Supplementary Table 1. Complete list of putative binding sites or regions

Among the 220 genomic strings mentioned in the RA literature, 44 are orthologs of others. Once they are removed, a nonredundant set of 176 elements remains, and is shown in the Table below.

Forty of the remaining sequences constitute Retinoic Acid Response Element-like motifs that have been noted but not tested. The were excluded from this paper, and are marked "Motif(s) only". We also eliminated elements found only in fish or birds, but not (so far) in mammals; genes for which we could not confidently find orthologs are marked "Orthologs", and intronic elements, which we did not include, are also marked. Finally, we excluded strings in repeats and a group of strings that are probably spurious or otherwise problematic. For example, one lies near a gap in the current rat assembly; another is in the coding region of an adjacent gene and is probably spurious; and there is a group of "sites" have been said to suppress transcription upon the addition of ligand—a phenomenon that is generally discounted in the field and for which no mechanism has been proposed. Most likely, they represent indirect activity (through one or more other transcription factors).

Gene is the symbol for the gene in nomenclature of the species (first) studied; ***Loc*** is with respect to the most relevant known transcription start site; ***Sp*** shows the species in which the site was first identified; ***Ref PMID*** is the Pub Med Identifier for the original paper(s) most relevant to the site; ***GI*** is the GenBank Gene Identifier; ***Reason Excluded***, where applicable, is self-explanatory.

No.	Gene	Loc	Sp	Ref PMID	GI	Reason Excluded
1	*Abcc2*	-413	Rn	10722729	3021574	NA
2	*ACADM*	-336	Hs	1328196	4557230	NA
3	*ADH1C*	-300	Hs	1996113	178116	Orthologs
4	*Adrb1*	-41	Rn	9448745	220670	NA
5	*Afp*	-6327	Rn	7525384	202783	NA
6	*Afp*	-2405	Rn	8945636	34876954	Rat assembly problem
7	*Afp*	-1961	Rn	7528016	202783	NA
8	*Afp*	-139	Rn	7512261	202783	NA
9	*Akp2*	-5095	Mm	8817450	20786316	Orthologs; Motif(s) only
10	*Aldh1a1*	-91	Hs	10995752	178358	NA
11	*Apoa1*	-210	Hs	8626539	178771	NA
12	*Apoa1*	-132	Hs	0008626539; 0001646397	178771	NA
13	*APOA2*	-733	Hs	8668150	28757	NA
14	*Apoc3*	-732	Hs	9893992	178839	NA
15	*Apoc3*	-672	Hs	9893992	178839	Probably spurious "site"

