

Supplementary Information

Systematic analysis of public domain compound potency data identifies selective molecular scaffolds across druggable target families

Ye Hu, Anne Mai Wassermann, Eugen Lounkine, and Jürgen Bajorath

Supplementary Tables S1-S6

Supplementary Figures S1-S4

Supplementary Table S1. Targets.

TargetID	BindingDB Target Name	Abbreviation
1	Caspase-1	CASP1
2	Caspase-2	CASP2
3	Caspase-3	CASP3
4	Caspase-4	CASP4
5	Caspase-5	CASP5
6	Caspase-6	CASP6
7	Caspase-7	CASP7
8	Caspase-8	CASP8
9	Cathepsin D	CatD
10	3-Phosphoinositide-Dependent Protein Kinase 1 (PDK1)	PDK1
11	Checkpoint Kinase (Chk1)	Chk1
12	Cyclin-Dependent Kinase 1 (CDK1)	CDK1
13	Cyclin-Dependent Kinase 2 (CDK2)	CDK2
14	Cyclin-Dependent Kinase 4 (CDK4)	CDK4
15	PIM-1 Kinase	PIM1
16	Phosphoinositide 3-Kinase (PI3K), gamma	PI3Kgamma
17	Polo-Like Kinase 1	PLK1
18	Polo-Like Kinase 2	PLK2
19	Polo-Like Kinase 3	PLK3
20	Polo-Like Kinase 4	PLK4
21	Protein Kinase C, gamma	PKCgamma
22	Protein Kinase C, zeta	PKCzeta
23	Protein kinase B (Akt 1)	Akt1
24	Protein kinase B (Akt 2)	Akt2
25	Protein kinase B (Akt 3)	Akt3
26	Tyrosine Kinase CSF1-R	CSF1R
27	Tyrosine Kinase CSF1-R Mutant (FGFR KID Chimera)	CDF1R*M1
28	Tyrosine Kinase CSF1-R Mutant (Tie2 KID Chimera)	CDF1R*M2
29	Tyrosine-protein kinase ITK	ITK
30	Insulin Receptor Tyrosine Kinase (IRK)	IRK
31	Insulin-like Growth Factor 1 Receptor (IGF-1R)	IGF1R
32	Mixed Lineage Kinase 1 (MLK1)	MLK1
33	Mixed Lineage Kinase 3 (MLK3)	MLK3
34	EGF-R Tyrosine Kinase	EGFR
35	Src	Src
36	FGFR-1 Tyrosine Kinase	FGFR1
37	PDGFR-beta	PDGFRbeta
38	Lck	Lck
39	Protein Kinase C, alpha	PKCalpha
40	Protein Kinase C, beta I	PKCbeta
41	Protein Kinase C, beta II	PKCbeta2
42	Protein Kinase C, delta	PKCdelta

43	Protein Kinase C, epsilon	PKCepsilon
44	Protein Kinase C, eta	PKCeta
45	ABL	ABL
46	ErbB-2 (HER2) Tyrosine Kinase	HER2
47	VEGFR-2 (KDR)	VEGFR2
48	VEGFR-1 (Flt-1)	VEGFR1
49	Carbonic Anhydrase I	CA1
50	Carbonic Anhydrase II	CA2
51	Carbonic Anhydrase IX	CA9
52	Carbonic Anhydrase VA	CA5A
53	Carbonic Anhydrase VB	CA5B
54	Carbonic Anhydrase VI	CA6
55	Tyrosine Kinase c-Kit	cKit
56	Tyrosine Kinase Syk	Syk
57	Tyrosine Kinase ZAP-70	ZAP70
58	GRB2	GRB2
59	Aurora Kinase A (Aurora-A)	AuroraA
60	ErbB-4 (HER4) Tyrosine Kinase	HER4
61	Cyclin-Dependent Kinase 5 (CDK5)	CDK5
62	Cyclin-Dependent Kinase 6 (CDK6)	CDK6
63	MAP Kinase p38 alpha	MAPK14
64	Tie-2	Tie2
65	FGFR-2 Tyrosine Kinase	FGFR2
66	Mixed Lineage Kinase 2 (MLK2)	MLK2
67	alpha-Amylase	alphaAmylase
68	Glycogen Synthase Kinase-3, alpha	GSK3alpha
69	Cyclin-Dependent Kinase 7 (CDK7)	CDK7
70	Cyclin-Dependent Kinase 9 (CDK9)	CDK9
71	ERK-2	ERK2
72	Tryptase	Tryptase
73	Thromboxane A2 Synthase (P450 TxA2)	TxA2
74	Glycogen Synthase Kinase-3, beta	GSK3beta
75	Liver X Receptor alpha (LXR-alpha)	LXRalpha
76	Liver X Receptor beta (LXR-beta)	LXRbeta
77	Matrix Metalloproteinase-1 (MMP-1)	MMP1
78	Matrix Metalloproteinase-2 (MMP-2)	MMP2
79	Matrix Metalloproteinase-3 (MMP-3)	MMP3
80	Matrix Metalloproteinase-9 (MMP-9)	MMP9
81	Aldosterone Synthase (CYP11B2)	CYP11B2
82	Aromatase (CYP19)	CYP19
83	Steroid 11-beta-Hydroxylase (CYP11B1)	CYP11B1
84	Cytochrome P450 3A4	CYP3A4
85	Steroid 17-alpha-Monooxygenase (CYP17)	CYP17
86	Tyrosine Kinase FLT3	FLT3
87	Acetylcholinesterase (AChE)	AChE
88	Butyrylcholinesterase (BuChE)	BuChE
89	Carboxylesterase 1	CE1
90	Carboxylesterase 2 (intestine, liver)	CE2

91	Quinone Reductase 2 (QR2)	QR2
92	Steroid Sulfatase (STS)	STS
93	Caspase-9	CASP9
94	Aldose Reductase (ALR2)	ALR2
95	Monoamine Oxidase Type A (MAO-A)	MAOA
96	Monoamine Oxidase Type B (MAO-B)	MAOB
97	Phosphodiesterase Type 4 (PDE4A)	PDEA4
98	Carbonic Anhydrase III	CA3
99	Carbonic Anhydrase IV	CA4
100	Carbonic Anhydrase VII	CA7
101	Carbonic Anhydrase XII	CA12
102	Carbonic Anhydrase XIV	CA14
103	Dipeptidyl Peptidase IV (DPP-IV)	DPP4
104	DPP8	DPP8
105	Dipeptidyl Peptidase VII (DPP-VII) (QPP)	DPP7
106	DPP9	DPP9
107	Fibroblast Activation Protein (FAP)	FAP
108	Matrix Metalloproteinase-8 (MMP-8)	MMP8
109	Matrix Metalloproteinase-13 (MMP-13)	MMP13
110	Cytochrome P450 2A6	CYP2A6
111	Factor Xa (fXa)	fXa
112	Thrombin	Thrombin
113	Urokinase (uPA)	uPA
114	Urokinase (uPA) Mutant (S190A)	uPA*M1
115	t-Plasminogen Activator (tPA)	tPA
116	Factor IXa (fIXa)	fIXa
117	Factor VIIa (fVIIa)	fVIIa
118	Phosphoinositide 3-Kinase (PI3K), C2beta	PI3KC2beta
119	Phosphoinositide 3-Kinase (PI3K), alpha	PI3Kalpha
120	Phosphoinositide 3-Kinase (PI3K), beta	PI3Kbeta
121	Phosphoinositide 3-Kinase (PI3K), delta	PI3Kdelta
122	Factor XIa (fXIa)	fXIa
123	Aurora Kinase B (Aurora-B)	AuroraB
124	XIAP-BIR3	XIAPBIR3
125	cIAP-1 BIR3	cIAP1BIR3
126	cIAP-2 BIR3	cIAP2BIR3
127	Protein Farnesyltransferase (PFT)	PFT
128	Protein-Tyrosine Phosphatase 1B (PTP1B)	PTP1B
129	T-Cell Protein-Tyrosine Phosphatase (TCPTP)	TCPTP
130	11-beta-Hydroxysteroid Dehydrogenase 1 (11-beta-HSD1)	11betaHSD1
131	11-beta-Hydroxysteroid Dehydrogenase 2 (11-beta-HSD2)	11betaHSD2
132	Trypsin	Trypsin
133	Geranylgeranyl Transferase (GGTase-I)	GGTase1
134	Rho-kinase (ROCK I)	ROCK1
135	Hepsin	Hepsin
136	Plasmin	Plasmin

137	beta-Secretase (BACE-1)	BACE1
138	Phosphodiesterase Type 4 (PDE4B)	PDE4B
139	Phosphodiesterase Type 4 (PDE4D)	PDE4D
140	Phosphodiesterase Type 1 (PDE1B)	PDE1B
141	Phosphodiesterase Type 10 (PDE10A)	PDE10A
142	Phosphodiesterase Type 11 (PDE11A)	PDE11A
143	Phosphodiesterase Type 2 (PDE2A)	PDE2A
144	Phosphodiesterase Type 3 (PDE3B)	PDE3B
145	Phosphodiesterase Type 5 (PDE5A)	PDE5A
146	Phosphodiesterase Type 7 (PDE7B)	PDE7B
147	Phosphodiesterase Type 3 (PDE3A)	PDE3A
148	cAMP-Dependent Protein Kinase (PKA)	PKA
149	Tyrosine Kinase JAK3	JAK3
150	Rho-kinase (ROCK II)	ROCK2
151	Mitogen-Activated Protein Kinase 10 (JNK3)	JNK3
152	Mitogen-Activated Protein Kinase 8 (JNK1)	JNK1
153	Mitogen-Activated Protein Kinase 9 (JNK2)	JNK2
154	beta-Secretase (BACE-2)	BACE2
155	Chymotrypsin	Chymotrypsin
156	Aldehyde Reductase (ALR1)	ALR1
157	Aldose Reductase (ALR2) Mutant (C298A/W219Y)	ALR2*M1
158	B-Raf Protein Kinase Mutant (V600E)	bRaf*M1
159	C-Raf Protein Kinase	cRaf
160	Integrin alpha4beta1 (VLA-4)	VLA4
161	Integrin alpha4beta7	LPAM1
162	17-beta-Hydroxysteroid Dehydrogenase 1 (17-beta- HSD1)	17betaHSD1
163	Estrogen Receptor (ER-alpha)	ERalpha
164	Estrogen Receptor (ER-beta)	ERbeta
165	ML-IAP-BIR	MLIAPBIR
166	cIAP1-BIR3	cIAP1BIR3
167	Methionine Aminopeptidase Type 2 (MetAP2)	MetAP2
168	Methionine Aminopeptidase Type 1 (MetAP1)	MetAP1
169	Tyrosine Kinase ACK1	ACK1
170	Tyrosine Kinase c-Met	cMet
171	Dihydrofolate Reductase (DHFR)	DHFR
172	Thymidylate Synthase (TS)	TS
173	Androgen Receptor (AR)	AR
174	Glucocorticoid Receptor (GR)	GR
175	Progesterone Receptor (PR)	PR
176	Mineralocorticoid Receptor (MR)	MR
177	Glucocerebrosidase (Acid beta-Glucosidase)	Alglucerase
178	alpha-Glucosidase (alpha-Glu)	Maltase
179	Thyroid Hormone Receptor (TR-alpha)	TRalpha
180	Thyroid Hormone Receptor (TR-beta)	TRbeta
181	Integrin alpha2beta1	CD29
182	Histone Deacetylase 1 (HDAC1)	HDAC1

