

Supplementary Materials

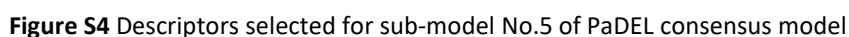
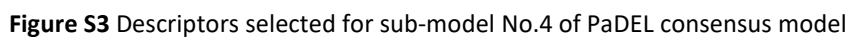
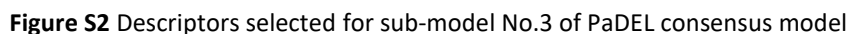
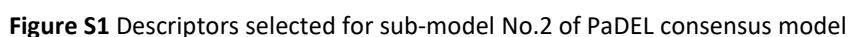
Joint Decision-Making Model Based on Consensus Modeling technology for the Prediction of Drug-induced Liver Injury

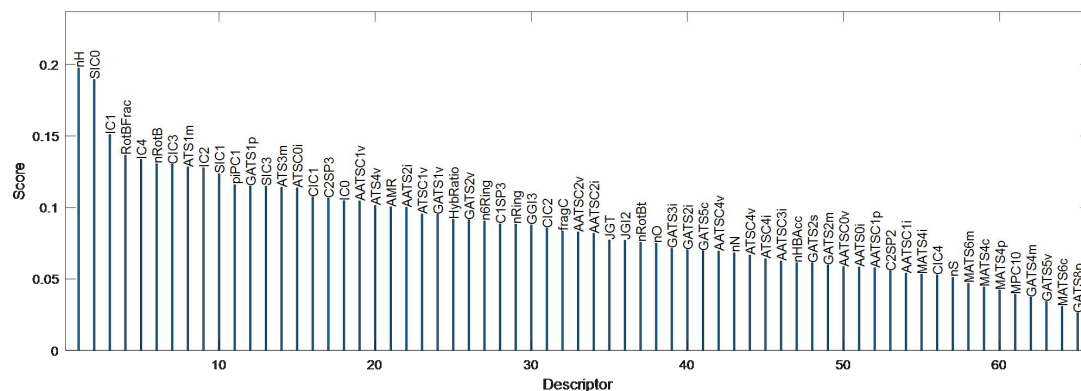
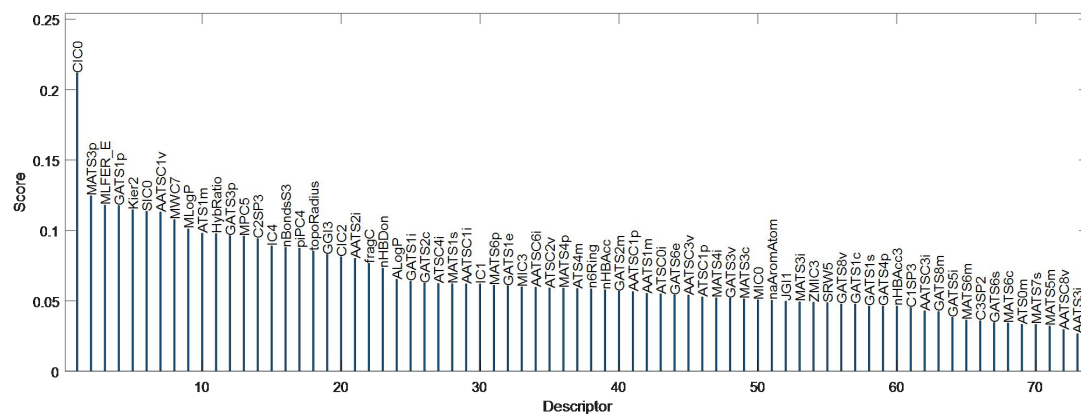
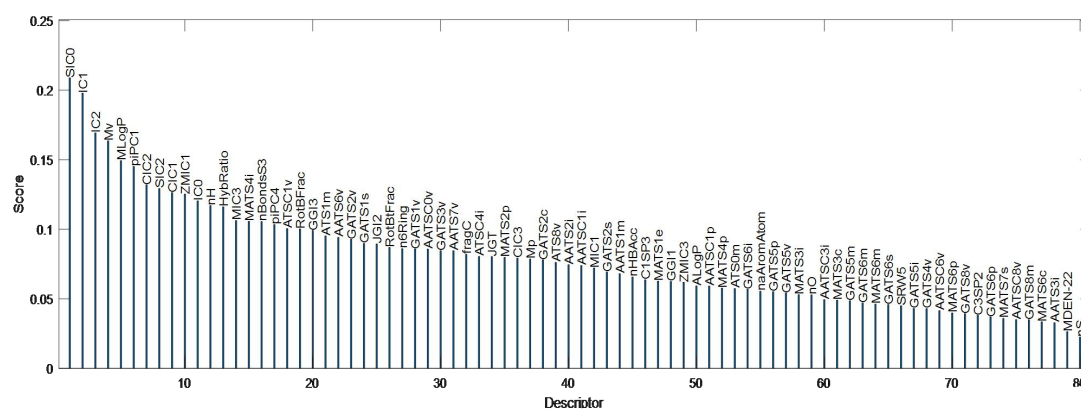
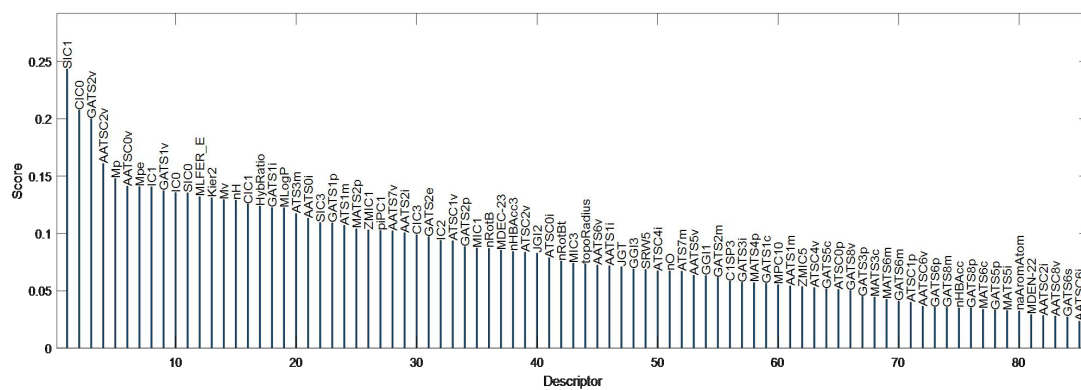
^{a,b}Yukun Wang and ^{b*}Xuebo Chen

