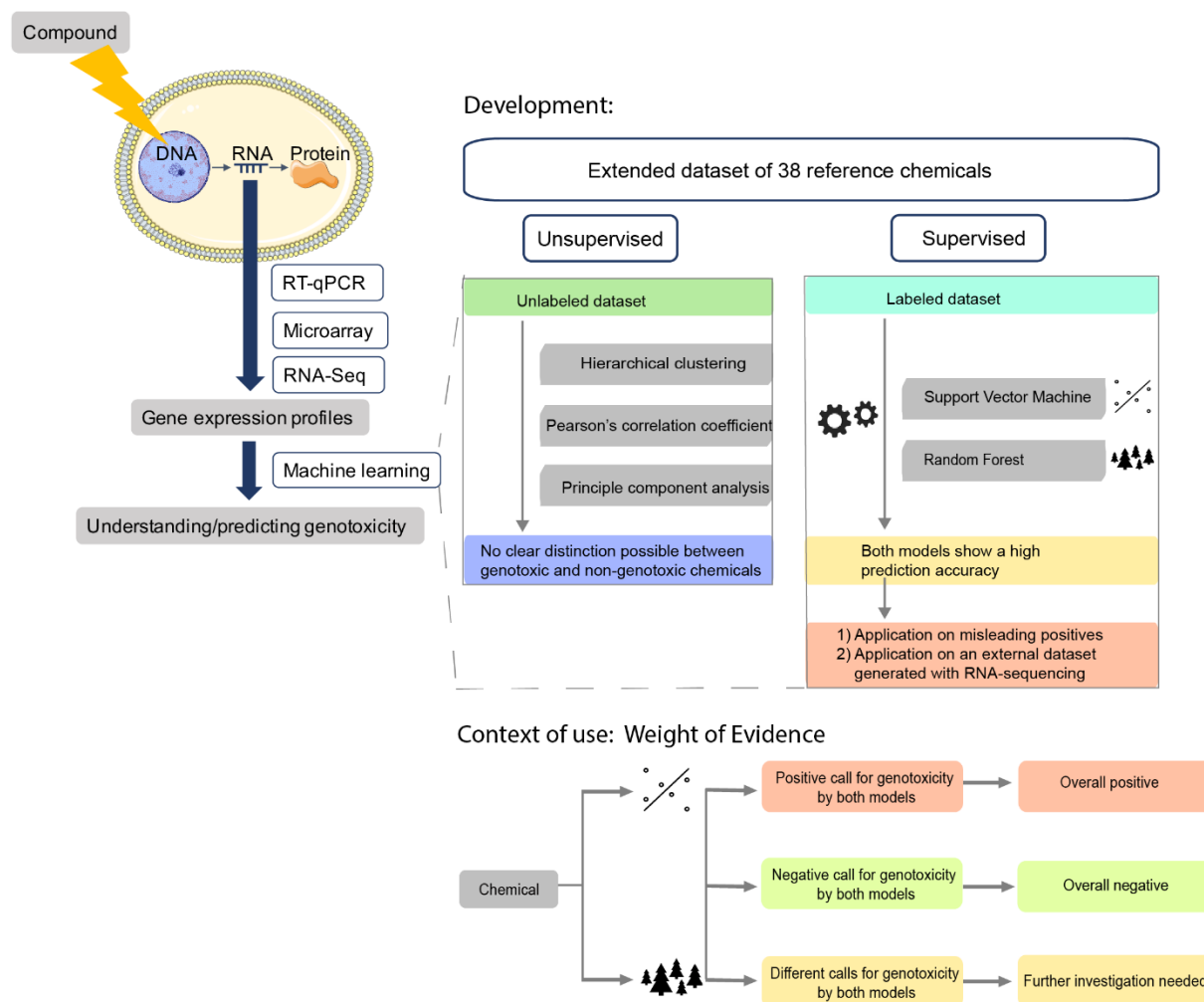


Novel Prediction Models for Genotoxicity Based on Biomarker Genes in Human HepaRG™ Cells

