Organisation for Economic Co-operation and Development

ENV/CBC/MONO(2026)17

Unclassified

English - Or. English

29 June 2026

ENVIRONMENT DIRECTORATE
CHEMICALS AND BIOTECHNOLOGY COMMITTEE

Peer Review Report of the Toxtracker Assay

OECD Series on Testing and Assessment

E-mail: ehscont@oecd.org.

JT03590129

This work was approved and declassified by the Chemicals and Biotechnology Committee on [02/07/2026].

Please cite this publication as:

OECD (2026), *Peer Review Report of the Toxtracker Assay*, OECD Series on Testing and Assessment, No. 440, OECD Environment, Health and Safety, Paris, link to ONE M&P document https://one.oecd.org/document/ENV/CBC/MONO(2026)17/en/pdf

Contact us

OECD Environment Directorate,
Environment, Health and Safety Division
2 rue André-Pascal
75775 Paris Cedex 16
France

E-mail: ehscont@oecd.org

About the OECD

The Organisation for Economic Co-operation and Development (OECD) is an intergovernmental organisation in which representatives of 38 countries in North and South America, Europe and the Asia and Pacific region, as well as the European Union, meet to co-ordinate and harmonise policies, discuss issues of mutual concern, and work together to respond to international problems. Most of the OECD's work is carried out by more than 200 specialised committees and working groups composed of member country delegates. Observers from several Partner countries and from interested international organisations attend many of the OECD's workshops and other meetings. Committees and working groups are served by the OECD Secretariat, located in Paris, France, which is organised into directorates and divisions.

The Environment, Health and Safety Division publishes free-of-charge documents in twelve different series: **Testing and Assessment**; **Good Laboratory Practice and Compliance Monitoring**; **Pesticides**; **Biocides**; **Risk Management**; **Harmonisation of Regulatory Oversight in Biotechnology**; **Safety of Novel Foods and Feeds; Chemical Accidents; Pollutant Release and Transfer Registers; Emission Scenario Documents; Safety of Manufactured Nanomaterials;** and **Adverse Outcome Pathways.** More information about the Environment, Health and Safety Programme and EHS publications is available on the OECD's World Wide Web site (https://www.oecd.org/en/topics/chemical-safety-and-biosafety.html).

This publication was developed in the IOMC context. The contents do not necessarily reflect the views or stated policies of individual IOMC Participating Organizations.

The Inter-Organisation Programme for the Sound Management of Chemicals (IOMC) was established in 1995 following recommendations made by the 1992 UN Conference on Environment and Development to strengthen co-operation and increase international co-ordination in the field of chemical safety. The Participating Organisations are FAO, ILO, UNDP, UNEP, UNIDO, UNITAR, WHO, World Bank, Basel, Rotterdam and Stockholm Conventions and OECD. The purpose of the IOMC is to promote co-ordination of the policies and activities pursued by the Participating Organisations, jointly or separately, to achieve the sound management of chemicals in relation to human health and the environment.

Foreword

This document contains the Peer review report of the ToxTracker gene reporter assay, an in vitro assay for the identification of genotoxic hazards and genotoxic/non-genotoxic mechanisms of action.

The scientific validity of the ToxTracker assay was assessed by an international Peer Review Panel in 2023. The peer review report supports the development of the corresponding test method for inclusion in a new Test Guideline, TG No 446A. The Peer review report and the Validation report of the ToxTracker assay were circulated for comments to the Working Party of the National Coordinators of the Test Guidelines Programme (WNT) in July and December 2025. The WNT approved the peer review report at its 38th meeting in April 2026. The WNT also approved the Validation report of the ToxTracker assay.

This report is published under the responsibility of the Chemicals and Biotechnology Committee.

Summary of the Peer Review of the Validation Study for ToxTracker Assay

19 December 2023

Table of Content:

Summary

1. This report presents the summary of the review of the validation study for the ToxTracker test method by independent Peer Review Panel (PRP).

2. The validation ring trial for ToxTracker was performed between 2017 and 2022. It was coordinated by an expert validation management team (VMT) supported by Toxys B.V. (The Netherlands) for training of seven naïve laboratories. A report of the validation ring trial study: "The ToxTracker assay: a stem cell-based reporter assay for mechanistic carcinogenicity hazard screening" (Attachment A) was made available to the OECD Secretariat for expert review in 2023.

3. The expert review was conducted between August and November 2023. The PRP evaluated the validation study considering eight charge questions based on the principles of the OECD GD34: *Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment*, in the context of the goals outlined in the validation study: (i) primary goal was to determine if ToxTracker is able to correctly predict the genotoxic properties of compounds; (ii) secondary goal was to investigate the MoA information that is provided by ToxTracker and the relevance for the genotoxicity prediction of the tested compounds.

4. Overall, the peer review panel considers that the validation study of the ToxTracker assay is consistent with the validation principles of the OECD GD 34 and that the validation principles are mainly met considering the primary goal of the validation study. Specifically, the PRP considered that the technical transferability, the intra- and inter-laboratory reproducibility of the ToxTracker assay were high and adequately demonstrated. The PRP also considered that the test protocol is quite comprehensive needing few clarifications. The validation study is based on application of the assay to a set of reference chemicals that cover a broad chemical space representative of the types of substances for which the test method is intended to be used in regulatory context. However, the PRP noted that the long-term commercial availability of the cell lines used for the assay needs to be considered. Also, raw data from the validation study (in addition to the results: positive, negative, equivocal) need to be made publicly available to fully address validation principle 8.

5. Based on the reference set of chemicals tested, it appears that there is a reliable relationship between the differential expression of the six selected genes measured with the ToxTracker and genotoxicity of tested substances (direct and/or indirect). Therefore, the PRP considers that development of a test guideline (TG) for assessment of genotoxicity based on the ToxTracker assay method can be supported by the validation study, considering also the findings and recommendations of the peer review.

6. With regard to the secondary goal of the validation study, the PRP considers that the assay may provide valuable MoA information, but further analysis and development is expected, for example to identify reporter gene(s) that could discriminate clastogenic and aneugenic genotoxicants. Additional analysis (e.g. Principal Component Analysis) with the raw data from the validation study would also be useful in this respect.

7. The PRP noted that in the introductory paragraphs and the title of the Validation Study Report claims are made to the use of ToxTracker for carcinogenicity hazard screening and assessment. However, given that the objectives of the validation study analysis deal with genotoxicity (direct and indicted), both in terms of prediction and MoA, the PRP highlighted their view that the authors' claims are not substantiated by the validation study.

8. The review and recommendations by the PRP considering the validation principles of transferability, reproducibility, biological relevance, regulatory need and transparency are expected to be informative for the future development of the assay and a potential TG for ToxTracker.

9. Recommendations:

- Provide more in-depth and better referenced description of the background of the development of the assay and selection of the six reporter genes. A signal transduction diagram (from sensors, transducers, and effectors, up to cell event(s)) would be helpful.
- Ensure easy public access to all raw reporter response data of the validation study
- Further improve protocol clarity (see CQ3 and Conclusions section)
- Further develop the assay to support discrimination between clastogenic and aneugenic MoA (e.g. by identification and testing of relevant reference chemicals and/or by applying additional analysis to existing and newly generated raw reporter data)
- Further development/optimisation of the assay and the prediction models to ensure good performance within the range of intermediate reporter response levels
- Consider Confidence Intervals (CIs) when calculating performance statistics (see Annex 5)
- Consider long-term commercial availability of the cell lines for ToxTracker
- Include academic research laboratories in future validation studies (general recommendation not strictly related to ToxTracker)

Background

10. Genotoxicity testing is an essential part of chemical safety assessment. Regulatory schemes across different sectors (from human and veterinary medicines via pesticides and biocides to industrial chemicals and cosmetics) rely on a set of methods for testing genotoxicity of chemicals that span from bacterial to whole animal in vivo test systems, to assess mutagenic, clastogenic and carcinogenic potential of chemicals. Alongside, a strong call for a major paradigm shift in toxicity testing drives the development of New Approach Methodologies (NAMs) that would allow chemical risk assessment to move towards the use of biologically relevant, scientifically robust, humane and more efficient, high throughput tools and practices for evaluating toxicity and safety of chemicals. NAMs which enable detecting and deciphering mechanistic aspects of genotoxicity have potential to drive the paradigm shift for the endpoints of genotoxicity and carcinogenicity in the safety assessment of chemicals.

11. ToxTracker is a mouse embryonic stem cell-based reporter assay that monitors differential expression of six biomarker genes shown to be preferentially activated upon exposure to different classes of stressors associated with genotoxicity and carcinogenicity (Hendriks et al. 2011 & 2012). Moreover, the activation of each of the six biomarkers has been associated with different cellular pathways and modes of action (MoAs) (Hendriks et al. 2016) as outlined in the table below.

Biological damage	Cellular pathway	Biomarker gene	MOA
DNA damage	ATR/Chk1 DNA damage signaling	Bscl2	Mutagenic DNA lesions
	NF-kB signaling	Rtkn	DNA double-strand breaks
Oxidative stress	Nrf2 antioxidant response	Srxn1	ROS production
	Nrf2-independent	Blvrb	ROS production
Protein damage	Unfolded protein response	Ddit3	Protein damage
Cellular stress	p53 signaling	Btg2	Cytotoxicity

12. The reporter cells for the ToxTracker assay were created after transfection of mouse embryonic stem cells with bacterial artificial chromosomes (BACs) containing C-terminal green fluorescent protein (GFP)-tagged biomarker genes, including their promoter and regulatory elements. Thus, ToxTracker test assay can detect the transcriptional regulation of the biomarker genes by their own physiological promoter in response to various stressors. The differential induction of the different GFP reporters in response to exposure to chemicals is quantified in 96-well plate format using flow-cytometry.

13. The ToxTracker assay was developed within the Department of Toxicogenetics, Leiden Amsterdam Center for Drug Research (LACDR), at Leiden University Medical Center (Hendriks et al. 2011). It has been further optimised by the biotech spin-off Toxys B.V. Leiden, The Netherlands (Brandsma I., et al. 2020).

14. In 2016, ToxTracker was introduced to the OECD Test Guideline Programme (TPG) by Netherlands. It was accepted on the TPG workplan (Project 4.125) with the outline for validation study based on the OECD Guidance Document 34 - *Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment* (OECD GD 34), followed by an expert peer review to assess the validation status

of the assay and the readiness of the ToxTracker assay for development of an OECD Test Guideline based on the results from the Validation Study Report (Attachment A).

15. The validation study was performed between 2017 and 2022 coordinated by an independent Validation Management Team (VMT) supported by the developer Toxys for the initial training/onboarding of the assay in the participating laboratories (7 'naïve' laboratories from different industries in the USA, UK, Canada and Switzerland) and for compiling the Validation Study Report. The VMT, excluding Toxys, was responsible for defining the validation project structure, selection of the partner laboratories, setting the different milestones for the project and analysis of the test results.

16. The purpose of the validation study, as outlined in the study report, was to establish transferability and reproducibility of the assay with regard to two related but distinct goals. The primary goal was to determine if ToxTracker is able to correctly predict the genotoxic properties of compounds. For this a broad selection of well-established genotoxic and non-genotoxic substances were selected to establish the transferability, sensitivity and specificity of ToxTracker. The secondary goal was to investigate the mechanistic/MoA information provided by ToxTracker and its relevance for genotoxicity predictions of the tested substances (direct and/or indirect genotoxicity). For this, the results from ToxTracker concerning the molecular perturbations induced by the reference chemicals were analysed in relation to results from the current standard in vitro and in vivo genotoxicity assays and existing knowledge on MoA (see paragraphs 4 and 5 of the Validation Study Report in Attachment A).

The peer review process

17. The Peer Review Panel (PRP) was established in August 2023 in order to provide an independent review of the validation of the ToxTracker assay according to the principles of the OECD GD 34 (OECD, 2005) considering the purpose and goals of the validation study (see above paragraphs 3 and 16 of this report and paragraphs 4/5 of the Validation Study Report).

18. The review experts were nominated by the OECD Working Party of the National Coordinators of the Test Guidelines Programme (WNT) and were then approached by an independent Peer Review Manager (PRM) who was contracted by the OECD to coordinate the work. The members of the Panel are listed in Annex 1.

19. The PRP was asked to review the document: '*The ToxTracker assay: a stem cell-based reporter assay for mechanistic carcinogenicity hazard screening dated 2023 Version 2.3*' (made available on the OECD public website in July 2023), and respond to charge questions (outlined in Annex 2) based on the OECD GD 34: *Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment*.

20. In addition to the validation study, the following articles were provided to the PRP as supporting information:

    Hendricks et al., 2012 - doi:10.1093/toxsci/kfr281
    Hendricks et al., 2016 - doi: 10.1093/toxsci/kfv323
    Wills et al., 2021 - doi: 10.1093/mutage/geab020
    Boisvert et al., 2023 - doi: 10.1002/em.22525

21. Each Panel member provided written responses to the charge questions to the PRM by 20 September 2023. PRM provided the PRP with the compilation of the written comments (Annex 3) and highlighted points requiring further discussion or clarification. The developers of the ToxTracker assay were also asked to provide some written responses (Annex 4) before the end-of-review teleconference on 23 October 2023.

22. Based on the teleconference discussion and the written responses by the assay developers, the PRM prepared a draft report that was reviewed and agreed by the PRP.

23. This report presents the summary of the assessment of the validation study report on the ToxTracker and the resulting agreed responses of the PRP to each of the charge questions.

Scientific basis and Relevance to an endpoint of interest (Validation principles 1 and 2)

Charge Question 1: *Do you consider that the rationale for the ToxTracker test method is clearly elaborated in the validation study and the other supporting documents in terms of its scientific basis, regulatory purpose and need? Consider the potential added value of the mechanistic information for genotoxicity generated with this assay.*

24. The written comments reflect agreement that the rationale for the ToxTracker assay in relation to identification of genotoxic substances is clearly elaborated. The assay is based on expression profile of six genes that are specifically activated by different cellular signalling responses associated with (direct and indirect) genotoxicity. As such it can represent a valuable expansion of the in vitro genotoxicity testing battery for regulatory applications.

25. The additional information generated by ToxTracker and relating to mechanistic aspects of cellular perturbations is considered valuable for developing AOPs and to support regulatory conclusions by integrating it into a weight-of-evidence approach. It was noted that the additional mechanistic information provided by ToxTracker is not intended to replace the established in vitro and in vivo genotoxicity tests, but rather to clarify and/or elucidate the mechanism of (geno)toxicity (mutagenicity, clastogenicity, oxidative stress, protein stress and p53 activation). While the value of the additional mechanistic information was recognised, considering the goals of the validation study, the focus of the reviewer's comments was on elucidating mechanisms of genotoxicity (direct and indirect) and not on non-genotoxic mechanisms that may also be related to carcinogenicity and other adverse cellular processes.

26. At the teleconference the PRP discussed the usefulness of the ToxTracker information for establishing relevant reference doses (e.g. BMD or other values) for regulatory health assessment. It was concluded that the assay and the protocol/data interpretation model applied within the validation study is not suitable, at present, for developing reference values for regulatory assessment. However, the additional mechanistic information generated by ToxTracker, particularly when indirect genotoxicity is indicated, could inform further testing for deriving threshold values.

27. In their written comments the PRP also provided some considerations for the context of potential use of ToxTracker in the future:

    •Consider comparison of the advantages and limitations of the ToxTracker assay and other similar methods that provide mechanistic information related to genotoxicity (e.g. TGx-DDI and DNA damage gene expression assays, FISH assay for aneuploidy, methods targeting ROS, and oxidation of DNA bases such as 8-OH-dG, etc.).

    •Clarify whether mechanistic data are required as part of the basic testing battery for genotoxicity or only when a positive response is observed

    •Consider additional information for usefulness of mechanistic information in regulatory application (Menz et al., 2023).

Overall, the peer review panel considers that the Validation Principle 1 has been partially met. The scientific basis of the assay (induction of reporter genes by different genotoxic substances) is elaborated, but further illustration and reference to primary literature is needed. The mechanistic information provided by the assay could be useful for moving towards the new paradigm of a mechanistically based toxicity assessment in the future following further development and analysis. However at the present level of development the assay can only be a complement to the existing battery of in vitro genotoxicity assays.

Charge Question 2: *Is the relationship between the endpoint(s) measured with the ToxTracker (differential expression of the six selected genes) and the (biological) phenomenon of interest (genotoxicity) described adequately and using relevant scientific information/references?*

28. The written comments reflect agreement that the validation study shows a reliable relationship between the differential expression of the six selected genes measured with the ToxTracker and genotoxicity (direct and/or indirect).

29. Data generated by ToxTracker could also potentially provide useful mechanistic information relevant to elucidate the MoA of genotoxic substances. However, PRP highlighted that the validation study itself does not address the relevance of the ToxTracker assay to distinguish between aneugenicity and clastogenicity. According to the ToxTracker data interpretation procedure (page 70 of the Validation Report – Attachment A), substances with Rtkn positive and Bscl2 negative result are considered potentially aneugenic substances. However, Rtkn positive substances are also predicted as clastogenic as Rtkn activation is thought to be associated with double strand breaks (page 27 of the Validation Report – Attachment A). In addition, reviewers noted that specific aneugenic substances (indirect genotoxins) were not included in the validation study (i.e. tubulin poisons or kinase inhibitors) which was also acknowledged in the developers' response (Annex 4). Furthermore, PRP noted that oxidative stress is considered as mainly an indirect threshold effect in the validation analysis, but peroxides could also induce direct DNA damage and therefore the oxidative stress mechanism may also be relevant for direct genotoxicants.

30. The developers were asked to provide more information in writing about the selection of the reference chemicals for testing in the validation study overall and, in particular, in relation to distinguishing between clastogens and aneugens. The responses are included in Annex 4 (see responses to question 1 and 3). They are in line with the PRP's view that the assay and the protocol with the particular interpretation procedure used in the validation study does not allow for clear discrimination of clastogens and aneugens. However, some reviewers noted that an extended version of the assay, ToxTracker ACE (Brandsma et al., 2020) (considering cell cycle analysis and aneuploidy assessment) may provide more adequate information for the specific mechanisms underlying the different types of genotoxicants in the future.

31. PRP supports future development of the extended version of the ToxTracker but noted that it was not covered by the validation study under review.

32. One reviewer specifically noted that even if carcinogenesis and mutagenesis (as well as some other mechanisms) are obviously highly linked, ToxTracker cannot be considered as a method for "*in vitro carcinogenicity hazard identification*" or "*cancer hazard/risk assessment*" as stated in paragraphs 1 and 2 of the validation study. ToxTracker method should be strictly dedicated to genotoxicity and/or MoA of genotoxic substances. In addition, some reviewers noted that the information supporting the linkage/relationship between the mechanisms associated with genotoxicity (i.e. formation of bulky DNA adducts and subsequent inhibition of DNA replication; induction of DNA double-strand breaks; arresting of cells in mitosis; G1/S cell cycle checkpoint regulation; oxidative stress; and activation of the unfolded protein response) and the expression of the different reporter genes is not sufficiently detailed and includes very few references. Different reviewers emphasized the need for more detailed information on different relationships, e.g. the relationship between Bscl2 activation and formation of bulky DNA adducts and subsequent inhibition of DNA replication, or the relationship of Rtkn activation and arresting of cells in mitosis). The PRP agreed that more in-depth and better referenced description of the background of the development of the assay and selection of each of the six reporter genes would be very useful for the future development of the assay and potentially a test guideline. Inclusion of an overall signal transduction diagram (from sensors, transducers, and effectors, up to cell event(s)) is recommended.

33. One reviewer pointed out the potential of ToxTracker to provide a 'rich' pattern of results to help with MoA explanations if the expression patterns of the suite of all six reporters is analysed by Principal Component Analysis (PCA). Example of an exploratory PCA analysis based on the final call (positive/negative) for the different reporter genes for each of the 64 substances used on the validation study was presented at the teleconference (Annex 3). The review panel agreed that PCA analysis would be a good approach for identification of clusters of substances with similar responses and also detection of outliers and trends. More specific analysis using raw data would be even more informative and help better understand the false negative and false positive results for some chemicals. The PRM noted that the supporting document Boisvert et al., 2023 - doi: 10.1002/em.22525 reports a PCA analysis performed following the validation study. After the teleconference the study by Boisvert et al., 2023 was considered by some members of the panel with specific expertise in statistics and found it to be complementary to the exploratory PCA analysis provided by the reviewer (Annex 3). In particular, the PCA analysis by the developers is based upon BMD estimates of a doubling (2-fold increase expression) in the values for each of the six reporters. Some estimated values were used when BMDs could not be determined or were very large. These results were used to analyse the set of chemicals with differential responses of the six marker reporters to help differentiate chemicals largely associated with DNA damage and genotoxic stress from those largely associated with oxidative damage, demonstrating the value of the PCA analysis for extracting valuable information relevant to mechanistic aspects of toxicity and MoA.

The peer review panel agreed that the Validation Principle 2 has been met considering the primary goal of the validation study. Specifically the endpoint(s) measured with the ToxTracker (differential expression of the six selected genes) are relevant for evaluation of genotoxicity. However, the ToxTracker assay with the prediction model used in the validation study does not support distinguishing between clastogenic and aneugenic genotoxicants nor specific identification of carcinogens.

Technical Characterisation
(Validation principles 3 and 4)

Charge Question 3: *Do you consider that the protocol description in the Validation Study (including also section 15 in the Validation Report on Discussion and learnings) is sufficiently detailed, providing clear coverage of all the materials and procedures for data generation and interpretation?*

34. Written comments by reviewers reflect overall agreement that the protocol description and learnings from the validation study, are well-detailed and provide clear coverage of all the materials and procedures for data generation and interpretation. Reviewers find that careful thought has been put into the choice of positive controls for each of the cell lines. The clear instructions on the preparation of solutions, especially for poorly soluble substances, are particularly noteworthy. Reviewers emphasize the importance of inclusion of this instruction in all OECD TGs.

35. It was also noted, that at least one laboratory had some problems with earlier protocols and that the protocol needed some improvements over time which are detailed in Section 15 of the validation report. These improvements need to be incorporated for transferring the assay to 'routine' users in the future.

36. Reviewers suggested adding the following specific information/considerations for clarity of the protocol:

    1. Clearly specify the fold increase for significant response of each of the GFP reporter genes (in the validation study it is not clear that the same, two-fold increase is applicable to reporters other than Bscl2-GFP and Rtkn-GFP)
    2. Explain the basis for the choice of the 2-fold increase in GFP threshold for a significant response
    3. What is the rationale for selecting the highest testing concentration of 1 mg/mL, given that (other OECD TGs for testing genotoxicity in vitro) recommends 2 mg/mL as the highest test concentration when no precipitation or cytotoxicity occurs
    4. Use of solvent control (when different solvents are used) in addition to negative/nongenotoxic controls
    5. Specify cell seeding density for each cell line
    6. More comprehensive description of the flow cytometer parameters: calibration steps, number of events to be read. Also the description of adjustment of compensations following analysis of experimental samples, with screen copies of appropriate or inappropriate cytometer quadrants to help the operator.
    7. A clear definition of "independent repeat experiments" is needed
    8. A more detailed explanation of specific issues affecting quality and reproducibility of cell counting

37. The developers were asked to provide more information about points 1, 2, 3 and 4 above. The responses are included in Annex 4 (see responses to question 2, 4 and 5).

38. The review panel reviewed the responses at the teleconference and considers that the choice of testing concentration needs further explanation/justification. One reviewer pointed out that the ICH S2R1 (step 5) Maximum concentration specifies: "The maximum top concentration recommended is 1 mM or 0.5 mg/ml, whichever is lower, when not limited by solubility in solvent or culture medium or by cytotoxicity (note 8)". In addition the choice of 2 mg/ml or 10mM (if MW is known) within the OECD TG context has had a long history that started even

higher at 5mg/ml. It was agreed that the choice of testing dose may require additional discussion for the context of the development of the OECD TG.

39. The choice of a 2-fold increase in reporter genes expression as a cut-off for positive response was further discussed at the teleconference in the context of the detailed response by the developers as to how this cut-off was determined (Annex 4). PRP notes that while this may be an optimal cut-off level for an assay that aims to dichotomise the data set and successfully distinguish between genotoxic and non-genotoxic substances, some weakly positive substances may be 'lost' due to a lack of clear reproducibility when values fall close to this cut-off borderline. This could be better evaluated from the analysis of the raw data for reporter responses (see CQ8). In addition, testing of wider range of genotoxic substances would be valuable to show appropriate sensitivity of the test system for genotoxins with weak response (see also paragraph 50). If appropriate reference chemicals are identified, additional testing and optimisation of the prediction model would be crucial as most real world genotoxic chemicals are unlikely to react like positive control substances that normally elicit strong responses.

40. The description of use of solvent control(s) was also raised with a need for better description and identification within the protocol, particularly when test chemicals with different solvents are used.

41. One reviewer noted that high throughput application of ToxTracker may be limited given that mES cells are adherent cells requiring detachment before analysis,

42. The reviewers also suggested specific editorial changes for the protocol but also for sections of the validation report:

    - Point 14: It seems that in the last sentence the "survival level of 75%" should actually be 25%.
    - Point 13: "Induction of the fluorescent ToxTracker reporters as well as the number of viable cells is determined by flow cytometry after 24 h exposure of the cells to the test compounds." It is not clear what is meant by viable cells in this sentence, as this would suggest that the cells are stained to assess viability. Actually, the proliferation of the cells is used as cytotoxicity marker and cells are only counted. The word "viable" should be deleted.
    - Figure 3: The x-axis label for tunicamycin is in different units for GFP induction and cell survival. This should be harmonized.
    - Table 8; 6/7 positive results should be 85.7%?.
    - Table 11: The final prediction for sodium diclofenac should be "E".
    - Table footnotes for the letters (e.q. P, N, I, E) used in the tables could be helpful.
    - Adding a 2-fold increase threshold line and 75% cytotoxicity threshold line in the data figures (e.q. Figure 4, 5,6) would be helpful.
    - Table 20 –marked wrongly Table 19 on page 38 of the validation report
    - Annex 5: There is no key to the colours explaining the results.

Overall, the peer review panel agreed that the Validation Principle 3 has been met. However, several recommendations are provided for specific clarifications in the protocol description (above numbered list and also Conclusions Section below).

Charge Question 4: *Do you consider that the intra-, and inter-laboratory reproducibility of the ToxTracker test method results is adequately demonstrated considering also the effects of biological variability of the test method.*

43. The written review comments reflect agreement that the intra- (WLR), and inter- (BLR) laboratory reproducibility of the method were adequately demonstrated. It is noted that both WLR and BLR are high. In addition, standard deviations are relatively low, which shows that the system is working properly.

44. The WLR in six out of the seven participating laboratories varied between 80% and 96.7%. Overall, WLR (average of all reporter genes) varies from 97.5% for the lab with the best overall performance and 71.1% for the lab with the lowest reproducibility. At the teleconference one reviewer noted that analysis of the WLR for each individual reporter gene would be very useful to provide.

45. The significantly lower WLR for one laboratory despite appreciable (and clear) protocol descriptions, specialised training and the use of standard materials was noted by one reviewers as a flag for consideration how robust the method would be for 'general,' or for use in academic research laboratories. However the panel agreed that generally, for regulatory use and following a standard test guideline protocol it is likely that the assay would be performed by specialised, well trained commercial laboratories that will show proficiency and GLP compliance. The potential issue of assay transferability to academic laboratories was interesting from the point of view of future development of the validation framework and not related to the current validation study of the ToxTracker assay for the purposes of development of an OECD test guideline.

46. The BLR in the ToxTracker validation trial for the seven participating laboratories varied between 83% for the genotoxicity predictions (Bscl2 / Rtkn) and 71% for oxidative stress (Srxn1 / Blvrb) (Table on page 42 of the validation report). It was noted that direct comparisons between laboratories are difficult because different concentrations of the test substances were applied for testing in each laboratory based on the results of the dose finding pre-test for cytotoxicity.

47. It was also noted that only selected representative data is shown in the validation report. Namely, the raw data for GFP induction and cell survival are shown for three laboratories for only one substance and the pattern of the different reporters is very similar (Figure 4 in the validation report). However, graphical representation of the test results for all substances and all laboratories would be useful. One laboratory had particular challenges with accurate measurements of cell numbers which resulted in an overestimation of the cytotoxicity of compounds.

Overall the peer review panel agreed that the Validation Principle 4 has been met. However, it was noted that transferability may be challenging in the context of non-commercial academic research laboratories.

Fitness for Purpose
(Validation principles 5 and 6)

Charge Question 5: *Are the reference chemicals (applicability domain) used to evaluate the laboratory transferability and reproducibility of the ToxTracker test method representative of the types of substances for which the test method will be used?*

48. The panel considers that the reference chemicals used for the validation study cover a broad chemical space and are representative of the types of substances for which the test method would be used in regulatory context. The list of 64 compounds used in the validation study consists of organic and inorganic, aromatic and aliphatic molecules to cover a broad chemical space. Also, a number of compounds that require metabolic activation were selected. (Selection of the chemicals and the databases used are elaborated in detail by the developers in Annex 4, response to question 1).

49. However, as mentioned above, to clearly distinguish aneugens from clastogens, specific consideration of aneugenic compounds is needed. In addition, the specific MoA is not considered for the different 64 substances in the validation set. One reviewer specifically questioned how the relevance of the mechanistic and MoA info could be evaluated in the absence of good understanding of the MoA applicable to the reference chemicals (e.g. no chemical associated with the mechanism of unfolding protein is identified in the reference set).

50. One reviewer noted that, from a statistical practice point of view the total number of substances used in the study is small and, that the confidence intervals (CIs) associated with the performance of the test will be relatively wide when sample size of the two groups (genotoxic and non-genotoxic) is small (i.e. n=32). Therefore it should be kept in mind that the evaluation of the sensitivity and specificity of the assay are associated with some uncertainty. However, it was recognised that the number and choice of chemicals seems realistic to provide sufficient information and evidence for a validation study given the resources available for a study of this type. Related to this, it was noted that the chemicals used in the validation study are probably at the extremes of the spectrum of genotoxicity and, therefore, it may be expected that the ToxTracker assay will allow clear discrimination between positive and negative results, but it is not clear how effective this assay would be for genotoxic chemicals with more intermediate level of response (see also paragraph 39). Further analysis of the pattern of results for the suite of reporters and also testing of additional relevant substances is needed, if available.

The peer review panel agreed that the Validation Principle 5 has been met for distinguishing between genotoxic and non-genotoxic substances (the primary goal of the validation study). However, in order to evaluate and ensure good predictive performance for substances across the spectre of responses of the reporter genes and for specific MoA (e.g. aneugenic vs clastogenic), further analysis and testing of relevant substances is needed. In addition, the panel agreed that it would be useful to represent the CIs associated with the performance statistics.

Charge Question 6: *Considering the primary and secondary goal of the validation study (section 3 of the validation report), do you consider that the performance of the test method has been evaluated in relation to information from existing toxicity testing data for biologically relevant context (species, cell type, mechanism of toxicity etc)*

51. The review panel agrees that the performance of the test method is considered to have been evaluated in relation to information from relevant existing toxicity testing data for the biologically-relevant context (reference data are taken from test systems using bacteria to rodents).

52. In addition, the primary objective of the study is broadly met. Within the constraints of such a project, a high discrimination between genotoxic and non-genotoxic chemicals was achieved.

53. The secondary/subsidiary objective is partly met. The assay based on six reporter genes, which are relevant for aspects of genotoxicity, has high potential to provide useful evidence on modes of action. However, further research is needed in this area if the approach is to provide more than supplementary information. Further development is needed on the inclusion and analysis of more chemicals with well characterised MoA (for example: specific aneugenic substances, and substances with known MoAs relevant to the specific reporter genes such is Ddit3).

The peer review panel agreed that the Validation Principle 6 has been met with regard to the primary goal of the validation study i.e. to distinguish genotoxic from non-genotoxic substances. However, further work is needed to explore the full potential of the mechanistic information generated by the assay for better characterisation of relevant MOAs for genotoxicity and its extended use in safety assessment (also discussed under CQ 2 and 5).

Data integrity and transparency (Validation principles 7 and 8)

Charge Question 7: *Assess the goodness of the experimental/validation practices (for example: the coding of test chemicals for blind testing and analysis of test results, the testing for interference with the test system/medium components/measurement device/method as appropriate, or any others you considered relevant) under which the data have been generated for the validation study of the ToxTracker test method?*

54. Written comments reflect agreement that the data integrity and transparency appear satisfactory and this was confirmed at the teleconference discussion.

55. The procedures in place for collecting, analysis and interpreting were suitable for a validation trial. Chemical selection was done by an independent group, samples were not identified to the testers. Blinding was carried out during testing and not broken until after results had been released.

56. Data acceptance criteria for negative and positive controls were fixed. Interpretation criteria were fixed.

57. The VMT provided some technical support. Where the developer was involved in discussion over problems with the technical application of the protocol, no interference appears to have occurred in relation to discussions or release of details allowing chemical identification.

The peer review panel agreed that the Validation Principle 7 has been met.

Charge Question 8: *Do you consider that all the data supporting the assessment of the validity of the test method are easily available for expert review? These include: a detailed and readily available test method protocol to the public and independent laboratories; organised and easily accessible data to permit independent review(s); benchmarks by which an independent laboratory can itself assess its proper adherence to the protocol.*

58. Peer reviewers considered that the authors have made a good effort to present the data and information clearly. All final results (positive/negative calls) from the ToxTracker ring trial are collected in a large database and were available to the reviewers.

59. However, availability of raw data (GFP induction/cell survival) and figures for all substances from all laboratories was not available. Reviewers emphasized the importance of availability of the raw data for ease of any independent review of the results but also for better analysis of the MoA information generated by the ToxTracker assay.

60. Some specific recommendations were made to add clarity to the available information:

    1. Tables 10-12: It is not easily discernable from Tables 10-12 which 5 compounds did not have acceptable data sets to include in the summary statistic described in paragraph 36. The developers provided this info and criteria for data acceptance (Annex 4, see responses to question 6);
    2. Annex 5: There is no key to the colours explaining the results. Colours presumably reflect the call for each of the 6 GEP from each of the 7 laboratories in Phase I.

The peer review panel agreed that the Validation Principle 8 has been partially met as the raw data for reporter stimulation could not be examined. Public availability of raw data of the validation trial (GFP induction and cell survival) in addition to final assay result which were available to the panel (positive, negative, equivocal) is strongly suggested.

Conclusions

61. The peer review panel considered that the validation study of the ToxTracker assay was correctly designed to address the main goal of the study, i.e. to determine if ToxTracker is able to correctly predict the genotoxic properties of substances and to provide additional mechanistic information relevant for genotoxicity prediction. However, more in-depth and better referenced description of the background of the development of the assay and selection of the six reporter genes would be helpful for users. Inclusion of an overall signal transduction diagram (from sensors, transducers, and effectors, up to cell event(s)) is recommended.

62. Importantly, the PRP noted that in the introductory paragraphs and the title of the Validation Study Report claims are made regarding a potential application of ToxTracker for carcinogenicity hazard screening and assessment. Given that the objectives of the validation study analysis deal with genotoxicity (direct and indicted), both in terms of prediction and mechanistic/MoA information, the PRP considered that these authors' claims are not substantiated by the validation study.

63. The validation study was conducted appropriately in terms of data integrity and transparency, but limited availability of all raw data was noted, possibly due to the limitations of the publicly available data storage capacity.

64. The intra- (WLR), and inter- (BLR) laboratory reproducibility of ToxTracker results were high, except for one laboratory where for one reporter gene the WLR was 71.1%. However, the reason for the lower WLR in that laboratory was identified by the VMG and the protocol was updated to address the issue. Overall, the standard deviations from the available row data are low, which indicates that the test system is working properly.

65. The protocol described in the validation study report was considered quite clear but some specific suggestions for further improvement of clarity are provided:

    1. Clearly specify the fold increase for significant response of each of the GFP reporter genes (not just Bscl2-GFP and Rtkn-GFP)
    2. Explain the basis for the choice of the 2-fold increase in GFP threshold for a significant response
    3. Justify the rationale for selecting the highest testing concentration of 1 mg/mL, (compared to other OECD TGs for testing genotoxicity in vitro which recommend 2 mg/mL)
    4. Clarify use of solvent control (when different solvents are used)
    5. Specify cell seeding density for each cell line
    6. More comprehensive description of the flow cytometer parameters: calibration steps, number of events to be read. Also the description of adjustment of compensations following analysis of experimental samples, with screen copies of appropriate or inappropriate cytometer quadrants to help the operator.
    7. A clear definition of "independent repeat experiments" is needed
    8. A more detailed explanation of specific issues affecting quality and reproducibility of cell counting

66. Regarding technical availability and transferability of the assay protocol, the PRP noted that long-term commercial availability of the cell lines necessary for the assay needs to be considered for the development of the TG. In addition, for future validation studies in general, and for any potential updates of the validation of ToxTracker assay in particular, inclusion of academic, in addition to commercial, laboratories may need to be considered. This would

ensure transferability of assays relevant for regulatory context, to any laboratory based only on the protocol and independent of the training or support by the assay developer.

67. Regarding the choice of reference chemicals for the validation study, the number and choice is realistic but probably more representative of the extremes of the spectrum of genotoxicity and, therefore, appropriate for discrimination between positive and negative results and addressing the primary goal of the validation study. However, additional testing/data with different relevant substances (if available) would be valuable to allow analysis of the intermediate responses for the suite of all six reporters. In addition, it would be useful to represent the performance statistics of the assay together with the Confidence Intervals (CIs).

68. Overall, while the assay provides valuable mechanistic information related to genotoxic properties of the tested chemicals, further development is needed, for example to identify reporter gene(s) that could discriminate between clastogens and aneugens. Additional analysis (e.g. PCA) with the data in the validation study is recommended. The analysis carried out by the test developers (Boisvert et al., (2023) was noted but not reviewed in depth. It was considered complementary to that provided by one of the reviewers (Annex 3) performed using the final call results (positive/negative) from the validation study.

Acknowledgements:

On behalf of the OECD Secretariat the PRM thanks the Peer Review Panel for their review and valuable comments.

References

OECD (2005). Series on Testing and Assessment No. 34: Guidance Document on the validation and international acceptance of new or updated test methods for hazard assessment. OECD Environment, Health and Safety Publications. Paris. Available at: http://www.oecd.org/env/ehs/testing/series-testing-assessment-publications-number.htm.

Hendriks, G., Atallah, M., Raamsman, M., Morolli, B., van der Putten, H., Jaadar, H., Tijdens, I., Esveldt-van Lange, R., Mullenders, L., van de Water, B., et al. (2011). Sensitive DsRed fluorescence-based reporter cell systems for genotoxicity and oxidative stress assessment. Mutat. Res. Fundam. Mol. Mech. Mutagen. 709–710, 49–59.

Hendriks, G., Atallah, M., Morolli, B., Calleja, F., Ras-Verloop, N., Huijskens, I., Raamsman, M., van de Water, B., and Vrieling, H. (2012). The ToxTracker assay: Novel GFP reporter systems that provide mechanistic insight into the genotoxic properties of chemicals. Toxicol. Sci. 125, 285–298.

Hendriks, G., Derr, R. S., Misovic, B., Morolli, B., Calleja, F. M. G. R., and Vrieling, H. (2016). The extended ToxTracker assay discriminates between induction of DNA damage, oxidative stress, and protein misfolding. Toxicol. Sci. 150, 190–203.

Inger Brandsma, Nynke Moelijker, Remco Derr, Giel Hendriks, Aneugen Versus Clastogen Evaluation and Oxidative Stress-Related Mode-of-Action Assessment of Genotoxic Compounds Using the ToxTracker Reporter Assay, *Toxicological Sciences*, Volume 177, Issue 1, September 2020, Pages 202–213, https://doi.org/10.1093/toxsci/kfaa103

Wills JW, Halkes-Wellstead E, Summers HD, Rees P, Johnson GE. Empirical comparison of genotoxic potency estimations: the in vitro DNA-damage ToxTracker endpoints versus the in vivo micronucleus assay. Mutagenesis. 2021 Aug 27;36(4):311-320. doi: 10.1093/mutage/geab020. PMID: 34111295; PMCID: PMC8391785.