^a School of Chemical Engineering, University of Science and Technology Liaoning, No. 185, Qianshan, Anshan 114051, Liaoning, China; wyk410@163.com

^b School of Electronic and Information Engineering, University of Science and Technology Liaoning, No. 185, Qianshan, Anshan 114051, Liaoning, China

* Correspondence: xuebochen@126.com; Tel.: +864125928367





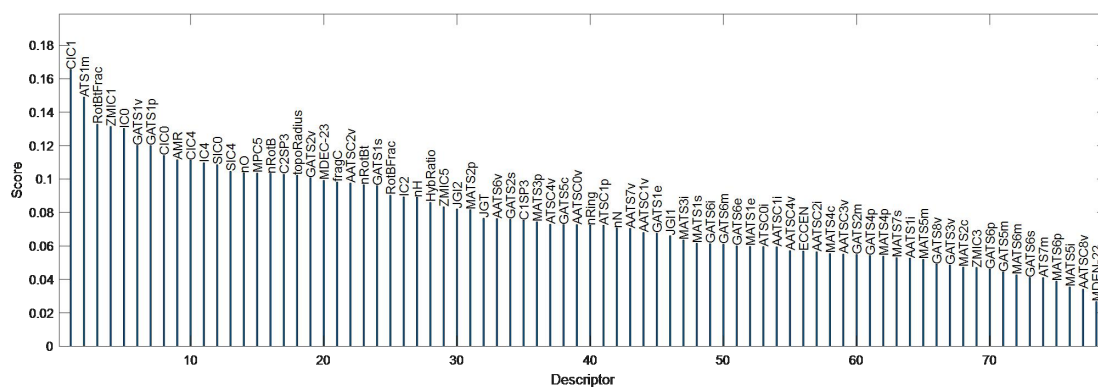


Figure S9 Descriptors selected for sub-model No.10 of PaDEL consensus model

