

Combination of multiple neural crest migration assays to identify environmental toxicants from a proof-of-concept chemical library

Johanna Nyffeler^{1,2} · Xenia Dolde^{1,3} · Alice Krebs^{1,3} · Kevin Pinto-Gil⁴ · Manuel Pastor⁴ · Mamta Behl⁵ · Tanja Waldmann¹ · Marcel Leist^{1,2,3}

Abstract Many in vitro tests have been developed to screen for potential neurotoxicity. However, only few cell function-based tests have been used for comparative screening, and thus experience is scarce on how to confirm and evaluate screening hits. We addressed these questions for the neural crest cell migration test (cMINC). After an initial screen, a hit follow-up strategy was devised. A library of 75 compounds *plus* internal controls (NTP80-list), assembled by the National Toxicology Program of the USA (NTP) was used. It contained some known classes of (developmental) neurotoxic compounds. The primary screen yielded 23 confirmed hits, which comprised ten flame retardants, seven pesticides and six drug-like compounds. Comparison of concentration–response curves for migration and viability showed that all hits were specific. The extent to which

migration was inhibited was 25–90%, and two organochlorine pesticides (DDT, heptachlor) were most efficient. In the second part of this study, (1) the cMINC assay was repeated under conditions that prevent proliferation; (2) a transwell migration assay was used as a different type of migration assay; (3) cells were traced to assess cell speed. Some toxicants had largely varying effects between assays, but each hit was confirmed in at least one additional test. This comparative study allows an estimate on how confidently the primary hits from a cell function-based screen can be considered as toxicants disturbing a key neurodevelopmental process. Testing of the NTP80-list in more assays will be highly interesting to assemble a test battery and to build prediction models for developmental toxicity.

Keywords Cell migration · Cell tracking · Cytotoxicity · High content imaging · Developmental toxicity · Human stem cells

Abbreviations

AOP	Adverse outcome pathways
AraC	Cytosine arabinoside
cMINC	Circular MINC
CytoD	Cytochalasin D
DMSO	Dimethyl sulfoxide
DNT	Developmental neurotoxicity
EC	Effective concentration
EC90V	EC90 of viability
EC75M	EC75 of migration
EGCG	Epigallocatechin gallate
EGF	Epidermal growth factor
FBS	Foetal bovine serum
FGF	Fibroblast growth factor
FR	Flame retardant
GRIND2	Grid-independent descriptors

✉ Marcel Leist
marcel.leist@uni-konstanz.de

- ¹ In Vitro Toxicology and Biomedicine, Department inaugurated by the Doerenkamp-Zbinden Foundation, University of Konstanz, Box 657, Universitaetsstr. 10, 78457 Konstanz, Germany
- ² Research Training Group RTG1331, Konstanz, Germany
- ³ Konstanz Research School Chemical Biology (KoRS-CB), Konstanz, Germany
- ⁴ Research Programme on Biomedical Informatics (GRIB), Department of Experimental and Health Sciences, Universitat Pompeu Fabra, Dr. Aiguader 88, 08003 Barcelona, Spain
- ⁵ Division of National Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA

hESC	Human embryonic stem cell
log <i>P</i>	Octanol–water distribution coefficient
MeHgCl	Methylmercury(II) chloride
MINC	Migration of neural crest cell
MW	Molecular weight
NCC	Neural crest cell
NHBA	Number of hydrogen bond acceptors
NOAEL _v	No observed adverse effect level for viability
NRB	Number of rotatable bonds
NTP	National Toxicology Program
PAH	Polycyclic aromatic hydrocarbon
PBDE	Polybrominated diphenyl ether
PBS	Phosphate buffered saline
PCA	Principal component analysis
QSAR	Quantitative structure-activity relationship
ROI	Region of interest
TB-BPA	Tetrabromobisphenol A
TPSA	Polar surface

Introduction

During the past 10 years, many new assays have been developed to identify developmental toxicants and developmental neurotoxicants (Bal-Price et al. 2012; Robinson and Piersma 2013; Schmidt et al. 2016; Smirnova et al. 2014; van der Laan et al. 2012). For such non-animal test methods, there are two fundamentally different applications: the first is for hazard identification of defined compounds, and potential further use of the data in risk assessment; the second is for screening of larger sets of compounds to identify those that require further testing. The latter approach requires a relatively high throughput, but has little stringent requirements for specificity and formal test validation. Examples for large screening programmes are ToxCast and Tox21 (Judson et al. 2014; Tice et al. 2013).

There has always been a strong focus of safety sciences on in depth characterization of individual compounds, either to elucidate mechanisms of toxicity or to provide sufficient data for regulatory purposes. In contrast to this, the idea of screening large numbers of compounds is a relatively new development (Collins et al. 2008; Leist et al. 2008; Shukla et al. 2010). It has become possible through adoption of high-throughput technology originally developed for drug discovery in pharmaceutical industry. Such screen data are frequently linked to large uncertainties. Moreover, the assays are generally established in a way that allows for many false-positives. Therefore, different strategies have been developed in the pharmaceutical industry to further filter, confirm and characterise the hits. A present challenge in toxicology is to establish corresponding quality evaluation and follow-up strategies for toxicity screens

(Judson et al. 2013; Leist et al. 2014; van der Burg et al. 2015).

For drug discovery screens, the first step is invariably a confirmation testing based on the same assay as used for initial screening. In general, this is followed by secondary, tertiary and higher level tests to reduce the level of technical and biological uncertainty. Such assays use different technologies, different analytical readouts or other biological test systems. Moreover, the follow-up tests often increase in complexity to get closer to providing human relevant information. In parallel, investigations are initiated to provide answers for the question on how broadly applicable a screen finding is, with respect to cell types, tissues, developmental stages and other important biological determinants. Although a lot can be learned from drug discovery, the development of own strategies for toxicology is important, because the objectives of safety sciences and drug discovery are different.

One approach of *in vitro* toxicology is to use several parallel tests (test batteries) instead of a single test. This provides information on broader sets of biological functions or pathways that may be disturbed. This principle has been used successfully for the definition of non-genotoxicants. If a compound is found negative in three non-animal tests that cover different types of biological processes relevant to DNA damage, then the chemical is judged to be non-mutagenic (Adler et al. 2011; Combes et al. 2008). Recently, also a test battery for dermal sensitization has been approved. In this case, the three tests currently forming the battery cover key events of an adverse outcome pathway (Basketter et al. 2015; Ezendam et al. 2016; Reisinger et al. 2015). A third successful example of *in vitro* test battery use, in another domain of toxicology, is the potential substitution of the *in vivo* uterotrophic assay of the endocrine disruptor screening program by a set of ToxCast assays directed to various steps relevant for oestrogen signalling (Browne et al. 2015).

In the fields of neurotoxicity and developmental neurotoxicity (DNT) first suggestions for test batteries have been emerging (Schmidt et al. 2016; Schultz et al. 2015; Shinde et al. 2016; van Thriel et al. 2012; Zimmer et al. 2014), but further information on the contribution of individual assays is required, and this is also true for reproductive toxicity. In the latter field, reporter assays for nuclear receptors and a Wnt-signalling assay have been combined with a zebrafish embryotoxicity test and the murine embryonic stem cell test within the ReProTect and ChemScreen projects (Kroese et al. 2015; Piersma et al. 2013; Schenk et al. 2010; Sonneveld et al. 2005; van der Burg et al. 2015).

An important step in the further development of test batteries for DNT testing is an evaluation of the technical performance and throughput feasibility of individual assays. A standard approach for this has been the selection of a

small set of compounds, mostly less than ten, that are considered positive and negative controls for the test endpoint chosen. While this is an important first step (Crofton et al. 2011; Leist et al. 2010), relying on this strategy only has made a comparison of the results of cell-based tests difficult. Therefore, further progress depends on the assembly of larger sets of compounds (called a compound library) to be tested in all tests of a battery. Large libraries of several thousand compounds have been established in the Tox21 program (Huang et al. 2016; Richard et al. 2016). Examples for smaller libraries are the 150 compounds of the TG-GATES project for hepatotoxicity (Grinberg et al. 2014; Igarashi et al. 2015), the 30 compounds of the ESNATS battery (Colaïanna et al. 2016; Zimmer et al. 2014) or the set of about 75 compounds assembled by the National Toxicology Program (NTP) of the USA (Pei et al. 2015; Ryan et al. 2016).

Concerning the test endpoints, biochemical measures or reporter gene expression are particularly easy to upscale to high throughput, but the data generated by such tests are difficult to link to classical endpoints of toxicity. The reverse is true for tests of complex and superordinate cell functions such as differentiation, proliferation or migration. Such assays are typically based on laborious cell systems (i.e. long differentiation protocol from pluripotent stem cells) or they require complicated measurement techniques (e.g. electrical activity of neuronal cells) that limit throughput. Nevertheless, it has been repeatedly suggested that such cell function assays will be the first to be used in DNT *in vitro* batteries (Aschner et al. 2016; Bal-Price et al. 2015; Crofton et al. 2011; Hirsch et al. 2016). However, until now the use and inter-laboratory comparison of such assays for screening is still rare. Neurite outgrowth is one of the few examples of a neuronal cell function assay that has been successfully run at high throughput in different tests (Crofton et al. 2012; Hoelting et al. 2016; Krug et al. 2013; Ryan et al. 2016; Stiegler et al. 2011).

Another important function for the development and maintenance of the nervous system is cell migration. For instance, in the developing vertebrae central nervous system, neural precursor cells migrate along radial glial cells from the ventricular zone to outer cortical layers (Luhmann et al. 2015). Moreover, throughout life, neural precursors are formed in the subventricular zone, and these then migrate as doublecortin positive neural precursors to their final tissue location (Kaneko et al. 2017). Toxic influences on human central nervous system precursor migration have been studied in the neurosphere migration assay (Fritsche et al. 2005; Moors et al. 2009).

Neural crest cells (NCC) are another migratory cell population arising during early neurodevelopment. We previously established a (low-throughput) assay that measures migration inhibition of neural crest cells (MINC

assay) (Zimmer et al. 2012). Initial results of the MINC assay showed that some exemplary compounds specifically affect NCC function (Dreser et al. 2015; Zimmer et al. 2014). Subsequently, a modified version of the assay (cMINC) was adapted for higher throughput by making it experimenter-independent and enabling automated image acquisition (Nyffeler et al. 2016). In the present study, we explored the suitability of the cMINC assay for screening purposes. Moreover, we devised a hit follow-up strategy. In this context, we established and used a panel of secondary assays to obtain further information on the robustness of primary hits and to get an estimate on how confidently the primary hits from a cell function-based screen can be considered as toxicants disturbing a key neurodevelopmental process.

Materials and methods

NCC differentiation

NCCs were differentiated from H9 (WA09; WiCell, Wisconsin) human embryonic stem cells (hESC). The modified line expressed GFP under the control of the *Dll1* promoter (H9-*Dll1*; provided by M. Tomishima, Memorial Sloan Kettering Cancer Centre, New York, NY, USA) as described earlier (Zimmer et al. 2012). Import of the hESC and all experiments utilising them were carried out according to German legislation and under the licence of the Robert-Koch Institute (license number 1710-79-1-4-27).

Differentiation to NCCs was performed exactly as described earlier (Nyffeler et al. 2016; Zimmer et al. 2012). Briefly, the protocol involves differentiation of hESC to rosettes, manual picking of rosettes and FACS-sorting for positive expression of the NCC marker HNK-1 and negative expression of *Dll1*. The sorted cells were expanded in N2-S medium, consisting of DMEM/F12 (Gibco) and supplemented with 100 µg/ml apotransferrin, 25 µg/ml insulin, 8.6 mM glucose, 100 µM putrescine, 30 nM selenite, 20 nM progesterone (all purchased from Sigma-Aldrich), 1× GlutaMax and 1% penicillin/streptomycin (both from Gibco), 20 ng/ml epidermal growth factor (EGF) and 20 ng/ml fibroblast growth factor (FGF) 2 (both from R&D). After 27–30 days, cells were cryopreserved until further use. For all experiments, freshly thawed NCCs were used and cultured on poly-L-ornithine/fibronectin/laminin coated plates in N2-S medium supplemented with 20 ng/ml EGF and 20 ng/ml FGF2.

NCC migration setup ('cMINC')

The NCC migration assay was performed exactly as described in Nyffeler et al. (2016). Briefly, cells were

seeded (95,000 cells/cm²) in poly-L-ornithine/fibronectin/laminin coated 96-well plates containing silicon stoppers (Platypus Technologies, Madison, WI, US) to create a circular cell-free area. One day after seeding, stoppers were removed to initiate migration and medium was refreshed. After 24 h, test compounds were added as 5× concentrate (relative to the final test concentration), and after additional 24 h of incubation, migration and viability were measured. A positive control, cytochalasin D (CytoD, 200 nM) was run on every plate for quality control. For standard testing, toxicants were diluted in medium *plus* dimethyl sulfoxide (DMSO) so that the final DMSO concentration was 0.1% in all wells.

