

## Supplementary Information

Perturbation efficiency resolves target-count bias in network proximity metrics: A controlled audit

Antony Bevan

## Supplementary Tables

Table S1: **Hyperforin target genes and literature sources.** All 14 raw targets with UniProt IDs, gene symbols, and primary literature sources. Targets marked with \* are present in the liver-expressed LCC (STRING  $\geq 900$ , GTEx TPM  $\geq 1$ ).

UniProt	Gene	In LCC	Source
O75469	NR1I2 (PXR)	Yes*	[1, 2]
P08684	CYP3A4	Yes*	[1]
P11712	CYP2C9	Yes*	[3]
P20813	CYP2B6	Yes*	[4]
P08183	ABCB1	Yes*	[5]
Q9UNQ0	ABCG2	Yes*	[6]
O15440	ABCC2	Yes*	[7]
P31749	AKT1	Yes*	[8]
P08253	MMP2	Yes*	[8]
P14780	MMP9	Yes*	[9]
Q9Y210	TRPC6	No	[10]
P15692	VEGFA	No	[9]
Q13794	PMAIP1	No	[11]
Q12879	GRIN1	No	[12]

Table S2: **Quercetin target curation summary.** Target counts at each processing stage.

Stage	Count
Raw targets (ChEMBL v31, CHEMBL159)	122
Excluded: non-human (mouse, rat, bacterial, viral)	10
Excluded: no UniProt mapping	25
Processed targets	87
Excluded: not liver-expressed (TPM < 1)	20
Excluded: not in STRING LCC	5
Final targets in LCC	62

Table S3: **DILI gene set curation.** Genes associated with drug-induced liver injury from DisGeNET (UMLS C0860207).

Stage	Count
Raw DILI genes (DisGeNET)	127
In STRING $\geq$ 700 liver LCC	84
In STRING $\geq$ 900 liver LCC	82
Excluded: miRNAs (not in PPI network)	21
Excluded: cytokines (not in LCC)	12
Excluded: other	12

Table S4: **Genes targeted by both compounds.** Five genes present in both Hyperforin and Quercetin target sets.

Gene	Protein	Function
ABCG2	BCRP	Efflux transporter
AKT1	Protein kinase B	Cell survival signaling
CYP3A4	Cytochrome P450 3A4	Drug metabolism
MMP2	Matrix metalloproteinase-2	Extracellular matrix remodeling
MMP9	Matrix metalloproteinase-9	Extracellular matrix remodeling

## References

- [1] Linda B Moore, Bryan Goodwin, Stacey A Jones, G Bruce Wisely, Connie J Serabjit-Singh, Timothy M Willson, John L Collins, and Steven A Kliewer. St. John’s wort induces hepatic drug metabolism through activation of the Pregnan X Receptor. *Proceedings of the National Academy of Sciences*, 97(13):7500–7502, 2000. doi: 10.1073/pnas.130155097.
- [2] Reginald E Watkins, G Bruce Wisely, Linda B Moore, John L Collins, Millard H Lambert, Shawn P Williams, Timothy M Willson, Steven A Kliewer, and Matthew R Redinbo. The human nuclear xenobiotic receptor PXR: structural determinants of directed promiscuity. *Science*, 292(5525):2329–2333, 2001. doi: 10.1126/science.1060762.
- [3] R Scott Obach. Inhibition of human cytochrome P450 enzymes by constituents of St. John’s Wort, an herbal preparation used in the treatment of depression. *Journal of Pharmacology and Experimental Therapeutics*, 294(1):88–95, 2000. doi: 10.1124/jpet.294.1.88.
- [4] Bernard J Komoroski, Shuyan Zhang, Steven A Wrighton, Stephen C Strom, Raman Venkataraman, and Erin G Schuetz. Induction and inhibition of cytochromes P450 by the St. John’s wort constituent hyperforin in human hepatocytes. *Drug Metabolism and Disposition*, 32(5):512–518, 2004. doi: 10.1124/dmd.32.5.512.

Table S5: **Direct DILI gene connectivity.** Hyperforin targets with first-order (distance = 1) connections to DILI genes in the STRING network ( $\geq 900$ ). DILI neighbors are genes present in the 82-gene DILI set.

