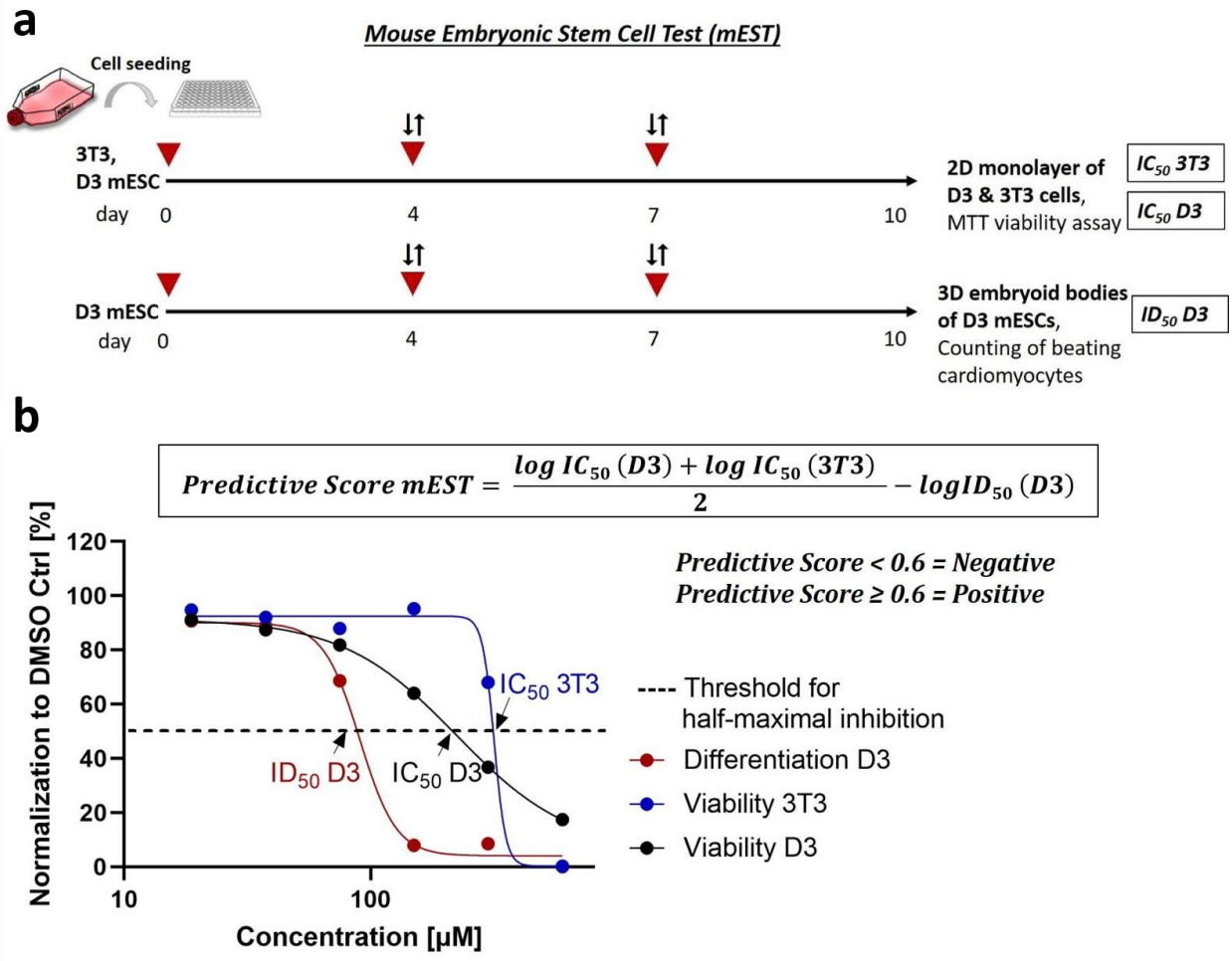
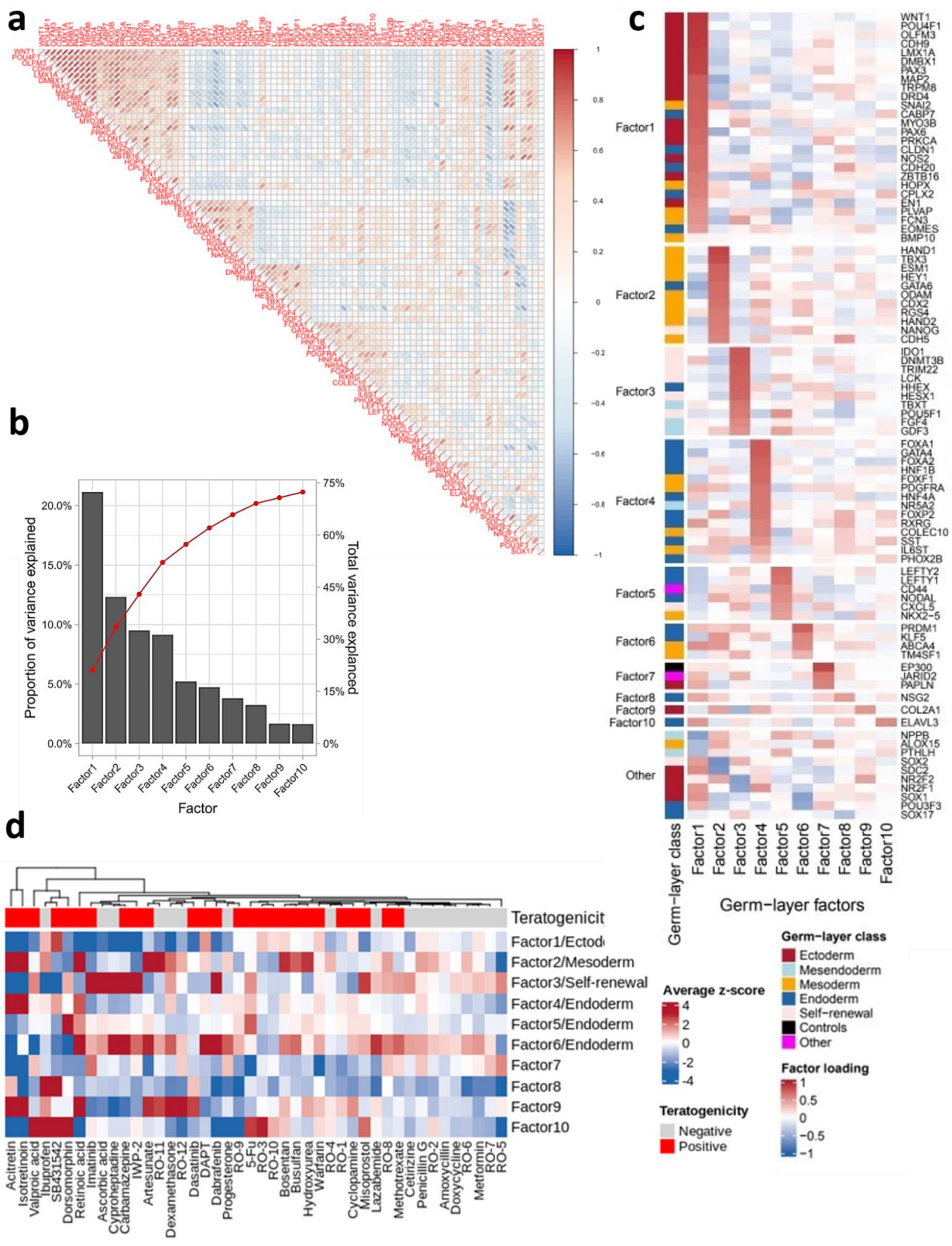
Supplementary Data