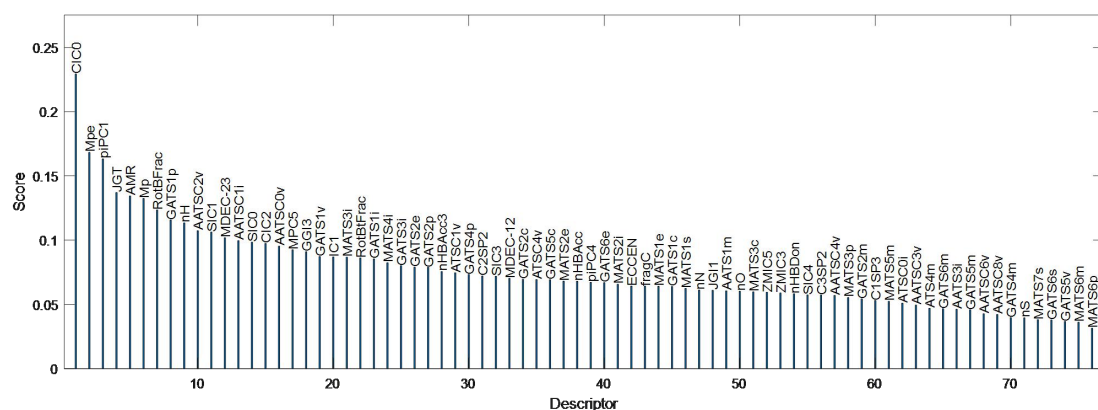


Figure S10 Descriptors selected for sub-model No.11 of PaDEL consensus model

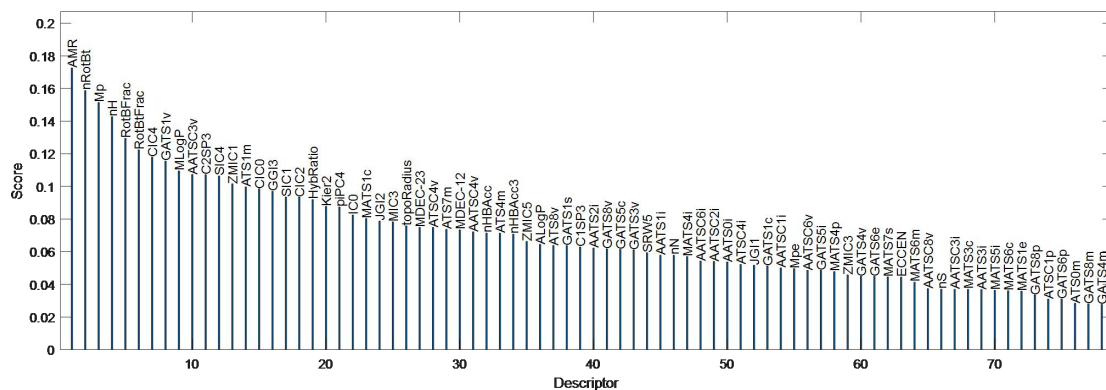


Figure S11 Descriptors selected for sub-model No.12 of PaDEL consensus model

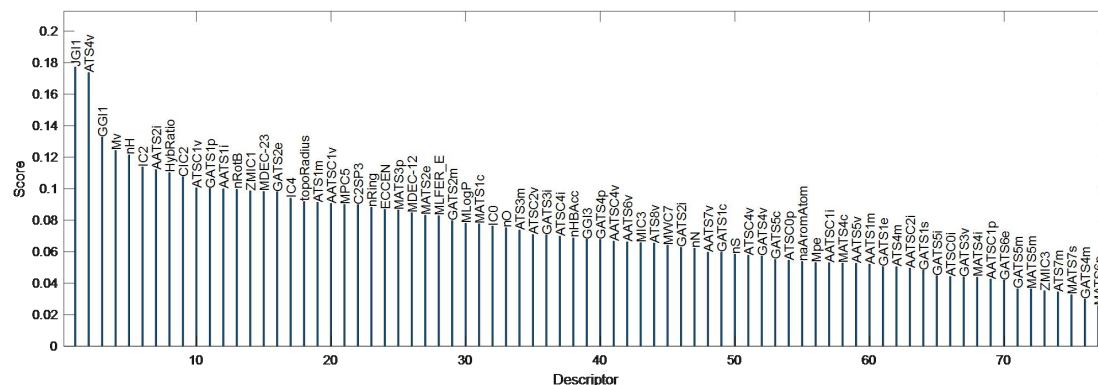
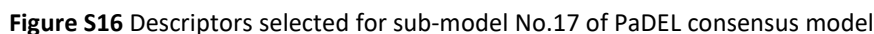
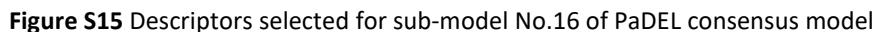
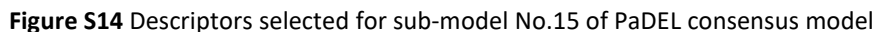
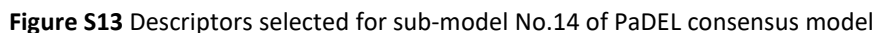


