127.
9. Wu, Y., Connors, D., Barber, L., Jayachandra, S., Hanumegowda, U.M. & Adams, S.P. (2009). Multiplexed assay panel of cytotoxicity in HK-2 cells for detection of renal proximal tubule injury potential of compounds. *Toxicology in Vitro* **23**, 1170–1178.
10. Li, Y., Kandasamy, K., Chuah, J.K.C., Lam, Y.N.,

Toh, W.S., Oo, Z.Y. & Zink, D. (2014). Identification of nephrotoxic compounds with embryonic stem cell-derived human renal proximal tubular-like cells. *Molecular Pharmaceutics* **11**, 1982–1990.
11. Li, Y., Oo, Z.Y., Chang, S.Y., Huang, P., Eng, K.G., Zeng, J.L., Kaestli, A.J., Gopalan, B., Kandasamy, K., Tasnim, F. & Zink, D. (2013). An *in vitro* method for the prediction of renal proximal tubular toxicity in humans. *Toxicology Research* **2**, 352–362.
12. Plummer, D.T., Leathwood, P.D. & Blake, M.E. (1975). Urinary enzymes and kidney damage by aspirin and phenacetin. *Chemico-Biolological Interactions* **10**, 277–284.
13. Guo, X. & Nzerue, C. (2002). How to prevent, recognize, and treat drug-induced nephrotoxicity. *Cleveland Clinic Journal of Medicine* **69**, 289–290, 293–294, 296–297 passim.
14. Delzell, E. & Shapiro, S. (1998). A review of epidemiologic studies of nonnarcotic analgesics and chronic renal disease. *Medicine (Baltimore)* **77**, 102–121.
15. Schnellmann, R.G. (1998). Analgesic nephropathy in rodents. *Journal of Toxicology & Environmental Health Part B: Critical Reviews* **1**, 81–90.
16. Anzai, N., Jutabha, P., Kanai, Y. & Endou, H. (2005). Integrated physiology of proximal tubular organic anion transport. *Current Opinion in Nephrology & Hypertension* **14**, 472–479.
17. Lohr, J.W., Willsky, G.R. & Acara, M.A. (1998). Renal drug metabolism. *Pharmacological Reviews* **50**, 107–141.
18. Morrissey, K.M., Stocker, S.L., Wittwer, M.B., Xu, L. & Giacomini, K.M. (2013). Renal transporters in drug development. *Annual Review of Pharmacology & Toxicology* **53**, 503–529.
19. Adler, M., Ramm, S., Hafner, M., Muhlich, J.L., Gottwald, E.M., Weber, E., Jaklic, A., Ajay, A.K., Svoboda, D., Auerbach, S., Kelly, E.J., Himmelfarb, J. & Vaidya, V.S. (2016). A quantitative approach to screen for nephrotoxic compounds *in vitro*. *Journal of the American Society of Nephrology* **27**, 1015–1028.
20. Grigoryev, D.N., Liu, M., Hassoun, H.T., Cheadle, C., Barnes, K.C. & Rabb, H. (2008). The local and systemic inflammatory transcriptome after acute kidney injury. *Journal of the American Society of Nephrology* **19**, 547–558.
21. Tramma, D., Hatzistylianou, M., Gerasimou, G. & Lafazanis, V. (2012). Interleukin-6 and interleukin-8 levels in the urine of children with renal scarring. *Pediatric Nephrology* **27**, 1525–1530.
22. Niemir, Z.I., Stein, H., Ciechanowicz, A., Olejniczak, P., Dworacki, G., Ritz, E., Waldherr, R. & Czekalski, S. (2004). The *in situ* expression of interleukin-8 in the normal human kidney and in different morphological forms of glomerulonephritis. *American Journal of Kidney Diseases* **43**, 983–998.
23. Gerritsma, J.S., Hiemstra, P.S., Gerritsen, A.F., Prodjosudjadi, W., Verweij, C.L., Van Es, L.A. & Daha, M.R. (1996). Regulation and production of IL-8 by human proximal tubular epithelial cells *in vitro*. *Clinical & Experimental Immunology* **103**, 289–294.
24. Akcay, A., Nguyen, Q. & Edelstein, C.L. (2009). Mediators of inflammation in acute kidney injury. *Mediators of Inflammation* **2009**, 137072.
25. Su, R., Li, Y., Zink, D. & Loo, L.H. (2014). Supervised prediction of drug-induced nephrotoxicity based on interleukin-6 and -8 expression levels. *BMC Bioinformatics* **15**, Suppl. 16, S16.
26. Narayanan, K., Schumacher, K.M., Tasnim, F., Kandasamy, K., Schumacher, A., Ni, M., Gao, S., Gopalan, B., Zink, D. & Ying, J.Y. (2013). Human embryonic stem cells differentiate into functional renal proximal tubular-like cells. *Kidney International* **83**, 593–603.
27. Kandasamy, K., Chuah, J.K., Su, R., Huang, P., Eng, K.G., Xiong, S., Li, Y., Chia, C.S., Loo, L.H. & Zink, D. (2015). Prediction of drug-induced nephrotoxicity and injury mechanisms with human induced pluripotent stem cell-derived cells and machine learning methods. *Scientific Reports* **5**, 12337.
28. Freedman, B.S., Brooks, C.R., Lam, A.Q., Fu, H., Morizane, R., Agrawal, V., Saad, A.F., Li, M.K., Hughes, M.R., Werff, R.V., Peters, D.T., Lu, J., Baccei, A., Siedlecki, A.M., Valerius, M.T., Musunuru, K., McNagny, K.M., Steinman, T.I., Zhou, J., Lerou, P.H. & Bonventre, J.V. (2015). Modelling kidney disease with CRISPR-mutant kidney organoids derived from human pluripotent epiblast spheroids. *Nature Communications* **6**, 8715.
