

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第3670030号  
(P3670030)

(45) 発行日 平成17年7月13日(2005.7.13)

(24) 登録日 平成17年4月22日(2005.4.22)

(51) Int. Cl.<sup>7</sup>

F I

C O 7 K 14/47

C O 7 K 14/47

C O 7 K 16/18

C O 7 K 16/18

C 1 2 N 1/21

C 1 2 N 1/21

C 1 2 N 15/09

C 1 2 P 21/02

C

C 1 2 P 21/02

C 1 2 P 21/08

請求項の数 11 (全 72 頁) 最終頁に続く

(21) 出願番号 特願平6-84470  
 (22) 出願日 平成6年4月22日(1994.4.22)  
 (65) 公開番号 特開平7-330799  
 (43) 公開日 平成7年12月19日(1995.12.19)  
 審査請求日 平成13年4月4日(2001.4.4)  
 (31) 優先権主張番号 特願平5-136602  
 (32) 優先日 平成5年5月14日(1993.5.14)  
 (33) 優先権主張国 日本国(JP)  
 (31) 優先権主張番号 特願平5-257455  
 (32) 優先日 平成5年9月22日(1993.9.22)  
 (33) 優先権主張国 日本国(JP)  
 (31) 優先権主張番号 特願平6-49904  
 (32) 優先日 平成6年2月23日(1994.2.23)  
 (33) 優先権主張国 日本国(JP)

(73) 特許権者 000173588  
 財団法人癌研究会  
 東京都豊島区上池袋1丁目37番1号  
 (73) 特許権者 000000217  
 エーザイ株式会社  
 東京都文京区小石川4丁目6番10号  
 (74) 代理人 100087642  
 弁理士 古谷 聡  
 (74) 代理人 100076680  
 弁理士 溝部 孝彦  
 (74) 代理人 100091845  
 弁理士 持田 信二  
 (72) 発明者 中村 祐輔  
 東京都豊島区北大塚3-17-5

最終頁に続く

(54) 【発明の名称】 MDC蛋白質およびそれをコードするDNA

(57) 【特許請求の範囲】

【請求項1】

配列番号2に示される蛋白質の全部を含むものからなるMDC蛋白質。

【請求項2】

配列番号3に示される蛋白質の全部を含むものからなるMDC蛋白質。

【請求項3】

配列番号6に示されるDNAの全部を含むものからなるDNA。

【請求項4】

配列番号7に示されるDNAの全部を含むものからなるDNA。

【請求項5】

配列番号9に示される、エクソン/イントロンを含むDNAの全部を含むものからなるDNA。

【請求項6】

請求項3、4または5のいずれかに記載のDNAを含むプラスミド。

【請求項7】

請求項6記載のプラスミドを保持する形質転換体。

【請求項8】

請求項7記載の形質転換体を培養し、発現産物を回収することを含む請求項1または2のいずれかに記載の蛋白質の製造方法。

【請求項9】

10

20

請求項 1 または 2 のいずれかに記載の蛋白質と結合する抗体。

【請求項 1 0】

請求項 3、4 または 5 のいずれかに記載の D N A 配列を含むプライマーまたはプローブ。

【請求項 1 1】

請求項 1 0 記載のプライマ - またはプロ - ブを被検 D N A とハイブリダイズさせることを含む遺伝子解析方法。

【発明の詳細な説明】

【0 0 0 1】

【産業上の利用分野】

本発明は M D C 蛋白質、それをコードする遺伝子 D N A および該 D N A を用いる遺伝子解析法に関し、医療、診断等の分野で利用される。

【0 0 0 2】

【従来の技術】

癌の発症には細胞の蛋白質の変異が重要な役割を演ずるという考え方は古くから知られている。近年の遺伝子工学の発達は特定蛋白質をコードする遺伝子 D N A の増幅や癌細胞における遺伝子変異の解析を可能にし、癌研究の分野においても飛躍的な発展をもたらした。

これまでに細胞の癌化、癌細胞の異常増殖に関与すると考えられている蛋白質をコードする遺伝子（癌遺伝子）の解析、同定が進み、その数は数十に及んでいる。一方これと反対に作用するもの（癌抑制遺伝子）がここ数年脚光を浴びており、これまでに癌抑制遺伝子として、網膜芽細胞腫の R b 遺伝子（Friend, S.H. et al., Proc. Natl. Acad. Sci. USA. 84, 9095, 1987）、大腸癌の p 5 3 遺伝子（Lane, D.P. et al., Nature, 278, 261, 1979）および A P C 遺伝子（Kenneth, W.K., et al, Science 253, 661, 1991）、W i l m s 腫瘍の W T 1 遺伝子（Call, K.M. et al., Cell, 60, 509, 1990）などが発見されており、p 5 3 遺伝子の場合には、変異遺伝子を家系を通じて伝えている例がある（" Li-Fraumeni 症候群 "（Makin, D. et al, Science 250, 1233, 1990; Srivastava, S. et al, Nature 348, 747, 1990））ことが知られている。また、癌の発生、進展悪性化、転移などには一つの遺伝子の異常だけでなく複数の遺伝子の異常が関与していることが次第に明らかになりつつあり、さらに多くの未同定の癌遺伝子、癌抑制遺伝子が存在するものと考えられている。それらの発見、解明は研究、臨床の専門家はもとより全世界の人々から期待されているところである。

【0 0 0 3】

乳癌は遺伝性（家族性）乳癌と非遺伝性（散发性）乳癌に分けられ、遺伝性乳癌は発症年齢の区分から早発型と晩発型に分けられる。このうち少なくとも遺伝性で早発型の乳癌においては、家系解析によって第 1 7 番染色体のごく一部が欠けている頻度が極めて高いことが明らかとなっている（Hall, J.M. et al., Science, 250, 1684-1689, 1990）。また、これと同じ部位は遺伝性卵巣癌でも高頻度に欠失していることが示されている（Narod, S.A. et al., Lancet, 338, 82-83, 1991）。

従って、この欠失部位に癌抑制遺伝子が存在し、その欠損や変異によって引き起こされる蛋白質の欠損や異常が乳癌や卵巣癌の原因の一つであると考えられている。

また一般の（散发性）乳癌の発生においても、この部位の遺伝子の後天的な変異や欠損による蛋白質の異常や欠損が起こり、その細胞が癌となるケースとして関与しているものと考えられている（Sato et al., Cancer Res., 51, 5794-5799, 1991）。このためこの部位に存在する原因遺伝子を単離し、蛋白質を同定することは、今や世界中の医師や研究者のみならず一般の人々、特に乳癌患者の多い欧米では女性にとって切実な問題として期待されている。

【0 0 0 4】

【発明が解決しようとする課題】

本発明は、乳癌および卵巣癌に関わる新規な蛋白質と、それをコードする遺伝子、さらに

10

20

30

40

50

それらを用いた癌の検査方法、診断方法等を提供することにある。

【 0 0 0 5 】

【課題を解決するための手段】

本発明者らは、ヒト第 17 番染色体の DNA 断片を導入したコスミド・クローンを多数作成した。さらにその多数のコスミド・クローンについて、その DNA 断片をプローブとした蛍光 in situ ハイブリダイゼーション法 (FISH 法; Inazawa et al., Genomics, 10, 1075-1078, 1991) によって染色体上での存在位置を決定した。これら染色体上の位置を決定したコスミド・クローン (コスミド・マーカー) によって第 17 番染色体についての非常に詳細な物理的染色体地図を作成した。プローブとした各コスミドのクローン名と、決定された染色体上の位置についての一覧を表 1 ~ 3 および図 1 に示した。図 1 ではクローン名はクローン番号でのみ示してある。

10

【 0 0 0 6 】

【表 1】

No.	プローブ名	ローカス シンボル	染色体上の位置	No.	プローブ名	ローカス シンボル	染色体上の位置
1	cCI17-1		17q21.1	67	cCI17-501		17q21.3
2	cCI17-7		17q22	68	cCI17-502		17p11.2
3	cCI17-11		17p11.2	69	cCI17-504		17q24
4	cCI17-24		17q21.1-q21.2	70	cCI17-505	D17S544	17p12-p11.1
5	cCI17-25		17q12	71	cCI17-506	D17S545	17q21
6	cCI17-28		17q21.3	72	cCI17-507		17q21.3
7	cCI17-32		17q11.2	73	cCI17-508	D17S546	17q25.1-q25.2
8	cCI17-35		17q21.3	74	cCI17-509		17q24
9	cCI17-44		17q23.1	75	cCI17-510		17q23
10	cCI17-50		17q23.1	76	cCI17-511		17q25.1-q25.2
11	cCI17-57		17q21	77	cCI17-513	D17S548	17q11.2
12	cCI17-63		17q21.3	78	cCI17-514		17q25.1
13	cCI17-90		17q12-q21.1	79	cCI17-515		17q22
14	cCI17-95		17q23.1	80	cCI17-516	D17S550	17q25.1
15	cCI17-96		17q21.3	81	cCI17-517		17q21.3
16	cCI17-97		17q21.31	82	cCI17-518		17q25.3
17	cCI17-315	D17S521	17q25.1-q25.2	83	cCI17-519	D17S551	17q25.2-q25.3
18	cCI17-316		17q12-q21.1	84	cCI17-520		17q11.2
19	cCI17-317		17q11.2	85	cCI17-521		17q25.1-q25.2
20	cCI17-321		Centromere	86	cCI17-523		17q22
21	cCI17-403		17q21.2-q21.3	87	cCI17-524		17q21.1-q21.2
22	cCI17-412		17q11.2	88	cCI17-525		17p13.1
23	cCI17-415		17q21.1-q21.2	89	cCI17-526		17q11.2-q12
24	cCI17-422		17q22	90	cCI17-527		17q21.3
25	cCI17-425		17q11.2	91	cCI17-528		17q22
26	cCI17-428		17q23.1	92	cCI17-529	D17S552	17q25.1-q25.2
27	cCI17-451		17q21.1-q21.2	93	cCI17-530		17q23
28	cCI17-452	D17S524	17q25	94	cCI17-532		17p11.2
29	cCI17-453	D17S525	17p13	95	cCI17-533		17q21.3
30	cCI17-454	D17S526	17q23	96	cCI17-535		17q12-q21.1
31	cCI17-456	D17S527	17q23.1-q23.2	97	cCI17-536		17p11.2
32	cCI17-457		17q11.2	98	cCI17-539		17q21.3
33	cCI17-458	D17S528	17q21.1-q21.2	99	cCI17-540		17q25
34	cCI17-460	D17S529	17q11.2	100	cCI17-541		17q21.3
35	cCI17-462		17q23	101	cCI17-542		17q21.3
36	cCI17-463		17q21	102	cCI17-543		17q11.2-q12
37	cCI17-464		17q24.3-q25.1	103	cCI17-544		17q25.1
38	cCI17-465	D17S531	17q25.1-q25.2	104	cCI17-545		17q25.1
39	cCI17-466		17q25.1-q25.2	105	cCI17-546		17q24.3-q25.1
40	cCI17-467		17q25	106	cCI17-547		17q21.3
41	cCI17-468	D17S532	17q11.2	107	cCI17-548		17q23
42	cCI17-469	D17S533	17q25.2-q25.3	108	cCI17-549		17q25.3
43	cCI17-471		17p13.3-p13.2	109	cCI17-550		17q23
44	cCI17-473	D17S534	17q11.2	110	cCI17-551		17q25.1-q25.2
45	cCI17-475	D17S535	17q11.2	111	cCI17-552		17q12
46	cCI17-477		17q21.3	112	cCI17-553		17q23
47	cCI17-479		17q21.3	113	cCI17-554		17q25.1
48	cCI17-480		17q25.1-q25.2	114	cCI17-557		17q25.1-q25.2
49	cCI17-482	D17S536	17q11.2	115	cCI17-559		17q24.3-q25.1
50	cCI17-483		17p13	116	cCI17-560		17q25.1-q25.2
51	cCI17-484	D17S537	17p13.1	117	cCI17-561		17q25
52	cCI17-485		17q12	118	cCI17-562		17q11.2
53	cCI17-486		17q25.1	119	cCI17-563		17q25.1
54	cCI17-487	D17S538	17q25.1	120	cCI17-564		17q25.1
55	cCI17-488	D17S539	17p13.2-p13.1	121	cCI17-565		17q23
56	cCI17-489	D17S540	17q23	122	cCI17-567		17q21.3
57	cCI17-490		17q11.2	123	cCI17-568		17q25
58	cCI17-491		17p13.1	124	cCI17-569		17q12
59	cCI17-492	D17S542	17q11.2	125	cCI17-570		17q11.1
60	cCI17-493		17q25.1-q25.2	126	cCI17-571		17p13
61	cCI17-494		17q22	127	cCI17-572		17q25.1
62	cCI17-495		17q25	128	cCI17-573		17q25.3
63	cCI17-497		17q11.2	129	cCI17-574		17q12-q21.2
64	cCI17-498		17p11.2	130	cCI17-576		17q21.1-q21.2
65	cCI17-499		17q21.1-q21.2	131	cCI17-577		17q25.1
66	cCI17-500		17p.12	132	cCI17-578		17q11.2-q12