Supplementary Figure S1: Mouse embryonic stem-cell test (mEST): assay workflow and prediction model.

(a) Assay workflow. Mouse D3 and 3T3 cells are differentiated over a time course of 10 days to determine mEST endpoints (IC_{50} and ID_{50}) for the calculation of the teratogenicity prediction score. Cells are treated with compounds in six-concentrations at day 0, day 4 and day 7. The endpoints are cell viability (D3 and 3T3 cells) and counting of beating cardiomyocytes at day 10 (derived from D3 EBs).

(b) The prediction model of the mEST assay and normalized concentration-response curves. Values are normalized to solvent (DMSO) controls to determine the half-maximal concentration of viability for D3 ESC (IC_{50} D3) and 3T3 fibroblasts (IC_{50} 3T3) and half-maximal concentration of cardiomyocyte differentiation (ID_{50} D3), respectively. A predictive score of <0.6 indicates non-teratogens whereas a predictive score ≥0.6 indicates teratogens.


a
b
Proportion of variance explained
Total variance explained
Factor
c
Germ−layer class
Germ−layer factors
Average z-score
Germ−layer class
Ectoderm
Mesendoderm
Mesoderm
Endoderm
Self−renewal
Controls
Other
Factor loading
Teratogenicity
Negative
Positive
d
Teratogenicit
Factor1/Ectod
Factor2/Mesoderm
Factor3/Self-renewal
Factor4/Endoderm
Factor5/Endoderm
Factor6/Endoderm
Factor7
Factor8
Factor9
Factor10