Screen compounds were obtained diluted in DMSO from the NTP as described in Ryan et al. (2016). For the hit-confirmation, all compounds were ordered independently and diluted in DMSO. A list of the test chemicals is given in the Supporting Information. For both the screening and hit-confirmation part, aliquots were stored at -80 °C and for each biological replicate, a fresh aliquot was used. The screening was performed with 4–5 technical replicates, whereas in the hit-confirmation 3–4 technical replicates were done.

Image-based measurement of cell migration and viability

Image acquisition was performed exactly as described in Nyffeler et al. (2016). Briefly, cells were stained with H-33342 and calcein-AM and imaged on a high content imaging microscope (Cellomics ArrayScanVTI, Thermo Fischer). To obtain data on NCC viability, four fields (647 × 647 μm²) outside the migration area were imaged with a 10× lens (approximately 500 cells/field). Viability was defined as the number of H-33342 and calcein double-positive cells as determined by an automated algorithm described earlier (Krug et al. 2013; Stiegler et al. 2011).

To obtain data on cell migration, the centre of the well was imaged in four tiles with the 5× lens. The four images were joined to obtain one micrograph covering an area of 2590 × 2590 μm². For quantification of migration, a software tool (freely accessible at <http://invitrotox.uni-konstanz.de/>) was used to estimate the most likely position of the previously cell-free area, to set thresholds for colour intensity for both dyes, and to count the number of H-33342 and calcein double-positive cells in the region of interest (ROI). The radius of the ROI was reduced to 0.7–0.9 mm to account for the fact that cells could migrate into the cell-free area for 24 h before the treatment was started. The radius of the reduced ROI was chosen in a way that at least 150, but not more than 300 cells, were in the ROI in the untreated condition.

Analysis of migration and viability data

If not mentioned otherwise, viability and migration were normalised to solvent control (0.1% DMSO). Concentration–response curves from averaged data ($n = 3$ experiments) were fitted using R (R Core Team 2015) and the package ‘drc’ (Ritz and Streibig 2005) to a log-logistic function with four parameters, where the upper asymptote was set to 100%, and appropriate constraints were used for the other parameters (see Fig. S4 and S5). The data points within a curve were weighted with 1/SD; effective concentration (EC) values of the fitted curves were estimated. Efficacy was defined as the amount of migration-inhibition at the EC90 of viability (EC90V). If the EC90V was above the highest tested concentration, the extent of migration at the highest tested concentration was used instead and the greater-than sign was introduced. Specificity was defined as the ratio of EC90V and EC75 of migration (EC75M).

Immunocytochemical characterization

For immunocytochemical staining, cells were seeded on glass as drops (10,000 cells per 10 μl drop) to observe migrating cells at the border. After 2 days in culture, cells were fixed with 4% formaldehyde and 2% sucrose in phosphate-buffered saline (PBS). For intracellular epitopes, cells were permeabilized for 15 min with 0.1% Triton, followed by blocking with 10% foetal bovine serum (FBS) in PBS for 1 h before primary antibodies were applied in 10% FBS in PBS overnight. The next day, cells were washed with 0.05% Tween in PBS and secondary antibodies were applied (1:1000) for 1 h. Cell nuclei were counterstained with 1 μg/ml H-33342 for 10 min. Images were acquired using a point laser scanning confocal microscope Zeiss LSM 700 (Zeiss, Oberkochen, Germany). The list of antibodies used is given in the Supporting Information.

Time-lapse experiment and cell tracking

Cells were seeded under standard cMINC conditions and treated with the test compounds as described in Nyffeler et al. (2016). The DMSO concentration in the medium did not exceed 1%. During the 24-h treatment period, cells were imaged using an Axio Observer.Z1 microscope (Zeiss, Oberkochen, Germany), equipped with an AxioCam MRm camera and an incubation system for constant temperature and CO₂. Phase contrast images of the migration zone were acquired every 10 min using a 5× objective. After the time-lapse, migration and viability were determined as described in the ‘NCC migration setup’ section to ensure that the toxicants worked as under standard cMINC conditions with exposure in a standard incubator.

Cells were manually tracked using the ImageJ plugin, 'Manual Tracking' in reverse time. Cells were selected for tracking if they were (1) viable (i.e. not rounded up) at the end of the time-lapse, (2) located in the circular migration zone at the end of the time-lapse and (3) trackable for the entire time-lapse (i.e. did not leave the imaging frame).

For each biological replicate, ten individual cells were tracked per condition (and 20–30 cells for the untreated control). The cumulated distance of each cell was normalised to the median of the untreated control population. Notably, cells were only tracked if they were viable at the end of the assay. By this way, we purposely selected for 'healthy' cells to exclude cell death as a confounding factor. This likely leads to an overestimation of the speed of the cell population. Additionally, only those that were in the migration zone at the end of the assay were tracked. This might result in preferential counting of faster-than-average, mobile cells, which again would overestimate the overall speed. Statistical significance of differences between treated and non-treated conditions was calculated on the combined data of two biological replicates using a Wilcoxon rank sum test. In general, tracking was performed at the EC90V (to ensure that a non-cytotoxic, but rather high concentration was used) except for: LiCl, CdCl₂, BDPD were tested at the highest concentration that was used in the cMNC (which was non-cytotoxic); EGCG and berberine were measured at 20 µM; valinomycin at 1.25 µM.

Transwell experiments

The transwell inserts (pore size 8 µm, polycarbonate membrane, Corning, catalog no. 3422) were pre-coated with poly-L-ornithine/fibronectin/laminin. Cells were seeded at a density of 50,000 cells per insert (150,000 cells/cm²) into the upper chamber. The cells were stimulated to migrate through the membrane by addition of medium with 5% FBS to the lower chamber (reservoir). Notably, migration did not occur, when FBS was added to the upper reservoir or to both reservoirs. Test compounds were applied to both the lower and upper chambers. In general, the EC90V concentration was used except for CytoD (200 nM), taxol (10 nM), EGCG (20 µM), berberine chloride (20 µM), PBDE-99 (20 µM) and valinomycin (1.25 µM). The cells were incubated for 6 h at 37 °C and 5% CO₂ and allowed to migrate.

After incubation, the medium was aspirated from inserts and reservoirs and then the upper side of each insert was gently swabbed, using cotton swabs, to remove cells that had not migrated through the membrane. Reservoirs and inserts were washed once with PBS and then the migrated cells were fixed with 3.7% formaldehyde and stained with crystal violet for 10 min. Then, the

inserts were thoroughly rinsed with water and dried for at least 24 h.

To evaluate the number of migrated cells through the membrane, five pictures per condition were taken under the light microscope. In a typical experiment, stimulation with FBS lead to ca. 100 migrated cells per picture. The number of migrated cells is normalised to that of cells stimulated only with FBS in absence of toxicants.

Analysis of the chemical space

The structures of the compounds were obtained from the SMILES provided in the original table, converted to SDF file format using RDKit version 0.9.2 (<http://www.rdkit.org>) and protonated to pH 7.4 using Moka version 1.1 (Milletti et al. 2007). The values of the molecular weight (MW) and octanol-water distribution coefficient ($\log P$) were obtained using RDKit.

The structures were normalized using standardizer (<https://github.com/flatkinson/standardiser>, accessed: 21st November 2016) and converted to 3D using Corina version 3.494 (Sadowski et al. 1994). These were then used to generate GRIND2 descriptors (Duran et al. 2009; Pastor et al. 2000) making use of Pentacle software version 1.0.6 (<http://www.moldiscovery.com/software/pentacle>), with default settings. The resulting molecular descriptors were then projected into the principal component analysis (PCA) scores obtained for a collection of 8298 ToxCast and Tox21 compounds (USEPA 2016) characterised using a similar procedure (see Supporting Information Excel File).

From the original 75 library compounds, the following compounds had to be removed because they are salts or contain metallic elements not supported by our methods: methylmercury(II) chloride (MeHgCl), acetic acid manganese (2+) salt, bis(tributyltin)oxide, methylcyclopentadienyl manganese tricarbonyl. For the remaining 71 compounds, values of $\log P$ and MW were obtained. In the process of computing the GRIND2 molecular descriptors, two more compounds had to be removed: saccharin sodium salt hydrate and benzo(b)fluoranthene. Thus, the final series projected in the ToxCast and Tox21 space contains 69 compounds.

Data handling and statistics

If not mentioned otherwise, values are expressed as means \pm SD. If not indicated otherwise, experiments were performed at least three times (i.e. using three different cell preparations), with at least three technical replicates per condition.

Results and discussion

Introduction to the biological system

NCCs were differentiated from pluripotent stem cells via a rosette stage, sorted for positive expression of the NCC marker HNK-1, expanded and cryopreserved as described previously (Zimmer et al. 2012). The resulting cells have been extensively characterised for NCC marker expression elsewhere (Nyffeler et al. 2016). We continued this characterisation here by staining for some cell organelles that are relevant for migration or that provide targets for toxicants. This high-resolution imaging showed that NCCs have varying shapes, with the polygonal type being most frequent. As in fibroblasts, a pronounced cortical actin skeleton that extended into cellular processes was conspicuous. The microtubules stabilised the interior part of the cells, but did not extend into the projections, while nestin, as the typical intermediary filament expressed in these cells, extended throughout the cytoplasm and right into the cell extensions. The few, and quite distinct, mitochondria formed a network around the nucleus, but sometimes extended into protrusions of the leading edge. The Golgi apparatus was always sharply delineated and clearly situated to one side of the nucleus. Its position was often, but not always, in the direction of the leading edge, as observed in some other migratory cells (Fig. 1a).

After this confocal imaging of the cells, conventional, but highly automated fluorescence microscopy was used to obtain high-throughput screening data. Cells were seeded in 96-well plates containing round silicon stoppers, and after one day in culture, the stoppers were removed to allow migration into the circular area (Fig. 1b). After 48 h, the cultures were stained with the viability dye calcein-AM (cytoplasmic localization in live cells) and with Hoechst 33342 to visualise the nuclei of all cells. The number of viable cells (defined as the number of calcein-positive nuclei), and the number of migrated cells were determined using previously published algorithms (Krug et al. 2013; Nyffeler et al. 2016; Stiegler et al. 2011).

Using this procedure, several published endpoint-specific controls (Aschner et al. 2016; Barenys et al. 2016; Zimmer et al. 2012) were tested for their effect on cell migration. The JNK inhibitor SP600125 and the Src-kinase inhibitor PP2 inhibited migration in a concentration-dependent manner at non-cytotoxic concentrations in the micro-molar range. As expected, also, the G α i signalling inhibitor pertussis toxin, as well as the polyphenol EGCG (preventing laminin from binding to the integrin β 1 receptor) inhibited NCC migration (Fig. 1c). All these largely different positive controls yielded highly

reproducible concentration-dependent inhibition data in three separate experiments and allowed to exemplify effect quantification according to our earlier developed preliminary prediction model for the cMINC test. We therefore decided that the assay was ready for a screen of multiple unknown compounds, and we adopted our earlier definition of compound efficacy as ‘the extent of migration inhibition at the highest non-cytotoxic test concentration (EC90V)’ and of compound specificity as ‘the ratio of the EC90V and the concentration inhibiting migration by 25% (EC75M)’.

Overview of the testing strategy

The library to be tested has been assembled at the NTP of the USA to probe the feasibility and performance of screening approaches in the fields of neurotoxicity and developmental neurotoxicity (Pei et al. 2015; Ryan et al. 2016). It contains 75 unique compounds, five of which (saccharin, deltamethrin, triphenyl phosphate, MeHgCl and phenobarbital) were included as duplicates for internal consistency control; it is thus short-named as ‘NTP80-list’ here. The library contained some known (developmental) neurotoxic compounds of the group of pesticides, drug-like compounds and industrial chemicals, in addition to large sets of polycyclic aromatic hydrocarbons (PAH) and flame retardants (FR) (Fig. S1). Moreover, five negative controls were included (acetylsalicylic acid, acetaminophen, D-glucitol, L-ascorbic acid and saccharin sodium salt). The library was delivered in the form of 20 mM stocks in DMSO, except for few compounds of lower solubility.