【 0 0 0 7 】

【 表 2 】

No.	プローブ名	ローカス シンボル	染色体上の位置	No.	プローブ名	ローカス シンボル	染色体上の位置
133	cCI17-579		17q25.1	198	cCI17-652		17q22
134	cCI17-581		17q11.2	199	cCI17-653		17q22
135	cCI17-582		17q21.3	200	cCI17-654		17p13
136	cCI17-583		17q12-q21.1	201	cCI17-655		17q23
137	cCI17-584		17q21.3	202	cCI17-656		17q25.1
138	cCI17-586		17p13	203	cCI17-657		17p13
139	cCI17-587		17p13	204	cCI17-658		17q21.3
140	cCI17-588		17p13.1	205	cCI17-659		17q25.1
141	cCI17-590		17q12	206	cCI17-660		17q25.1
142	cCI17-591		17q24	207	cCI17-662		17p12
143	cCI17-592		17q21.3	208	cCI17-663		17q25.1
144	cCI17-593		17q25.1-q25.2	209	cCI17-664		17q25.1-q25.2
145	cCI17-594		17q25.2-q25.3	210	cCI17-665		17q23
146	cCI17-595		17q25.1-q25.2	211	cCI17-666		17q23.1
147	cCI17-596		17q11.2	212	cCI17-667		17q24
148	cCI17-597		17q25.3	213	cCI17-668		17q22
149	cCI17-598		17q12-q21.1	214	cCI17-669		17p13
150	cCI17-599		17q12	215	cCI17-670		17q21.3
151	cCI17-600		17q23	216	cCI17-671		17q11.2
152	cCI17-601		17q21.1-q21.2	217	cCI17-672		17q25.1-q25.2
153	cCI17-602		17q11.2-q12	218	cCI17-673		17q12-q21.1
154	cCI17-603		17p11.2	219	cCI17-674		17q21.3
155	cCI17-604		17q23	220	cCI17-675		17q21.3
156	cCI17-605		17q21.1-q21.2	221	cCI17-676		17q23
157	cCI17-606		17p13	222	cCI17-677		17q12-q21.1
158	cCI17-607		17q25	223	cCI17-678		17q23
159	cCI17-608		17p11.2	224	cCI17-679		17q23.1
160	cCI17-609		17q21.3	225	cCI17-680		17p13
161	cCI17-610		17q12-q21.1	226	cCI17-681		17p11.1-p11.2
162	cCI17-611		17q22	227	cCI17-683		17q11.2
163	cCI17-612		17q21.3	228	cCI17-684		17q25.1-q25.2
164	cCI17-613		17q25.1	229	cCI17-685		17p13
165	cCI17-614		17q21.3	230	cCI17-687		17q12
166	cCI17-615		17q21.1	231	cCI17-688		17p11.2
167	cCI17-616		17q25.1	232	cCI17-690		17q11.2-q12
168	cCI17-617		17q21.3	233	cCI17-691		17q25.1
169	cCI17-618		17q23.1	234	cCI17-692		17q23
170	cCI17-619		17q21.3	235	cCI17-693		17p11.2
171	cCI17-621		17q25.1-q25.2	236	cCI17-694		17p11.2
172	cCI17-622		17q12	237	cCI17-695		17p11.2
173	cCI17-623		17q25.1-q25.2	238	cCI17-696		17q23.3
174	cCI17-624		17p13	239	cCI17-697		17q25
175	cCI17-625		17q23	240	cCI17-698		17q11.2
176	cCI17-626		17q23	241	cCI17-699		17q23
177	cCI17-627		17p13	242	cCI17-700		17q23
178	cCI17-628		17q23	243	cCI17-701		17q21.3
179	cCI17-630		17q11.2	244	cCI17-702		17q25.2-q25.3
180	cCI17-631		17p11.2	245	cCI17-703		17p13
181	cCI17-662		17q22	246	cCI17-704		17q23
182	cCI17-633		17q12	247	cCI17-705	D17S554	17p11.2
183	cCI17-634		17q21.3	248	cCI17-706	D17S555	17q12
184	cCI17-636		17p13	249	cCI17-707	D17S556	17q25.1-q25.2
185	cCI17-637		17q12	250	cCI17-708		17p13
186	cCI17-638		17p11.2	251	cCI17-709		17p12
187	cCI17-639		17q12	252	cCI17-710	D17S557	17q25.3
188	cCI17-640		17q11.2	253	cCI17-711		17q32.1
189	cCI17-641		17q25.1	254	cCI17-712	D17S558	17p11.2
190	cCI17-642		17q12-q21.1	255	cCI17-713	D17S559	17p13
191	cCI17-643		17q21.3	256	cCI17-714	D17S560	17q25.3
192	cCI17-644		17q23	257	cCI17-715		17q21.3
193	cCI17-645		17p13	258	cCI17-716	D17S561	17p13
194	cCI17-646		17p13	259	cCI17-717		17p13
195	cCI17-647		17q25.1-q25.2	260	cCI17-719		17q25
196	cCI17-650		17q12	261	cCI17-721		17q23
197	cCI17-651		17q25.1	262	cCI17-722	D17S563	17q25.2-q25.3

10

20

30

40

【 0 0 0 8 】

【 表 3 】

No.	プローブ名	ローカス シンボル	染色体上の位置	No.	プローブ名	ローカス シンボル	染色体上の位置
263	cCI17-723		17p13	304	cCI17-834		17q11.2-q12
264	cCI17-724	D17S564	17p11.2	305	cCI17-835		17q21.3
265	cCI17-726		17q25	306	cCI17-841		17p12
266	cCI17-727	D17S566	17p13	307	cCI17-1005		17q21.3
267	cCI17-728	D17S567	17p12	308	cCI17-1008		17q21.3
268	cCI17-729	D17S568	17q11.2	309	cCI17-1016		17q23.1
269	cCI17-730		17q21.3	310	cCI17-1018		17q21.2-21.3
270	cCI17-732	D17S570	17p13.2	311	cCI17-1019		17q23.1
271	cCI17-733		17q25.1	312	cCI17-1024		17q12
272	cCI17-735	D17S572	17q25.3	313	cCI17-1029		17q11.2
273	cCI17-736	D17S573	17q21.3	314	cCI17-1030		17q22
274	cCI17-737	D17S557	17q25.2-q25.3	315	cCI17-1031		17q11.2
275	cCI17-739	D17S575	17q25.1	316	cCI17-1032		17q23.1-23.2
276	cCI17-741		17q25.3	317	cCI17-1049		17q21.3
277	cCI17-742		17q25	318	cCI17-1055		17q21.3
278	cCI17-743		17q23.3	319	cCI17-1059		17q21.1-q21.2
279	cCI17-744		17q23	320	cCI17-1063		17q12
280	cCI17-745	D17S577	17p13	321	cCI17-1073		17q11.2
281	cCI17-801		17q11.2-q12	322	cCI17-1079		17q12
282	cCI17-802		17p11.2	323	cCI17-1082		17q22
283	cCI17-808		17q25.1-q25.2	324	cCI17-1094		17q21.1
284	cCI17-809		17q23	325	cCI17-1101		17q12
285	cCI17-810		17p13.2-p13.1	326	cCI17-1103		17q11.2
286	cCI17-812		17q25.1	327	cCI17-1106		17q11.2
287	cCI17-813		17q23	328	cCI17-1702		17q21.1-q21.2
288	cCI17-814		17p11.2	329	cCI17-1705		17q21.2-q21.3
289	cCI17-815		17q24	330	cCI17-1706		17q21.1-q21.2
290	cCI17-816		17q23	331	cCI17-1707		17q21.2-q21.3
291	cCI17-817		17q23	332	cCI17-1709		17q12
292	cCI17-818		17p11.2	333	cCI17-1710		17q21.3
293	cCI17-820		17q12	334	cCI17-1711		17q12
294	cCI17-821		17p13	335	cCI17-1715		17q12
295	cCI17-822		17q11.1	336	cCI17-1717		17q21.3
296	cCI17-823		17q12	337	cCI17-1719		17q11.2
297	cCI17-825		17p11.2	338	cCI17-1720		17q24.3-q25.1
298	cCI17-826		17q11.2	339	cCI17-1722		17q23
299	cCI17-827		17p11.2	340	cCI17-1723		17q21.3
300	cCI17-828		17p11.2	341	cCI17-1724		17q11.2
301	cCI17-831		17q25.1-q25.2	342	cCI17-1725		17q21.1
302	cCI17-832		17p11.2	343	pCMM86		17q23
303	cCI17-833		17q23				

10

20

## 【 0 0 0 9 】

これらのマーカーの中から、個体によって制限酵素断片の長さが異なる性質（RFLP：Restriction Fragment Length Polymorphism：制限酵素断片長多型）を持つもの、すなわちRFLPマーカーを選び出した。表4～6に、選ばれたマーカークローンと、使用する制限酵素名、およびそれによって検出される複数種の断片の具体的断片長を記載した。