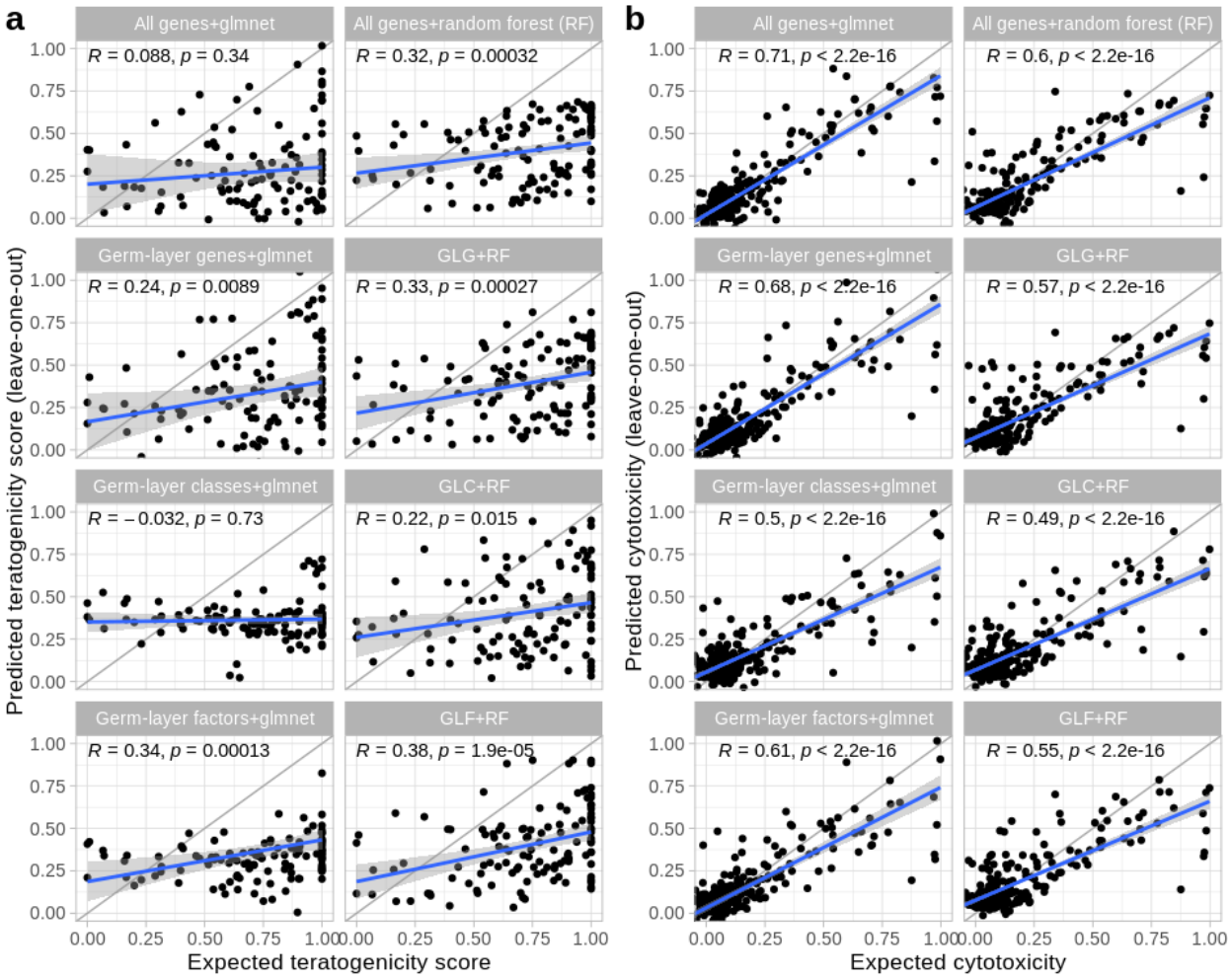
Supplementary Figure S2: Factor analysis of germ-layer genes.

(a) Pairwise Pearson correlation coefficients between germ-layer genes. Each row and each column indicate one gene. Red colors represent a strong correlation between expressions of genes whereas blue colors represent anti-correlation.

(b) Proportion of (co)variance explained by the first ten germ-layer factors (grey bars). Cumulatively, they explain more than 70% of total data variance (red line).

(c) Factor loading (like Figure 2b), with all gene symbols shown.