Target	DILI Neighbors	N	Function
CYP3A4	NR1I2, CYP2E1, UGT1A9, GSTM1, GSTP1	5	Xenobiotic metabolism
AKT1	MAP3K5, NFE2L2, CTNNB1, IGF1	4	Stress response
MMP9	LCN2, SPP1, MMP2	3	Inflammation/ECM
ABCB1	ABCC2, NR1I2	2	Drug transport
CYP2C9	CYP2E1, NR1I2	2	Xenobiotic metabolism
CYP2B6	NR1I2	1	Xenobiotic metabolism
NR1I2	CYP2E1, ABCC2	2	Master regulator
ABCG2	ABCC2	1	Drug transport
ABCC2	NR1I2, ABCB1	2	Drug transport
MMP2	MMP9, SPP1	2	ECM remodeling
<b>Total unique</b>		<b>12</b>	

Table S6: **Quercetin direct DILI gene connectivity summary.** Summary statistics for first-order DILI connections across Quercetin's 62 targets.

Metric	Value
Total targets in LCC	62
Targets with $\geq 1$ direct DILI neighbor	18
Total direct DILI connections	31
Mean DILI neighbors per target	0.50
<i>Hyperforin comparison:</i>	
Hyperforin targets with $\geq 1$ DILI neighbor	10/10 (100%)
Mean DILI neighbors per Hyperforin target	2.4

- [5] M Hennessy, D Kelleher, JP Lloyd, A Alrajhi, O Meenaghan, C McDonald, F Mulcahy, JP Spiers, and J Feely. St John's wort increases expression of P-glycoprotein: implications for drug interactions. *British Journal of Clinical Pharmacology*, 53(1):75–82, 2002. doi: 10.1046/j.1365-2125.2002.01512.x.
- [6] H Assefa and V Butterweck. The role of hyperforin in the metabolic and transport-mediated drug interactions of St. John's wort. *Planta Medica*, 70(4):291–300, 2004. doi: 10.1055/s-2004-818938.
- [7] Er-Jia Wang, Mary Barecki-Roach, and William W Johnson. Quantitative characterization of direct P-glycoprotein inhibition by St John's wort constituents hypericin and hyperforin. *Journal of Pharmacy and Pharmacology*, 56(11):1451–1456, 2004. doi: 10.1211/0022357044736.
- [8] C Quiney, C Billard, A M Faussat, C Salanoubat, and J P Kolb. Hyperforin directly inhibits AKT1

Table S7: **Quercetin target genes in the liver-expressed network.** All 62 Quercetin targets in STRING v12.0 LCC (confidence  $\geq 900$ ) with liver TPM  $\geq 1$  (GTEx v8). Sorted by descending liver expression.

Gene	TPM	Gene	TPM	Gene	TPM	Gene	TPM
CFB	1115	CYP3A4	335	FN1	229	ALDH2	183
ANPEP	160	PPIA	112	SERPINA5	104	CYP1A2	72
CA2	64	APP	63	PYGL	55	HDAC6	45
ESRRA	42	MAOA	35	AKR1C2	33	AKT1	33
CTSH	28	XDH	26	CHRNA4	25	PIK3R1	24
PIM1	24	LDLR	23	EGFR	17	ELOVL1	18
PKN1	16	GSK3A	13	YES1	13	MET	12
DAPK1	12	BACE1	11	CSNK2A1	10	FSTL1	9
SIRT6	8	GSK3B	7	CDK7	7	CAV2	7
PTPN2	6	CYP1A1	5	PRMT7	5	MMP2	5
AKR1B1	5	PDE6D	5	PTK2	4	ABCG2	4
IQGAP1	4	ADRB2	3	BRAF	4	KDR	3
SRC	3	ALOX5	3	CYP1B1	3	TLR4	3
NUAK1	3	AXL	2	ADA	2	LCK	2
ABCC1	2	PLK1	1	ACHE	1	MMP9	1
SYK	1	PDZK1IP1	1				

Table S8: **DILI gene set (82 genes).** Genes in STRING v12.0 LCC (confidence  $\geq 900$ ) with liver TPM  $\geq 1$  (GTEx v8). Source: DisGeNET (UMLS C0860207). Sorted alphabetically.