## 【 0 0 1 0 】

## 【表4】

30

No.	プローブ名	ローカス シンボル	制限酵素	対立遺伝子 (頻度)	染色体上の位置	
2	cCI17-7	D17S860	PvuII	3.0 kb(0.33)		
16	cCI17-97	D17S861	PstI	1.8+1.2 kb(0.67)	17q21.3	
17	cCI17-315	D17S521	TaqI	8.2 kb(0.92)		
18	cCI17-316	D17S862	MspI	4.7+3.5 kb(0.08)	17q25.1-q25.2	
19	cCI17-317	D17S522	TaqI	2.0 kb(0.67)		
29	cCI17-453	D17S525	BglII	1.8 kb(0.33)	17q12-q21.1	
42	cCI17-469	D17S533	MspI	3.1 kb(0.33)		
54	cCI17-487	D17S538	EcoRI	2.7 kb(0.67)	17q11.2	
56	cCI17-489	D17S540	MspI	TaqI 2.6-3.9 kb 4 alleles VNTR,60% heterozygosity also polymorphic with MspI,PstI,PvuII	17p13	10
58	cCI17-491	D17S863	TaqI	BglII 5.8-7.5 kb 4 alleles VNTR,50% heterozygosity also polymorphic with EcoRI,TaqI,PstI,PvuII,MapI	17q25.2-q25.3	
59	cCI17-492	D17S542	BglII	MspI 2.0-2.6 kb 5 alleles VNTR,83% heterozygosity also polymorphic with EcoRI,TaqI,PvuII	17q25.1	
61	cCI17-494	D17S865	EcoRI	5.8 kb(0.75)	17q23	
70	cCI17-505	D17S544	MspI	3.3 kb(0.25)		
71	cCI17-506	D17S545	MspI	2.1 kb(0.50)		
73	cCI17-508	D17S546	MspI	TaqI 1.5 kb(0.50)	17p13.1	
80	cCI17-516	D17S550	TaqI	PvuII 1.35 kb(0.50)		
88	cCI17-525	D17S866	MspI	1.2 kb(0.50)	17q11.2	20
118	cCI17-562	D17S5867	TaqI	0.7 kb(0.50)		
137	cCI17-584	D17S868	MspI	3.6 kb(0.75)	17p12-p11.1	
166	cCI17-615	D17S869	PstI	3.3 kb(0.25)		
243	cCI17-701	D17S870	TaqI	2.1 kb(0.40)	17q21	
244	cCI17-702	D17S871	MspI	1.4 kb(0.60)	17q25.1-q25.2	
			TaqI	10.3 kb(0.92)	17q25.1	
			MspI	7.8 kb(0.008)		
			TaqI	3.1 kb(0.58)		
			PvuII	3.0 kb(0.42)		
			MspI	4.1 kb(0.67)		
			TaqI	2.7+1.4 kb(0.33)		
			PstI	3.0 kb(0.33)		
			MspI	2.6 kb(0.67)		
			TaqI	4.6 kb(0.50)		
			PvuII	4.0 kb(0.50)		
			MspI	4.1 kb(0.25)		
			TaqI	2.4+1.7 kb(0.75)		
			PstI	3.4 kb(0.83)		
			MspI	2.2 kb(0.17)		
			TaqI	2.7 kb(0.42)		
			PvuII	2.3 kb(0.58)		
			MspI	3.5 kb(0.42)		
			TaqI	3.2 kb(0.58)		
			PstI	7.1 kb(0.92)		
			MspI	6.6 kb(0.08)		
			TaqI	3.8 kb(0.25)		
			PstI	3.6 kb(0.75)		
			MspI	5.2 kb(0.42)		
			TaqI	4.7 kb(0.58)		
			RsaI	4.1 kb(0.83)		
			BglII	3.4 kb(0.17)		
			PvuII	5.2 kb(0.83)		
			MspI	4.1 kb(0.17)		
			TaqI	6.6 kb(0.83)		
			PstI	5.6 kb(0.17)		
			MspI	2.9 kb(0.83)		
			TaqI	2.2 kb(0.17)		

【 0 0 1 1 】

【 表 5 】

No.	プローブ名	ローカス シンボル	制限酵素	対立遺伝子 (頻度)	染色体上の位置	
245	cCI17-703	D17S877	TaqI 2.6-3.8 kb 4 alleles VNTR, 50% heterozygosity also polymorphic with MspI, RsaI, PstI, PvuII		17p13	
247	cCI17-705	D17S554	PstI	4.3 kb(0.50) 2.3+2.0 kb(0.50)	17p11.2	
250	cCI17-708	D17S878	PvuII 2.6-9.0 kb 10 alleles VNTR, 87% heterozygosity also polymorphic with MspI, TaqI, BglII, PstI, EcoRI		17p13	
252	cCI17-710	D17S557	MspI 2.0-2.6 kb 5 alleles VNTR, 100% heterozygosity also polymorphic with RsaI, TaqI, PstI, PvuII, EcoRI		17q25.3	
254	cCI17-712	D17S558	MspI	3.1 kb(0.58) 2.9 kb(0.42) 6.6 kb(0.67) 4.3+2.3 kb(0.33) 7.1 kb(0.50) 3.9+3.2 kb(0.50)	17p11.2	10
255	cCI17-713	D17S559	MspI 2.2-2.8 kb 3 alleles VNTR, 50% heterozygosity also polymorphic with PstI		17p13	
256	cCI17-714	D17S560	RsaI	4.5 kb(0.58) 4.3 kb(0.42) 3.8 kb(0.75) 2.8 kb(0.25) 3.8 kb(0.58) 3.5 kb(0.42) 2.6 kb(0.58) 2.4 kb(0.42) 1.5 kb(0.58) 1.4 kb(0.42)	17q25.3	20
257	cCI17-715	D17S872	PstI	3.3 kb(0.17) 3.0 kb(0.83) 3.6 kb(0.87) 3.3 kb(0.13)	17q21.3	
258	cCI17-716	D17S561	TaqI	2.4 kb(0.87) 1.3+1.1 kb(0.13)	17p13	
261	cCI17-721	D17S864	RsaI	2.9 kb(0.25) 1.6 kb(0.75) 4.4 kb(0.83) 3.9 kb(0.17)	17q22-q23	
262	cCI17-722	D17S563	MspI	4.1 kb(0.83) 3.4 kb(0.17) 5.2 kb(0.83) 4.1 kb(0.17) 6.6 kb(0.83) 5.6 kb(0.17) 2.9 kb(0.83) 2.2 kb(0.17) 13.0 kb(0.75) 12.5 kb(0.25)	17q25.2-q25.3	30
263	cCI17-723	D17S873	MspI	3.0 kb(0.33) 1.7 kb(0.67) 0.8 kb(0.70) 0.5 kb(0.30) 3.6 kb(0.33) 1.9 kb(0.67) 5.8 kb(0.50) 5.3 kb(0.50) 4.6 kb(0.58) 4.2 kb(0.42)	17p13	
266	cCI17-727	D17S566	PvuII 2.6-9.0 kb 10 alleles VNTR, 87% heterozygosity also polymorphic with MspI, TaqI, BglII, PstI, EcoRI		17p13	40
268	cCI17-729	D17S568	MspI	4.6 kb(0.58) 2.6 kb(0.42)	17q11.2	

【 0 0 1 2 】

【 表 6 】



No.	プローブ名	ローカス シンボル	制限酵素	対立遺伝子 (頻度)	染色体上の位置
269	cCI17-730	D17S874	MspI	2.2-3.5 kb 4 alleles VNTR, 83% heterozygosity also polymorphic with TaqI, BglII, PstI, PvuII	17q21.3
270	cCI17-732	D17S570	RsaI	3.2 kb(0.50)	17p13.2
				2.7 kb(0.50)	
		BglII		8.5 kb(0.50)	
				3.2 kb(0.50)	
		PstI		2.5 kb(0.58)	
				1.7 kb(0.42)	
		PvuII		4.2 kb(0.50)	17q25.1
				4.1 kb(0.50)	
271	cCI17-733	D17S875	MspI	3.4 kb(0.75)	
				2.6 kb(0.25)	
272	cCI17-735	D17S572	MspI	4.1 kb(0.83)	17q25.3
				3.4 kb(0.17)	
		RsaI		5.2 kb(0.83)	
				4.1 kb(0.17)	
		PvuII		2.9 kb(0.83)	
				2.2 kb(0.17)	
273	cCI17-736	D17S573	TaqI	1.7-2.5 kb 7 alleles VNTR, 100% heterozygosity also polymorphic with MspI, RsaI, PstI, PvuII	17q21.3
275	cCI17-739	D17S575	MspI	3.3 kb(0.33)	17q25.1
				2.4 kb(0.67)	
278	cCI17-743	D17S876	TaqI	4.3 kb(0.17)	
				2.8 kb(0.83)	

10

20

## 【0013】

RFLPマーカーは、これを用いることによって両親から受け継いだ2本の相同染色体を多型性の差により識別できる (informative, ただし、両方が同じ多型パターンを示す場合には識別できない: not informative) という特徴を有している。この2本の相同染色体の多型性パターンの差 (ヘテロ接合性) が、正常組織で存在し、癌組織で消失している現象 (LOH: loss of heterozygosity) が検出されることは、片方の染色体でそのRFLPマーカー部位が欠失していることを意味する。また一般的に、一对の染色体のうち一方の欠失と他方の変異などによる両方の染色体上の癌抑制遺伝子の不活化が癌化につながるものと考えられており、多くの癌に共通して欠失している領域に癌抑制遺伝子が存在しているものと考えられている。

30

本発明者らは、このようにして得た詳細な染色体地図とRFLPマーカーを使用して、約300例の乳癌と約100例の卵巣癌について第17番染色体のLOHを調べた結果、ヘテロ接合性が識別可能なケースにおいて、17q21付近のcCI17-701とcCI17-730の二つのコスミド・マーカーの間の領域 (2.4 cM) が高頻度に欠失していることが明らかとなった。

卵巣癌における第17番染色体染色体長腕の部分的欠失を図2に示した。黒丸はヘテロ接合性の消失 (LOH)、白丸は保持を表している。2カ所の共通欠失領域は傍線で示した。

乳癌における第17番染色体染色体長腕の部分的欠失を図3に示した。黒丸はヘテロ接合性の消失 (LOH)、白丸は保持を表している。2カ所の共通欠失領域は傍線で示した。

40

## 【0014】

この領域は遺伝性乳癌の家系解析から原因遺伝子の存在が示唆されている領域と一部が重なりあっていた。両者の重なりあう領域の中のコスミド・クローンをプローブとして、650例の乳癌症例における変異を、サザンブロット解析によって調べたところ、2例において、上記で選別されたコスミド・クローンcCI17-904に含まれるDNAの一部の領域が数倍に増幅していることが明らかとなった。この異常について詳細に調べたところ、約6~9 Kbの断片が4~6コピー程度の異常な繰り返しとしてつながっていることが分かった。さらに、このコスミド・クローンの制限酵素断片のうち種を越えて保存されている配列を有する断片をプローブとして、cDNA (メッセンジャーRNAから逆転写された相補配列DNA) ライブラリーをスクリーニングした結果、新規な蛋白質をコード

50

する遺伝子を単離した。この遺伝子の配列構造を決定し、乳癌についての遺伝子異常の有無を調べたところ、明らかな遺伝子変異が同定された。これらのことから、本蛋白質の欠損や異常およびこれをコードする遺伝子DNAの欠失や変異が乳癌・卵巣癌の発生に深く関与していることが明らかとなった。

上記の、マーカー群から該遺伝子の単離までの過程、さらに乳癌組織で遺伝子の異常が起きていた箇所を図4に示した。クローン名はクローン番号でのみ示してある。

#### 【0015】

本発明は、例えば本蛋白質の欠損や異常の有無あるいはそれをコードする遺伝子の欠失や変異の有無を調べることによって、少なくとも、乳癌・卵巣癌の一部についてのリスク診断、早期発見、経過観察、治療方針の決定、予後の推定などの困難な問題に対する解決方法と材料を提供し、この分野の技術を飛躍的に進歩させ得るという点できわめて重要である。

10

すなわち本発明は、(1)配列番号1、2、3または4に示されたアミノ酸配列を有する蛋白質の、全部または一部を含むものからなる蛋白質、(2)配列番号5、6、7、8または9に示された塩基配列を有するDNAの全部または一部を含むものからなるDNA、(3)配列番号5、6、7、8または9に示された塩基配列を有するDNAの、全部または一部を組み込んだプラスミド、該プラスミドで形質転換された宿主細胞および該宿主細胞培養物より発現産物を回収することを含む蛋白質の製造方法、(4)配列番号1、2、3または4で表される蛋白質の全部または一部を含むものを抗原とする抗体、(5)配列番号5、6、7、8または9で表されるDNA配列の一部を含むプライマー、プローブまたはマーカー、またはこれらを使用することを特徴とする遺伝子解析方法、からなる。これら蛋白質とDNAに関しては、それと実質的に同等であるものも本発明に含まれる。