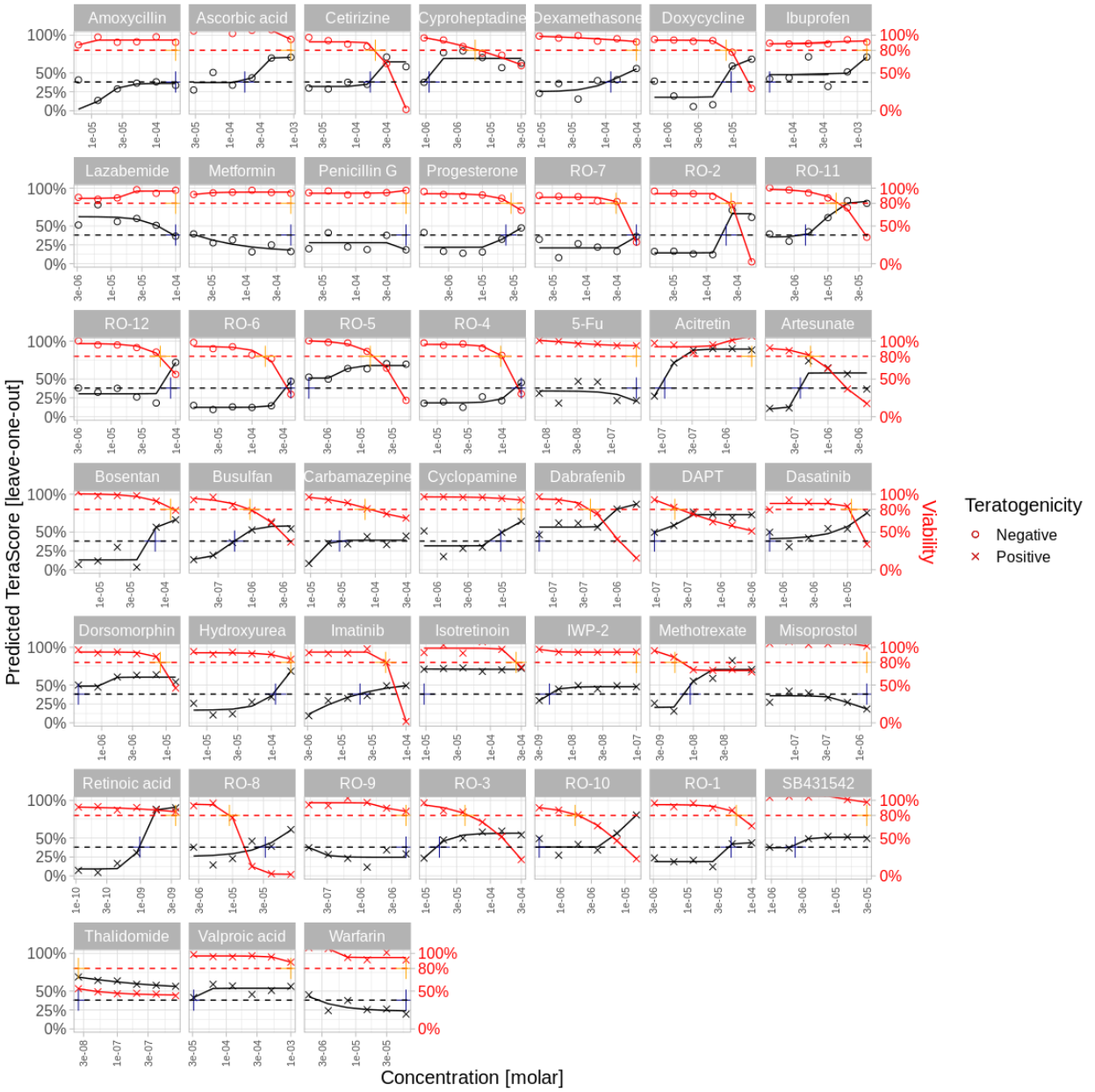
(d) Expression levels of germ-layer factors, represented as average z-scores of associated germ-layer genes, induced by compound treatments in the highest non-cytotoxic concentration (viability ≥ 80%). Blue colors indicate downregulation whereas red colors indicate upregulation of factors. The top side bar uses color to indicate compound classification: grey=non-teratogens, red=teratogens.



Supplementary Figure S3: Leave-one-out prediction of teratogenicity scores and of cytotoxicity.

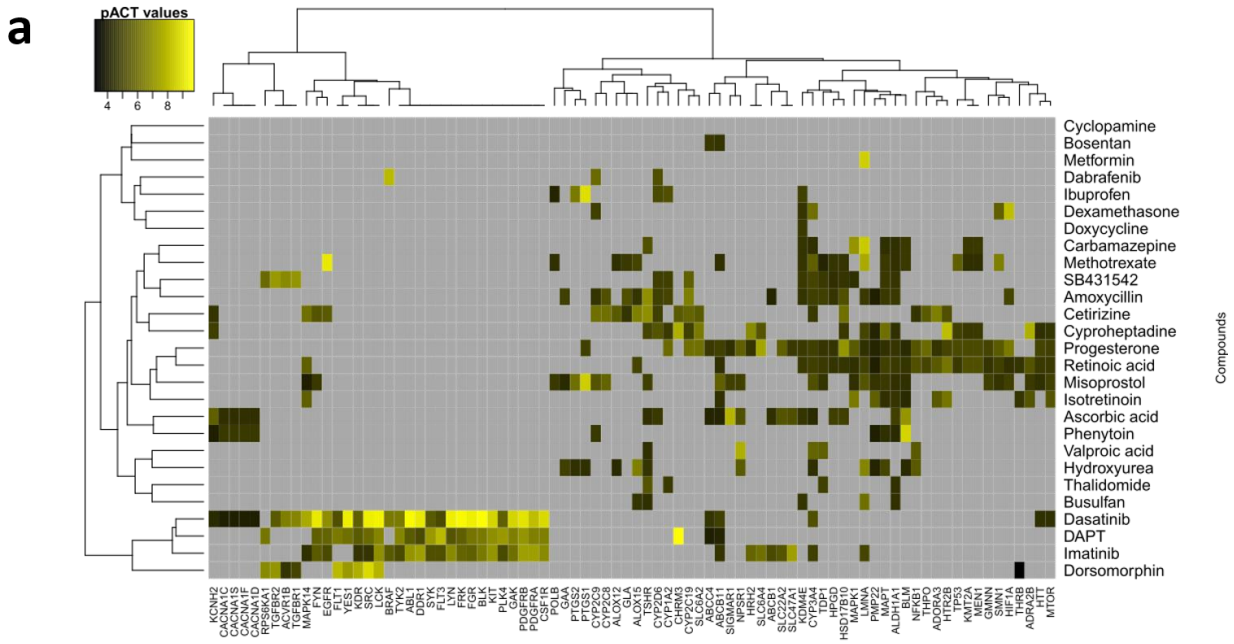
(a) Prediction of teratogenicity scores using the leave-one-out scheme. For each combination of features and machine learning models, the teratogenicity score of each compound is predicted based on the scores of all other compounds. The predicted scores (y-axis) are compared with expected scores as defined in Fig. 3b. Each dot represents one concentration of one compound. The gray line indicates *y*=*x*. *R* gives the Spearman correlation coefficient, and *p* values are derived from the Spearman correlation test. The gray diagonal line represents *y*=*x*. The blue line indicates linear regression, with 95% confidence intervals in the gray area.