Figure S12 Descriptors selected for sub-model No.13 of PaDEL consensus model



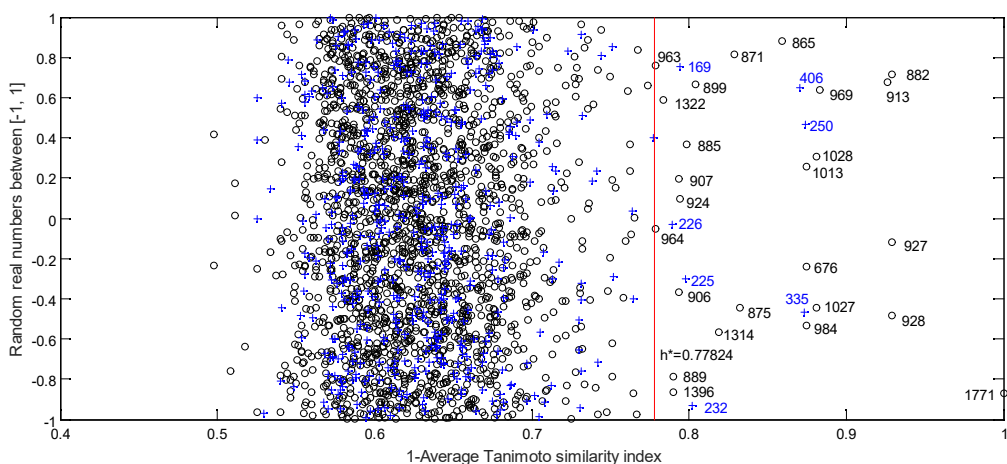


Figure S21 Application domain (AD) plot of the SubFP consensus model.

Table S1. The parameters and performance of the 20 PubchemFP-Based classification model

No.	Parameters		Performance											
			Training set				5-fold cross validation set				Test set			
	C	g	ACC	SE	SP	MCC	ACC	SE	SP	MCC	ACC	SE	SP	MCC
1	6.53E+07	24.0	96.8	97.9	95.1	93.2	70.7	73.2	66.3	38.8	76.5	79.2	71.9	50.5
2	1.36E+07	8.1	96.9	97.8	95.4	93.4	71.0	77.3	60.2	37.6	77.5	81.6	70.4	51.9
3	7.71E+07	5.0	97.0	98.0	95.3	93.6	72.1	81.5	56.2	38.8	77.8	84.0	67.3	52.0
4	2.42E+07	8.1	97.0	98.1	95.2	93.6	71.0	77.6	59.8	37.5	75.9	80.4	68.4	48.6
5	5.65E+07	7.3	96.9	98.0	95.1	93.4	72.6	78.6	62.5	41.2	75.9	81.0	67.3	48.4
6	6.06E+07	6.2	97.1	98.2	95.2	93.7	70.9	79.3	56.7	36.6	75.2	81.3	64.8	46.5
7	7.15E+07	10.4	97.1	97.8	96.0	93.8	71.3	76.7	62.0	38.5	76.5	80.1	70.4	50.1
8	8.23E+07	30.5	96.9	97.5	95.8	93.3	72.1	75.0	67.2	41.5	74.8	77.7	69.9	47.0
9	7.36E+07	5.7	96.8	97.8	95.1	93.1	72.7	80.6	59.3	40.6	76.9	83.4	65.8	49.9
10	6.37E+07	8.3	97.0	98.0	95.3	93.6	71.1	76.7	61.4	38.0	76.9	80.7	70.4	50.8
11	4.47E+07	99.9	96.9	97.9	95.1	93.3	71.0	72.7	68.0	39.8	74.8	77.7	69.9	47.0
12	5.32E+07	5.7	97.1	98.2	95.2	93.8	71.2	79.3	57.2	37.2	75.4	81.0	65.8	47.0
13	5.09E+07	6.2	96.9	98.0	95.1	93.4	72.6	79.8	60.3	40.6	77.1	83.4	66.3	50.4
14	5.49E+07	24.0	96.5	97.6	94.8	92.6	70.0	73.5	64.0	36.9	75.4	79.5	68.4	47.6
15	8.05E+07	30.5	96.9	97.8	95.3	93.3	70.9	73.5	66.6	39.2	75.0	78.6	68.9	47.0
16	5.37E+07	6.7	97.1	98.2	95.2	93.8	70.6	77.7	58.5	36.5	75.4	80.7	66.3	47.1
17	7.23E+07	88.2	96.9	98.0	94.9	93.3	70.0	73.3	64.5	37.1	76.1	78.3	72.4	50.0
18	2.28E+07	6.5	97.0	97.9	95.4	93.6	72.0	78.9	60.2	39.5	75.9	80.4	68.4	48.6
19	7.80E+07	77.0	96.6	97.6	95.1	92.8	69.9	72.3	65.8	37.3	75.6	78.6	70.4	48.4
20	9.64E+07	10.9	96.8	98.0	94.8	93.2	70.2	73.8	64.1	37.3	77.3	81.0	70.9	51.6