As the production of NCC is cost- and time-demanding, a tiered screening strategy was chosen (Fig. 2). In a first step, all compounds were tested at the technically highest concentration (1:1000 of the stock). If cytotoxicity was observed (viability <85%), the compound was retested at progressively twofold dilutions, until a non-cytotoxic concentration (NOAEL_V) was determined. Note that in this pre-screen, we deliberately chose a non-stringent viability threshold (normally, a cutoff of 90% is used) to include a high percentage of potentially specific toxicants. At the NOAEL_V, migration was measured in three independent experiments. Compounds were classified as ‘negative’ if average migration was \geq 80% (again, the threshold for migration-inhibition was chosen less stringent than in the standard procedure). All other compounds were considered as ‘potential positive hits’ and underwent hit confirmation testing in the second tier. For this purpose, all relevant compounds were purchased from a different source (confirmation of compound identity and integrity), and re-tested in the migration assay. At this stage, six dilutions were tested simultaneously. A compound was classified as ‘positive

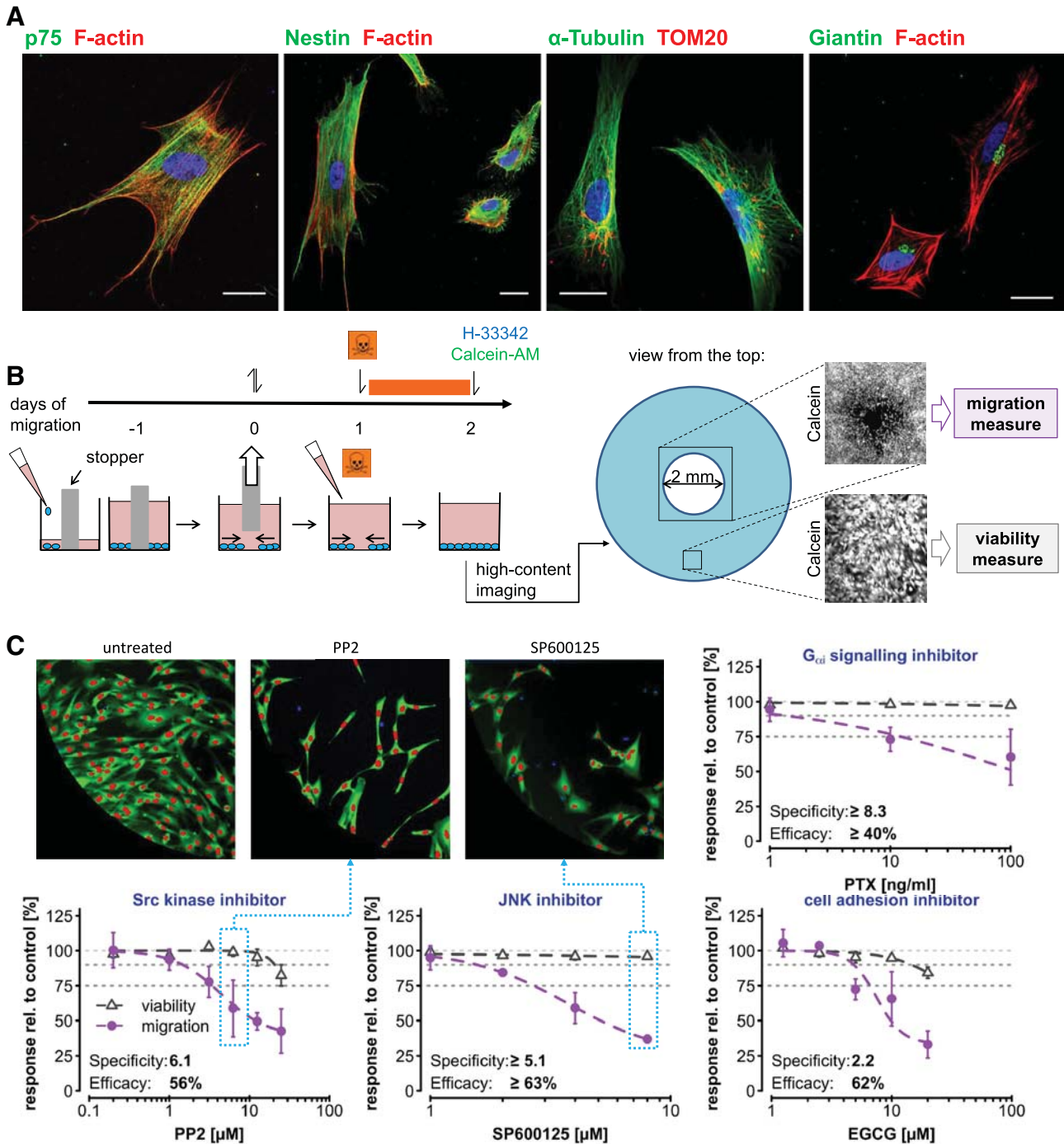


Fig. 1 Introduction to the biological system. **a** Cryopreserved NCCs were thawed and plated on coverslips. After 2 days, they were fixed, and various cellular structures and markers were stained; H-333342 (depicted in blue) was used to visualise the nuclei. NCCs migrating into cell-free zones were imaged using a confocal microscope. The scale bar corresponds to 25 μ m. *p75* NCC marker; *Nestin* stem cell marker; *TOM20* mitochondrial marker; *Giantin* golgi marker. **b** Graphical representation of the assay scheme: cells were seeded in 96-well plates containing silicone stoppers to create a cell-free circular area. After 1 day, the stoppers were removed, and cells were allowed to migrate for 24 h before the test compounds were added for additional 24 h. For imaging with an automated screening micro-

scope, cells were visualised using H-333342 (intercalates into DNA) and calcein-AM (fluorescent in viable cells only). Imaging of the central zone allowed the quantification of cell migration; the cell viability was determined from images of the well periphery. **c** Examples for endpoint-specific controls. The known biological activities of the test compounds are indicated on top of the diagrams in blue. The light grey dotted line indicates the 100% value. The other two grey lines are drawn at 90 and 75% to indicate thresholds for reduced viability and migration, respectively. All values are normalised to untreated controls. Data are mean \pm SD from three experiments. In the microscopy images, the viable cells are displayed in green and nuclei of viable cells in red (colour figure online)

Fig. 2 Overview of the testing scheme leading from initial screen to hazard statement. In a first phase, the highest non-cytotoxic concentration was assessed in an iterating procedure. Then, the migration assay was performed (three independent runs) and compounds were classified either as potential hits or as 'inactives'. Subsequently, concentration–response curves were obtained for the potential hit compounds. The confirmed hits were then further characterised using three follow-up assays. The decision points required that at least one concentration had to fit the indicated decision criterion. The “hazard statement” is meant to imply a concentration range at which effects on NCC are to be expected and a measure of confidence in the screen assay results. *PBDE* polybrominated diphenyl ether

hit' if migration was $<75\%$ at EC90V. In the third tier of this study, these 'positive hits' were further studied: (1) the cMINC assay was repeated under conditions that prevent proliferation; (2) a transwell migration assay was used to test whether a different migration assay would confirm the result; (3) cells were traced to assess cell speed as a different endpoint. This final tier was included to obtain information in how far technically-defined 'screen hits' can indeed be considered as compounds associated with the toxicological hazard of 'interfering with the fundamental neurodevelopmental process' of NCC migration (Fig. 2).

Results of the screening

During the screening of the NTP80-list, only 12 compounds proved to be cytotoxic at the highest tested concentration. For these substances, a NOAEL_V was determined, before migration data were obtained. For all other compounds, migration was measured at the highest possible concentration (usually $20\ \mu\text{M}$). On each plate, CytoD was run as a positive control and reduced migration to 30–70% compared to solvent control (Fig. S2). There was only limited variation among assays and cell lots, indicating that the assay worked reproducibly.

In tier 1, out of the 80 compounds, 27 'potential positive hits' were identified (Fig. 3), and none of the compounds triggered an increase in migration. For one compound (benzo[b]fluoranthene), the standard quantification procedure was not possible, as the fluorescent compound interfered with image acquisition of the H-33342 channel. However, manual inspection of the calcein-stained cells did not indicate any inhibition of migration, and hence no further experiments were conducted.

Among the 'potential positive hits', there were ten of 12 tested FR, seven of 17 pesticides, six of the 15 drug-like compounds as well as two PAHs and one industrial chemical. None of the five negative controls inhibited migration.

Hit confirmation with primary assay

In tier 2, new stocks were prepared for all 'potential positive hit' compounds, and full concentration–response

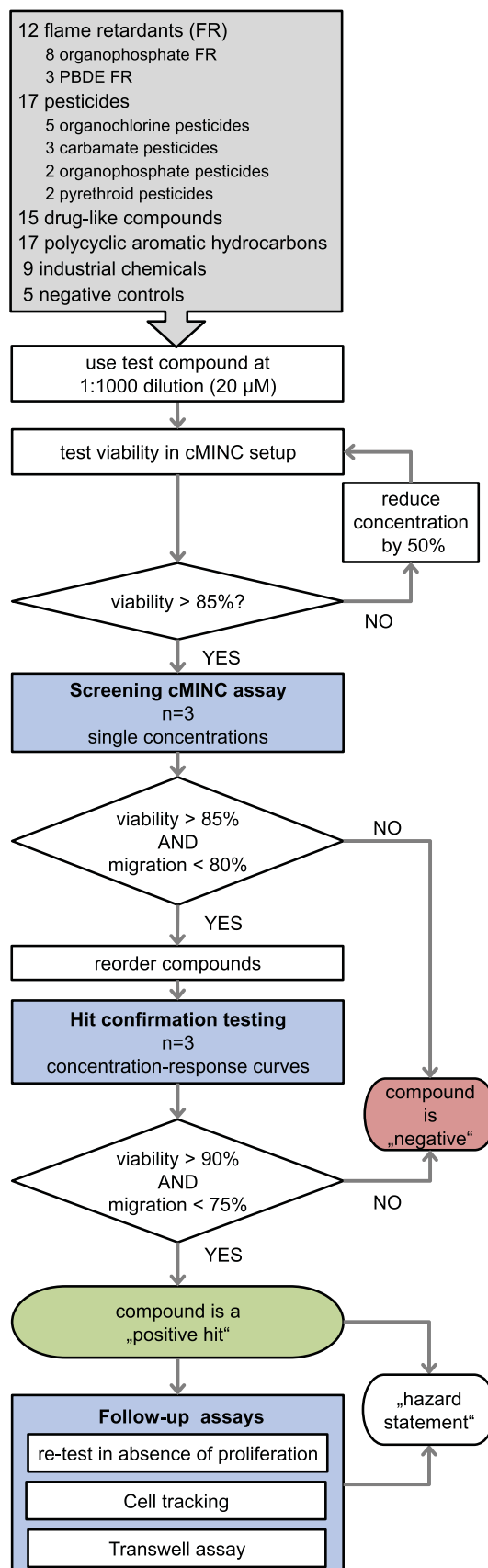
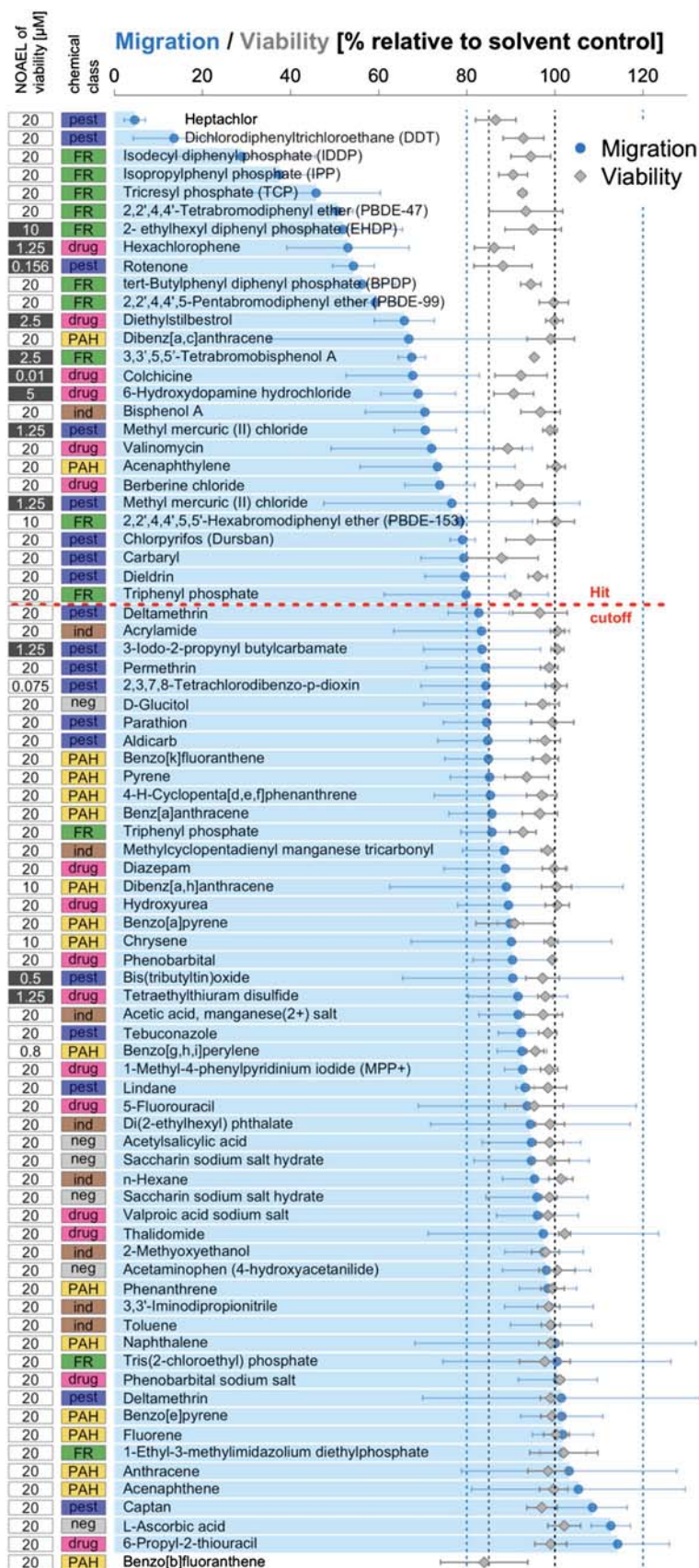