Boisvert, Lorrie ; Remco Derr; Torben Osterlund ; Giel Hendriks ; Inger Brandsma (2023), Quantitative interpretation of ToxTracker dose–response data for potency comparisons and mode-of-action determination, Environ Mol Mutagen.; 64:132–143, doi: 10.1002/em.22525

Menz J, Götz ME, Gündel U, Gürtler R, Herrmann K, Hessel-Pras S, Kneuer C, Kolrep F, Nitzsche D, Pabel U, Sachse B, Schmeisser S, Schumacher DM, Schwerdtle T, Tralau T, Zellmer S, Schäfer B. Genotoxicity assessment: opportunities, challenges and perspectives for quantitative evaluations of dose-response data. Arch Toxicol. 2023 Sep;97(9):2303-2328. doi: 10.1007/s00204-023-03553-w. Epub 2023 Jul 5. PMID: 37402810; PMCID: PMC10404208.

Annex 1 – Peer review Panel

Expert Name	Affiliation	Country
Fabrice Nesslany	ERBC Group, Baugy	France
Masataka Tsuda	Division of Genetics and Mutagenesis National Institute of Health Sciences	Japan
David Lovell	Emeritus Reader St George's Medical School, University of London	United Kingdom
Hong Xie	U.S. Food and Drug Administration Center for Food Safety and Applied Nutrition Cosmetics Regulatory Science Branch	USA
Jeannette Koenig	German Federal Institute for Risk Assessment, Dept. of Pesticide Safety	Germany

Peer Review Manager	Julija Filipovska, Independent Consultant

Annex 2 – Charge Questions

Charge Questions for the Peer Review of the Validation Report on the ToxTracker assay: a stem cell-based reporter assay for mechanistic carcinogenicity hazard screening

Questions based on the validation principles and criteria in OECD GD34 (link)

Scientific basis and Relevance to an endpoint of interest
CQ1. Do you consider that the rationale for the ToxTracker test method is clearly elaborated in the validation study and the other supporting documents in terms of its scientific basis, regulatory purpose and need? Consider the potential added value of the mechanistic information for genotoxicity generated with this assay.
[Your review comments]
CQ2. Is the relationship between the endpoint(s) measured with the ToxTracker (differential expression of the six selected genes) and the (biological) phenomenon of interest (genotoxicity) described adequately and using relevant scientific information/references?
[Your review comments]
Technical Characterisation
Q3. Do you consider that the protocol description in the Validation Study (including also section 15 in the Validation Report on Discussion and learnings) is sufficiently detailed, providing clear coverage of all the materials and procedures for data generation and interpretation? *e.g.*, a description of the materials needed, such as specific cell types, constructs, a description of what is measured and how it is measured, a description of how data will be analysed, controls and decision criteria for evaluation of data and what are the criteria for acceptable test performance.
[Your review comments]
Q4. Do you consider that the intra-, and inter-laboratory reproducibility of the ToxTracker test method results is adequately demonstrated considering also the effects of biological variability of the test method. Please comment on the reproducibility of the test method in relation the measurement(s) for the differential expression of each gene and also overall for the ToxTracker method.
[Your review comments]
Fitness for Purpose
Q5. Are the reference chemicals (applicability domain) used to evaluate the laboratory transferability and reproducibility of the ToxTracker test method representative of the types of substances for which the test method will be used?
[Your review comments]
Q6. Considering the primary and secondary goal of the validation study (section 3 of the validation report), do you consider that the performance of the test method has been evaluated in relation to information from existing toxicity testing data for biologically relevant context (species, cell type, mechanism of toxicity etc)
[Your review comments]
Data integrity and transparency
Q7. Assess the goodness of the experimental/validation practices (for example: the coding of test chemicals for blind testing and analysis of test results, the testing for interference with the test system/medium components/measurement device/method as appropriate, or

any others you considered relevant) under which the data have been generated for the validation study of the ToxTracker test method?
[Your review comments]
Q8. Do you consider that all the data supporting the assessment of the validity of the test method are easily available for expert review? These include: a detailed and readily available test method protocol to the public and independent laboratories; organised and easily accessible data to permit independent review(s); benchmarks by which an independent laboratory can itself assess its proper adherence to the protocol.
[Your review comments]

Overall conclusion
Based on your assessment above for: • The completeness of the protocol for the ToxTracker test method, the clarity of the elements composing the method and the clarity of the data interpretation procedures; • The biological relevance of the assay and selected genes to the biological effect of interest; • The capacity to discriminate genotoxic chemicals from non-genotoxic chemical within the scope defined in the Validation Study under review; • The level of reproducibility and transferability of the results of the ToxTracker assay; please provide your assessment of the level of readiness of the ToxTracker assay for development of a Test Guideline that would address genotoxicity and could also add valuable MoA information for the assessment of different types of genotoxicants.
[Your review comments]

Annex 3 – Compilation Responses

Peer Review of the Validation Report on the ToxTracker Assay

August-September 2023

Scientific basis and Relevance to an endpoint of interest	
CQ1. Do you consider that the rationale for the ToxTracker test method is clearly elaborated in the validation study and the other supporting documents in terms of its scientific basis, regulatory purpose and need? Consider the potential added value of the mechanistic information for genotoxicity generated with this assay.	**PR1** The rationale for the ToxTracker test method is clearly stated. However, additional description of any advantage of ToxTracker to other similar methods, such as TGx-DDI and DNA damage gene expression assay, could be helpful.
	PR2 Regarding the scientific basis, the rationale for the ToxTracker test method is clearly elaborated (it combines 6 fluorescent reporter genes that are specifically activated by different cellular signalling responses associated with (direct and indirect) genotoxicity. Given that ToxTracker can provide an overview of the MoA of genotoxic compounds, the main question that arises from a regulatory point of view is whether mechanistic data are required as soon as the basic battery is produced or only when a positive response is observed? The implementation of methods aimed at defining a genotoxic mode of action already exists, e.g., FISH for aneuploidy, methods targeting ROS, oxidation of DNA bases for oxidation induction (e.g., 8-OH-dG, use of FPG/hOGG1, etc.). The induction of genes targeting oxidation and aneuploidy in the Toxtracker would in no way obviate the need to carry out these more specific methods (for instance, the reference method for defining aneugenic activity is the FISH). Given that the applicants position the regulatory application of Toxtracker to reveal genotoxic modes of action, a comparison with these reference methods would be useful. The Applicants state that the application of BMD to human risk assessment and the calculation of health-based guidance values are already used for regulatory genotoxicity testing (25-27). These publications mainly refer to the potential use of **in vivo** genotoxicity results in the construction of health values, not in vitro data. The development of risk assessment methods using in vitro data has emerged, particularly in the field of cosmetics, where the use of animals is prohibited, but these methods refer to toxicity data, not genotoxicity data (which remains a no go in cosmetics). While this remains a prospect, it seems all the more compromised with methods investigating primary DNA damage.

	PR3 Yes, I consider that the rationale for the ToxTracker test method is clearly elaborated in the validation study.
	PR4 Yes. The case for using the ToxTracker test method is made well. The report is presented well, particularly the potential for its role in meeting a niche in the regulations for genotoxicology. ToxTracker is proposed as a novel approach for in vitro carcinogenicity hazard identification and is, consequently, a valuable expansion of the in vitro genotoxicity testing battery for regulatory applications. The objective of providing relevant mechanistic information, which can add value, as opposed to being just another test, is well supported. However, it should not be considered as an alternative to the already existing in vitro genotoxicity assays, but as an additional tool that provides mode-of-action (MOA) information that would be useful in the move towards the new paradigm of a mechanistically based risk assessment.
	PR5 Yes, the regulatory purpose and need is clearly described in the validation report. It is clear that the intention of the test is to identify genotoxic substances and to provide additional information on the genotoxic and/or non-genotoxic MoA of substances. ToxTracker is not intended to replace the established in vitro and in vivo genotoxicity tests, but rather to clarify the mechanism of genotoxicity (mutagenicity and clastogenicity as well as indirect genotoxicity due to oxidative stress as example) and/or non-genotoxicity (oxidative stress, protein stress and p53 activation). The additional information generated by ToxTracker is particularly valuable for developing AOPs and to support regulatory conclusions by integrating it into a weight-of-evidence approach. In particular, when genotoxicity is indirectly induced and thus a threshold for genotoxic damage exists, reference values may be derived. In addition, mechanistic information on genotoxicity can be used to decide on appropriate follow-up genotoxicity testing strategies. On page 8, it is described that ToxTracker can be used to distinguish between clastogenicity and aneugenicity is possible. According to the table "Interpretation of ToxTracker data" on page 73, aneugenic substances are Rtkn positive and Bscl2 negative. However, Rtkn positive substances are also predicted as clastogenic. It is therefore not clear how ToxTracker can distinguish between aneugenicity and clastogenicity. This issue needs further clarification and a more detailed description with substance examples. To demonstrate that ToxTracker also identifies aneugenic substances, experimental evidence should be provided. Thus, it would be necessary to include known aneugens in the validation experiment and to present and discuss the results in more detail. However, I could not identify any aneugen in the test group. As far as I know, another type of ToxTracker, namely ToxTracker ACE includes, DNA staining and is able to identify aneugens. I recommend to cite the following reference to justify the regulatory purpose:

	Menz J. et al. Genotoxicity assessment: opportunities, challenges and perspectives for quantitative evaluations of dose-response data. Arch Toxicol. 2023 Sep;97(9):2303-2328. doi: 10.1007/s00204-023-03553-w. Epub 2023 Jul 5. PMID: 37402810; PMCID: PMC10404208.
CQ2. Is the relationship between the endpoint(s) measured with the ToxTracker (differential expression of the six selected genes) and the (biological) phenomenon of interest (genotoxicity) described adequately and using relevant scientific information/references?	**PR1** Based on the description in section 4.7, activation of the Rtkn-GFP genotoxicity reporter is associated with both clastogens (induction of DNA double-strand breaks) and aneugens. It is not clearly stated if this method is able to differentiate between clastogens and aneugens which is very important in safety evaluation and risk assessment.
	PR2 There is a reliable relationship between the differential expression of the six selected genes measured with the ToxTracker and (direct and indirect) genotoxicity: - Bscl2 is activated upon the formation of **bulky DNA adducts** and subsequent inhibition of DNA replication. - Activation of Rtkn is associated with induction of **DNA double-strand breaks**. - Activation of the Rtkn and aresting of cells in mitosis is often observed for aneugens - Btg2, a component of the P53-dependent DNA damage response, is involved in regulation of the G1/S cell cycle checkpoint but is also induced by various other cellular stressors. - Srxn1 and Blvrb are activated by the 2 major cellular antioxidant pathways. - Activation of the unfolded protein response is detected by Ddit3 ⇨ Toxtracker must be strictly designated as a method able investigating DNA primary damage and predict a genotoxic mode of action ⇨ Not clear how presumed clastogenicity and aneuploidy are differentiated (Rtkn-GFP targets both DNA Rtkn-GFP and "aneuploidy") ⇨ Oxidation is a thresholded but not always an indirect MoA (for instance, peroxides could be direct genotoxins) ⇨ It would be useful to position the 6 reporter genes within an overall signal transduction diagram (from sensors, transducers, and effectors, up to cell event(s)) Other general comments: Even if carcinogenesis and mutagenesis (as well as some other mechanisms) are obviously highly linked, it is not acceptable to label the Toxtracker method as an "*in vitro carcinogenicity hazard identification*" or to state that it is used for "*cancer hazard/risk assessment*". Toxtracker method should be strictly dedicated to genotoxicity and/or MoA of genotoxic substances.

	By the way, the ToxTracker trial was solely focused on genotoxicity prediction and carcinogenicity was not considered as a criterion for compound selection! Under no circumstances can this type of methods reveal a carcinogenic activity.
	PR3 The relationship between the endpoints and genotoxicity is well explained. However, there is little information about the mechanism of how the GFP gene is expressed upon the formation of bulky adducts, subsequent inhibition of DNA replication, and the induction of DNA double-strand breaks in Bscl2 and Rtkn cells.
	PR4 From my perspective, the mechanisms underlying the relationship between the six biomarker endpoints and potential genotoxicity are described well. The potential relationships between the results from the six genes and use in MOA approaches is clearly outlined. It appears that the mouse stem cell-based reporter assay can identify genotoxic compounds with appreciable accuracy while the use of the fluorescent reporter genes in the assay may provide insight into the MoA of genotoxic and non-genotoxic substances.
	PR5 Yes, the relationship between the different biomarker proteins and their role in DNA damage, oxidative stress, p53, and unfolded protein response is adequately described. However, a brief description of the basis on which the various reporters were selected would be appreciated. The scientific basis for the assay method is briefly described on page 5, but the reference given in this context (Reference No. 9) does not explain the entire background.
Technical Characterisation	
Q3. Do you consider that the protocol description in the Validation Study (including also section 15 in the Validation Report on Discussion and learnings) is sufficiently detailed, providing clear coverage of all the materials and procedures for	**PR1** Some clarification is needed. In section 7.16, it is not clear if a 2-fold increase in GFP reporter genes other than Bscl2-GFP and Rtkn-GFP can be used to classify genotoxic compounds other than indicate MoA. The interpretation of the other four report genes should be more specific. OECD recommends 2 mg/mL as the highest test concentration in in vitro genotoxicity assays when no precipitation or cytotoxicity. Any explanation why in this method the highest concentration is only 1 mg/mL? Besides a negative control, a solvent control should be included. Any data show that the same cell seeding density can be used for all six cell lines?

data generation and interpretation? *e.g.*, a description of the materials needed, such as specific cell types, constructs, a description of what is measured and how it is measured, a description of how data will be analysed, controls and decision criteria for evaluation of data and what are the criteria for acceptable test performance.	**PR2** The experimental phases of ToxTracker protocol are well described is Annex 1 of the validation report. Regarding analysis, details are missing for setting flow cytometer parameters (up to date, only important parameters are mentioned: - separation of intact cells from broken cells and debris - appropriate fluorescence levels of the untreated control cells - detection of fluorescence induction in treated reporter cells It would be needed to have details on the calibration steps, number of events to be read... and for adjustment of compensations following analysis of experimental samples, with screen copies of appropriate or inappropriate cytometer quadrants to help the operator. Controls and decision criteria for evaluation of data are described and argued but it remains unclear why the 2-fold increase in GFP threshold is the threshold for a significant response (empiric and/or statistics and/or AI approach, e.g., Deep Learning…).
	PR3 Yes, I think the protocol description in the Validation Study, including section 15 in the Validation Report on Discussion and learnings, is well-detailed and provides clear coverage of all the materials and procedures for data generation and interpretation.
	PR4 The full ToxTracker protocol (Annex 1 dated October 2018) provides details of the materials, equipment, preparation of cell culture plates, preparation of culture plates, seeding and exposure of reported cell lines and analysis of reporter induction. Annexes include checklists for the assay. The detail of the technical characteristics is clearly and exhaustively described. The complete methodology is, in effect, broken down into a series of specific actions. There is a considerable degree of what could be called a 'pipeline' to be followed to achieve a standardised method which is relatively easily transferable. The pipeline of procedures from sample derivation to data collection, analysis and interpretation is standardized with checklists related to the steps. This should, in general, allows consistent and objective interpretation of results. Methods for integrating and interpretating the results are in place. This approach aids the production of high-quality data for the validation. It was noted, though, that at least one laboratory had some problems and the protocol needed to be modified at various times. Transferability to 'routine' users may be an issue in the future.

	The high quality of data produced may not be maintainable in the future if the assay is in routine use by a large number of laboratories. Careful thought went into the choice of positive controls for each of the cell lines. Viability was considered. Only GFP results at compound concentrations that showed <75% cytotoxicity were considered suitable for use to analyse for results. Data from measurements with >75% cytotoxicity were discarded as they were considered uninterpretable in a meaningful way. Clear cut-off criteria were developed. A 2-fold GFP reporter induction was the threshold for a positive ToxTracker result. The 2-fold increase was the equivalent of a 3 standard deviation (SD) increase over fluorescence levels in solvent control cell cultures. Compounds were classified as genotoxic in ToxTracker when either one or both Bscl2-GFP and Rtkn-GFP reporters were induced above the 2-fold increase in GFP threshold. Test results for every reporter gene from three independent repeat experiments carried out by a laboratory are weighted according to a prediction model. For every compound, which was tested in each of three independent laboratories, the classifications were compared, and an overall classification was made based on a weighted calculation. In some cases. the Validation Management Team (VMT) used expert judgement to accept or reject some controls and designated this as "acceptable with restrictions".
	PR5 Yes, the protocol description is very detailed and clear. The clear instructions on the preparation of solutions, especially for poorly soluble substances, are particularly noteworthy. This instruction should be included in all OECD TGs on genotoxicity testing, as it is a very important factor. Some further comments: - There is no explanation why the maximum dose is set at 1 mg/mL. In other in vitro genotoxicity tests the maximum dose is 2 mg/mL. This should be consistent if there is no justification for setting the maximum concentration at 1 mg/mL in this particular test. - Strategy for setting flow cytometry gates on page 74: recommendation to exclude doublets would be reasonable. Same applies on page 68, 11. Analysis of ToxTracker reporter induction (day 5), please include: - doublet discrimination to count only single cells - Calibration of the flow cytometer prior to measurement should also be recommended. - A clear definition of "independent repeat experiments" is needed, as researchers may have different understandings. - Page 11, point 14: It seems that in the last sentence the "survival level of 75%" should actually be 25%. - Pages 10-11: "Induction of the fluorescent ToxTracker reporters as well as the number of viable cells is determined by flow cytometry after 24h exposure of the cells to the test compounds." It is not clear what is meant by viable cells in this sentence, as this

	would suggest that the cells are stained to assess viability. Actually, the proliferation of the cells is used as cytotoxicity marker and cells are only counted. The word "viable" should be deleted. - Figure 3, page 12: The x-axis label for tunicamycin is in different units for GFP induction and cell survival. This should be harmonized.
Q4. Do you consider that the intra-, and inter-laboratory reproducibility of the ToxTracker test method results is adequately demonstrated considering also the effects of biological variability of the test method.	**PR1** The intra-, and inter-laboratory reproducibility is adequately demonstrated.
	PR2 During the 3rd phase of validation, regarding the sensibility: 25 out of the 32 genotoxic compounds (78%) were found positive in the 3 labs. On the remaining genotoxic substances: - For three compounds (9%), two labs reported a positive classification, but one lab classified the compounds as nongenotoxic. Differences could be due to a difference in concentration selection or related to inadequate metabolization by S9 without certainty. - Four (12%) of the expected genotoxic compounds were classified as non-genotoxic by all validation laboratories: acrylonitrile, benzene, cadmium chloride and NDMA. If it is right that benzene needs a specific metabolization to reveal its mutagenicity (CYP2E1 not expressed in current ametabolic activation systems) and that acrylonitrile is an epigenetic carcinogen, cadmium chloride is found positive in many in vitro test systems and NDMA, a potent mutagen, induces a clear mutagenic activity in the Ames test even with rat S9 (hamster S9 allows optimizing the formation of the ultimate mutagenic metabolite). Regarding the specificity, from the 32 expected non-genotoxic compounds included in the validation trial, none were classified overall as genotoxic, although for a number of compounds, a positive genotoxicity result was reporter by one laboratory (not always the same). For these unexpected results, the activation of the Rtkn-GFP reporter which indicates the formation of DNA strand breaks was observed but not the Bscl2-GFP reporter. ⇨ A specific analysis of this kind of results and the refinement of the interpretation criteria are required to confirm or slightly modify them.
	PR3 Yes. The BLR in the ToxTracker validation trial across the seven validation laboratories ranged from 83% for genotoxicity predictions to 71% for oxidative stress. The overall consistency in classifying

	genotoxic compounds was 91%. These results mean that Toxtracker is a relatively reproducible experiment.
	PR4 All the laboratories were able to perform the ToxTracker assay according to the assay protocol. One laboratory reported issues with solubility of one of the compounds. A laboratory had some challenges with accurate measurements of cell numbers which resulted in an overestimation of the cytotoxicity of compounds. No other significant issues were observed. Reproducibility was, in general, high. The results of the intra- (WLR), and inter-laboratory reproducibility (BLR) of the method results were adequately demonstrated. Some degree of biological variability was found, occasionally complicating interpretation. There was some degree of intra-laboratory variability (restricted to a small number of laboratories). This arose despite appreciable (and clear) descriptions, specialised training and the use of proprietary material. It raises the important question of how robust the method would be for 'general use'. In the first phase (training) a set of 8 compounds was tested independently by seven laboratories. This was followed by followed by a second phase where 6 chemicals were selected for proficiency testing by the 7 laboratories (results in Figure 4). (There seem to be a minor transcription error in Table 8; 6/7 positive results should be 85.7%?). In the third phase (validation or ring trial). 64 chemicals were selected for study. To prevent any selection bias or conflicts of interest, Toxys, as developer of the ToxTracker assay, was not involved in compound selection. Seven industrial/commercial laboratories took part in the validation process. No academic laboratories participated. Each compound was tested independently in three laboratories. None of the participating laboratories were aware of the compound selection. Compounds remained coded during the data analysis. The WLR and VLR were described as follows: : "Overall, the WLR in 6 of the 7 participating laboratories varied between 80% and 96.7%." "The overall WLR (average of all reporters) varies from 97.5% for the lab with the best overall performance and 71.1% for the lab with the lowest reproducibility." "The BLR in the ToxTracker validation trial for the seven validation laboratories varied between 83% for the genotoxicity predictions and 71% for oxidative stress (Table 20). The lowest laboratory had some

	challenges with accurate measurements of cell numbers which resulted in an overestimation of the cytotoxicity of compounds." There were some striking clear false negative results. Four of the expected genotoxic compounds were classified as non-genotoxic in ToxTracker by all the laboratories in the validation study. Acrylonitrile, Benzene and NDMA were negative and part of a group of 19 chemicals which had no positive results for any combination of biomarkers and +/- S9 mix. A fourth chemical, cadmium chloride, was negative for the two genotoxicity biomarkers but positive for nearly all the other biomarkers. 19 genotoxic chemicals were positive for both genotoxic biomarkers Bslc2 and Rtkn. A further 9 genotoxic chemicals were discordant for the genotoxic biomarkers: positive for Bslc2 but negative for Rtkn. These chemicals were also positive for some or all the other biomarkers. No non-genotoxic chemical was positive for either of the genotoxic biomarkers (Bslc2 and Rtkn). 16 were negative for all 6 biomarkers while the other 16 were positive for some or all the other four biomarkers. These patterns of results can be visualized using a Principal Component Analysis (PCA) on the matrix of test results. The range of results illustrates the potential of the method for provided a 'rich' pattern of results to help with MOA explanations.
	PR5 Yes, the intra- and interlaboratory reproducibility is predominantly demonstrated. It is clear from the validation report that the choice of test substance concentration is crirtical and in this context the accurate measurement of cell count is important to measure viability/cytotoxicity. However, in laboratory 4, there were some unspecified problems with the cell counting that eventually led to a relatively high number of inconclusive results. A more detailed explanation of what exactly happened would be useful, as such problems should be avoided by appropriate instructions in the TG. According to the figures in the validation report, the intra-laboratory reproducibility of the different reporter genes is good. The standard deviations are relatively low, which shows that the system is working properly. Direct comparisons between laboratories are difficult because different concentrations of the test substances were used. For one compound (compound 1), a comparison of GFP induction is shown for three laboratories and the pattern of the different reporters is very similar. However, I miss the graphical representation of the test results for all substances and laboratories. As I understand, the validation project was more focused on the correct identification of genotoxic and non-genotoxic substances and not the identification of the MoA by different laboratories. Comment on table 11: The final prediction for sodium diclofenac should be "E".

Fitness for Purpose	
Q5. Are the reference chemicals (applicability domain) used to evaluate the laboratory transferability and reproducibility of the ToxTracker test method representative of the types of substances for which the test method will be used?	**PR1** The reference chemicals are considered representative. But if activation of the Rtkn-GFP reporter is also observed for aneugens, some aneugenic compounds could be included in the test.
	PR2 During the interlaboratory validation trial, the compound list (64 compounds) consists of organic and inorganic, aromatic and aliphatic molecules to cover a broad chemical space. Also, a number of compounds were selected that require metabolic activation in the liver. Each reference compound was tested independently in three laboratories. <u>Cytotoxicity</u>: Differences in dose selection were found for some reference compounds (e.g., for cisplatin, the maximum concentration tested ranged from 0.24 µg/mL (Lab 4) to 62 µg/mL (Lab 7)). Authors claim that it could often be explained by either the approaches used for definition of cytotoxicity or by different criteria for solubility of compounds. It meant that both cytotoxicity and criteria for solubility should be clearly and more precisely defined. Interestingly, several modifications were made to the ToxTracker protocol for the third phase of the ToxTracker ring trial. <u>Genotoxicity</u>: During the second phase of the ring trial project, the prediction of genotoxicity by the seven labs was 100% in absence of S9-mix and 97,9% in presence of S9-mix showing a very good between-lab reproducibility for almost all of the reporter genes. For Cisplatin, the main difference was noticed on Srxn1 and Blvrb (see Table 4c, page 19). ⇨ It would be interesting to perform a specific analysis on both these genes.
	PR3 Yes, the reference chemicals used to evaluate the laboratory transferability and reproducibility of the ToxTracker test method are representative of the types of substances for which the test method will be used. They encompass a range of chemical classes and toxicological mechanisms that are relevant to the intended applications of the method.
	PR4 The chemicals chosen covered a broad spectrum of chemical classes with representatives from four groups: I) genotoxic carcinogens (22),

	II) genotoxic non-carcinogens (10), III) non-genotoxic carcinogens (10) and IV) non-genotoxic non-carcinogens (22). The chemicals used are probably at the extremes of the spectrum of genotoxicity and might, therefore, be expected to give more accurate predictions with a clear discrimination between positive and negative results. It is not obvious how effective the assay would be for chemicals with more intermediate level of genotoxicity. It was argued that the use of 64 chemicals would give sufficient statistical power for the study. However, it is important to appreciate that, in practice. this 'n' is small and that the confidence intervals (CIs) associated with the performance of the test will be wide when the genotoxic and non-genotoxic groups had relatively small sample sizes (i.e., n=32). Estimate of 95% CIs for n=32 (Wilson method) Sensitivity +ve n LCL UCL 62.5% 20 32 0.45 0.77 75% 24 32 0.58 0.87 81.25% 26 32 0.66 0.91 87.5% 28 32 0.72 0.95 93.75 30 32 0.81 0.98 However, the number and choice of chemicals seems realistic to provide sufficient information and evidence for a validation study given the resources available for a study of this type.
	PR5 A relatively broad chemical space with a balanced data set (32 genotoxic and 32 non-genotoxic substances) was included in the validation. However, since the aim of the ToxTracker test is to identify not only genotoxic substances but also different genotoxic modes of action, it would be useful to take this into account when selecting the test substances. As mentioned above, it would be advisable to expand the data set with aneugenic substances. Also, the specificity for identifying clearly mutagenic and clastogenic substances should be demonstrated, which was only partially done in the current validation report.
Q6. Considering the primary and secondary goal of the validation study (section 3 of the validation report), do you consider that the performance of the test method has been evaluated in relation to information from existing toxicity testing data for biologically relevant context (species, cell type,	**PR1** The data were analyzed for their genotoxicity predictions and compared with the standard in vitro and in vivo genotoxicity assay data. The specificity of the report genes for MOA is not well explained.
	PR2 The performance of the test method has been investigated. The comparison between ToxTracker and the standard in vitro and in vivo genotoxicity assays for compounds with an oxidative stress or

mechanism of toxicity etc)	protein reactive MoA is not sufficiently thorough for published results of MN and CA studies. For instance, relevant experimental conditions (e.g., type of cells used, duration of treatment…) should be known because they may also have a direct influence on the nature of the results (sometimes leading to known misleading results). Otherwise, Cadmium does not "only" induce oxidative stress. Following the literature, it also leads to the destruction of the spindle after chromosomal super contraction, thus, it induces aneuploidy as notices in many publications. Even if the induction of aneuploidy by cadmium is not systematic, depending on the test system used, this would warrant further discussion (in any case, do not limit the genotoxicity of Cd solely to oxidative stress). <u>Side comment</u> Authors wrote: "*For a number of compounds, only activation of the Rtkn, but not the Bscl2 genotoxicity reporter was observed. This is often observed for compounds causing indirect genotoxic effects, including aneugens or oxidative stress-inducing compounds*". Aneuploidy is not an indirect genotoxic effect but a direct effect on component(s)*, that allow preserving the integrity of the ploidy. It corresponds to non-DNA binding genotoxicity. * Examples of mechanisms leading to aneuploidy: *Misalignment of spindle orientation and contraction and more specifically of microtubule synthesis and assembly during spindle division, unequal chromosome alignment, abnormal kinetochore development and/or function, centromeric DNA assembly, disruption of sister chromatid separation, synthesis, division and function of centrioles and polar bodies of centromeric organizing regions...*
	PR3 Yes, I believe that the performance of the test method has been evaluated in relation to information from existing toxicity testing data for biologically relevant context.
	PR4 The protocol says: "the primary goal of this validation project … to determine if ToxTracker is able to correctly predict the genotoxic properties of compounds." The "secondary goal … was to investigate the MoA information that is provided by ToxTracker and the relevance for the genotoxicity prediction of the tested compounds." The primary objective is broadly met. Within the constraints of such a project a high discrimination between positive and negative genotoxic chemicals was achieved. The secondary/subsidiary model is partly met showing the potential of the 6 markers to provide useful evidence on modes of action which may be useful in the assessment of this goal; but, it also showed that further research is needed in this area if the approach is to provide more than supplementary information.

	PR5 The validation clearly shows that the **ToxTracker** assay can reliably distinguish between genotoxic and non-genotoxic substances. With regard to the unravelling of genotoxic MoA, further information is required, i.e. the correct identification of the genotoxic/non-genotoxic MoA by different laboratories and the inclusion of aneugenic substances.
Data integrity and transparency	
Q7. Assess the goodness of the experimental/validation practices (for example: the coding of test chemicals for blind testing and analysis of test results, the testing for interference with the test system/medium components/measurement device/method as appropriate, or any others you considered relevant) under which the data have been generated for the validation study of the ToxTracker test method?	**PR1** The protocol seems well designed. How much difference were the setting parameters of flow cytometer from each lab?
	PR2 All compounds were coded (the laboratories received blinded compounds). Due to the possibility of autofluorescence by the test compound (that may interfere with the measurement of GFP reporter induction), the GFP levels were measured in wild-type mES during the dose range finding (a correction is foreseen in case of proven interference) Data acceptance criteria for neg and pos controls were fixed Interpretation criteria were fixed All results from the ToxTracker ring trial are collected in a large database and are available for review.
	PR3 ToxTracker is an excellent system for rapidly assessing the genotoxicity induced by compounds, including their modes of action (MoA), within a short period of time. While this assay system utilizes a specialized instrument called FACS, it excels in enabling even beginners to conduct objective evaluations.
	PR4 Data integrity and transparency seemed satisfactory. Validation study seems to have been conducted appropriately. The procedures in place for collecting, analysis and interpreting were suitable for a validation test. Chemical selection was done by an independent group, samples were not identified to the testers. Blinding was carried out during testing and not broken until after results had been released. The VMT provided some technical support. Where the developer was involved in discussion over problems with conduct of studies no influence seems to have been put into discussions or release of details allowing chemical identification.

	PR5 It is very much appreciated that the substances were coded throughout the data analysis. Autofluorescent substances could interfere with the test system, but the protocol describes in detail how to handle such substances.
Q8. Do you consider that all the data supporting the assessment of the validity of the test method are easily available for expert review? These include: a detailed and readily available test method protocol to the public and independent laboratories; organised and easily accessible data to permit independent review(s); benchmarks by which an independent laboratory can itself assess its proper adherence to the protocol.	**PR1** Not enough description in the data file which makes it not easy to review the data. Table footnotes for the letters (e.q. P, N, I, E) used in the tables could be helpful. Adding a 2 fold increase threshold line and 75% cytotoxicity threshold line in the data figures (e.q. Figure 4, 5,6) would be helpful.
	PR2 Yes. No comments
	PR3 Yes. In this assay, a specialized device called FACS is used. In the past, this device was expensive, so the facilities that had it were limited. In recent years, the number of facilities with this device has increased, making it easier to use this assay.
	PR4 In general, yes. Protocol provided with checklists etc for monitoring progress with carrying out assay. Appreciable amounts of results are provided. Tables of results and Excel sheets show patterns of results. There are good presentations of dose-response relationships showing representative patterns of GFP reporter increases and viability in the report. It is not clear if the original assay data are available and, if so, how easy it would be to carry out a full confirmatory statistical analysis if it was considered useful. Table 12 is a summary of the combined ToxTracker reporter activations from the different validation laboratories in the ring trial. Results in the form of the positive/negative 'calls' are available in various places in the report. Tables 10 for positive call for the genotoxic compounds (28/32) & Table 11 for the non-genotoxic compounds (0/32 with 2 equivocal). Results in the report for sensitivity of 87% are based upon 26/30 and, for specificity, as 90% from 26/29 compounds. These 59 compounds are those where acceptable data were collected from at least 2 laboratories. It is not clear from Tables 10-12 which 5 compounds did not have acceptable data sets to include in the summary statistic. A number of annexes show results. Annex 3 is an MS Excel spreadsheet excel spreadsheet template for ToxTracker data analysis.

	Annex 4 is an example data set. Annex 5 shows "Results from phase 2 proficiency testing of the ToxTracker validation" with green, yellow and red for 6 +S9 and 6 -S9 compounds and has the comment "Description automatically generated with medium confidence". There is no key to the colours explaining the results. Colours presumably reflect the call for each of the 6 GEP from each of the 8 laboratories in Phase I. Annex 6: shows a Summary of the ToxTracker validation results for all reporter activations.
	PR5 As the validation report is publicly available, I consider the data/protocol to be easily accessible to independent laboratories and experts. However, it would be beneficial to also have access to all raw data and figures for all substances from all laboratories.

Example of an exploratory PCA analysis

D.P. Lovell 26th September 2023

Principal Component Analysis (PCA) of ToxTracker results of 6 reporter genes/biomarkers

A Principal Component Analysis (PCA) was carried out on the results for the 12 values (Positive or negative) of the 6 reporter genes with and without S9 mix. PCA reduces a set of highly correlated measures into a series of uncorrelated components. It is a powerful technique for dimension reduction of a data set, identification of clusters, detection of outliers and trends.

The table below shows the % variability 'explained' by the uncorrelated Principal Components (PCs).

In the Eigenanalysis table below, PC I 'explains' about 55% of the total variability, PC II a further 18%. Nearly 90% is 'explained' by the first 4 PCs. Given the relatively high number of PCs contributing to the total variance it is possible that some of these 'later' PC's will be identifying individuals with one or more outlying values or unusual patterns of results.

Th figures below show the scatterplots of PCII and PCIII v. PCI showing the degree of variability between the 64 chemicals. The groups numbers in the figures (1 to 28) are based upon the groups of chemicals which show the same pattern of results for the various reporter biomarkers (see table of chemicals, with group number and biomarker results). Red circles denote predominantly groups of genotoxic and blue predominantly non-genotoxic substances.

The loadings for PCI show a high number of positive results give a compound a high positive PC I score and a location to the right on the PCI axis; a high number of negative results give a high negative PC I score (left of PC I axis). PC II provides a contrast between results of for Srnx1, Blvrb and Dd1t3 compared with the other three biomarkers (for example Group 11 v. Groups 22 & 23).

PCIV emphasises the contrast between BScl2 and Blvrb with the other biomarkers while PCV contrasts results with added S9 mix with results without S9 mix.

Principal Component Analysis: BScl2-S9, BScl2+S9, Rtkn-S9, Rtkn+S9, Srnx1-S9, S

Eigenanalysis of the Correlation Matrix

Eigenvalue	6.5817	2.1789	1.0711	0.7343	0.5555	0.3386	0.2389	0.1466
Proportion	0.548	0.182	0.089	0.061	0.046	0.028	0.020	0.012
Cumulative	0.548	0.730	0.819	0.881	0.927	0.955	0.975	0.987

Eigenvalue	0.0710	0.0578	0.0194	0.0061
Proportion	0.006	0.005	0.002	0.001
Cumulative	0.993	0.998	0.999	1.000

Principal Component weightings/loadings

Variable	PC1	PC2	PC3	PC4	PC5
BScl2-S9	0.328	0.139	0.278	-0.332	-0.284
BScl2+S9	0.322	0.141	0.289	-0.303	0.070
Rtkn-S9	0.346	0.135	0.057	0.260	-0.128
Rtkn+S9	0.341	0.154	0.081	0.283	0.321
Srnx1-S9	0.281	-0.287	-0.350	0.158	-0.468
Srnx1+S9	0.260	-0.272	-0.472	0.289	0.058
Blvrb-S9	0.301	-0.122	-0.209	-0.552	-0.159
Blvrb+S9	0.283	-0.198	-0.289	-0.344	0.564
Dd1t3-S9	-0.022	-0.595	0.361	0.019	-0.142
Dd1t3+S9	-0.017	-0.592	0.364	0.022	0.210
Btg2-S9	0.345	0.028	0.216	0.200	-0.270
Btg2+S9	0.340	0.062	0.217	0.280	0.304

Table of grouping and biomarker results for 64 chemicals. (1 is genotoxicity and 2 non-genotoxicity) (Biomarker results are in the order in the table above.)

Genotoxicity	Compound	Group	Biomarker Results
1	Acrylonitrile	1	NNNNNNNNNNNN
1	Benzene	1	NNNNNNNNNNNN
1	Dimethylnitrosamine	1	NNNNNNNNNNNN
2	Ropinirolehydrochloride	1	NNNNNNNNNNNN
2	Methylcarbamate	1	NNNNNNNNNNNN
2	Diethanolamine	1	NNNNNNNNNNNN
2	Hexachloroethane	1	NNNNNNNNNNNN
2	Melamine	1	NNNNNNNNNNNN
2	Benzylalcohol	1	NNNNNNNNNNNN
2	o-anthranilicacid	1	NNNNNNNNNNNN
2	Ampicillintrihydrate	1	NNNNNNNNNNNN
2	Sodiumchloride	1	NNNNNNNNNNNN
2	D-mannitol	1	NNNNNNNNNNNN
2	Allylalcohol	1	NNNNNNNNNNNN
2	Chlormequatchloride	1	NNNNNNNNNNNN
2	Sulfisoxazole	1	NNNNNNNNNNNN
2	Sucrose	1	NNNNNNNNNNNN
2	Cyclohexanone	1	NNNNNNNNNNNN
2	1-Nitropropane	1	NNNNNNNNNNNN

2	Di(2-ethylhexyl)phthalate	2	NNNNNNNNPNNN
2	Tunicamycin110	3	NNNNNNNNPPNN
2	p-Nitrophenol	3	NNNNNNNNPPNN
2	PhenforminHCl	3	NNNNNNNNPPNN
2	2-ethyl-1,3-hexanedio	4	NNNNPPNNNNNN
2	Erythromycin	5	NNNNPPNNPNNN
2	Tolbutamide	6	NNNNPPNNPPNN
2	Sodiumdiclofenac	7	NNNNPPNNPPPE
2	Vanilin	8	NNNNPPNPNNNN
2	2-Phenylphenol	9	NNNNPPNPPPNN
2	Leadacetate	10	NNNNPPPPNNNN
2	Sodiumsaccharin	10	NNNNPPPPNNNN
2	Chlorpheniraminemaleate	10	NNNNPPPPNNNN
1	CadmiumChloride	11	NNNNPPPPPPNN
2	CyclosporinA598	11	NNNNPPPPPPNN
2	Phenanthrene	11	NNNNPPPPPPNN
2	Tertiarybutylhydroquinone	12	NNNNPPPPPPPP
1	12-Dimethyl-benzanthracene	13	NNNPNNNPNNNP
1	p-Chloroaniline	14	NNPPNPNPNNNP
1	2-Acetylaminofluorene	15	NNPPPPNNNNNN
1	6-Mercaptopurine	16	NNPPPPNNNNPP
1	9-Aminoacridine	16	NNPPPPNNNNPP
1	8-Hydroxyquinoline	17	NNPPPPNPPPPP
1	1,2-dibromoethane	18	NNPPPPPPNNNN
1	o-Anisidine	19	NNPPPPPPNNPP
1	Phenol	19	NNPPPPPPNNPP
1	Cyclophosphamide	20	NPNPNPNPNNNP
1	Benzo[a]pyrene	21	NPNPNPNPNPNP
1	Cytosinearabinose	22	PPPPNNPPNNPP
1	4-nitroquinoline-1-oxide	23	PPPPPNPNNNPP
1	2,4-Diaminotoluene	24	PPPPPNPPPPPP
1	Etoposide	25	PPPPPPNNNNPP
1	MitomycinC	26	PPPPPPPPNNPP
1	Cisplatin	26	PPPPPPPPNNPP
1	1,2-Dimethylhydrazine	26	PPPPPPPPNNPP
1	Azidothymidine	26	PPPPPPPPNNPP
1	ENU	26	PPPPPPPPNNPP
1	4,4’-Oxydianiline	26	PPPPPPPPNNPP
1	Busulfan	26	PPPPPPPPNNPP
1	Ethylmethanesulfonate	26	PPPPPPPPNNPP
1	p-Phenylenediamine2HCl	26	PPPPPPPPNNPP
1	3-Nitropropionicacid	26	PPPPPPPPNNPP
1	5-fluorouracil	26	PPPPPPPPNNPP
1	p-Anisidine	27	PPPPPPPPPNPP
1	2,6-Diaminotoluene	28	PPPPPPPPPPPP

Annex 4 - Developers' responses to questions

Developers' responses to questions based on written review comments provided 20 October 2023

1. Could you provide a focused summary of how the reference set of chemicals was chosen. I know you list some databases used by most developers of similar assays, but it would be good to have a consolidated list with references and then **in addition**, what was the strategy for choosing these particular chemicals for your assay.