Supplementary Data



Graphical abstract

Tab. S1: List of 19 genotoxic reference chemicals with main mechanism of action

Chemical name	In vitro genotoxicity		In vivo genotoxicity	Main mechanism of action	CAS number	References
	Ames	MNvit/CAvit				
Aflatoxin B (AFB) (pro-gtx)	+	+	+	Bulky adduct formation	1162-65-8	Vinken et al., 2008; Kirkland et al., 2016
Benzoapyrene (BAP) (pro-gtx)	+	+	+	Polycyclic aromatic hydrocarbon, bulky adduct formation	50-32-8	Vinken et al., 2008; Kirkland et al., 2016
Bleomycin sulfate (BLE)	+	+	+	Radical generator causing DNA strand breaks (radiomimetic)	9041-93-4	Miller, M. G. et al., 1986; Mozdarani and Saberi, 1994; Kirkland et al., 2008; Miller, K., 1991
Cadmium chloride (CAD)	-	+	+	DNA repair inactivator, cell cycle inducer, p53 inhibitor	10108-64-2	Kirkland et al., 2016
Cisplatin (CIS)	+	+	+	Cross-linking agent	15663-27-1	Kirkland et al., 2016
Cyclophosphamide (CPM) (pro-gtx)	+	+	+	Alkylating agent	50-18-0	Vinken et al., 2008; Kirkland et al., 2016
Dimethylnitrosamine (DMN) (pro-gtx)	+	+	+	Alkylating agent	62-75-9	Vinken et al., 2008; Kirkland et al., 2016
Methyl methanesulfonate (MMS)	+	+	+	N7 alkylation, replication fork impairment	66-27-3	Kirkland et al., 2016
2-Nitrofluorene (2-NF)	+	+	+	Bulky adduct formation	607-57-8	Vinken et al., 2008
4-Methylnitrosamino-1-3-pyridyl-1-butanone (NNK) (pro-gtx)	+	+	+	Bulky adduct formation	64091-91-4	Vinken et al., 2008
Vinblastine sulfate (VIN)	-	+	+	tubulin polymerization inhibitor	143-67-9	Kirkland et al., 2016
Zidovudine (ZID)	+	+	+	Antimetabolite, nucleoside analogue	30516-87-1	Kirkland et al., 2016
Glycidol (GLY)	+	+	+	Clastogen that binds to DNA	556-52-5	Mertens et al, 2017; IARC, 2000
4-Aminophenol (4AP)	-	+	+	Alkylating agent	123-30-8	Mertens et al., 2017; EFSA, 2007; SCCP, 2005; SCCS, 2011b
2,4 Diaminotoluene (DAT) (pro-gtx)	+	+	+	DNA adduct formation	95-80-7	Kirkland et al., 2016
Ethyl methanesulfonate (EMS)	+	+	+	Alkylating agent	62-50-0	Miller, 1991; Kamber, et al. 2009
1-Ethyl-1-nitrosourea (ENU)	+	+	+	Alkylating agent	759-73-9	Kirkland et al., 2016
Etoposide (ETO)	+	+	+	Topoisomerase II inhibitor	33419-42-0	Kirkland et al., 2016
Chloramphenicol (CAM)	-	+	+	Clastogen that binds to DNA	56-75-7	Kirkland et al., 2008

Tab. S2: List of 19 non-genotoxic reference chemicals with applicability domain

Chemical name	In vitro genotoxicity		In vivo genotoxicity	Main mechanism of action	CAS number	References
	Ames	MNvit/CAvit				
Ampicillin trihydrate (AMP)	-	-	-	Drug/beta-lactam antibiotic	7177-48-2	Kirkland et al., 2016
Caprolactam (CAP)	-	-	-	Synthetic polymer	105-60-2	Greene et al., 1979; Hastwell et al., 2006
Clonidine (CND)	-	-	-	Drug/ α 2 adrenergic agonist	4205-90-7	Vinken et al., 2008
D-Mannitol (DMO)	-	-	-	Sweetener	69-65-8	Vinken et al., 2008; Kirkland et al., 2016
Hydroxybenzomorpholine (HBM)	+	-	-	Oxidative hair dye	26021-57-8	SCCP, 2006
1-naphthol (NAP)	-	+	-	Oxidative hair dye/ insecticide	90-15-3	SCCP, 2008
Sodium chloride (NCL)	-	-	-	Condiment/food preservative	7647-14-5	Hastwell et al., 2006
Nifedipine (NFE)	-	-	-	Drug/antihypertensive	21829-25-4	Vinken et al., 2008
Sodium diclofenac (SDF)	-	-	-	Non-steroidal anti-inflammatory drug	15307-79-6	Vinken et al., 2008
Sorbitol (SOR)	-	-	-	Sweetener/sugar alcohol	50-70-4	Pottenger et al., 2007
Tolbutamide (TBT)	-	-	-	Drug/hypoglycemic potassium channel blocker	64-77-7	Vinken et al., 2008
Triclosan (TRI)	-	+	-	Cosmetic preservative	3380-34-5	SCCS, 2011a
4-Methyl-2-pentanol (4M2P)	-	NA	NA	Food contact material	108-11-2	EFSA, 2017
Melamine (MELA)	-	-	-	Additive in resins, plasticizers, paint	108-78-1	Kirkland et al., 2016
2-Methyl-1-propanol (2M1P)	-	-	-	Food contact material	78-83-1	Api et al., 2019; ECHA, 2011
Phthalimide (PHTH)	-	-	NA	Food contact material	85-41-6	ECHA, 2007
4-Chlororesorcinol (4CR)	-	+	-	Oxidative hair dye component	95-88-5	SCCS, 2010b
Anthranilic acid (ANT)	-	+	-	Intermediate in azo dyes and saccharin production	118-92-3	Kirkland et al., 2016
Basic orange 31 (BOR)	+	-	-	(Oxidative) hair dye component	97404-02-9	SCCS, 2010a