183	Histone Deacetylase 10 (HDAC10)	HDAC10
184	Histone Deacetylase 2 (HDAC2)	HDAC2
185	Histone Deacetylase 3 (HDAC3)	HDAC3
186	Histone Deacetylase 4 (HDAC4)	HDAC4
187	Histone Deacetylase 6 (HDAC6)	HDAC6
188	Histone Deacetylase 8 (HDAC8)	HDAC8
189	Histone Deacetylase 4 (HDAC4) Mutant (H976Y)	HDAC4*M1
190	Inosine Monophosphate Dehydrogenase Type 1 (IMPDH1)	IMPDH1
191	Inosine Monophosphate Dehydrogenase Type 2 (IMPDH2)	IMPDH2
192	Cathepsin B	CatB
193	Cathepsin K	CatK
194	Cathepsin L	CatL
195	Cathepsin S	CatS
196	Cannabinoid Receptor 1	CB1
197	Cannabinoid Receptor 2	CB2
198	Estrogen-Related Receptor, gamma	ERRgamma2
199	Diacyl Glycerolacyltransferase 1 (DGAT-1)	DGAT1
200	Delta Opioid Receptor	DOR1
201	NK1 Receptor	NK1R
202	Mammalian Target of Rapamycin (mTOR)	mTOR
203	Cholecystokinin-1 Receptor	CCK1R
204	Cholecystokinin-2 Receptor	CCK2R
205	Cytochrome P450 2C9	CYP2C9
206	Adenosine Receptor A1	AdenosineA1R
207	Adenosine Receptor A2A	AdenosineA2R
208	hERG Potassium Channel 1	hERG
209	Adenosine A3 Receptor	AdenosineA3R
210	Cytochrome P450 2C19	CYP2C19
211	Cytochrome P450 2D6	CYP2D6
212	Dopamine Receptor D2	DopamineD2R
213	Serotonin Receptor 1A	5HT1A
214	Apoptosis Regulator Bcl-2	Bcl2
215	Apoptosis Regulator Bcl-xL	BclxL
216	Matriptase	Matriptase
217	Peroxisome Proliferator-Activated Receptor alpha	PPARalpha
218	Peroxisome Proliferator-Activated Receptor delta	PPARdelta
219	Peroxisome Proliferator-Activated Receptor gamma	PPARgamma
220	Kappa Opioid Receptor	KOR1
221	Mu Opioid Receptor	MOR1
222	Nociceptin Receptor (ORL1 Receptor)	ORL1R
223	Tumor Progression Loci-2 Kinase (COT)	COT
224	Sphingosine 1-Phosphate Receptor 1	S1P1
225	Sphingosine 1-Phosphate Receptor 3	S1P3
226	Sphingosine 1-Phosphate Receptor 4	S1P4
227	Sphingosine 1-Phosphate Receptor 5	S1P5

228	Dopamine Transporter (DAT)	DAT
229	Norepinephrine Transporter (NET)	NET
230	Serotonin Transporter (SERT)	SERT
231	Histamine H3 Receptor	HH3R
232	Histamine H4 Receptor	HH4R
233	Triacylglycerol Hydrolase	TGH
234	Fatty Acid Amide Hydrolase	FAAH
235	TNF-alpha-Convertng Enzyme	ADAM17
236	Matrix Metalloproteinase-14 (MMP-14)	MMP14
237	Matrix Metalloproteinase-7 (MMP-7)	MMP7
238	Kallikrein	Kallikrein
239	Neutrophil Elastase (NE)	NE
240	Kinesin Spindle Protein Eg5	Eg5
241	Leukotriene A4 Hydrolase	LTA4hydrolase
242	Matrix Metalloproteinase-12 (MMP-12)	MMP12
243	Protein-Tyrosine Phosphatase SHP-2	SHP2
244	Protein-Tyrosine Phosphatase SHP-1	SHP1
245	Mitogen- and Stress-Activated Protein Kinase 1 (MSK1)	MSK1
246	Ribosomal S6 Kinase 1 (RSK-1)	RSK1
247	MRCK alpha	MRCKalpha
248	17-beta-Hydroxysteroid Dehydrogenase 2 (17-beta-HSD2)	17betaHSD2
249	GABAA Receptor, alpha-1, beta-3, gamma-2	GABA-ARa1b3g2
250	GABAA Receptor, alpha-2, beta-3, gamma-2	GABA-ARa2b3g2
251	GABAA Receptor, alpha-3, beta-3, gamma-2	GABA-ARa3b3g2
252	GABAA Receptor, alpha-5, beta-3, gamma-2	GABA-ARa5b3g2
253	GABAA Receptor, alpha-6, beta-3, gamma-2	GABA-ARa6b3g2
254	I-kappa-B Kinase 1 (IKK-alpha)	IKKalpha
255	I-kappa-B Kinase 2 (IKK-beta)	IKKbeta
256	PIM-2 Kinase	PIM2
257	Acyl Cholesterol Acyltransferase 1 (ACAT-1)	ACAT1
258	Gastric Inhibitory Polypeptide Receptor (GIPR)	GIPR
259	Glucagon Receptor	GLR

The list of 259 targets extracted from BindingDB is shown. The abbreviations are used in the text and heat maps.

Supplementary Table S2. Target communities.

Community	#Molecules	#Targets	#Target Pairs	Target List
1a	2252	59	123	11;12;13;14;15;26;27;28;29;30;31;34;35;36;37; 38;45;46;47;48;55;56;57;58;59;60;61;62;63;64; 65;67;68;69;70;71;73;74;81;82;83;84;85;86;91; 92;110;123;149;151;152;153;169;170;205;210; 211;223;256
1b	501	21	91	49;50;51;52;53;54;77;78;79;80;98;99;100;101;102; 108;109;235;236;237;242
2	182	17	44	21;22;23;24;25;39;40;41;42;43;44;134;148;150;245; 246;247
3	557	16	47	72;111;112;113;114;115;116;117;122;132;135;136; 155;216;238;239
4	106	7	10	200;201;203;204;220;221;222
5	117	11	39	97;138;139;140;141;142;143;144;145;146;147
6	37	5	6	124;125;126;165;166
7	115	9	31	1;2;3;4;5;6;7;8;93
8	292	5	7	103;104;105;106;107
9	122	8	22	182;183;184;185;186;187;188;189
10	72	4	4	128;129;243;244
11	107	6	7	206;207;208;209;240;241
12	46	6	10	16;118;119;120;121;202
13	8	5	9	249;250;251;252;253
14	20	4	6	224;225;226;227
15	316	4	6	192;193;194;195
16	35	4	6	173;174;175;176
17	239	4	6	87;88;89;90
18	35	4	5	17;18;19;20

The number of molecules (#Molecules), targets (#Targets) and target pairs (#Target Pairs) for each community is reported. Target IDs are provided according to Table S1.

Supplementary Table S3. Scaffolds.

ID	Scaffold	#Molecules
1	c1ccccc1	182
2	C(Nc1ccccc1N2CCCC2)c3cccc(Oc4ncccc4c5ccnnc5)c3	5
3	C(NC1C[NH2+])CC1OCc2ccc(Cc3ccccc3)cc2)c4ccccc4	7
4	c1ccc(cc1)c2cccc(c2)c3cc4ccccc4[nH]3	13
5	C(Nc1ccc(cc1)n2cnc3ccccc23)c4ccnnc4c5ccccc5	10
6	C(Nc1ccccc1CNc2ccccc2)c3ccccc3	13
7	C(CNc1ncc2cc(cnc2n1)c3ccccc3)C[NH+]4CC[NH2+]CC4	15
8	o1c(c2ccccc2)c(c3ccccc3)c4ennc14	7
9	C1Nc2ccccc2C1=Cc3ccc[nH]3	19
10	C(c1ccccc1)c2cnc(NC3CCNCC3)nc2	18
11	C(C[NH+]1CCCC1)Nc2cc(nc(n2)c3occcc3)n4cccn4	9
12	C(Oc1ccccc1)C2CCCN2Sc3ccc4N(Cc5ccccc5)CCc4c3	7
13	C(NCc1ccccc1)c2ccccc2	7
14	C(Nc1ccc(Oc2ccnc3ccccc23)cc1)C4=CN(C4)c5ccccc5	6
15	C(CN1CCc2ccccc12)NCC(CC3CCCC3)Nc4oc5ccccc5n4	6
16	C(Oc1ccccc1)N(c2ccccc2)c3ccnc(Nc4ccc(cc4)N5CC[NH2+]CC5)n3	8
17	C(NCc1ccc[nH+]1)c2ccccc2	7
18	C(c1ccccc1)c2ccc3nc(Nc4ccccc4)cn3c2	11
19	C(Cc1ccccc1)c2ccccc2	23
20	C([NH+]1CCC2(CC1)CNCN2c3ccccc3)c4ccccc4	8
21	C1Nc2ccc(cc2C=C1)c3ccccc3	14
22	C1Nc2cc(Nc3ccccc3)[nH+]cc2C=C1c4ccccc4	6
23	C(Nc1ccccc1c2c1Cc3c2n[nH]c3c4ccccc4)[N-]Nc5ccccc5	15
24	c1ccc2c(c[nH]c2c1)c3ccnnc3	7
25	C(CNc1cc2ncc(cc2cn1)c3ccccc3)C[NH+]4CC[NH2+]CC4	8
26	C(CNc1ccccc1)NCC(CC2CCOCC2)CC3CCCC3	6
27	c1ccc(cc1)c2cccc(c2)c3nc4ccccc4[nH]3	14
28	C(COCc1ccc(Cc2ccccc2)cc1)Cc3ccccc3	5
29	C1C=C(Oc2ccccc12)c3ccccc3	8
30	C(Cc1ccccc1)Nc2cccc[nH+]2	21
31	C(Nc1ccccc1)c2ccccc2	16
32	c1ccc(cc1)c2nc3ccccc3[nH]2	13
33	C(Cc1ccccc1)N2Cc3[nH]nc(NCc4ccc(cc4)N5CC[NH2+]CC5)c3C2	6
34	C(Nc1ccccc1)c2cccc(OC3CCOC3)c2	7
35	C1CCc2[nH]c(C=C3CNc4ccccc34)cc2C1	10
36	c1cnnc1	5
37	C(NCc1cc2ccccc2n1Cc3ccccc3)c4ccccc4	15
38	c1ccc(cc1)c2cc3ccccc3[nH]2	8
39	C1c2ccccc2c3n[nH]c(c13)c4ccccc4	9
40	C(Oc1ccc2c(Nc3ccccc3)ncnc2c1)C4CC[NH2+]CC4	13
41	C(N(c1ccccc1)n2cnc2)c3ccccc3	9
42	N(c1ccccc1)c2nccc(n2)c3ccccc3	19
43	C(Nc1n[nH]c2nc(ccc12)c3ccccc3)C4CC4	5
44	C1CC(c2ccncc2)c3ccc(cc3C1)c4ccccc4	7