Strategy for choosing these particular chemicals for your assay

The standard Ames bacterial reverse mutation test detects direct genotoxins where it can be deduced whether the mechanism for the induced mutagenic effect is base-pair substitution or frameshift mutation. The added value of ToxTracker is that this test not only detects direct genotoxins, but also detects indirect genotoxins by mechanisms of oxidative stress, cellular stress and protein damage. ToxTracker is important specifically with regard to the indirect mechanisms of genotoxicity, because there is a lack of test models. The aim of this extensive validation study (7 laboratories involved) was to include, in addition to direct genotoxins, a large number of chemicals that induce a genotoxic effect via oxidative stress, cellular stress and protein damage.

At the UKEMS/DEMS/BEMS meeting in Leuven on the 26th June 2017, the validation management team (VMT) came together to discuss and select a broad set of chemicals (organic, inorganic, aromatic and aliphatic molecules) with sufficient *in vitro* and *in vivo* genotoxicity data available to be able to predict the expected outcome in ToxTracker with confidence. Also a number of chemicals were selected that require metabolic activation in the liver. Four classes of chemicals (genotoxic carcinogens, genotoxic non-carcinogens, non-genotoxic carcinogens, non-genotoxic non-carcinogens) were considered. Chemical selection was based on publicly available lists and databases (Kirkland et al., 2011; Kirkland et al., 2016; Madia et al., 2020) and from the database of Toxys. Also the principles of chemical selection for the JaCVAM validation of the in vivo comet assay was considered (Morita et al., 2015).

Summary of how the reference set of chemicals was chosen

At the UKEMS/DEMS/BEMS meeting, 99 substances were identified by the VMT as potential candidates to be tested in the ToxTracker assay. The distribution of the 99 substances among the four classes was as follow:

- Genotoxic carcinogens: 34
- Genotoxic non-carcinogens: 11
- Non-genotoxic carcinogens: 22
- Non-genotoxic non-carcinogens: 32

Thereafter, information on IARC class, *in vitro* (Ames, MLA/HPRT, MNvit, CAvit) and *in vivo* (MN, CA, Comet, UDS) genotoxicity results and, if results on the outcome in the ToxTracker assay were already available (DNA damage, oxidative stress and protein damage) was gathered and tabulated to be further discussed by the VMT. For the feasibility of the validation study, the number of substances to be tested had to be reduced. Furthermore, the VMT found it not necessary to select 25% of the chemicals to be included in each of the four classes. Therefore, fewer chemicals were selected for the classes of genotoxic non-carcinogens and non-genotoxic carcinogens since many of these chemicals

have undefined modes of action and hence difficult to allocate to one of these classes. A total of 72 chemicals, well balanced in terms of chemical classes, direct versus indirect action and 50% positives and 50% negatives, were selected. The final distribution was as follows:

- Genotoxic carcinogens: 26
- Genotoxic non-carcinogens: 10
- Non-genotoxic carcinogens: 10
- Non-genotoxic non-carcinogens: 26

Upon purchasing the chemicals, a lot of discussion was required for chemicals that were not commercially available anymore, and so it was necessary to have them replaced by alternatives if possible. Furthermore, chemicals showing conflicting data or which were extremely expensive to be included in this extensive validation study were excluded. The list of chemicals was finally reduced to 64 chemicals showing a distribution of:

- Genotoxic carcinogens: 22
- Genotoxic non-carcinogens: 10
- Non-genotoxic carcinogens: 10
- Non-genotoxic non-carcinogens: 22

References

- Kirkland D, Reeve L, Gatehouse D, Vanparys P. A core in vitro genotoxicity battery comprising the Ames test plus the in vitro micronucleus test is sufficient to detect rodent carcinogens and in vivo genotoxins. Mutation Research/Genetic Toxicology and Environmental Mutagenesis. 2011 Mar;721(1):27–73.
- Kirkland D, Kasper P, Martus HJ, Müller L, van Benthem J, Madia F, et al. Updated recommended lists of genotoxic and non-genotoxic chemicals for assessment of the performance of new or improved genotoxicity tests. Mutat Res Genet Toxicol Environ Mutagen [Internet]. 2016 Jan 1 [cited 2022 Jul 14];795:7–30. Available from: https://pubmed.ncbi.nlm.nih.gov/26774663/
- Madia F, Kirkland D, Morita T, White P, Asturiol D, Corvi R. EURL ECVAM Genotoxicity and Carcinogenicity Database of Substances Eliciting Negative Results in the Ames Test: Construction of the Database. Mutation Research/Genetic Toxicology and Environmental Mutagenesis. 2020 Jun;854–855:503199.
- Morita T, Uno Y, Honma M, Kojima H, Hayashi M, Tice RR, Corvi R, Schechtman L. The JaCVAM international validation study on the in vivo comet assay: Selection of test chemicals. Mutat Res Genet Toxicol Environ Mutagen. 2015 Jul;786-788:14-44.

2. What is the basis for the choice of the 2-fold increase in GFP for a significant response (empiric and/or statistics and/or AI approach, e.g., Deep Learning…)?

Originally, the threshold for a positive response in ToxTracker was set at a 1.5-fold increase in GFP fluorescence levels. GFP reporter expression was always measured in intact cells after 24 hour exposure to the compound by flow cytometry. GFP fluorescence after exposure was expressed as mean fluorescence intensity and compared to vehicle control cell cultures. The 1.5-fold limit was based on 3 times the standard deviation of the fluorescence levels in vehicle-exposed cells (Hendriks et al., 2016). During earlier validation projects where different flow cytometers were used for measurement of the ToxTracker reporters, we observed that the variation in the background levels increased. For this reason, the threshold for a positive

ToxTracker result was set at 2-fold. This threshold is still based on 3 times the standard deviation of vehicle exposed cultures.

More recently, we applied the benchmark dose response modeling (BMD) approach to quantify the ToxTracker responses (Boisvert et al., 2022). To determine the BMD, we first needed to set the critical effect size/bench mark response that defines an increase in ToxTracker reporter signals that could not be explained by the variation in the background fluorescence levels.
We performed a bootstrapping analyses of 1500 vehicle exposed ToxTracker cell cultures and determined the geometric 95th and 99th percentile of the control ratio distributions across all studies investigated. The resulting cut-off values for a significant ToxTracker reporter activation across all reporters were determined as the geometric 95th percentile of the control ratio distributions across all studies investigated, ranged from 1.46 to 1.62. The average value across all reporters is approximately 1.5, i.e., a 1.5-fold change indicates a significant positive response. Thus, the cut-off fold-change value for a weakly positive response determined herein is identical to that already used by Toxys, i.e., a 1.5-fold increase over background. Furthermore, the cut-off values identified using the geometric 99th percentile of the control ratio distribution range from 1.64 to 1.88, with an average value of 1.74 for all the reporters. With respect to delineation of a strong positive response, the value currently used (i.e., a 2-fold increase) is 15% greater than that determined herein. This indicates that the approach for identification of strong positive responses is conservative, i.e., would be less than a 1% probability that a GFP induction of 2-fold change above control could be obtained by chance alone.

References

- Hendriks G, Derr RS, Misovic B, Morolli B, Calléja FM, Vrieling H. The Extended ToxTracker Assay Discriminates Between Induction of DNA Damage, Oxidative Stress, and Protein Misfolding. Toxicol Sci. 2016 Mar;150(1):190-203. doi: 10.1093/toxsci/kfv323. Epub 2015 Dec 29. PMID: 26719371; PMCID: PMC5009621.

- Boisvert L, Derr R, Osterlund T, Hendriks G, Brandsma I. Quantitative interpretation of ToxTracker dose-response data for potency comparisons and mode-of-action determination. Environ Mol Mutagen. 2023 Feb;64(2):132-143. doi: 10.1002/em.22525. Epub 2023 Feb 2. PMID: 36645179.

3. Reviewers commented on lack of aneugenic chemicals in the set of chemicals used for development, given at some part of the report it is stated that "ToxTracker was able to correctly predict the genotoxic MoA of the tested compounds, including (……) and differentiation between a clastogenic or aneugenic MoA". Would you like to provide some input on this aspect?

Indeed, the VMT did not include any classical aneugenic substances, i.e. tubulin poisons or kinase inhibitors, to the ToxTracker validation. The primary objective was to determine whether ToxTracker could discriminate between DNA-reactive (genotoxic) substances and non-DNA-reactive (non-genotoxic) substances. ToxTracker does not contain a dedicated reporter gene that is activated upon exposure to aneugenic substances. However, ToxTracker typically classifies aneugenic substances as genotoxic (Hendriks et al., 2016). Genotoxic compounds with an aneugenic MOA in general selectively activate the Rtkn-GFP reporter in ToxTracker but not the Bscl2-GFP reporter. In contrast, DNA-reactive chemicals typically

activate both the Rtkn-GFP and Bscl2-GFP reporters. Confirmation of the aneugenic MOA could be obtained from the kinetics of Rtkn-GFP reporter activation. For directly DNA-reactive chemicals, activation of Rtkn-GFP was readily observable after 6-8 h. For tubulin poisons, the activation of the Rtkn-GFP took 12-14 h (Hendriks et al., 2016). Including analysis of the kinetics of ToxTracker reporter activation in the interlaboratory validation was considered too complex and out-of-scope.
More recently, an extended ToxTracker assay protocol was developed to extend the classification of aneugenic and clastogenic compounds by including cell cycle analysis and aneuploidy assessment (Brandsma et al., 2020). The extended ToxTracker ACE protocol was validated using a selection of 16 genotoxic compounds with a well-established clastogenic or aneugenic MOA. The ACE protocol does not improve the genotoxicity prediction by ToxTracker, but can aid the mode-of-action assessment of genotoxic substances.

References

- Brandsma I, Moelijker N, Derr R, Hendriks G. Aneugen Versus Clastogen Evaluation and Oxidative Stress-Related Mode-of-Action Assessment of Genotoxic Compounds Using the ToxTracker Reporter Assay. Toxicol Sci. 2020 Sep 1;177(1):202-213. doi: 10.1093/toxsci/kfaa103. PMID: 32617558.

4. What is the rationale selecting the highest testing concentration of 1 mg/mL, given that (other OECD TGs for testing genotoxicity in vitro) recommends 2 mg/mL as the highest test concentration when no precipitation or cytotoxicity?

Selection of the highest testing concentration was based on ICH S2 (R1) and recommendations from the IWGT. In ICH S2 (R1), itis recommended to use 1 mM or 0.5 mg/ml as top concentration for *in vitro* mammalian cell assays. Only in case of a very low molecular weight, higher test concentrations should be considered. In the ToxTracker validation, all test substances were blinded before distribution to the different laboratories. To avoid the possibility to reveal the compound names, no molecular weights could be shared. For this reason, a top concentration was set in mg/ml. And because some compounds have a low molecular weight and we wanted to avoid the possibility that these would be tested at a too low concentration, we set the top concentration for all compounds at 1 mg/ml or the maximum soluble concentration.

5. Was solvent control ever part of the test development? Some reviewers suggest that solvent control in addition to negative control is needed (which I believe is not part of the protocol now, so it would be good to know if this was considered/evaluated and found to not be necessary, or less informative then the negative controls only.

In every ToxTracker experiment, a solvent control is included as negative control. In every well of the ToxTracker plates, the same amount of solvent is added. By default each ToxTracker cell line is exposed to 1% DMSO, with or with the test compound. Induction of the ToxTracker reporters is calculated by comparing the mean GFP fluorescence intensity in

exposed cultures and solvent-exposed cultures. This is defined in the ToxTracker protocol and applied to all solvents used, DMSO and PBS.

6. Which 5 compounds (Tables 10-12) did not have acceptable data sets to include in the summary statistic?

The following critera were applied for data acceptance:

- Did the positive controls (cisplatin for DNA damage, DEM for oxidative stress, tunicamycin for protein unfolding and aflatoxin B1 for S9) induce the mimimum reporter activation in the relevant cell lines
- Was the cytotoxicity of the positive controls within the expected range
- Did the test compound induce <75% cytotoxicity. All concentrations that induce more than 75% cytotoxicity are omitted from data analysis

In cases where the positive controls did not meet the acceptance critereria listed above, the data for all the compounds in that experiment were omitted from data analysis.

Compounds for which we obtained acceptable data from two or three laboratories were used for calculations of between-lab reproducibility (BLR). Compound for which acceptable data were reviewed from one lab were excluded.
The following five compounds had only valid data for one laboratory:

- No. 11: Benzene (CAS 71-43-2)
- No. 23: 6-Mercaptopurine (CAS 50-44-2)
- No. 35: 2-Phenylphenol sodium salt (CAS 6152-33-6)
- No. 38: Cyclosporin A (CAS 59865-13-3)
- No. 57: D-mannitol (CAS 69-65-8)

Annex 5 - Some resources for estimating of sensitivity and specificity with confidence intervals

"No standardized guidelines exist for the biostatistical methods appropriate for studies evaluating diagnostic tests. Publication recommendations such as the STARD statement provide guidance for the analysis of data, but biostatistical advice is minimal and application is inconsistent. This article aims to provide a self-contained, accessible resource on the biostatistical aspects of study design and reporting for investigators. **For all dichotomous diagnostic tests, estimates of sensitivity and specificity should be reported with confidence intervals.** Power calculations are strongly recommended to ensure that investigators achieve desired levels of precision."
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3661219/#:~:text=The%20CRS%20method%20suggests%20that,(95.7%20%25%2C%20100.6%20%25)

"Statistical uncertainty about diagnostic performance parameters, e.g**. DSe and DSp, should be presented as confidence intervals (CI). Typically, a 95% CI is used** and its width (precision of the estimated value) depends strongly on the sample size used for parameter estimation. Exact CI are preferred to normal approximations because they avoid upper limits that exceed 100%."
https://www.woah.org/fileadmin/Home/fr/Health_standards/tahm/2.02.05_STATISTICAL_VALIDATION.pdf

"Confidence intervals: It should be mandatory to report confidence intervals for any measure of diagnostic accuracy"'
https://core.ac.uk/download/pdf/159149781.pdf

"We calculated sensitivity and specificity and 95% confidence intervals (exact binomial method) for each assay, using the manufacturers' thresholds and in line with the current MHRA TPP criteria."
https://assets.publishing.service.gov.uk/media/5f05887ad3bf7f2be6e02161/Evaluation__of_sensitivity_and_specificity_of_4_commercially_available_SARS-CoV-2_antibody_immunoassays.pdf

"Measures that describe diagnostic accuracy: There are different ways to describe diagnostic accuracy. Appropriate measures include estimates of sensitivity and specificity pairs, likelihood ratio of positive and negative result pairs, and ROC (Receiver Operating Characteristic) analysis along with confidence intervals."
https://www.fda.gov/media/71147/download

The ToxTracker assay: a stem cell-based reporter assay for mechanistic carcinogenicity hazard screening

Validation report

Report prepared by:

Giel Hendriks[1,*], David Kirkland[2], Philippe Vanparys[3], Jan van Benthem[4] and Els Adriaens[5]

for the Organization for Economic Cooperation and Development Working Group of the National Coordinators of the Test Guidelines Program.

[1] Toxys, Oegstgeest, The Netherlands
[2] Kirkland Consulting, United Kingdom
[3] Consultant Genetic Toxicology, Belgium
[4] National Institute for Public Health and the Environment (RIVM), The Netherlands
[5] Adriaens Consulting, Aalter, Belgium

*Direct correspondence to:

Giel Hendriks, Ph.D.
Toxys
Leiden Bio Science Park
De Limes 7, 2342 DH, Oegstgeest
The Netherlands
+31 71 3322470
g.hendriks@toxys.com

Table of contents

1. List of abbreviations and acronyms

AOP	Adverse Outcome Pathway
BLR	Between-laboratory reproducibility
BMD	Benchmark Dose response
CA	Chromosomal Aberration assay
ECHA	European Chemicals Agency
EFSA	European Food Safety Authority
ECVAM	European Center for the Validation of Alternative Methods
GD	Guidance Document
GTTC	Genetic Toxicology Technical Committee at HESI
IATA	Integrated Approached to Testing and Assessment
ICH	International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use
IWGT	International Working group on Genetic Toxicology
IVIVE	*In Vitro* to *In Vivo* Extrapolation
LOAEL	Lowest observed adverse effect level
MLA	Mouse Lymphoma Assay
MN	Micronucleus Assay (*in vivo*)
MNvit	*in vitro* Micronucleus Assay
MoA	Mode of Action
NAM	New Approach Method
NOAEL	No observed adverse effect level
NRC	U.S. National Research Council
OECD	Organization for Economic Co-operation and Development
PARC	European Partnership for the Assessment of Risks from Chemicals
SCCS	Scientific Committee on Consumer Safety
SD	Standard deviation
SPSF	Standard Project Submission Form
TG	(OECD) Test Guideline
TGR	Transgenic rodent assay
TPF	Test Presubmission Form
VMT	Validation Management Team
WLR	Within-laboratory reproducibility
WNT	Working Group of National Coordinators of the OECD TGs program
WoE	Weight of Evidence

2. Introduction

1. In November 2016, an SPSF for a new test guideline (TG) for the ToxTracker assay was submitted to the OECD by The Netherlands. The SPSF application was primarily aimed at informing the OECD members on the ToxTracker assay and its applications for mechanistic genotoxicity and cancer hazard/risk assessment. The proposal was to perform an extensive prospective inter-laboratory validation study of the ToxTracker assay that should eventually lead to an official ToxTracker TG. The interlaboratory validation was performed according to OECD Guidance Document 34 (GD 34; OECD, 2005). On April 2017, the WNT accepted the SPSF on ToxTracker.

2. In parallel to the OECD submission, a Test Presubmission Form (TPF) for ToxTracker was submitted to the European Union Reference Laboratory for Alternatives to Animal Testing (EURL ECVAM) on November 2016 (TM2016-03). In the TPF, ToxTracker was proposed as a novel approach for *in vitro* carcinogenicity hazard identification. The ToxTracker presubmission was reviewed by PARERE on May 2017. The preliminary response from PARERE on the presubmission was that ToxTracker would be a valuable expansion of the *in vitro* genotoxicity testing battery for regulatory applications. ToxTracker should not be considered as an alternative to the already existing *in vitro* genotoxicity assays, but as an additional tool that provides genotoxic mode-of-action information that would be useful to move towards the new paradigm of a mechanistically based risk assessment.

3. Genotoxicity testing is an essential part of chemical safety assessment. Recommendations for *in vitro* and *in vivo* testing are provided by various international guidelines, including for human pharmaceuticals in ICH S2(R1) (1), for industrial chemicals from ECHA (2), for foods and food additives from EFSA (3) and for cosmetic ingredients from SCCS (4). With the report from the U.S. National Research Council (NRC) in 2018 on Toxicity Testing in the 21st Century (5) as well as from various international initiatives to improve the safety testing approaches (i.e. US Tox21, EuToxRisk, PARC) comes a strong call for a major paradigm shift in toxicity testing. The use of novel approaches should allow chemical risk assessment to move beyond the classical endpoints and non-relevant test systems and extend the integration of mechanism-based toxicity test strategies that includes quantitative assessment of the mode-of-action (MoA) for (geno)toxic chemicals. In fact, in various legislations, there is already room for the use of New Approach Methods (NAMs) for chemical safety assessment. Data from these NAMs, for instance about the Mode-of-Action of genotoxic compounds, can be used in a Weight-of-Evidence (WoE) approach (4,6). The ToxTracker assay nicely fits into such an approach, as it provides mechanistic insight into the genetic toxicity and cancer hazard of chemicals. ToxTracker is a mouse stem cell-based reporter assay that can identify genotoxic compounds with high accuracy (7,8). By combining different fluorescent reporter genes, the assay is able to provide insight into the MoA of genotoxic substances. The mechanism-based genotoxicity information can be particularly useful in AOP and WoE approaches and

can contribute to the development of an Integrated Approach to Testing and Assessment (IATA).

3. Goals of the interlaboratory validation project

4. Although the primary use of Tox Tracker will be to provide insight into the MoA of genotoxic substances, such information is not valuable unless it has been demonstrated that the test system can reliably discriminate between mutagenic/genotoxic DNA reactive and not-genotoxic substances. Therefore, the primary goal of this validation project described in this report was to determine if ToxTracker is able to correctly predict the genotoxic properties of compounds. A broad selection of well-established genotoxic and non-genotoxic compounds was tested in ToxTracker to establish the sensitivity and specificity of the assay. In this interlaboratory validation study, the transferability and reproducibility of the ToxTracker assay was also established. The genotoxicity predictions for the tested compounds were compared from various repeat tests within a laboratory and between laboratories to calculate the within-lab and between-lab reproducibility.

5. ToxTracker is considered as an expansion of the toolbox of *in vitro* genotoxicity test to provide insight into the MoA of genotoxic substances. The secondary goal of the validation project was to investigate the MoA information that is provided by ToxTracker and the relevance for the genotoxicity prediction of the tested compounds. Information about the MoA of genotoxic and non-genotoxic compounds was applied to discriminate between direct and indirect genotoxic compounds and to better understand the results from ToxTracker in relation to results from the current standard *in vitro* and *in vivo* genotoxicity assays.

4. Scientific basis for the test method

6. ToxTracker combines six fluorescent reporter genes that are specifically activated by different cellular signaling responses that are associated with genotoxicity and carcinogenicity. The biomarker genes that are applied in ToxTracker were selected from toxicogenomics studies in which mouse stem cells were exposed to forty different genotoxic and non-genotoxic carcinogens (9). By detecting the activation of these reporter genes following chemical exposure, ToxTracker can discriminate between the induction of DNA damage, oxidative stress and protein damage and provide insight into the MoA of genotoxic substances in a single test (Figure 1) (10). An accurate genotoxicity prediction in ToxTracker, and relevant MoA information is based on the combined profile of all six fluorescent reporter genes.

7. In ToxTracker, genotoxicity is primarily predicted by the activation of either of the two independent fluorescent reporters Bscl2-GFP and Rtkn-GFP. The Bscl2-GFP reporter is activated upon the formation of bulky DNA adducts and subsequent inhibition of DNA replication which is a potent activator of the DNA damage response. These replication-blocking DNA lesions often lead to the formation of mutations. Activation of

the Rtkn-GFP genotoxicity reporter is associated with induction of DNA double-strand breaks. Many clastogenic compounds cause DNA damage by direct interaction with the DNA and typically activate both the Bscl2-GFP and Rtkn-GFP ToxTracker reporters. For such DNA reactive compounds, which may pose a cancer risk even at very low doses, a linear approach is typically used for risk assessment (11). In contrast, there are also substances that are genotoxic without directly interacting with the DNA (indirect genotoxicity) (11). Tubulin poisons, which interfere with chromosome segregation during mitosis, are indirect genotoxins and exposure to such agents can lead to aneuploidy (12). Activation of the Rtkn-GFP reporter and arresting of cells in mitosis is often observed for aneugens (13). The genotoxicity of compounds is further confirmed in ToxTracker by the Btg2-GFP reporter. Btg2 is a component of the P53-dependent DNA damage response and is involved in regulation of the G1/S cell cycle checkpoint but is also induced by various other cellular stressors. Also compounds that cause high levels of oxidative stress in the cells can indirectly affect the DNA. Insufficient or faulty repair of oxidative DNA damage can lead to mutations or chromosomal aberrations. Induction of oxidative stress is detected in ToxTracker by the Srxn1-GFP and Blvrb-GFP reporters that are activated by the two major antioxidant pathways in the cell. Also, protein misfolding or damage can lead indirectly to DNA damage. Activation of the unfolded protein response following chemical exposure is a potent trigger of apoptosis, leading to DNA breaks and chromosome fragmentation. Activation of the unfolded protein response is detected in ToxTracker by the Ddit3-GFP reporter. For these types of indirect damage, compensatory mechanisms exist, such as the presence of endogenous antioxidants. These compensatory mechanisms might give rise to a non-linear dose response and therefore a threshold approach is considered appropriate for such indirect genotoxins (14). To accurately predict the genotoxicity of compounds, both direct and indirect genotoxic effects should be considered. For this reason, integration of all six reporter genes in ToxTracker is required for reliable chemical safety assessment as well as for providing mechanistic information.

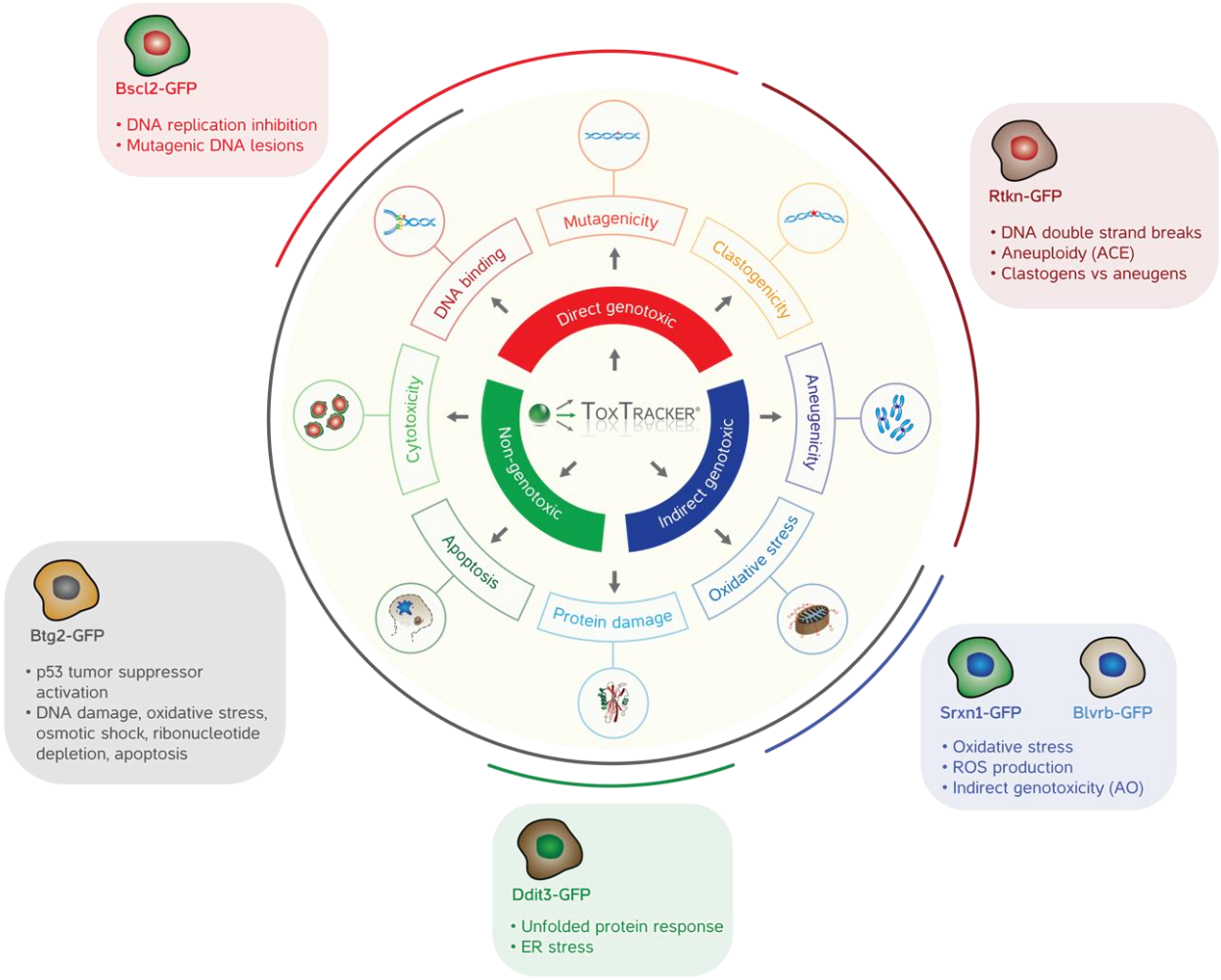


Figure 1: Schematic representation of the genotoxic and non-genotoxic endpoints covered in the ToxTracker assay.

8. ToxTracker is a GFP-based reporter assay, using untransformed mouse embryonic stem cells (mESC). Mutations in stem cells have been shown to play a crucial role in tumorigenesis. Stem cells are highly relevant for carcinogenicity hazard assessment and are highly suitable for mutation and genotoxicity testing due to their high cell proliferation rate (15). mESC are primary cells that are genetically stable, have an infinite life span and are proficient in all major DNA damage signaling and cell cycle regulation pathways. Activation of the fluorescent reporters in ToxTracker is measured by flow cytometry. GFP signals in individual cells are quantified in exposed and non-exposed cultures. Simultaneously, cytotoxicity of the tested compounds is determined by relative cell count.

5. The relationship between ToxTracker and *in vivo* biological effects

9. Extensive technical validation (>400 compounds) showed that ToxTracker combines a very high sensitivity (94%) and specificity (95%) for the detection of *in vivo* genotoxicity with the ability to provide insight into the MOA of genotoxic agents (7,8,13). ToxTracker contains reporters that predict induction of gene mutations and chromosomal damage with a high accuracy (>90%). In many cases, ToxTracker was able to correctly predict the genotoxic MoA of the tested compounds, including discrimination between direct and indirect genotoxicity, genotoxicity related to oxidative stress and differentiation between a clastogenic or aneugenic MoA. For the genotoxicity prediction, the differential activation of all six ToxTracker reporters was assessed. Activation of the ToxTracker reporter genes does show a strong correlation with the standard *in vitro* and *in vivo* genotoxicity assays. In a correlation study with 66 compounds from the ECVAM library of reference compounds (16), activation of the Bscl2-GFP reporter gene for induction of mutagenic DNA adducts showed a 93% correlation with a positive result in the Ames and/or MLA mutation assays. A negative result for the Bscl2-GFP ToxTracker reporter showed a 91% correlation with negative Ames and MLA results. Activation of the Rtkn-GFP reporter gene in ToxTracker that indicated the induction of DNA double strand breaks shows a very strong correlation of 92% with a positive result in the *in vivo* MN assay (17). A negative Rtkn-GFP reporter response correlated in 91% of the cases with a negative in vivo MN result. Interestingly, 30-40% of compounds that were negative in ToxTracker and negative in the *in vivo* MN assay did induce MN *in vitro*, underscoring the limited specificity of the *in vitro* MN assay (18). In many cases, the discrepancy between the *in vitro* MN and ToxTracker or *in vivo* MN assay could be explained by high levels of oxidative stress induction by the compound *in vitro*, e.g. tert-butylhydroquinone and resorcinol. Both compounds induce MN *in vitro*. ToxTracker was able to confirm the induction of oxidative stress and classified tert-butylhydroquinone and resorcinol as non-genotoxic, in line with the *in vivo* genotoxicity classification (Figure 2). Also for non-genotoxic compounds the mode-of-action information can be valuable. Induction of oxidative stress and the unfolded protein response have been shown to play a role chemical carcinogenesis (19,20).

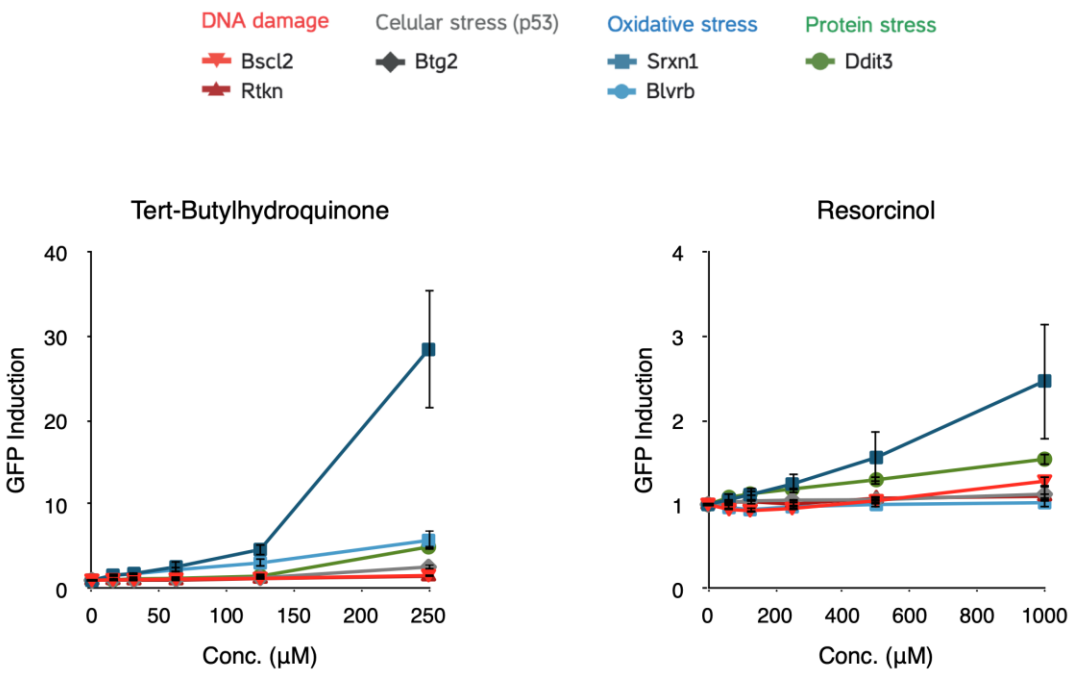


Figure 2: ToxTracker results for tert-butylhydroquinone and resorcinol. Both compounds activated the oxidative stress reporters (Srxn1-GFP; blue). None of the compounds activated the genotoxicity reporters (Bscl2-GFP/Rtkn-GFP; red)

6. Regulatory applications

10. The current standard battery of *in vitro* genotoxicity assays generally lacks the ability to provide information about the MoA of genotoxic substances. None of the existing OECD TG programs cover the same endpoints as ToxTracker. The battery of *in vitro* assays that are currently accepted for regulatory genetic toxicity testing include the bacterial Ames (TG471), mouse lymphoma (TG490) and mammalian cell (TG476) *in vitro* mutation tests. Chromosome damage is assessed by the *in vitro* micronucleus (TG487) and chromosome aberration (TG473) tests. This battery of *in vitro* genotoxicity assays generally has a sufficient sensitivity for genotoxicant identification but occasionally suffers from a relative high frequency of misleading positive test results (i.e., low specificity). The lack of mechanistic insight into these positive test results can be a serious challenge in the hazard assessment of chemicals and pharmaceuticals. By integrating the different reporter genes, ToxTracker can provide insight into the MoA of genotoxic compounds by predicting the induction of mutagenic DNA lesions, clastogenic effects, induction of aneugenicity and indirect genotoxicity caused by oxidative stress and protein damage in a single assay.

11. ToxTracker is already widely used by industry and research organizations for genotoxicity assessment (21–23). The ToxTracker assay is typically applied as (i) a screening assay for early *in vitro* genotoxicity prediction or (ii) as a follow-up test to assess the MoA of compounds that give a positive or equivocal result in the regulatory

accepted *in vitro* and *in vivo* genotoxicity assays. Application of ToxTracker as a screening assay can provide mechanistic information and indications on the compounds' genotoxicity that may indicate which classical *in vitro* genotoxicity tests would be most relevant to perform. Alternatively, ToxTracker would be recommended following a positive result from the current *in vitro* genotoxicity battery (Ames/MLA for mutagenicity and MNvit/CA for chromosome damage). ToxTracker is currently not intended to replace the existing TGs for *in vitro* genotoxicity testing. ToxTracker proved to be a valuable addition to the standard battery of *in vitro* genotoxicity assays for regulatory applications as it has the unique ability to reveal genotoxic modes of action and a number of non-genotoxic modes of action, such as oxidative stress and protein damage. ToxTracker results have already been included successfully in regulatory submissions to the FDA and US-EPA in the United States and to EMA, ECHA and EFSA in Europe as part of a WoE approach.

12. There is currently regulatory demand for more quantitative approaches for human risk assessment. In the standard battery of *in vitro* genotoxicity assays, often the no observed adverse effect level (NOAEL) or lowest observed adverse effect level (LOAEL) concentrations are calculated. However, the NOAEL/LOAEL can be adversely affected by study design and dose selection. As an alternative, quantitative dose-response modeling can be used to determine a robust potency metric such as the Benchmark Dose (BMD) (24). Application of the BMD for human risk assessment and calculation of health-based guidance values (HBGV) are already applied for regulatory genotoxicity testing (25–27). Data from the ToxTracker assay are highly suited for quantitative dose-response modeling using the BMD approach. By calculation the BMD for the different ToxTracker reporters, chemicals can be quantitatively ranked according to their potency to induce genotoxic through a specific MoA (26)(28). The BMD values from the ToxTracker reporters can be applied in IVIVE approaches (29), supporting the regulatory demands for quantitative assessment of human risk assessment.

7. Protocol for conducting the ToxTracker assay

13. The general protocol for conducting the ToxTracker assay has been published previously (8). See Annex 1 for the full ToxTracker protocol. The assays are performed in 96-well plates that contain the six different GFP reporter cell lines. The cells are seeded 24h prior to exposure to the test substances. Typically, five concentrations of a compound are tested in ToxTracker and four positive control compounds for each of the different GFP reporters are included in every test. Induction of the fluorescent ToxTracker reporters as well as the number of viable cells is determined by flow cytometry after 24h exposure of the cells to the test compounds. Each compound is tested in three independent biological repeat experiments. The relative induction of the GFP reporters is calculated from the mean fluorescence from the three repeats by comparing treated cultures with the related vehicle control cultures. Cytotoxicity is determined based on relative cell count in exposed and control cultures.

14. Four positive control compounds have been selected to ensure the proper response of the different reporter genes. Compounds were selected based on their different MoA. Cisplatin is a DNA cross-linking agent that is included as a positive control for activation of the Bscl2-GFP, Rtkn-GFP and Btg2-GFP genotoxicity reporters. Aflatoxin B1 is a mutagenic substance that requires metabolic activation and is applied as positive control for S9 activity, Diethyl maleate activates the Nrf2-dependent antioxidant response and is applied as a positive control for the oxidative stress reporters Srxn1-GFP and Blvrb-GFP. Tunicamycin is a specific activator of the unfolded protein response and is the positive control for the protein damage reporter Ddit3-GFP. Each of the positive controls selectively activates the associated reporter genes, underscoring the specificity of the different reporters (Figure 3). The specificity of the reporters is not impacted by cytotoxicity, at least not up to the maximum acceptable cell survival level of 75%.