Tab. S3: List of down- and upregulated GENOMARK biomarker genes and corresponding ensembl gene IDs according to HC analysis after exposure to a genotoxicant

Downregulated genes		Upregulated genes	
GENOMARK gene	Ensembl gene ID	GENOMARK gene	Ensembl gene ID
MUC15	ENSG00000169550	ABCA12	ENSG00000144452
GPD1	ENSG00000167588	ADORA2B	ENSG00000170425
TNFRSF11B	ENSG00000164761	AKR1B1	ENSG00000085662
ENSG0261578	ENSG00000261578	AREG	ENSG00000109321
FOLH1	ENSG00000086205	ARL14	ENSG00000179674
RARRES1	ENSG00000118849	CCDC178	ENSG00000166960
HLF	ENSG00000108924	CDIP1	ENSG00000089486
PPP1R1A	ENSG00000135447	CEL	ENSG00000170835
PALMD	ENSG00000099260	CLCA2	ENSG00000137975
MRC1	ENSG00000260314	DGKA	ENSG00000065357
PPP1R3B	ENSG00000173281	DRAXIN	ENSG00000162490
PIK3C2G	ENSG00000139144	DSE	ENSG00000111817
C19ORF80	ENSG00000130173	DUSP14	ENSG00000276023
EXT1	ENSG00000182197	E2F7	ENSG00000165891
SPP2	ENSG00000072080	ENSG0260912	ENSG00000260912
CYP39A1	ENSG00000146233	FDXR	ENSG00000161513
ADH4	ENSG00000198099	FOSL1	ENSG00000175592
C8B	ENSG00000021852	G6PC	ENSG00000131482
FREM2	ENSG00000150893	GDF15	ENSG00000130513
GPAM	ENSG00000119927	GLIPR2	ENSG00000122694
C10ORF11	ENSG00000148655	GPR87	ENSG00000138271
NCKAP5	ENSG00000176771	HERC5	ENSG00000138646
THRSP	ENSG00000151365	IL12A	ENSG00000168811
AQP9	ENSG00000103569	KRT17	ENSG00000128422
HSD11B1	ENSG00000117594	LAMP3	ENSG00000078081
G0S2	ENSG00000123689	LINC01503	ENSG00000233901
PAGE4	ENSG00000101951	MFSD2A	ENSG00000168389
ONECUT1	ENSG00000169856	MICALL1	ENSG00000100139
SLC22A7	ENSG00000137204	NBEAP1	ENSG00000258590
SLC39A11	ENSG00000133195	PMAIP1	ENSG00000141682
ADH1C	ENSG00000248144	PPP1R14C	ENSG00000198729
CFHR3	ENSG00000116785	PROX1	ENSG00000117707
SLC25A33	ENSG00000171612	PVRL4	ENSG00000143217
ENSG0261051	ENSG00000261051	RRAD	ENSG00000166592
TTPA	ENSG00000137561	S100A2	ENSG00000196754
HFE2	ENSG00000168509	SFN	ENSG00000175793
CYP2A6	ENSG00000255974	SPAG1	ENSG00000104450
INHBE	ENSG00000139269	SYNC	ENSG00000162520
CYP8B1	ENSG00000180432	SYTL1	ENSG00000142765
ENSG0259347	ENSG00000259347	TMEM185B	ENSG00000226479
LIPC	ENSG00000166035	TREM2	ENSG00000095970
SLC22A1	ENSG00000175003	VSIG1	ENSG00000101842