Fig. 3 Results of the screening. Compounds were screened according to the procedure indicated in Fig. 2 at a 1:1000 dilution (mostly 20 μM) if they were not cytotoxic at this concentration. Otherwise, they were tested at the highest non-cytotoxic concentration. Compounds are ordered according to the screen results from most migration-inhibiting (*top*) to least migration-inhibiting. For each compound, viability (*grey diamonds*) and migration (*blue circles*) at the highest tested non-cytotoxic concentration are displayed. The *vertical lines* are at 80, 85, 100, 120% of control and indicate the thresholds for viability (*grey*) and migration (*blue*). Compounds reducing migration by $\geq 20\%$ underwent subsequent hit confirmation testing. For one compound (benzo[b]fluoranthene), migration could not be evaluated due to interference of the chemical with the fluorophores of the test. The chemical classes of the compounds were *FR* flame retardant, *PAH* polycyclic aromatic hydrocarbon, *ind* industrial chemical, *pest* pesticide, *drug* drug-like compound, *neg* negative control. On the *left side* of the overview, the NOAEL of viability is displayed in a *white box* for compounds not cytotoxic at the highest tested concentration (i.e. the NOAEL is greater than the given value) and in a *grey box* for all other compounds (colour figure online)



curves were obtained. Both, the viability and migration data, were fitted to a log-logistic function (Fig. S4 and S5) and specificity and efficacy were calculated as described in Nyffeler et al. (2016).

Of the 26 ‘potential positive hits’ (MeHgCl was twice in the library), 23 were confirmed (Figs. 4, 5, S3A). The three non-confirmed compounds were the two PAHs acenaphthylene and dibenz[a,c]anthracene as well as bisphenol A (Fig. S3B). At this stage, all 23 confirmed hits fell into the chemical classes of FR (10 hits), pesticides (7 hits) and drug-like compounds (6 hits).

The hits belonging to the FR group showed a rather large specificity (Fig. 4), and a potency in the low (single-digit) micromolar range. Tetrabromobisphenol A (TB-BPA) was the most potent FR, followed by the organophosphorus compounds and polybrominated diphenyl ethers (PBDEs). The six hits amongst the eight organophosphorus FR have in common that they contain at least one aromatic side chain, whereas the two non-hits did not show this feature. For PBDEs, it seems that an increasing number of bromine atoms reduces the effect, as PBDE-47 (4 bromines) was more effective than PBDE-99 and PBDE-153 (5 and 6 bromines, respectively).

Of 15 tested drug-like compounds, six were found to be migration-inhibiting after the hit confirmation phase (Fig. 5a; Fig. S3A). Colchicine, previously known as a NCC migration-inhibitor (Nyffeler et al. 2016), was the most potent hit (low nanomolar range). Hexachlorophene, a chlorinated compound used as a disinfectant, and diethylstilbestrol (previously used as a contraceptive) were amongst the hits. For valinomycin, the hit confirmation showed a non-orthodox and biphasic concentration–response behaviour (confirmed in altogether six experiments). In the low nanomolar range, the cell viability dropped reproducibly by about 10%, but was not affected by increasing concentrations up to the micromolar range. At concentrations of about 2–4 μM , a pronounced effect on migration was measured. To display the effect of the compound in the micromolar range in a way comparable to other substances, the data were normalised to the data obtained in the lower nM range instead of the usual comparison to solvent controls. Amongst the drug-like compounds, also berberine chloride and 6-hydroxydopamine were confirmed as migration-inhibitors, but they had a rather low specificity and efficacy (Fig. S3A).

Seven out of 17 tested pesticides inhibited NCC migration (Fig. 5b; Fig. S3A). MeHgCl and rotenone were confirmed as known migration-inhibitors (Dresler et al. 2015; Moors et al. 2009; Zimmer et al. 2012), while pyrethroids had no effect at all. Concerning a role of choline esterase inhibition, carbaryl (Fig. S3A) was the only positive hit of the carbamate group ($n = 3$) and it had a low efficacy (29%). Chlorpyrifos showed specific effects at relatively

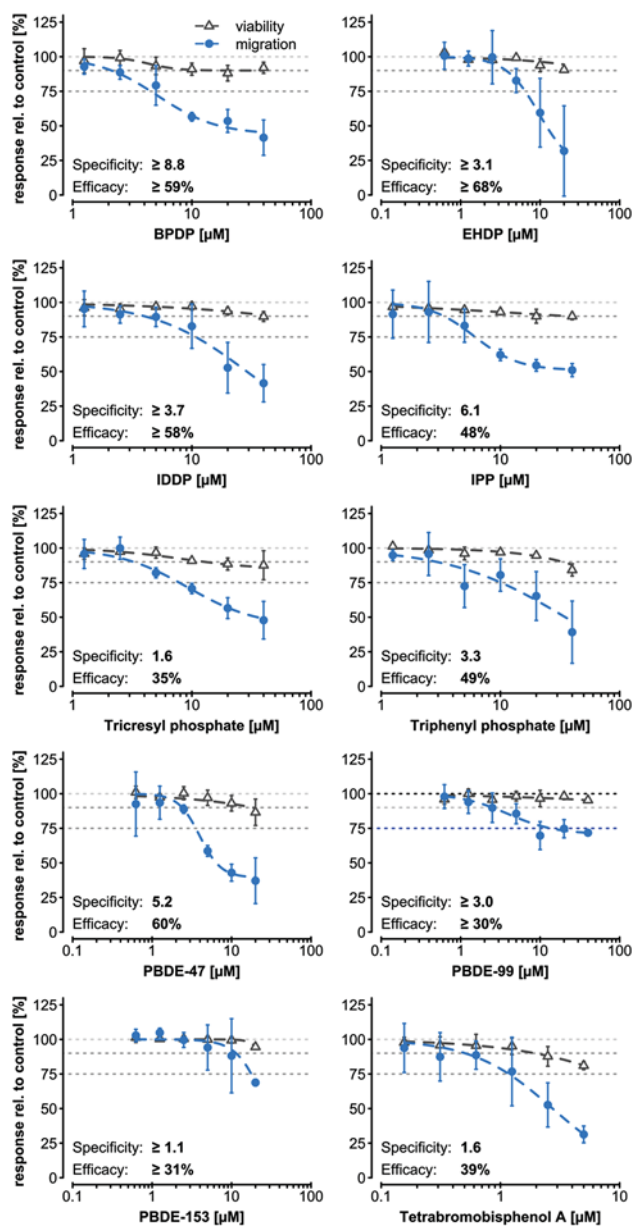


Fig. 4 Hit confirmation with the primary assay: flame retardants. Compounds identified as potential hits in the screen were re-ordered and concentration–response curves were obtained for viability (grey triangles) and migration (blue circles). All values are normalised to the solvent control (0.1% DMSO). The horizontal light grey dotted line indicates 100% value for easier reading of the diagrams. The other two grey lines are drawn at 90 and 75% to indicate the threshold for reduced viability and migration, respectively. A log-logistic function with constraints was fitted to the concentration–response curve and the EC90 of viability and the EC75 of migration were intrapolated. The ratio between these two values was termed ‘specificity’, whereas ‘efficacy’ was defined as the amount of migration-inhibition at the EC90 of viability. Data are mean \pm SD from three experiments (colour figure online)

high concentrations of 10–20 μM . The three organochlorine pesticides DDT, dieldrin and heptachlor were among the most efficient compounds in the library, although

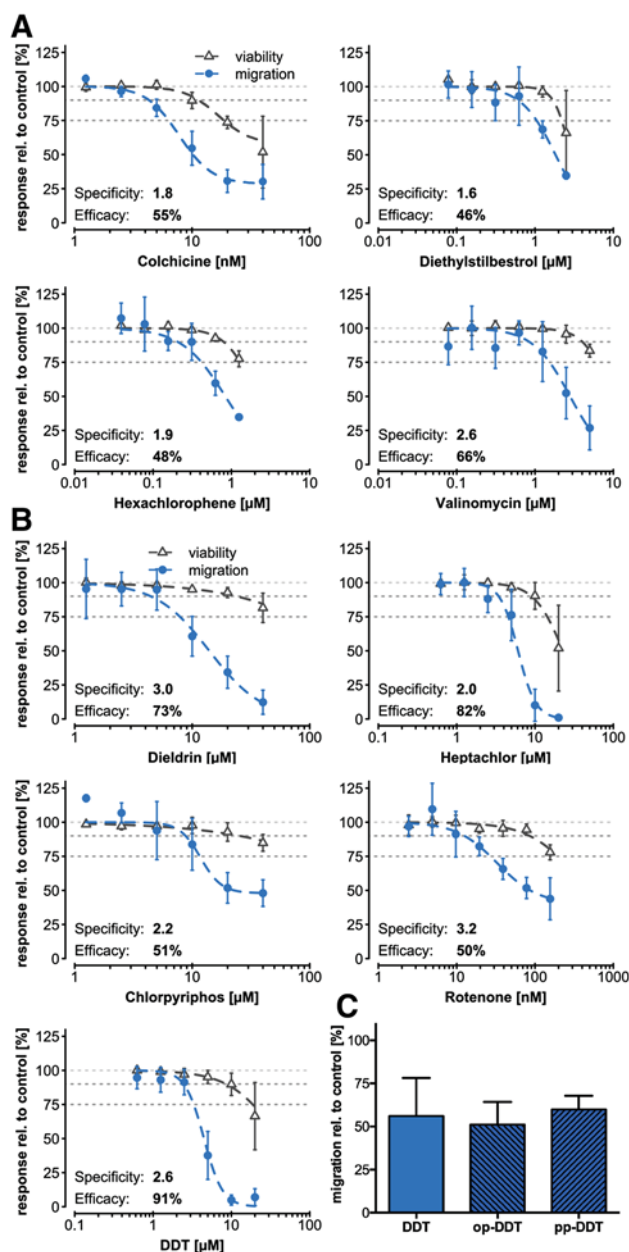


Fig. 5 Hit confirmation with the primary assay: drug-like compounds and pesticides. **a, b** Compounds identified as potential hits in the screen were re-ordered and concentration–response curves were obtained for viability (grey triangles) and migration (blue circles). All values are normalised to the solvent control (0.1% DMSO). The horizontal light grey dotted line indicates the 100% value for easier reading of the diagrams. The other two grey lines are drawn at 90 and 75% to indicate the threshold for reduced viability and migration, respectively. A log-logistic function with constraints was fitted to the concentration–response curve and the EC90 of viability and the EC75 of migration were interpolated. The ratio between these two values was termed ‘specificity’, whereas ‘efficacy’ was defined as the amount of migration-inhibition at the EC90 of viability. To compare the results with other compounds, the data of valinomycin were normalised to 100% (i.e. multiplied with 100/89). Data are mean \pm SD from three experiments. **c** Two isomers of dichlorodiphenyltrichloroethane (DDT) and the commercial DDT mixture were retested in the cMINC assay (5 μ M). Data are mean \pm SD from three experiments. *op*-DDT 2,4'-DDT, *pp*-DDT 4,4'-DDT

concentrations of several micromolar were required to trigger specific effects. In the case of DDT, we followed up on the screen finding by examining effects of the two congeners 2,4'-DDT (*op*-DDT; 1,1,1-Trichloro-2-(2-chlorophenyl)-2-(4-chlorophenyl)ethane) and 4,4'-DDT (*pp*-DDT; 1,1,1-Trichloro-2,2-bis(4-chlorophenyl)ethane). The initial finding of a highly efficient inhibition of migration was confirmed with both compounds (Fig. 5c).