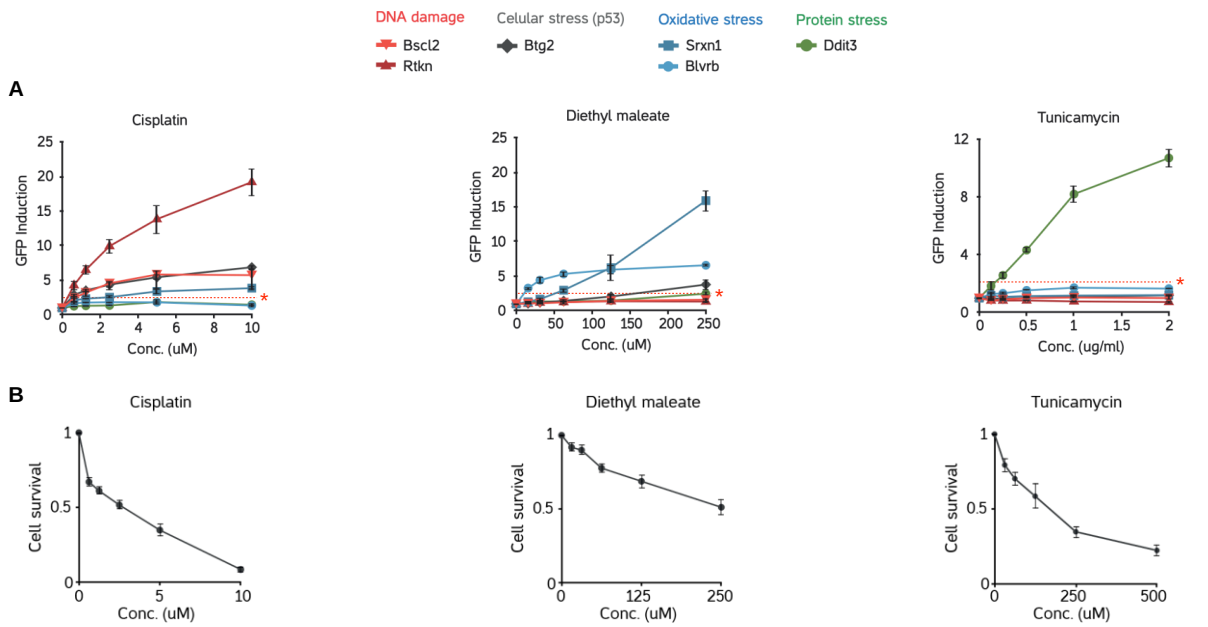


Figure 3: Selective activation of the ToxTracker reporter genes following exposure to the positive control compounds. A) The ToxTracker cells were exposed to increasing concentrations of the DNA damaging agent cisplatin, the oxidative stress-inducing agent diethyl maleate (DEM) or the unfolded protein response-activating compound tunicamycin. GFP induction levels in intact cells were determined by flow cytometry at 24 h after initiation of the exposure. B) Cell survival was determined by flow cytometry after 24-h exposure as the relative decrease in cell concentration compared with untreated controls. * 2-fold GFP reporter induction is the threshold for positive ToxTracker result.

15. Metabolization of compounds by the liver can have a major impact on their genotoxic properties. In common with the standard *in vitro* genotoxicity assays, ToxTracker uses S9 liver extract from Aroclor1254- or phenobarbital/β-naphthoflavone-induced rats for metabolic activation of compounds. The S9 metabolization protocol has previously been optimized for ToxTracker to allow for assessment of all the different endpoints

that are covered by the assay. The ToxTracker assay is by default always performed in the absence and presence of a S9 metabolizing system.

16. The criteria for a positive or negative test result are defined in the ToxTracker protocol (Table 1). Only GFP reporter inductions at compound concentrations that induce <75% cytotoxicity are acceptable for ToxTracker analysis. A compound is classified as genotoxic if at least a 2-fold increase in expression of the Bscl2-GFP and/or Rtkn-GFP reporters is induced. Also, for the other four reporter genes, a 2-fold increase is GFP signals is applied as the threshold for a result to be considered positive. The 2-fold increase in GFP reporter activation as cut-off for a positive response is based on 3 times the standard deviation (SD) of the fluorescence levels in solvent control cell cultures. The validity of the 2-fold induction threshold for a positive ToxTracker result was recently confirmed using a bootstrapping analysis of more than a 1000 solvent control ToxTracker cultures (28). For each ToxTracker reporter the distribution of background fluorescence levels was determined (30). Compounds are classified as non-genotoxic in ToxTracker if the induction levels of the genotoxicity reporter genes (Bscl2-GFP and Rtkn-GFP) is less than 1.5-fold. In the case that fluorescent reporter activation exceeds 1.5-fold but remains below a 2-fold increase, a weak positive score (+) is applied, but only on case of a clear dose response.

17. For assessment of the mode-of-action of compounds, also induction of the other ToxTracker reporters is included. Btg2 induction is associated with activation of the p53 tumor suppressor gene. Srxn1-GFP and Blvrb-GFP activation are indicative for the induction of cellular oxidative stress and Ddit3-GFP activation is associated with the unfolded protein response. For each independent ToxTracker experiment, activation of the different reporters is assessed, and compounds are classified accordingly. The criteria for a positive and negative test result are equal for all the ToxTracker reporters. For a final conclusion, the test results for every reporter gene from three independent repeat experiments are weighted according to the prediction model below (Table 2).

Table 1: Classification criteria for a positive ToxTracker test result and assessment calls for a single experiment.

GFP induction factor	Viability	Dose-response	Call
≥2.0 at 1 or more concentrations	≥0.25	Yes	+
≥2.0 at 1 or more concentrations	≥0.25	No	+
<1.5 at all concentrations	≥0.25 but approaches 0.25	No	-
<1.5 at all concentrations	≥0.25 but limited by precipitation	No	-
<1.5 at all concentrations	≥0.25 but with limited toxicity and not limited by precipitation	No	(-)
>1.5 but <2.0	≥0.25	Yes	(+)
>1.5 but <2.0	≥0.25	No	-

Table 2: Prediction model for ToxTracker.

Calls in 3 experiments (in any order)	Overall call for reporter with – or +S9 condition
+ + +	+
+ + (+)	+
+ + -	+
+ (+) (+)	+
- - -	-
- - (+)	-
- - +	-
- (+) (+)	-
(+) (+) (+)	E
+ (+) -	E
+ (+) (-)	E

18. To ensure a reliable classification of compounds, various quality controls and data acceptance criteria have been defined for the ToxTracker assay (see annex 1). Proper growth of the GFP reporter cell lines is monitored by tracking their proliferation rate. Criteria for the minimal proliferation rate of the cell lines is defined. In every ToxTracker experiment, positive controls for induction of DNA damage (cisplatin), S9

metabolization (aflatoxin B1), oxidative stress (diethyl maleate) and protein unfolding (tunicamycin) are included. Minimum induction levels for the different reporters have been defined. In the case that these minimum induction levels are not reached, an experiment should be discarded. Also, instructions for preparing compound dilutions and handling solubility issues are defined in the protocol.

8. Interlaboratory validation study management

19. Following the SPSF application by The Netherlands in 2016, an international interlaboratory ring trial for the ToxTracker assay was organized. The purpose of the trial was to establish the transferability and reproducibility of the assay. The goals of the interlaboratory trial were (i) to evaluate the accuracy of ToxTracker to predict *in vivo* genotoxicity and (ii) to validate the application of the mechanistic information that is provided by ToxTracker. Genotoxicity was defined as giving a positive *in vivo* result in the transgenic rodent (TgR), the PigA mutation, the MN, the CA and/or the Comet assays. A good intra- and inter-laboratory reproducibility is essential for any regulatory applications of the assay. The accuracy to identify genotoxic compounds and identify their MoA will determine the impact and positioning of the assay in the standard strategy for genotoxicity testing.

20. The ToxTracker trial was organized according to OECD guidance document 34. The ring trial was organized by the validation management team (VMT) together with Toxys B.V. (The Netherlands). The VMT consisted of several recognized experts with established expertise in genetic toxicology and experience with interlaboratory validation studies for the OECD (Table 3). The VMT was responsible for defining the validation project structure, selection of the partner laboratories, setting the different milestones for the project and analysis of the test results. The VMT, excluding Toxys, was also responsible for the selection of compounds that would be included in the trial as well as for defining the test criteria and data acceptance. Toxys, as developer of the ToxTracker assay, was not involved in compound selection to prevent any selection bias or conflicts of interest.

Table 3: ToxTracker validation management team.

Validation Management Team	Affiliation	Country
David Kirkland	Kirkland consulting	UK
Philippe Vanparys	Gentoxicon	BE
Jan van Benthem	RIVM	NL
Els Adriaens	Adriaens consulting	BE
Giel Hendriks	Toxys	NL

21. The ToxTracker interlaboratory validation was performed by seven experienced laboratories from different industries (Table 4). All leading scientists in the ring trial are experienced in regulatory genotoxicity testing and many actively participate in scientific expert groups at the HESI GTTC and IWGT. The VMT regularly consulted with representatives from the participating laboratories about the progress of the ring trial. All important decisions about the validation study plan, study protocols, timelines, data analysis and acceptance were discussed and approved by the full ToxTracker validation team, including VMT and laboratories.

Table 4: ToxTracker ring trail laboratories

Partner	Industry	Country	Scientists involved
Pfizer	Pharma	US	Maik Schuler Maria Engel
Proctor & Gamble	Cosmetics	US	Stefan Pfuhler Ashley Allemang
GenenTech	Pharma	US	Tomomi Kiyota Jennifer Vogt Gabrielle Cole
Corteva agriscience	Agrochemicals	US	Raja Settivari Abby Myhre Stephanie Kellum
Roche	Pharma	CH	Andreas Zeller Valerie Naëssens
Labcorp	CRO	UK	Julie Clements Darren Kidd
Charles River Labs	CRO	CA	Annie Hamel Marise Roy Renato Cardoso

9. Compound selection

22. The selection of compounds for the interlaboratory validation trial was done by the VMT, excluding Toxys to prevent any conflict of interest or selection bias. The aim was to have a broad selection of compounds to cover as many chemical classes as reasonably possible with sufficient *in vitro* and *in vivo* genotoxicity data available. Compound selection was based on publicly available lists and databases (16,31,32). The procedure and considerations for selection of the compounds were similar as described for the JaCVAM international validation of the *in vivo* comet assay (33). The selected compounds can be divided into four groups: I) genotoxic carcinogens, II) genotoxic non-carcinogens, III) non-genotoxic carcinogens and IV) non-genotoxic non-carcinogens. The ToxTracker trial was solely focused on genotoxicity prediction and

carcinogenicity was not considered as a criterion for compound selection. The compound list consists of organic and inorganic, aromatic and aliphatic molecules to cover a broad chemical space. Also, a number of compounds were selected that require metabolic activation in the liver. A full list of the selected compounds and their genotoxicity classification can be found in Annex 2. Genotoxicity of the compounds has been previously established and is based on the WoE from the various *in vitro* and *in vivo* mutation and genotoxicity assays that are publicly available. To have sufficient statistical power in the study, in total 64 chemicals were selected. Each compound was tested independently in three laboratories. All compounds were coded and distributed exclusively by Els Adriaens. None of the participating laboratories were aware of the compound selection.

23. The ToxTracker interlaboratory validation project was divided into three phases. During the first phase, the ToxTracker assay was installed in the seven participating laboratories. The ToxTracker reporter cell lines, cell culture media and positive/negative control compounds were supplied by Toxys. Assay protocols were provided and an experienced scientist from Toxys visited each lab to set up the assay. During the lab training, proper culture of the reporter stem cell lines was ensured and the flow cytometer that was available at the laboratories was configured to run ToxTracker. A training set of 8 well-established genotoxic and non-genotoxic compounds was used to validate proper installation of the assay (Table 5). During a five-day training, each laboratory performed ToxTracker in three independent repeat experiments. The eight compounds were tested in the absence and presence of a rat liver extract (S9)-based metabolizing system. A template for ToxTracker data analysis (MS excel spreadsheet, see Annex 3) was shared by Toxys with the participating laboratories. An example data set from the installations in the laboratories can be found in Annex 4.

Table 5: Training set of compounds for ToxTracker installation

Compound	Cas number	Genotoxicity classification
Cisplatin	15663-27-1	Genotoxin, DNA cross-linker
Etoposide	33419-42-0	Genotoxin, Topo II-inhibitor
Aflatoxin B1	1162-65-8	Mutagen, requires S9 metabolization
Benzo[a]pyrene	50-32-8	Mutagen, requires S9 metabolization
Sodium arsenite	7784-46-5	Possible genotoxin, oxidative stress
Diethyl maleate	141-05-9	Non-genotoxin, oxidative stress
Tunicamycin	11089-65-9	Non-genotoxin, activation of UPR
Rosuvastatin	287714-41-4	Non-genotoxin

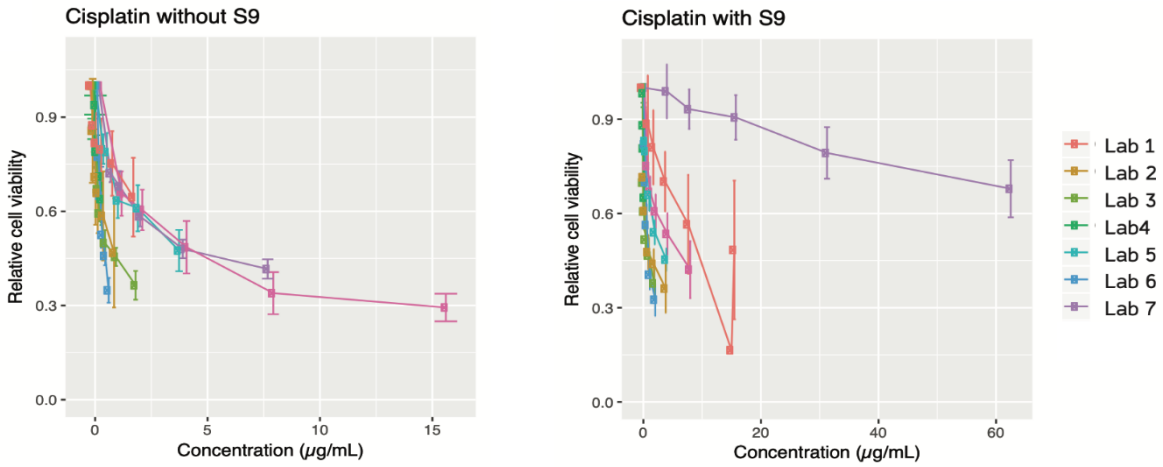
24. After the successful installation of ToxTracker, the participating labs were requested to run a limited proficiency test to show their ability to run the assay. For this, the VMT selected six compounds that were coded (Table 6). The laboratories received the blinded compounds and instructions for which solvent should be used. Each lab was requested to run a dose range finding experiment to determine which concentrations to apply in ToxTracker. The top concentration for each compound should not exceed the 75% cytotoxicity induction threshold. Next, each compound was tested in the full ToxTracker protocol in the absence and presence of S9-mix.

Table 6: Compound selection for ToxTracker proficiency testing

Compound	Cas number	Genotoxicity classification
Cisplatin	15663-27-1	Genotoxin, DNA cross-linker
Ethyl methanesulphonate	62-50-0	Genotoxin, DNA methylation agent
Benzo[a]pyrene	50-32-8	Mutagen, requires S9 metabolization
Ampicillin	69-53-4	Non-genotoxin, antibiotic
D-mannitol	69-65-8	Non-genotoxin
o-anthranilic acid	118-92-3	Non-genotoxin

25. All test results from the second phase (proficiency testing) of the validation project were collected and analyzed by the VMT. First, the labs determined the cytotoxicity of the six compounds. Relative cell survival was determined after 24h exposure by relative cell count (Figure 4A). The labs selected the top concentration to apply in ToxTracker that induced 50-75% cytotoxicity, as instructed by the assay protocol (Figure 4B). In most of the cases, the compound concentrations that were selected by the different labs were in the same order of magnitude. Differences in dose selection could often be explained by either the approaches used for definition of cytotoxicity (i.e. differences in cell count by flow cytometers) and different criteria for solubility of compounds. Next, the labs performed the ToxTracker assay for each of the six reporter cell lines and indicated whether a positive response was observed (Figure 4C), according to the prediction model for ToxTracker (Table 1 and 2). All the laboratories were able to perform the ToxTracker assay according to the assay protocol. One lab reported issues with solubility of one of the compounds. No other significant issues were observed. The VMT collected the results and calculated the between-lab reproducibility for each of the toxicological endpoints that are included in ToxTracker (Figure 4D).

A



B **Maximum concentration tested (µg/mL)**

Laboratory	Without S9	With S9
Lab 1	1.9	15.6
Lab 2	0.97	3.9
Lab 3	1.9	1.9
Lab 4	0.24	0.24
Lab 5	3.9	3.9
Lab 6	0.48	1.9
Lab 7	7.81	62

C **Summary GFP induction**

Without S9	Bscl2	Rtkn	Srxn1	Blvrb	Ddit3	Btg2	With S9	Bscl2	Rtkn	Srxn1	Blvrb	Ddit3	Btg2
Lab 1	P	P	P	P	N	P	Lab 1	P	P	N	N	N	P
Lab 2	P	P	P	P	N	P	Lab 2	P	P	P	P	N	P
Lab 3	P	P	P	P	N	P	Lab 3	P	P	P	P	N	P
Lab 4	P	P	P	P	N	P	Lab 4	P	P	P	N	N	P
Lab 5	P	P	P	N	N	P	Lab 5	P	P	P	N	N	P
Lab 6	P	P	P	P	N	P	Lab 6	P	P	P	P	N	P
Lab 7	P	P	P	P	N	P	Lab 7	P	P	N	N	N	N

D

	Without S9					With S9				
Reporters	Genotoxicity	Oxidative stress	Prot.	p53	Reporters	Genotoxicity	Oxidative stress	Prot.	p53
BLR	100%	100%	100%	100%	BLR	100%	75%	100%	87.5%

Figure 4: Example of ToxTracker results from the validation laboratories. A) Cytotoxicity of the compound was determined by relative cell count after 24h exposure of the ToxTracker stem cells. B) The top concentration for ToxTracker was selected based on cytotoxicity of the compound. C) Classification of the compound for the different ToxTracker reporters. A positive result (P) was recorded in the case of at least a 2-fold increase of GFP expression. GFP induction levels lower than 2-fold resulted in a negative score (N), according to the criteria defined in tables 1 and 2. D) Between-lab reproducibility for the different toxicological endpoint that are investigated in ToxTracker.

26. Following the laboratory proficiency testing during the second phase of the ring trial project, the VMT concluded that all laboratories were able to perform ToxTracker according to the assay protocol. In this blinded study, all laboratories correctly identified ethyl methanesulphonate (EMS) and cisplatin (CIS) as genotoxic compounds (Table 8). Benzo[a]pyrene (BAP) when tested in the presence of S9 rat liver extract, was classified as a genotoxic compound in 6 of the 7 labs as expected. Lab 4 experienced solubility issues with BAP (likely related to prolonged storage) and therefore tested a lower concentration than the other labs which resulted in a non-genotoxic classification. Ampicillin, mannitol and anthranilic acid were correctly classified as non-genotoxic by all 7 labs. The overall between-lab reproducibility of ToxTracker in the proficiency test for the prediction of genotoxicity by the seven labs was 100% in absence of S9-mix and 97,9% in presence of S9-mix.

27. All six compounds were also analyzed for their genotoxic MoA (see Annex 5 for an overview of the results). CIS and EMS also induced the reporters for oxidative stress (Srxn1/Blvrb) and p53 activation (Btg2) but not the protein damage reporter (Ddit3). Together with the observed induction of both genotoxicity reporters (Bscl2/Rtkn), the ToxTracker results indicate that CIS and EMS are genotoxins that directly bind to the DNA. BAP activated all ToxTracker reporters indicating that the compound is a genotoxin with a broader MoA. The three tested non-genotoxic compounds did not induce any of the ToxTracker reporters.

Table 8: Genotoxicity classification of the phase 2 validation compounds.

	Bscl2 / Rtkn Without S9							Bscl2 / Rtkn With S9					
Laboratory	AMP	MAN	ANT	EMS	BaP	CIS	**Laboratory**	AMP	MAN	ANT	EMS	BaP	CIS
Lab 1	N	N	N	P	N	P	Lab 1	N	N	N	P	P	P
Lab 2	N	N	N	P	N	P	Lab 2	N	N	N	P	P	P
Lab 3	N	N	N	P	N	P	Lab 3	N	N	N	P	P	P
Lab 4	N	N	N	P	N	P	Lab 4	N	N	N	P	N	P
Lab 5	N	N	N	P	N	P	Lab 5	N	N	N	P	P	P
Lab 6	N	N	N	P	N	P	Lab 6	N	N	N	P	P	P
Lab 7	N	N	N	P	N	P	Lab 7	N	N	N	P	P	P
BLR (%)	100	100	100	100	100	100	**BLR (%)**	100	100	100	100	87.5	100
Overall (%)	**100**						**Overall (%)**	**97.9**					

N: Negative, GFP induction at all tested concentrations was <2 for the Bscl2 and Rtkn reporter cell lines.
P: Positive, GFP induction for at least one tested concenntration was ≥2 for the Bscl2 and/or Rtkn cell lines

28. After the ToxTracker installation and training of the laboratories, a number of modifications were made to the ToxTracker protocol by the VMT (See annex 1). Specific instructions were added about how to deal with poorly soluble compounds.

Warming and sonication of poorly soluble compounds was added to the test protocol. In addition, specific instructions how to assess precipitation of the compounds was added. Precipitation of the test material should be addressed in the cell culture plates by microscopy at the end of exposure. Also, further instructions were added about quality controls for culture of the reporter cell lines, minimum induction levels of the GFP reporter cell lines for the positive control and criteria for a positive test result. Instructions for potentially required additional repeat experiments were added. Finally, a checklist for the scientists to make sure that all the crucial steps in the ToxTracker protocol were followed was added. The final ToxTracker validation protocol was approved by the VMT and all participating laboratories.

29. In the third phase of the ToxTracker ring trial, the seven laboratories received 24 or 30 coded compounds. Each compound was tested in three independent repeat experiments in the absence and presence of S9-mix. The laboratories received only instructions on which solvents to apply for preparing stock solutions. They therefore first needed to run a dose range finding experiment to determine the top concentration that should be tested in ToxTracker. The highest concentration should induce 50-75% cytotoxicity. In case of precipitation of the compound, the maximum soluble concentration would be applied. For non-cytotoxic compounds, the maximum concentration was set at 1 mg/ml. The mean GFP expression of the six ToxTracker reporters as well as cell concentrations was determined by flow cytometry and the results were collected in a standard data analysis template (Annex 3). From this, the induction levels of the fluorescent reporters, as well as cytotoxicity of the compounds, were calculated. During the validation experiments, some technical support was provided by the VMT. Most of the questions were related to data acceptance and the requirement to perform a fourth repeat experiment. In two instances, a videoconference was organized with the VMT and all the participating laboratories to provide feedback and guidance on data acceptance.

30. The ToxTracker ring trial was performed between 2017 and 2022. Timelines for the validation project are summarized in Table 9. Assembly of the VMT and onboarding of the participating laboratories was completed in 2017. Technical training of the laboratories was completed in 2018. In 2018, all laboratories performed the second phase of the project, the proficiency testing. After approval of all phase 2 results by the VMT and establishment of the final validation protocol by the full ToxTracker consortium, the phase 3 validation was started in 2018. The timelines for completion of the experiments varied significantly between the laboratories. The Covid pandemic that started at the end of 2019 also had an impact on the performance of the validation experiments. When the labs completed their experiments, the test results were submitted to Els Adriaens for data analysis and compilation of the validation results. All test results were submitted in spring 2021 and analysis of the data was performed by the VMT. Compounds remained coded throughout the data analysis.

Table 9: Timeline of the interlaboratory validation ring trial of ToxTracker.

	Activity	Date	Status
1.	Assemble validation team	2017	✓
2.	Technical training of labs (phase 1)	2017-2018	✓
3.	Establish validation protocol	2018	✓
4.	Proficiency testing of labs (phase 2)	2018	✓
5.	Establish final ToxTracker protocol	2018	✓
6.	ToxTracker validation 24/30 compounds (phase 3)	2018-2021	✓
7.	Data analysis	2021	✓
8.	Draft validation report	2022	✓
9.	Review by OECD	2023	

10. Validation data analysis

31. Analysis of the ToxTracker validation trial results was performed by the VMT. All test results were compiled into a large database using the R programming language for statistical computing and graphics (34). Throughout the analysis, all compounds remained coded, except for the assay positive control compounds. During this analysis, the VMT first focused on data acceptance. The VMT first verified if the positive controls that were included in every experiment met the acceptance criteria, meaning did the control compounds induce the fluorescent reporters above the minimum threshold as set in the protocol at acceptable cytotoxicity levels (See protocol in Annex 2)? In some cases, one of the reporters did not meet all the acceptance criteria, e.g. cytotoxicity of the compounds was higher than the cut-off of 75% (Figure 3A) or activation of one of the genotoxicity reporters did not meet the minimal induction level (2-fold or 3-fold increase in GFP for the Bscl2-GFP or Rtkn-GFP respectively) (Figure 3B). In those cases, the VMT assessed if there was sufficient evidence that all cell lines were performing correctly, that the positive control compounds were active and if S9 metabolization was sufficient. In those cases, the VMT used their expert judgement to accept or reject the controls. In the data analysis database, these experiments were marked as “acceptable with restrictions”. In case the positive controls in a certain experiment did not meet the acceptance criteria, all the results from the test compounds in that experiment were discarded.

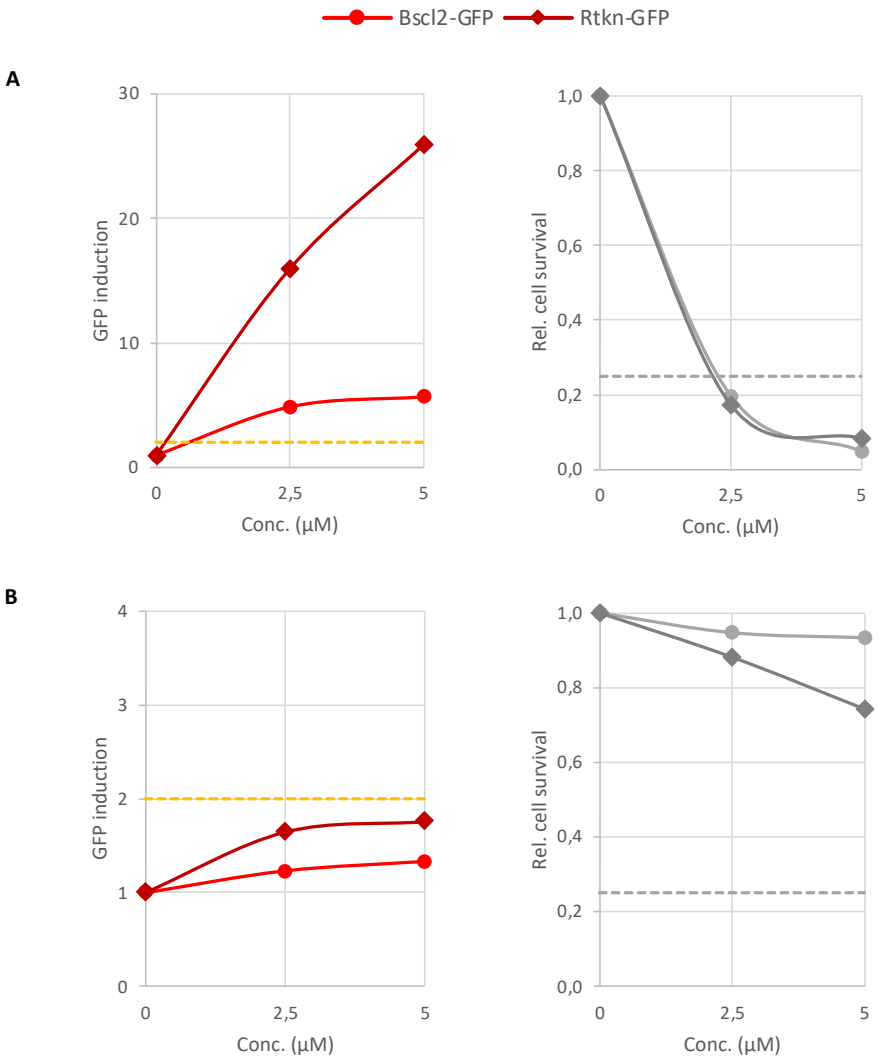


Figure 3: Examples of non-acceptable ToxTracker results. ToxTracker cells were exposed to the positive control compound cisplatin for 24h and activation of the Bscl2-GFP and Rtkn-GFP reporters as well as cytotoxicity was determined. **(A)** The ToxTracker reporters were activated above the 2-fold induction cut-off (left) but only at non-acceptable cytotoxicity levels (75%) (right). **(B)** Cytotoxicity levels were acceptable (right) but none of the ToxTracker reporters were activated above the 2-fold induction threshold.

32. From every participating laboratory, the ToxTracker results were analyzed for their acceptability according to the data acceptance criteria, with inclusion of expert judgment on occasions as discussed above. Inductions of the six fluorescent reporters following exposure to the test compounds as well as the relative cell survival was calculated. Figure 4 illustrates this for one of the test compounds. From each experiment, reporter activation was classified as positive, negative or inconclusive according to the ToxTracker prediction model (Table 1). Next, the classifications from the three independent repeat tests were weighted into an overall classification according to the prediction model (Table 2). Compounds were classified for their

properties to induce genotoxicity (Bslc2/Rtkn), oxidative stress (Srxn1/Blvrb), protein damage (Ddit3) and p53 activation (Btg2). In case no clear overall classification could be made, compounds were classified as equivocal. If the results did not meet the data acceptance criteria, even after expert judgement, the compound was classified as inconclusive. For every compound, which was tested in each of three independent laboratories, the classifications were compared, and an overall classification was made based on a weighted calculation. After compiling the overview of the test results, the data analysis was approved by the VMT and the full validation consortium. After this approval, the compounds were decoded in May 2022.

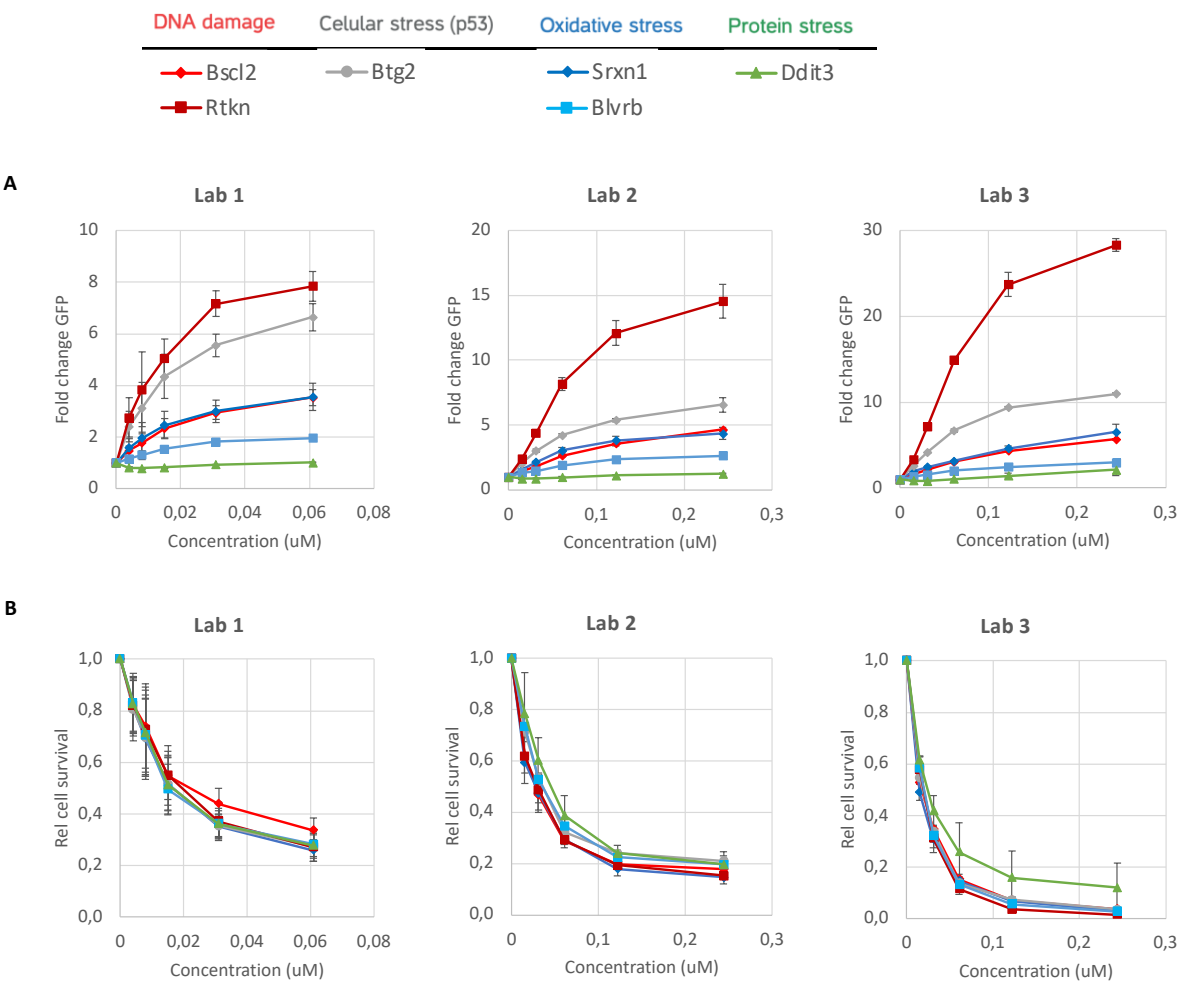


Figure 4: ToxTracker results for compound 1 from three independent laboratories. **(A)** Activation of the six different ToxTracker reporters following exposure to increasing concentration of the test compound. The Bscl2 and Rtkn reporters indicate genotoxicity, Srxn1 and Blvrb are induced by oxidative stress, Btg2 is associated with the p53 tumor suppressor response and Ddit3 is induced by protein misfolding. **(B)** Cytotoxicity of the compound is determined by relative cell count in cultures exposed to the compound and their related vehicle control cultures.

33. First, the data from the seven validation laboratories were analyzed for their genotoxicity predictions. To this end, we first focused on the genotoxicity reporters (Bscl2/Rtkn) in ToxTracker and compared their responses to the expected genotoxicity classification of the 64 compounds that were included in this validation trial. The test compounds in this study were previously categorized as genotoxic or non-genotoxic by different expert committees and working groups based on the WoE from the standard battery of *in vitro* (Ames, MN, CA) and *in vivo* (TGR, MN) assays. Compounds were classified as genotoxic in ToxTracker when either one or both Bscl2-GFP and Rtkn-GFP reporters were induced above the 2-fold increase in GFP threshold. The Bscl2-GFP reporter indicated formation of bulky DNA adducts and the Rtkn-GFP reporter is activated upon formation of DNA double strand breaks. For 25 of the 32 expected genotoxic compounds, there was full concordance between the ToxTracker validation results and the genotoxicity classification from the standard genotoxicity testing battery. For three compounds (1,2-dimethylhydrazine, benzo[a]pyrene and 2,6-diaminotoluene), two labs reported a positive classification, but one lab classified the compounds as non-genotoxic. The negative result for 1,2-dimethylhydrazine in one of the labs was likely caused by a difference in concentration selection. The lab selected an 8-fold lower concentration to test in ToxTracker than the other two laboratories (31.5 µM instead if 250 µM) from the dose range finding experiment. The negative result for benzo[a]pyrene in one laboratory was probably related to inadequate metabolization by S9. There was also no cytotoxicity reported for the compound in presence of S9-mix, in contrast to the other laboratories. However, the positive control compound Aflatoxin B1 for S9 metabolization did result in the expected ToxTracker activation. Also 2,6-diaminotoluene was tested at a 15-fold lower concentration in one of the laboratories, likely causing the negative ToxTracker result (Table 10).

34. Four of the expected genotoxic compounds were classified as non-genotoxic in ToxTracker by all validation laboratories. Acrylonitrile was reported positive in the Ames mutation assay and showed mixed results in other *in vitro* genotoxicity assays (35). However, the *in vivo* MN and CA assays were negative for acrylonitrile and no DNA adduct could be detected following *in vivo* exposure. The carcinogenicity of acrylonitrile was suggested to be related to epigenetic mechanisms (36). Benzene is a very potent human carcinogen and *in vivo* mutagen (37,38). However, benzene is generally negative in the standard *in vitro* genotoxicity assays. Some benzene metabolites do induce MN or CA (39). Also, the oxidative stress and induction of oxidative DNA lesions has been reported. In general, *in vitro* metabolization by S9 does not support the genotoxic effects of benzene. The lack of *in vitro* metabolization of benzene is the likely cause for the negative ToxTracker result. Cadmium chloride induces MN and CA *in vitro* and *in vivo* (40). Oxidative stress was reported to be an important mechanism for the genotoxicity of cadmium chloride. The ToxTracker validation laboratories classified cadmium chloride as non-genotoxic but did observe significant levels of oxidative stress. Finally, NDMA was classified as non-genotoxic whereas the nitrosamine compound is a very potent mutagen *in vivo*. *In vitro* metabolization of nitrosamines is poorly supported by S9 rat liver extract. NDMA was positive in the Ames mutation assay as well as the *in vitro* MN at concentration above 25 mM (41) In the ToxTracker validation, the maximum concentration that was tested

by the laboratories was set at 1 mg/ml, thereby limiting the NDMA exposures to non-cytotoxic concentrations. NDMA has previously been classified as genotoxic in ToxTracker when tested up to 25 mM in the presence of hamster S9-mix.

Table 10: Genotoxicity classification, meaning activation of the Bscl2 or/and Rtkn genotoxicity reporters, of 32 *in vivo* genotoxic compounds in the ToxTracker trial.

										Final prediction	Weighted calculation	
	Compound		Lab 1	Lab 2	Lab 3	Lab 4	Lab 5	Lab 6	Lab 7	Overall	Pos	Neg
1	Etoposide	33419-42-0	P	P	P					P	1,00	
2	Mitomycin C	50-07-7	P		P	P				P	1,00	
3	Cisplatin	15663-27-1	P		P			P		P	1,00	
4	1,2- Dimethylhydrazine	306-37-6	P	P				N		P	0,67	0,33
5	1,2-dibromoethane	106-93-4				I	P	P		P	1,00	
6	Cyclophosphamide	6055-19-2	P		P	I				P	1,00	
7	2-Acetylaminofluorene	53-96-3				P	P		P	P	1,00	
8	Azidothymidine	30516-87-1				I	P		P	P	1,00	
9	ENU	759-73-9		P				P	P	P	1,00	
10	Acrylonitrile	107-13-1				N	N		N	N		1,00
11	Benzene	71-43-2		I	I		N			N		1,00
12	4,4' -Oxydianiline	101-80-4		P				P	P	P	1,00	
13	Busulfan	55-98-1	P			P		P		P	1,00	
14	Ethyl methanesulfonate	62-50-0	P		I	P				P	1,00	
15	p-Chloroaniline	106-47-8				I	P	P		P	1,00	
16	7,12-Dimethyl-benzanthracene	57-97-6		P	P		I			P	1,00	
17	Benzo[a]pyrene	50-32-8		P	P		N			P	0,67	0,33
18	Cadmium Chloride	10108-64-2		N				N	N	N		1,00
19	Dimethylnitrosamine	62-75-9	N		I				N	N		1,00
20	2,4-Diaminotoluene	95-80-7				I	P	P		P	1,00	
21	o-Anisidine	90-04-0	P		I				P	P	1,00	
22	4-nitroquinoline-1-oxide	56-57-5			P			P	P	P	1,00	
23	6-Mercaptopurine	50-44-2	P		I	I				P	1,00	
24	Cytosine arabinose	147-94-4		P				P	P	P	1,00	
25	p-Phenylenediamine 2HCl	624-18-0	P				P	P		P	1,00	
26	8-Hydroxyquinoline	148-24-3			P	P		P		P	1,00	
27	9-Aminoacridine	90-45-9	P	P				P		P	1,00	
28	2,6-Diaminotoluene	823-40-5				N	P		P	P	0,67	0,33
29	3-Nitropropionic acid	504-88-1	P		P				P	P	1,00	
30	p-Anisidine	104-94-9		P	I				P	P	1,00	
31	5-fluorouracil	51-21-8	P	P		P				P	1,00	
32	Phenol	108-95-2	P				P	P		P	1,00	

35. From the 32 expected non-genotoxic compounds that were included in the validation trial, none were classified overall as genotoxic (Table 11), although for a number of compounds, a positive genotoxicity result was reporter by one laboratory.

Table 11: Genotoxicity classification, meaning activation of the Bscl2 or/and Rtkn genotoxicity reporters of 32 *in vivo* non-genotoxic compounds in the ToxTracker trial.