45	C(CNCCNCCNC(CNCCNCCNCCc1cccc1)Cc2c[nH]c3cccc23)Cc4cccc4	9
46	C(Nc1cccc1)Nc2ccc(cc2)c3cccc4oncc34	32
47	C(CNCc1cccc1)C[NH2+]CCN2CCCC2	6
48	C(Nc1cccc1CNc2cccn2)c3cc(Cn4ccnc4)cs3	6
49	C(NC1CCCC1)C(Cc2cccc2)c3csc4cccc34	6
50	C(Oc1ccc(Nc2nenc3cccc23)cc1)c4cccc4	9
51	c1ccc(cc1)c2cccc2	31
52	C(Nc1n[nH]c2CNCC12)c3cccc3	5
53	C1Nc2ccc3nsc3c2C1=CNc4cccc4	14
54	C(c1cccc1)c2cnc3[nH]ncc3c2	13
55	C1Nc2cc[nH+]cc2C=C1c3cccc3	7
56	C(CN1CCCC1)Cc2cccc2	18
57	c1ccc2nnc2c1	7
58	C(Oc1ccc(Nc2nenc3cc(C#CC4CCC[NH2+]4)sc23)cc1)c5cccc5	14
59	C(NCc1cc(c[nH]1)c2n[nH]cc2c3cccc3)c4cccc4	13
60	C(CCCNCCc1c[nH]c2cccc12)CCCNc3c4CCCCc4[nH+]c5cccc35	5
61	N(c1cccc1)c2nenc3cccc23	38
62	C(CNc1cccc1)NCC(CC2CCCC2)Nc3oc4cccc4n3	9
63	N(c1cccc1)c2nccc(n2)c3cc[nH]c3	12
64	C(CCc1cccc1)CN2CCc3[nH]ncc3C2	6
65	C(Cc1nsc1CNc2cccc2)c3ccncc3	9
66	C(CN1CC=NC=C1c2cccc2)NCc3cccc3	26
67	C(NSc1ccc2NCC(=NNc3cccc3)c2c1)c4cccc4	14
68	C(OC1CCCCO1)c2en(nn2)c3cccc3	6
69	C(C[NH2+]CCN1CCCC1)CN2CCc3cccc3C2	8
70	C([NH+]1CCC(CC1)c2nnc([nH]2)c3cccn3)c4ccc(cc4)c5nc6nnc6cc5c7cccc7	5
71	C(Cc1enccc1c2oc(Nc3cccc3)nn2)c4ccncc4	7
72	C1Nc2cc(ccc2C1=Cc3ccc[nH]3)c4cccc4	10
73	C(CNc1mnc1)NCC(Cc2cccc2)NC[N-]Sc3cccc3	5
74	C1CCc2[nH]c(C=C3CNc4cc(ccc34)c5cccc5)cc2C1	5
75	C(C1=C(CCCC1)Nc2ccc(cc2)C3=NNCCC3)c4cccc4	16
76	C1Nc2nenc(Nc3cccc3)c2C1=Cc4ccc[nH]4	6
77	C(Oc1esc(c1)n2cnc3cccc23)c4cccc4	8
78	C(N(Cc1ccc(Oc2cccc2)cc1)c3cccc3)c4cccc4	5
79	N(Sc1cccc1)c2cccc2	7
80	C(Nc1nc(Nc2cccc2)cc3ncn13)c4cccc4	17
81	C(Sc1cccc1)C2CCNCC2	11
82	C(Cc1cccc1)NCNc2ccc(Oc3ccnc4ccsc34)cc2	8
83	C(CNc1ncc2cc(c(NCNe3cccc3)nc2n1)c4cccc4)C[NH+]5CC[NH2+]CC5	11
84	C1Cc2cccc2C1=Cc3cnc[nH]3	10
85	C(C[NH2+]Cc1cccc1)NC2CC2	8
86	C(Nc1cccc1)Oc2ccc3NC4[NH2+]CCC4c3c2	20
87	C(CNCOc1cccc1)NCCc2cccc2	5
88	C1c2cccc2c3n[nH]c(c13)c4cccc4	24
89	C(C1=Nc2c(CN1)enn2c3cccc3)c4cccc4	22
90	C(CN1CCCC1)Cc2ccc(cc2)c3cccc3	45
91	C(CNCCc1ncc1)NCCNCCc2cccc2	11
92	N(Sc1cccc1)c2mnc2	23

93	C1Cc2ccccc2N1	28
94	C1CC2(CC[NH+]1C(c3ccccc3)c4ccccc4)CNCN2c5ccccc5	21
95	C1NCC(=C1c2c[nH]c3ccccc23)c4c[nH]c5ccccc45	7
96	C(c1ccccc1)n2ncc3cc(Nc4[nH+]enn5ccccc45)ccc23	14
97	C1CC2CCC3=C(CC=C4CCOCC34)C2C1	16
98	C(CCc1ccccc1)CN2CCn3cnnc3C2	6
99	C(CN1CCc2ccccc12)NCC(CCC3CCCCC3)NCc4ccccc4	6
100	C1NC(C=C1c2ccccc2)c3ccccc3	5
101	C(CNc1nncs1)NCC(Cc2ccccc2)NSc3ccccc3	15
102	C(CNCc1onc(Cc2ccccc2)n1)Cc3ccccc3	26
103	c1cc[nH]c1	5
104	C(Cc1ccccc1)Oc2ccccc(CNCC3CCN(CC3)c4cc[nH+]cc4)c2	18
105	C(Nc1ccccc1)c2ccccc(Nc3ncccc3c4cncn4)c2	5
106	C1Nc2ccccc2C1=CNc3ccccc3	14
107	C1Cc2ccccc3c4c5CNCc5c6c7ccccc7[nH]c6c4n(C1)c23	6
108	C1Cc2ccc(Nc3nc(Nc4ccccc4)c5cc[nH]c5n3)cc2N1	10
109	C(CN1CCCC1)Cc2onc(n2)c3ccccc3	28
110	N(Sc1ccc(cc1)[n+]2ccccc2)c3nncs3	5
111	C(Cc1ccccc1)NC(CNCc2onc(Cc3ccccc3)n2)Cc4ccccc4	5
112	C1NCc2c1c3c(ccc4cc[nH]c34)c5[nH]c6ccccc6c25	32
113	C(Nc1ccc2NC(=NCc3ccc(C=Cc4ccccc4)s3)Nc2c1)c5ccccc5	6
114	N(c1ccccc1)c2nenc3[nH]nc(Nc4ccccc4)c23	10
115	C(Sc1cnsc1)c2ocn2	9
116	C(CNc1ccccc1)NCC(Cc2ccccc2)NCc3ccccc3	8
117	c1encc(c1)c2ccnc(n2)n3cccn3	7
118	N(Sc1ccc(cc1)[n+]2ccc(cc2)c3ccccc3)c4nncs4	6
119	C(Nc1ccccc1)[N-]NCc2ccccc2	5
120	C1NCc2c1c3c4ccccc4[nH]c3c5[nH]c6ccccc6c25	12
121	C(N1Cc2ccccc2C3(CC[NH+](CC3)C4CCCC4)C1)c5ccccc5	6
122	C(Cc1cncnc1c2oc(Nc3ccccc3)nn2)c4cncnc4	6
123	C1c2ccccc2c3[nH+]c4ccccc4cc13	5
124	C(Nc1nncnc2oc(cc12)c3ccccc3)C4CC4	5
125	C(OCc1cn(nn1)C2CCCCO2)c3ccccc3	10
126	N(c1ccccc1)c2nenc3[nH]nc(c4ccccc4)c23	6
127	C(c1ccccc1)c2ccc3ncc(Cc4ccccc4)n3c2	14
128	C(Nc1ccc2NC(=NCc3ccccc3)Nc2c1)c4ccccc4	9
129	C(NC1CC1)C(CSc2ccccc2)[NH2+]Cc3ccccc3	5
130	C(Cc1ncccc1c2oc(Nc3ccccc3)nn2)c4cncnc4	12
131	N(c1ccccc1)c2cc(n[nH]2)c3ccccc3	6
132	C1NC=Nc2c1enn2c3ccccc3	11
133	C1CCC(C1)N2CC=Cc3cnc(Nc4ccc(cc4)N5CC[NH2+]CC5)nc23	8
134	C1Nc2ccccc2C1=NNc3ccccc3	35
135	N(c1ccccc1)c2nc(Nc3ccccc3)c4cc[nH]c4n2	8
136	C1Cc2ccccc2C1	21
137	C(C[NH+]1CC[NH2+]CC1)Nc2cc(nc(n2)c3occcc3)n4cccn4	6
138	C([NH+]1CCCC1)c2ccccc2c3ccc(cc3)N4CCc5enn(c5C4)c6ccccc6	12
139	c1cc[nH+]cc1	12
140	C1Cc2c(CO1)c[nH]c2C=C3CNc4ccccc34	9