82 DILI-Associated Genes							
ABCB1	AHR	ALB	ALDOB	AMBP	APOA1	APOE	APOH
ARG1	ARNT	ATG5	BAX	BTD	C3	CAT	CCL2
CLU	COL3A1	CTNNB1	CXCL1	CXCL10	CYP2A6	CYP2C19	CYP2C9
CYP2E1	DGAT2	ENO1	FGA	FLT1	FMO3	GADD45A	GC
GCLC	GPT	GSN	GSTM1	GSTM2	GSTP1	HLA-A	HLA-B
HLA-DQB1	HLA-DRB1	HMGB1	HMOX1	HPD	HPX	IGF1	IL18
IL1R2	KRT18	LCN2	LGALS3	MAP3K5	MED1	MMP2	MTHFR
NAT2	NFE2L2	NR1H3	NR1H4	NR1I2	NR1I3	PLAT	PLG
PNP	POLG	PON1	PPARA	PRKDC	PTGS2	RBP1	SLPI
SNX18	SOD1	SOD3	SPP1	TALDO1	TBXA2R	TCTN1	TF
TTR	UGT1A9						

kinase activity and promotes apoptosis in AML cells. *Leukemia*, 21(10):2101–2111, 2007. doi: 10.1038/sj.leu.2404834.

[9] C Quiney, C Billard, C Salanoubat, J D Fourneron, and J P Kolb. Hyperforin inhibits MMP-9 secretion by B-cell chronic lymphocytic leukemia cells. *Leukemia*, 20(8):1514–1521, 2006. doi: 10.1038/sj.leu.2404283.

[10] Kristian Leuner, Viacheslav Kazanski, Marina Müller, Kirill Essin, Britta Henke, Martina Gassen, Christopher Koch, Christina Bulut, Karola Silbermann, Annette Kopp-Schneider, Gerald Thiel, Vladimir Laketa, Inna Gorshkova, Valentina Przetheskikh, Christian Harteneck, Wolfgang F Graier, Vadym Degtyar, Peter Lipp, Axel Lückhoff, and Walter E Müller. Hyperforin—a key constituent of St. John’s wort specifically activates TRPC6 channels. *The FASEB Journal*, 21(14): 4101–4111, 2007. doi: 10.1096/fj.07-8110com.

[11] Katarina Hostanska, J Reiher, S Jessenmeyer, J Reichling, and R Saller. Hyperforin and hypericin:

Table S9: **Null distribution parameters from permutation testing.** Null distribution parameters (mean and standard deviation) from  $n = 1,000$  degree-matched permutations. Note the tightening of the Quercetin null distribution as the number of targets increases, which drives the inflation of proximity Z-scores.

Metric	Compound	$\mu_{null}$	$\sigma_{null}$	$x_{obs}$	Z-score
<i>Shortest-path proximity (at <math>\geq 900</math>)</i>					
	Hyperforin (10)	2.21	0.235	1.30	-3.86
	Quercetin (62)	2.17	0.091	1.68	-5.44
<i>Random walk influence (at <math>\geq 900</math>)</i>					
	Hyperforin (10)	0.0147	0.0098	0.1138	+10.12
	Quercetin (62)	0.0148	0.0038	0.0322	+4.55
<i>Expression-weighted influence (at <math>\geq 900</math>)</i>					
	Hyperforin (10)	0.0205	0.0125	0.1330	+8.98*
	Quercetin (62)	0.0209	0.0049	0.0493	+5.79*

\*Significance remains high despite tissue-specific attenuation.  $\mu_{null}$  = null distribution mean;  $\sigma_{null}$  = null distribution standard deviation;  $x_{obs}$  = observed metric value.

Table S10: **Bootstrap sensitivity excludes target-count confounding.** Random 10-target subsets ( $n = 100$ ) sampled without replacement from Quercetin's 62-target pool. Hyperforin's observed influence exceeds the entire bootstrap distribution.