										Final prediction	Weighted calculation	
	Compound		Lab 1	Lab 2	Lab 3	Lab 4	Lab 5	Lab 6	Lab 7	Overall	Pos	Neg
33	Di(2-ethylhexyl)phthalate	117-81-7				I	N		N	N		1,00
34	Lead (ii) acetate	6080-56-4	P	N			I			E	0,50	0,50
35	2-Phenylphenol sodium salt	6152-33-6		N	I		I			N		1,00
36	Ropinirole hydrochloride	91374-20-8		I				N	N	N		1,00
37	Methyl carbamate	598-55-0		N				N	N	N		1,00
38	Cyclosporin A	59865-13-3	N		I	I				N		1,00
39	Sodium saccharin	128-44-9		N	N		N			N		1,00
40	Diethanolamine	111-42-2	N		N			N		N		1,00
41	Hexachloroethane	67-72-1			N			N	N	N		1,00
42	Melamine	108-78-1	P			I		N		E	0,50	0,50
43	Tunicamycin	11089-65-9	N					N	N	N		1,00
44	p-Nitrophenol	100-02-7		N				P	N	N	0,33	0,67
45	Phenanthrene	85-01-8	N	N	N					N		1,00
46	Tertiarybutylhydroquinone	1948-33-0	P			N		N		N	0,33	0,67
47	Benzyl alcohol	100-51-6		N	I		N			N		1,00
48	Vanilin	121-33-5	P		N				N	N	0,33	0,67
49	Erythromycin stearate	114-07-8	N					N	P	N	0,33	0,67
50	Sodium diclofenac	15307-79-6			I			P	N	N	0,50	0,50
51	o-anthranilic acid	118-92-3	N		N				N	N		1,00
52	Tolbutamide	64-77-7	N		N	I				N		1,00
53	2-ethyl-1,3-hexanediol	94-96-2	N	N	I					N		1,00
54	Chlorpheniramine maleate	113-92-8	N	N				N		N		1,00
55	Ampicillin trihydrate	7177-48-2				I	N		N	N		1,00
56	Sodium chloride	7647-14-5				N	N		N	N		1,00
57	D-mannitol	69-65-8	N		I	I				N		1,00
58	Allyl alcohol	107-18-6				I	N	N		N		1,00
59	(2-chloroethyl)trimethyl-NH3Cl	999-81-5				I	N	N		N		1,00
60	Sulfisoxazole	127-69-5				I	N	N		N		1,00
61	Sucrose	57-50-1				I	N		N	N		1,00
62	Cyclohexanone	108-94-1	N		N	I				N		1,00
63	1-Nitropropane	108-03-2		N	N		N			N		1,00
64	Phenformin HCl	834-28-6		N				N	N	N		1,00

In most cases (lead acetate, tert-butyl hydroquinone, vanillin, erythromycin stearate and diclofenac), the lab only observed activation of the Rtkn-GFP reporter which indicates the formation of DNA strand breaks but not the Bscl2-GFP reporter for the formation of bulky DNA lesions and DNA replication inhibition. This pattern of reporter activation is typically observed for compounds that are indirect genotoxins, often secondary to the induction of oxidative stress. Indeed, for lead acetate and tert-butyl hydroquinone, indirect genotoxicity due to oxidative stress has been reported (42)(43). It is therefore interesting and relevant that, for all of these Rtkn-GFP reporter positive compounds, the laboratories reported activation of the Srxn1 and Blvrb reporters for oxidative stress. P-nitrophenol activated both Bscl2 and Rtkn genotoxicity reporters in one laboratory, indicating direct DNA reactivity, but this result could not be confirmed by the other laboratories. P-nitrophenol is negative in the standard battery of *in vitro* genotoxity assays, but there are various reports of positive CA and MN tests *in vivo* (16). Also, melamine, a non-genotoxic compound, was classified as genotoxic by one laboratory but was negative in the other two labs.

36. From the weighted calculations, the overall sensitivity and specificity of ToxTracker to identify (*in vivo*) genotoxic compounds was calculated. The calculations were done for 59 compounds for which acceptable data was collected from at least 2 labs. Compounds for which data was only available from 1 laboratory were excluded from the calculations. In this validation trial, the ToxTracker assay correctly identified genotoxic compounds with a sensitivity of 87% (26 of 30 expected positives) and a specificity of 90% (26 of 29 expected negatives). The accuracy of identifying genotoxic compounds in this validation study was very much in line with previous validation reports by Toxys (8).

11. Genotoxic mode-of-action assessment in ToxTracker

37. The second main objective of this interlaboratory validation study was to investigate the added value of the MoA information that is provided by ToxTracker for the genotoxicity prediction of compounds, and this was a meaningful objective since, as described above, it was shown that ToxTracker could reliably discriminate between genotoxic and non-genotoxic substances. ToxTracker is proposed as an expansion of the *in vitro* genotoxicity test battery to provide insight into the MoA of genotoxic compounds. Especially for compounds for which conflicting results have been reported from the various *in vitro* genotoxicity assays or between the *in vitro* and *in vivo* tests, insight into the MoA of compounds can help to explain the differences and to better classify compounds (13,44). The ToxTracker assay combines six different reporters to investigate the induction of DNA damage, oxidative stress and protein damage. Activation of the Bscl2 and/or Rtkn reporters indicate genotoxicity, Srxn1 and Blvrb activation shows induction of oxidative stress, Ddit3 is associated with protein unfolding and Btg3 activation is linked to the p53-associated cellular stress response. By assessing the differential induction of these reporters, the assay can provide insight into the MoA of genotoxic compounds. It is therefore essential to include the full panel of six ToxTracker reporters for genotoxicity prediction and MoA assessment. In the validation trial, induction of all six reporters was determined by the laboratories following exposure to the 32 genotoxic and 32 non-genotoxic compounds. A summary of the results is shown in Table 12. Every compound was tested in three laboratories. The results from the different labs for every reporter were combined into an overall classification using a weighted calculation. An overview of all the test results can be found in Annex 6.

Table 12: Summary of the combined ToxTracker reporter activations from the different validation laboratories in the ring trial.

Code	Compound	Cas#	Bscl2 -S9	Bscl2 +S9	Rtkn -S9	Rtkn +S9	Overall	Srxn1 -S9	Srxn1 +S9	Blvrb -S9	Blvrb +S9	Overall	Ddit3 -S9	Ddit3 +S9	Overall	Btg2 -S9	Btg2 +S9	Overall
1	Etoposide	33419-42-0	P	P	P	P	P	P	P	N	N	P	N	N	N	P	P	P
2	Mitomycin C	50-07-7	P	P	P	P	P	P	P	P	P	P	N	N	N	P	P	P
3	Cisplatin	15663-27-1	P	P	P	P	P	P	P	P	P	P	N	N	N	P	P	P
4	1,2- Dimethylhydrazine	306-37-6	P	P	P	P	P	P	P	P	P	P	N	N	N	P	P	P
5	1,2-dibromoethane	106-93-4	N	N	P	P	P	P	P	P	P	P	N	N	N	N	N	N
6	Cyclophosphamide	6055-19-2	N	P	N	P	P	N	P	N	P	P	N	N	N	N	P	P
7	2-Acetylaminofluorene	53-96-3	N	N	P	P	P	P	P	N	N	P	N	N	N	N	N	N
8	Azidothymidine	30516-87-1	P	P	P	P	P	P	P	P	P	P	N	N	N	P	P	P
9	ENU	759-73-9	P	P	P	P	P	P	P	P	P	P	N	N	N	P	P	P
10	Acrylonitrile	107-13-1	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
11	Benzene	71-43-2	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
12	4,4' -Oxydianiline	101-80-4	P	P	P	P	P	P	P	P	P	P	N	N	N	P	P	P
13	Busulfan	55-98-1	P	P	P	P	P	P	P	P	P	P	N	N	N	P	P	P
14	Ethyl methanesulfonate	62-50-0	P	P	P	P	P	P	P	P	P	P	N	N	N	P	P	P
15	p-Chloroaniline	106-47-8	N	N	P	P	P	N	P	N	P	P	N	N	N	N	P	P
16	7,12-Dimethyl-benzanthracene	57-97-6	N	N	N	P	P	N	N	N	P	P	N	N	P	N	P	P
17	Benzo[a]pyrene	50-32-8	N	P	N	P	P	N	P	N	P	P	N	P	P	N	P	P
18	Cadmium Chloride	10108-64-2	N	N	N	N	N	P	P	P	P	P	P	P	P	N	N	P
19	Dimethylnitrosamine	62-75-9	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
20	2,4-Diaminotoluene	95-80-7	P	P	P	P	P	P	N	P	P	P	P	P	P	P	P	P
21	o-Anisidine	90-04-0	N	N	P	P	P	P	P	P	P	P	N	N	N	P	P	P
22	4-nitroquinoline-1-oxide	56-57-5	P	P	P	P	P	P	N	P	N	P	N	N	N	P	P	P
23	6-Mercaptopurine	50-44-2	N	N	P	P	P	P	P	N	N	P	N	N	N	P	P	P
24	Cytosine arabinose	147-94-4	P	P	P	P	P	N	N	P	P	P	N	N	N	P	P	P
25	p-Phenylenediamine 2HCl	624-18-0	P	P	P	P	P	P	P	P	P	P	N	N	N	P	P	P
26	8-Hydroxyquinoline	148-24-3	N	N	P	P	P	P	P	N	P	P	P	P	P	P	P	P
27	9-Aminoacridine	90-45-9	N	N	P	P	P	P	P	N	N	P	N	N	N	P	P	P
28	2,6-Diaminotoluene	823-40-5	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P
29	3-Nitropropionic acid	504-88-1	P	P	P	P	P	P	P	P	P	P	N	N	N	P	P	P
30	p-Anisidine	104-94-9	P	P	P	P	P	P	P	P	P	P	P	N	P	P	P	P
31	5-fluorouracil	51-21-8	P	P	P	P	P	P	P	P	P	P	N	N	N	P	P	P
32	Phenol	108-95-2	N	N	P	P	P	P	P	P	P	P	N	N	N	P	P	P
33	Di(2-ethylhexyl)phthalate	117-81-7	N	N	N	N	N	N	N	N	N	N	P	N	P	N	N	N
34	Lead acetate	6080-56-4	N	N	N	N	N	P	P	P	P	P	N	N	N	N	N	N
35	2-Phenylphenol	6152-33-6	N	N	N	N	N	P	P	N	P	P	P	P	P	N	N	P
36	Ropinirole hydrochloride	91374-20-8	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
37	Methyl carbamate	598-55-0	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
38	Cyclosporin A	59865-13-3	N	N	N	N	N	P	P	P	P	P	P	P	P	N	N	N
39	Sodium saccharin	128-44-9	N	N	N	N	N	P	P	P	P	P	N	N	N	N	N	N
40	Diethanolamine	111-42-2	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
41	Hexachloroethane	67-72-1	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
42	Melamine	108-78-1	N	N	N	N	N	N	N	N	N	P	N	N	N	N	N	N
43	Tunicamycin	11089-65-9	N	N	N	N	N	N	N	N	N	N	P	P	P	N	N	N
44	p-Nitrophenol	100-02-7	N	N	N	N	N	N	N	N	N	N	P	P	P	N	N	N
45	Phenanthrene	85-01-8	N	N	N	N	N	P	P	P	P	P	P	P	P	N	N	N
46	Tertiarybutylhydroquinone	1948-33-0	N	N	N	N	N	P	P	P	P	P	P	P	P	P	P	P
47	Benzyl alcohol	100-51-6	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
48	Vanilin	121-33-5	N	N	N	N	N	P	P	N	P	P	N	N	N	N	N	P
49	Erythromycin	114-07-8	N	N	N	N	N	P	P	N	N	P	P	N	P	N	N	N
50	Sodium diclofenac	15307-79-6	N	N	N	N	N	P	P	N	N	P	P	P	P	P	E	P
51	o-anthranilic acid	118-92-3	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
52	Tolbutamide	64-77-7	N	N	N	N	N	P	P	N	N	P	P	P	P	N	N	N
53	2-ethyl-1,3-hexanediol	94-96-2	N	N	N	N	N	P	P	N	N	P	N	N	N	N	N	N
54	Chlorpheniramine maleate	113-92-8	N	N	N	N	N	P	P	P	P	P	N	N	N	N	N	N
55	Ampicillin trihydrate	7177-48-2	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
56	Sodium chloride	7647-14-5	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
57	D-mannitol	69-65-8	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
58	Allyl alcohol	107-18-6	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
59	Chlormequat chloride	999-81-5	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
60	Sulfisoxazole	127-69-5	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
61	Sucrose	57-50-1	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
62	Cyclohexanone	108-94-1	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
63	1-Nitropropane	108-03-2	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
64	Phenformin HCl	834-28-6	N	N	N	N	N	N	N	N	N	N	P	P	P	N	N	N

38. In general, for the genotoxic compounds with an expected direct DNA damaging MoA, both Bscl2 and Rtkn genotoxicity reporters were activated, e.g. etoposide, mitomycin C, cisplatin, 5-fluorouracil. Activation of the Bscl2 reporter indicates the formation of DNA adducts that inhibit DNA replication. Bscl2 activation generally correlates with a positive result in the standard *in vitro* bacterial (Ames) and mammalian cell gene mutation assays (MLA/HPRT) and indicates a mutagenic MoA. The Rtkn reporter is activated upon the formation of DNA double strand breaks and indicates a clastogenic MoA. Rtkn activation correlates strongly with a positive response in the *in vitro* and *in vivo* MN and CA clastogenicity assays. Compounds that require metabolization in the liver activated the genotoxicity reporters only in the presence of S9-mix, e.g. cyclophosphamide, benzo[a]pyrene and 7,12-Dimethyl-benzanthracene. For most of the genotoxic compounds, activation of the p53 tumor suppressor-associated Btg2 reporter and occasionally the oxidative stress reporters Srxn1 and Blvrb was also

observed. However, from the dose response graphs for the genotoxic compounds, induction of the genotoxicity reporters is clearly the primary response (example for etoposide shown in Figure 3). For a number of compounds, only activation of the Rtkn, but not the Bscl2 genotoxicity reporter was observed. This is often observed for compounds causing indirect genotoxic effects, including aneugens or oxidative stress-inducing compounds. For example, the major mechanistic pathway for the genotoxicity of 1,2-dibromoethane is through binding to the cellular antioxidant GSH (45). Also, for 8-hydroquinoline, the primary genotoxic MoA was reported to occur through induction of oxidative stress (46). Accordingly, 1,2-dibromoethane and 8-hydroquinoline primarily activated the Srxn1 oxidative stress reporter as well as the Rtkn reporter for clastogenic DNA lesions, suggesting that genotoxicity (clastogenicity) of these compounds is caused by oxidative stress (Figure 5).

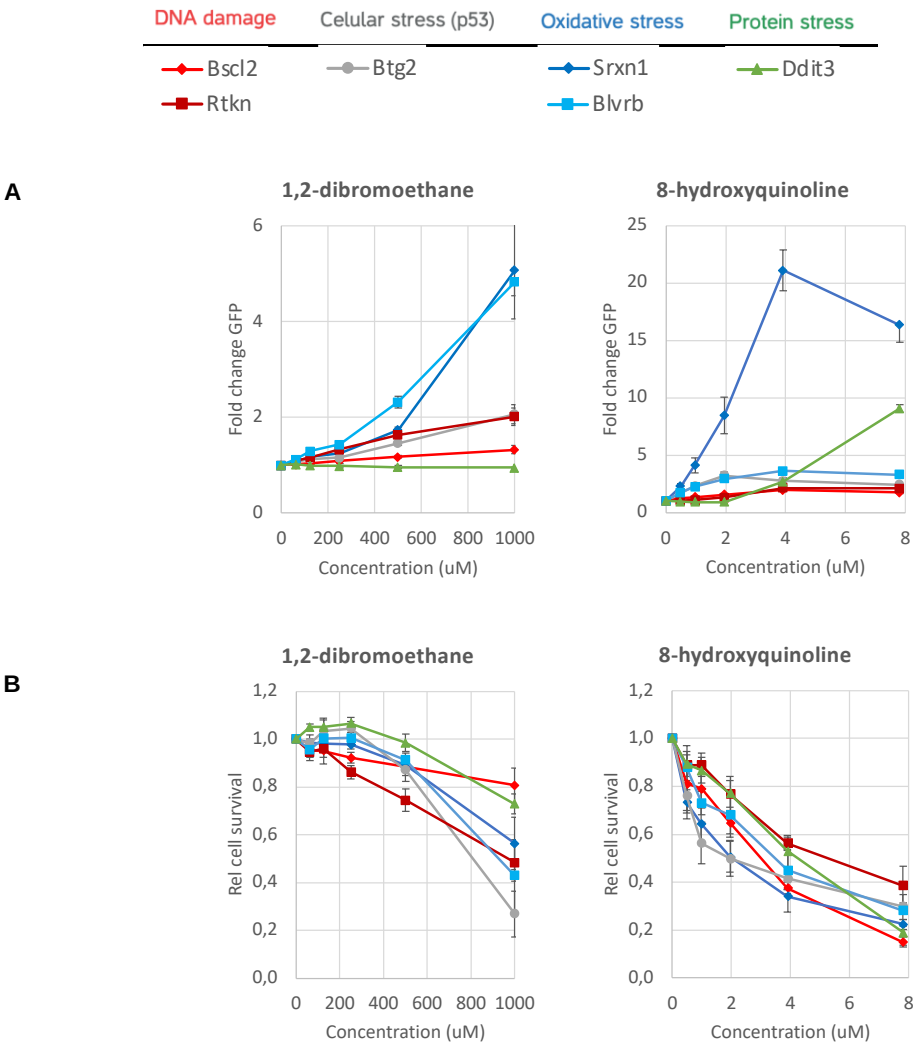
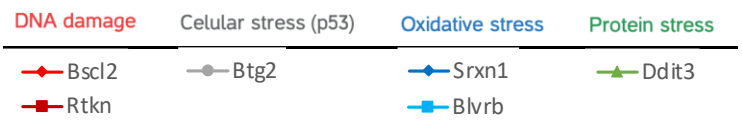


Figure 5: ToxTracker results for 1,2-dibromoethane and 8-hydroxyquinoline. **(A)** Activation of the six different ToxTracker reporters following exposure to increasing concentration of the test compound. The Bscl2 and Rtkn reporters indicate genotoxicity, Srxn1 and Blvrb are induced by oxidative stress, Btg2 is associated with the p53 tumor suppressor response and Ddit3 is induced by protein misfolding. **(B)** Cytotoxicity of the compounds is determined by relative cell count in cultures exposed to the compound and their related vehicle control cultures.

39. Overall, all tested non-genotoxic compounds in the validation trial were correctly predicted as non-genotoxic in ToxTracker. In a weighted approach, none of these compounds induced the Bscl2 and Rtkn genotoxicity reporters (Table 12). However, approximately 50% of the tested non-genotoxic compounds induced oxidative stress or protein unfolding which have been associated with carcinogenicity (15,47). All compounds that were classified as non-genotoxic in ToxTracker were negative in the Ames bacterial mutation assay. However, a number of compounds that were predicted non-genotoxic in ToxTracker (no activation of Bscl/Rtkn reporters) have been reported to induce positive results in the *in vitro* MN or CA assay (16,18). Occasionally, positive results from these *in vitro* clastogenicity assays do not correctly predict *in vivo* genotoxicity. Various reasons for this discrepancy have been proposed, including misleading *in vitro* positive responses caused by high levels of cytotoxicity (48). Also, indirect genotoxicity caused by high levels of oxidative stress can cause positive results in the *in vitro* MN assay but are often not observed *in vivo* due to lower *in vivo* exposure levels and more efficient anti-oxidant systems (48). For example, lead acetate was shown to induce DNA strand breaks *in vitro* due to oxidative stress and also tert-butyl hydroquinone gave positive results in the *in vitro* CA assay (43,49). In this ToxTracker validation trial, the different laboratories reported high levels of oxidative stress for these compounds. Based on the dose response curves for the genotoxicity reporters (Bscl2/Rtkn) and oxidative stress (Srxn1/Blvrb), oxidative stress induction appeared to be the primary mechanism of toxicity for these compounds (Figure 5). In the validation trial, lead acetate and tert-butyl hydroquinone were classified as non-genotoxic although 1 lab reported a positive result for the Rtkn reporter for DNA strand breaks after exposure to lead acetate and tert-butyl hydroquinone. Also, induction of protein damage has been associated with induction of genotoxic effects, primarily *in vitro* (48). In the ToxTracker trial, tunicamycin and p-nitrophenol were classified as non-genotoxic but both compounds activated the Ddit3 reporter for protein unfolding. Tunicamycin is non-genotoxic *in vivo*, but induced MN *in vitro*, p-nitrophenol in positive in the *in vitro* CA and MN assays. Together these examples indicate that information about the genotoxic and non-genotoxic effects of compounds can be valuable to more accurately predict the *in vivo* genotoxic effects of compounds. The genotoxic MoA information can also be used to explain discrepancies between various *in vitro* and *in vivo* genotoxicity assays in a WoE approach.



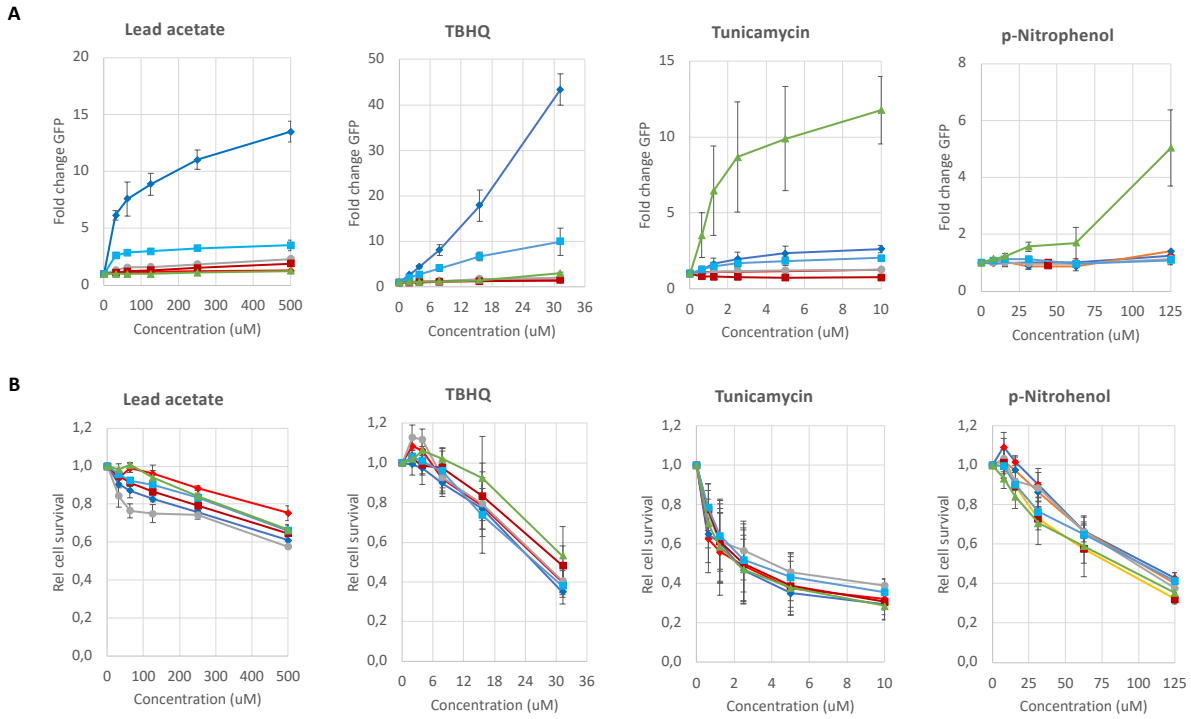


Figure 5: ToxTracker results for the non-genotoxic compounds lead acetate, tert-butyl hydroquinone, tunicamycin and p-nitrophenol. **(A)** Activation of the six different ToxTracker reporters following exposure to increasing concentration of the test compound. The Bscl2 and Rtkn reporters indicate genotoxicity, Srxn1 and Blvrb are induced by oxidative stress, Btg2 is associated with the p53 tumor suppressor response and Ddit3 is induced by protein misfolding. **(B)** Cytotoxicity of the compounds is determined by relative cell count in cultures exposed to the compound and their related vehicle control cultures.

40. To further explore the contribution of the different reporter genes in the genotoxicity prediction, the ToxTracker results were compared to outcomes of the standard *in vitro* and *in vivo* genotoxicity assays (Table 13). For this comparison, a number of compounds were selected for which data are available in the public domain that allow a WoE-based classification as compounds with MoAs involving either oxidative stress (blue) or unfolded protein response (green). For all ten compounds that were expected to induce oxidative stress, activation of the oxidative stress reporters (Srxn1/Blvrb) was observed by the laboratories. Four compounds that have been shown to induce the unfolded protein response, activated the Ddit3 reporter in ToxTracker. Importantly, nearly all of the selected compounds were predicted to be non-genotoxic in ToxTracker and were also negative in the standard *in vivo* genotoxicity assays. In contrast, many of these compounds were classified as genotoxic in at least one of the standard *in vitro* genotoxicity assays (Ames, MN, CA). The MoA information that is obtained from the ToxTracker assay could help to gain mechanistic insight into the hazardous properties of compounds and to improve the *in vivo* genotoxicity prediction.

Table 13: Comparison between ToxTracker and the standard in vitro and in vivo genotoxicity assays for compounds with an oxidative stress or protein reactive MoA.

	Compound	CAS number	In vitro				In vivo			ToxTracker	
			Ames	MLA	MN	CA	MN	CA	TgR	Genotoxic	MoA
Group I: Genotoxic carcinogens											
18	Cadmium Chloride	10108-64-2	E		P	P	P	P		N	Oxidative stress
22	4-nitroquinoline-1-oxide	56-57-5	P	P	P	P	P	P	P	P	DNA reactive, oxidative stress
Group II: Genotoxic non-carcinogens											
25	p-Phenylenediamine 2HCl	624-18-0	P	P	P	P	N			P	DNA reactive, oxidative stress
26	8-Hydroxyquinoline	148-24-3	P			P	N			P	Indirect genotoxin, oxidative stress, protein reactive
32	Phenol	108-95-2	N		P	P	P	N		P	Indirect genotoxin, oxidative stress
Group III: Non-genotoxic carcinogens											
34	Lead (ii) acetate trihydrage	6080-56-4	N		P	E	E	P		E	Oxidative stress
35	2-Phenylphenol sodium salt	6152-33-6	P			E	N	N		N	Oxidative stress, protein reactive
38	Cyclosporin A (CsA)	59865-13-3	N				N			N	Protein reactive
Group IV: Non-genotoxic non-carcinogens											
43	Tunicamycin	11089-65-9	N		P		N			N	Protein reactive
44	p-Nitrophenol (4-nitrophenol)	100-02-7	N		E	P	N			N	Protein reactive
45	Phenanthrene	85-01-8	P		E	E				N	Oxidative stress, protein reactive
46	Tertiarybutylhydroquinone	1948-33-0	N		P	P	N	N		N	Oxidative stress, protein reactive
54	Chlorpheniramine maleate	113-92-8	N	P		P	N			N	Oxidative stress
58	Allyl alcohol	107-18-6	P	P		P	N			N	Oxidative stress

12. Within-lab and between-lab reproducibility for genotoxicity predictions

41. One of the primary objectives of the ToxTracker ring trial was to establish the transferability and reproducibility of the assay. We first focused on the reproducibility of the genotoxicity prediction in ToxTracker. For each of the participating laboratories, the WLR was determined. Every compound was tested in three independent repeat experiments for activation of the Bscl2-GFP and Rtkn-GFP genotoxicity reporters. For each of these biomarkers, the results from the repeat experiments were analyzed for their acceptability according to the criteria set in the ToxTracker protocol, with expert judgement where appropriate. Next, from every acceptable experiment, the positive or negative classifications for the different reporters were compared (Table 14 provides an example for one of the laboratories). The experiments were considered reproducible if the laboratory came to the same conclusion in the three independent repeat tests. For the example shown in Table 14, the reproducibility was 96.7% for the genotoxicity classification. For some compounds, e.g. phenol, lead acetate and tert-butyl hydroquinone, the three repeat experiments gave slightly different results but the classification of the compounds was identical between the experiments.

Table 14: Example of the WLR of the two genotoxicity reporters in ToxTracker within one of the validation laboratories

			Genotoxicity			
			Bscl2	Bscl2	Rtkn	Rtkn
Code	Compound	Cas#	Result -S9	Result +S9	Result -S9	Result +S9
1	Etoposide	33419-42-0	+++	+++	+++	+++
2	Mitomycin C	50-07-7	+++	+++	+++	+++
3	Cisplatin	15663-27-1	+++	+++	+++	+++
4	1,2- Dimethylhydrazine	306-37-6	+++	+++	+++	+++
6	Cyclophosphamide	6055-19-2	---	+++	---	+++
13	Busulfan	55-98-1	+++	+++	+++	+++
14	Ethyl methanesulfonate	62-50-0	+++	+++	+++	+++
19	Dimethylnitrosamine	62-75-9	---	---	---	---
21	o-Anisidine	90-04-0	---	+++	+++	+++
23	6-Mercaptopurine	50-44-2	---	(+)-(+)	+++	+-+
25	p-Phenylenediamine 2HCl	624-18-0	+++	+++	+++	+++
27	9-Aminoacridine	90-45-9	---	---	+++	+++
29	3-Nitropropionic acid	504-88-1	+++	+++	+++	+++
31	5-fluorouracil	51-21-8	-++	-++	-++	+++
32	Phenol	108-95-2	---	-(+)-	-++	+++
34	Lead (ii) acetate	6080-56-4	-(+)(+)	---	(+)++	---
38	Cyclosporin A	59865-13-3	---	---	---	---
40	Diethanolamine	111-42-2	---	---	---	---
42	Melamine	108-78-1	+++	+++	+++	+++
43	Tunicamycin	11089-65-9	---	---	---	---
45	Phenanthrene	85-01-8	---	---	---	---
46	Tertiarybutylhydroquinone	1948-33-0	(+)-(+)	(+)(+)-	(+)+-	+++
48	Vanilin	121-33-5	---	---	---	+++
49	Erythromycin stearate	114-07-8	---	---	---	---
51	o-anthranilic acid	118-92-3	---	---	---	---
52	Tolbutamide	64-77-7	---	---	---	---
53	2-ethyl-1,3-hexanediol	94-96-2	---	---	---	---
54	Chlorpheniramine maleate	113-92-8	---	---	---	---
57	D-mannitol	69-65-8	---	---	---	---
62	Cyclohexanone	108-94-1	---	---	---	---

42. For every laboratory, the WLR was calculated. We determined the WLR for the genotoxicity classification of the compounds, meaning the activation of the Bscl2 and Rtkn genotoxicity reporters was assessed in the absence and presence of S9-mix to classify the compounds as genotoxic. The reproducibility of genotoxicity classification for every compound was determined between repeat experiments (Table 14). The number of compounds that were tested in the different laboratories varied between 24 and 30. Results were considered reproducible if the laboratory gave the compound the same classification in the three repeat experiments. Non-reproducible (in Table 15) means that at least in one experiment, the laboratory came to a different classification of the compounds. Overall, the WLR in 6 of the 7 participating laboratories varied between 80% and 96.7%. However, Laboratory 4 had some challenges with accurate measurements of cell numbers which resulted in an overestimation of the cytotoxicity

of compounds. Following the data acceptance criteria for ToxTracker, data obtained at concentrations that induce >75% cytotoxicity should be discarded from analysis. The cytotoxicity assessment issues in lab 4 therefore resulted in a relatively high number of inconclusive tests which led to lack of reproducibility and negatively impacted their WLR calculations.

Table 15: Within-laboratory reproducibility for the Bscl2 and Rtkn genotoxicity reporters in the seven ToxTracker validation laboratories.

Lab	Tested compounds	Reproducible	Non-reproducible	WLR
1	30	29	1	96,7%
2	24	22	2	91,7%
3	25	23	2	92,0%
4	26	19	7	73,1%
5	24	20	4	83,3%
6	30	24	6	80,0%
7	27	22	5	81,5%

43. Finally, also the BLR in the ToxTracker ring trial for the prediction of genotoxicity was determined. For this, the classifications of the tested chemicals for their induction of the Bscl2 and Rtkn genotoxicity reporters were compared between the participating laboratories. For each lab, an overall classification was made from the three repeat experiments. These overall classifications were compared between laboratories to establish the BLR. As an example, the ToxTracker results for etoposide from the three laboratories are summarized in Table 16. In this example, all three labs classified etoposide as genotoxic compound in all three biological repeat experiments. Results were reproducible with each lab as well as between labs.

Table 16: Example of the between-laboratory reproducibility for the activation of the different genotoxicity reporters in ToxTracker following exposure to etoposide (compound 1).

	Bscl2							
	-S9	+S9	Final call -S9	Final call +S9	Overall Final call	No. Valid runs	WLR	BLR
LAB 1	PPP	PPP	P	P	P	3	Yes	
LAB 2	PPP	NNN	P	N	P	3	Yes	Yes
LAB 3	PPP	P(P)(P)	P	P	P	3	Yes	

	Rtkn							
	-S9	+S9	Final call -S9	Final call +S9	Overall Final call	No. Valid runs	WLR	BLR
LAB 1	PPP	PPP	P	P	P	3	Yes	
LAB 2	PPP	PPN	P	P	P	3	Yes	Yes
LAB 3	PPP	PPP	P	P	P	3	Yes	

44. For the genotoxicity BLR calculation, the results for the Bscl2 and Rtkn reporters were combined as a positive call for each of these biomarkers would lead to a positive

genotoxicity classification for a compound. The BLR was determined for 59 compounds in the ring trial for which acceptable data from at least two laboratories were available. Compounds for which only acceptable data was obtained from one laboratory were excluded from the BLR calculations. The BLR for the genotoxicity predictions in the ToxTracker ring trial between the seven validation laboratories was 83.1%.

13. WLR and BLR for MoA assessment from all ToxTracker reporters

45. There are currently no consolidated databases available with compounds that specifically induce oxidative stress or the unfolded protein response. It is therefore not possible to calculate the sensitivity and specificity of the ToxTracker reporters for prediction of oxidative stress (Srxn1/Blvrb) or protein damage (Ddit3). We therefore focused on the transferability and reproducibility of all six ToxTracker reporters by the different laboratories. All compounds were tested for activation of the biomarkers that indicate induction of DNA damage, oxidative stress, protein damage and p53-associated cellular stress. For each of these biomarkers, the results from the repeat experiments were analyzed and the positive or negative classifications for each of the reporters were compared (Table 17 provides an example for one of the laboratories). The experiments were considered reproducible if the laboratory came to the same conclusion in the different repeats. For the example shown in Table 17, the reproducibility varied between 96.7% for induction of DNA damage (activation of Bscl2 and/or Rtkn), protein damage (Ddit3) and p53-associated cellular stress (Btg2), and 100% for oxidative stress (Srxn1 and/or Blvrb). For some compounds, the three repeat experiments gave slightly different results, but the classification of the compounds was identical between the experiments. For example, cyclophosphamide was negative in two repeats for induction of oxidative stress in the absence of S9-mix, but in the third repeat a weak positive (+) was recorded. Nevertheless, the overall call for oxidative stress was negative according to the ToxTracker prediction model (Table 2 and 3). In case that induction of a reporter was different between repeats, but the results -S9 and +S9 together resulted in similar calls in the three repeats, the results were indicated as reproducible.

Table 17: Example of the intra-lab reproducibility of the six different ToxTracker reporters within one of the validation laboratories

			Genotoxicity				Oxidative stress				Protein damage		Cellular stress	
			Bscl2	Bscl2	Rtkn	Rtkn	Srxn1	Srxn1	Blvrb	Blvrb	Ddit	Ddit	Btg2	Btg2
Code	Compound	Cas#	Result -S9	Result +S9	Result -S9	Result +S9	Result -S9	Result +S9	Result -S9	Result +S9	Result -S9	Result +S9	Result -S9	Result +S9
1	Etoposide	33419-42-0	+++	+++	+++	+++	+++	+++	(+)(+)(+)	--(+)	---	---	+++	+++
2	Mitomycin C	50-07-7	+++	+++	+++	+++	+++	+++	+++	+++	---	---	+++	+++
3	Cisplatin	15663-27-1	+++	+++	+++	+++	+++	+++	+++	+++	---	---	+++	+++
4	1,2- Dimethylhydrazine	306-37-6	+++	+++	+++	+++	+++	+++	+++	+++	---	---	+++	+++
6	Cyclophosphamide	6055-19-2	---	+++	---	+++	--(+)	+++	---	+++	---	---	---	+++
13	Busulfan (Myleran)	55-98-1	+++	+++	+++	+++	+++	+++	+++	+++	---	---	+++	(+)++
14	Ethyl methanesulfonate (EMS)	62-50-0	+++	+++	+++	+++	+++	+++	+++	+++	---	---	+++	+++
19	Dimethylnitrosamine (N-nitrosodimethy	62-75-9	---	---	---	---	---	---	---	---	---	---	---	---
21	o-Anisidine	90-04-0	---	+++	+++	+++	+++	+++	+++	+++	---	---	+++	+++
23	6-Mercaptopurine	50-44-2	---	(+)-(+)	+++	+-+	+++	+(+)+	-(+)(+)	---	---	---	+++	+++
25	p-Phenylenediamine 2HCl	624-18-0	+++	+++	+++	+++	+++	+++	+++	+++	---	---	+++	+++
27	9-Aminoacridine	90-45-9	---	---	+++	+++	+ + +	+ + +	- - -	- - -	---	---	+++	+++
29	3-Nitropropionic acid	504-88-1	+++	+++	+++	+++	+++	+++	+++	+++	---	---	+++	+++
31	5-fluorouracil	51-21-8	-++	-++	-++	+++	-++	+++	-+-	-+(+)	---	---	-++	+++
32	Phenol	108-95-2	---	-(+)-	-++	+++	+++	+++	---	+++	---	---	+-+	+++
34	Lead (ii) acetate trihydrage available	6080-56-4	-(+)(+)	---	(+)++	---	+++	+++	+++	+--	---	---	+++	---
38	Cyclosporin A (CsA)	59865-13-3	---	---	---	---	+++	+++	(+)++	+++	+++	+++	---	---
40	Diethanolamine (DEA or DEOA)	111-42-2	---	---	---	---	---	---	---	---	---	---	---	---
42	Melamine	108-78-1	+++	+++	+++	+++	---	---	---	---	---	---	---	---
43	Tunicamycin	11089-65-9	---	---	---	---	+++	+-+	+(+)+	+(+)-	+++	+++	---	---
45	Phenanthrene	85-01-8	---	---	---	---	+-+	+++	---	(+)-+	--+	+++	---	---
46	Tertiarybutylhydroquinone	1948-33-0	(+)-(+)	(+)(+)-	(+)+-	+++	+++	+++	+++	+++	++ +	+- +	+++	+++
48	Vanilin	121-33-5	---	---	---	+++	+-+	--(+)	---	+++	---	---	+--	---
49	Erythromycin (Erythromycin stearate 6-	114-07-8	---	---	---	---	++-	+++	-(+)-	(+)--	+++	---	---	---
51	o-anthranilic acid	118-92-3	---	---	---	---	---	---	---	---	---	---	---	---
52	Tolbutamide	64-77-7	---	---	---	---	+++	+++	---	---	+++	++-	---	---
53	2-ethyl-1,3-hexanediol	94-96-2	---	---	---	---	+++	+++	---	---	---	+--	---	---
54	Chlorpheniramine maleate	113-92-8	---	---	---	---	+++	+++	+++	-++	---	---	---	---
57	D-mannitol	69-65-8	---	---	---	---	---	---	---	---	---	---	---	---
62	Cyclohexanone	108-94-1	---	---	---	---	---	---	---	---	---	---	---	---

46. In addition to reproducibility of the genotoxicity prediction (discussed above), WLR was also determined for the 4 other reporter genes following the same criteria (Table 18). The WLR was comparable to that seen for Bscl2 and Rtkn for 6 of the 7 labs, with Laboratory 4 again showing low WLR values due to the cytotoxicity issues discussed above. The overall WLR (average of all reporters) varies from 97.5% for the lab with the best overall performance and 71.1% for the lab with the lowest reproducibility. Together, these WLR calculations, together with the successful proficiency tests during the second phase of the validation trial confirm the excellent transferability of the ToxTracker assay.

Table 18: Within-laboratory reproducibility for the various ToxTracker reporters in the different validation laboratories.

	Lab 1	Lab 2	Lab 3	Lab 4	Lab 5	Lab 6	Lab 7
DNA damage	96,7	91,7	92,0	73,1	83,3	80,0	81,5
Oxidative stress	100,0	75,0	73,3	77,8	91,7	70,0	81,5
Protein damage	96,7	91,7	87,5	74,1	100,0	86,7	75,0
p53 activation	96,7	70,8	82,6	59,3	87,5	83,3	81,5

47. Finally, also the BLR in the ToxTracker validation trial for all reporters was determined. For this, the classifications of the tested chemicals for their induction of DNA damage, oxidative stress, protein damage and p53-associated cellular stress were compared

between the participating laboratories. For each lab, an overall classification was made for every endpoint in ToxTracker from the three repeat experiments. These overall classifications were compared between laboratories to establish the BLR. As an example, the ToxTracker results for etoposide from the three laboratories are summarized in Table 19. Five out of six ToxTracker reporters gave a similar result in the three laboratories. Only the result for the Blvrb reporter was not reproducible in this example.