141	c1ccc2cnccc2c1	5
142	C1CN(CC1OC2CCCC2)c3nenc4cccc34	7
143	N(Sc1ccc(cc1)[n+]2ccc(cc2c3cccc3)c4cccc4)c5nncs5	6
144	C1Nc2cccc2C=C1	10
145	C(Nc1cccc(CNc2enc(Nc3ccc(cc3)N4CC[NH2+]CC4)nc2)c1)c5cccc5	8
146	C(CNCc1cccc1)NCCNc2cccc2	9
147	C(NC1CCCC1OCc2ccc(Cc3cccc3)cc2)c4cccc4	13
148	C1Cc2cccc2C=C1c3ccnc3	7
149	c1ccc(cc1)c2enc3nenc3c2	45
150	C1Nc2ccc3c4cccc4OC(=Cc5cccs5)c3c2C=C1	6
151	C1CCCC1	14
152	C(C1C[NH2+]C(CN2CCCC2)C1)N3Cc4cccc4C3	7
153	c1ccc(cc1)c2enc3cncc3c2	22
154	C(Nc1cccc1)N2CCc3cccc23	8
155	C(Nc1cccc1N2CCCC2)c3cccc(c3)C#Cc4cnnc4	9
156	C(Nc1cccc1)c2ccc(Oc3ncccc3c4cnnc4)c2	5
157	C(c1cccc1)n2ccn2	23
158	C([N-]NCc1cccc1)[N-]Sc2cccc2	10
159	o1ncc(n1)c2nc3encc3[nH]2	8
160	N(c1cccc1)c2cc(Nc3cccc3)ncn2	17
161	C(Oc1cccc1)C2CCN2Sc3ccc4N(Cc5cccc5)CCc4c3	5
162	C([NH+]1CCC1)c2ccc(cc2)c3noc(n3)c4cccc4	10
163	C(c1cncc1)c2nnc(Nc3cccc3)c4cccc24	17
164	C(Nc1cccc1CNc2cccn2)c3cccs3	18
165	C(Nc1cccc1)Nc2ccc(Oc3ccc4nc[nH]c4c3)cc2	16
166	C1CC(CC[NH2+]1)Sc2ccc(Sc3cccc3)cc2	5
167	C(Oc1cnnc1)c2cccc2	5
168	C(NSc1cccc1)c2cccc2	20
169	C(Cc1nsc1CNc2cccc2)c3conc3	5
170	C(OCc1nc2ennc(Cc3cccc(OC4CCCC4)c3)c2n1)c5cccc5	21
171	C1[N-]CC(=Cc2cccc2)S1	21
172	N(c1cccc1)c2ccnc(Nc3cccc3)n2	23
173	C(C1=C(CCCC1)Nc2ccc(cc2)C3=NNCC3)c4cccc4	14
174	c1ccc(cc1)c2c[nH]c3cccc23	15
175	C1c2cccc2c3n[nH]c(c13)c4ccsc4	5
176	C1Nc2nenc2C=C1c3cccc3	17
177	C(NCc1en(mn1)C2CCCCO2)c3cccc3	10
178	C(CCc1cccc1)CN2CCCNCC2	21
179	N(c1cccc1)c2nccc(n2)c3enn4ncccc34	29
180	C(NC1C[NH2+]CCCC1OCc2ccc(Cc3cccc3)cc2)c4cccc4	7
181	C(CNCCNCCc1cccc1)CSCc2cccc2	26
182	C(C(Cc1cccc1)NSc2cccc2)N3CCCC3	12
183	C(CN1CC(=NC=C1c2cccc2)NCCc3cccc3)NCc4cccc4	24
184	C1CC([NH+]2CCC3(CC2)CNCN3c4cccc4)c5cccc5C1	17
185	C(Nc1ccc2[nH+]ccc(Nc3cccc3)c2c1)c4c[nH]cn4	7
186	C1Cc2cccc2C1=Cc3ccnc3	17
187	C(Nc1cccc1)Nc2ccc(cc2)c3csc4nenc34	37
188	C(CN1CCCC1)Cc2ccc(cc2)C3=CNCC=C3	17

189	<chem>C1Nc2nc(Nc3ccccc3)ncc2C=C1c4ccccc4</chem>	38
190	<chem>C(CNC(CNCCNC(CNCCNCCNCCc1ccccc1)Cc2c[nH]c3ccccc23)Cc4ccccc4)Cc5ccccc5</chem>	5
191	<chem>C(CNC1CCCN(CC1)Sc2cccn2)NCc3oc4ccccc4c3</chem>	13
192	<chem>C(NC1=NCCO1)c2csc(CNc3ccccc3CNc4cccn4)c2</chem>	7
193	<chem>o1c2ccccc2c3ccccc13</chem>	18
194	<chem>N(c1ccccc1)c2ccccc2</chem>	17
195	<chem>C(Nc1ccccc1)Nc2ccc(cc2)c3coc4ncccc34</chem>	5
196	<chem>C(Nc1ccc(cc1)n2ccnc2)c3ccnn3c4ccc5onccc5c4</chem>	10
197	<chem>C(CNCc1ccccc1)NCCN2CCc3ccccc23</chem>	6
198	<chem>C1CC2C(Nc3ccccc23)[NH2+]1</chem>	7
199	<chem>c1ccc2cc(ccc2c1)c3ccccc3</chem>	11
200	<chem>C(C[NH+]1CCCCC1)Nc2cccc3c2Cc4c3n[nH]c4c5ccccc5</chem>	5
201	<chem>c1ccc(cc1)N=Nc2cn[nH]c2</chem>	6
202	<chem>C(Oc1ccc2c(Nc3ccccc3)cc[nH+]c2c1)C4CC[NH2+]CC4</chem>	8
203	<chem>C1CCC(CC1)[NH+]2CCC3(CC2)CNCc4ccccc34</chem>	10
204	<chem>C1Cc2ccccc2C1=Cc3ccccc3</chem>	7
205	<chem>C(Cc1ccccc1)NC2=C(CNC=C2)c3nc4ccc(cc4[nH]3)C5CC[NH2+]CC5</chem>	5
206	<chem>c1ccc(cc1)n2cnc3ccccc23</chem>	39
207	<chem>c1ccc(cc1)n2cccn2</chem>	11
208	<chem>C1CN(CC[NH2+]1)c2ccc(cc2)c3ccccc3</chem>	7
209	<chem>C(C(Cc1ccccc1)NSc2ccccc2)N3CCNCC3</chem>	10
210	<chem>C(COc1ccc2c(Nc3cc(CCNe4ccccc4)[nH]n3)ncnc2c1)C[NH+]5CCCCC5</chem>	12

All scaffolds occurring in at least five active compounds are reported in SMILES format and the number of active compounds containing each scaffold is provided.

Supplementary Table S4. Target-selective scaffolds.

Community	Target	#Scaffolds	Scaffold List
1a	CDK1	2	18,54
1a	CDK2	17	39,42,52,53,57,63,88,89,106,115,127,131,132,134,140,175,200
1a	CDK4	9	10,23,95,107,112,120,133,160,172
1a	PIM1	2	29,171
1a	ITK	2	113,128
1a	IGF1R	4	103,108,135,205
1a	EGFR	5	50,76,96,114,126
1a	Src	7	7,16,22,55,83,176,189
1a	FGFR1	2	25,149
1a	PDGFRbeta	3	9,72,206
1a	Lck	1	145
1a	ABL	1	202
1a	HER2	1	58
1a	VEGFR2	16	8,35,40,61,65,71,74,105,122,130,153,156,165,169,187,195
1a	VEGFR1	1	163
1a	ZAP70	3	80,102,111
1a	AuroraA	1	33
1a	CDK5	1	24
1a	Tie2	3	2,105,155
1a	GSK3alpha	1	43
1a	CDK9	1	201
1a	ERK2	1	59
1a	GSK3beta	2	124,179
1a	CYP11B2	4	148,186,199,204
1a	CYP19	2	41,144
1a	CYP11B1	1	84
1a	CYP17	1	44
1a	FLT3	1	46
1a	STS	1	130
1a	AuroraB	1	210
1a	JNK1	3	17,30,139
1a	cMet	2	14,82
1a	COT	1	185
1b	CA2	12	1,31,68,79,92,110,118,125,130,136,143,177
1b	CA9	2	119,158
1b	MMP2	4	51,101,166,168
1b	CA14	1	174
1b	MMP8	1	73
1b	ADAM17	1	81
1b	MMP12	1	193
2	Akt1	2	70,159

2	PKCbeta2	1	180
2	PKCdelta	1	28
2	PKCeta	2	3,147
3	fXa	11	5,6,37,38,48,49,104,138,164,192,196
3	Thrombin	1	183
3	uPA	5	4,27,31,32,141
3	FVIIa	2	34,66
3	fXIa	1	91
3	Matriptase	2	182,209
3	NE	1	157
4	DOR1	2	45,190
4	ORL1R	5	20,94,121,184,203
5	PDE4B	1	170
5	PDE4D	1	207
5	PDE3B	2	75,173
5	PDE3A	1	142
7	CASP3	4	1,12,161,181
8	DPP4	10	47,56,64,69,90,98,109,152,178,188
9	HDAC4	1	151
9	HDAC4*M1	1	13
10	TCPTP	1	194
10	SHP2	2	67,134
11	AdenosineA2R	3	11,117,137
11	AdenosineA3R	2	36,167
11	Eg5	1	100
12	PI3Kalpha	1	97
14	S1P1	1	162
15	CatK	3	87,191,208
15	CatL	3	116,146,197
15	CatS	8	1,15,26,51,62,85,99,129
16	GR	3	21,78,150
17	AChE	3	60,86,123
17	BuChE	1	198
17	CE2	3	19,93,154
18	PLK1	1	77

For each scaffold, the target for which it has the strongest selective tendency is reported.

For each target, the number of scaffolds with selective tendency is also given.

Supplementary Table S5. Community-selective scaffolds.

Scaffold	Rank	Community	Median TS
48	1	3	5.176
164	2	3	4.799
178	3	8	4.032
196	4	3	4.000
192	5	3	3.539
100	6	11	3.529
141	7	3	3.431
126	8	1a	3.352
162	9	14	3.094
182	10	3	3.079
138	11	3	2.950
167	12	11	2.819
109	13	8	2.806
60	14	17	2.796
205	15	1a	2.699
59	16	1a	2.686
14	17	1a	2.633
37	18	3	2.628
6	19	3	2.624
145	20	1a	2.590
133	21	1a	2.572
104	22	3	2.562
98	23	8	2.531
9	24	1a	2.522
85	25	15	2.519
113	26	1a	2.388
129	27	15	2.312
16	28	1a	2.230
135	29	1a	2.201
154	30	17	2.197
152	31	8	2.174
99	32	15	2.151
41	33	1a	2.119
190	34	4	2.076
144	35	1a	2.033
11	36	11	2.020
119	37	1b	2.003
5	38	3	1.984
128	39	1a	1.965
188	40	8	1.910
68	41	1b	1.906
117	42	11	1.900
137	43	11	1.899
108	44	1a	1.899
193	45	1b	1.818

81	46	1b	1.814
155	47	1a	1.789
78	48	16	1.761
115	49	1a	1.758
160	50	1a	1.743
112	51	1a	1.72
69	52	8	1.713
131	53	1a	1.670
136	54	1b	1.620
79	55	1b	1.618
150	56	16	1.608
203	57	4	1.607
90	58	8	1.605
210	59	1a	1.594
158	60	1b	1.583
47	61	8	1.573
91	62	3	1.564
102	63	1a	1.498
35	64	1a	1.479
171	65	1a	1.479
40	66	1a	1.477
12	67	7	1.458
191	68	15	1.445
132	69	1a	1.439
125	70	1b	1.400
53	71	1a	1.380
64	72	8	1.375
173	73	5	1.362
94	74	4	1.350
208	75	15	1.345
26	76	15	1.343
88	77	1a	1.331
19	78	17	1.319
184	79	4	1.307
172	80	1a	1.301
179	81	1a	1.299
175	82	1a	1.297
20	83	4	1.289
7	84	1a	1.279
107	85	1a	1.277
93	86	17	1.25
15	87	15	1.245
49	88	3	1.243
111	89	1a	1.243
116	90	15	1.230
177	91	1b	1.216
70	92	2	1.197
57	93	1a	1.180