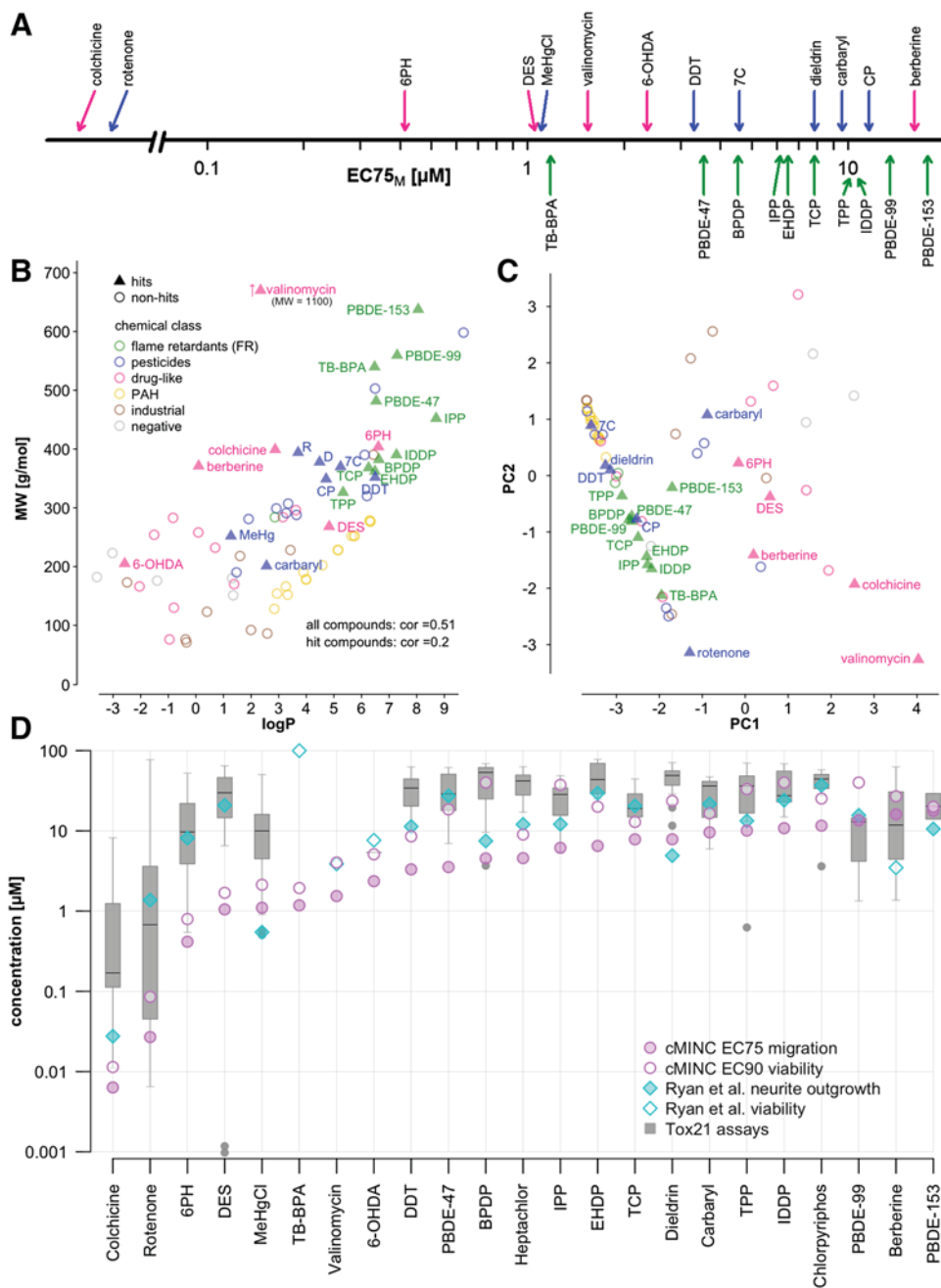
In summary, most of the screen hits were confirmed, and additional information on potency and efficacy could be obtained (Fig. S7 and Table S1). The cMINC screen detected compounds with a potency range over three orders of magnitude (Fig. 6a), and a higher dynamic range may still be obtained using higher start concentrations for the hit finding. The yet little-characterised group of FR had the most consistent effect on NCC migration, whereas all PAHs and industrial chemicals tested were found to be negative.

Representation of chemical space by the library and comparison of hits to published data

As the screen hits were unevenly distributed amongst the groups of compounds assembled in the NTP80-list, it was interesting to see whether some physicochemical properties of chemicals were particularly overrepresented amongst the hits. On the first view, it appeared as if lipophilic compounds (e.g. organochlorines or FR with aromatic substituents) may be particularly prone to be hits in the cMINC. To test this hypothesis, a systematic cheminformatics approach was used. First, we investigated whether the hits and non-hits significantly differ in their MW, $\log P$, number of rotatable bonds (NRB), number of hydrogen bond acceptors (NHBA) and polar surface (TPSA) using a Wilcoxon rank sum test. The group of hits had significantly higher MW ($p < 10^{-6}$), $\log P$ and NRB ($p < 0.01$) (data not displayed). Also, some parameter combinations (e.g. $\log P$ AND MW) differed significantly between hits and non-hits, but none of these combinations was able to specifically predict a hit compound.

Representation of all library compounds in the MW– $\log P$ space (Fig. 6b) illustrated that both properties were weakly correlated ($r = 0.51$), while MW and $\log P$ were largely independent ($r = 0.2$) for the screen hits. To obtain a less biased overview of compound properties, grid-independent descriptors (Duran et al. 2009; Pastor et al. 2000) (GRIND2) were computed as described in the ‘‘Methods’’ section. In order to anchor the display of the NTP80-list according to these descriptors, the same parameters were also obtained for a set of ToxCast and Tox21 compounds ($n = 8298$). Then, two principal components (PC1, PC2) were obtained for this large chemical group using PCA, and the NTP80-list was projected in this principal component space (Fig. 6c). The hits obtained in our screen

Fig. 6 Representation of the library in the chemical space and comparison to published data. **a** Schematic indicating the potencies of the confirmed hits displayed as EC75 of migration, i.e. the concentration at which migration was reduced by 25%. **b** Library compounds are represented according to their molecular weight (MW) and the partition coefficient ($\log P$). **c** Grid-independent descriptors were calculated for 8298 compounds of the Tox21 library and for the screen library used here. A principal component analysis was performed for the Tox21 compounds and the library compounds tested in our study were displayed in the background of the Tox21 principal component space along the first two dimensions. *6-OHDA* 6-hydroxydopamine hydrochloride, *6PH* hexachlorophene, *7C* heptachlor, *CP* chlorpyrifos, *D* dieldrin, *DES* diethylstilbestrol, *R* rotenone, *TB-BPA* 3,3',5,5'-tetrabromobisphenol A. **d** Comparison of the hit-confirmation data (violet circles) with results from a neurite outgrowth assay (turquoise diamonds) and with a broad set of Tox21 assays (grey boxes and points), with individual endpoint data of active assays compiled in the form of boxplots (Ryan et al. 2016). For neurite outgrowth data, the most sensitive endpoint from various neurite outgrowth endpoints and viability measurement is displayed (colour figure online)



occupied a large area of this chemical space, i.e. they were unlikely to reflect only a very narrow chemical subgroup. In a next step, a large number of classification models were built from the screen hits, and they were validated by leave-one-out procedures. However, a convincing separation of hits and non-hits could not be obtained, based on the use of these chemical descriptors (not shown), i.e. cMINC toxicity was not predictable based on the structural information available.

For further information on the cMINC hits, they were compared to available data from a similar screen for neurite outgrowth of induced pluripotent stem cell-derived

neurons and to screen data obtained within the Tox21 program (Ryan et al. 2016). In general, the potency ranking of compounds correlated well between cMINC and the 'Ryan assay', but the apparent toxicant potencies were mostly three- to tenfold higher for the migration assay (Fig. 6d). The cMINC also appeared to be more sensitive than the majority of Tox21 assays, at least for the hit compounds identified in the migration assay. In fact, for most compounds that specifically blocked NCC migration, the cMINC was the most sensitive of all the tests that were compared. Amongst these particularly potent migration inhibitors were rotenone, hexachlorophene, DDT,

PBDE-47, BPDP, heptachlor, EHDP and dieldrin. In contrast to this, hits like PBDE-99, PBDE-153 and berberine had EC75M values in the cMINC that fully overlapped with the range of concentrations that affected most Tox21 assays.

Altogether, the hits identified here were generally bioactive in other tests. Concerning the prioritisation and follow-up of such compounds, the cMINC appeared to be particularly sensitive for certain toxicants, and such toxicants came from a broad range of structural templates, not easily covered by a quantitative structure-activity relationship (QSAR) approach.

Cell tracking as follow-up assay

After the comparison of the cMINC hits to data in the literature, we tested in a second step the biological consistency of screen hits. For this purpose, we examined in how far the

inhibition of migration of NCC was confirmed, when other (secondary) tests for the same endpoint were used.

The first secondary assay assessed the migration speed of single cells. The test conditions were identical to the cMINC, but the analytical endpoint was time-lapse imaging of the cells. The average cell speed (expressed as cumulated travelled distance within 24 h) was determined from the tracks of the cells on the dish (Nyffeler et al. 2016). Untreated cells typically covered a distance of about 1000 μm in 24 h. Cells treated with the exemplary migration toxicant LiCl covered only about half the distance (Fig. 7a). All six positive controls of this test, including the recently characterised migration-inhibitor EGCG (Barenys et al. 2016) were used at non-cytotoxic concentrations, and they reduced the speed by 20–45% (Fig. 7b).

The ‘cMINC hits’ were also tested at their EC90V (or at the highest non-cytotoxic concentration reached). To reduce the workload, only two exemplary FR were tested: BPDP was chosen as representative for the group

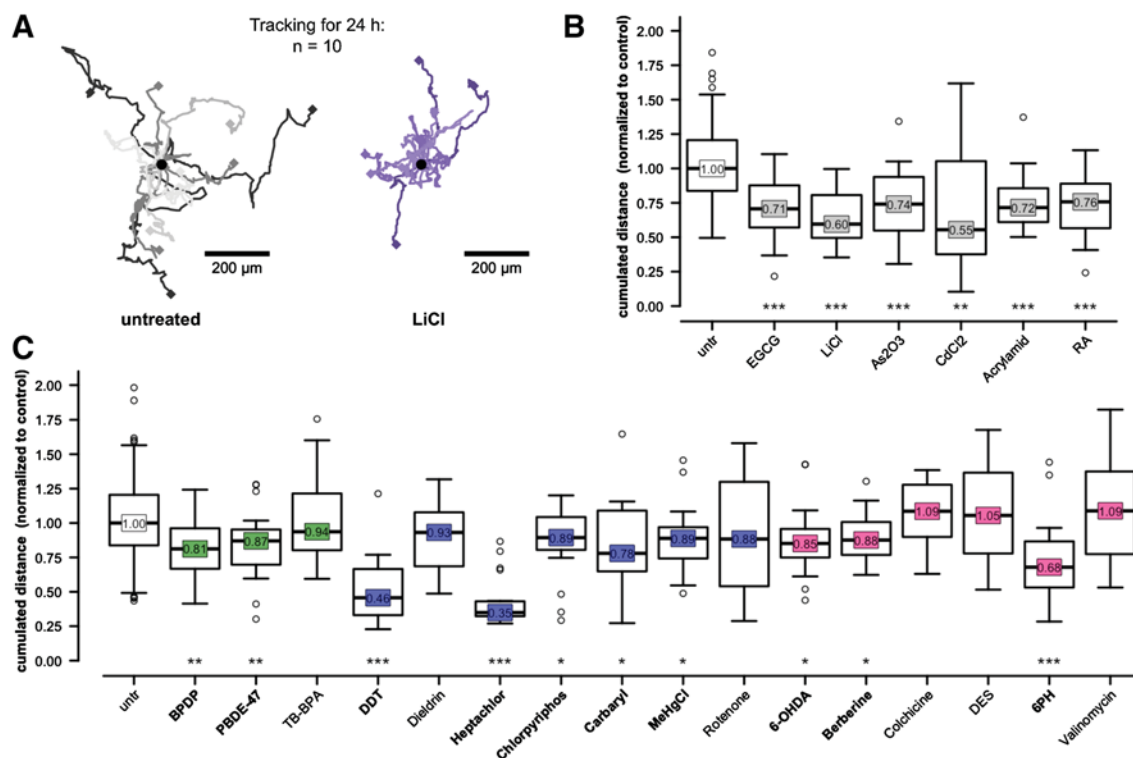


Fig. 7 Cell tracking as follow-up assay. Cells (plated at standard cMINC conditions) were imaged during the 24-h treatment period to assess the cumulated travelled distance of individual cells by manual tracking. **a** Visualisation of cell tracks for untreated cells and cells treated with 20 mM LiCl. **b** Establishment of the cell tracking assay using endpoint-specific controls and compounds known to inhibit NCC migration in the cMINC assay. For each treatment condition, 20 individual cells were tracked and the cumulated travelled distance was normalised to the untreated control. The box represents the first and third quartiles, the black line the median (number given in the rectangle) and the whiskers are at an interquartile range of 1.5. **c** Per-