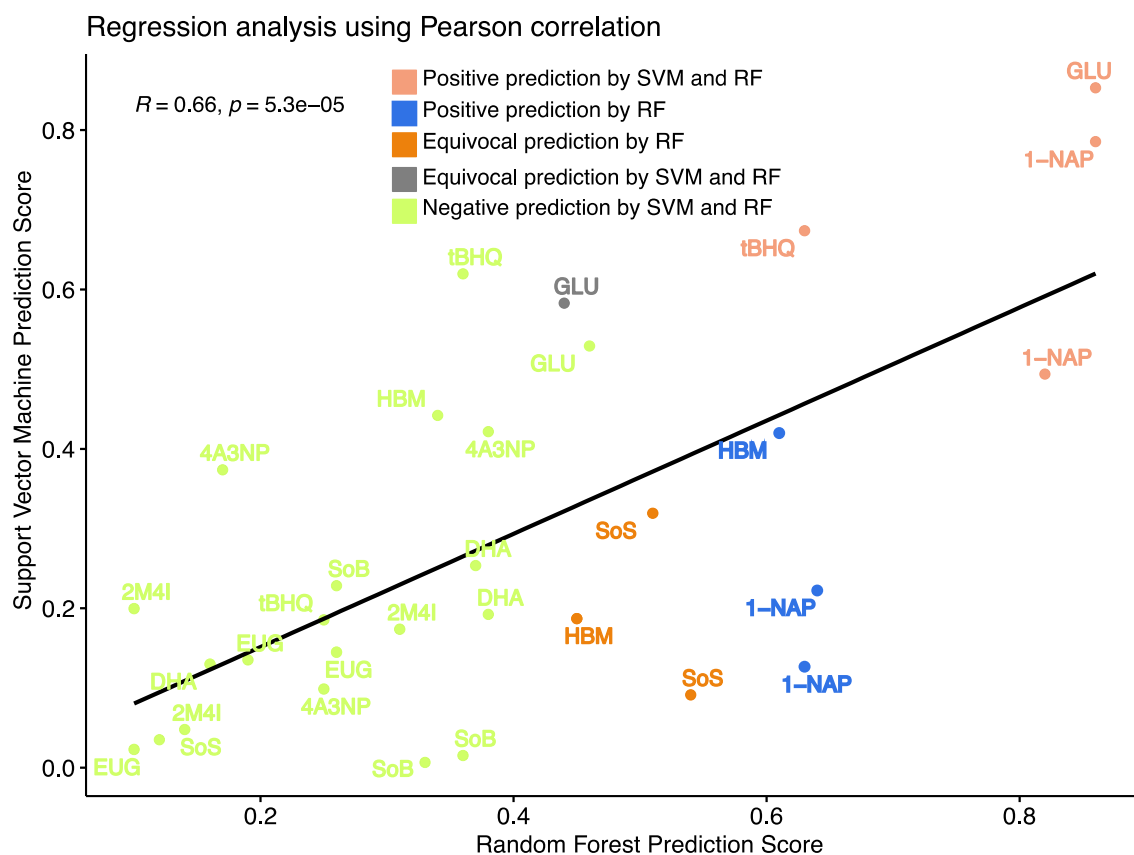


Fig. S1: Pearson's correlation analysis on the individual predictions for the 10 "misleading positives" using support vector machine (SVM) and random forest (RF)

The dots represent the predictions made by SVM and RF. HBM, 2M4I, 4A3N, SoB, DHA, tBHQ, GLU, SoS and EUG (n = 3) and for 1-NAP (n = 4).