No.	Gene	Loc	Sp	Ref PMID	GI	Reason Excluded
16	*Apoc3*	-83	Hs	0009893992; 0009691099	178839	NA
17	*APP*	Intron 7	Hs	10727079	2429080	Motif(s) only; Intronic
18	*AQP1*	-2218	Hs	12051745	3135285	NA
19	*ARR3*	-727	Hs	11980849	11414583	Motif(s) only; Probably spurious
20	*ASB2*	-133	Hs	11566180	15777898	Repeat
21	*BGLAP*	-511	Hs	2159384	3252209	Orthologs
22	*Bmp2*	-2373	Mm	9880512	51712651	NA
23	*BTK*	-201	Hs	7927535	17486439	Motif(s) only
24	*C3*	-268	Hs	11368771	29544	Motif(s) only
25	*Ca2*	-1211	Gg	10799323	63115	Bird only
26	*CAMK2A*	5′ UTR	Hs	8795626	51511721	Motif(s) only; Probably spurious
27	*Cd38*	745	Hs	9624127	3273305	Intronic
28	*CDKN1A*	-1212	Hs	8940196	902576	Orthologs; Repeat
29	*Cdx1*	-511	Mm	10938132	402307	NA
30	*CEBPE*	42	Hs	10330422	1399174	Intronic
31	*CETP*	-158	Hs	10329401	1732066	Orthologs; Motif(s) only
32	*Cfh*	-56	Mm	1700780	193805	Orthologs
33	*Chat*	various	Mm	7790895	220364	Motif(s) only
34	*Clu*	-2810	Rn	9547504	207525	Motif(s) only
35	*Cntn1*	various	Mm	9332725	861196	Motif(s) only; Probably spurious
36	*Col1a1*	-1345	Rn	9077477	467809	Down regulation
37	*Col1a1*	-562	Mm	10729205	191856	Down regulation
38	*Col1a2*	-900	Mm	1429872; 0010729205	15291107	Down regulation
39	*Crabp1*	-1008	Mm	9392513	237797	NA
40	*CRABP2*	-5654	Hs	0008071361; 0009856825	497656	Coding region of another gene
41	*Crabp2*	-1090	Mm	1327537	192707	NA
42	*Crabp2*	-586	Hs	9856825	18549451	NA
43	*CRABP2*	-454	Hs	1334086	18549451	Motif(s) only
44	*Cryab*	-132	Mm	9651402	191551	NA
45	*Cryab*	-69	Mm	9651402	191551	NA
46	*Crygf*	-207	Mm	8436299	192784	Orthologs
47	*CSH1*	-969	Hs	7589779	183148	Orthologs; Repeat
48	*CSH1*	-538	Hs	7589779	183148	Orthologs
49	*CYP1A1*	-468	Hs	8024563	30346	NA
50	*CYP24A1*	-291	Hs	9228086	1418240	NA
51	*CYP24A1*	-269	Hs	9228086	1418240	NA
52	*CYP4F2*	-708	Hs	11162441	3347821	Orthologs

No.	Gene	Loc	Sp	Ref PMID	GI	Reason Excluded
53	*Cyp7a1*	-139	Mm	9799805	797278	NA
54	*DIO1*	-105	Hs	9492050	2315182	NA
55	*Drd2*	-66	Rn	9405615	34865068	NA
56	*Ebaf*	-70	Mm	9496783	2347041	NA
57	*EGFR*	67	Hs	2169350	4885198	Intronic
58	*Egr1*	-469	Rn	7877619	26017808	NA
59	*Epo*	3198	Hs	11050012; 1448072	3135305	Intronic
60	*ETS1*	-432	Hs	11327309	31263	NA
61	*Ets1*	41	Mm	10773887	11761707	Intronic
62	*FBP1*	-340	Hs	10731708	1000070	Repeat
63	*Fgf8*	-4542	Hs	12054865	5801866	NA
64	*Foxa1*	-1094	Mm	10388516	22138698	NA
65	*FSCN2*	-24	Hs	10783262	3172526	Motif(s) only
66	*Gh1*	-190	Rn	8384845	56221	NA
67	*Gnrh1*	-1490	Mm	11245923	915407	NA
68	*GPX2*	-1021	Hs	10498757	4504102	Motif(s) only
69	*GPX2*	460	Hs	10498757	4504102	Motif(s) only; Intronic
70	*GSTP1*	1521	Hs	9679546	726097	Intronic
71	*GSTP1*	1552	Hs	9679546	726097	Intronic
72	*H1fo*	-531	Mm	8078070	193849	NA
73	*H2 region*	-210	Mm	8413217	199413	Orthologs; Motif(s) only
74	*H2 region*	562	Mm	8604312	51154	Orthologs; Intronic
75	*Hnf4a*	-295	Mm	11585914	998414	NA
76	*Hoxa1*	4661	Mm	1360810	14626313	Intronic
77	*Hoxa4*	-3054	Hs	8759021	1399048	NA
78	*Hoxa7*	-238	Hs	12068955	1781003	NA
79	*Hoxa7*	-175	Hs	12068955	1781003	NA
80	*Hoxb1*	-1623	Mm	0007916164; 0007831296	11024922	NA
81	*Hoxb1*	3208	Mm	0007914354; 0007831297	11024922	Intronic
82	*Hoxb1*	7045	Mm	8999919	11024922	Intronic
83	*Hoxb4*	5835	Mm	9697850	11024922	Intronic
84	*Hoxb5*	8141	Hs	12514132	29799031	Intronic
85	*Hoxd4*	-2551	Mm	0002159384; 0008093325	1791004	NA
86	*Hoxd4*	3379	Mm	10940626	17530755	Intronic
87	*HSD17B1*	-503	Hs	8614400	806392	Repeat
88	*IBSP*	-1038	Hs	8061918	438617	Motif(s) only; Repeat