Table 19: Example of the between-laboratory reproducibility for the activation of the different ToxTracker reporters following exposure to etoposide.

	Bscl2							
	-S9	+S9	Final call -S9	Final call +S9	Overall Final call	No. Valid runs	WLR	BLR
LAB 1	PPP	PPP	P	P	P	3	Yes	
LAB 2	PPP	NNN	P	N	P	3	Yes	Yes
LAB 3	PPP	P(P)(P)	P	P	P	3	Yes	
	Rtkn							
	-S9	+S9	Final call -S9	Final call +S9	Overall Final call	No. Valid runs	WLR	BLR
LAB 1	PPP	PPP	P	P	P	3	Yes	
LAB 2	PPP	PPN	P	P	P	3	Yes	Yes
LAB 3	PPP	PPP	P	P	P	3	Yes	
	Srxn1							
	-S9	+S9	Final call -S9	Final call +S9	Overall Final call	No. Valid runs	WLR	BLR
LAB 1	PPP	PPP	P	P	P	3	Yes	
LAB 2	PPP	N(P)N	P	N	P	3	Yes	Yes
LAB 3	PPP	NNP	P	N	P	3	Yes	
	Blvrb							
	-S9	+S9	Final call -S9	Final call +S9	Overall Final call	No. Valid runs	WLR	BLR
LAB 1	(P)(P)(P)	NN(P)	E	N	E	3	Yes	
LAB 2	PNN	IPN	N	E	E	3	No	No
LAB 3	NNN	NNN	N	N	N	3	Yes	
	Ddit							
	-S9	+S9	Final call -S9	Final call +S9	Overall Final call	No. Valid runs	WLR	BLR
LAB 1	NNN	NNN	N	N	3	N	Yes	
LAB 2	NNN	NNN	N	N	3	N	Yes	Yes
LAB 3	NNN	NNN	N	N	3	N	Yes	
	Btg2							
	-S9	+S9	Final call -S9	Final call +S9	Overall Final call	No. Valid runs	WLR	BLR
LAB 1	PPP	PPP	P	P	P	3	Yes	
LAB 2	PPP	NPP	P	P	P	3	Yes	Yes
LAB 3	PPP	PPP	P	P	P	3	Yes	

48. For the BLR calculation, the results for the Bscl2 and Rtkn reporters were combined since a positive call for either of these biomarkers would lead to a positive genotoxicity classification for a compound. The same approach was used for the Srxn1and Blvrb oxidative stress reporters. The BLR was determined for 59 compounds in the ring trial

for which acceptable data from at least two laboratories were available. The BLR was calculated for the different toxicological endpoints in ToxTracker. The BLR in the ToxTracker validation trial for the seven validation laboratories varied between 83% for the genotoxicity predictions and 71% for oxidative stress (Table 20). The overall reproducibility of the classification of genotoxic compounds was 91%. The overall reproducibility of predicting protein damage and p53-associated cellular stress was comparable for the genotoxic and non-genotoxic compounds.

Table 19: Between-lab reproducibility for the different toxicological endpoints that are assessed in ToxTracker.

	DNA damage	Oxidative stress	Protein damage	Cell stress
	Bscl2 / Rtkn	Srxn1 / Blvrb	Ddit3	Btg2
BLR	83,1	71,0	82,5	78,3

14. Availability of data for expert review

49. All results from the ToxTracker ring trial are available for review. The test results are collected in a large database, but this is also accessible through an excel spreadsheet. Full statistical analysis has been performed on the primary data sets. Publications about the development and technical validation of the ToxTracker assay have been published by Toxys and may be made available for expert review upon request. Various case studies using ToxTracker for genotoxicity testing and MoA assessment of pharmaceutical, (agro)chemical, cosmetic and flavor/fragrance compounds have been published in peer-reviewed journals.

15. Discussion and learnings

50. The interlaboratory validation of the ToxTracker assay was performed according to OECD guidance document 34 wherever reasonably possible. The project was coordinated by the experts in the VMT. Toxys, the developer of ToxTracker, provided technical support to the project, but was not involved in selection of the compounds for the validation trial or in data analysis. Also coding and distribution of the compounds to the participating labs was done by the VMT, excluding Toxys. The participating laboratories received the blinded compounds and instructions how to perform the tests. Data analysis was started when all the laboratories provided their results. Data analysis was performed by the VMT. First the results were analyzed for their acceptability, based on the data acceptance criteria that were defined by the VMT and described in the validation protocol. Next the acceptable results for the tested compounds were analyzed for the induction of genotoxicity. This was an important first step since information on MoA would only be valuable if it was shown that ToxTracker could reliably distinguish between genotoxic and non-genotoxic substances. The WLR

and BLR was determined, as well as the overall sensitivity and specificity for identification of genotoxic compounds. Finally, also the additional reporters that provide further insight into the MoA of genotoxic compounds were analyzed. Throughout the data analysis, compounds remained blinded to prevent any bias in the analysis. After sharing and discussing the results from the validation trial with the full validation consortium and approval of the data, the compounds were decoded.

51. During the training phases of the validation trial, a number of changes and clarifications were made to the ToxTracker protocol. The most important modification was a change in the S9 metabolization protocol. ToxTracker relies on S9 rat liver extract for metabolization of compounds. Originally, the standard S9 protocol that is also used in the Ames and in vitro MN assay was included in ToxTracker (7,8). In this protocol, cells are exposed to the compounds in presence of 1% S9-mix for 3 hours, followed by a 21 hour culture period without S9-mix before analysis of the fluorescent reporters. Exposure times are limited to 3 hours because of the potential toxicity of S9-mix. Although when using this S9 protocol, genotoxic compounds are effectively metabolized and correctly identified as genotoxic by ToxTracker, the recovery time after exposure resulted in a strong reduction in signals for oxidative stress and protein damage induction. To improve the sensitivity for detection of all cellular responses in ToxTracker, the S9 protocol for ToxTracker was optimized by Toxys. In this improved protocol, cells are exposed to 0.25% S9-mix for 24 hours continuously and ToxTracker reporter activation is analyzed immediately after exposure without a recovery period. This improved protocol was also implemented in the ToxTracker validation protocol during the validation trial. The laboratories were requested to first test the genotoxic compound benzo[a]pyrene during the second phase of the validation (proficiency testing) using the original S9 protocol (Figure 6A). Benzo[a]pyrene was correctly classified as genotoxic by the laboratories. After adoption of the updated S9 protocol by the VMT, all labs were required to retest benzo[a]pyrene. With the new S9 protocol, the genotoxicity of benzo[a]pyrene was confirmed, and also oxidative stress induction was readily detected (figure 6B). The updated S9 protocol was therefore applied throughout the third phase of the validation trial.

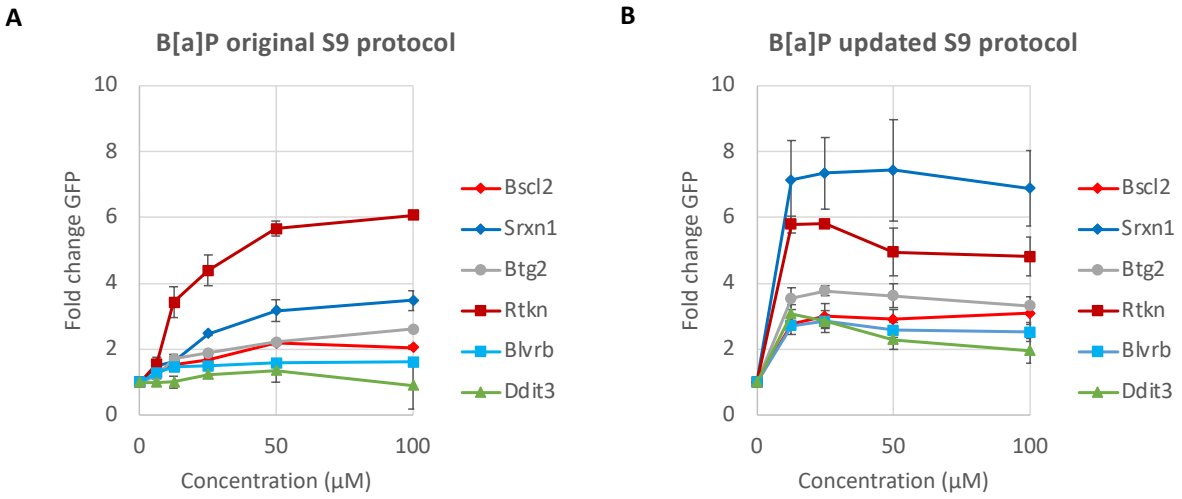


Figure 6: Genotoxicity of benzo[a]pyrene in ToxTracker in the original and updated S9 metabolization protocol **(A)** Activation of the six different ToxTracker reporters following exposure to increasing concentration of the test compound. The Bscl2 and Rtkn reporters indicate genotoxicity, Srxn1 and Blvrb are induced by oxidative stress, Btg2 is associated with the p53 tumor suppressor response and Ddit3 is induced by protein misfolding. **(B)** Cytotoxicity of the compounds is determined by relative cell count in cultures exposed to the compound and their related vehicle control cultures.

52. During the second phase of the validation project, a few aspects in the ToxTracker protocol needed clarification, mostly related to compound handling and stock solution preparation. Also, there were a number of questions about quality control and data acceptance. Therefore, after the second phase (proficiency testing), a number of instructions on preparing compound solution, establishing precipitation and quality controls for the assay were included in the ToxTracker protocol by the validation team. These instructions helped the laboratories to increase the reproducibility of their repeat experiments. The most important reason for variation between repeats in the ToxTracker validation was the accurate measurement of cytotoxicity which could unnecessarily invalidate experiments. In addition, clear instructions were provided how to make the compound stock solutions and how to handle compounds that did not dissolve. During the proficiency testing, it became apparent that laboratories had different approaches to establish precipitation of compounds in the cell culture plates. Clear instructions how to determine the maximum soluble concentrations were added to the protocol by the VMT.

53. After completion of the ToxTracker validation trial and analysis of the results, the VMT came to a number of conclusions and learnings for future improvements of ToxTracker. A recommendation for future applications of ToxTracker would be to improve the instructions and quality controls for relative cell counting using the flow cytometer. Another learning from the ring trial was that the positive control compound Aflatoxin B1, included to ensure proper activity of S9-mix, was not very stable and was a source for variation. A modification to the ToxTracker protocol to use

cyclophosphamide as an alternative positive control compound requiring metabolic activation has already been made.

16. Summary and conclusions

54. ToxTracker is a high content *in vitro* reporter assay which has been shown to be very useful for the accurate prediction of *in vivo* genotoxicity. By combining various reporters that indicate different toxicological effects relevant for genetic toxicology, ToxTracker has the advantage that it can provide insight into the mode of action of genotoxic and non-genotoxic chemicals. The regulatory need and applications for ToxTracker are outlined in this document. In order to investigate how ToxTracker may complement the standard battery of *in vitro* genotoxicity assays, a comprehensive interlaboratory validation trial of the ToxTracker assay was organized. The validation was conducted using the principles outlined in OECD GD 34. The primary objectives of the validation were to establish the transferability and reproducibility of the assay, and to confirm the ability of ToxTracker to correctly classify compounds as genotoxic. In addition, the reproducibility to predict the genotoxic MoA was investigated and how this information can be applied to improve in vitro prediction of in vivo genotoxicity. During the validation trial, ToxTracker was successfully installed in the laboratories of the seven partner laboratories. A limited proficiency test confirmed the ability of the laboratories to perform the assay. Although no major problems occurred, a number of improvements was made to the ToxTracker protocol. During the validation trial, the seven laboratories tested 64 chemicals (32 expected to be positive and 32 expected to be negative) that together cover a broad spectrum of chemical spaces. Also metabolic activation of chemicals by liver enzymes was included in the study. Each compound was tested in three laboratories. During the validation trial, we determined the within and between-lab reproducibility of ToxTracker. The WLR varied between 80% and 96.7% for 6/7 of the different laboratories confirming the good transferability of the assay when cytotoxicity is accurately assessed. One laboratory suffered from some technical issues, resulting in a lower WLR of 73.1%. The BLR for genotoxicity classification of chemicals was in the region of 83%. The interlaboratory validation confirmed the accuracy of ToxTracker to correctly identify genotoxic compounds with a sensitivity of 87% and a specificity of 90%. Together, from this validation trial we concluded that ToxTracker is a robust *in vitro* assay for the accurate prediction of *in vivo* genotoxicity. With information on the MoA of chemicals that is provided by the assay, ToxTracker would be a valuable addition to the battery of genotoxicity assays that is applied for regulatory applications.

17. References

1. ICH. International conference on harmonisation of technical requirements for registration of pharmaceuticals for human use ICH harmonised tripartite

guideline guidance on genotoxicity testing and data interpretation for pharmaceuticals intended for human use S2(R1). 2011.

2. https://echa.europa.eu/guidance-documents/guidance-on-information-requirements-and-chemical-safety-assessment. ECHA GD for chemical safety assessment.
3. Hardy A, Benford D, Halldorsson T, Jeger M, Knutsen HK, More S, et al. Clarification of some aspects related to genotoxicity assessment. EFSA Journal. 2017 Dec;15(12).
4. SCCS (Scientific Committee on Consumer Safety), SCCS Notes of Guidance for the Testing of Cosmetic Ingredients and their Safety Evaluation 11th revision, 30-31 March 2021, SCCS/1628/21. 2021.
5. Thomas R. The US Federal Tox21 Program: A strategic and operational plan for continued leadership. ALTEX. 2018;163–8.
6. How to use alternatives to animal testing to fulfil your information requirements for REACH registration [Internet]. 2016 [cited 2022 Sep 28]. Available from: https://echa.europa.eu/documents/10162/17250/practical_guide_how_to_use_alternatives_en.pdf/148b30c7-c186-463c-a898-522a888a4404
7. Hendriks G, Atallah M, Morolli B, Calléja F, Ras-Verloop N, Huijskens I, et al. The ToxTracker assay: Novel GFP reporter systems that provide mechanistic insight into the genotoxic properties of chemicals. Toxicological Sciences. 2012 Jan 1;125(1):285–98.
8. Hendriks G, Derr RS, Misovic B, Morolli B, Calléja FMGR, Vrieling H. The Extended ToxTracker Assay Discriminates Between Induction of DNA Damage, Oxidative Stress, and Protein Misfolding. Toxicological Sciences. 2016 Mar 1;150(1):190–203.
9. Hendriks G, Atallah M, Raamsman M, Morolli B, van der Putten H, Jaadar H, et al. Sensitive DsRed fluorescence-based reporter cell systems for genotoxicity and oxidative stress assessment. Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis. 2011 May 10;709–710:49–59.
10. Hendriks G, Derr RS, Misovic B, Morolli B, Calléja FMGR, Vrieling H. The Extended ToxTracker Assay Discriminates Between Induction of DNA Damage, Oxidative Stress, and Protein Misfolding. Toxicological Sciences. 2016;150(1):190–203.
11. Kirsch-Volders M, Vanhauwaert A, Eichenlaub-Ritter U, Decordier I. Indirect mechanisms of genotoxicity. Toxicol Lett. 2003 Apr;140–141:63–74.
12. Lynch AM, Eastmond D, Elhajouji A, Froetschl R, Kirsch-Volders M, Marchetti F, et al. Targets and mechanisms of chemically induced aneuploidy. Part 1 of the report of the 2017 IWGT workgroup on assessing the risk of aneugens for carcinogenesis and hereditary diseases. Mutat Res Genet Toxicol Environ Mutagen [Internet]. 2019 Nov 1 [cited 2022 Jul 13];847. Available from: https://pubmed.ncbi.nlm.nih.gov/31699346/
13. Brandsma I, Moelijker N, Derr R, Hendriks G. Aneugen Versus Clastogen Evaluation and Oxidative Stress-Related Mode-of-Action Assessment of Genotoxic Compounds Using the ToxTracker Reporter Assay. Toxicological Sciences. 2020 Sep 1;177(1):202–13.
14. Nohmi T. Thresholds of Genotoxic and Non-Genotoxic Carcinogens. Toxicol Res. 2018 Oct 15;34(4):281–90.

15. Hendriks G, van de Water B, Schoonen W, Vrieling H. Cellular-signaling pathways unveil the carcinogenic potential of chemicals. Journal of Applied Toxicology. 2013;33(6):399–409.

16. Kirkland D, Kasper P, Martus HJ, Müller L, van Benthem J, Madia F, et al. Updated recommended lists of genotoxic and non-genotoxic chemicals for assessment of the performance of new or improved genotoxicity tests. Mutat Res Genet Toxicol Environ Mutagen [Internet]. 2016 Jan 1 [cited 2022 Jul 14];795:7–30. Available from: https://pubmed.ncbi.nlm.nih.gov/26774663/

17. Wills JW, Halkes-Wellstead E, Summers HD, Rees P, Johnson GE. Empirical comparison of genotoxic potency estimations: The in vitro DNA-damage ToxTracker endpoints versus the in vivo micronucleus assay. Mutagenesis. 2021 Jul 1;36(4):311–20.

18. Fowler P, Smith K, Young J, Jeffrey L, Kirkland D, Pfuhler S, et al. Reduction of misleading ("false") positive results in mammalian cell genotoxicity assays. I. Choice of cell type. Mutat Res [Internet]. 2012 Feb 18 [cited 2022 Jul 22];742(1–2):11–25. Available from: https://pubmed.ncbi.nlm.nih.gov/22138618/

19. Kakehashi A, Wei M, Fukushima S, Wanibuchi H. Oxidative Stress in the Carcinogenicity of Chemical Carcinogens. Cancers (Basel) [Internet]. 2013 Dec 1 [cited 2023 Mar 27];5(4):1332. Available from: /pmc/articles/PMC3875942/

20. Madden E, Logue SE, Healy SJ, Manie S, Samali A. The role of the unfolded protein response in cancer progression: From oncogenesis to chemoresistance. Biol Cell [Internet]. 2019 Jan 1 [cited 2023 Mar 27];111(1):1–17. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/boc.201800050

21. Allemang A, Lester C, Roth T, Pfuhler S, Peuschel H, Kosemund K, et al. Assessing the genotoxicity and carcinogenicity of 2-chloroethanol through structure activity relationships and in vitro testing approaches. Food and Chemical Toxicology. 2022 Oct;168:113290.

22. Smart DE, Bozhilova S, Miazzi F, Haswell LE, Gaca MD, Thorne D, et al. Application of ToxTracker for the toxicological assessment of tobacco and nicotine delivery products. Toxicol Lett [Internet]. 2022 Jan 19 [cited 2022 Jan 26]; Available from: http://www.ncbi.nlm.nih.gov/pubmed/35065211

23. Mesnage R, Ibragim M, Mandrioli D, Falcioni L, Tibaldi E, Belpoggi F, et al. Comparative toxicogenomics of glyphosate and Roundup herbicides by mammalian stem cell-based genotoxicity assays and molecular profiling in Sprague-Dawley rats. Toxicol Sci [Internet]. 2021 Nov 29 [cited 2022 Jan 26]; Available from: http://www.ncbi.nlm.nih.gov/pubmed/34850229

24. White PA, Long AS, Johnson GE. Quantitative Interpretation of Genetic Toxicity Dose-Response Data for Risk Assessment and Regulatory Decision-Making: Current Status and Emerging Priorities. Environ Mol Mutagen [Internet]. 2020 Jan 1 [cited 2022 Jul 14];61(1):66–83. Available from: https://pubmed.ncbi.nlm.nih.gov/31794061/

25. More S, Bampidis V, Benford D, Bragard C, Halldorsson T, Hougaard Bennekou S, et al. Statement on the derivation of Health-Based Guidance Values (HBGVs) for regulated products that are also nutrients. EFSA Journal. 2021 Mar;19(3).

26. White PA, Long AS, Johnson GE. Quantitative Interpretation of Genetic Toxicity Dose-Response Data for Risk Assessment and Regulatory Decision-Making: Current Status and Emerging Priorities. Environ Mol Mutagen. 2020 Jan 19;61(1):66–83.

27. Steiblen G, Benthem J van, Johnson G. Strategies in genotoxicology: Acceptance of innovative scientific methods in a regulatory context and from an industrial perspective. Mutation Research/Genetic Toxicology and Environmental Mutagenesis. 2020 May;853:503171.

28. Boisvert L, Derr R, Osterlund T, Hendriks G, Brandsma I. Quantitative Interpretation of ToxTracker Dose Response Data for Potency Comparisons and Mode-of-Action Determination. Environ Mol Mutagen. 2023 Jan 16;

29. Marc A. Beal, Marc Audebert, Tara Barton-Maclaren, Hannah Battaion, Jeffrey C. Bemis, Xuefei Cao, et al. Application of Quantitative in vitro to in vivo Extrapolation (IVIVE) Modeling to in vitro Genotoxicity Data Provides Conservative Estimates of in vivo Genotoxicity.

30. Boisvert L, Derr R, Osterlund T, Hendriks G, Brandsma I. Quantitative Interpretation of ToxTracker Dose Response Data for Potency Comparisons and Mode-of-Action Determination Environmental and Molecular Mutagenesis.

31. Kirkland D, Reeve L, Gatehouse D, Vanparys P. A core in vitro genotoxicity battery comprising the Ames test plus the in vitro micronucleus test is sufficient to detect rodent carcinogens and in vivo genotoxins. Mutation Research/Genetic Toxicology and Environmental Mutagenesis. 2011 Mar;721(1):27–73.

32. Madia F, Kirkland D, Morita T, White P, Asturiol D, Corvi R. EURL ECVAM Genotoxicity and Carcinogenicity Database of Substances Eliciting Negative Results in the Ames Test: Construction of the Database. Mutation Research/Genetic Toxicology and Environmental Mutagenesis. 2020 Jun;854–855:503199.

33. Morita T, Uno Y, Honma M, Kojima H, Hayashi M, Tice RR, et al. The JaCVAM international validation study on the in vivo comet assay: Selection of test chemicals. Mutation Research/Genetic Toxicology and Environmental Mutagenesis. 2015 Jul;786–788:14–44.

34. R Core Team. A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. [Internet]. 2022 [cited 2022 Oct 10]. Available from: https://www.R-project.org/

35. Whysner J, Ross PM, Conaway CC, Verna LK, Williams GM. Evaluation of possible genotoxic mechanisms for acrylonitrile tumorigenicity. Regul Toxicol Pharmacol [Internet]. 1998 [cited 2022 Jul 20];27(3):217–39. Available from: https://pubmed.ncbi.nlm.nih.gov/9693074/

36. Whysner J, Steward RE, Chen D, Conaway CC, Verna LK, Richie JP, et al. Formation of 8-oxodeoxyguanosine in brain DNA of rats exposed to acrylonitrile. Arch Toxicol [Internet]. 1998 [cited 2022 Jul 20];72(7):429–38. Available from: https://pubmed.ncbi.nlm.nih.gov/9708882/

37. Eastmond DA. Benzene-induced genotoxicity: a different perspective. J Toxicol Environ Health A [Internet]. 2000 [cited 2022 Jul 20];61(5–6):353–6. Available from: https://pubmed.ncbi.nlm.nih.gov/11086939/

38. Whysner J. Benzene-induced genotoxicity. J Toxicol Environ Health A. 2000;61(5–6):347–51.

39. IARC. Benzene. Benzene [Internet]. 2018 [cited 2022 Jul 20]; Available from: https://pubmed.ncbi.nlm.nih.gov/31769947/

40. Fahmy MA, Aly FAE. In vivo and In vitro Studies on the Genotoxicity of Cadmium Chloride in Mice. JOURNAL OF APPLIED TOXICOLOGY J Appl Toxicol. 2000;20:231–8.

41. Westphal GA, Bünger J, Schulz TG, Müller MM, Hallier E. Mutagenicity of N - nitrosodiethylamine in the Ames test with S. typhimurium TA1535 is due to volatile metabolites and is not dependent on cytochrome P4502E1 induction. Arch Toxicol. 2000 Dec 11;74(10):638–41.

42. Liu X, Wu J, Shi W, Shi W, Liu H, Wu X. Lead Induces Genotoxicity via Oxidative Stress and Promoter Methylation of DNA Repair Genes in Human Lymphoblastoid TK6 Cells. Med Sci Monit [Internet]. 2018 Jun 22 [cited 2022 Jul 21];24:4295. Available from: /pmc/articles/PMC6045917/

43. NTP. NTP website.

44. Allemang A, de Abrew KN, Shan YK, Krailler JM, Pfuhler S. A comparison of classical and 21st century genotoxicity tools: A proof of concept study of 18 chemicals comparing in vitro micronucleus, ToxTracker and genomics-based methods (TGx-DDI, whole genome clustering and connectivity mapping). Environ Mol Mutagen. 2021 Feb 1;62(2):92–107.

45. Cho SH, Guengerich FP. In vivo roles of conjugation with glutathione and O6-alkylguanine DNA-alkyltransferase in the mutagenicity of the bis-electrophiles 1,2-dibromoethane and 1,2,3,4-diepoxybutane in mice. Chem Res Toxicol [Internet]. 2013 Nov 18 [cited 2022 Jul 22];26(11):1765–74. Available from: https://pubmed.ncbi.nlm.nih.gov/24191644/

46. Barilli A, Atzeri C, Bassanetti I, Ingoglia F, Dall'Asta V, Bussolati O, et al. Oxidative Stress Induced by Copper and Iron Complexes with 8-Hydroxyquinoline Derivatives Causes Paraptotic Death of HeLa Cancer Cells. Mol Pharm. 2014 Apr 7;11(4):1151–63.

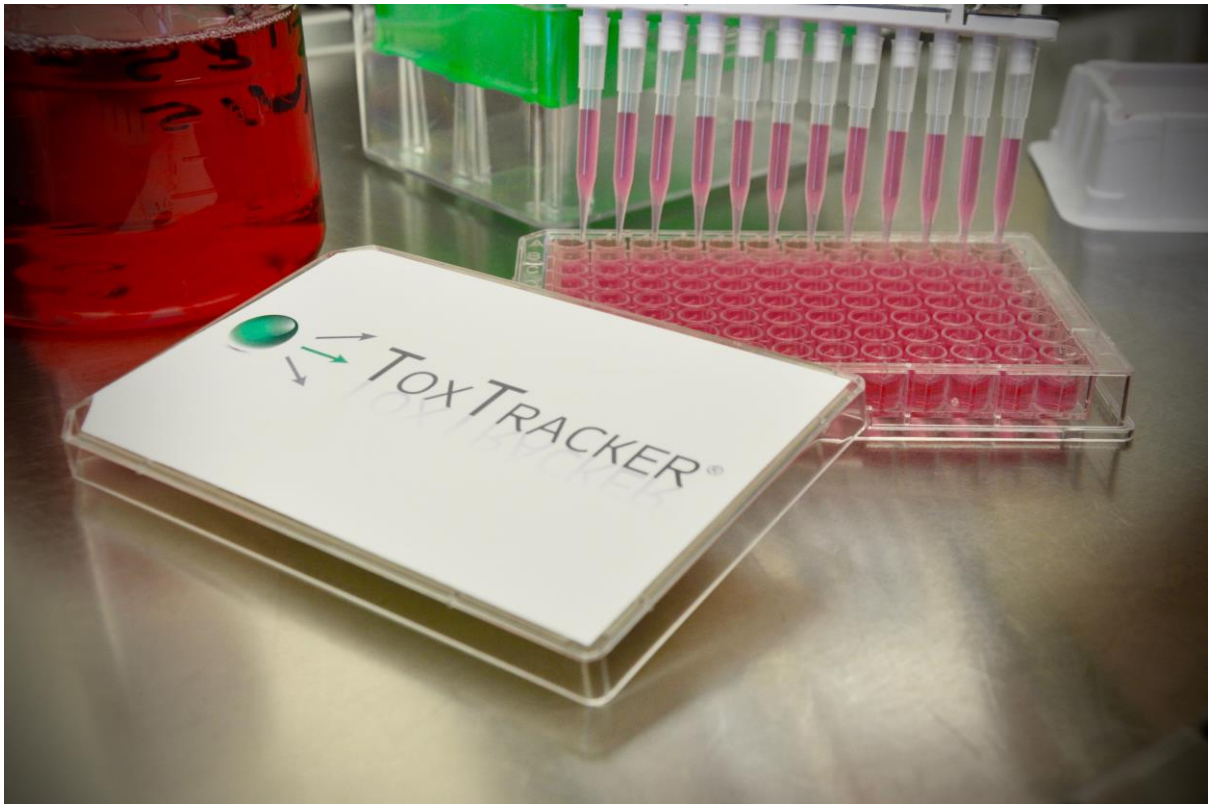
47. Hernández LG, van Steeg H, Luijten M, van Benthem J. Mechanisms of non-genotoxic carcinogens and importance of a weight of evidence approach. Mutation Research/Reviews in Mutation Research. 2009 Sep 1;682(2–3):94–109.

48. Fowler P, Smith R, Smith K, Young J, Jeffrey L, Kirkland D, et al. Reduction of misleading ("false") positive results in mammalian cell genotoxicity assays. II. Importance of accurate toxicity measurement. Mutat Res [Internet]. 2012 Aug 30 [cited 2022 Jul 22];747(1):104–17. Available from: https://pubmed.ncbi.nlm.nih.gov/22595317/

49. Woźniak K, Blasiak J. In vitro genotoxicity of lead acetate: induction of single and double DNA strand breaks and DNA-protein cross-links. Mutat Res [Internet]. 2003 Mar 3 [cited 2022 Jul 22];535(2):127–39. Available from: https://pubmed.ncbi.nlm.nih.gov/12581530/

Annex 1: Full ToxTracker protocol

ToxTracker assay protocol



Authors: Remco Derr
Giel Hendriks

Version: Final validation protocol (v2.3)
Date: 26th of October 2018

Table of content.

1. Background

The ToxTracker assay is a panel of GFP-based mouse embryonic stem (mES) reporter cell lines that can be used to identify the biological reactivity and potential carcinogenic properties of newly developed chemicals in a single test (Hendriks, 2012; Hendriks, 2016). ToxTracker is a mammalian stem cell-based assay that monitors activation of specific cellular signalling pathways for detection of the biological reactivity of compounds (Hendriks, 2013). In contrast to the cancer-derived cell lines that are currently used for *in vitro* genotoxicity testing, stem cells are genetically stable and proficient in all cellular pathways required for accurate detection of potentially carcinogenic properties of compounds. Extensive whole-genome transcription profiling has led to identification of a panel of biomarker genes that are preferentially activated upon exposure to different classes of carcinogens and toxicants (Hendriks, 2011). To allow easy assessment of the activation status of these biomarker genes, we have generated green fluorescent (GFP) mES reporter cell lines. These reporters were created using artificial chromosomes that contain the complete biomarker gene including promoter and regulatory elements ensuring physiological regulation of the GFP reporters following transfection into stem cells.

ToxTracker consists of a panel of six different mES GFP reporter cell lines representing four distinct biological responses that are associated with carcinogenesis, i.e. general cellular stress, DNA damage, oxidative stress and the unfolded protein response (Table 1).

Table 1: Specificity of the ToxTracker reporters.

Biological damage	Cellular pathway	Biomarker gene	MOA
DNA damage	ATR/Chk1 DNA damage signaling	Bscl2	Mutagenic DNA lesions
	NF-kB signaling	Rtkn	DNA double-strand breaks
Oxidative stress	Nrf2 antioxidant response	Srxn1	ROS production
	Nrf2-independent	Blvrb	ROS production
Protein damage	Unfolded protein response	Ddit3	Protein damage
Cellular stress	p53 signaling	Btg2	Cytotoxicity

2. Materials and equipment

- 6 ToxTracker reporter cell lines
- Wild type mES cells (B4418)
- Cell culture 96-wells plates
- Round bottom 96-wells plates (for compound dilutions)

- Sterile 0.1% gelatin in water solution (Sigma-Aldrich, G1890-100G)
- mES cell culture medium (see annex I)
- 0.05% trypsin-EDTA solution (Gibco 25300096)
- Sterile PBS (Gibco 14190094)
- 2% foetal bovine serum (FBS) in PBS
- Aroclor-1254 induced male Sprague Dawley rat liver S9, in 0.15 M KCl (Moltox 11-101.5)
- NADPH Regensys™ A solution (Moltox 60-200.5)
- NADPH Regensys™ B (Moltox 60-201.5L)
- DMSO
- Multichannel liquid reservoirs (12 well)
- 50 ml liquid reservoirs

- CO_2 incubator (37℃, 5% CO_2)
- Laminar flow cabinet
- Centrifuge with rotor for 15 ml tubes
- Water bath
- Multichannel pipet (20-200 µl)
- Multichannel pipet (2-20 µl)
- Pipets + tips
- Cell counter

3. Preparation of cell culture plates (day 1)

All 96-wells plates should be coated with gelatin before seeding of the ToxTracker reporter cell lines

- Add 50 µl 0.1% gelatin solution to each well of a 96-wells plate using a multichannel pipet
- Incubate for at least 5 min at RT (up to few hours)
- Start preparation of the cell lines

Instructions on cell viability
Only start the dose finding if the wild type mES cells are growing properly. The condition of the cell culture can be assessed by checking the morphology of the cells and by cell count. Two days after seeding the cells, the culture dish should be 70-90% confluent. The stem cells should be undifferentiated meaning that they , have a uniform morphology, it's difficult to identify individual cells and when the dishes get confluent often grow in dense clusters.
Proper cell growth is also addressed by cell count. When $5x10^6$ cells (p90) or $2x10^6$ cells (p60) were seeded on day 1, a healthy culture should result in $20\text{-}25x10^6$ (p90) or $8\text{-}12x10^6$ (p60) cells on day 3. If these cell numbers are not reached, the dose finding should be postponed. In that case, cells should be tripsinised and passed to fresh culture dishes according to the standard ToxTracker cell culture protocol.

4. Seeding wild-type mESC for dose-finding (day 1)

- Warm mESC medium in a 37°C water bath
- Aspirate medium from the mESC plates
- Wash cells twice with 3.5 ml (60 mm dishes) or 7 ml (90 mm dishes) PBS
- Add 0.5 ml (60 mm dishes) or 1 ml (90 mm dishes) trypsin-EDTA solution
- Incubate 5 min at room temperature (until the cells detach)
- Prepare a 15 ml tube with 2 ml (60 mm dishes) or 4 ml (90 mm dishes) warm mESC medium
- Resuspend the wild type mES cells in the trypsin-EDTA solution (DO NOT ADD MEDIUM)
- Transfer the cell suspension to the 15 ml tube
- Mix thoroughly
- Determine cell concentration
- Calculate the number of cells that is required for the assay (Table 2), 1 row per compound is required. Add +/-10% extra cells/volume for pipetting errors.
- Always include 2 extra rows for the control compounds (Cisplatin (5µM -S9 and +S9) & Aflatoxin B1 (5µM -S9 and +S9). See table 3 for a sample of plate lay-out.
- Include always 1 extra row of each cell line in a separate plate to count on the day of treatment for RPD / RICC calculations!
- Discard gelatin solution completely from the plate (no washing of the plate).
- Add 200 µl wild type mES cell suspension to each well of the 96-wells plate. See sample plate layout below (Table 3).
- Place cells in a cell culture incubator at 37°C, 5% CO2

Table 2: Cell numbers required for dose-finding

No. of rows	Number of cells	Volume mESC medium
1	$0.48x10^6$	2.4 ml
2	$0.96x10^6$	4.8 ml
3	$1.44x10^6$	7.2 ml
4	$1.92x10^6$	9.6 ml
5	$2.40x10^6$	12.0 ml
6	$2.88x10^6$	14.4 ml
7	$3.36x10^6$	16.8 ml
8	$3.84x10^6$	19.2 ml
9	$4.32x10^6$	21.6 ml
10	$4.80x10^6$	24.0 ml

Table 3: Sample plate design for dose-finding in wild type mESCs. Compound concentrations in µg/ml.

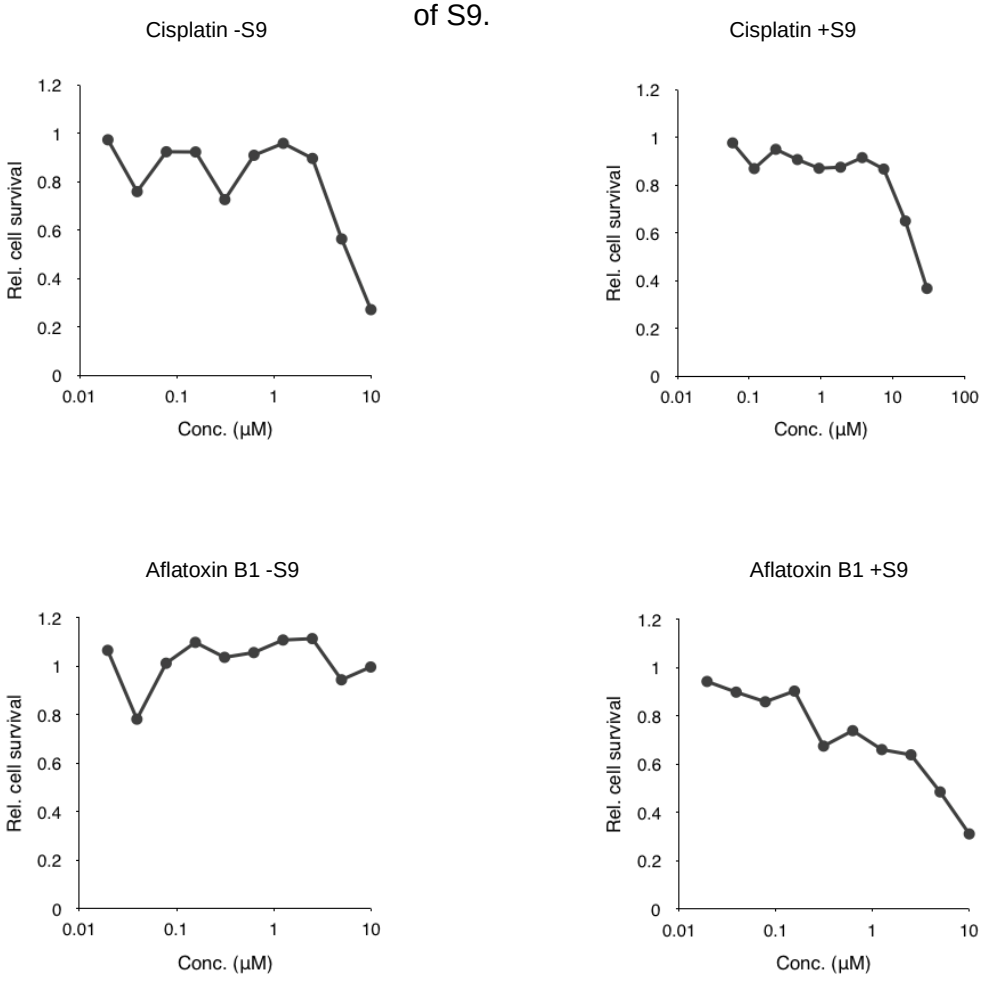
	1	2	3	4	5	6	7	8	9	10	11	12	
A	0	0,001	0,004	0,02	0,06	0,2	1	3,9	15,6	62,5	250	1000	Compound 1
B	0	0,001	0,004	0,02	0,06	0,2	1	3,9	15,6	62,5	250	1000	Compound 2
C	0	0,001	0,004	0,02	0,06	0,2	1	3,9	15,6	62,5	250	1000	Compound 3
D	0	0,001	0,004	0,02	0,06	0,2	1	3,9	15,6	62,5	250	1000	Compound 4
E	0	0,001	0,004	0,02	0,06	0,2	1	3,9	15,6	62,5	250	1000	Compound 5
F	0	0,001	0,004	0,02	0,06	0,2	1	3,9	15,6	62,5	250	1000	Compound 6
G	0	0,001	0,004	0,02	0,06	0,2	1	3,9	15,6	62,5	250	1000	Compound 7
H	0	0,001	0,004	0,02	0,06	0,2	1	3,9	15,6	62,5	250	1000	Compound 8

Table 4: Sample plate design for the dose-finding controls. Compound concentrations in µM.