20

#### 【0016】

以下に本発明を詳細に説明する。

##### (1) cDNAクローンの単離

ヒト第17番染色体のコスミド・クローンは、例えばヒト第17染色体だけを含むヒト・マウスの雑種細胞から染色体DNAを抽出し、時野らの報告した方法(Tokino et al., Am. J. Hum. Genet., 48, 258-268, 1991)により、その染色体DNA断片をpWEX15などのベクターに組み込むことにより作成することができる。この中から、全ヒトDNAをプローブとしてコロニーハイブリダイゼーションを行って、ヒト染色体由来の挿入配列(インサート)を持つクローンを選び出すことができる。

30

こうして得られるヒト第17染色体由来DNAを含む多数のコスミド・クローンについて、FISH法により、染色体上の位置を決定してマーカーとし、詳細な物理的染色体地図を作成することができる。また、制限酵素を用いて切断した断片長のパターンからRFLPマーカーを選別することができる(Nakamura et al., Am. J. Hum. Genet. 43, 854-859, 1988)。この地図とRFLPマーカーを利用して癌患者癌組織のDNAについてLOH(ヘテロ接合性の消失)を調べ、癌組織における染色体上の共通欠失領域を第17番染色体のq21付近のきわめて小さい領域に限局化することができる。

この限局化された領域内に存在するコスミド・クローンをプローブとして癌組織DNAのサザンブロット解析を行うことにより、癌組織における遺伝子異常に関わっている配列を有するクローンを選択することができる。さらにコスミド・クローンの制限酵素断片をプローブとして、種々の哺乳動物の染色体DNAに対するサザンブロット法を行うことにより、種を越えて保存されている基本的な細胞機能に関わるDNA配列を含む断片を選択することができる。重要な蛋白質をコードしているDNA配列領域は種を越えて保存されていることが多く、実際、現在までに単離された遺伝性疾患の遺伝子はその多くが種を越えて保存されている(Call, K.M. et al. Cell, 60, 509-520, 1990)。

40

このようにして得られたDNA断片をプローブとして用いることにより、ヒト第17番染色体のq21付近の限局化された領域に存在する遺伝子のcDNAをクローニングすることができる。cDNAの塩基配列は常法に従って決定することができる(Maniatis, J. et al. Molecular Cloning 2nd.ed., Cold Spring Harbor Laboratory Press, N.Y.1989)。

50

このようにして得られた遺伝子DNAのクローンは、その配列についてのSSCP法(Orita, M. et al, Genomics, 5, 874-879, 1984)(Orita, M. et al. Cell, 60, 509-520, 1990)、RNase Protection法(Winter, E., Perucho, M. et al. Proc. Natl. Acad. Sci. USA, 82, 7575-7579, 1985)(Myers, R.M., et al. Science, 230, 1242-1246, 1985)などを用いた癌患者での異常の有無、異常の出現頻度の検索により、目的の原因遺伝子のクローンであることを確かめることができる。

#### 【0017】

##### (2) 遺伝子の全構造の確認

上記の方法により得られたcDNAは、配列番号6および7の新規なDNA配列であることが確認され、対応する配列番号2および3のアミノ酸配列が判明した。さらに5 RACE法、RT-PCR法などによって、配列番号8に示されるDNA配列と対応する配列番号4に示されるアミノ酸配列が明らかにされた。さらに染色体DNAについては、もとのコスミドクローンcCI17-904の塩基配列を解析し、cDNAの塩基配列と比較することによって、イントロン-エクソン結合部が確認され、イントロン/エクソンを含む配列番号9の構造が明らかにされた。

本発明者らは、これらすべてに共通のアミノ酸配列であるところの配列番号1で示される配列の全部または一部を含む蛋白質をMDC蛋白質と命名し、以下MDC蛋白質と記載する。同様に配列番号5には、MDC蛋白質をコードするDNAに共通のDNA配列が示されている。

本発明のDNAは、その一部をプライマーまたはプローブとして用いることにより、遺伝子解析、診断に利用することができる。「一部」とは、プライマーまたはプローブとして使用するオリゴヌクレオチドが本発明のDNA配列と相補的な少なくとも約6個の塩基配列からなり、好ましくは少なくとも約8個の塩基配列、さらに好ましくは約10~12個の、さらに好ましくは約15~25個の塩基配列からなる対応するポリヌクレオチドを意味する。

このDNAがコードするMDC蛋白質は全部または一部をエピトープとして用いて、抗体の作成およびその抗体を用いる研究用、診断用試薬として利用することができる。「エピトープ」とは、ポリペプチドの抗原決定基を意味し、一般に少なくとも5個のアミノ酸で構成される。(例えば、6個のアミノ酸で構成されるポリペプチドが抗体と結合することは公知である。公表特許公報60-500684号)MDC蛋白質の一部とは、本発明のMDC蛋白質の連続するアミノ酸配列を有する少なくとも約3~5個のアミノ酸、好ましくは少なくとも約8~10個のアミノ酸、さらに好ましくは少なくとも約11~20個のアミノ酸からなるポリペプチドを意味する。また約20個以上のアミノ酸からなるポリペプチドであっても使用できることは言うまでもない。

「実質的に同等である」とは、MDC蛋白質およびその一部からなるものにおいて、それらのアミノ酸配列が1つまたは複数個のアミノ酸の置換、欠失、挿入を伴うものであって、MDC蛋白質を用いる研究、診断に同等の効果を奏するものを意味し、これら均等物も本発明に含まれる。また本発明のDNAの場合も、「実質的に同等である」とは上記蛋白質の場合と同様の意味(ただし、1つまたは複数個の塩基配列の置換、欠失、挿入を伴うもの)を示すものである。

#### 【0018】

##### (3) 組換え発現ベクターとその形質転換体

上記記載の方法により得られたヒトMDC蛋白質をコードする遺伝子DNAあるいはその断片を適切なベクターに組み込み、該ベクターを適切な宿主細胞に移入することにより形質転換体を得ることができる。これを常法により培養し培養物よりヒトMDC蛋白質を大量に生産することができる。さらに具体的には、ヒトMDC蛋白質をコードするDNAまたはその断片を、その発現に適したベクターのプロモーター下流に制限酵素とDNAリガーゼを用いる公知の方法により再結合して組換え発現ベクターを作成することができる。使用できるベクターとしては、例えば大腸菌由来のプラスミドpRB322, pUC18, 枯草菌由来のプラスミドpUB110, 酵母菌由来のプラスミドpRB15, ファージ

10

20

30

40

50

ベクタ - g t 1 0 , g t 1 1 あるいは動物ウイルス由来のベクタ - S V 4 0 などが挙げられるが宿主内で複製、増幅可能なベクターであれば特に限定されない。プロモーターおよびターミネーターに関してもヒト M D C 蛋白質をコードする D N A 塩基配列の発現に用いられる宿主に対応したものであれば特に限定されず、宿主に応じて適切な組み合わせも可能である。用いる D N A はヒト M D C 蛋白質をコードする D N A であれば何れでも良く、配列番号 1、2、3 または 4 記載の塩基配列に限定されるものではなく、意図的であるか否かにかかわらず塩基配列の一部が置換、欠損、挿入、あるいはこれらが組み合わされた塩基配列を有する D N A であってもよい。また化学合成によって合成されたものでも良い。

#### 【 0 0 1 9 】

このようにして得られた組換え発現ベクターはコンピテント細胞法 ( J. Mol. Biol., 53, 154, 1970 )、プロトプラスト法 ( Proc. Natl. Acad. Sci. USA, 75, 1929, 1978 )、リン酸カルシウム法 ( Science, 221, 551, 1983 )、インビトロパッケージング法 ( Proc. Natl. Acad. Sci. USA, 72, 581, 1975 )、ウイルスベクター法 ( Cell, 37, 1053, 1984 ) などにより宿主に導入し、形質転換体が作製される。宿主としては大腸菌、枯草菌、酵母および動物細胞などが用いられ、得られた形質転換体はその宿主に応じた適切な培地中で培養される。培養は通常 20 ~ 45 °C、pH 5 ~ 8 の範囲で行われ、必要に応じて通気、攪拌が行われる。培養物からの M D C 蛋白質の分離・精製は公知の分離・精製法を適宜組み合わせて実施すれば良い。これらの公知の方法としては塩析、溶媒沈殿法、透析ゲルろ過法、電気泳動法、イオン交換クロマトグラフィー、アフィニティクロマトグラフィー、逆相高速液体クロマトグラフィーなどが挙げられる。

#### 【 0 0 2 0 】

##### ( 4 ) 抗体の作成

抗体は、M D C 蛋白質の全部あるいは一部分を抗原として、通常の方法で作成することができる。例えば、ポリクローナル抗体はマウス、モルモット、ウサギ等の動物の皮下、筋肉内、腹腔内、静脈に複数回接種し十分に免疫した後、斯かる動物から採血、血清分離して作製する。なお、市販のアジュバントも使用できる。

モノクローナル抗体は、例えば、M D C 蛋白質で免疫したマウスの脾細胞と市販のマウスミエローマ細胞との細胞融合により得られるハイブリドーマを作成後、該ハイブリドーマ培養上清、または該ハイブリドーマ投与マウス腹水から調製することができる。

抗原とする M D C 蛋白質は必ずしも全アミノ酸構造を有する必要はなく、部分構造を有するペプチド、その変異体、誘導体、あるいは他のペプチドとの融合ペプチドであってもよく、調製法は生物学的手法、化学合成手法いずれでもよい。

これら抗体はヒト生体試料中の M D C 蛋白質の同定や定量を可能とし癌診断試薬などに使用できる。

M D C 蛋白質の免疫学的測定法は、公知の方法に準ずればよく、たとえば蛍光抗体法、受身凝集反応法、酵素抗体法などいずれの方法においても実施できる。

#### 【 0 0 2 1 】

##### ( 5 ) ヒト癌組織の遺伝子解析

遺伝子解析される生体試料はヒト正常組織、各種ヒト癌組織をはじめヒト血液、体液、分泌液などを用いることができる。D N A の抽出・調製は、たとえば ( Sato T., et al. Cancer Res., 50, 7184, 1990 ) の方法で行う。

本発明により提供されるヒト M D C 蛋白質をコードする遺伝子 D N A の制限酵素断片をプローブとして、または該遺伝子 D N A の中から適切な位置の塩基配列を適宜選択し、その合成オリゴヌクレオチドをプライマーとして用いることにより該遺伝子の変異の有無を解析することができる。

また、試料中の該遺伝子における挿入、欠失などの異常もこれらの解析により検知することができる。

選択する塩基配列の部位は該遺伝子のエクソン部分、イントロン部分あるいは両部分の結合部分など、いずれも選択し得る。また用いる塩基配列を人為的に改変したものを用いる

10

20

30

40

50

ことができるのは言うまでもなく、これにより対応する遺伝子変異を検出することができる。

解析の方法としては、例えば選ばれた2種の配列のプライマーによりPCR法で部分配列を増幅させ、増幅産物の塩基配列を直接解析するか、あるいはこの増幅産物を前記と同様にプラスミドに組み込み、宿主細胞を形質転換させて培養し、得られるクローンの塩基配列を解析することができる。あるいはまた、Ligase Chain Reaction法(Wu et al., Genomics, 4, 560-569, 1989)、さらに、変異配列特異的PCR法(Ruano and Kidd, Nucleic Acid Research, 17, 8392, 1989)、(C.R. Newton et al., Nucleic Acid Research, 17, 2503-2517, 1989)を利用することにより試料中の該遺伝子の特定の変異の有無を直接検出することができる。

10

また同様に、選ばれたDNA配列あるいはこれに由来するRNA配列を含むプローブを用いて、SSCP法、RNase-Protection法により点突然変異の検出を行うことができる。またこれらのプローブを用いることにより、サザンハイブリダイゼーション法での試料中の該遺伝子の変異の検出、ノーザンハイブリダイゼーション法での試料中の該遺伝子の発現量の異常を検出することができる。