formance of the screen hits in the cell tracking assay: green flame retardants; blue pesticides; pink drug-like compounds. Data are from two independent biological experiments, i.e. the box represents data from 20 tracked cells. Statistical significance of differences between treated and non-treated conditions was calculated using a Wilcoxon rank sum test. Statistical indicators are given below the boxes $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. Names of screen compounds that reduced the speed of migration are bold. EGCG epigallocatechin gallate, RA retinoic acid, TB-BPA 3,3',5,5'-tetrabromobisphenol A, 6-OHDA 6-hydroxydopamine hydrochloride, DES diethylstilbestrol, 6PH hexachlorophene (colour figure online)

of organophosphorus FR and PBDE-47 for the group of PBDEs. The test revealed that only four of the cMINC hits reduced cell speed to a similar extent as the positive controls (by $\geq 20\%$): DDT, heptachlor, carbaryl and hexachlorophene (Fig. 7c). For these toxicants, a reduction of cell speed appears to be a major mechanism responsible for reduced migration. Compounds like TB-BPA, rotenone or valinomycin did not affect cell speed at all. For such toxicants, the effect measured on a population level in the cMINC is likely caused by a change of the percentage of migratory cells, by altered directionality behaviour or by indirect effects that would need further investigation. For a large group of compounds, the reduction of speed was small, but statistically significant. It is not clear at present whether such relatively minor effects are of biological significance. More elaboration on the underlying mechanisms, extensive testing of additional positive and negative controls to establish a robust prediction model and more refined statistics would be required to come to a clearer conclusion on the impact that toxicants have on NCC migration speed.

However, also the presently available data set clearly indicated that the choice of the analytical endpoints can lead to different conclusions (hit subsets) in chemical screens. Moreover, it became clear that one key neurodevelopmental process, such as NCC migration, has a complex underlying biology that may be only partially captured by any given single test. In the case examined here, there were, for instance, compounds that reduced single-cell migration speed, while others did not. Compounds from both groups reduced the number of cells found in the circular area of the cMINC and had thus been classified as migration-inhibitors.

Re-testing of confirmed hits under conditions that prevent proliferation

Differences between the cMINC and the cell tracking assay may be re-conciled if compounds not affecting the migration speed could be shown to inhibit cell proliferation, and thus to be false positives in the cMINC. Therefore, the cMINC assay was repeated for all confirmed hits under conditions that do not allow any NCC proliferation (1 μM cytosine arabinoside, AraC, (Nyffeler et al. 2016). All toxicants were tested at the EC90V (same concentration as for cell tracking), and a similar threshold for impairment of function (i.e. 20% reduction) was set. A full prediction model has not been developed for the cMINC in the presence of AraC, but for the purpose of this comparison, we considered it plausible that a reduction of cells in the migration area to 80% or below can be considered

(1) biologically relevant, and (2) comparable to the tracking assay. Most of the compounds were confirmed to be migration-inhibitors under these conditions (Fig. S6). Thus, the number of artefacts induced by effects on proliferation in the standard cMINC was low, and the data in the absence and presence of AraC were well-correlated (Fig. S6B). Only TPP, PBDE-153, MeHgCl, berberine chloride and carbaryl were not confirmed as hits. This puts a direct effect of these compounds on NCC migration at question. In order to measure the impact of chemicals on migration, independent of proliferation, and in the absence of an additional toxicant (AraC), a short-term migration test was established.

Re-testing in a short-term transwell migration assay

Migration through a membrane, along a chemotactic gradient can be tested within few hours, and is thus little influenced by proliferation in a cell population. Therefore, a transwell migration assay was established. For this purpose, cells were seeded in inserts (upper compartment of a two-compartment dish). FBS was added to the lower compartment to stimulate the cells to migrate into the direction of this chemoattractant. Under these conditions, control NCC translocated through the porous membrane into the lower compartment (Fig. 8a). After 6 h, migrated cells were fixed, stained, imaged and manually counted. In a typical experiment, stimulation with 5% FBS led to the migration of 100 cells/image field, while < 3 cells migrated in the absence of FBS (Fig. 8b). The endpoint-specific controls CytoD, taxol and EGCG reduced transwell migration by $> 50\%$, while several known unspecific toxicants (Nyffeler et al. 2016) reduced transwell migration by $\leq 25\%$ (Fig. 8c). Based on these data, the threshold for migration-inhibition was set at 25% inhibition of migration. As next step, five well-established cMINC toxicants (Nyffeler et al. 2016) were tested at their EC90V or a corresponding reasonable test concentration. LiCl strongly ($> 50\%$) inhibited transwell migration, As_2O_3 , CdCl_2 and acrylamide had a moderate effect (25–50% range). Only retinoic acid (albeit a strong migration-inhibitor in the cMINC assay) did not specifically affect transwell migration ($< 25\%$). Thus, this small evaluation of the assay showed that the hits largely overlapped with the cMINC, but that some toxicants showed clear differences. One potential reason may be that the assays measure short-term vs long-term consequences of chemical exposure, which may be particularly important for slow-acting nuclear receptor ligands, such as retinoic acid. Another conceptual difference of the assays is the distance of migration required to lead to a positive effect. In the transwell assay, mainly the ‘mobility’ of a cell is assessed. Cells either translocate across the very thin membrane (at any

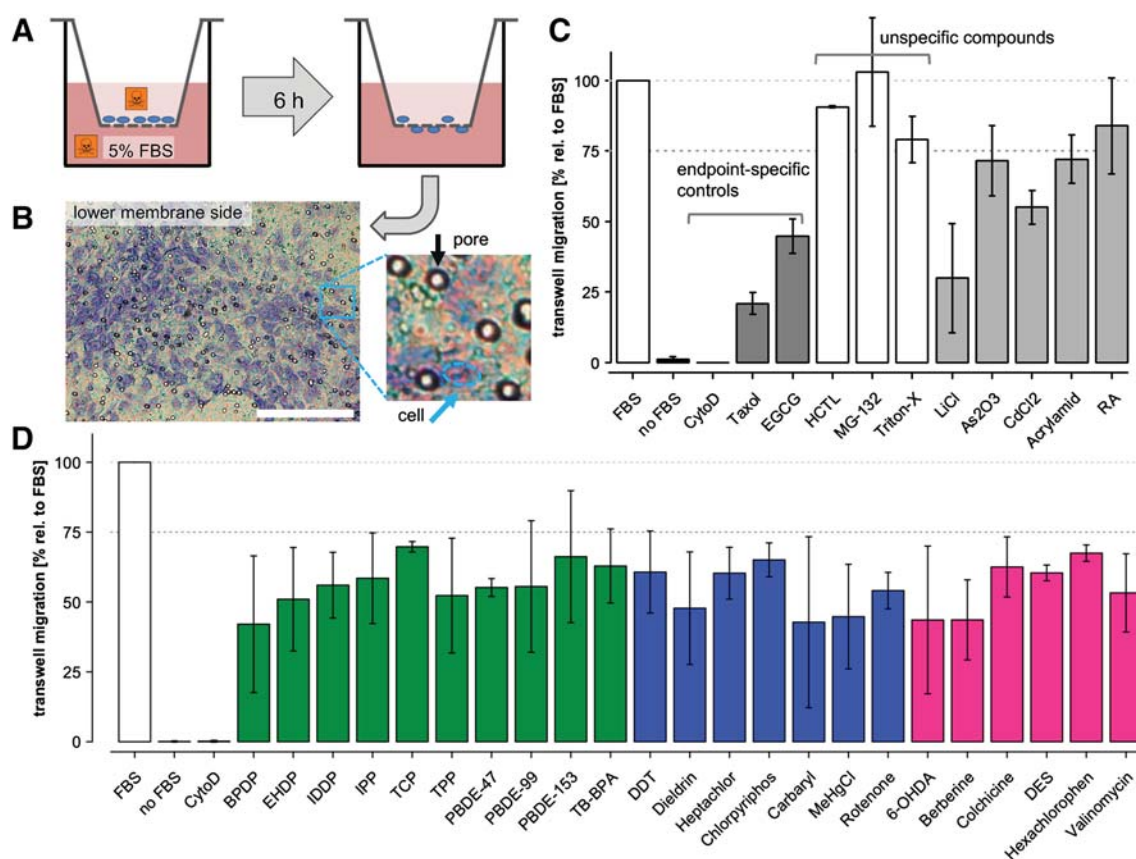


Fig. 8 Transwell migration assay as follow-up assay. **a** Schematic indicating the assay principle: cells were plated into transwell inserts and stimulated to migrate with 5% foetal bovine serum (FBS) in the lower compartment. Test items were added to both compartments. After 6 h, cells were fixed, stained and the number of cells that migrated through the transwell membrane were manually counted. **b** Exemplary microscopy image of migrated cells (*purple blotches*) on the lower surface of the membrane. The *scale bar* corresponds to 150 μm . **c** Benchmarking of the transwell assay with endpoint-specific controls (*dark grey*), unspecific compounds (*white*) and compounds well established to inhibit NCC migration in the cMINC assay (*light grey*). Transwell migration is measured as the number

of migrated cells relative to cells stimulated with 5% FBS alone. The *dotted grey line* at 75% separates unspecific compounds from endpoint-specific controls. Data are expressed as mean \pm SD from at least two independent biological experiments. **d** Performance of the screen hits in the transwell migration assay: *green* flame retardants; *blue* pesticides; *pink* drug-like compounds. Data are expressed as mean \pm SD from three independent biological experiments. *FBS* foetal bovine serum, *CytoD* cytochalasin D, *EGCG* epigallocatechin gallate, *HCTL* L-homocysteine thiolactone, *RA* retinoic acid, *TB-BPA* 3,3',5,5'-tetrabromobisphenol A, *6-OHDA* 6-hydroxydopamine hydrochloride, *DES* diethylstilbestrol (colour figure online)

time point of the 6-h assay window), or they remain in the upper compartment. Thus, we hypothesise that compounds that slow down cells rather than immobilise them would not reduce transwell migration.

Eventually, all cMINC hits were tested at their EC90V (or highest available concentration), and all compounds reduced transwell-migration by 30–45% (Fig. 8d). Notably, also compounds with small effects in the cMINC assay like berberine chloride, PBDE-99 and MeHgCl had a robust effect in the transwell assay.

In summary, all hits from the cMINC were confirmed in this third follow-up migration assay, and in general, it is likely that most cMINC hits from this screen do indeed specifically affect NCC migration.

Conclusions and outlook

The screen of the NTP80-list of 75 compounds for inhibition of migration of NCCs yielded 26 initial screen hits, 23 of which were confirmed and quantified as to their efficacy, potency and specificity. All hits fell into the classes of FR, pesticides or drug-like compounds, whereas all PAHs and industrial chemicals were negative. All of the confirmed hits in our screen (with the exception of TB-BPA) were among the active compounds identified in a neurite-outgrowth screen (Ryan et al. 2016), and five of the six compounds affecting neurites selectively were also among the hits that specifically affected NCC migration.