References

- Api, A. M., Belmonte, F., Belsito, D. et al. (2019). RIFM fragrance ingredient safety assessment, isobutyl alcohol, CAS Registry Number 78-83-1. *Food Chem Toxicol* 134, Suppl 2, 110999. doi:10.1016/j.fct.2019.110999
- ECHA (2007). Registration Dossier phthalimide. *European Chemicals Agency*. <https://echa.europa.eu/da/registration-dossier/-/registered-dossier/13146/7/7/2>
- ECHA (2011). Registration Dossier isobutyl alcohol. *European Chemicals Agency*. <https://echa.europa.eu/lv/registration-dossier/-/registered-dossier/15092/7/7/1>
- EFSA (2007). Opinion of the scientific panel on food additives, flavourings, processing aids and materials in contact with food (AFC) on the food colour Red 2G (E128) based on a request from the commission related to the re-evaluation of all permitted food additives. *EFSA J* 5, 515. doi:10.2903/j.efsa.2007.515
- EFSA (2017). Scientific Opinion on Flavouring Group Evaluation 7, Revision 5 (FGE.07Rev5): Saturated and unsaturated aliphatic secondary alcohols, ketones and esters of secondary alcohols and saturated linear or branched-chain carboxylic acids from chemical group 5. *EFSA J* 15, e04725. doi:10.2903/j.efsa.2017.4725
- Greene, E. J., Friedman, M. A. and Sherrod, J. A. (1979). In vitro mutagenicity and cell transformation screening of caprolactam. *Environ Mutagen* 1, 399-407. doi:10.1002/em.2860010413
- Hastwell, P. W., Chai, L.-L., Roberts, K. J. et al. (2006). High-specificity and high-sensitivity genotoxicity assessment in a human cell line: Validation of the greenscreen HC GADD45a-GFP genotoxicity assay. *Mutat Res* 607, 160-175. doi:10.1016/j.mrgentox.2006.04.011
- IARC (2000). Glycidol. Some industrial chemicals. In I. A. F. R. o. Cancer, *IARC Monogr Eval Carcinog Risks Hum* 77. Lyon, France: IARCPress. <https://publications.iarc.fr/Book-And-Report-Series/Iarc-Monographs-On-The-Identification-Of-Carcinogenic-Hazards-To-Humans/Some-Industrial-Chemicals-2000>
- Kamber, M., Fluckiger-Isler, S., Engelhardt, G. et al. (2009). Comparison of the Ames II and traditional Ames test responses with respect to mutagenicity, strain specificities, need for metabolism and correlation with rodent carcinogenicity. *Mutagenesis* 24, 359-366. <https://doi.org/10.1093/mutage/geb017>
- Kirkland, D., Kasper, P., Müller, L. et al. (2008). Recommended lists of genotoxic and non-genotoxic chemicals for assessment of the performance of new or improved genotoxicity tests: A follow-up to an ECVAM workshop. *Mutat Res* 653, 99-108. doi:10.1016/j.mrgentox.2008.03.008
- Kirkland, D., Kasper, P., Martus, H. J. et al. (2016). 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* 795, 7-30. doi:10.1016/j.mrgentox.2015.10.006

- Mertens, B., Van Bossuyt, M., Fraselle, S. et al. (2017). Coatings in food contact materials: Potential source of genotoxic contaminants? *Food Chem Toxicol* 106, 496-505. doi:10.1016/j.fct.2017.05.071
- Miller, K. (1991). Clastogenic effects of bleomycin, cyclophosphamide, and ethyl methanesulfonate on resting and proliferating human B- and T-lymphocytes. *Mutat Res* 251, 241-251. doi:10.1016/0027-5107(91)90079-4
- Miller, M. G., Rodgers, A. and Cohen, G. M. (1986). Mechanisms of toxicity of naphthoquinones to isolated hepatocytes. *Biochem Pharmacol* 35, 1177-1184. doi:10.1016/0006-2952(86)90157-7
- Mozdarani, H. and Saberi, A. H. (1994). Induction of cytogenetic adaptive response of mouse bone marrow cells to radiation by therapeutic doses of bleomycin sulfate and actinomycin D as assayed by the micronucleus test. *Cancer Lett* 78, 141-150. doi:10.1016/0304-3835(94)90043-4
- Pottenger, L. H., Bus, J. S. and Gollapudi, B. B. (2007). Genetic toxicity assessment: Employing the best science for human safety evaluation part VI: When salt and sugar and vegetables are positive, how can genotoxicity data serve to inform risk assessment? *Toxicol Sci* 98, 327-331. doi:10.1093/toxsci/kfm068
- SCCP (2005). Opinion on para-Aminophenol. SCCP/0867/05.
https://ec.europa.eu/health/ph_risk/committees/04_sccp/docs/sccp_o_00e.pdf
- SCCP (2006). Opinion on Hydroxybenzomorpholine. SCCP/0965/05.
https://ec.europa.eu/health/ph_risk/committees/04_sccp/docs/sccp_o_066.pdf
- SCCP (2008). Opinion on 1-Naphthol. SCCP/1123/07.
https://ec.europa.eu/health/ph_risk/committees/04_sccp/docs/sccp_o_125.pdf
- SCCS – Scientific Committee on Consumer Safety (2010a). Opinion on Basic Orange 31. SCCS/1334/10. doi:10.2772/31576
- SCCS (2010b). Opinion on 4-Chlororesorcinol. SCCS/1224/09. doi:10.2772/22514
- SCCS (2011a). Opinion on Triclosan. SCCS/1414/11. doi:10.2772/96027
- SCCS (2011b). Opinion on p-Aminophenol. SCCS/1409/11.
https://ec.europa.eu/health/scientific_committees/consumer_safety/docs/sccs_o_078.pdf
- Vinken, M., Doktorova, T., Ellinger-Ziegelbauer, H. et al. (2008). The carcino- genomics project: Critical selection of model compounds for the development of omics-based in vitro carcinogenicity screening assays. *Mutat Res* 659, 202-210. doi:10.1016/j.mrrev.2008.04.006