Table S2. The parameters and performance of the 20 SubFP-Based classification model

No.	Parameters		Performance(%)											
			Training set				5-fold cross validation set				Test set			
	C	g	ACC	SE	SP	MCC	ACC	SE	SP	MCC	ACC	SE	SP	MCC
1	7.13E+07	17.0	91.6	94.6	86.6	81.9	67.9	72.0	60.9	32.4	71.2	74.1	66.3	39.7
2	1.88E+07	46.4	86.2	88.3	82.6	70.6	67.1	69.4	63.1	31.7	69.7	71.7	66.3	37.2

3	4.66E+07	11.8	87.3	91.8	79.6	72.5	66.4	72.3	56.3	28.4	72.9	78.0	64.3	42.2
4	7.30E+07	6.2	89.3	94.1	81.1	76.9	67.0	73.3	56.2	29.4	72.9	77.4	65.3	42.4
5	8.00E+06	8.7	90.4	93.6	85.0	79.3	70.4	75.0	62.5	37.2	72.2	78.3	61.7	40.2
6	4.36E+07	20.5	90.5	94.4	83.9	79.5	69.5	73.9	61.9	35.4	73.7	78.0	66.3	44.0
7	2.37E+07	16.6	89.6	93.1	83.6	77.5	66.1	71.1	57.5	28.2	74.8	77.1	70.9	47.2
8	3.98E+07	65.7	85.1	90.5	76.1	67.7	65.9	70.1	58.8	28.4	69.5	75.0	60.2	35.0
9	5.74E+07	4.9	89.8	92.8	84.8	78.0	66.6	72.2	57.2	29.1	73.1	78.3	64.3	42.5
10	9.00E+05	18.5	88.2	94.3	77.8	74.3	66.6	71.8	57.7	29.2	70.6	76.2	61.2	37.3
11	3.55E+07	56.0	87.2	91.1	80.5	72.3	67.6	71.4	61.1	32.0	69.9	72.6	65.3	37.2
12	7.86E+07	32.2	88.4	92.8	80.8	74.8	66.6	72.4	56.8	29.0	71.4	77.4	61.2	38.7
13	6.19E+07	41.6	87.4	90.8	81.5	72.7	66.2	70.1	59.4	29.0	70.8	72.9	67.3	39.4
14	5.84E+07	20.9	90.0	93.0	84.8	78.3	70.0	73.5	64.0	36.8	71.6	75.9	64.3	39.8
15	1.12E+08	8.7	90.4	93.6	85.0	79.3	70.4	75.0	62.5	37.2	72.2	78.3	61.7	40.2
16	1.02E+08	8.7	88.7	92.4	82.4	75.6	69.8	75.4	60.2	35.5	70.8	75.6	62.8	38.1
17	1.01E+08	8.7	90.4	93.6	85.0	79.3	70.9	75.4	63.1	38.1	72.3	78.0	62.8	40.8
18	3.34E+07	58.1	86.4	87.9	83.9	71.2	65.7	70.2	58.1	27.9	71.0	70.8	71.4	41.0
19	9.56E+07	86.8	83.8	89.6	73.7	64.7	65.5	70.7	56.6	26.9	68.9	75.6	57.7	33.3
20	5.55E+07	56.0	85.7	89.6	78.9	69.1	65.8	70.0	58.6	28.2	71.2	74.4	65.8	39.6