#### 【0022】

なお、このMDC蛋白質をコードする遺伝子DNAを含むプラスミドを保有するEscherichia coli DH5/pBR1、Escherichia coli XL1-Blue MRF'Kan/pCR-5P2、および該遺伝子の染色体DNAを含むコスミドを保有するEscherichia coli 490A/cCl 17-904は通商産業省工業技術院生命工学工業技術研究所にそれぞれ寄託番号FERM BP-4286、FERM BP-4555、FERM BP-4287、として平成5年4月28日、平成6年2月8日、平成5年4月28日寄託された。

20

#### 【0023】

##### 【発明の効果】

本発明の、ヒトMDC蛋白質および該蛋白質をコードする遺伝子DNAの、全部またはその一部を含むものは、癌の研究試薬、検査・診断試薬および治療薬として期待される。

#### 【0024】

##### 【実施例】

以下の実施例により本発明を詳細に且つ具体的に説明するが、本発明はこれらの実施例に限定されるものではない。

30

#### 【0025】

(実施例1) ヒト第17番染色体特異的コスミド・クローンの単離と染色体地図の作成  
時野らの方法(Tokino et al., Am.J.Hum.Genet., 48, 258-268, 1991)によりヒトの正常細胞とマウスの株化細胞とを融合させた雑種細胞の中から、ヒト染色体のうち第17番染色体のみを保有するヒト・マウス雑種細胞株(GM10331)を選択した。この雑種細胞株の染色体DNAを制限酵素Sau3AIで適度に切断し、断片の末端をdATPとdGTPを用いた部分的フィルインによって処理した。このうち35~42kbのサイズの断片を分画し、予め制限酵素XhoIで切断し同様に末端をdCTPとdTTPを用いた部分的フィルインによって処理したコスミドベクターpWEX15に挿入した。こうして得られたコスミド・クローンの中から、<sup>32</sup>Pラベルしたヒト染色体DNAをプローブとしたコロニーハイブリダイゼーション法で、ヒトのDNA断片を含むクローンを選び出すことによって342個のヒト第17番染色体特異的コスミド・クローンを単離した。

40

これらそれぞれのヒト第17番染色体特異的コスミド・クローンについて、そのDNA断片がハイブリダイズする染色体上の位置をFISH法(Inazawa et al., Genomics, 10, 1075-1078, 1991)によって決定し、第17番染色体の物理的染色体地図を作成した(表1~3)(図1)。

染色体上の位置を決定したコスミド・クローン(コスミド・マーカー)について、RFLPを検出するか否かを、6人の非血縁者のDNAを用い、既知の方法(Nakamura et al., Am.J.Hum.Genet., 43, 854-859, 1988)によって検索した。制限酵素は、MspI, TaqI, BglII, PstI, PvuII, RsaI, EcoRIのいずれかを用いた。こ

50

の結果 43 個のクローンが RFLP を検出した (表 4 ~ 6)、すなわち RFLP マーカーとして使用可能であった。

【0026】

(実施例 2) 卵巣癌および乳癌におけるヒト第 17 番染色体長腕の共通欠失領域の検索  
卵巣癌 94 症例と乳癌 246 症例の手術材料より腫瘍組織を得た。対応する正常組織あるいは末梢血サンプルを各々の患者より得た。これらの組織、サンプルより DNA を既知の方法 (Sato et al., Cancer Res., 50, 7184-7189, 1990) で抽出した。DNA は適当な制限酵素で切断し、1.0% アガロースゲル電気泳動を行い、ナイロンメンブレンに 0.1 N - NaOH / 0.1 M - NaCl でサザントランスファーした (Sato et al., Cancer Res., 50, 7184-7189, 1990)。

10

これらのメンブレンに対して、実施例 1 の方法で得られたヒト第 17 番染色体長腕 (17q) 上に位置付けられた RFLP マーカーをプローブとしてサザンハイブリダイゼーション (Sato et al., Cancer Res., 50, 7184-7189, 1990) を行うことにより LOH (ヘテロ接合性の消失) を検索した。結果を表 7 に示す。

【0027】

【表 7】

プローブ	染色体上の 位置	制限 酵素	卵 巢 癌				乳 癌			
			被 検 患 者 数		対立遺伝子の消失／識別可能例 (%)		被 検 患 者 数		対立遺伝子の消失 ／識別可能例 (%)	
			漿膜細胞	粘液細胞	透明細胞	その他	漿膜細胞	粘液細胞	透明細胞	その他
CI17-316	q12-21.1	MspI	32	15	12	22	6/13(46.2)	0/1(0.0)	0/3(0.0)	1/9(11.1)
CI17-592	q21.3	EcoRI	14	13	9	15	2/3(66.7)	0/1(0.0)	0/1(0.0)	2/4(50.0)
CI17-701	q21.3	TaqI	24	14	13	19	9/15(60.0)	2/12(16.7)	0/7(0.0)	5/12(41.7)
CI17-730	q21.3	TaqI	29	15	13	19	6/12(50.0)	0/4(0.0)	0/4(0.0)	2/4(50.0)
CI17-507	q21.3	MspI	22	14	11	20	6/7(85.7)	1/3(33.3)	1/3(33.3)	2/5(40.0)
CI17-533	q21.3	TaqI	22	13	11	16	6/11(54.5)	3/9(33.3)	1/7(14.3)	4/9(44.4)
CI17-7	q22	PvuII	21	8	9	15	4/5(80.0)	0/1(0.0)	0/3(0.0)	1/3(33.3)
CI17-489	q23	MspI	26	13	11	21	5/5(100.0)	0/2(0.0)	0/3(0.0)	3/8(37.5)
CMM86	q23	TaqI	28	13	10	17	6/17(35.3)	1/8(12.5)	0/6(0.0)	2/10(20.0)
CI17-516	q25.1	TaqI	29	14	14	21	6/17(35.3)	1/10(10.0)	0/7(0.0)	6/11(54.5)
CI17-710	q25.3	TaqI	18	13	10	12	4/8(50.0)	3/8(37.5)	0/6(0.0)	3/7(42.9)

## 【 0 0 2 8 】

卵巣癌においては94例中84例で少なくとも1つのマーカーによりヘテロ接合性を識別可能(informative)であり、そのうち33例(39.3%)で少なくとも1つのマーカーによりLOHが検出された。乳癌では246例中214例が少なくとも1つのマーカーについてヘテロ接合性が識別可能であり、そのうち88例(41.4%)で少なくとも1つ

10

20

30

40

50

のマーカ－がLOHを検出した。

#### 【0029】

上記の結果からヒト第17番染色体長腕(17q)に部分的欠失があると判定できる例、すなわち2つ以上のマーカ－でヘテロ接合性を識別可能(informative)であり、なおかつ欠失部分(LOH検出)と保持部分(LOH不検出)とを示す例を集めて解析した。

その結果、卵巣癌では8例で2つの共通欠失領域が見出された(図2)。1つはCI17-316(17q12-21.1)とCI17-507(17q21.3)の2つのマーカ－間の領域であり、もう一つはCI17-516(17q25.1)マーカ－より末端側の領域である。

同様に、乳癌では35例で2つの共通欠失領域が見いだされた(図3)。1つは卵巣癌においても見いだされたCI17-701(17q21.3)とCI17-730(17q21.3)の2つのマーカ－間の領域でありさらに狭い領域に限局化されている、もう1つはCI17-516(17q25.1)マーカ－より末端側の領域であり、これも卵巣癌で欠失の認められた領域である。

上記の欠失領域の解析(deletion mapping)で得られた2つの共通欠失領域のうち、CI17-701とCI17-730の2つのマーカ－の間の領域は、遺伝性乳癌および卵巣癌の家系解析(linkage mapping)の結果から発症と強く連関を示す領域17q21(Hall et al., Am.J.Hum.Genet., 50, 1235-1242, 1992)に近いものであった。この領域の長さ(2つのマーカ－の間の遺伝学的距離)はリンケージ解析法(Lathrop et al., Am.J.Hum.Genet., 37, 482-498, 1985 ;Donis-Keller et al., Cell, 51, 319-337, 1987)により2.4 cMと推定された。

#### 【0030】

(実施例3) 最小限局領域に含まれるコスミド・クロ－ンの単離

家系解析の結果から限局化された領域は17q21上のTHRA1とMfd188の2つのマーカ－間の領域であることが示されている(Hall et al., Am.J.Hum.Genet., 50, 1235-1242, 1992)(Bowcock A.M. et al., Am.J.Hum.Genet., 52, 718-22, 1993)ので、これらのマーカ－とCI17-701とCI17-730マーカ－の相対的順序を決定することにより2つの別々の戦略によって得られたマッピング情報を統合することを試みた。各マーカ－の相対的順序の決定は、本発明者らが新たに開発した2色FISH法(2-color Fish method)によって行った。この方法は、細胞の同調操作によって高度に伸展した染色体標本を作製して用いることにより精密度を増し、さらに異なる色の蛍光体で標識したプローブによっておこなうFISH法であり、これにより極めて近接したマーカ－の相対的順序の決定が可能である。

#### 【0031】

その結果、Mfd188マーカ－はCI17-701とCI17-730のマーカ－の間に位置し、THRA1マーカ－はCI17-701よりもセントロメア側に位置することが見出された(図4a)。すなわち、家系解析によって限局化された遺伝性乳癌と連関する領域と、欠失領域の解析で限局化された散发性乳癌における共通欠失領域とは互いに重なり合っており、重なり合っている最小領域はCI17-701とMfd188の2つのマーカ－にはさまれる領域であることを見出した(図4a)。パルスフィールドゲル電気泳動によりこの領域の物理的地図を作製したところ、重複領域の長さは約500 kbと非常に限局化された。

さらに、実施例1の方法で得られたコスミド・クロ－ンのうち17q21.3に位置づけられた37個のクロ－ンと既知のマーカ－THRA1、Mfd188およびPPYについて2色FISH法による精密なマッピングを行った。その結果、CI17-701とCI17-730の2つのマーカ－に囲まれる領域に15個のコスミド・クロ－ンが位置づけられ、そのうち、上記重複領域上にはCI17-527とCI17-904の2つのコスミド・クロ－ンが位置していることが判明した(図4a、b)。

#### 【0032】

(実施例4) 乳癌における遺伝子異常の検出



重複領域約500kbのうち約150kbはすでに4つのコスミド・クローンC I 17 - 701、C I 17 - 527、C I 17 - 904およびM f d 188でカバーされているので、まず650例の散発性乳癌組織のDNAの制限酵素(S a c I、P v u IIまたはP s t I)断片に対してこれらのコスミド・クローンのDNAあるいはその断片をプローブとしたサザンブロット解析によって、癌細胞に生じている欠失、重複、増幅、転座などの大きな構造的遺伝子異常、いわゆる遺伝子再構成を検出することを試みた。

その結果、C I 17 - 904のDNAまたはその9.5kb H i n d III断片(図4c)をプローブとしたとき、2例の乳癌組織において遺伝子再構成が検出された(図5a, b)。これらの遺伝子再構成は癌組織においてのみ起きており、正常組織では認められない、サイズの異なる余分のバンドが認められ、さらにいくつかのバンドの濃度が増大していた。すなわち、このプローブに対応する一定のDNA領域に遺伝子増幅が生じていた。2例のうち、1例においては、9.5kb H i n d III断片(図4c)と隣接するE - H 5.2もしくはH i n d 6.1断片をプローブとして乳癌組織のDNA S a c I断片に対してサザンブロット解析を行った時には遺伝子増幅は検出されなかった(図6、Case 1)。すなわち、このCaseの遺伝子増幅は9.5kb H i n d III断片に対応する領域内に起きており、4~5倍に増幅していることが示された。