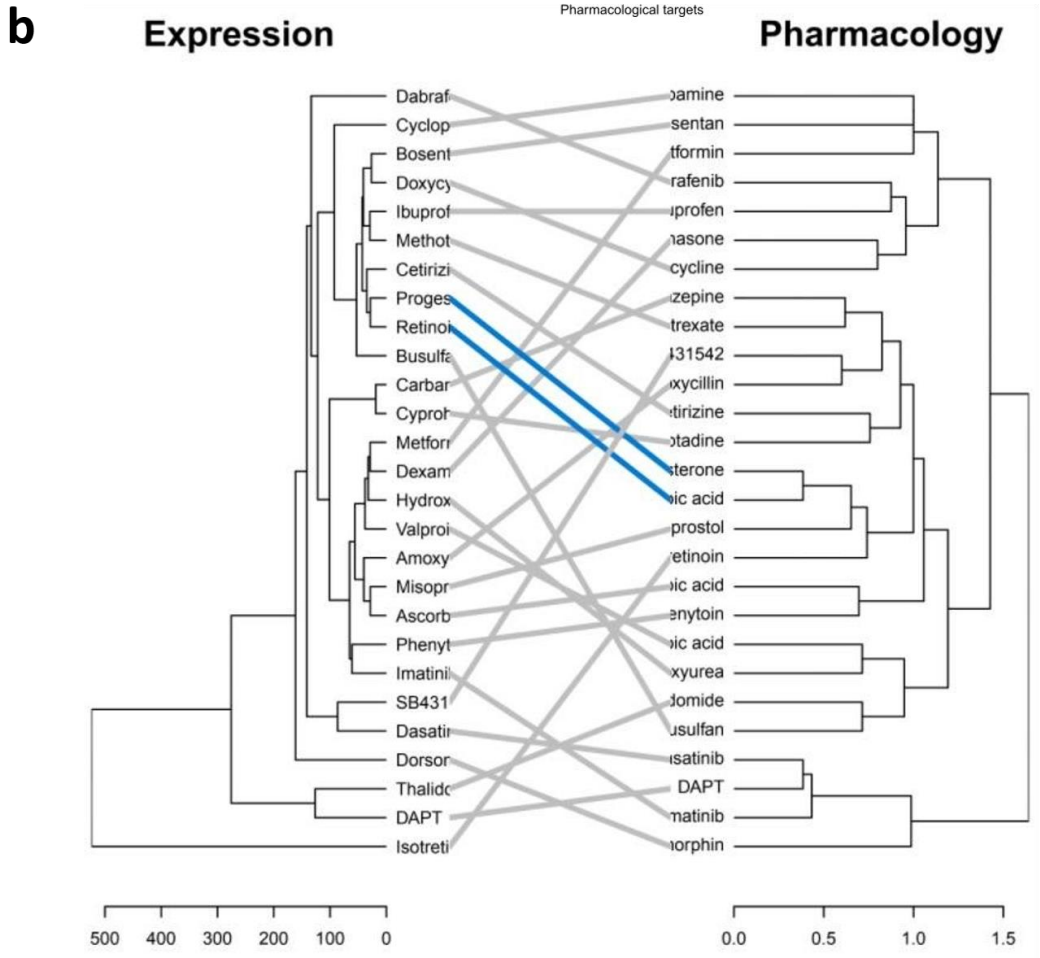
(b) Prediction of cytotoxicity using the leave-one-out scheme. All legends follow the definitions in Suppl. Fig. S3a.



Supplementary Figure S4: Classification of teratogenicity by Random Forest models.

Concentration response curves of 45 reference compounds tested with the human *TeraTox* assay. Values for predicted teratogenicity were obtained by leave-one-out training/testing (median values of n=2), values for cytotoxicity were measured in the assay and normalized by DMSO (which is set as 100%). NCC_{max} and TC_{min} are plotted based on at least 80% viability threshold and the optimal teratogenicity score threshold of 0.38. Curve fit was performed with the four-parameter model offered by the *drc* package.