	1	2	3	4	5	6	7	8	9	10	11	12	
A	0	0,01	0,02	0,04	0,08	0,2	0,3	0,6	1,3	2,5	5	10	CisPt -S9
B	0	0,01	0,02	0,04	0,08	0,2	0,3	0,6	1,3	2,5	5	10	AFB1 -S9
C	0	0,01	0,02	0,04	0,08	0,2	0,3	0,6	1,3	2,5	5	10	CisPt +S9
D	0	0,01	0,02	0,04	0,08	0,2	0,3	0,6	1,3	2,5	5	10	AFB1 +S9
E													
F													
G													
H													

Control plate

Figure 1: Example dose response for the control compounds CisPt and AFB1 in absence and presence of S9.



5. Preparation of compound dilution series (day 1)

Every chemical/substance that is tested in the ToxTracker assay is first analysed for cytotoxicity in a broad dose range finding. For the dose range finding, a maximum concentration of 1 mg/ml or 1 µl/ml is used. In case concentrations are limited by solubility or the occurrence of precipitation in the culture medium, the maximum soluble concentration will be used in the assay (see below). 11 different concentrations for the test substance will be tested, starting at the maximum concentration and ten consecutive 4-fold dilutions. Later in the ToxTracker assay, compounds will be tested in 2-fold dilution series. Therefore, already during compound dilution preparations for the dose finding, 22 serial dilutions in 2-fold dilution steps will be made. From these 22 dilutions series, the appropriate dilutes will be used for the dose finding. The 2-fold dilution series will be prepared in 96-wells round bottom plates for easy handling of the dilutions using a multi channel pipet. Please find the overview of the compound concentration range and plate layout below.

Instructions on compound solubility
Solubility is initially assessed by eye when the stock solutions are prepared. In case a compound does not dissolve completely in the proposed solvent at room temperature, the solution can be warmed at 37 ℃ or placed in a sonication bath for some time. When the compounds still does not dissolve at all in the proposed solvent, please contact Els for instructions. When a compound partly dissolves, continue with preparations of the compound dilution series as indicated below, even if you see some precipitation at higher concentrations. Make sure you create a homogeneous suspension before you prepare the dilution range. In case of partial soluble, it is highly recommend to prepare the dilution series in 1.5 ml Eppendorf tubes in stead of a multi-well plate.
The maximum test concentration in the dose range finding is 1 mg/ml. Expose the cells to the compound dilutions. Solubility of the compounds will be assessed in the cell culture medium **at the end of the 24 h. exposure**. Precipitation in the cell cultures should be observed under a microscope. Define the maximum soluble concentration at the end of the dose range finding. The top doses for the ToxTracker analysis should be based on cytotoxicity as described below, but is limited by the maximum soluble concentrations in cell culture medium after 24 h. incubation at 37 ℃.

Table 5: Compound concentration range that will be prepared for the dose finding and ToxTracker assay. Concentrations that will be applied in the dose finding are indicated in red.

Compound dilution	Prepared stock dilution (µg/ml)	Final concentration in well
1	0,0476837158203125	0,000476837158203125
2	**0,095367431640625**	**0,00095367431640625**
3	0,19073486328125	0,0019073486328125
4	**0,3814697265625**	**0,003814697265625**
5	0,762939453125	0,00762939453125
6	**1,52587890625**	**0,0152587890625**
7	3,0517578125	0,030517578125
8	**6,103515625**	**0,06103515625**
9	12,20703125	0,1220703125
10	**24,4140625**	**0,244140625**
11	48,828125	0,48828125
12	**97,65625**	**0,9765625**
13	195,3125	1,953125
14	**390,625**	**3,90625**
15	781,25	7,8125
16	**1.562,5**	**15,625**
17	3.125	31,25
18	**6.250**	**62,5**
19	12.500	125
20	**25.000**	**250**
21	50.000	500
22	**100.000**	**1000**

Table 6: Plate layout for preparations of the 2-fold dilution series. 22 serial dilutions are prepared. Dilutions that will be applied in the dose range finding are indicated in red.

Dilution plate 1

	1	2	3	4	5	6	7	8	9	10	11	12
A	0	1	2	3	4	5	6	7	8	9	10	11
B	0	1	2	3	4	5	6	7	8	9	10	11
C	0	1	2	3	4	5	6	7	8	9	10	11
D	0	1	2	3	4	5	6	7	8	9	10	11
E	0	1	2	3	4	5	6	7	8	9	10	11
F	0	1	2	3	4	5	6	7	8	9	10	11
G	0	1	2	3	4	5	6	7	8	9	10	11
H	0	1	2	3	4	5	6	7	8	9	10	11

Dilution plate 2

	1	2	3	4	5	6	7	8	9	10	11	12
A	12	13	14	15	16	17	18	19	20	21	22	
B	12	13	14	15	16	17	18	19	20	21	22	
C	12	13	14	15	16	17	18	19	20	21	22	
D	12	13	14	15	16	17	18	19	20	21	22	
E	12	13	14	15	16	17	18	19	20	21	22	
F	12	13	14	15	16	17	18	19	20	21	22	
G	12	13	14	15	16	17	18	19	20	21	22	
H	12	13	14	15	16	17	18	19	20	21	22	

- Prepare for each compound a 100 mg/ml solution in DMSO or water according to the MSDS sheet or instructions from the Compound Selection Team.
- In case of liquid compounds, prepare a 100 µl/ml stock solution. The compound is diluted in DMSO unless instructed otherwise.
- Dilute the compound in 22 consecutive 2-fold dilutions in the appropriate solvent
- Have at least 125 µl for each dilution available
- Dilutions can be prepared in 96-wells round bottom plates or multi-tube strips as shown in table 6
- For convenience, dilution series can be prepared a number of days before the dose range finding and stored at -20℃ for future testing.

6. Exposure of the wild-type mESC for dose-finding (day 2)

24 h after seeding of the wild type mES cells in the 96-well plates, fresh ES cell medium containing the diluted chemicals is added to the cells. For all compounds 11 4-fold dilutions are tested. The maximum tested concentration that will be tested in the dose finding is 1 mg/ml. Also a vehicle control is included. The reference compounds cisplatin and aflatoxin B1 are included as positive controls (maximum tested concentration 5 µM) **The dose range finding is performed in absence and presence of S9.**

Exposure in absence of S9

- Warm mESC medium in a 37°C water bath
- Prepare mESc medium with the dilution series of the test compound (198 µl medium + 2 µl compound). 11 consecutive 4-fold dilutions (see tables 5 and 6).
- Aspirate medium from the reporter cell 96-wells plates
- Add 200 µl of the compound dilution in medium to the cells. A sample plate design can be found in Tables 3 and 4.
- Store plates for 24 h in a cell culture incubator at 37°C, 5% CO2
- Important: Perform a cell count in the row of cells that was seeded separately to find the cell concentration at the moment of treatment for calculation of RPD and RICC.
- Warm up the Trypsin-EDTA solution in water bath
- Aspirate cell culture medium
- Wash cells twice with 200 µl PBS
- Completely aspirate PBS and add 40 µl trypsin-EDTA solution
- Incubate for 5 min at room temperature
- Resuspend cells in trypsin-EDTA solution (DO NOT ADD MEDIUM)
- Add 110 µl of cold PBS supplemented with 2% FBS
- Perform a cell count by flow cytometry. Alternatively, manual or automated cell count (e.g. Coulter counter) can be used.

Exposure in presence of S9

- Warm mESC medium in a 37°C water bath
- Prepare mESC medium with the dilution series of the test compound (193 µl medium + 2 µl compound). 11 consecutive 4-fold dilutions (see tables 5 and 6).
- Prepare a 10% S9 rat liver (aroclor-1254 induced rats) solution with the RegenSysA/B cofactor solutions (according to the Moltox manufacturers protocols)
- Add 5 µl of the 10% S9 solution to every compound dilution
- Aspirate medium from the reporter cell 96-wells plates
- Add 200 µl of the compound dilutions in medium containing 0.25% S9 to the cells. A sample plate design can be found in Tables 3 and 4.
- Store plates for 24 h in a cell culture incubator at 37°C, 5% CO2

7. Analysis of dose-finding by Flow Cytometry (day 3)

Cell concentrations in each well are determined after 24h exposure in absence or presence of S9. Cell count in the wells is used to estimate the relative cell survival after exposure.

- Warm up the Trypsin-EDTA solution in water bath
- **CHECK CELL CULTURES FOR COMPOUND PRECIPITATION**
- Aspirate cell culture medium
- Wash cells twice with 200 µl PBS

- Completely aspirate PBS and add 40 µl trypsin-EDTA solution
- Incubate for 5 min at room temperature
- Resuspend cells in trypsin-EDTA solution (DO NOT ADD MEDIUM)
- Add 110 µl of cold PBS supplemented with 2% FBS
- Analyse samples by flow cytometry
- IMPORTANT: check samples for GFP induction of selected concentrations for autofluorescence of the compound. If an increase of 2-fold or more is observed, seed wild-type mESC during seeding of ToxTracker reporter cell lines for **both -S9 and +S9** treatments to correct for autofluorescence of compound.

Selection of the compound concentration for ToxTracker analysis.

Selection of the maximum concentration that will be applied in the ToxTracker assay will be determined based on cell count. The cell concentration after compound exposure is divided by the cell concentration of the vehicle control exposed cells. Cytotoxicity based on RPD will be calculated for data analysis, but will not be used for selection of the maximum tested concentration. The top concentration that is selected for ToxTracker induces 50-75% cytotoxicity. Concentration selection for -S9 and +S9 should be equal. Only in cases where a more >4-fold difference in top concentration is observed, you can chose a different top concentration for -S9 and +S9 treatment. Concentrations that induce >75% cytotoxicity, corresponding to <0.25 relative cell survival in the ToxTracker results spreadsheet, should not be included in this study. In the case that a tested concentration gives <50% cytotoxicity and the following testing concentration (4-fold higher) induces >75% cytotoxicity, the non-tested intermediate concentration should be selected as top concentration for the ToxTracker analysis.
In case no cytotoxicity is observed, the maximum concentration of 1 mg/ml is applied in ToxTracker. Selection of the top concentration for the ToxTracker analysis can also be limited by solubility of the compound (see instructions above).

8. Preparation of cell culture plates (day 3)

All 96-wells plates should be coated with gelatin before seeding of the ToxTracker reporter cell lines

- Add 50 µl 0.1% gelatin solution to each well of a 96-wells plate using a multichannel pipet
- Incubate for at least 5 min at room temperature (up to few hours)
- Start preparation of the ToxTracker reporter cell lines

Instructions on cell viability
Only start the ToxTracker assay if the reporter cells are growing properly. The condition of the cell culture can be assessed by checking the morphology of the cells and by cell count. Two days after seeding the reporter cells, the culture dish should be 70-90% confluent. The stem cells should be undifferentiated meaning that they have a uniform morphology, it's difficult to identify individual cells and when the dishes get confluent often grow in dense clusters.
Proper cell growth is also addressed by cell count. When $5x10^6$ cells (p90) or $2x10^6$ cells (p60) were seeded on day 1, a healthy culture should result in $20\text{-}25x10^6$ (p90) or $8\text{-}12x10^6$ (p60) cells on day 3. If these cell numbers are not reached, the dose finding should be postponed. Cells should be tripsinised and passed to fresh culture dishes according to the standard ToxTracker cell culture protocol.

9. Seeding ToxTracker reporter cell lines (day 3)

- Warm mESC medium in a 37°C water bath
- Aspirate medium from the mESC plates
- Wash cells twice with 3.5 ml (60 mm dishes) or 7 ml (90 mm dishes) PBS
- Add 0.5 ml (60 mm dishes) or 1 ml (90 mm dishes) trypsin-EDTA solution
- Incubate 5 min at room temperature (until the cells detach)
- Prepare for each cell line a 15 ml tube with 2 ml (60 mm dishes) or 4 ml (90 mm dishes) warm mESC medium
- Resuspend the mES cells in the trypsin-EDTA solution (DO NOT ADD MEDIUM)
- Transfer the cell suspension to the 15 ml tube
- Mix thoroughly
- Determine cell concentration
- Calculate the number of cells that is required for the assay (Table 4). Add 10% extra cells/volume to ensure sufficient cell suspension for seeding.
- Always include an extra control plate for the four reference compounds.

- Include a second extra plate to count cell concentration at the day of treatment for RPD and RICC calculations.
- Discard gelatin solution completely from the plate (no washing of the plate).
- For each reporter cell line, add 200 µl cell suspension to each well of the 96-wells plate. Seed each row of the plate with a different reporter cell line. See sample plate layout below (Table 8).
- DO NOT ADD G418 to the plates
- Place cells in a cell culture incubator at 37°C, 5% CO2

Table 7: Cell numbers required for ToxTracker analysis

No. of plates	Number of cells	Volume mESC medium
1	$0.48x10^6$	2.4 ml
2	$0.96x10^6$	4.8 ml
3	$1.44x10^6$	7.2 ml
4	$1.92x10^6$	9.6 ml
5	$2.40x10^6$	12.0 ml
6	$2.88x10^6$	14.4 ml
7	$3.36x10^6$	16.8 ml
8	$3.84x10^6$	19.2 ml
9	$4.32x10^6$	21.6 ml
10	$4.80x10^6$	24.0 ml

Table 8: Sample plate design for the ToxTracker assay

	1	2	3	4	5	6	7	8	9	10	11	12
A	Bscl2-GFP											
B	Srxn1-GFP											
C	Btg2-GFP											
D	Rtkn-GFP											
E	Blvrb-GFP											
F	Ddit3-GFP											
G	wild type mESC (required only in case of autofluorescence)											
H												
	x	x	x	x	x	x	x	x	x	x	x	x

10. Exposure of the ToxTracker reporter cell lines (day 4)

24h after seeding the cells in the 96-well plates, fresh ES cell medium containing the diluted chemicals is added to the cells. The compound dilutions were already prepared during the dose range finder. For each tested compound, five concentrations are tested in 2-fold dilutions. The highest compound concentration will induce significant cytotoxicity (50-75% cell death), or will be the maximum soluble compound concentration or 1 mg/ml (in case of no/low cytotoxicity). Solubility is determined during the dose range finding as described above. Positive reference treatments with cisplatin (DNA damage), diethyl maleate (oxidative stress), tunicamycin (unfolded protein response) and aflatoxin B1 (S9 metabolism) are included in all experiments.

Exposure in absence of S9

- Warm mESC medium in a 37°C water bath
- Prepare mESC medium with the dilution series of the test compound (1.5 ml medium + 15 µl compound). Five consecutive 2-fold dilutions.
- For each experiment, a control plate should be prepared with the positive controls cisplatin, diethyl maleate, tunicamycin and Aflatoxin B1. For the control compounds, two concentrations are included. See table 9 for a standard control plate layout.
- Aspirate medium from the reporter cell 96-wells plates
- Add 200 µl of the compound dilutions to the reporter cell lines. A sample plate design can be found in Table 10.
- Store plates in a cell culture incubator at 37°C, 5% CO2
- Perform a cell count on the extra plate that was seeded on day 3 for the calculation of RPD and RICC. Cell count can be performed manually, in an automated cell counter or by using a flow cytometer. For flow cytometry-based cell count, counting beads can be added to the cell suspension for a reliable absolute cell number calculation.
- Warm up the Trypsin-EDTA solution in water bath
- Aspirate cell culture medium
- Wash cells twice with 200 µl PBS
- Completely aspirate PBS and add 40 µl trypsin-EDTA solution
- Incubate for 5 min at room temperature
- Resuspend cells in trypsin-EDTA solution (DO NOT ADD MEDIUM)
- Add 110 µl of cold PBS supplemented with 2% FBS
- Analyse samples by flow cytometry

Table 9: Sample plate design for the ToxTracker control plate.

	1	2	3	4	5	6	7	8	9	10	11	12
Bscl2-GFP	0	2,5	5	0	125	250	0	2	4	0	2,5	5
Srxn1-GFP	0	2,5	5	0	125	250	0	2	4	0	2,5	5
Btg2-GFP	0	2,5	5	0	125	250	0	2	4	0	2,5	5
Rtkn-GFP	0	2,5	5	0	125	250	0	2	4	0	2,5	5
Blvrb-GFP	0	2,5	5	0	125	250	0	2	4	0	2,5	5
Ddit3-GFP	0	2,5	5	0	125	250	0	2	4	0	2,5	5
	x	x	x	x	x	x	x	x	x	x	x	x
	x	x	x	x	x	x	x	x	x	x	x	x

CisPt DEM Tunica AFB1 +S9

Table 10: Sample plate design for the compound testing in the ToxTracker assay

	1	2	3	4	5	6	7	8	9	10	11	12
Bscl2-GFP	0	0,625	1,25	2,5	5	10	0	6,25	12,5	25	50	100
Srxn1-GFP	0	0,625	1,25	2,5	5	10	0	6,25	12,5	25	50	100
Btg2-GFP	0	0,625	1,25	2,5	5	10	0	6,25	12,5	25	50	100
Rtkn-GFP	0	0,625	1,25	2,5	5	10	0	6,25	12,5	25	50	100
Blvrb-GFP	0	0,625	1,25	2,5	5	10	0	6,25	12,5	25	50	100
Ddit3-GFP	0	0,625	1,25	2,5	5	10	0	6,25	12,5	25	50	100
mESC (when AF)	0	0,625	1,25	2,5	5	10	0	6,25	12,5	25	50	100
	x	x	x	x	x	x	x	x	x	x	x	x

Compound 1 Compound 2

Exposure in presence of S9

- Warm mESC medium in a 37°C water bath
- Prepare mESC medium with the dilution series of the test compound (1.5 ml medium + **15.4 µl** compound). Five consecutive 2-fold dilutions
- Prepare a 10% S9 rat liver (aroclor-1254 induced rats) solution with the RegenSysA/B cofactor solution (according to the Moltox manufacturers protocols)

- Add 37.5 µl to each medium + compound mix (1.5 ml + 15.4 µl mix) to obtain a 0.25% S9 concentration in the medium + compound mix solution.
- Aspirate medium from the reporter cell 96-wells plates
- Add 200 µl of the compound dilutions with S9 to the reporter cell lines. A sample plate design can be found in Table 10 (the positive control Aflatoxin B1 +S9 is included on the control plate, Table 9).
- Store plates for 24 h in a cell culture incubator at 37°C, 5% CO2

11. Analysis of ToxTracker reporter induction (day 5)

Induction of the GFP reporters is determined after 24 h exposure using a flow cytometer. Only GFP expression in intact single cells is determined. Mean GFP fluorescence in each well is measured. During GFP detection, also cell concentration in each well is determined and used for cytotoxicity assessment.

ToxTracker analyses can be performed on various flow cytometers with 96-well plate capabilities. Proper settings for the flow cytometer should be verified before running ToxTracker by testing the control compounds in the ToxTracker cell lines. Important parameters when setting the proper instrument settings are:

- separation of intact cells from broken cells and debris
- appropriate fluorescence levels of the untreated control cells
- Detection of fluorescence induction in treated reporter cells

Guidance on setting the appropriate cell gates and fluorescence levels for the BD FacsCanto, Millipore Guava and Miltenyi MacsQuant can be found in Annex III.

- Warm up the Trypsin-EDTA solution in a 37ºC water bath
- Aspirate cell culture medium
- Wash twice with 200 µl PBS
- Completely remove PBS and add 40 µl of trypsin-EDTA solution
- Incubate for 5 min at room temperature
- Resuspend cells in trypsin-EDTA solution (DO NOT ADD mESC MEDIUM)
- Add 110 µl cold 2% FBS in PBS solution
- Analyse samples by flow cytometry.
- Determine mean GFP fluorescence in 5000 intact mES cells as well as the cell concentration in each well. Alternatively on the BD FacsCanto flow cytometers, a fixed volume (**30 µl**) of cell suspension can be analysed to measure GFP induction and to estimate the cell concentration in the wells.
- Calculate relative GFP induction levels in exposed reporter cells compared to the vehicle control exposed cells.

- Calculate relative cell survival in exposed reporter cells compared to the vehicle control exposed cells.

12. Quality controls for the ToxTracker assay

To ensure proper performance of the ToxTracker assay, various quality control checks are included in the protocol. In case the quality limits are not met, the results from this ToxTracker test should be discarded and a repeat experiment should be conducted.

The quality of the cells that are applied in the ToxTracker assay is assessed from **cell count** of vehicle control treated samples in the assay. At the start of the ToxTracker analysis, 40,000 cells are seeded per well of the 96-wells plate. After two days when the ToxTracker reporter cell lines are analysed for GFP expression, the **minimum cell concentration** per well should be $4x10^5$ cells/ml for all the cell lines (-20% for the Srxn1-GFP reporter because of slower cell growth).

GFP induction and cytotoxicity levels of the positive controls validate the overall quality of the ToxTracker assay. Treatment with Cisplatin (DNA damage), Diethyl maleate (oxidative stress), Tunicamycin (protein unfolding) and Aflatoxin B1 (S9 metabolism) are standard controls for theToxTracker assay and should be included in every experiment. For each control, two concentrations are included in the control plates as indicated in Table 9. Control compounds should results in a **minimal GFP induction levels** in the relevant ToxTracker reporter cell lines as indicated in Table 11. In case the GFP induction levels are below these thresholds, all test results from this experiment should be discarded. Cytotoxicity for Cisplatin at 5 µM is 50% (+/-20%). For Diethyl maleate cytotoxicity levels should be around 50% (+/-30%) at 250 µM. Tunicamycin induces a cytotoxicity level of 50% (+/-20%) at a concentration of 4 µg/ml. For Aflatoxin B1 (5 µM) cytotoxicity levels should be around 50% (+/-20%) in presence of S9. Discard the experiment if those levels of cytotoxicity are not reached. If the quality of the cells was not acceptable, prepare fresh control compound dilutions.

Table 11: Minimum induction levels for the ToxTracker reporter cell lines by the relevant control compound.

Cell line	Compound treatment	Min. relative GFP induction level
Bscl2-GFP	Cisplatin/Aflatoxin B1+S9	2
Srxn1-GFP	DEM	8
Btg2-GFP	Cisplatin/Aflatoxin B1+S9	2
Rtkn-GFP	Cisplatin/Aflatoxin B1+S9	3
Blvrb-GFP	DEM	5
Ddit3-GFP	Tunicamycin	4

Autofluorescence by the test compound may interfere with the measurement of GFP reporter induction. Therefore, GFP levels are measured in wild-type mES during the dose range finding. If a relative induction of 2 or more is found at concentrations that will be applied during the ToxTracker assay, wild-type mES cells should be included in the ToxTracker assay to correct afterwards for the level of autofluorescence caused by the compound. In case of autofluorescence, the green fluorescence signals are subtracted from the GFP reporter fluorescence levels in the ToxTracker assay. After green fluorescence correction, relative inductions of the GFP reporters in exposed cells is calculated as described above in the protocol. In case autofluorescence is observed for a compound, wild type mES cells should be included in all tests -S9 and +S9 for this compound to perform fluorescence correction.

Additional checks:

- Changes of batches of material should be monitored with regard to their influence on principal endpoints in use in a study.
- Check cells each day for confluence, adherence, cell morphology and contamination.
- The impact of variation of cell proliferation and cell differentiation should be monitored and documented.
- Perform checks to control functionality of FACS.
- Solubility of the test compounds should be verified during the dose range finding, prior to the ToxTracker analysis.
- Changes in pH of the cell culture medium by the compounds should be monitored by the phenol red pH indicator in the mESC medium. In case changes in pH of the mESC medium, pH should be measured and reported.
- During the ToxTracker analysis, changes in cell viability and cell morphology should be monitored after 6h exposure and after 24h, just before preparation of the cells for flow cytometry analysis.

Criteria for a positive ToxTracker result

The ToxTracker assay is considered to have a positive response when a compound induces at least a 2 fold increase in GFP expression in any of the reporters. Activation of the Bscl2-GFP or Rtkn-GFP reporters indicate induction of DNA damage. Btg2 induction is associated with activation of the p53 tumor suppressor. Srxn1-GFP and Blvrb-GFP indicated induction of cellular oxidative stress and Ddit3-GFP activation is associated with the unfolded protein response. Only GFP inductions at compound concentrations that showed <75% cytotoxicity are used for the ToxTracker analysis. Data from measurements >75% cytotoxicity can not be interpreted in a meaningful way and are therefore discarded.

Requirements for repeat experiments

ToxTracker is standard performed in three independent repeats. When the three repeats give the same ToxTracker results, positive or negative, for the various reporters, the analysis can be considered highly reliable and reproducible. In case one of the repeat experiments gives a different results than the other two repeats (two clear positives and one clear

negative, or two clear negatives and one clear positive), **a fourth repeat experiment is demanded**. In case the repeat experiments give comparable results that are around the cut-off values for a positive results (2-fold increase in fluorescence), a fourth repeat experiment might be waived.

ANNEX I: Composition of mES cell culture medium

- 450 ml Knockout ES medium (KO-DMEM, Gibco 10829018)
- 50 ml foetal bovine serum (QC-ed and certified for ES cells by Toxys)
- 5 ml Glutamax (Gibco 35050061)
- 5 ml 100 mM Sodium Pyruvate (Gibco 11360039)
- 5 ml Non-essential amino acids (Gibco 11140035)
- 1 ml 50 mM 2-mercaptoethanol (Gibco 31350010)
- 5 ml Penicillin-Streptomycin (10,000 U/mL) (Gibco 15140122)
- 500 µl Leukemia Inhibitory Factor (LIF, Toxys home-made)

ANNEX II: Checklists for the ToxTracker assay

ToxTracker assay data check list

Dose range finding			Checked	Remarks
	- Performed in wild type mES cells - Test in 96-wells plate - Test 11 compound concentrations in 4-fold dilution steps and a vehicle control - Dose range finding performed in absence and presence of S9 - Cytotoxicity calculated as relative cell count after 24 h. exposure			
	Before exposure	Were the cells growing propertly	☐	Do not continue if cells are dying, differentiating or show slow growth rate
		Were the compounds properly dissolved	☐	Only use soluble compound concentrations
		Start dose finding in absence and presence of S9	☐	
	At moment of exposure	Dd you see precipitation of the compound	☐	In case of precipitation, adjust your test concentrations
		Did you perform a cell count at the moment of exposure	☐	Cell count can be used to calculate cytotoxicity based on RPD
	After exposure	Did you notice issues with the untreated control cells	☐	
		Check for precipitation of the compound	☐	Use microscope to assess compound precipitation in the wells. In case of no/limited cytotoxicity, select the maximum compound concentration that does not precipitate as top doses in ToxTracker
		Did you encounter any problems with cell count	☐	
	Data analysis	Calculate cell viability after treatment based on relative cell count	☐	
		Determine the maximum concentration for ToxTracker	☐	Repeat the dose finding in case of "unusual" dose response curve, i.e. survival curve going up and down
		Maximum concentration results in 30-50% viability (50-70% cytotoxicity)	☐	
		Check if compounds are autofluorescent	☐	If yes, than include wt mES cells in the ToxTracker assay to compensate for autofluorescence
	If you did not encounter any issues during the dose range finding and you we able to determine the proper concentration to apply in the ToxTracker assay, please proceed.			

ToxTracker assay data check

ToxTracker assay			Checked	Remarks
	- Performed in 6 ToxTracker reporter cell lines - Include wild type mES cells in the assay in case of autofluorescence of compounds at concentrations that are applied in ToxTracker (for -S9 AND +S9 treatments - Test in 96-wells plates - Test 5 compound concentrations in 2-fold dilution steps and a vehicle control - Toxtracker assay is performed in absence and presence of S9. - Compound concentrations in ToxTracker -S9 can vary from test +S9 - Determine induction of GFP reporter expression using flow cytometry - Cytotoxicity calculated as relative cell count after 24 h. exposure			
	Before exposure	Were the cells growing propertly	☐	Do not continue if cells are dying, differentiating or a show slow growth rate
		Were the compounds properly dissolved (in case you made fresh solutions)	☐	Try to get the compound into solution by warming the solution or sonication in case you see precipitation of the compound after freeze-thawing the solution. If compound is partly soluble, make a homogenous suspension before preparing the dilutions.
		Start dose finding in absence and presence of S9	☐	
	At moment of exposure	Did you see precipitation of the compound	☐	
		Did you perform a cell count at the moment of exposure	☐	Cell count can be used to calculate cytotoxicity based on RPD
		Did you refresh the medium after 3 h. for the +S9 exposures	☐	
	After exposure	Did you notice issues with the untreated control cells	☐	
		Did you encounter any problems with analysis of the ToxTracker reporters by flow cytometry	☐	
	Data analysis	Calculate cell viability after treatment based on relative cell count	☐	CYTOTOXICITY IN TOXTRACKER SHOULD BE COMPARABLE TO THE DOSE FINDING. IF NOT, RESULTS SHOULD BE DISCARDED
		Determine the relative induction levels of the different ToxTracker reporters	☐	
		In case of autofluorescence, correct the GFP reporter levels using the fluorescence measurement in wt mES cells	☐	See ToxTracker protocol for instructions how to perform the autofluorescence correction
		Did the assay control compounds provide comparable results as your historical controls	☐	
		Did you get comparable results as from previous repeat experiments	☐	

Data acceptance criteria

Cytotoxicity	
	- Only data at compound concentrations that induce less than 75% cytotoxicity (more than 25% relative cell viability) should be considered for ToxTracker analysis - Data from compound concentrations that cause >75% cytotoxicity (meaning less than 25% cell viability) should be discarded - Only data from soluble compounds concentrations should be considered for ToxTracker analysis - Data from compound concentrations that precipitate should be discarded - In case cytotoxicity in ToxTracker deviates significantly from the dose range finder, data should be critically reviewed or discarded
ToxTracker	
	- A 2-fold increase in GFP reporter induction is required for a positive ToxTracker result - Induction levels of <2-fold for the GFP reporters should be considered as a negative ToxTracker result - Induction of the ToxTracker reporters of >1.5-fold but lower than 2-fold indicates potential (geno)toxic properties but further testing will be required. - A >2-fold GFP induction at one concentration is sufficient for a positive ToxTracker result - In case no logical dose response is observed, data should be critically reviewed - Activation of the Bscl2-GFP and Rtkn-GFP reporters indicates genotoxicity of a compound - Activation of the Srxn1-GFP and Blvrb-GFP reporters indicates induction of oxidative stress - Activation of the Ddit3 reporter indicates induction of protein damage - Activation of the Btg2-GFP reporters indicates induction of the p53 response - In case autofluorescence levels of compounds is >10-fold, even after fluorescence compensation ToxTracker results should be interpreted with extreme caution

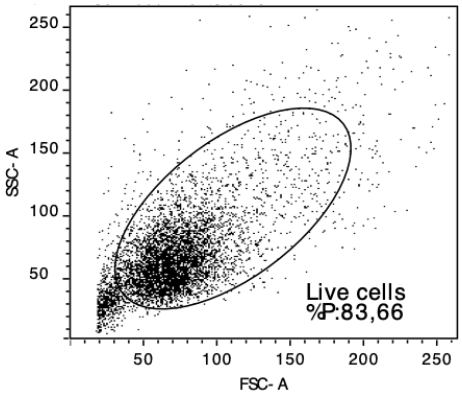
ToxTracker data interpretation

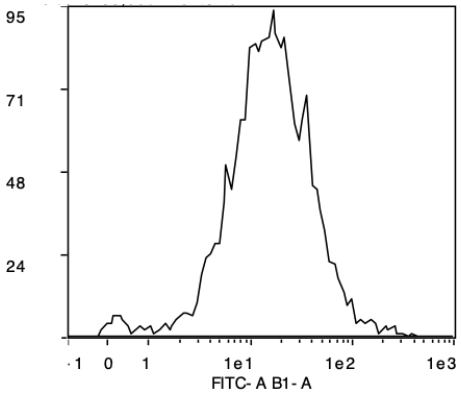
ToxTracker	
Bscl+Rtkn positive	- Compound is classified as genotoxic
	- Activation of Bscl2-GFP indicates DNA reactivity of a compound, induction of bulky DNA lesions and DNA replication stress. Compound likely mutagenic
	- Activation of Rtkn-GFP indicates induction of DNA double strand breaks. Compound likely clastogenic
Rtkn positive, Bscl2 negative	- Compound is classified as genotoxic
	- Compound has potential aneugenic properties
	- Potential indirect genotoxicity related to high levels of oxidative stress
Bscl2+Rtkn negative	- Compound is classified as non-genotoxic
Btg2 positive	- Compound activates the p53 response
	- Compound is potentially genotoxic but only when Bscl2-GFP and/or Rtkn-GFP reporters are activated
Srxn1 and/or Blvrb positive	- Compound induces oxidative stress
Ddit3 positive	- Compound induces the unfolded protein response

ANNEX III: Guidance on flow cytometer settings

Setting gates for living cells.

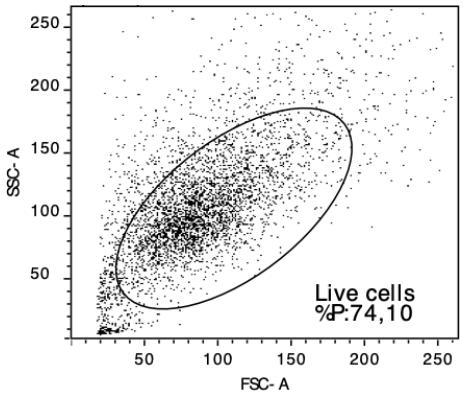
Untreated

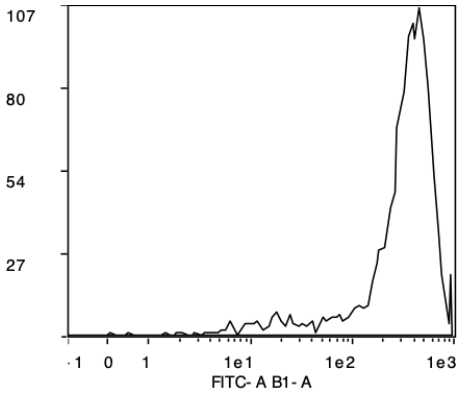




Cisplatin

Make that the intact cells are properly separated from the dead cells and debris. The fluorescence histograms should only be based in the “living cells” gate.





Background GFP levels for each cell line

The proper flow cut-meter settings should be set in a test run, before you start the validation experiments. If you do not get all fluorescence values within the ranges given, please reduce or increase the laser power. Setting will vary between brands and individual flow cytometers.

In the case you obtain much lower levels in untreated cells in 96-well plate please increase the power of the blue laser (GFP signal) until you are around these levels. In case when you are much higher in background levels, please reduce the power of the laser to match the values.

Using laser power settings that result in background too high or too low fluorescence levels in the reporter cell lines can affect the fold changes in ToxTracker reporter activation upon exposure to the test compounds.