これを、より詳細に検討するため9.5kb H i n d III断片内の6個のS a c I断片、A、B、C、D、E、F(図4c)の各々をプローブとして乳癌組織のDNA S a c I断片に対してサザンブロット解析を行った。その結果、プローブA、Bで2.5kb、プローブB、C、Dで3.0kb、プローブE、Fで2.5kb、プローブFで0.9kbの異常サイズの増幅バンドが認められた(図7)。

#### 【0033】

他の1例においては、乳癌組織のDNA S a c I断片に対してE - H 5.2をプローブとした時に遺伝子増幅が検出された(図6、Case 2)が、H i n d 6.1断片をプローブとした時には検出されなかった(図6、Case 2)。この時、E - H 5.2をプローブとした時にはサイズ異常を示すバンドは認められず、増幅のみが認められた。すなわち、このCaseの遺伝子増幅は9.5kb H i n d III断片に対応する領域内からE-H5.2に対応する領域のさらに外側(テロミア側)にわたる範囲に起きていることが示された。

#### 【0034】

(実施例5) cDNAの単離と構造決定

2例の乳癌組織において遺伝子再構成の認められた領域内あるいはその近傍から、発現されている遺伝子を単離するために、コスミド・クローンC I 17 - 904のDNA断片について、種を越えて保存されている基本的な細胞機能に関わるDNA配列を含むものの選択を行った。すなわち、コスミド・クローンC I 17 - 904のDNA断片の各々をプローブとして、ウシ、ブタ、マウス、ラット、ニワトリのDNA断片のサザンブロットに対してハイブリダイゼーションを行った。その結果、コスミド・クローンC I 17 - 904の3.5kb H i n d III - K s p I断片(図4c)がウシ、ブタ、マウス、ラットのDNAとハイブリダイズし、種を越えてよく保存されていた。

この3.5kb H i n d III - K s p I断片をプローブとして乳腺、乳癌細胞株、胎児脳、大脳、小脳の5つの異なる臓器由来のヒトcDNAライブラリーをスクリーニングした。このうち、小脳cDNAライブラリーから最長のcDNAがクローニングされた。このcDNAはコスミド・クローンC I 17 - 904の3.5kb H i n d III - K s p I断片および隣接する複数の制限酵素断片とハイブリダイズし、その範囲は染色体上で20kbを超える領域に広がっていた。

このcDNAの塩基配列を解析した結果、2923bpよりなり、27bpの5'非翻訳領域、1575bpのコーディング領域、1306bpの3'非翻訳領域および15bpのpoly(A)テールを含む新規なDNA塩基配列であることを確認した(配列番号6)。このcDNA配列に含まれるオープンリーディングフレームは、524アミノ酸よりなる新規な蛋白質(MDC蛋白質:配列番号2)をコードしていた。オープンリーディングフレームの最初のATGのすぐ上流にはイン・フレームの停止コドンが存在した。ポリ

10

20

30

40

50

アデニル化シグナル A A T A A A はポリアデニル化部位から約 20 bp 上流に認められた。

#### 【0035】

(実施例6) 染色体DNAの構造決定

実施例5で得られたcDNAに対応する染色体DNAの構造を明らかにするために、コスミド・クローンC I 17-904について、このcDNAの塩基配列を含む部分とその周辺の塩基配列を決定し、両者の配列を比較することにより、エクソン・イントロン結合部を決定した。その結果、実施例5で得られたcDNAに対応する25個のエクソンを含む新規な遺伝子DNAの配列構造が明らかとなった(配列番号9)。25個のエクソンは比較的小さなサイズで、染色体上約20kbの領域にまたがって存在していることが示された。

10

#### 【0036】

(実施例7) 乳癌における該遺伝子エクソン構造の異常の検出

実施例6で明らかとなったエクソン/イントロンを含む該遺伝子DNAの構造から、実施例4で2例の乳癌組織に共通して異常を検出したプローブ(コスミド・クローンC I 17-904の9.5kb Hind III 断片)の配列領域中に、第2、第3、第4エクソンが存在していることが明らかとなった。さらに、第2エクソンはプローブEの配列領域中に存在し、第3、第4エクソンはプローブFの配列領域中に存在する(図4c)。従って、実施例4の9.5kbのHind III 断片領域を含む遺伝子再構成は該遺伝子の3つのエクソンを含む領域において正常なエクソン構造を破壊していると考えられる。このことを確認するために、第2、第3、第4エクソンに対応するDNA配列のプローブによって、上記2例の乳癌組織の染色体DNAに対するサザンブロット解析を行った結果、先に示したプローブEあるいはプローブFを用いて行った結果(図7)と同様の異常サイズの増幅バンドが認められた。

20

#### 【0037】

(実施例8) 遺伝子発現の組織特異性

実施例5で得られたcDNAをプローブとして、種々のヒト組織(脳、心臓、腎臓、肝臓、肺、脾臓、胎盤、骨格筋、大腸、末梢血リンパ球、卵巣、小腸、脾臓、睾丸、胸腺)由来のmRNAについてのノーザンブロット解析を行った結果、脳に最も強い発現が、心臓、卵巣、睾丸に弱い発現が認められた。

30

さらに、より微弱な発現を検出するためにRT-PCR(reverse-transcriptase PCR)による増幅での検出をおこなった。すなわち、種々のヒト組織由来のmRNAから逆転写酵素反応により、ランダムヘキサマーをプライマーとして1本鎖cDNAを合成した。これを鋳型として実施例6で明らかとなったエクソン21および23の配列にそれぞれ由来するプライマーBCO9およびBCO12を用いてPCRをおこなった。その結果、期待される大きさのPCR産物は主に中枢神経系(大脳、小脳および胎児脳)と内分泌系または生殖系臓器(睾丸、卵巣、乳腺、副腎、胸腺および脾臓)に認められた。

用いたプライマーの配列は下記のとおりである。

BCO9 5'-GCACCTGCCCCGGCAGT-3'

(コーディング鎖、配列番号6の塩基番号1764-1780に相当)

40

BCO12 5'-CCAGGACAGCCCCAGCGATG-3'

(アンチセンス鎖、配列番号6の塩基番号1976-1957に相当)。

#### 【0038】

(実施例9) RT-PCRによるmRNAの直接塩基配列決定

ヒト胎児脳およびヒト睾丸mRNAをもとに、エクソン19上の配列に由来するプライマー-GMA701とエクソン21上の配列に由来するプライマー-GMB704を用いてRT-PCRを行い、増幅されたDNAの塩基配列をプライマー-GMA702またはプライマー-GMB703を用いて直接決定した。その結果、実施例5で得られた配列番号6の小脳由来のcDNA配列から10塩基(塩基番号1512-1521の欠失した配列が発見され、配列番号7のDNA配列に対応するmRNAが発現していることが明らかとなった。胎児脳、

50

睾丸のいずれの mRNA でも同一の結果が得られた。配列番号 7 の cDNA 配列に含まれるオープンリーディングフレームは、670 アミノ酸よりなる MDC 蛋白質 (配列番号 3) をコードしていた。

これは配列番号 9 の染色体 DNA 上のエクソン 20 の開始が塩基番号 6073 番ではなく塩基番号 6083 番から始まる別の RNA スプライシングによっているものと考えられた。このようなスプライシングのバリエーションは例えば Oda らの報告によっても知られている (Biochem. Biophys. Res. Commun. 193, 897-904 (1993))。これにより配列番号 6 の cDNA と配列番号 7 の cDNA のコードするアミノ酸配列 (配列番号 2 と配列番号 3) は、この部分以降で異なっている。すなわち配列番号 6 の cDNA ではエクソン 20 内の部分で停止コドンを生じるが、配列番号 7 の cDNA ではリーディングフレームがシフトしてさらに下流までオープンリーディングフレームが続いている。

PCR および塩基配列決定に用いたプライマーの配列は下記のとおりである。

GMA701 5'-GGCTGCTGATCGCTTCTGCTAC-3'

(コーディング鎖、配列番号 6 の塩基番号 1413-1434に相当)

GMA702 5'-GAGAAGCTGAATGTGGAGGG-3'

(コーディング鎖、配列番号 6 の塩基番号 1435-1456に相当)

GMB703 5'-GTCAGAGCCGTCGCCAGC-3'

(アンチセンス鎖、配列番号 6 の塩基番号 1675-1657に相当)

GMB704 5'-GCCATCCTCCACATAGCTCAGG-3'

(アンチセンス鎖、配列番号 6 の塩基番号 1696-1675に相当)。

#### 【0039】

(実施例 10) RACE 法による 5' 末端配列の増幅

配列番号 7 の cDNA の全長 cDNA を得るため、5' - cDNA 末端の PCR 増幅 (5' - RACE; Frohman, et al, Proc. Natl. Acad. Sci. USA, 85, 8998-9002, 1988; Belyavski, et al, Nucleic Acid Res. 17, 2919-2932, 1988) を行った。ヒト脳由来 poly A(+) RNA 2 µg (Clontech社) をもとに特異的オリゴマー SGN012 をプライマーとし市販の合成キットを使用して一本鎖 cDNA を合成した。一本鎖 cDNA 末端にアンカーオリゴマーを結合する Edwards らの方法 (Nucleic Acid Res. 19, 5227-5232, 1991) を応用した市販キットを用いて 5' RACE を行った。キットのアンカーオリゴマーと特異的オリゴマー SGN012 をプライマーとして PCR を行った結果、約 580 bp の増幅産物を電気泳動法で検出した。

この増幅産物を電気泳動ゲルより抽出、精製し、プラスミドベクター pCR-Script (Stratagene 社) の Srf I 切断部位に挿入しクローニングした。各クローンよりプラスミド DNA を精製し、塩基配列を決定した。そのうちの一つである pCR-5P2 はアンカーオリゴマーの配列に続いて ATG から始まる 501 bp の cDNA インサートを有し、その 315 番目以降の塩基配列は配列番号 7 の塩基番号 45 (エクソン 2 の開始位置) 以降の配列と、後述する 1 塩基を除いて、まったく一致した。さらに、pCR-5P2 の最初の ATG から始まるリーディングフレームは、配列番号 7 の cDNA のコードするポリペプチドとフレームが一致した。また、これによって得られたポリペプチド配列の N 末端領域には疎水性に富むアミノ酸が連続するシグナルペプチドがコードされていた。

#### 【0040】

以上の、5' RACE によって得られた、5' 末端配列が mRNA 上で真に配列番号 7 の塩基番号 45 以降の配列とつながっていることを確認するために RT-PCR を行った。ヒトの脳、胎児脳、卵巣、および睾丸由来の poly A(+) RNA (Clontech社) をもとにランダムヘキサマーをプライマーとして一本鎖 cDNA を合成し、次いで pCR-5P2 の最初の 20 塩基の配列を有するオリゴマー (SGN013) をセンスプライマー、SGN011 または SGN012 をアンチセンスプライマーとした。その結果、いずれの組織の RNA を用いた場合においても、期待される増幅産物 (SGN013 / SGN011 で約 500 bp、SGN013 / SGN012 で約 750 bp) を電気泳動法で検出した。従って、

5' RACE によって得られた pCR-5P2 の 5' 末端配列が mRNA 上で配列番号 7 の塩

10

20

30

40

50

基番号 45 以降の配列とつながっていることが確認され、配列番号 8 の cDNA が構成された。配列番号 8 の cDNA のオープンリーディングフレームは、769 アミノ酸よりなる MDC 蛋白質 (配列番号 4) をコードしている。

用いた特異的オリゴマーの配列は以下の通りである。

S G N 0 1 1    5' -GATGTAAGTCAAGTTCCCATCAGAGA- 3'

( アンチセンス鎖、配列番号 7 の塩基番号 231-206 に相当 )