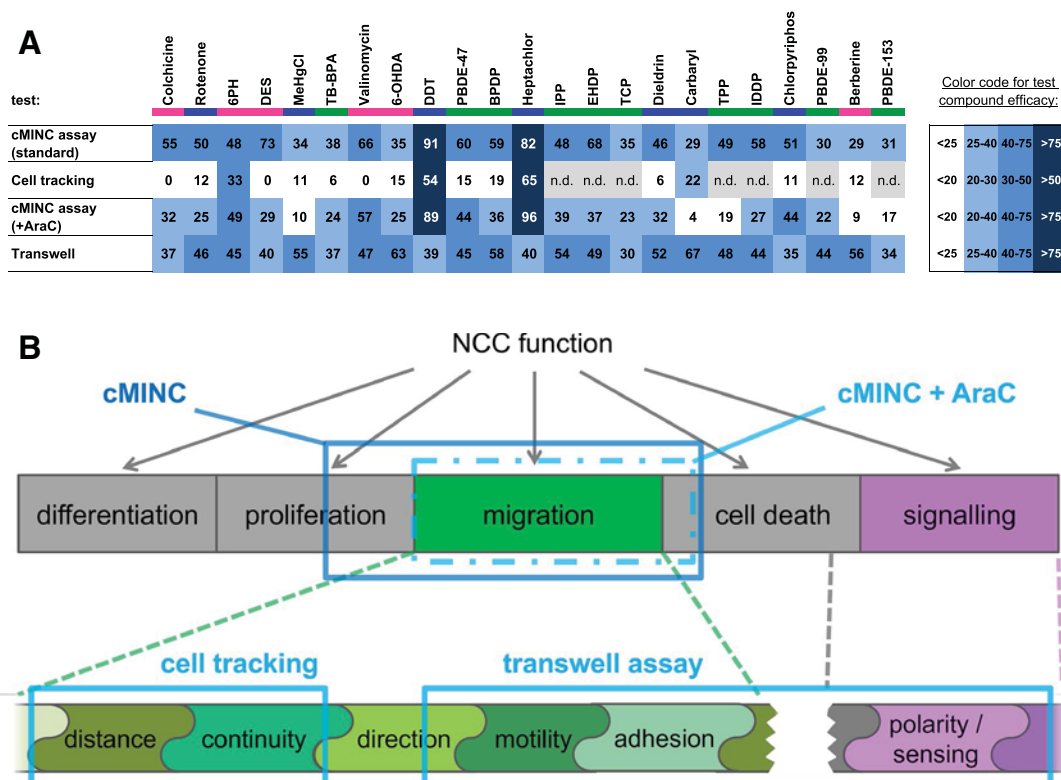


Fig. 9 Summary of assay data and biological features captured. The results of the cMINC assay, the re-testing in absence of proliferation, the transwell assay and the cell tracking were summarised in *tabular form*. The *colour code* is explained in the legend to the right. Briefly: *white* indicates a lack of effect. The intensity of *blue* indicates the efficacy. *n.d.* not determined, *TB-BPA* 3,3',5,5'-tetrabromobisphenol A, *6-OHDA* 6-hydroxydopamine hydrochloride, *DES* diethylstilbes-

tol, *6PH* hexachlorophene. **b** The *diagram* indicates biological processes captured by the different assays. The *top band* indicates various NCC functions. The *blue boxes* indicate which NCC functions are captured by the different assays. The *bottom band* is an expansion from the upper band, with a focus on signalling and different migration subfunctions. The *blue boxes* indicate which functions are measured by the cell tracking and transwell assays

The potential developmental neurotoxicants were then investigated in three follow-up tests, and all compounds were positive in at least one additional assay (Fig. 9a). A comparison of the hits showed that the two organochlorines, DDT and heptachlor, had the strongest effects on NCC migration, albeit low micromolar concentrations were required for this manifestation of toxicity. Comparison of hits across multiple other published assays indicated that NCCs seem to be particularly sensitive to hexachlorophene and diethylstilbestrol as compared to other cell types or biological processes (Fig. 6d). The chemical groups with the highest hit rate were the organophosphorus FR. Our data suggest that such compounds, if foetal concentrations >1 μM are reached, might disturb development similar to the PBDE FR that are being phased out from the markets because of their hazard (Linares et al. 2015). In this context, it is noteworthy that novel templates for FR have been recently characterised and that these were considerably less toxic to NCCs (Hirsch et al. 2016).

An important question raised by our study is, "why do different migration assays yield different hits?". We favour

the explanation that (1) each 'fundamental neurodevelopmental process' is a complex assembly of cellular functions; (2) it combines multiple adverse outcome pathways (AOPs) and can, therefore, be affected by many chemicals in different ways; and (3) each test method focusses on different aspects within a fundamental neurodevelopmental process and, therefore, reflects key events of different AOPs. Some empirical evidence suggests that this hypothesis holds true in many cases: for instance, 'neurite outgrowth' could actually mean in a given test: dendrite growth, axon growth, axon branching, growth cone extension, etc. It could also be the extension of already well-established neurites, or rather be the initiation of neurite formation; or it could refer to peripheral neurons or central neurons. There is clear evidence that such processes may be affected differentially (Hoelting et al. 2016; Kim et al. 2002). A similar heterogeneity holds true for neuronal network activity (e.g. type of neurons used; presence or absence of glia), and also for migration (Fig. 9b).

Different migration assays are complementary or partially overlapping, but never identical. Apart from the

above arguments, identity is also not possible, because each experimental system and test protocol is prone to different sets of artefacts, false positives and false negatives. For instance, the cMINC not only measures migration, but the test endpoint may also be affected by proliferation, cell death, and signalling disturbances (Fig. 9b, top row—blue box). The contributions of various biological processes are different for each migration assay. Moreover, migration itself is a highly complex process, involving several cell biological functions (Fig. 9b, bottom row). In the cMINC assay, migration is measured as the number of cells in the migration zone at the end of the assay. A reduction can be caused by several mechanisms and the assay does not provide information about which subfunction was affected. Successful migration requires that cells adhere to the substrate, but that they can also sever bonds to the extracellular matrix (adhesion). A cell also needs to be able to remodel the cytoskeleton and cytoplasmic structure in the right way (motility). Moreover, cell migration normally is a cyclic sequence of moving phases and rest phases, and the ratio of these phases is described by ‘continuity’. Finally, cells can change direction randomly, or rather keep an overall direction (directionality) and this has in the end a large effect on the overall net distance covered.

The four assays used here measure different sets of the above migration features. The cell tracking assay captures ‘distance’ and ‘continuity’ but excludes cell proliferation and cell death. The transwell assay captures ‘motility’ and ‘adhesion’. As the cells need to react to the chemoattractant stimulus, also some aspects of cell signalling, (‘polarity’ and ‘sensing’) are important, whereas the sub-functions that dominate the cell tracking assay are negligible. As a confounding factor, cell death might also be involved in the transwell assay, whereas proliferation effects are negligible due to the short assay time. The assays presented in this study, therefore, cover different aspects of NCC migration and NCC function. They provide thus complementary information. A single compound may have several simultaneous effects on cells, e.g. change adhesion, change cell stiffness, change cell speed and hence have an effect in several assays, but not necessarily in all. This explains the differential activity pattern observed here (Fig. 9a). With respect to a test strategy, we conclude from this that it is dangerous to rely on a single test if a big and complex biological process like cell migration is investigated. For practical purposes, we suggest to use the cMINC as primary screening assay, and to characterise the hits with at least one further (better more) secondary assay.

In order to assemble more and more information on screen hits, three forms of follow-up need to be considered, and some examples for them have been provided in our study: (1) technical follow-up (confirmation of findings in the same assay): in the present study, 23 of 26

screen hits were confirmed in the hit-confirmation phase. This high level of confirmation (88%) indicates that the assay is technically robust, but not perfect; (2) biological consistency (confirmation of findings in another assay for the same endpoint): the primary hits were tested in three follow-up assays, and the transwell assay confirmed findings from the cMINC. The other assays showed a more heterogeneous picture. As a general strategy, a given test could be combined with other cell function assays (Hirsch et al. 2016), with biochemical assays (Ryan et al. 2016), with tests on simple model organisms (Behl et al. 2016) or with transcriptome analysis (Pallocca et al. 2016) to provide more biological information; (3) mechanistic consistency amongst related compounds: for drug discovery screens it is common to look for structurally related compounds that give consistent hit patterns. For hit singletons, the likelihood of artefacts and wrong conclusions is known to be high. For toxicological screens, there are usually not enough compounds for direct QSAR or read-across approaches, but corresponding compounds may be selected for a follow-up. In the present study, all PBDEs were among the hits and all organophosphorus FR with aromatic side chains, whereas aliphatic ones were negative. Thus, for such compounds, there is higher confidence that this type of basic structure is linked to a hazard. For the organochlorines, there were clearly positive and negative compounds (for example, lindane was negative). Clear consistency has been shown here for DDT congeners that showed comparable efficacy.

Another form of mechanistic consistency testing involves exploration of a suggested mode of action. For instance, identification of chlorpyrifos as a hit would suggest that acetylcholine esterase may be a target. This could be tested by examination of other (structurally unrelated) choline esterase inhibitors. Alternatively, chlorpyrifos oxone may be tested and should be more potent if the common target was choline esterase. Similar considerations would apply to the follow-up of parathion as a hit by testing of paraoxon, or to the consideration whether the oestrogen receptor plays a role in the toxicity of diethylstilbestrol. One may follow up with several estrogenic compounds, but the high concentration of diethylstilbestrol required to inhibit NCC migration already suggests now that this effect is independent of oestrogen receptor activation.

In summary, our study has provided rich information on issues and potential solutions associated with the use of cell function assays for screening. For the future, we would like to pinpoint some important areas of work. For instance, a hit confirmation rate of 80% may sound good. However, if five such assays with 20% false positives are combined, the overall false positive rate (probability of a compound being a hit in at least one of the five assays) would be 67%. How will the field deal with this? One approach could be

to combine tests in biological networks guided by molecular mechanisms/AOPs (Jaworska et al. 2015; Juberg et al. 2016), or to use biological information across assays in form of a biological read across (Hartung 2016; Patlewicz and Fitzpatrick 2016; Strickland et al. 2016a, b). Another old, but continuously present and topical issue (Judson et al. 2016; Leist et al. 2010) is the problem of cytotoxicity. The definition of hit specificity is often arbitrary, and more experience needs to be collected for each test to differentiate real hits from those that only reflect cytotoxicity. To advance on all these questions, it will be most important to run the same set of compounds described here in as many tests as possible, or to take similar approaches with other compound sets.

Acknowledgements This work was supported by the Land BW, the Doerenkamp-Zbinden Foundation, the DFG (RTG1331, KoRS-CB) and the European Project EU-ToxRisk. We are grateful to M. Kapitzka, H. Leisner and the staff of the bioimaging center (BIC) and the flow cytometry center (FlowKon) for invaluable experimental support.

Compliance with ethical standards

Conflict of interest The authors declare no conflict of interest.