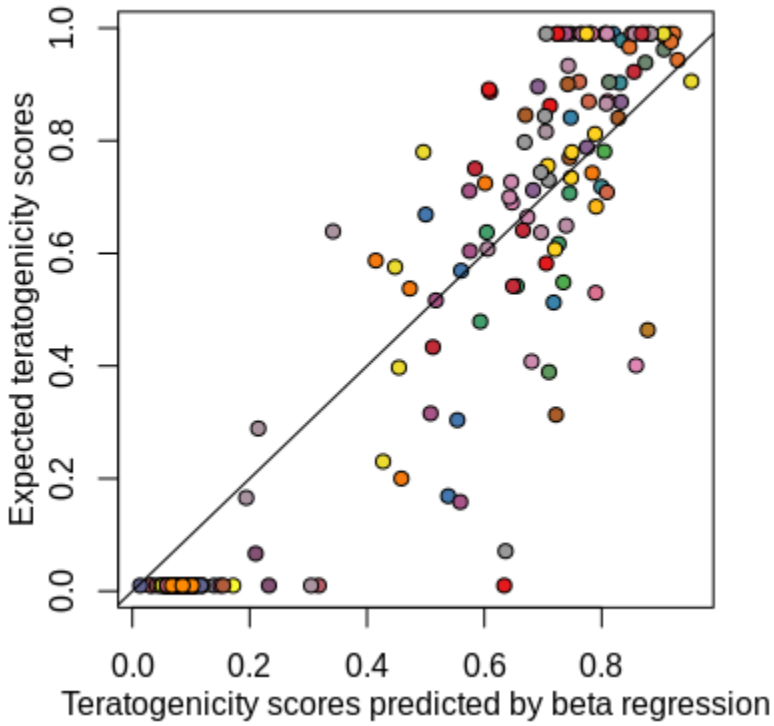

a
pACT values
4
6
8
Cyclopamine
Bosentan
Metformin
Dabrafenib
Ibuprofen
Dexamethasone
Doxycycline
Carbamazepine
Methotrexate
SB431542
Amoxycillin
Cetirizine
Cyproheptadine
Progesterone
Retinoic acid
Misoprostol
Isotretinoin
Ascorbic acid
Phenytoin
Valproic acid
Hydroxyurea
Thalidomide
Busulfan
Dasatinib
DAPT
Imatinib
Dorsomorphin
Compounds
Pharmacological targets


b
Expression
Pharmacology
500
400
300
200
100
0
0.0
0.5
1.0
1.5

Supplementary Figure S5: Pharmacological profiles of the drugs.

(a) Heatmap of target profiles of compounds. Only compounds with annotated targets in ChEMBL are shown. Each row represents one compound, and each column represents a human protein target. Colors indicate pACT values, which are absolute log10 transformed assay values (Ki, IC_{50}, *etc.*). Gray cells mean that data is not available. Compounds are clustered by binary distance and the Ward method.

(b) Two aligned dendrograms of differential gene expression and of pharmacology, linked by compounds. Left: dendrogram of differential gene expression profiles induced by compounds (average across concentration). Right: dendrogram of pharmacological profiles derived from Suppl. Fig. S5a. Lines connect the same compounds in two dendrograms. Gray lines: lack of correspondence. Blue lines: the same cluster is found in both dendrograms.



Supplementary Figure S6: Performance of the generalized linear model with beta regression.

Input to the regression model of ten germ-layer factors and significant interactions between them identified by a Bayesian network. The target variable is the teratogenicity score. We observe a good correlation

between predicted (x-axis) and expected (y-axis) teratogenicity scores that are defined by cosine similarity. Each dot represents one concentration of one compound. Colors are used to represent different drugs. The diagonal line indicates *y=x*.

Supplementary Table S1: Human teratogenicity classification for commercial compounds Classifications and human therapeutic plasma concentrations (C_{max}) were obtained from official data of the U.S. food and drug administration (FDA USPI) or from literature.

Reference Compound	*Teratogenicity Classification*	*Maximal therapeutic plasma concentration (C_{max}) [µM]*	*Reference*
Acitretin	**Positive**	2.4	(3,88)
Amoxicillin	Negative	13.8	(52-54)
Artesunate	**Positive**	1.16	(38,89)
Ascorbic Acid	Negative	49,000	(56,90-92)
Bosentan	**Positive**	1.1	(31,95)
Busulfan	**Positive**	0.52	(3,39)
Carbamazepine	**Positive**	49	(3,32,96-98)
Cetirizine	Negative	0.84	(3,57,58)
Cyclopamine	**Positive**	n/a	(43,44,47)
Cyproheptadine	Negative	104.4	(59,60)
Dabrafenib	**Positive**	2.8	(3,40)
DAPT	**Positive**	n/a	(46)
Dasatinib	**Positive**	0.22	(3,41,100)
Dexamethasone	Negative	0.48	(61)