83	94	1a	1.163
61	95	1a	1.155
166	96	1b	1.125
181	97	7	1.122
174	98	1b	1.103
76	99	1a	1.097
209	100	3	1.087
199	101	1a	1.082
114	102	1a	1.082
34	103	3	1.069
156	104	1a	1.061
197	105	15	1.061
56	106	8	1.055
87	107	15	1.046
62	108	15	1.030
153	109	1a	1.025
54	110	1a	1.019
110	111	1b	1.000
127	112	1a	0.987
147	113	2	0.987
55	114	1a	0.966
27	115	3	0.933
36	116	11	0.928
22	117	1a	0.925
121	118	4	0.909
206	119	1a	0.892
66	120	3	0.891
159	121	2	0.885
45	122	4	0.878
200	123	1a	0.872
186	124	1a	0.870
148	125	1a	0.854
157	126	3	0.851
28	127	2	0.844
170	128	5	0.844
185	129	1a	0.831
18	130	1a	0.829
2	131	1a	0.826
39	132	1a	0.820
74	133	1a	0.818
77	134	18	0.816
183	135	3	0.778
67	136	10	0.775
63	137	1a	0.773
89	138	1a	0.771
72	139	1a	0.760
163	140	1a	0.739
42	141	1a	0.712

58	142	1a	0.706
118	143	1b	0.699
8	144	1a	0.676
4	145	3	0.673
38	146	3	0.653
65	147	1a	0.651
187	148	1a	0.648
120	149	1a	0.644
84	150	1a	0.629
80	151	1a	0.598
204	152	1a	0.588
106	153	1a	0.578
92	154	1b	0.564
189	155	1a	0.561
143	156	1b	0.554
124	157	1a	0.550
43	158	1a	0.544
29	159	1a	0.541
82	160	1a	0.530
13	161	9	0.508
44	162	1a	0.471
207	163	5	0.469
17	164	1a	0.467
30	165	1a	0.463
21	166	16	0.452
146	167	15	0.451
103	168	1a	0.446
161	169	7	0.442
50	170	1a	0.431
3	171	2	0.410
10	172	1a	0.398
96	173	1a	0.372
149	174	1a	0.367
46	175	1a	0.366
32	176	3	0.362
52	177	1a	0.361
168	178	1b	0.354
201	179	1a	0.339
142	180	5	0.332
176	181	1a	0.326
198	182	17	0.301
151	183	9	0.301
75	184	5	0.283
180	185	2	0.279
97	186	12	0.271
169	187	1a	0.260
165	188	1a	0.255
140	189	1a	0.254

73	190	1b	0.250
33	191	1a	0.250
86	192	17	0.192
71	193	1a	0.179
195	194	1a	0.170
130	195	1a	0.140
25	196	1a	0.136
105	197	1a	0.130
122	198	1a	0.129
139	199	1a	0.127
123	200	17	0.090
95	201	1a	0.089
101	202	1b	0.073
202	203	1a	0.030
194	204	10	0.026
24	205	1a	0.000
23	206	1a	0.000

The 206 community-selective scaffolds are ranked in the order of descending median *TS* values. Community and scaffold IDs are provided.

Supplementary Table S6. Drugs containing community-selective scaffolds.

ScaffoldID	Drug_Name
19	Dienestrol
19	Propoxyphene
61	Erlotinib
92	Sulfamethizole
93	Ropinirole
98	Sitagliptin
136	Rasagiline
139	Pyridostigmine
144	Procaterol
151	Dinoprostone
151	Dinoprost
151	Misoprostol
151	Alprostadil
151	Carboprost
174	Fluvastatin
194	Meclofenamic acid
194	Diclofenac
194	Lumiracoxib
194	Mefenamic acid

For the 11 of 206 community-selective scaffolds that are also found in DrugBank, drugs containing each scaffold are listed.

Supplementary Figure Legends

Supplementary Figure S1. Distribution of targets and scaffolds over communities.

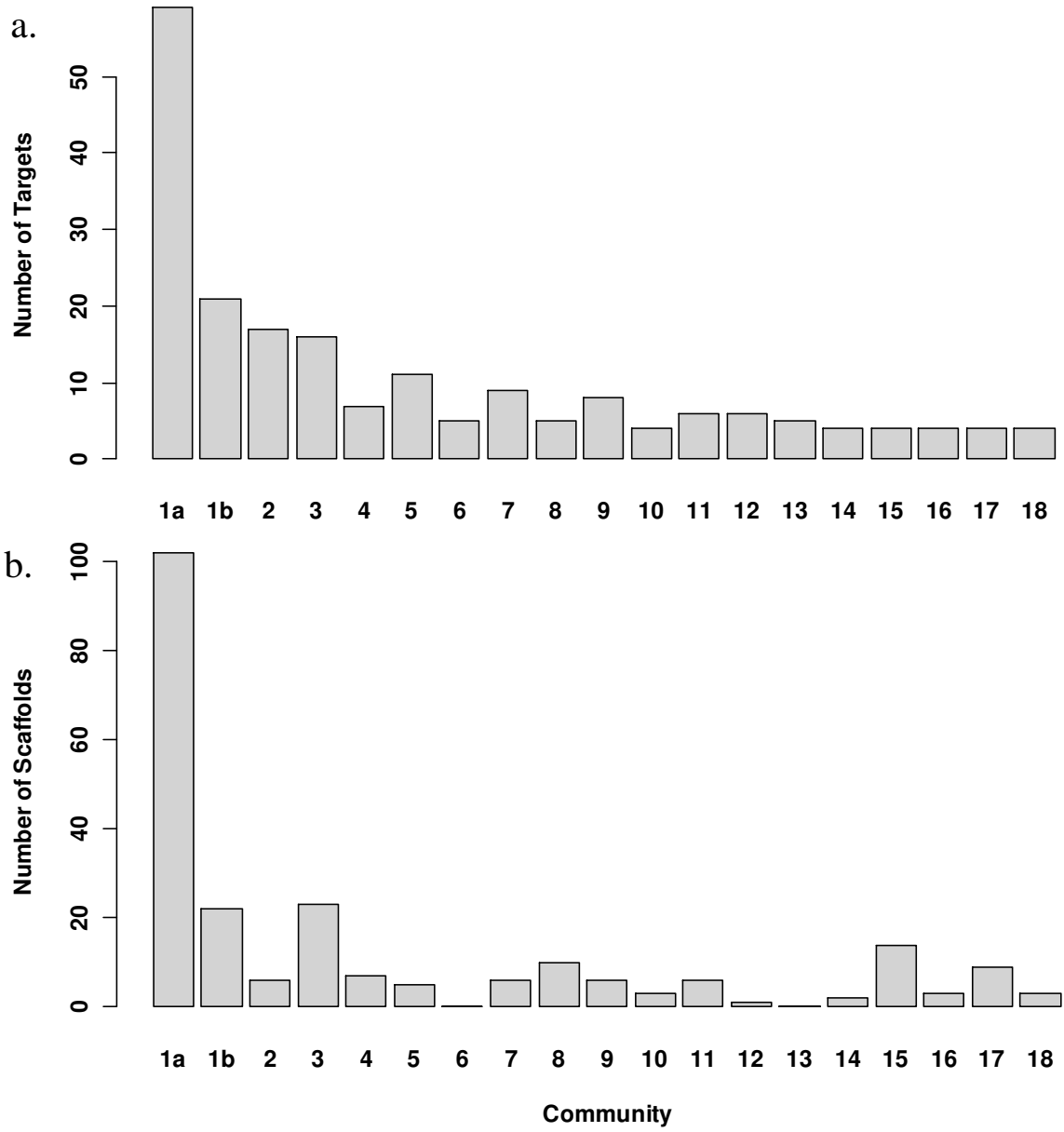
The number of targets and scaffolds are reported for each community.

Supplementary Figure S2. Target selectivity profiles. Heat maps representing the target selectivity profiles of all communities are shown. Targets form columns and scaffolds rows. A cell corresponding to a scaffold-target combination is filled if the scaffold is present in at least five compounds active against the target and color-coded according to median *TS* values

Supplementary Figure S3. Community selectivity profiles. Heat maps representing the community selective profiles of all scaffolds are shown (in segments of 30 scaffolds). Columns represent communities and rows scaffolds. Cells are color-coded according to median *TS* value.

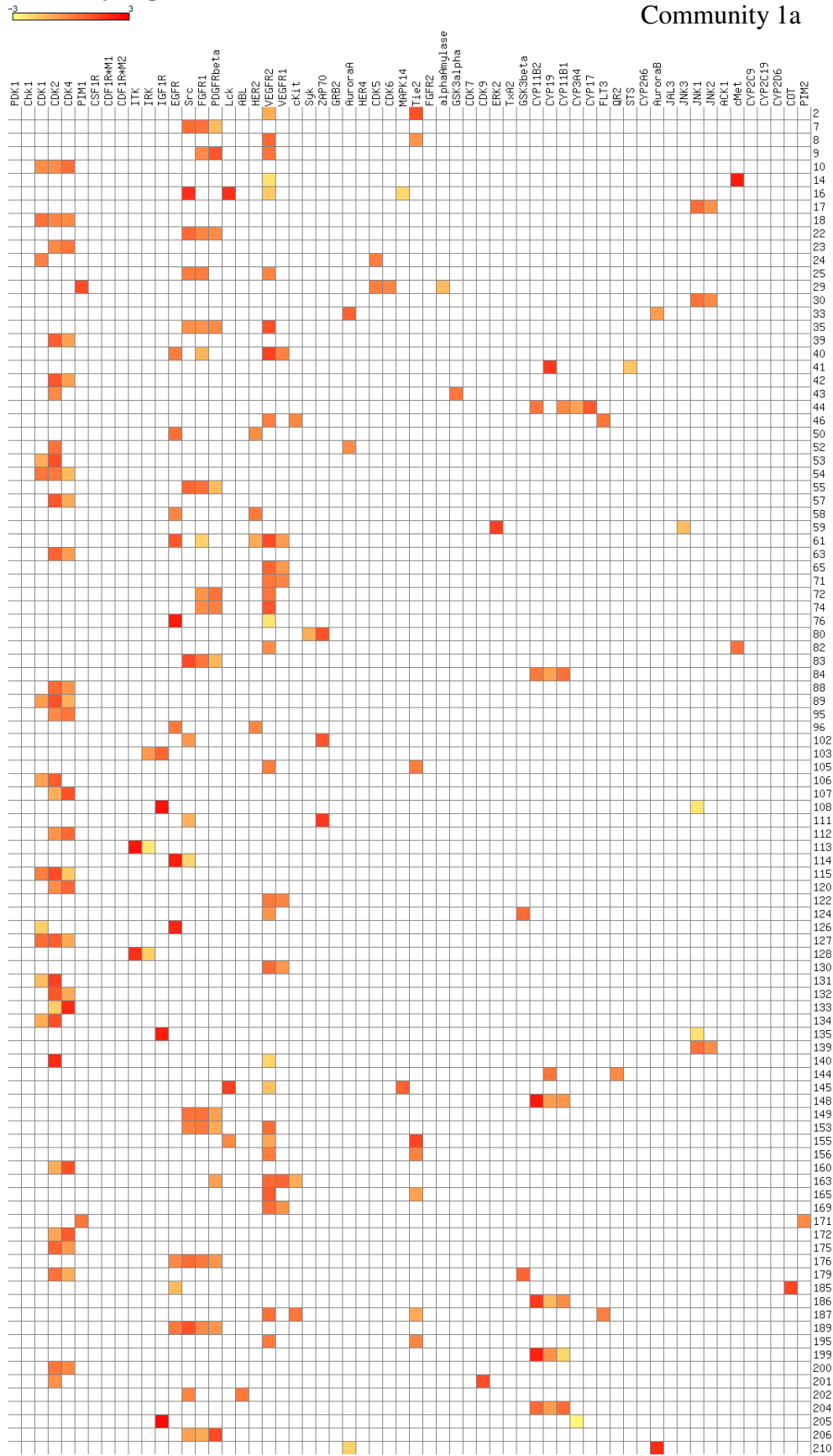
Supplementary Figure S4. Pubchem analysis. Bioassay data from human confirmatory assays available in Pubchem were collected. Active compounds with reported potency values (> 1 millimolar) were selected and assay pairs were generated by determining shared active compounds. The inserted histogram shows distribution of number of shared compounds. Only three assay/target pairs were found to contain at least five common active compounds.