References

- Adler S et al (2011) Alternative (non-animal) methods for cosmetics testing: current status and future prospects-2010. *Arch Toxicol* 85:367–485. doi:10.1007/s00204-011-0693-2
- Aschner M et al (2016) Reference compounds for alternative test methods to indicate developmental neurotoxicity (DNT) potential of chemicals: example lists and criteria for their selection and use. *Altex*. doi:10.14573/altex.1604201
- Bal-Price AK et al (2012) Advancing the science of developmental neurotoxicity (DNT): testing for better safety evaluation. *Altex* 29:202–215
- Bal-Price A et al (2015) International STakeholder NETwork (IST-NET): creating a developmental neurotoxicity (DNT) testing road map for regulatory purposes. *Arch Toxicol* 89:269–287. doi:10.1007/s00204-015-1464-2
- Barenys M et al (2016) Epigallocatechin gallate (EGCG) inhibits adhesion and migration of neural progenitor cells in vitro. *Arch Toxicol*. doi:10.1007/s00204-016-1709-8
- Basketter DA, White IR, McFadden JP, Kimber I (2015) Skin sensitization: implications for integration of clinical data into hazard identification and risk assessment. *Hum Exp Toxicol* 34:1222–1230. doi:10.1177/0960327115601760
- Behl M et al (2016) Editor's highlight: Comparative Toxicity of Organophosphate Flame Retardants and Polybrominated Diphenyl Ethers to *Caenorhabditis elegans*. *Toxicol Sci* 154:241–252. doi:10.1093/toxsci/kfw162
- Browne P, Judson RS, Casey WM, Kleinstreuer NC, Thomas RS (2015) Screening chemicals for estrogen receptor bioactivity using a computational model. *Environ Sci Technol* 49:8804–8814. doi:10.1021/acs.est.5b02641
- Colaianna M et al (2016) Fingerprinting of neurotoxic compounds using a mouse embryonic stem cell dual luminescence reporter assay. *Arch Toxicol*. doi:10.1007/s00204-016-1690-2
- Collins FS, Gray GM, Bucher JR (2008) Toxicology. Transforming environmental health protection. *Science* 319:906–907. doi:10.1126/science.1154619
- Combes R, Grindon C, Cronin MT, Roberts DW, Garrod JF (2008) Integrated decision-tree testing strategies for mutagenicity and carcinogenicity with respect to the requirements of the EU REACH legislation. *Altern Lab Anim* 36(Suppl 1):43–63
- Crofton KM et al (2011) Developmental neurotoxicity testing: recommendations for developing alternative methods for the screening and prioritization of chemicals. *Altex* 28:9–15
- Crofton KM, Mundy WR, Shafer TJ (2012) Developmental neurotoxicity testing: a path forward. *Congenit Anom (Kyoto)* 52:140–146. doi:10.1111/j.1741-4520.2012.00377.x
- Dresler N et al (2015) Grouping of histone deacetylase inhibitors and other toxicants disturbing neural crest migration by transcriptional profiling. *Neurotoxicology* 50:56–70. doi:10.1016/j.neuro.2015.07.008
- Duran A, Zamora I, Pastor M (2009) Suitability of GRIND-based principal properties for the description of molecular similarity and ligand-based virtual screening. *J Chem Inf Model* 49:2129–2138. doi:10.1021/ci900228x
- Ezendam J, Braakhuis HM, Vandebriel RJ (2016) State of the art in non-animal approaches for skin sensitization testing: from individual test methods towards testing strategies. *Arch Toxicol* 90:2861–2883. doi:10.1007/s00204-016-1842-4
- Fritsche E, Cline JE, Nguyen N-H, Scanlan TS, Abel J (2005) Polychlorinated biphenyls disturb differentiation of normal human neural progenitor cells: clue for involvement of thyroid hormone receptors. *Environ Health Perspect* 113:871–876. doi:10.1289/ehp.7793
- Grinberg M et al (2014) Toxicogenomics directory of chemically exposed human hepatocytes. *Arch Toxicol* 88:2261–2287. doi:10.1007/s00204-014-1400-x
- Hartung T (2016) Making big sense from big data in toxicology by read-across. *Altex* 33:83–93. doi:10.14573/altex.1603091
- Hirsch C et al (2016) Multiparameter toxicity assessment of novel DOPO-derived organophosphorus flame retardants. *Arch Toxicol*. doi:10.1007/s00204-016-1680-4
- Hoelting L et al (2016) Stem cell-derived immature human dorsal root ganglia neurons to identify peripheral neurotoxicants. *Stem Cells Transl Med* 5:476–487. doi:10.5966/sctm.2015-0108
- Huang R et al (2016) Modelling the Tox21 10 K chemical profiles for in vivo toxicity prediction and mechanism characterization. *Nat Commun* 7:10425. doi:10.1038/ncomms10425
- Igarashi Y, Nakatsu N, Yamashita T, Ono A, Ohno Y, Urushidani T, Yamada H (2015) Open TG-GATEs: a large-scale toxicogenomics database. *Nucleic Acids Res* 43:D921–D927. doi:10.1093/nar/gku955
- Jaworska JS, Natsch A, Ryan C, Strickland J, Ashikaga T, Miyazawa M (2015) Bayesian integrated testing strategy (ITS) for skin sensitization potency assessment: a decision support system for quantitative weight of evidence and adaptive testing strategy. *Arch Toxicol* 89:2355–2383. doi:10.1007/s00204-015-1634-2
- Juberg DR et al (2016) FutureTox III: bridges for translation. *Toxicol Sci*. doi:10.1093/toxsci/kfw194
- Judson R et al (2013) Perspectives on validation of high-throughput assays supporting 21st century toxicity testing. *Altex* 30:51–56
- Judson R et al (2014) In vitro and modelling approaches to risk assessment from the U.S. Environmental Protection Agency ToxCast programme. *Basic Clin Pharmacol Toxicol* 115:69–76. doi:10.1111/bcpt.12239
- Judson R et al (2016) Analysis of the effects of cell stress and cytotoxicity on in vitro assay activity across a diverse chemical and assay space. *Toxicol Sci* 153:409. doi:10.1093/toxsci/kfw148
- Kaneko N, Sawada M, Sawamoto K (2017) Mechanisms of neuronal migration in the adult brain. *J Neurochem*. doi:10.1111/jnc.14002

- Kim JJ, Beck HN, Lein PJ, Higgins D (2002) Interferon gamma induces retrograde dendritic retraction and inhibits synapse formation. *J Neurosci* 22(11):4530–4539
- Kroese ED et al (2015) Evaluation of an alternative in vitro test battery for detecting reproductive toxicants in a grouping context. *Reprod Toxicol* 55:11–19. doi:10.1016/j.reprotox.2014.10.003
- Krug AK, Balmer NV, Matt F, Schonenberger F, Merhof D, Leist M (2013) Evaluation of a human neurite growth assay as specific screen for developmental neurotoxicants. *Arch Toxicol* 87:2215–2231. doi:10.1007/s00204-013-1072-y
- Leist M, Hartung T, Nicotera P (2008) The dawning of a new age of toxicology. *Altex* 25:103–114
- Leist M, Efremova L, Karreman C (2010) Food for thought ... considerations and guidelines for basic test method descriptions in toxicology. *Altex* 27:309–317
- Leist M et al (2014) Consensus report on the future of animal-free systemic toxicity testing. *Altex* 31:341–356. doi:10.14573/altex.1406091
- Linares V, Belles M, Domingo JL (2015) Human exposure to PBDE and critical evaluation of health hazards. *Arch Toxicol* 89:335–356. doi:10.1007/s00204-015-1457-1
- Luhmann HJ, Fukuda A, Kilb W (2015) Control of cortical neuronal migration by glutamate and GABA. *Front Cell Neurosci* 9:4. doi:10.3389/fncel.2015.00004
- Milletti F, Storchi L, Sforna G, Cruciani G (2007) New and original pKa prediction method using grid molecular interaction fields. *J Chem Inf Model* 47:2172–2181. doi:10.1021/ci700018y
- Moors M et al (2009) Human neurospheres as three-dimensional cellular systems for developmental neurotoxicity testing. *Environ Health Perspect* 117:1131–1138. doi:10.1289/ehp.0800207
- Nyffeler J, Karreman C, Leisner H, Kim YJ, Lee G, Waldmann T, Leist M (2016) Design of a high-throughput human neural crest cell migration assay to indicate potential developmental toxicants. *Altex*. doi:10.14573/altex.1605031
- Pallocca G et al (2016) Identification of transcriptome signatures and biomarkers specific for potential developmental toxicants inhibiting human neural crest cell migration. *Arch Toxicol* 90:159–180. doi:10.1007/s00204-015-1658-7
- Pastor M, Cruciani G, McLay I, Pickett S, Clementi S (2000) GRIND-INdependent descriptors (GRIND): a novel class of alignment-independent three-dimensional molecular descriptors. *J Med Chem* 43:3233–3243
- Patlewicz G, Fitzpatrick JM (2016) Current and future perspectives on the development, evaluation, and application of in silico approaches for predicting toxicity. *Chem Res Toxicol* 29:438–451. doi:10.1021/acs.chemrestox.5b00388
- Pei Y et al (2015) Comparative neurotoxicity screening in human iPSC-derived neural stem cells, neurons and astrocytes. *Brain Res*. doi:10.1016/j.brainres.2015.07.048
- Piersma AH et al (2013) Evaluation of an alternative in vitro test battery for detecting reproductive toxicants. *Reprod Toxicol* 38:53–64. doi:10.1016/j.reprotox.2013.03.002
- R Core Team (2015) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
- Reisinger K et al (2015) Systematic evaluation of non-animal test methods for skin sensitisation safety assessment. *Toxicol In Vitro* 29:259–270. doi:10.1016/j.tiv.2014.10.018
- Richard AM et al (2016) ToxCast chemical landscape: paving the road to 21st century Toxicology. *Chem Res Toxicol* 29:1225–1251. doi:10.1021/acs.chemrestox.6b00135
- Ritz C, Streibig JC (2005) Bioassay analysis using R. *J Stat Softw* 12:1–22
- Robinson JF, Piersma AH (2013) Toxicogenomic approaches in developmental toxicology testing. *Methods Mol Biol* 947:451–473. doi:10.1007/978-1-62703-131-8_31
- Ryan KR, Sirenko O, Parham F, Hsieh JH, Cromwell EF, Tice RR, Behl M (2016) Neurite outgrowth in human induced pluripotent stem cell-derived neurons as a high-throughput screen for developmental neurotoxicity or neurotoxicity. *Neurotoxicology* 53:271–281. doi:10.1016/j.neuro.2016.02.003
- Sadowski J, Gasteiger J, Klebe G (1994) Comparison of automatic three-dimensional model builders using 639 X-ray structures. *J Chem Inf Comput Sci* 34:1000–1008. doi:10.1021/ci00020a039
- Schenk B et al (2010) The ReProTect Feasibility Study, a novel comprehensive in vitro approach to detect reproductive toxicants. *Reprod Toxicol* 30:200–218. doi:10.1016/j.reprotox.2010.05.012
- Schmidt BZ et al (2016) In vitro acute and developmental neurotoxicity screening: an overview of cellular platforms and high-throughput technical possibilities. *Arch Toxicol*. doi:10.1007/s00204-016-1805-9
- Schultz L et al (2015) Evaluation of drug-induced neurotoxicity based on metabolomics, proteomics and electrical activity measurements in complementary CNS in vitro models. *Toxicol In Vitro* 30:138–165. doi:10.1016/j.tiv.2015.05.016
- Shinde V et al (2016) Definition of transcriptome-based indices for quantitative characterization of chemically disturbed stem cell development: introduction of the STOP-Toxukn and STOP-Toxukk tests. *Arch Toxicol*. doi:10.1007/s00204-016-1741-8
- Shukla SJ, Huang R, Austin CP, Xia M (2010) The future of toxicity testing: a focus on in vitro methods using a quantitative high-throughput screening platform. *Drug Discov Today* 15:997–1007. doi:10.1016/j.drudis.2010.07.007
- Smirnova L, Hogberg HT, Leist M, Hartung T (2014) Developmental neurotoxicity—challenges in the 21st century and in vitro opportunities. *Altex* 31:129–156. doi:10.14573/altex.1403271
- Sonneveld E, Jansen HJ, Riteco JA, Brouwer A, van der Burg B (2005) Development of androgen- and estrogen-responsive bioassays, members of a panel of human cell line-based highly selective steroid-responsive bioassays. *Toxicol Sci* 83:136–148. doi:10.1093/toxsci/kfi005
- Stiegler NV, Krug AK, Matt F, Leist M (2011) Assessment of chemical-induced impairment of human neurite outgrowth by multiparametric live cell imaging in high-density cultures. *Toxicol Sci* 121:73–87. doi:10.1093/toxsci/kfr034
- Strickland J et al (2016a) Integrated decision strategies for skin sensitization hazard. *J Appl Toxicol* 36:1150–1162. doi:10.1002/jat.3281
- Strickland J et al (2016b) Multivariate models for prediction of human skin sensitization hazard. *J Appl Toxicol*. doi:10.1002/jat.3366
- Tice RR, Austin CP, Kavlock RJ, Bucher JR (2013) Improving the human hazard characterization of chemicals: a Tox21 update. *Environ Health Perspect* 121:756–765. doi:10.1289/ehp.1205784
- USEPA (2016) ToxCast & Tox21 Chemicals from DSS-Tox_20151019. Retrieved from <https://www.epa.gov/chemical-research/toxicity-forecaster-toxcasttm-data> on 29th July, 2016. Data released 19th Oct 2015
- van der Burg B et al (2015) The ChemScreen project to design a pragmatic alternative approach to predict reproductive toxicity of chemicals. *Reprod Toxicol* 55:114–123. doi:10.1016/j.reprotox.2015.01.008
- van der Laan JW, Chapin RE, Haenen B, Jacobs AC, Piersma A (2012) Testing strategies for embryo-fetal toxicity of human pharmaceuticals. Animal models vs. in vitro approaches: a workshop report. *Regul Toxicol Pharmacol* 63:115–123. doi:10.1016/j.yrtph.2012.03.009
- van Thriel C, Westerink RH, Beste C, Bale AS, Lein PJ, Leist M (2012) Translating neurobehavioural endpoints of developmental neurotoxicity tests into in vitro assays and readouts. *Neurotoxicology* 33:911–924. doi:10.1016/j.neuro.2011.10.002

Zimmer B, Lee G, Balmer NV, Meganathan K, Sachinidis A, Studer L, Leist M (2012) Evaluation of developmental toxicants and signaling pathways in a functional test based on the migration of human neural crest cells. *Environ Health Perspect* 120:1116–1122. doi:[10.1289/ehp.1104489](https://doi.org/10.1289/ehp.1104489)

Zimmer B et al (2014) Profiling of drugs and environmental chemicals for functional impairment of neural crest migration in a novel stem cell-based test battery. *Arch Toxicol* 88:1109–1126. doi:[10.1007/s00204-014-1231-9](https://doi.org/10.1007/s00204-014-1231-9)