S G N 0 1 2    5' -AACAGCTGGTGGTCGTTGATCACAA- 3'

( アンチセンス鎖、配列番号 7 の塩基番号 485-461 に相当 )

S G N 0 1 3    5' -ATGAGGCTGCTGCGGCGCTG- 3'

( コーディング鎖、配列番号 8 の塩基番号 1-20 に相当 )

10

なお、エクソン 2 の開始位置以降で配列番号 8 が配列番号 6 または 7 と相違している 1 塩基は、エクソン 2 の開始位置から 4 番目の塩基 (配列番号 8 の 318 番目の C、配列番号 6 および 7 の 48 番目の A) に当たり、コードするアミノ酸は配列番号 4 の 106 番目 His、配列番号 2 および 3 の 7 番目 Glu である。これは多型を反映しているものと考えられる。

これら 3 つの MDC 蛋白質 (配列番号 2、3 および 4) のバリエーションに共通なアミノ酸配列は、488 アミノ酸からなる配列 (配列番号 1) であり、この部分をコードする DNA 配列も共通の配列 (配列番号 5) である。

#### 【 0 0 4 1 】

( 実施例 1 1 ) 既知蛋白質とのホモロジー

20

MDC 蛋白質のアミノ酸配列は、蛇毒の出血性毒素である HR1B (Takeya et al., J. Biol. Chem., 265, 16068-16073, 1990)、Prorhodostomin (Au et al., Biochem. Biophys. Res. Commun., 181, 585-593, 1991)、Protrigramin (Neeper et al., Nucleic Acid Res., 18, 4255, 1990) とホモロジーを有していた。また、モルモット精子表層蛋白である PH30 (Blobel et al., Nature, 356, 248-252, 1992)、ラットまたはサル副睾丸の蛋白である EAP1 (Perry et al., Biochem. J., 286, 671-675, 1992) ともホモロジーを有していた。MDC 蛋白質、配列番号 2 (524 アミノ酸)、配列番号 4 (769 アミノ酸) とのホモロジーの「一致% / 対象領域アミノ酸数」は次のとおりであった。左が配列番号 2 について、右が配列番号 4 についての値である。

HR1B	32.5 / 335	32.2 / 379
Prorhodostomin	29.0 / 420	29.0 / 420
Protrigramin	27.7 / 430	28.1 / 438
PH30b	38.1 / 147	30.8 / 302
EAP1 rat	36.0 / 364	33.1 / 475
EAP1 monkey	30.4 / 503	29.9 / 599

30

#### 【 0 0 4 2 】

( 実施例 1 2 ) 形質転換体の作製

MDC 蛋白質 (配列番号 2) をコードする DNA (配列番号 6) を基質として、プライマー SGN006 と SGN008 を用いて、MDC 蛋白質の一部をコードする DNA 断片を PCR で増幅した。用いたプライマーの配列は以下の通りである。

40

S G N 0 0 6    5' -CACAGATCTGGGGGCATATGCTCCCTG- 3'

( コーディング鎖、配列番号 6 の塩基番号 766-783 に相当 )

S G N 0 0 8    5' -AACAAGCTTCTACTGATGTCTCCACC- 3'

( アンチセンス鎖、配列番号 6 の塩基番号 1602-1585 に相当、下線は終止コドンを示す )  
ベクター構築用に、プライマーの 5' 端にはそれぞれ Bgl II、Hind III の切断部位配列を付加してある。

PCR 増幅産物をアガロースゲル電気泳動によって分取し、Bgl II と Hind III で切断した。こうして得られた MDC 蛋白質の一部をコードする DNA 断片を、予め BamH I と Hind III で切断しておいた pMAL-c2 (New England Biolabs 社製) ベクターに結合してプラスミド pMAL-MDC(C1) を構築した。

50

同様に、予め BamH I と Hind III で切断しておいた pQE-13 (Diagen 社製) ベクターに結合してプラスミド pH6-MDC(C1) を構築した。

さらに、pMAL-MDC(C1) の M D C 蛋白質コーディング領域から BamH I 切断部位 (配列番号 6 の塩基番号 1483) 以降を除き、M D C 蛋白質の 2 つのパリエーション (配列番号 2 および 3) に共通する部分のコーディング領域となるようにした。すなわち pMAL-MDC(C1) を BamH I と Hind III で切断し、末端を平滑化したのち再結合して、プラスミド pMAL-MDC(dC1) を構築した。

#### 【 0 0 4 3 】

pMAL-c2 ベクターに組み込んだ断片は、N 末端側にマルトース結合蛋白質 (M B P) を有する融合蛋白質として発現されるので、その融合蛋白質はアミロースカラムによってアフィニティー精製した。また、pQE-13 ベクターに組み込んだ断片は、N 末端側に 6 個のヒスチジン残基よりなるペプチド (His 6) を有する融合蛋白質として発現されるので、その融合蛋白質は金属キレートカラムによってアフィニティー精製した。

各プラスミド pMAL-MDC(C1) pMAL-MDC(dC1) および pH6-MDC(C1) を用いて E.coli JM 109 をトランスフォームし、アンピシリン耐性で選択してそれぞれの形質転換体を得た。

#### 【 0 0 4 4 】

(実施例 1 3) 組換え M D C 蛋白質の発現と精製

実施例 1 2 で得られたそれぞれの形質転換体を培養し、培養物より組換え M D C 融合蛋白質を抽出、精製した。

すなわち、各形質転換体を 100ml の L B 培地 (1%ポリペプトン、0.5% 酵母抽出物、1% NaCl) で 37℃ 一夜振とう培養した。培養液を予め 37℃ に加温した L B 培地で 10 倍に希釈したうえ、さらに 30 分 ~ 90 分培養して、対数増殖期の培養物を得た。培養物 1 リットル - に I P T G ( Isopropyl-beta-D-thiogalactopyranoside ) を終濃度 1 mM となるように添加して 3 ~ 4 時間培養した。培養物から遠心分離により菌体を集めた。

プラスミド pMAL-MDC(C1)、pMAL-MDC(dC1) による形質転換体の場合は、菌体に 10 ml のカラムバッファー (20mM Tris-HCl pH7.4, 200mM NaCl) を加え超音波によって破碎した。組換え M D C 融合蛋白質は、破碎液の不溶性画分に存在したので、これを遠心分離して変性バッファー (8M 尿素、20mM Tris-HCl pH8.5, 10mM ジチオスライトール) に溶解した。次いで、これをカラムバッファーに透析後、遠心分離して上清可溶画分を集めた。透析不溶性画分は、さらに変性、透析、遠心分離を繰返して上清可溶画分を回収した。集めた可溶性画分をアミロースカラム (New England Biolabs 社製) につけ、カラムバッファーで洗浄後、10 mM マルトースを含むカラムバッファーで溶出した。溶出画分は、280 nm の吸光度および SDS ポリアクリルアミド電気泳動法 (クマシーブルー染色) で解析して分画した。この結果、プラスミド pMAL-MDC(C1) および pMAL-MDC(dC1) による形質転換体のそれぞれで、期待される約 68 Kd の M B P 融合蛋白質が主要バンドとして検出される画分が得られた。収量はそれぞれ 46.4 mg、10.0 mg であった (OD280 = 1 のとき、1 mg/ml としたとき)。これらの融合蛋白質を以下、それぞれ MBP-MDC(C1)、MBP-MDC(dC1) と称する。

#### 【 0 0 4 5 】

同様に、プラスミド pH6-MDC(C1) による形質転換体の場合は、菌体に 10 ml のソニケーションバッファー (10mM リン酸ナトリウム pH8.0, 200mM NaCl) を加え超音波によって破碎した。組換え M D C 融合蛋白質は、破碎液の不溶性画分に存在したので、これを遠心分離してバッファー A (6M塩酸グアニジン, 100mM NaH<sub>2</sub>PO<sub>4</sub>, 10mMTris-HCl, pH8.0) に溶解し、遠心分離して上清可溶画分を集め、Ni - N T A カラム (Diagen社製) につけ、バッファー A、次いでバッファー B (8M尿素, 100mM NaH<sub>2</sub>PO<sub>4</sub>, 10mMTris-HCl, pH8.0) で洗浄後、バッファー C (8M尿素, 100mM NaH<sub>2</sub>PO<sub>4</sub>, 10mMTris-HCl, pH6.3)、バッファー D (8M尿素, 100mM NaH<sub>2</sub>PO<sub>4</sub>, 10mMTris-HCl, pH5.9)、バッファー E (8M尿素, 100mM NaH<sub>2</sub>PO<sub>4</sub>, 10mMTris-HCl, pH4.5) およびバッファー F (6M塩酸グアニジン, 200mM酢酸) で段階的に溶出した。溶出画分は、280 nm の吸光度および SDS ポリアクリルアミド電気泳動法 (クマシーブルー染色) で解析して分画した。この結果、バッファー F によ

10

20

30

40

50

る溶出液に、期待される約 34 Kd の His6 融合蛋白質が単一バンドとして検出される画分が得られた。収量は 51.9 mg であった (OD280 = 1 のとき、1 mg/ml としたとき)。この融合蛋白質を以下、His 6-MDC(C1) と称する。

#### 【0046】

(実施例 14) モノクローナル抗体およびウサギポリクローナル抗体の作製

実施例 13 で得られた 3 種の組換え融合蛋白質、His 6-MDC(C1)、MBP-MDC(dC1)、および MBP-MDC(C1) を、それぞれ、免疫抗原、抗体精製・スクリーニング用抗原、測定用標準抗原として用いた。

抗 MDC 蛋白特異的モノクローナル抗体は、His 6-MDC(C1) をマウスに免疫して作製した。すなわち、His 6-MDC(C1) の 3 M 尿素 / PBS 溶液 (500 - 1000 µg/ml) を完全アジュバントと 1 : 1 の割合で混合しマウスの腹腔内に 100 µg / 匹にて 2 週間隔で 4 ~ 6 回免疫を行った。免疫終了後、P3U1 細胞と B 細胞とのハイブリドーマを PEG1500 を用いて作製し、培養上清中の抗体価をモニターし、抗 MDC 蛋白特異的抗体を産生するハイブリドーマの選択を行った。

抗体価の測定は、実施例 13 で得た MBP-MDC(dC1) 融合蛋白質を固相化 (5 µg/ml) したポリスチレン製カップに培養上清 100 µl を加え第一反応を行い、洗浄後、抗マウス IgG-HRP (Horse-raddish peroxidase) を加え第二反応を行った。洗浄後、酵素基質溶液 (過酸化水素水および ABTS (2, 2'-アジノ-ビス-(3-エチルベンゾチアゾリン-6-スルホン酸) 混合液) を添加し、発色反応 (第三反応) を行いモニターした。

ハイブリドーマを 96 ウエルマルチプレートにて培養し、HAT 選択を行い、約 2 週間後に培養上清中の抗体価を測定し抗原と特異的に反応するクローンを選択した。更に、クローニング操作を行い、3 クローン (G1-5A2-2C8, G2-2F2-3D11, G2-2D10-3F5) を抗体産生ハイブリドーマとして樹立した。樹立したクローンの産生抗体のクラスおよびサブクラスは、G1-5A2-2C8 は IgG<sub>1</sub>, G2-2F2-3D11 は、IgG<sub>2b</sub>, G2-2D10-3F5 は IgM であった。各ハイブリドーマ細胞 300 万個を、予め約 1 週間前に 0.5ml のプリスタンを腹腔内に投与しておいた BALB/c マウスの腹腔内に接種し、8 ~ 10 日後に腹水を採取した。各腹水よりプロテイン G カラムによるアフィニティークロマトグラフィーで抗体を精製した。