Supplementary Figure S1

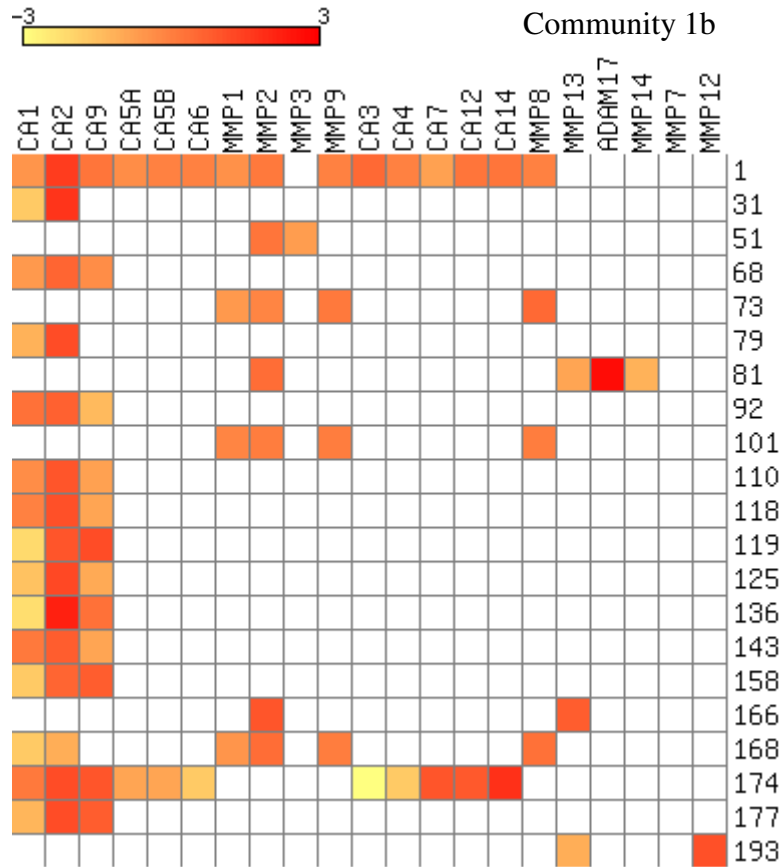


Supplementary Figure S2a

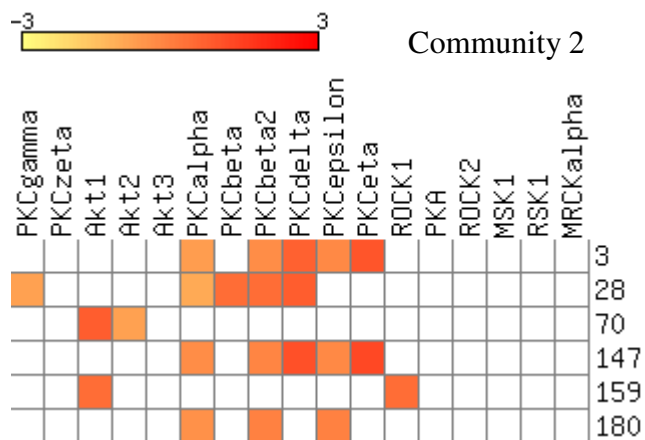
Community 1a



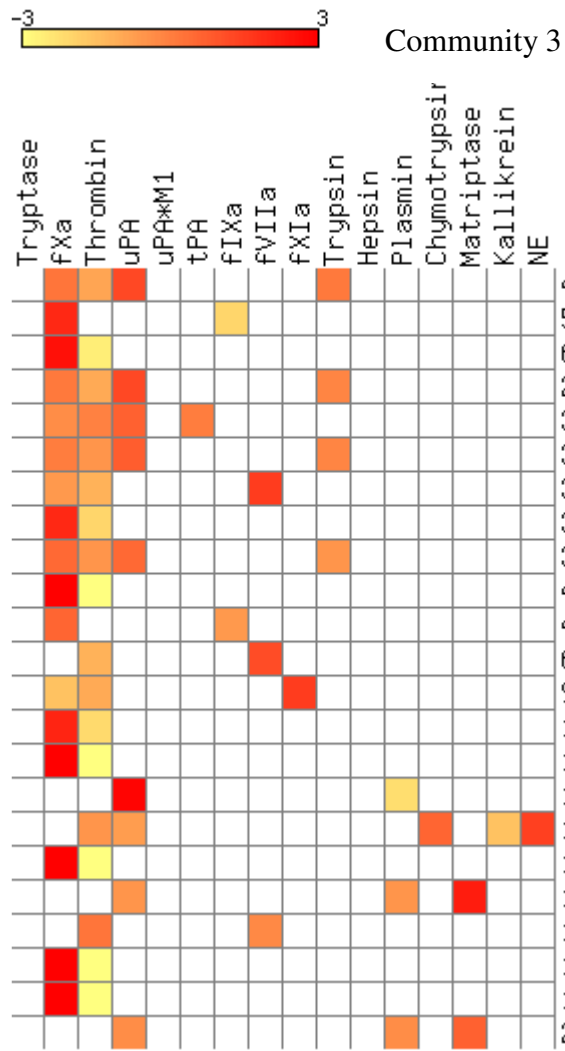
Supplementary Figure S2b



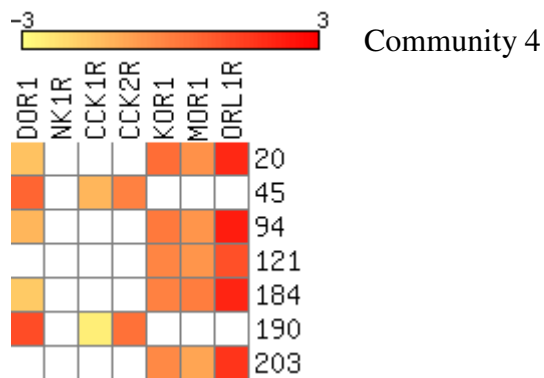
Supplementary Figure S2c



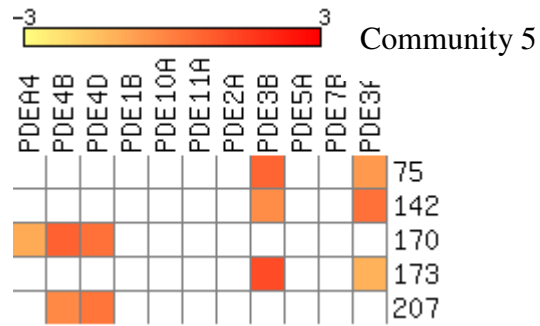
Supplementary Figure S2d



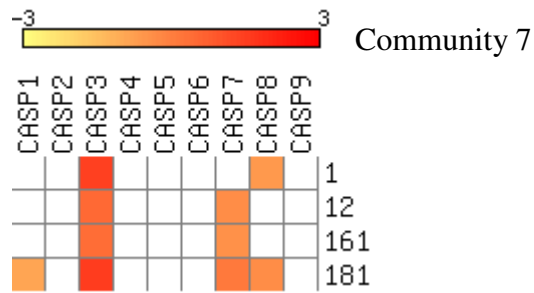
Supplementary Figure S2e



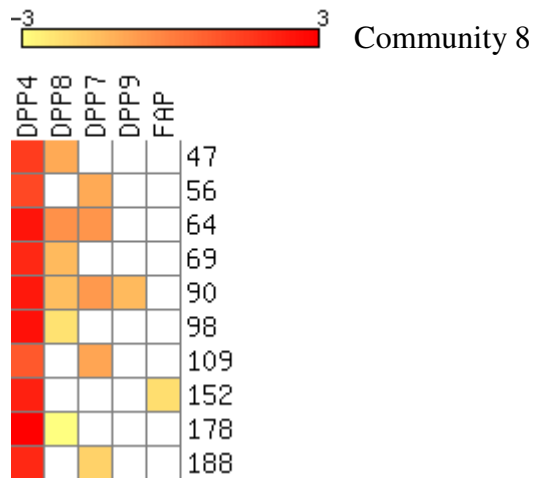
Supplementary Figure S2f



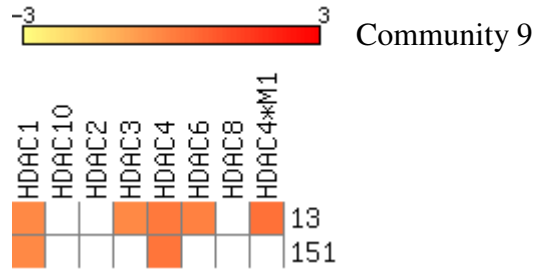
Supplementary Figure S2g