Table S3. The parameters and performance of the 20 KRFP-Based classification model

No.	Parameters		Performance(%)											
			Training set				5-fold cross validation set				Test set			
	C	g	ACC	SE	SP	MCC	ACC	SE	SP	MCC	ACC	SE	SP	MCC
1	2.10E+06	5.0	96.2	97.8	93.5	91.8	71.8	78.1	61.0	39.2	75.8	81.3	66.3	47.9
2	1.57E+07	5.8	96.4	97.5	94.7	92.4	71.3	78.4	59.3	38.0	77.3	81.6	69.9	51.4
3	3.60E+06	5.2	96.9	98.2	94.5	93.3	70.8	77.9	58.6	36.8	78.0	81.9	71.4	53.1
4	3.38E+07	17.0	96.7	97.6	95.2	92.9	69.9	73.3	64.0	36.7	76.3	78.9	71.9	50.2
5	4.80E+06	16.6	96.6	97.8	94.7	92.8	70.7	73.5	65.9	38.7	74.4	76.8	70.4	46.4
6	3.64E+07	32.6	96.9	98.0	94.9	93.3	69.7	71.7	66.3	37.1	77.7	81.0	71.9	52.6
7	7.00E+06	30.5	96.8	97.9	94.9	93.2	68.7	72.5	62.3	34.2	75.2	77.1	71.9	48.2
8	2.20E+06	6.9	96.5	97.9	94.1	92.6	69.8	75.1	60.6	35.5	76.5	83.7	64.3	48.9
9	2.03E+07	15.4	96.8	97.9	94.8	93.1	70.9	74.1	65.4	38.9	75.0	76.5	72.4	48.0
10	2.20E+06	6.9	96.5	97.9	94.3	92.6	69.6	74.9	60.5	35.1	76.5	83.1	65.3	49.1
11	4.58E+07	7.1	96.6	98.0	94.3	92.8	71.0	76.5	61.6	38.0	75.8	81.9	65.3	47.7
12	1.06E+07	6.5	96.5	97.6	94.5	92.5	68.8	76.4	55.7	32.4	76.3	82.2	66.3	48.9
13	3.69E+07	6.3	96.1	97.5	93.8	91.6	68.9	76.7	55.8	32.8	76.9	82.8	66.8	50.1
14	9.60E+06	5.8	96.7	98.1	94.4	93.0	68.2	75.8	55.1	31.2	75.6	81.0	66.3	47.5
15	1.16E+07	7.1	96.6	98.0	94.1	92.7	70.9	76.6	61.2	37.7	76.5	82.8	65.8	49.2
16	2.00E+06	11.1	96.4	97.6	94.3	92.2	69.6	75.0	60.5	35.2	75.4	78.9	69.4	47.8
17	1.94E+07	4.6	96.9	97.9	95.3	93.4	70.3	79.8	54.2	35.0	77.7	83.4	67.9	51.8
18	7.10E+06	9.2	96.8	97.9	95.1	93.2	67.6	74.0	56.8	30.7	77.5	82.8	68.4	51.5
19	5.70E+06	7.4	96.7	97.9	94.8	93.0	68.7	76.5	55.4	32.2	75.6	80.4	67.3	47.7
20	1.13E+07	14.7	96.6	97.8	94.5	92.7	68.4	73.5	59.8	33.0	75.4	79.5	68.4	47.6

Table S4. The parameters and performance of the 20 E-stateFP-Based classification model