	BD Facscanto	Guava	MacsQuant
Bscl2-GFP	300-500	50-80	10-13
Srxn1-GFP	200-400	15-35	4-7
Btg2-GFP	200-400	15-50	4-7
Rtkn-GFP	400-800	90-140	17-24
Blvrb-GFP	400-800	60-110	12-20
Ddit3-GFP	100-300	20-40	4-5

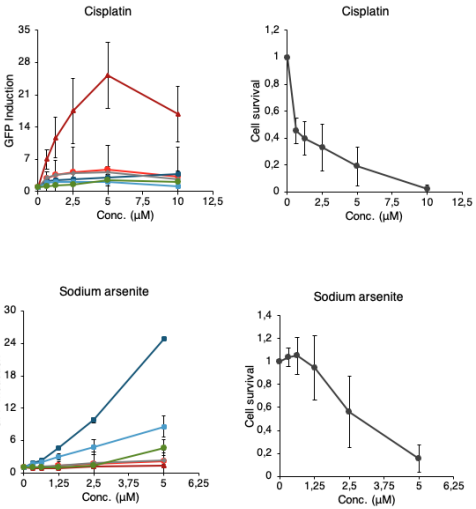
Annex 2: Compound selection for the ToxTracker validation study

	Compound	CAS number	In vitro				In vivo		
			Ames	MLA	MN	CA	MN	CA	TgR
Group I: Genotoxic carcinogens									
1	Etoposide	33419-42-0	P	P	P	P	P	P	N
2	Mitomycin C	50-07-7	P	P	P	P	P	P	P
3	Cisplatin	15663-27-1	P	P	P	P	P	P	P
4	1,2- Dimethylhydrazine	306-37-6	P				P		
5	1,2-dibromoethane	106-93-4	P	P		P	E	N	N
6	Cyclophosphamide	6055-19-2	P	P	P	P	P		P
7	2-Acetylaminofluorene	53-96-3	P	P	P	P	P	P	P
8	Azidothymidine (Zidovudine)	30516-87-1	P	P	P	P	P		
9	ENU	759-73-9	P	P	P	P	P	P	P
10	Acrylonitrile	107-13-1	P	P		P	E	N	N
11	Benzene	71-43-2	N		P	P	P	P	P
12	4,4′ -Oxydianiline	101-80-4	P	P		P	P	P	
13	Busulfan (Myleran)	55-98-1	P	P	P	P	P	P	
14	Ethyl methanesulfonate	62-50-0	P	P	P	P	P	P	P
15	p-Chloroaniline	106-47-8	P	E		P	P		N
16	7,12-Dimethyl-benzanthracene	57-97-6	P	P	P	P	P	P	P
17	Benzo[a]pyrene	50-32-8	P	P	P	P	P	P	P
18	Cadmium Chloride	10108-64-2	E		P	P	P	P	
19	Dimethylnitrosamine	62-75-9	P	P	P	P	P		P
20	2,4-Diaminotoluene	95-80-7	P	P		P	N		P
21	o-Anisidine	90-04-0	P	P		P	N		P
22	4-nitroquinoline-1-oxide	56-57-5	P	P	P	P	P	P	P
Group II: Genotoxic non-carcinogens									
23	6-Mercaptopurine	50-44-2	P	P	P	P	P		
24	Cytosine arabinose	147-94-4	E	P	P	P	P	P	
25	p-Phenylenediamine 2HCl	624-18-0	P	P	P	P	N		
26	8-Hydroxyquinoline	148-24-3	P			P	N		
27	9-Aminoacridine	90-45-9							
28	2,6-Diaminotoluene	823-40-5	P		P	P	P	N	N
29	3-Nitropropionic acid	504-88-1	P	P	N	E			
30	p-Anisidine	104-94-9	P	E		P			
31	5-fluorouracil	51-21-8	N		P	P	P		
32	Phenol	108-95-2	N		P	P	P	N	
Group III: Non-genotoxic carcinogens									
33	Di(2-ethylhexyl)phthalate	117-81-7	N		N	N	N	N	E
34	Lead (ii) acetate trihydrage	6080-56-4	N		P	E	E	P	

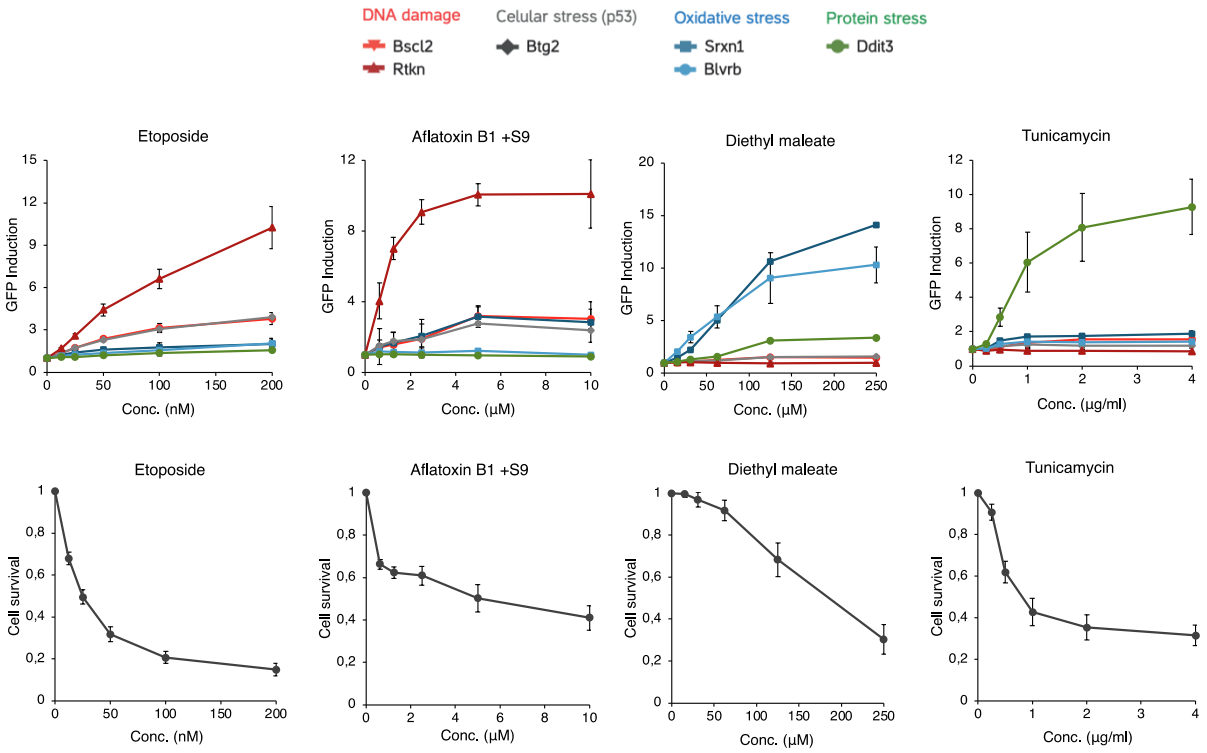
35	2-Phenylphenol sodium salt	6152-33-6	P			E	N	N	
36	Ropinirole hydrochloride	91374-20-8	N	N		N	N		
37	Methyl carbamate	598-55-0	N			N	N		
38	Cyclosporin A (CsA)	59865-13-3	N				N		
39	Sodium saccharin	128-44-9	N	N	N	N	N	N	N
40	Diethanolamine	111-42-2	N			N	N		
41	Hexachloroethane	67-72-1	N		N	N	N		
42	Melamine	108-78-1	N			N	N		
Group IV: Non-genotoxic non-carcinogens									
43	Tunicamycin	11089-65-9	N		P		N		
44	p-Nitrophenol (4-nitrophenol)	100-02-7	N		E	P	N		
45	Phenanthrene	85-01-8	P		E	E			
46	Tertiarybutylhydroquinone	1948-33-0	N		P	P	N	N	
47	Benzyl alcohol	100-51-6	N		N	N	N		
48	Vanilin	121-33-5	N			N	N		
49	Erythromycin	114-07-8	N			N			
50	Sodium diclofenac	15307-79-6	N	N		N	N	N	
51	o-anthranilic acid	118-92-3	N		N	N	N	N	
52	Tolbutamide	64-77-7	N			N	P		
53	2-ethyl-1,3-hexanediol	94-96-2	N	N	N	P	N	N	
54	Chlorpheniramine maleate	113-92-8	N	P		P	N		
55	Ampicillin trihydrate	7177-48-2	N			N	N	P	
56	Sodium chloride	7647-14-5							
57	D-mannitol	69-65-8	N			N	N	N	
58	Allyl alcohol	107-18-6	P	P		P	N		
59	(2-chloroethyl)trimethyl-ammonium chloride	999-81-5	N			N		N	
60	Sulfisoxazole	127-69-5	N		N	N	N	N	
61	Sucrose	57-50-1	N				N		
62	Cyclohexanone	108-94-1	N			P		E	
63	1-Nitropropane	108-03-2	N			N	N		
64	Phenformin HCl	834-28-6	N			N			

Annex 3: Example ToxTracker data analysis template

			Exp. 1				Exp. 2				Exp. 3								
Reporter	Treatment	Conc. (%)	GFP mean	inductie	Cell Conce	survival	GFP mean	inductie	Cell Concent	survival	GFP mean	inductie	Cell Conce	survival	GFP induction	S.D	Rel. cell viability	Summarized su	St. dev. survival
Bscl2	CisPt	0	1709	1	101068	1	1275	1	51773	1	1055	1	75069	1	1	0	1	1	0
Bscl2	CisPt	0,625	5039	2,94851	39675	0,392557486	3345	2,62353	26699	0,515694	3003	2,84645	30569	0,407212	2,806160936	0,16619232	0,438487675	0,453933371	0,097080763
Bscl2	CisPt	1,25	6788	3,97191	32316	0,319745122	4201	3,2949	23347	0,450949	3632	3,44265	26731	0,356086	3,56982313	0,35597066	0,375593398	0,39719126	0,124149773
Bscl2	CisPt	2,5	8381	4,90404	21830	0,215993193	4444	3,48549	24730	0,477662	4638	4,39621	14956	0,19923	4,261912059	0,71874595	0,297628445	0,332911844	0,173675273
Bscl2	CisPt	5	9688	5,66881	6162	0,060968853	4729	3,70902	17490	0,337821	5215	4,94313	3543	0,047197	4,773653247	0,99082692	0,148662101	0,189832504	0,146482771
Bscl2	CisPt	10	5407	3,16384	358	0,00354217	3930	3,08235	592	0,011435	3683	3,491	228	0,003037	3,245728901	0,21627914	0,006004636	0,020146324	0,031457378
Bscl2	NaAsO2	0	1709	1	101068	1	1331	1	66147	1	1087	1	70355	1	1	0	1	1	0
Bscl2	NaAsO2	0,3125	1770	1,03569	107874	1,0673408	1472	1,10594	60725	0,918031	1090	1,00276	69677	0,990363	1,048129555	0,0527	0,99191167	1,036433556	0,079276492
Bscl2	NaAsO2	0,625	1876	1,09772	119154	1,178948827	1562	1,17355	58758	0,888294	1163	1,06992	63360	0,900576	1,113729629	0,05364151	0,989272911	1,050957905	0,162016255
Bscl2	NaAsO2	1,25	2099	1,2282	105593	1,044771837	1946	1,46206	51239	0,774623	1455	1,33855	53189	0,756009	1,34293623	0,11698927	0,858467946	0,943623246	0,278842853
Bscl2	NaAsO2	2,5	2516	1,47221	60630	0,599893141	2440	1,83321	34073	0,51511	1761	1,62006	34951	0,496781	1,641823093	0,18148283	0,537261346	0,559892207	0,310996572
Bscl2	NaAsO2	5	3927	2,29783	22862	0,22620414	2914	2,18933	7887	0,119234	2012	1,85097	10466	0,14876	2,112710761	0,23307942	0,164732811	0,155116261	0,115998027
Snxn1	CisPt	0	2269	1	56546	1	1184	1	40769	1	884	1	49315	1	1	0	1		
Snxn1	CisPt	0,625	5260	2,3182	21702	0,383793725	2722	2,29899	18036	0,442395	1889	2,13688	25985	0,526919	2,251355389	0,09960493	0,451035823		
Snxn1	CisPt	1,25	5346	2,3561	18074	0,319633573	3109	2,62584	17998	0,441463	1978	2,23756	23516	0,476853	2,406501722	0,19898956	0,412649778		
Snxn1	CisPt	2,5	5447	2,40062	15035	0,265889718	3387	2,86064	15405	0,377861	2244	2,53846	20488	0,415452	2,599906814	0,23608762	0,353067345		
Snxn1	CisPt	5	6300	2,77655	9793	0,173186432	3979	3,36064	11531	0,282837	2503	2,83145	17863	0,362222	2,989547801	0,32254685	0,272748777		
Snxn1	CisPt	10	5457	2,40502	1932	0,034166873	4791	4,04645	2529	0,062032	4192	4,74208	6495	0,131704	3,73118613	1,20000161	0,075967883		
Snxn1	NaAsO2	0	1836	1	55859	1	1135	1	46625	1	795	1	51189	1	1	0	1		
Snxn1	NaAsO2	0,3125	3660	1,99346	60664	1,086020158	1835	1,61674	48733	1,045212	1461	1,83774	51725	1,010471	1,815979996	0,18930194	1,047234318		
Snxn1	NaAsO2	0,625	5208	2,8366	56495	1,011385811	2447	2,15595	47875	1,02681	1607	2,02138	58684	1,146418	2,337977364	0,43703113	1,061537879		
Snxn1	NaAsO2	1,25	9788	5,33115	32977	0,590361446	5148	4,53568	33755	0,723968	3042	3,82642	46838	0,915001	4,564417533	0,75278122	0,743110181		
Snxn1	NaAsO2	2,5	22519	12,2653	8347	0,149429814	10261	9,04053	16225	0,347989	6450	8,11321	22282	0,435289	9,806328909	2,17938129	0,310902641		
Snxn1	NaAsO2	5	48619	26,4809	4981	0,089170948	29123	25,659	3539	0,075903	17770	22,3522	7097	0,138643	24,83072297	2,18544841	0,101239167		
Btg2	CisPt	0	1975	1	64762	1	1188	1	61997	1	1234	1	41064	1	1	0	1		
Btg2	CisPt	0,625	7035	3,56203	22438	0,346468608	2916	2,45455	31805	0,513009	3868	3,13452	22909	0,557885	3,050364217	0,55851571	0,472454185		
Btg2	CisPt	1,25	9258	4,68759	16402	0,253265804	3528	2,9697	28603	0,461361	4209	3,41086	23710	0,577391	3,689383634	0,8921745	0,430672742		
Btg2	CisPt	2,5	9620	4,87089	7280	0,112411599	4085	3,43855	25146	0,4056	4774	3,86872	22609	0,55058	4,059385959	0,73495601	0,356197151		
Btg2	CisPt	5	8933	4,52304	1067	0,016475711	4628	3,89562	14082	0,22714	5074	4,11183	9660	0,235243	4,176830771	0,31871791	0,159619427		
Btg2	CisPt	10	4934	2,49823	271	0,004184553	3538	2,97811	632	0,010194	3452	2,79741	352	0,008572	2,757916378	0,24236835	0,007650193		
Btg2	NaAsO2	0	1794	1	66810	1	1056	1	70300	1	1348	1	38859	1	1	0	1		
Btg2	NaAsO2	0,3125	2119	1,18116	64035	0,958464302	1091	1,03314	71413	1,015832	1377	1,02151	45115	1,160992	1,078605571	0,08900442	1,045096252		
Btg2	NaAsO2	0,625	2574	1,43478	54111	0,809923664	1110	1,05114	78815	1,121124	1251	0,92804	50752	1,306055	1,137986838	0,26429885	1,079034215		
Btg2	NaAsO2	1,25	3404	1,89744	41880	0,626852268	1348	1,27652	59664	0,848706	1496	1,10979	54802	1,410278	1,427914445	0,4150746	0,961945334		
Btg2	NaAsO2	2,5	3932	2,19175	15036	0,225056129	2116	2,00379	39421	0,560754	1817	1,34792	36444	0,937852	1,847820335	0,44300744	0,574554092		
Btg2	NaAsO2	5	4827	2,69064	3504	0,052447238	2789	2,6411	8422	0,119801	2265	1,68027	11878	0,305669	2,337333666	0,56957517	0,159305769		
Rtkn	CisPt	0	852	1	75112	1	541	1	59895	1	645	1	50866	1	1	0	1		
Rtkn	CisPt	0,625	7937	9,31573	27469	0,365707211	2862	5,2902	29428	0,491326	4186	6,48992	33490	0,658397	7,031951169	2,06677477	0,505143423		
Rtkn	CisPt	1,25	14191	16,6561	20075	0,267267547	5701	10,5379	25863	0,431806	5147	7,97984	33216	0,65301	11,72461368	4,4582053	0,450694359		
Rtkn	CisPt	2,5	21422	25,1432	13780	0,183459367	8808	16,281	21965	0,366725	7196	11,1566	32334	0,63567	17,52691427	7,07605593	0,395284887		
Rtkn	CisPt	5	26887	31,5575	4012	0,053413569	14448	26,7061	16910	0,282327	11308	17,5318	24882	0,489168	25,2651315	7,1230302	0,274969531		
Rtkn	CisPt	10	9335	10,9566	518	0,006896368	8883	16,4196	898	0,014993	14790	22,9302	1403	0,027582	16,76879956	5,99446336	0,016490516		
Rtkn	NaAsO2	0	936	1	64434	1	499	1	65451	1	682	1	58537	1	1	0	1		
Rtkn	NaAsO2	0,3125	847	0,90491	67863	1,053217246	498	0,998	70307	1,074193	543	0,79619	66497	1,135982	0,899699402	0,10100518	1,087797508		
Rtkn	NaAsO2	0,625	857	0,9156	67833	1,052751653	444	0,88978	69308	1,05893	487	0,71408	76158	1,301023	0,839818032	0,1096581	1,137568173		
Rtkn	NaAsO2	1,25	908	0,97009	59268	0,919824937	467	0,93587	70586	1,078456	452	0,66276	89983	1,537199	0,856237937	0,16843075	1,17849309		
Rtkn	NaAsO2	2,5	1520	1,62393	16728	0,259614489	659	1,32064	41991	0,641564	498	0,73021	68259	1,166083	1,224926062	0,45448625	0,689087133		
Rtkn	NaAsO2	5	1249	1,3344	3528	0,054753701	785	1,57315	4926	0,075262	704	1,03226	17234	0,294412	1,313268689	0,27106267	0,141476064		
Blvrb	CisPt	0	2301	1	46509	1	1600	1	61833	1	1966	1	42475	1	1	0	1		
Blvrb	CisPt	0,625	5594	2,43112	15441	0,332000258	2751	1,71938	29426	0,475895	3565	1,81333	21973	0,517316	1,987939486	0,38666703	0,441737025		
Blvrb	CisPt	1,25	5745	2,49674	10885	0,234040723	3195	1,99688	25868	0,418353	3721	1,89268	20763	0,488829	2,128763677	0,32290806	0,380407369		
Blvrb	CisPt	2,5	5765	2,50543	3317	0,07131953	3440	2,15	22938	0,370967	3565	1,81333	24656	0,580483	2,156252991	0,3460953	0,340923042		
Blvrb	CisPt	5	5799	2,52021	424	0,009116515	3376	2,11	9788	0,158297	3742	1,90336	16018	0,377116	2,177855225	0,31397405	0,181509938		
Blvrb	CisPt	10	2815	1,22338	303	0,006514868	1712	1,07	687	0,011111	2489	1,26602	406	0,009559	1,18646784	0,10309286	0,009061335		
Blvrb	NaAsO2	0	2462	1	45774	1	1479	1	74602	1	2047	1	44701	1	1	0	1		
Blvrb	NaAsO2	0,3125	4437	1,80219	41077	0,897387163	2103	1,42191	76612	1,026943	3424	1,67269	42719	0,955661	1,632263925	0,19333983	0,959997029		
Blvrb	NaAsO2	0,625	6185	2,51219	37702	0,82365535	2594	1,75389	73424	0,98421	3827	1,86957	43734	0,978367	2,045212732	0,40852514	0,928744088		
Blvrb	NaAsO2	1,25	8928	3,62632	32040	0,699960676	4214	2,84922	62985	0,84428	5299	2,58867	49695	1,11172	3,021402951	0,53982952	0,885320362		
Blvrb	NaAsO2	2,5	14452	5,87002	16079	0,35126928	7695	5,20284	36601	0,490617	6729	3,28725	46770	1,046285	4,786704587	1,34073028	0,62939049		
Blvrb	NaAsO2	5	22945	9,31966	2409	0,05262813	14899	10,0737	7271	0,097464	13055	6,37763	16710	0,373817	8,590327684	1,95299288	0,174636382		
Ddit3	CisPt	0	560	1	70369	1	416	1	38656	1	398	1	59513	1	1	0	1		
Ddit3	CisPt	0,625	772	1,37857	18535	0,263397235	502	1,20673	17042	0,440863	436	1,09548	32135	0,539966	1,226926528	0,14262349	0,414742096		
Ddit3	CisPt	1,25	820	1,46429	11780	0,16740326	557	1,33894	13741	0,355469	472	1,18593	28359	0,476518	1,329719223	0,13940704	0,333129915		
Ddit3	CisPt	2,5	909	1,62321	3077	0,043726641	599	1,4399	10636	0,275145	530	1,33166	26438	0,444239	1,464925474	0,14737973	0,254370194		
Ddit3	CisPt	5	2185	3,90179	222	0,003154798	814	1,95673	3042	0,078694	602	1,51256	13248	0,222607	2,457026433	1,27075502	0,101485249		
Ddit3	CisPt	10	1331	2,37679	108	0,001534767	876	2,10577	365	0,009442	718	1,80402	365	0,006133	2,095525015	0,28652019	0,00570338		
Ddit3	NaAsO2	0	615	1	54303	1	383	1	42824	1	424	1	60414	1	1	0	1		
Ddit3	NaAsO2	0,3125	580	0,94309	56438	1,039316428	418	1,09138	43705	1,020573	376	0,88679	72485	1,199805	0,973755232	0,1056868	1,086564562		
Ddit3	NaAsO2	0,625	622	1,01138	50370	0,927573062	413	1,07833	44158	1,031151	375	0,88443	82770	1,370047	0,991381686	0,09848265	1,109590166		
Ddit3	NaAsO2	1,25	615	1	38672	0,712152183	476	1,24282	41729	0,97443	401	0,94575	85584	1,416625	1,062858187	0,15819382	1,034402565		
Ddit3	NaAsO2	2,5	1050	1,70732	11543	0,212566525	534	1,39426	23562	0,550205	530	1,25	65954	1,091701	1,450524316	0,23379335	0,618157539		
Ddit3	NaAsO2	5	2765	4,49593	4230	0,077896249	2368	6,18277	3239	0,075635	1424	3,35849	25035	0,414391	4,679064383	1,42101637	0,18930737		
Wt	CisPt	0	389	1	75874	1	234	1	48269	1	258	1	63864	1	1	0	1		
Wt	CisPt	0,625	529	1,3599	26928	0,354904183	326	1,39316	30448	0,630798	333	1,2907	29678	0,464706	1,34791908	0,05227199	0,483469556		
Wt	CisPt	1,25	536	1,37789	23176	0,305453779	357	1,52564	24166	0,500653	355	1,37597	29132	0,456157	1,426500683	0,08586344	0,420754402		
Wt	CisPt	2,5	504	1,29563	17953	0,236615969	328	1,40171	19876	0,411776	343	1,32946	24579	0,384865	1,342265529	0,05418723	0,344418785		
Wt	CisPt	5	468	1,20308	8735	0,115125076	310	1,32479	12332	0,255485	344	1,33333	14872	0,23287	1,287068164	0,07285714	0,201159937		
Wt	CisPt	10	312	0,80206	1868	0,024619764	217	0,92735	2098	0,043465	283	1,0969	1643	0,025727	0,942102069	0,14797384	0,031270353		
Wt	NaAsO2	0	395	1	82189	1	239	1	59351	1	256	1	55612	1	1	0	1		
Wt	NaAsO2	0,3125	345	0,87342	83281	1,01328645	244	1,02092	58066	0,978349	231	0,90234	62319	1,120603	0,932227325	0,07816034	1,03741302		
Wt	NaAsO2	0,625	361	0,91392	74965	0,912105026	239	1	55585	0,936547	217	0,84766	67369	1,211411	0,920526767	0,0763862	1,020021072		
Wt	NaAsO2	1,25	366	0,92658	61652	0,750124713	240	1,00418	53786	0,906236	217	0,84766	61454	1,105049	0,926140876	0,07826486	0,920469922		
Wt	NaAsO2	2,5	373	0,9443	19395	0,235980484	251	1,05021	35078	0,591026	242	0,94531	51948	0,934115	0,979941834	0,06085542	0,587040564		
Wt	NaAsO2	5	431	1,09114	7669	0,093309324	333	1,39331	7240	0,121986	285	1,11328	19814	0,35629	1,199241977	0,16842814	0,190528494		

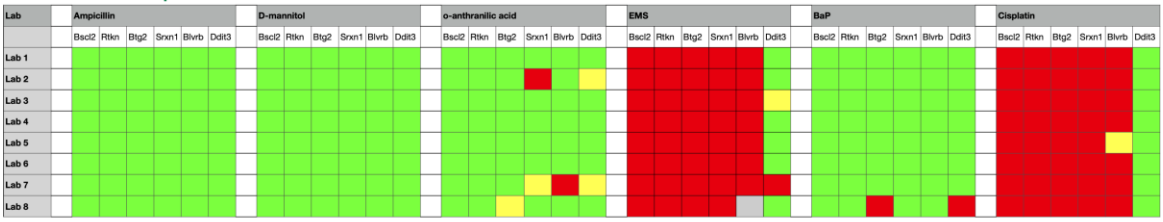


Annex 4: Example data set from ToxTracker installation training

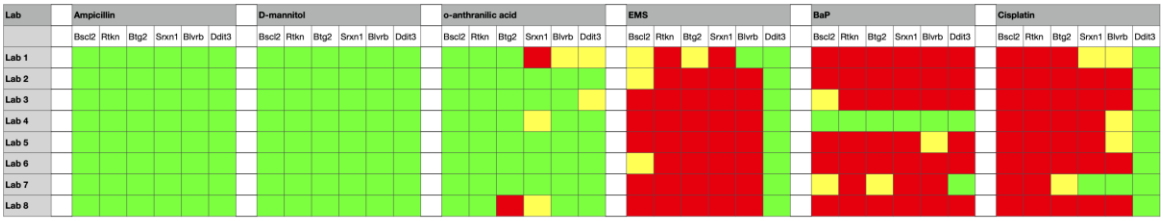


Annex 5: Results from phase 2 proficiency testing of the ToxTracker validation

Validation compounds -S9



Validation compounds +S9



Annex 6: Summary of the ToxTracker validation results for all reporters

35	2-Phenylphenol sodium salt tetrahydrat	6152-33-6	LAB 2	P	N	N	N	N	3	N	Yes		P	P	N	P	3	P	Yes		P	P	3	P	Yes		N	N	3	P	No	
35	2-Phenylphenol sodium salt tetrahydrat	6152-33-6	LAB 3	R	N	na	N	na	2*	na	Yes	No	N	na	N	na	3*	I	No	Yes	na	na	1*	I	I	Yes	N	na	2*	I	Yes	No
35	2-Phenylphenol sodium salt tetrahydrat	6152-33-6	LAB 5	P	N	N	N	E	3	E	No		P	P	N	P	3	P	Yes		P	P	3	P	Yes		N	N	3	N	No	
36	Ropinirole hydrochloride	91374-20-8	LAB 2	P	na	na	I	N	3	I	No		N	N	N	N	3	I	No		N	N	3	N	Yes		N	N	3	N	No	
36	Ropinirole hydrochloride	91374-20-8	LAB 6	P	N	N	N	N	3	N	Yes	Yes	N	N	N	N	3	N	Yes	Yes	N	N	3	N	Yes	Yes	N	N	3	N	Yes	Yes
36	Ropinirole hydrochloride	91374-20-8	LAB 7	P	N	N	N	N	3	N	Yes		N	N	N	N	3	N	Yes		N	N	3	N	Yes		N	N	3	N	Yes	
37	Methyl carbamate (Methylurethane)	598-55-0	LAB 2	P	N	N	N	N	3	N	Yes		N	N	N	N	3	N	Yes		N	N	3	N	Yes		N	N	3	N	Yes	
37	Methyl carbamate (Methylurethane)	598-55-0	LAB 6	P	N	N	N	N	3	N	Yes	Yes	N	N	N	N	3	N	Yes	Yes	N	N	3	N	Yes	Yes	N	N	3	N	Yes	Yes
37	Methyl carbamate (Methylurethane)	598-55-0	LAB 7	P	N	N	N	N	3	N	Yes		N	N	N	N	3	N	Yes		N	N	3	N	Yes		N	N	3	N	Yes	
38	Cyclosporin A (CsA)	59865-13-3	LAB 1	P	N	N	N	N	3	N	Yes		P	P	P	P	3	P	Yes		P	P	3	P	Yes		N	N	3	N	Yes	
38	Cyclosporin A (CsA)	59865-13-3	LAB 3	N	N	na	na	na	0	na	na	na	P	P	N	N	3	P	Yes	Yes	P	P	3	P	Yes	Yes	na	na	0	I	I	
38	Cyclosporin A (CsA)	59865-13-3	LAB 4	R	N	na	N	na	3*	na	No		P	na	P	na	3*	P	Yes		P	na	2*	P	Yes		N	na	3*	I	Yes	
39	Sodium saccharin	128-44-9	LAB 2	P	N	N	N	N	3	N	Yes		P	P	P	P	3	P	No		N	N	3	N	Yes		N	N	3	N	Yes	
39	Sodium saccharin	128-44-9	LAB 3	P	N	N	N	N	2	N	Yes	Yes	P	P	P	P	3	P	Yes	Yes	N	N	3	N	Yes	Yes	N	N	3	N	No	Yes
39	Sodium saccharin	128-44-9	LAB 5	P	N	N	N	N	3	N	Yes		P	P	P	P	3	P	Yes		N	N	3	N	Yes		N	N	3	N	Yes	
40	Diethanolamine (DEA or DEOA)	111-42-2	LAB 1	P	N	N	N	N	3	N	Yes		N	N	N	N	3	N	Yes		N	N	3	N	Yes		N	N	3	N	Yes	
40	Diethanolamine (DEA or DEOA)	111-42-2	LAB 3	P	N	N	N	N	3	N	Yes	Yes	N	N	N	N	3	N	Yes	Yes	N	N	2	N	Yes	Yes	N	N	3	N	Yes	Yes
40	Diethanolamine (DEA or DEOA)	111-42-2	LAB 6	P	N	N	N	N	3	N	Yes		N	N	N	N	3	N	Yes		N	N	3	N	Yes		N	N	3	N	Yes	
41	Hexachloroethane	67-72-1	LAB 3	P	N	N	N	N	2	N	Yes		N	N	N	N	2	N	Yes		N	N	2	N	Yes		N	N	2	N	Yes	
41	Hexachloroethane	67-72-1	LAB 6	P	N	N	N	N	3	N	Yes	Yes	N	P	N	N	2	P	Yes	No	N	N	3	N	Yes	Yes	N	N	3	N	Yes	Yes
41	Hexachloroethane	67-72-1	LAB 7	P	N	N	N	N	3	N	Yes		N	N	N	N	3	N	No		N	N	3	N	No		N	N	3	N	Yes	
42	Melamine	108-78-1	LAB 1	P	P	P	P	P	3	P	Yes		N	N	N	N	3	N	Yes		N	N	3	N	Yes		N	N	3	N	Yes	
42	Melamine	108-78-1	LAB 6	P	N	N	N	N	3	N	No	No	N	P	N	N	3	P	No	No	N	N	3	N	Yes	Yes	N	N	3	N	Yes	Yes
42	Melamine	108-78-1	LAB 4	P	N	na	N	N	2	I	No		N	N	N	N	3	N	No		N	N	3	N	Yes		N	N	2	N	Yes	
43	Tunicamycin	11089-65-9	LAB 1	P	N	N	N	N	3	N	Yes		P	P	P	E	3	P	Yes		P	P	3	P	Yes		N	N	3	N	Yes	
43	Tunicamycin	11089-65-9	LAB 6	P	N	N	N	N	3	N	Yes	Yes	N	N	N	N	3	N	No	No	P	P	3	P	Yes	Yes	N	N	3	N	Yes	Yes
43	Tunicamycin	11089-65-9	LAB 7	P	N	N	N	N	3	N	No		na	na	I	N	3	I	No		na	na	1	I	I		N	N	3	N	Yes	
44	p-Nitrophenol (4-nitrophenol)	100-02-7	LAB 2	P	N	N	N	N	3	N	Yes		N	N	N	N	3	N	No		P	P	3	P	Yes		N	N	3	N	Yes	
44	p-Nitrophenol (4-nitrophenol)	100-02-7	LAB 6	P	N	P	N	P	3	P	Yes	No	N	P	N	P	3	P	Yes	No	N	N	3	N	No	No	N	N	3	N	No	Yes
44	p-Nitrophenol (4-nitrophenol)	100-02-7	LAB 7	P	N	N	N	N	3	N	Yes		N	N	N	N	3	N	No		P	N	3	P	Yes		N	N	3	N	Yes	
45	Phenanthrene	85-01-8	LAB 1	P	N	N	N	N	3	N	Yes		P	P	N	E	3	P	Yes		N	P	3	P	Yes		N	N	3	N	Yes	
45	Phenanthrene	85-01-8	LAB 2	P	N	N	N	N	3	N	Yes	Yes	P	P	P	P	3	P	Yes	Yes	P	P	3	P	Yes	No	N	N	3	N	No	Yes
45	Phenanthrene	85-01-8	LAB 3	P	N	N	N	N	3	N	Yes		N	P	N	N	3	P	Yes		N	N	3	N	Yes		N	N	3	N	Yes	
46	Tertiarybutylhydroquinone	1948-33-0	LAB 1	P	N	N	E	P	3	P	Yes		P	P	P	P	3	P	Yes		P	P	3	P	Yes		P	P	3	P	Yes	
46	Tertiarybutylhydroquinone	1948-33-0	LAB 6	P	N	N	N	N	3	N	No	No	P	P	P	P	3	P	Yes	Yes	P	P	3	P	Yes	Yes	P	P	3	P	Yes	No
46	Tertiarybutylhydroquinone	1948-33-0	LAB 4	P	N	N	N	N	3	N	Yes		P	P	P	P	3	P	Yes		P	N	3	P	No		N	N	3	N	Yes	
47	Benzyl alcohol	100-51-6	LAB 2	P	N	N	N	N	3	N	Yes		N	N	N	N	3	N	No		N	N	3	N	No		N	N	3	N	Yes	
47	Benzyl alcohol	100-51-6	LAB 3	N	N	na	na	na	0	na	na	Yes	P	na	N	na	3*	P*	No	No	N	na	3*	I	Yes	Yes	na	na	1*	I	I	Yes
47	Benzyl alcohol	100-51-6	LAB 5	P	N	N	N	N	3	N	Yes		N	N	N	N	3	N	No		N	N	3	N	Yes		N	N	3	N	Yes	
48	Vanillin	121-33-5	LAB 1	P	N	N	N	P	3	P	Yes		P	N	N	P	3	P	Yes		N	N	3	N	Yes		N	N	3	N	No	
48	Vanillin	121-33-5	LAB 3	P	N	N	N	N	2	N	Yes	No	N	N	N	N	3	N	No	No	N	N	3	N	Yes	Yes	N	N	3	N	Yes	No
48	Vanillin	121-33-5	LAB 7	P	N	N	N	N	3	N	No		P	P	N	P	3	P	Yes		N	N	3	N	Yes		N	N	3	P	No	
49	Erythromycin (Erythromycin stearate 6	114-07-8	LAB 1	P	N	N	N	N	3	N	Yes		P	P	N	N	3	P	Yes		P	N	3	P	Yes		N	N	3	N	Yes	
49	Erythromycin (Erythromycin stearate 6	114-07-8	LAB 6	P	N	N	N	N	3	N	Yes	No	N	N	N	N	3	N	No	No	P	N	3	P	No	Yes	N	N	3	N	Yes	Yes
49	Erythromycin (Erythromycin stearate 6	114-07-8	LAB 7	P	N	N	N	P	3	P	No		P	P	N	N	3	P	No		na	na	1	I	I		N	N	3	N	Yes	
50	Sodium diclofenac	15307-79-6	LAB 3	R	N	na	N	na	2*	na	Yes		N	na	N	na	3*	I	No		na	na	1	I	I		N	na	2*	I	Yes	
50	Sodium diclofenac	15307-79-6	LAB 6	P	N	N	N	P	3	P	Yes	No	P	N	E	N	2	P	Yes	Yes	N	N	3	N	Yes	No	P	P	3	P	Yes	Yes
50	Sodium diclofenac	15307-79-6	LAB 7	P	N	N	N	N	3	N	No		P	P	N	N	3	P	Yes		P	P	3	P	Yes		P	E	3	P	Yes	
51	o-anthranilic acid	118-92-3	LAB 1	P	N	N	N	N	3	N	Yes		N	N	N	N	3	N	Yes		N	N	3	N	Yes		N	N	3	N	Yes	
51	o-anthranilic acid	118-92-3	LAB 3	P	N	N	N	N	2	N	Yes	Yes	N	N	N	N	3	N	Yes	Yes	N	N	3	N	Yes	Yes	N	N	3	N	Yes	Yes
51	o-anthranilic acid	118-92-3	LAB 7	P	N	N	N	N	3	N	Yes		N	N	N	N	3	N	No		N	N	3	N	Yes		N	N	3	N	Yes	
52	Tolbutamide	64-77-7	LAB 1	P	N	N	N	N	3	N	Yes		P	P	N	N	3	P	Yes		P	P	3	P	Yes		N	N	3	N	Yes	
52	Tolbutamide	64-77-7	LAB 3	P	N	N	N	N	2	N	Yes	Yes	P	P	N	N	2	P	Yes	Yes	P	P	2	P	Yes	Yes	N	P	2	P	Yes	No
52	Tolbutamide	64-77-7	LAB 4	R	N	na	N	na	3*	na	Yes		P	na	N	na	2*	P	Yes		N	na	3*	I	No		N	na	1*	I	I	
53	2-ethyl-1,3-hexanediol	94-96-2	LAB 1	P	N	N	N	N	3	N	Yes		P	P	N	N	3	P	Yes		N	N	3	N	No		N	N	3	N	Yes	
53	2-ethyl-1,3-hexanediol	94-96-2	LAB 2	P	N	N	N	N	3	N	Yes	No	N	P	N	N	3	P	No	Yes	N	N	3	N	Yes	No	N	N	3	N	Yes	Yes
53	2-ethyl-1,3-hexanediol	94-96-2	LAB 3	P	E	N	N	N	3	E	Yes		P	P	N	N	3	P	Yes		E	E	2	E	No		N	N	3	N	Yes	
54	Chlorpheniramine maleate	113-92-8	LAB 1	P	N	N	N	N	3	N	Yes		P	P	P	P	3	P	Yes		N	N	3	N	Yes		N	N	3	N	Yes	
54	Chlorpheniramine maleate	113-92-8	LAB 2	P	N	N	N	N	3	N	Yes	Yes	P	P	P	P	3	P	Yes	Yes	N	N	3	N	Yes	Yes	N	N	3	N	Yes	Yes
54	Chlorpheniramine maleate	113-92-8	LAB 6	P	N	N	N	N	3	N	Yes		P	N	N	N	2	P	Yes		N	N	3	N	Yes		N	N	3	N	Yes	
55	Ampicillin trihydrate	7177-48-2	LAB 7	P	N	N	N	N	3	N	Yes		N	N	N	N	3	N	Yes		N	N	3	N	Yes		N	N	3	N	Yes	
55	Ampicillin trihydrate	7177-48-2	LAB 4	R	N	na	N	na	3*	na	Yes	Yes	N	na	N	na	3*	I	Yes	Yes	N	na	2*	I	Yes	Yes	N	na	3*	I	Yes	Yes
55	Ampicillin trihydrate	7177-48-2	LAB 5	P	N	N	N	N	3	N	Yes		N	N	N	N	3	N	Yes		N	N	3	N	Yes		N	N	3	N	Yes	
56	Sodium chloride	7647-14-5	LAB 7	P	N	N	N	N	3	N	Yes		N	N	N	N	3	N	Yes		N	N	3	N	Yes		N	N	3	N	Yes	
56	Sodium chloride	7647-14-5	LAB 4	P	N	na	N	N	2	N	Yes	Yes	N	N	N	N	3	N	No	Yes	N	N	3	N	Yes	Yes	N	N	2	N	No	Yes
56	Sodium chloride	7647-14-5	LAB 5	P	N	N	N	N	3	N	Yes		N	N	N	N	3	N	Yes		N	N	3	N	Yes		N	N	3	N	Yes	
57	D-mannitol	69-65-8	LAB 1	P	N	N	N	N	3	N	Yes		N	N	N	N	3	N	Yes		N	N	3	N	Yes		N	N	3	N	Yes	
57	D-mannitol	69-65-8	LAB 3	R	N	na	N	na	2*	na	Yes	na	N	na	N	na	3*	I	Yes	Yes	na	na	1*	I	I	Yes	N	na	1	I	I	
57	D-mannitol	69-65-8	LAB 4	R	N	na	N	na	3*	na	Yes		N	na	N	na	2*	I	Yes		N	na	3*	I	Yes		N	na	3*	I	Yes	
58	Allyl alcohol	107-18-6	LAB 6	P	N	N	N	N	3	N	Yes		N	N	N	N	3	N	Yes		N	N	3	N	Yes		N	N	3	N	Yes	
58	Allyl alcohol	107-18-6	LAB 4	R	N	na	N	na	2*	na	Yes	Yes	N	na	P	na	3*	P	Yes	No	N	na	3*	I	Yes	Yes	N	na	2*	I	Yes	Yes
58	Allyl alcohol	107-18-6	LAB 5	P	N	N	N	N	3	N	No		P	P	P	P	3	P	Yes		N	N	3	N	Yes		N	N	3	N	Yes	
59	(2-chloroethyl)trimethyl-ammonium chl	999-81-5	LAB 6	P	N	N	N	N	3	N	Yes		N	N	N	N	3	N	Yes		N	N	3	N	Yes		N	N	3	N	Yes	
59	(2-chloroethyl)trimethyl-ammonium chl	999-81-5	LAB 4	P	N	na	N	N	2	I	No	Yes	N	N	N	N	3	N	Yes	Yes	N	N	3	N	No	Yes	N	N	2	I	No	Yes
59	(2-chloroethyl)trimethyl-ammonium chl	999-81-5	LAB 5	P	N	N	N	N	3	N	Yes		N	N	N	N	3	N	Yes		N	N	3	N	Yes		N	N	3	N	Yes	
60	Sulfisoxazole	127-69-5	LAB 6	P	N	N	N	N	3	N	Yes		N	N	N	N	3	N	No		N	N	3	N	Yes		N	N	3	N	Yes	
60	Sulfisoxazole	127-69-5	LAB 4	R	N	na	N	na	3*	na	Yes	Yes	N	na	N	na	3*	I	Yes	No	N	na	3*	I	Yes	Yes	N	na	2*	I	Yes	Yes
60	Sulfisoxazole	127-69-5	LAB 5	P	N	N	N	N	3	N	Yes		P	P	N	N	3	P	Yes		N	N	3	N	Yes		N	N	3	N	Yes	
61	Sucrose	57-50-1	LAB 7	P	N	N	N	N	3	N	Yes		N	N	N	N	3	N	Yes		N	N	3	N	Yes		N	N	3	N	Yes	
61	Sucrose	57-50-1	LAB 4	R	N	na	N	na	3*	na	Yes	Yes	N	na	N	na	2*	I	Yes	Yes	N	na	3*	I	Yes	Yes	N	na	3*	I	Yes	Yes
61	Sucrose	57-50-1	LAB 5	P	N	N	I	N	3	N	Yes		N	N	N	N	3	N	No		N	N	3	N	Yes		N	N	3	N	Yes	
62	Cyclohexanone	108-94-1	LAB 1	P	N	N	N	N	3	N	Yes		N	N	N	N	3	N	Yes		N	N	3	N	Yes		N	N	3	N	Yes	
62	Cyclohexanone	108-94-1	LAB 3	P	N	N	N	N	2	N	No	No	N	N	N	N	2	N	Yes	Yes	N	N	2	N	Yes	Yes	N	N	2	N	Yes	Yes
62	Cyclohexanone	108-94-1	LAB 4	R	N	na	N	na	3*	na	Yes		N	na	N	na	2*	I	Yes		N	na	3*	I	Yes		N	na	3*	I	Yes	
63	1-Nitropropane	108-03-2	LAB 2	P	N	N	N	N	3	N	Yes		N	N	N	N	3	N	Yes		N	N	3	N	Yes		N	N	3	N	Yes	
63	1-Nitropropane	108-03-2	LAB 3	P	N	N	N	N	2	N	Yes	Yes	N	N	N	N	2	N	Yes	Yes	N	N	2	N	Yes	Yes	N	N	2	N	Yes	Yes
63	1-Nitropropane	108-03-2	LAB 5	P	N	N	N	N	3	N	Yes		N	N	N	N	3	N	Yes		N	N	3	N	Yes		N	N	3	N	Yes	
64	Phenformin HCl	834-28-6	LAB 2	P	N	N	N	N	3	N	Yes		N	N	N	E	3	E	No		P	P	3	P	Yes		N	N	3	N	No	
64	Phenformin HCl	834-28-6	LAB 6	P	N	N	N	N	3	N	No	Yes	N	P	N	P	3	P	Yes	No	P	P	3	P	Yes	Yes	N	N	3	N	Yes	Yes
64	Phenformin HCl	834-28-6	LAB 7	P	N	N	N	N	3	N	No		N	N	N	N	3	E	No		P	P	3	P	Yes		N	N	3	N	Yes	

CONTROLS

P	Acceptable: proof that reporter cell lines provide responses within the historical control range; proof that the S9 metabolising system works sufficiently
R	Acceptable with restrictions: proof that reporter cell lines provide responses within the historical control range, but only acceptable data without S9
N	Not acceptable: insufficient evidence that the reporter cell lines show responses within the historical control range
*	Only data for test -S9