29. Morizane, R., Lam, A.Q., Freedman, B.S., Kishi, S., Valerius, M.T. & Bonventre, J.V. (2015). Nephron organoids derived from human pluripotent stem cells model kidney development and injury. *Nature Biotechnology* **33**, 1193–1200.
30. Takasato, M., Er, P.X., Chiu, H.S., Maier, B., Baillie, G.J., Ferguson, C., Parton, R.G., Wolvetang, E.J., Roost, M.S., Chuva de Sousa Lopes, S.M. & Little, M.H. (2015). Kidney organoids from human iPS cells contain multiple lineages and model human nephrogenesis. *Nature, London* **526**, 564–568.
31. Chuah, J.K. & Zink, D. (2017). Stem cell-derived kidney cells and organoids: Recent breakthroughs and emerging applications. *Biotechnology Advances* **35**, 150–167.
32. Su, R., Xiong, S., Zink, D. & Loo, L.H. (2016). High-throughput imaging-based nephrotoxicity prediction for xenobiotics with diverse chemical structures. *Archives of Toxicology* **90**, 2793–2808.
33. Davies, J.A. (2015). Biological techniques: Kidney tissue grown from induced stem cells. *Nature, London* **526**, 512–513.
34. Nouwen, E.J., Dauwe, S., van der Biest, I. & De Broe, M.E. (1993). Stage- and segment-specific expression of cell-adhesion molecules N-CAM, A-CAM, and L-CAM in the kidney. *Kidney International* **44**, 147–158.
35. Gomez-Lechon, M.J., Tolosa, L., Conde, I. & Donato, M.T. (2014). Competency of different cell models to predict human hepatotoxic drugs. *Expert Opinion on Drug Metabolism & Toxicology* **10**, 1553–1568.
36. Bens, M. & Vandewalle, A. (2008). Cell models for studying renal physiology. *Pflugers Archiv* **457**, 1–15.
37. Jenkinson, S.E., Chung, G.W., van Loon, E., Bakar, N.S., Dalzell, A.M. & Brown, C.D. (2012). The limitations of renal epithelial cell line HK-2 as a model of drug transporter expression and function in the proximal tubule. *Pflugers Archiv* **464**, 601–611.
38. Weiland, C., Ahr, H.J., Vohr, H.W. & Ellinger-Ziegelbauer, H. (2007). Characterization of primary rat proximal tubular cells by gene expression analysis. *Toxicology in Vitro* **21**, 466–491.

39. O'Brien, P.J., Irwin, W., Diaz, D., Howard-Cofield, E., Krejsa, C.M., Slaughter, M.R., Gao, B., Kaludercic, N., Angeline, A., Bernardi, P., Brain, P. & Hougham, C. (2006). High concordance of drug-induced human hepatotoxicity with *in vitro* cytotoxicity measured in a novel cell-based model using high content screening. *Archives of Toxicology* **80**, 580–604.
40. Ware, B.R., Berger, D.R. & Khetani, S.R. (2015). Prediction of drug-induced liver injury in micropatterned co-cultures containing iPSC-derived human hepatocytes. *Toxicological Sciences* **145**, 252–262.
41. Xu, J.J., Henstock, P.V., Dunn, M.C., Smith, A.R., Chabot, J.R. & de Graaf, D. (2008). Cellular imaging predictions of clinical drug-induced liver injury. *Toxicological Sciences* **105**, 97–105.
42. DesRochers, T.M., Suter, L., Roth, A. & Kaplan, D.L. (2013). Bioengineered 3D human kidney tissue, a platform for the determination of nephrotoxicity. *PLoS One* **8**, e59219.
43. Goncalves-Pereira, J., Martins, A. & Povoa, P. (2010). Pharmacokinetics of gentamicin in critically ill patients: Pilot study evaluating the first dose. *Clinical Microbiology & Infection* **16**, 1258–1263.
44. Bougen-Zhukov, N., Loh, S.Y., Lee, H.K. & Loo, L.H. (2017). Large-scale image-based screening and profiling of cellular phenotypes. *Cytometry Part A* **91**, 115–125.
45. Radio, N.M., Breier, J.M., Shafer, T.J. & Mundy, W.R. (2008). Assessment of chemical effects on neurite outgrowth in PC12 cells using high content screening. *Toxicological Sciences* **105**, 106–118.
46. Laksameethanasan, D., Tan, R., Toh, G. & Loo, L.H. (2013). cellXpress: A fast and user-friendly software platform for profiling cellular phenotypes. *BMC Bioinformatics* **14**, Suppl. 16, S4.
47. Loo, L.H., Wu, L.F. & Altschuler, S.J. (2007). Image-based multivariate profiling of drug responses from single cells. *Nature Methods* **4**, 445–453.
48. Loo, L.H., Laksameethanasan, D. & Tung, Y.L. (2014). Quantitative protein localization signatures reveal an association between spatial and functional divergences of proteins. *PLoS Computational Biology* **10**, e1003504.
49. Loo, L.H., Bougen-Zhukov, N.M. & Tan, W.C. (2017). Early spatiotemporal-specific changes in intermediate signals are predictive of cytotoxic sensitivity to TNFα and co-treatments. *Scientific Reports* **7**, 43541.
50. OECD (2007). *OECD Environment Health and Safety Publications: Series on Testing and Assessment No. 69. Guidance Document on the Validation of (Quantitative) Structure–Activity Relationship [(Q)SAR] Models. ENV/JM/MONO(2007)*2, 154pp. Paris, France: Organisation for Economic Co-operation & Development.