No.	Gene	Loc	Sp	Ref PMID	GI	Reason Excluded
89	*Icam1*	-225	Hs	0001983003; 0007647034; 0007737364	32621	NA
90	*Igfbp6*	-27	Hs	11267670	4210703	NA
91	*INS*	-1032	Hs	7639703	186429	In human repetitive polymorphism
92	*ITGAM*	various	Hs	1347945	188606	Motif(s) only
93	*ITGB2*	various	Hs	1346252	180022	Motif(s) only
94	*Itgb3*	-151	Gg	8702813	441308	Bird only
95	*Itgb7*	various	Mm	8318458	1682985	Motif(s) only
96	*KRT10*	-175	Hs	1284070	28316	Down regulation
97	*KRT14*	-91	Hs	1281867	533529	Down regulation
98	*KRT17*	-113	Hs	9326392	20560567	Down regulation
99	*KRT18*	-91	Hs	7667273	7339829	Repeat
100	*KRT5*	165	Hs	9326392	266236	Down regulation; Intronic
101	*KRT6A or B*	various	Hs	9326392	*See Note 1*	Down regulation; Orthologs
102	*Lamb1-1*	-467	Mm	1850696	254625	NA
103	*Laptm5*	-394	Mm	8839844	924940	Motif(s) only; Repeat
104	*LPA*	-1036	Hs	9535807	2323518	Orthologs; Motif(s) only
105	*Ltf*	-349	Mm	7623814	198789	Receptors bind but there is no transactivation
106	*Mdk*	-971	Mm	7982887	530257	NA
107	*MGP*	-138	Hs	9122176	3172535	Down regulation; Motif(s) only
108	*Mixl1*	-534	Mm	12095687	21744720	Motif(s) only
109	*Mixl1*	-416	Mm	12095687	21744720	Motif(s) only
110	*Mmp11*	-2141	Mm	10993903	10280608	Repeat
111	*MMP11*	-2030	Hs	10993903	16168698	Repeat
112	*Mmp11*	-385	Hs	7657606	984746	NA
113	*Mmp11*	-367	Mm	10993903	10280608	NA
114	*MPO*	-448	Hs	8662930	425516	Repeat
115	*Msx1*	-804	Mm	7916326	688052	Motif(s) only
116	*MSX1*	-336	Hs	7866431	20535188	No site found (although there is transactivation)
117	*MUC5B*	various	Hs	11713095	11275567	Motif(s) only
118	*Mycn*	-186	Hs	9570357	35074	NA
119	*Nes*	2378	Hs	10222142	2209203	Intronic
120	*Ngfr*	-860	Rn	1446821	56207	Repeat
121	*Nr2c1*	-136	Mm	10393558	2735500	Sequence problem
122	*OAS3*	various	Hs	11112351	8896067	Orthologs; Motif(s) only
123	*Oxt*	-162	Hs	1657967	20559789	NA
124	*Oxt*	-103	Hs	1657967	20559789	NA