Dorsomorphin	**Positive**	n/a	(47)
Doxycycline	Negative	13.6	(54,62,101-103)
5-Fluorouracil	**Positive**	222	(3,104)
Hydroxyurea	**Positive**	684	(3,33,106)
Ibuprofen	Negative	286	(3,63,64)
Isotretinoin	**Positive**	1	(3,35,107)
Imatinib	**Positive**	6.6	(3,34,108)
IWP-2	**Positive**	n/a	(47)
Lazabemide	Negative	n/a	(109)
Metformin	Negative	9	(65,110)
Methotrexate	**Positive**	4.7	(3,36)
Misoprostol	**Positive**	n/a	(111)
Penicillin G	Negative	1150	(66,112)
Progesterone	Negative	0.06	(68,114)
Retinoic Acid	**Positive**	1.31	(3,115-117)
SB431542	**Positive**	n/a	(51)
(±) Thalidomide	**Positive**	2.4	(3,118)
Valproic Acid	**Positive**	1423	(3,37)
Warfarin	**Positive**	6.8	(121,122)

Supplementary Table S2: *In vivo* EFD study data for developmental compounds.

Developmental compounds provided by F. Hoffmann- La Roche (compound annotation was blinded due to confidential regulations). Studies for embryo-fetal development were either performed at Roche or contract research organizations. Compounds were generally classified as positive if effects were obtained in at least one species. Maximal protein plasma binding for the lowest observed adverse effect levels (LOAEL C_{max}) from the lowest dose where effects have been observed were averaged. Data were obtained from the Master Thesis of Thomas Sergejew 2015 (73).

Reference Compound	*Teratogenicity Classification*	*in vivo data*			
		Rat	***LOAEL C_{max} [µM]***	***Rabbit***	***LOAEL C_{max} [µM]***
RO-1	**Positive**	**Positive**	18.5	**Positive**	37
RO-2	Negative	Negative	n/a	Negative	n/a
RO-3	**Positive**	**Positive**	91	n/a	n/a
RO-4	Negative	Negative	n/a	n/a	n/a
RO-5	Negative	Negative	n/a	n/a	n/a
RO-6	Negative	Negative	n/a	Negative	n/a
RO-7	Negative	Negative	n/a	n/a	n/a
RO-8	**Positive**	**Positive**	54	**Positive**	32
RO-9	**Positive**	Negative	n/a	**Positive**	0.94
RO-10	**Positive**	Negative	n/a	**Positive**	41
RO-11	Negative	Negative	n/a	n/a	n/a
RO-12	Negative	Negative	n/a	Negative	n/a

Supplementary Table S3: Germ-layer gene panel.

Representative developmental markers (germ-layers) are classified into endoderm, ectoderm, mesoderm, mesendoderm, pluripotency (self-renewal) and other categories, as described by Tsankov *et al*. (25,29,35).

Endoderm	Ectoderm	Mesoderm	Self-Renewal	Mesendoderm
CABP7	WNT1	SNAI2	NANOG	T
CLDN1	POU4F1	HOPX	IDO1	FGF4
CDH20	OLFM3	PLVAP	DNMT3B	GDF3
CPLX2	CDH9	FCN3	TRIM22	NR5A2
EOMES	LMX1A	BMP10	LCK	NPPB
GATA6	DMBX1	HAND1	HESX1	PTHLH
HHEX	PAX3	TBX3	POU5F1	
FOXA1	MAP2	ESM1	CXCL5	
GATA4	TRPM8	HEY1	SOX2	
FOXA2	DRD4	ODAM		
HNF1B	MYO3B	CDX2		
HNF4A	PAX6	RGS4		
FOXP2	PRKCA	HAND2		
RXRG	NOS2	CDH5		
SST	ZBTB16	FOXF1		
PHOX2B	EN1	PDGFRA		
LEFTY2	PAPLN	COLEC10		
LEFTY1	COL2A1	IL6ST		
NODAL	SDC2	NKX2-5		
PRDM1	NR2F2	ABCA4		
KLF5	NR2F1/NR2F2	TM4SF1		
HMP19	SOX1	ALOX15		
ELAVL3				
POU3F3				
SOX17				

Supplementary Table S4: Teratogenicity Prediction of Reference Compounds.

Classification of positive and negative reference compounds by the human *TeraTox* assay with associated maximum non-cytotoxic concentrations (NCC_{max}) and minimal teratogenic concentrations (TC_{min}) and its calculated predictive *TeraTox* score. Classification of the mouse EST with associated half-maximal concentrations for inhibition of growth for D3 mouse ESC ($IC_{50}D3$) and 3T3 fibroblasts ($IC_{50}3T3$) and half-maximal concentrations for inhibition of differentiation into beating cardiomyocytes for D3 mouse ESC ($ID_{50}D3$) and its calculated predictive score. TN= true negative, TP= true positive, FN= false negative, FP= false positive, values of several assay runs were averaged, n≥3, *highly cytotoxic