No.	Parameters		Performance(%)											
			Training set				5-fold cross validation set				Test set			
	5.11E+07	45.5	79.2	87.3	65.3	54.4	65.1	73.5	50.8	24.6	66.1	78.0	45.9	25.0
1	7.03E+07	45.5	80.0	86.8	68.4	56.4	64.9	73.6	50.1	23.9	66.9	77.1	49.5	27.4
2	2.27E+07	15.1	79.7	88.7	64.4	55.4	66.3	76.7	48.5	26.0	68.6	78.3	52.0	31.2
3	9.31E+07	34.0	82.3	88.2	72.2	61.4	61.6	67.3	51.9	18.9	67.8	74.1	57.1	31.2
4	6.76E+07	35.7	82.1	87.8	72.3	61.1	65.8	72.2	54.7	26.9	66.1	75.3	50.5	26.3
5	7.16E+07	15.6	82.3	88.8	71.3	61.5	63.3	72.0	48.5	20.7	67.8	75.3	55.1	30.6
6	5.52E+07	47.9	77.2	85.1	63.7	50.2	64.6	73.4	49.7	23.3	65.7	75.0	50.0	25.4
7	2.10E+06	14.1	82.4	87.3	73.9	61.8	63.2	67.5	55.9	23.0	69.9	78.6	55.1	34.4
8	7.56E+07	6.9	84.1	90.8	72.7	65.4	64.5	72.7	50.5	23.3	68.4	74.7	57.7	32.3
9	8.58E+07	6.9	83.5	91.2	70.2	63.9	64.9	74.1	49.0	23.5	68.0	76.5	53.6	30.5
10	9.26E+07	35.8	80.2	87.4	68.0	56.9	64.2	72.4	50.3	22.8	65.9	74.7	51.0	26.1
11	2.25E+07	9.8	76.9	82.5	67.4	50.1	62.3	69.8	49.5	19.3	64.6	73.8	49.0	23.1
12	3.85E+07	12.2	84.3	88.7	76.9	66.1	64.6	71.3	53.1	24.3	69.3	75.9	58.2	34.1
13	1.95E+07	15.1	81.3	88.5	68.9	59.0	65.0	73.8	49.9	24.1	66.1	74.7	51.5	26.6
14	2.94E+07	5.3	82.8	88.5	73.1	62.6	63.7	69.8	53.2	22.8	68.4	75.3	56.6	32.0
15	8.76E+07	38.7	78.3	82.8	70.7	53.5	64.0	71.2	51.9	23.0	66.1	73.5	53.6	27.2
16	8.76E+07	38.5	79.8	85.4	70.1	56.2	64.4	69.7	55.4	24.8	68.9	76.5	56.1	32.9
17	3.19E+07	8.7	79.0	88.0	63.6	53.8	65.6	75.0	49.7	25.1	65.9	75.9	49.0	25.5
18	4.49E+07	50.0	77.1	84.9	63.7	49.9	64.9	72.8	51.4	24.3	66.9	76.5	50.5	27.6
19	5.11E+07	45.5	79.2	87.3	65.3	54.4	65.1	73.5	50.8	24.6	66.1	78.0	45.9	25.0
20	5.11E+07	45.5	79.2	87.3	65.3	54.4	65.1	73.5	50.8	24.6	66.1	78.0	45.9	25.0

Table S5. Detailed information of privileged substructures responsible for DILI-positive compounds.

NO.	Type	Bit	N _{negative}	N _{positive}	F _{DP}	F _{DN}	F _{DN} /F _{DP}
1	SubFP	10	1	12	1.47	0.21	0.14
2		62	0	16	1.59	0	0.00
3		69	1	15	1.49	0.17	0.11
4		197	1	14	1.48	0.18	0.12
5	PubchemFP	329	1	10	1.44	0.25	0.17
6		648	1	11	1.46	0.23	0.16
7		40	1	11	1.46	0.23	0.16
8		748	1	11	1.46	0.23	0.16
9	KRPF	1575	1	12	1.47	0.21	0.14
10		1597	1	14	1.48	0.18	0.12
11		1756	0	10	1.59	0	0.00
12		1799*	0	14	1.59	0	0.00
13		3182*	2	22	1.46	0.23	0.16
14		3524	0	10	1.59	0	0.00
15		3953*	2	18	1.43	0.27	0.19
16		4232*	0	10	1.59	0	0.00
17		4252*	2	26	1.47	0.19	0.13

18	4387	0	11	1.59	0	0.00
19	4556	1	11	1.46	0.23	0.16
20	4651*	1	15	1.49	0.17	0.11
21	4689*	0	12	1.59	0	0
22	4692*	0	11	1.59	0	0
23	4708	1	11	1.46	0.23	0.16
24	4778*	1	13	1.47	0.19	0.13
25	4808*	0	11	1.59	0	0