No.	Gene	Loc	Sp	Ref PMID	GI	Reason Excluded
125	*Pck1*	-451	Rn	1848696	4070995	NA
126	*Pck1*	-337	Rn	8626419	4070995	NA
127	*Pcp2*	202	Mm	9224660	1679919	Orthologs; Intronic
128	*PECAM1*	-71	Hs	8955189	1743258	Motif(s) only
129	*Pit1 (Pou1f1)*	-10145	Mm	8504933	20598604	NA
130	*Pit1 (Pou1f1)*	-9999	Mm	8504933	20598604	NA
131	*PLAT*	-7319	Hs	7706255	683752	NA
132	*Pou5f1*	-45	Mm	0008289783; 0008289793	544952	NA
133	*Prkca*	-93	Mm	10486248	51766637	NA
134	*PTAFR*	-66	Hs	8570633	300030	NA
135	*Ptgds*	-586	Rn	9582446	206116	NA
136	*Pthr1*	-2714	Rn	10406468	15721867	NA
137	*Ranbp1/ Htf9*	27	Mm	9417108	6680314	Intronic
138	*Rara*	-55	Mm	1658797	200654	NA
139	*Rarb*	-53	Hs	0002153268; 0002177841; 0002164682	35881	NA
140	*Rarg*	-393	Hs	1320193	297145	NA
141	*RARRES3*	various	Hs	9843971	2281065	Motif(s) only
142	*Rbp1*	-1011	Mm	1648481	50547	NA
143	*Rbp2*	-617	Mm	0001651173; 0008288643	397366	Probably spurious
144	*Rbp2*	-125	Mm	8288643	397366	NA
145	*Rbp2*	-74	Mm	8288643	203586	NA
146	*Rbp4*	-194	Hs	8810324	51467897	NA
147	*Ren1*	-2674	Mm	11058598	1374947	NA
148	*Rrg1*	1	Mm	8754834	1245373	Orthologs; Intronic
149	*Rxrg*	-98	Hs	9006910	9366919	NA
150	*Sag*	-206	Mm	7708064	54037	Site only active in vivo
151	*SERPINB2*	-1641	Hs	*See Note* **2**	189558	Motif(s) only
152	*SERPINC1*	-92	Hs	8761481	511791	Motif(s) only; Probably spurious
153	*Sftpa1*	1491	Rn	8944731	1147797	Intronic
154	*SFTPB*	-458	Hs	10617585	338326	NA
155	*SHH*	-258	Dr	9233805	13992722	Fish only
156	*Slc10a1*	-53	Rn	10722729	1261825	NA
157	*Slc9a2*	various	Rn	9804979	3916117	Motif(s) only
158	*SMPD1*	-513	Hs	11788605	19703148	Motif(s) only

No.	Gene	Loc	Sp	Ref PMID	GI	Reason Excluded
159	*SPN*	-828	Hs	9174604	180125	Motif(s) only; Repeat
160	*SPN*	-178	Hs	9174604	180125	Motif(s) only
161	*Stat1*	-467	Mm	9092506	1916072	NA
162	*Tcf1*	-68	Hs	11027556	1854926	NA
163	*Tgm2*	-1745	Mm	8626785	1762323	NA
164	*THBD*	-1531	Hs	0008207015; 11036068	19909347	NA
165	*Thbd*	-941	Mm	8918245	54780	NA
166	*Tlx2*	-1163	Mm	10446220	5748488	NA
167	*TRIM31*	-400	Hs	12170760	22056653	Orthologs; Motif(s) only
168	*Tshb*	-204	Rn	9296372	207531	Down regulation
169	*TXN*	-626	Hs	11970916	3599465	Motif(s) only
170	*Tyr*	various	Mm	7620342	220624	Motif(s) only
171	*Ucp1*	-2480	Rn	8754778	57444	Motif(s) only
172	*Ucp3*	-71	Hs	11024001	9937094	NA
173	*VDR*	Intronic	Hs	0009212063; 10919269	22062255	Intronic
174	*VIPR1*	-747	Hs	7708752	508250	Motif(s) only
175	*Wnt1*	Region	Mm	7925022	2761013	Probably spurious
176	*ZNF42 (Zfp98)*	-830	Hs	8845378	1184871	Motif(s) only

Notes:

1. It is not possible to tell what gene the paper is about—either from the paper itself or from the sequence fragments it presents. The region BLASTS identially to *Krt6A*, *-B*, and *-C*.
2. There is no PMID for this paper (the journal was not indexed for Medline at the time the paper was written, and seems now to be defunct). See *Fibrinolysis* **8(2)**: 113-119 (1994).