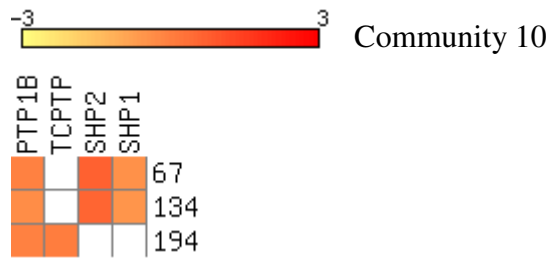
Supplementary Figure S2h



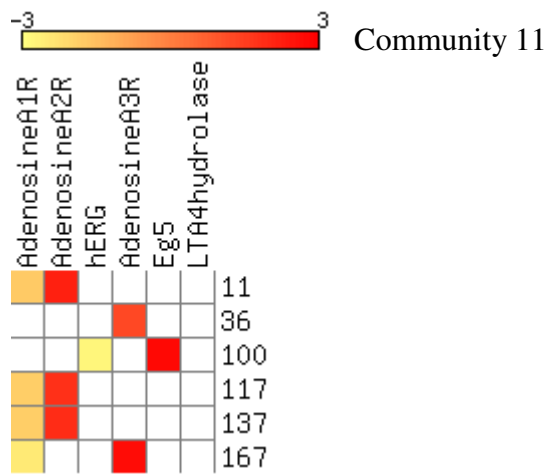
Supplementary Figure S2i



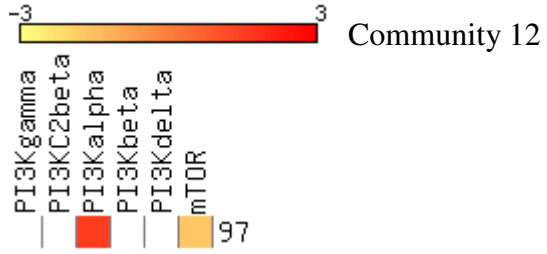
Supplementary Figure S2j



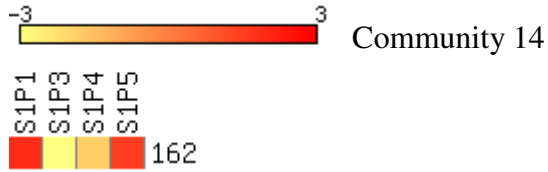
Supplementary Figure S2k



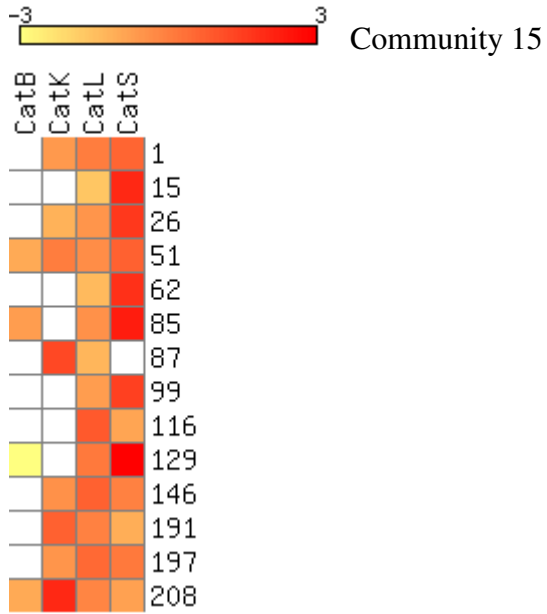
Supplementary Figure S2l



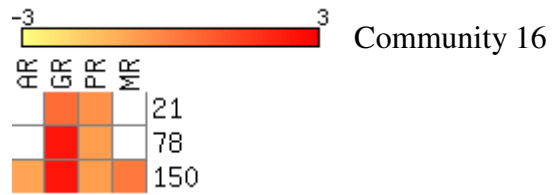
Supplementary Figure S2m



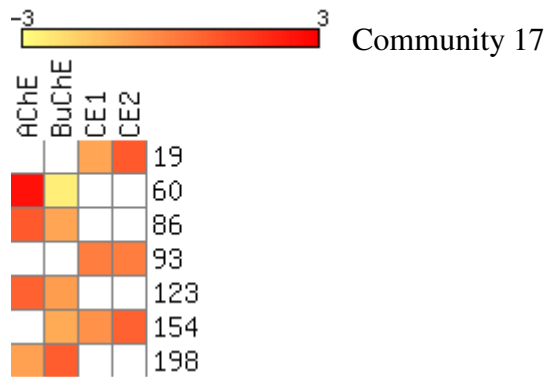
Supplementary Figure S2n



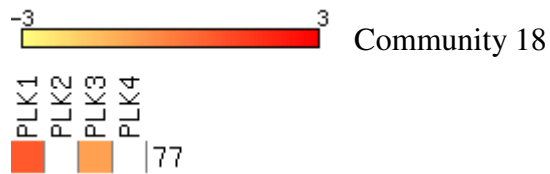
Supplementary Figure S2o