Reference Compound	Teratogenicity classification	Human *TeraTox* Assay				Mouse embryonic stem cell test				
		NCC_{max} [µM]	*TC_{min} [µM]*	*TeraTox score*	*Predicted Teratogenicity*	*$IC_{50}D3$ [µM]*	*$IC_{50}3T3$ [µM]*	*$ID_{50}D3$ [µM]*	*Predictive Score*	*Predicted Teratogenicity*
Amoxicillin	Negative	200	200	0.00	TN	2500	2500	2467	0.01	TN
Ascorbic acid	Negative	900	174	0.71	**FP**	1104	2000	2000	n/a	TN
Cetirizine	Negative	201	167	0.08	**FP**	332	500	215	0.27	TN
Cyproheptadine	Negative	5.8	1.1	0.72	**FP**	12	45	0.7	1.55	**FP**
Dexamethasone	Negative	300	120	0.40	**FP**	87	300	55	0.46	TN
Doxycycline	Negative	8.0	9.6	-0.08	TN	244	397	11.3	1.44	**FP**

Ibuprofen	Negative	1400	43.7	1.50	**FP**	2628	1380	1166	0.21	TN
Lazabemide	Negative	100	100	0.00	TN	59	235	106	0.05	TN
Metformin	Negative	500	500	0.00	TN	500	500	500	n/a	TN
Penicillin G	Negative	600	600	0.00	TN	2000	2000	2000	n/a	TN
Progesterone	Negative	28	23	0.08	**FP**	58	36	24	0.29	TN
RO-7	Negative	289	600	-0.32	TN	227	227	129	0.25	TN
RO-2	Negative	241	201	0.08	**FP**	47	68	90	n/a	TN
RO-11	Negative	13	4.5	0.46	**FP**	32	13	5.3	0.59	**BL**
RO-12	Negative	58	83	-0.16	TN	92	137	37	0.48	TN
RO-6	Negative	161	400	-0.40	TN	386	515	438	0.01	TN
RO-5	Negative	14	1.6	0.94	**FP**	9.1	5.8	4.1	0.25	TN

RO-4	Negative	96	200	-0.32	TN	39	170	2.5	1.51	**FP**
5-FU	**Positive**	0.3	0.3	0.00	**FN**	0.5	2.0	0.3	0.48	**FN**
Acitretin	**Positive**	2.5	0.1	1.39	TP	0.2	129	0.0	2.64	TP
Artesunate	**Positive**	0.5	0.4	0.09	TP	3.2	6.5	1.3	0.53	**FN**
Bosentan	**Positive**	125	72	0.24	TP	19.3	70.6	22.1	0.22	**FN**
Busulfan	**Positive**	0.9	0.5	0.25	TP	19.3	70.6	22.1	0.22	**FN**
Carbamazepine	**Positive**	70	28	0.40	TP	372	393	207	0.26	**FN**
Cyclopamine	**Positive**	20	9.6	0.32	TP	22	76	5.9	0.84	TP
Dabrafenib	**Positive**	0.4	0.06	0.82	TP	23	23	20	0.07	**FN**
DAPT	**Positive**	0.2	0.09	0.35	TP	324	176	35	0.83	TP
Dasatinib	**Positive**	12	0.6	1.31	TP	4.3	0.6	3.7	n/a	TP

Dorsomorphin	**Positive**	8.1	0.4	1.31	TP	1.9	1.4	0.7	0.34	**FN**
Hydroxyurea	**Positive**	200	116	0.24	TP	51	89	23	0.47	**FN**
Imatinib	**Positive**	48	19	0.40	TP	12	22	7	0.39	**FN**
Isotretinoin	**Positive**	250	9.4	1.42	TP	33	121	0.5	2.10	TP
IWP-2	**Positive**	0.1	4.5E3	1.35	TP	15.5	55	1	1.47	TP
Methotrexate	**Positive**	5.2E3	9E3	-0.23	**FN**	0.2	0.1	0.1	0.10	**FN**
Misoprostol	**Positive**	1.30	1.30	0.00	**FN**	23	13	40	n/a	**FN**
Retinoic acid	**Positive**	3.5E3	9.8E4	0.55	TP	0.004	77.9	0.014	1.60	TP
RO-8	**Positive**	9.0	32	-0.55	**FN**	110	104	85	0.10	**FN**
RO-9	**Positive**	5.0	5.0	0.00	**FN**	40	21	3.7	0.89	TP
RO-3	**Positive**	40	16	0.40	TP	77	147	18	0.77	TP

RO-10	**Positive**	1.7	0.5	0.53	TP	5.1	4.2	1	0.67	TP
RO1	**Positive**	58	48	0.08	TP	68	180	24	0.66	TP
SB431542	**Positive**	30	2.3	1.11	TP	36	21	5.8	0.68	TP
Thalidomide	**Positive**	0.03	0.03	n/a*	TP	2000	2000	2000	n/a	**FN**
Valproic acid	**Positive**	1000	31	1.51	TP	1252	2859	441	0.63	TP
Warfarin	**Positive**	60	60	0.00	**FN**	1892	895	974	0.13	**FN**