Statistic	Value	Interpretation
Hyperforin observed	0.1138	Reference
Bootstrap mean	0.0308	Expected if targets equivalent
Bootstrap SD	0.0100	Sampling variability
Bootstrap 95% CI	[0.0160, 0.0542]	2.5th–97.5th percentile
Hyperforin / mean	<b>3.7×</b>	Effect size
Exceeds 95% CI?	<b>Yes</b>	Not attributable to sampling

Random seed: 42. Note: Bootstrap confirms robustness to target selection; it does not constitute independent inferential evidence.

synergistic cytotoxicity and induced apoptosis in human malignant cell lines. *European Journal of Pharmaceutics and Biopharmaceutics*, 55(3):301–310, 2003. doi: 10.1016/s0939-6411(03)00021 -3.

- [12] Vikas Kumar, Alexander Mdzinarishvili, Thomas Kiewert, Maria P Abbracchio, Annalisa Pinna, Renata Ciccarelli, Walter E Müller, and Jochen Klein. NMDA receptor-antagonistic properties of hyperforin, a constituent of St. John's wort. *Journal of Pharmacological Sciences*, 102(1):47–54, 2006. doi: 10.1254/jphs.fp06041.

Table S11: **Chemical similarity excludes structural confounding.** Neither compound resembles known hepatotoxins ( $\text{Tanimoto} < 0.4$ ). Quercetin is more similar to DILI-positive drugs yet shows lower network influence.

Compound	Max Tanimoto (DILI+)	Max Tanimoto (DILI-)	Analog?*	Network rank
Hyperforin	0.154	0.202	No	1 (higher influence)
Quercetin	0.212	0.220	No	2 (lower influence)

\*Analog threshold:  $\text{Tanimoto} > 0.4$  (Maggiora et al., 2014). Morgan fingerprints (ECFP4, radius 2, 2048 bits). DILIranks: 542 DILI+, 365 DILI- drugs.

Table S12: **Hyperforin targets include regulatory hubs.** All 10 Hyperforin targets in the liver-expressed LCC, with liver expression (GTEx v8) and network degree. PXR (NR1I2) is the master regulator; CYP enzymes are downstream effectors.

Gene	Protein	TPM	Degree	Function	DILI link
NR1I2	PXR	43	28	Master regulator	Direct
CYP3A4	CYP3A4	335	89	Xenobiotic metabolism	Direct
CYP2C9	CYP2C9	434	76	Xenobiotic metabolism	Direct
CYP2B6	CYP2B6	125	42	Xenobiotic metabolism	Indirect
AKT1	PKB	33	<b>312</b>	Stress signaling	Indirect
ABCB1	P-gp	7	53	Drug efflux	Direct
ABCC2	MRP2	60	38	Drug efflux	Direct
ABCG2	BCRP	4	31	Drug efflux	Indirect
MMP2	MMP2	5	87	ECM remodeling	Indirect
MMP9	MMP9	1	94	ECM remodeling	Indirect

AKT1 is the highest-degree target (312 neighbors). Five of 10 targets (NR1I2, CYP3A4, CYP2C9, ABCB1, ABCC2) are directly connected to DILI genes. TPM = transcripts per million; DILI = drug-induced liver injury; LCC = largest connected component.

Table S13: **Influence ranking is robust to network construction parameters.** Hyperforin ranks first across all thresholds and influence metrics. Proximity Z-scores are unstable and reverse rankings between thresholds, failing to accurately reflect the physical distance advantage of Hyperforin.

Threshold	Compound	RWR Z	EWI Z	Proximity $d_c$	Proximity Z
$\geq 700$ (11,693 nodes)	Hyperforin	<b>+12.08</b>	+11.20	0.60	-6.04
	Quercetin	+5.53	+7.09	1.34	-5.46
$\geq 900$ (7,677 nodes)	Hyperforin	<b>+10.12</b>	+8.98	1.30	-3.86
	Quercetin	+4.55	+5.79	1.68	-5.44

Note: At  $\geq 900$ , Quercetin achieves a more "significant" proximity Z-score despite being physically more distant (1.68 vs 1.30) from DILI genes. RWR = random walk with restart; EWI = expression-weighted influence;  $d_c$  = mean minimum shortest-path distance; DILI = drug-induced liver injury.