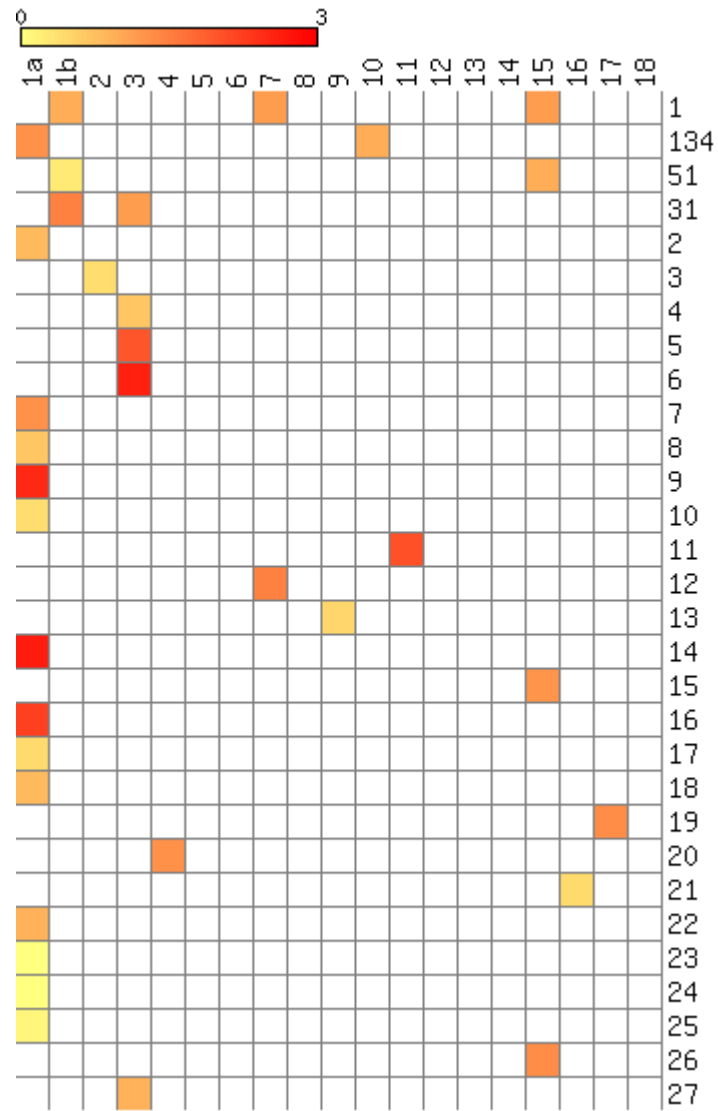
Supplementary Figure S2p



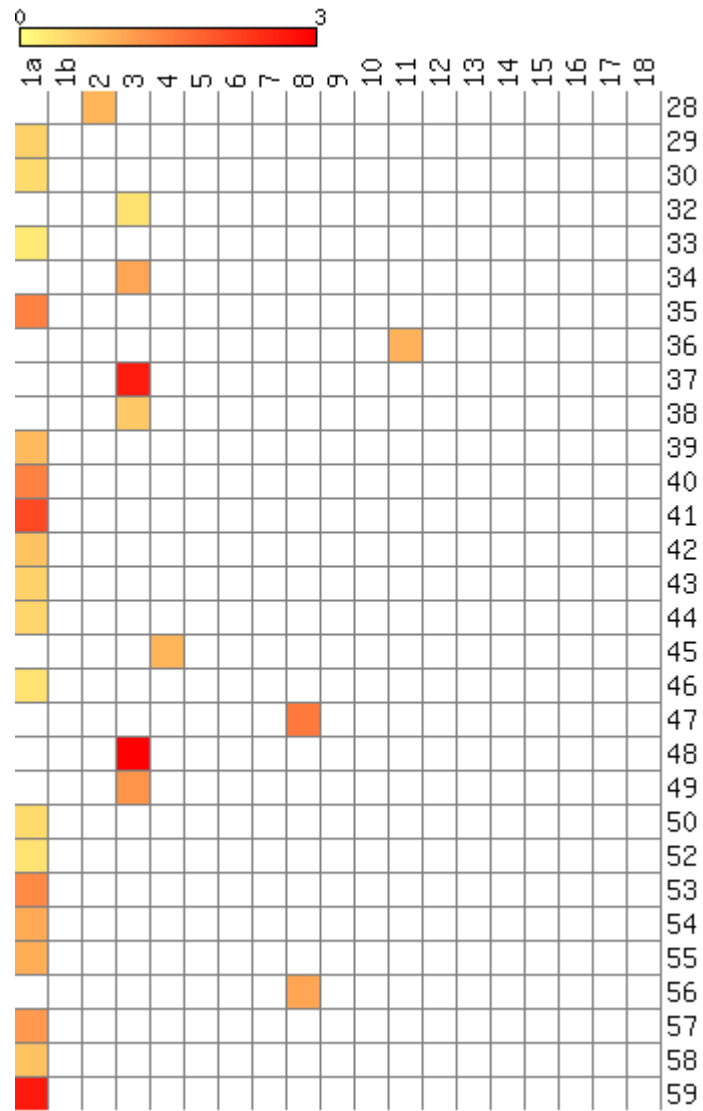
Supplementary Figure S2q



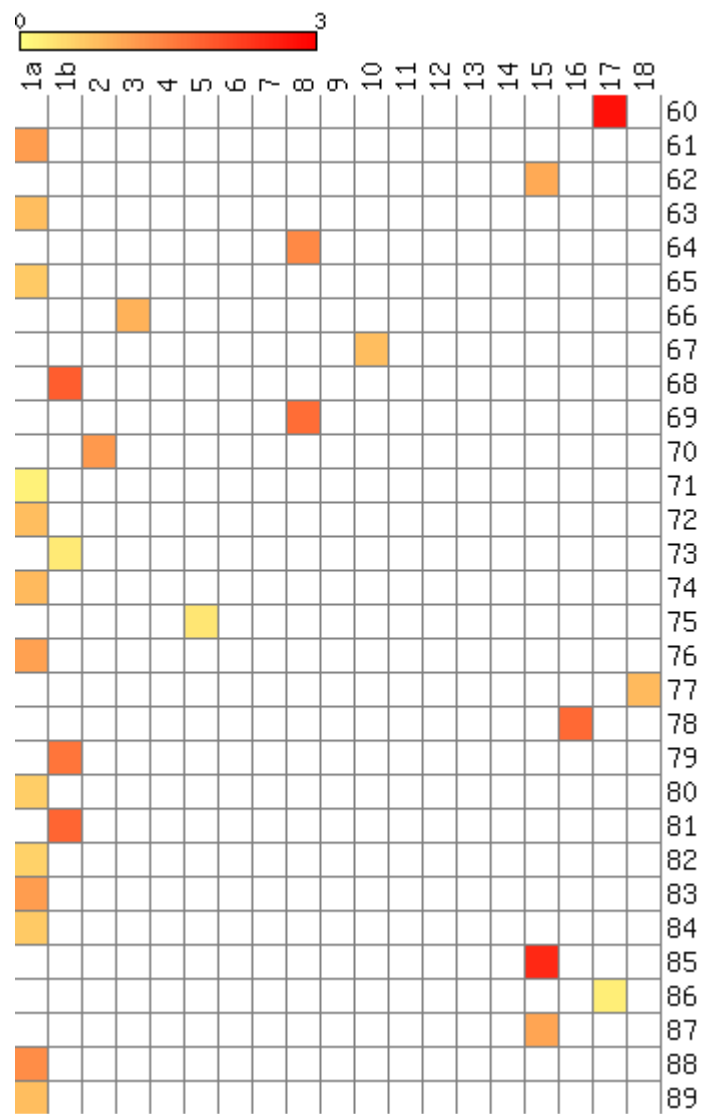
Supplementary Figure S3a



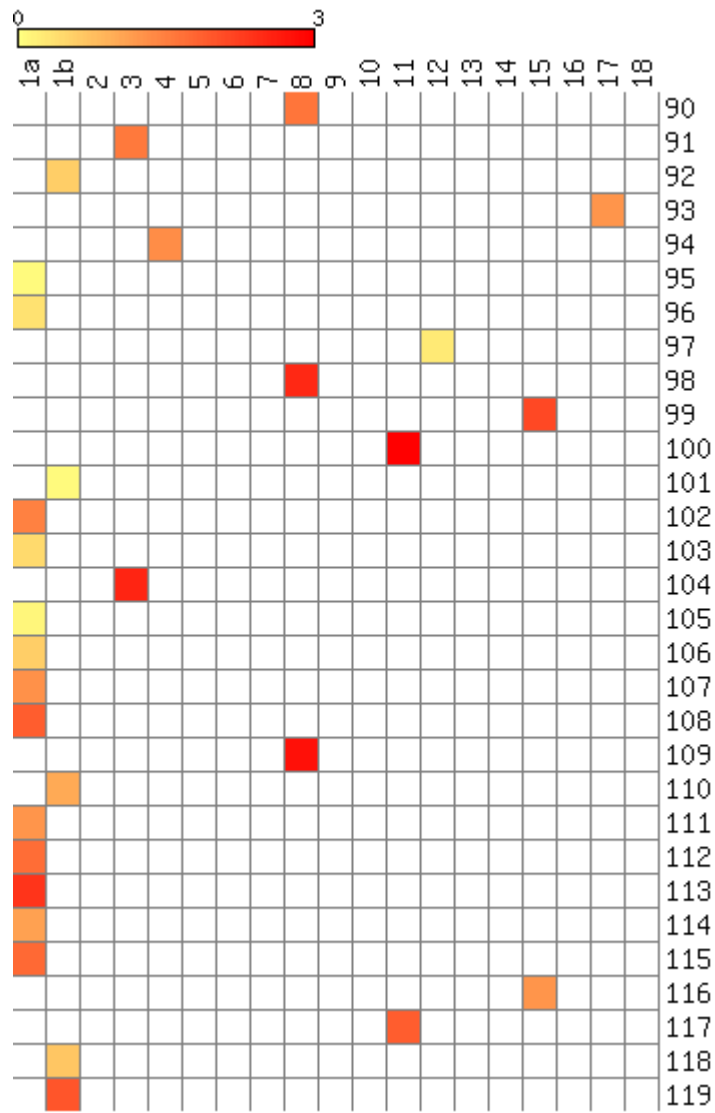
Supplementary Figure S3b



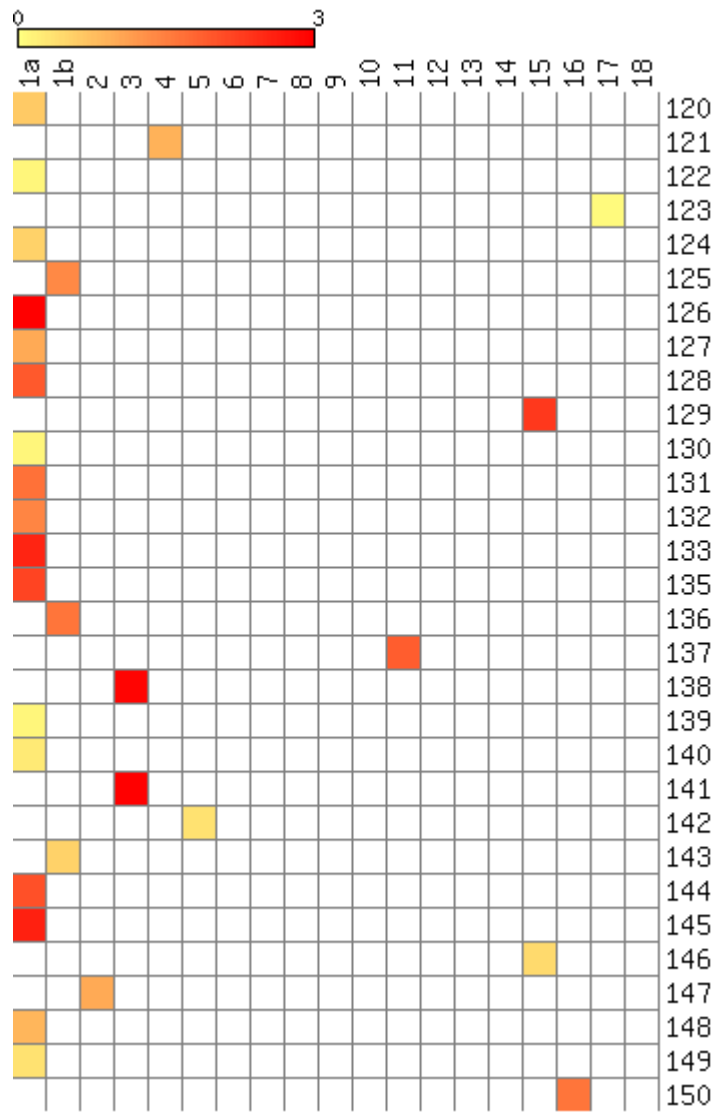
Supplementary Figure S3c



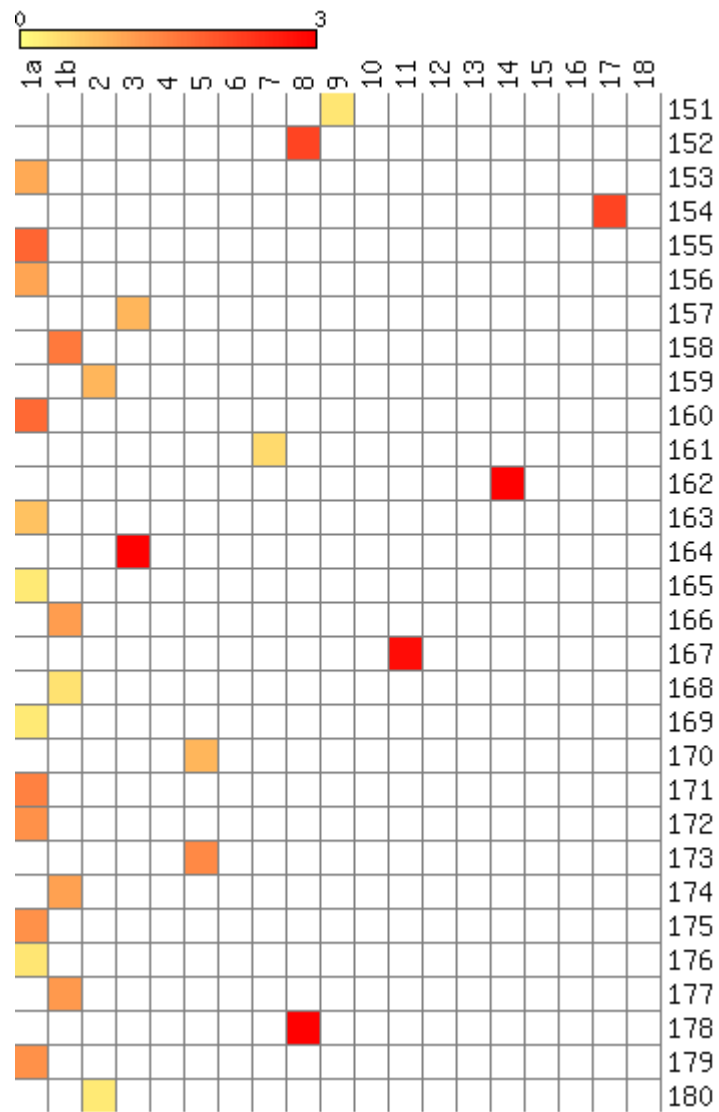
Supplementary Figure S3d



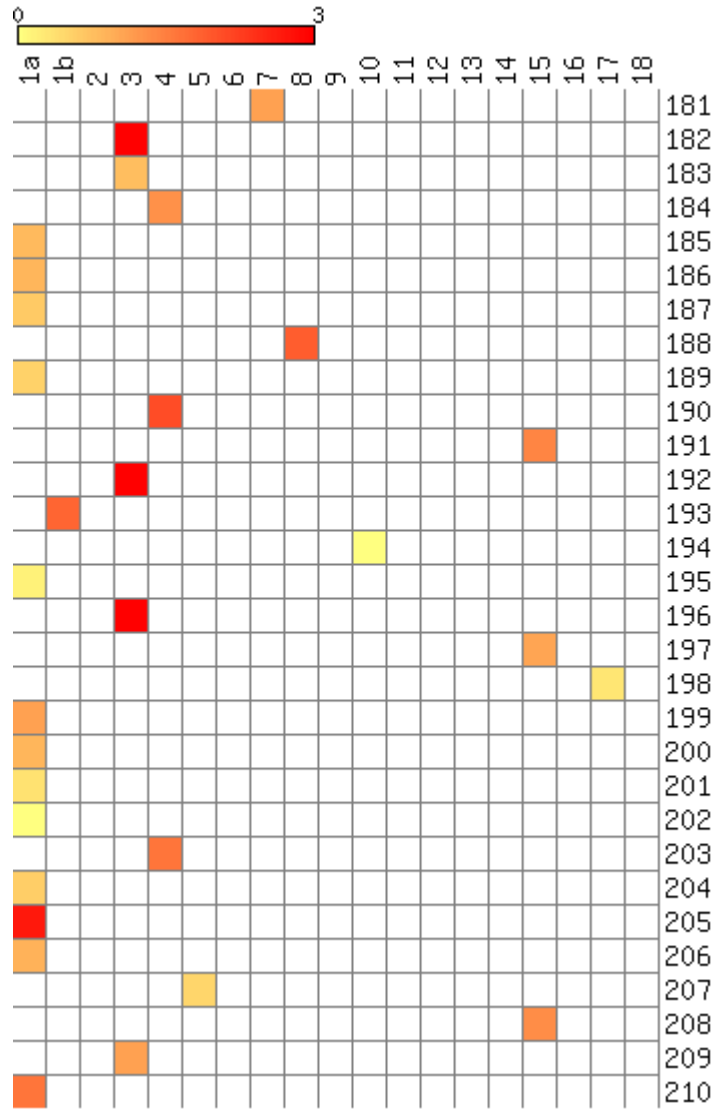
Supplementary Figure S3e



Supplementary Figure S3f



Supplementary Figure S3g



Supplementary Figure S4