#### 【0047】

同様に、実施例 13 で得られた His6-MDC(C1) を免疫抗原として、ウサギに免疫し、抗 MDC 蛋白ポリクローナル抗体を作製した。

すなわち、マウスと同様、His6-MDC(C1) の 3 M 尿素 / PBS 溶液 (500 - 1000 µg/ml) を完全アジュバントと 1 : 1 の割合で混合し免疫を行った。免疫終了後、抗血清を得、実施例 13 で得た MBP-MDC(dC1) 融合蛋白質を固相化したポリスチレン製カップを用いて、抗体価を測定した。抗血清を 500 倍から 64000 倍まで希釈し、その 100 µl をウエルに添加し、抗体価をやぎ抗ウサギ IgG-HRP を用いて検討したところ、64000 倍まで検出が可能であった。また、免疫前の血清中には、MBP-MDC(dC1) と反応する抗体が存在しなかったことから、MDC 蛋白質に特異的に反応する抗体が産生されていることが確認することができた。さらに、この抗血清を、プロテイン G カラムおよび MBP-MDC(dC1) 融合蛋白質を固相化したセファロースカラムによるアフィニティークロマトグラフィーで精製した。

#### 【0048】

このようにして得られた精製モノクローナル抗体および精製ウサギポリクローナル抗体を用いた ELISA 法による MDC 蛋白質の定量法を確立した。

すなわち、ハイブリドーマ (G2-2F2-3D11) 由来の精製モノクローナル抗体を 96 ウエルプレートに固相化し、BSA (Bovine serum albumin) でブロッキング後、精製 MBP-MDC(C1) 溶液を被検液として、0.156 ~ 5.00 µg/ml の範囲で 100 µl / ウエルずつ添加して室温 1 時間反応させた。ウエルを洗浄後、精製ウサギポリクローナル抗体溶液 (5 µg/ml) を 100 µl / ウエルずつ添加して室温 1 時間反応させた。ウエルを洗浄後、抗

ウサギ Ig G - HRP (5ug/ml) を 100 ul / ウエルずつ添加して室温 1 時間反応させた。反応終了後 2 mM アジ化ナトリウムを 100 ul / ウエルずつ添加し、405nm と 490nm の吸光値を測定した。得られた差分吸光値は、被検液濃度とよく相関した値を示し、0 ~ 2.5  $\mu$ g/ml の間でほぼ直線的な関係であることが確かめられた(図 8)。このことは、これらのモノクローナル抗体およびウサギポリクローナル抗体による ELISA 法が MDC 蛋白質の定量法として十分に使用できることを示している。

【 0 0 4 9 】

【配列表】

配列番号： 1

配列の長さ： 4 8 8

配列の型：アミノ酸

配列の種類：タンパク質

トポロジー：直鎖状

起源

生物名：ホモサピエンス

直接の起源

ライブラリー名：ヒト胎児脳 cDNA ライブラリー

## 配列

Leu Leu Ser Ser Gln Tyr Val Glu Arg His Phe Ser Arg Glu Gly Thr  
 1 5 10 15  
 Thr Gln His Ser Thr Gly Ala Gly Asp His Cys Tyr Tyr Gln Gly Lys  
 20 25 30  
 Leu Arg Gly Asn Pro His Ser Phe Ala Ala Leu Ser Thr Cys Gln Gly  
 35 40 45  
 Leu His Gly Val Phe Ser Asp Gly Asn Leu Thr Tyr Ile Val Glu Pro  
 50 55 60  
 Gln Glu Val Ala Gly Pro Trp Gly Ala Pro Gln Gly Pro Leu Pro His  
 65 70 75 80  
 Leu Ile Tyr Arg Thr Pro Leu Leu Pro Asp Pro Leu Gly Cys Arg Glu  
 85 90 95  
 Pro Gly Cys Leu Phe Ala Val Pro Ala Gln Ser Ala Pro Pro Asn Arg  
 100 105 110  
 Pro Arg Leu Arg Arg Lys Arg Gln Val Arg Arg Gly His Pro Thr Val  
 115 120 125  
 His Ser Glu Thr Lys Tyr Val Glu Leu Ile Val Ile Asn Asp His Gln  
 130 135 140  
 Leu Phe Glu Gln Met Arg Gln Ser Val Val Leu Thr Ser Asn Phe Ala  
 145 150 155 160  
 Lys Ser Val Val Asn Leu Ala Asp Val Ile Tyr Lys Glu Gln Leu Asn  
 165 170 175  
 Thr Arg Ile Val Leu Val Ala Met Glu Thr Trp Ala Asp Gly Asp Lys  
 180 185 190  
 Ile Gln Val Gln Asp Asp Leu Leu Glu Thr Leu Ala Arg Leu Met Val  
 195 200 205  
 Tyr Arg Arg Glu Gly Leu Pro Glu Pro Ser Asn Ala Thr His Leu Phe  
 210 215 220

10

20

30

40



Ser Gly Arg Thr Phe Gln Ser Thr Ser Ser Gly Ala Ala Tyr Val Gly  
 225 230 235 240  
 Gly Ile Cys Ser Leu Ser His Gly Gly Gly Val Asn Glu Tyr Gly Asn  
 245 250 255  
 Met Gly Ala Met Ala Val Thr Leu Ala Gln Thr Leu Gly Gln Asn Leu  
 260 265 270  
 Gly Met Met Trp Asn Lys His Arg Ser Ser Ala Gly Asp Cys Lys Cys  
 275 280 285  
 Pro Asp Ile Trp Leu Gly Cys Ile Met Glu Asp Thr Gly Phe Tyr Leu  
 290 295 300  
 Pro Arg Lys Phe Ser Arg Cys Ser Ile Asp Glu Tyr Asn Gln Phe Leu  
 305 310 315 320  
 Gln Glu Gly Gly Gly Ser Cys Leu Phe Asn Lys Pro Leu Lys Leu Leu  
 325 330 335  
 Asp Pro Pro Glu Cys Gly Asn Gly Phe Val Glu Ala Gly Glu Glu Cys  
 340 345 350  
 Asp Cys Gly Ser Val Gln Glu Cys Ser Arg Ala Gly Gly Asn Cys Cys  
 355 360 365  
 Lys Lys Cys Thr Leu Thr His Asp Ala Met Cys Ser Asp Gly Leu Cys  
 370 375 380  
 Cys Arg Arg Cys Lys Tyr Glu Pro Arg Gly Val Ser Cys Arg Glu Ala  
 385 390 395 400  
 Val Asn Glu Cys Asp Ile Ala Glu Thr Cys Thr Gly Asp Ser Ser Gln  
 405 410 415  
 Cys Pro Pro Asn Leu His Lys Leu Asp Gly Tyr Tyr Cys Asp His Glu  
 420 425 430  
 Gln Gly Arg Cys Tyr Gly Gly Arg Cys Lys Thr Arg Asp Arg Gln Cys  
 435 440 445  
 Gln Val Leu Trp Gly His Ala Ala Ala Asp Arg Phe Cys Tyr Glu Lys

10

20

30

40

450                      455                      460  
Leu Asn Val Glu Gly Thr Glu Arg Gly Ser Cys Gly Arg Lys Gly Ser  
465                      470                      475                      480  
Gly Trp Val Gln Cys Ser Lys Gln  
                    485                      488 。

【 0 0 5 0 】

配列番号：2

10

配列の長さ：5 2 4

配列の型：アミノ酸

配列の種類：タンパク質

トポロジー：直鎖状

起源

生物名：ホモサピエンス

直接の起源

ライブラリー名：ヒト胎児脳 c D N A ライブラリー

## 配列

Met Cys Trp Leu Ser His Gln Leu Leu Ser Ser Gln Tyr Val Glu Arg  
 1 5 10 15  
 His Phe Ser Arg Glu Gly Thr Thr Gln His Ser Thr Gly Ala Gly Asp  
 20 25 30  
 His Cys Tyr Tyr Gln Gly Lys Leu Arg Gly Asn Pro His Ser Phe Ala  
 35 40 45  
 Ala Leu Ser Thr Cys Gln Gly Leu His Gly Val Phe Ser Asp Gly Asn  
 50 55 60  
 Leu Thr Tyr Ile Val Glu Pro Gln Glu Val Ala Gly Pro Trp Gly Ala  
 65 70 75 80  
 Pro Gln Gly Pro Leu Pro His Leu Ile Tyr Arg Thr Pro Leu Leu Pro  
 85 90 95  
 Asp Pro Leu Gly Cys Arg Glu Pro Gly Cys Leu Phe Ala Val Pro Ala  
 100 105 110  
 Gln Ser Ala Pro Pro Asn Arg Pro Arg Leu Arg Arg Lys Arg Gln Val  
 115 120 125  
 Arg Arg Gly His Pro Thr Val His Ser Glu Thr Lys Tyr Val Glu Leu  
 130 135 140  
 Ile Val Ile Asn Asp His Gln Leu Phe Glu Gln Met Arg Gln Ser Val  
 145 150 155 160  
 Val Leu Thr Ser Asn Phe Ala Lys Ser Val Val Asn Leu Ala Asp Val  
 165 170 175  
 Ile Tyr Lys Glu Gln Leu Asn Thr Arg Ile Val Leu Val Ala Met Glu  
 180 185 190  
 Thr Trp Ala Asp Gly Asp Lys Ile Gln Val Gln Asp Asp Leu Leu Glu  
 195 200 205  
 Thr Leu Ala Arg Leu Met Val Tyr Arg Arg Glu Gly Leu Pro Glu Pro  
 210 215 220

10

20

30

40

Ser Asn Ala Thr His Leu Phe Ser Gly Arg Thr Phe Gln Ser Thr Ser  
 225 230 235 240  
 Ser Gly Ala Ala Tyr Val Gly Gly Ile Cys Ser Leu Ser His Gly Gly  
 245 250 255  
 Gly Val Asn Glu Tyr Gly Asn Met Gly Ala Met Ala Val Thr Leu Ala  
 260 265 270  
 Gln Thr Leu Gly Gln Asn Leu Gly Met Met Trp Asn Lys His Arg Ser  
 275 280 285  
 Ser Ala Gly Asp Cys Lys Cys Pro Asp Ile Trp Leu Gly Cys Ile Met  
 290 295 300  
 Glu Asp Thr Gly Phe Tyr Leu Pro Arg Lys Phe Ser Arg Cys Ser Ile  
 305 310 315 320  
 Asp Glu Tyr Asn Gln Phe Leu Gln Glu Gly Gly Gly Ser Cys Leu Phe  
 325 330 335  
 Asn Lys Pro Leu Lys Leu Leu Asp Pro Pro Glu Cys Gly Asn Gly Phe  
 340 345 350  
 Val Glu Ala Gly Glu Glu Cys Asp Cys Gly Ser Val Gln Glu Cys Ser  
 355 360 365  
 Arg Ala Gly Gly Asn Cys Cys Lys Lys Cys Thr Leu Thr His Asp Ala  
 370 375 380  
 Met Cys Ser Asp Gly Leu Cys Cys Arg Arg Cys Lys Tyr Glu Pro Arg  
 385 390 395 400  
 Gly Val Ser Cys Arg Glu Ala Val Asn Glu Cys Asp Ile Ala Glu Thr  
 405 410 415  
 Cys Thr Gly Asp Ser Ser Gln Cys Pro Pro Asn Leu His Lys Leu Asp  
 420 425 430  
 Gly Tyr Tyr Cys Asp His Glu Gln Gly Arg Cys Tyr Gly Gly Arg Cys  
 435 440 445  
 Lys Thr Arg Asp Arg Gln Cys Gln Val Leu Trp Gly His Ala Ala Ala

10

20

30

40

450                      455                      460  
Asp Arg Phe Cys Tyr Glu Lys Leu Asn Val Glu Gly Thr Glu Arg Gly  
465                      470                      475                      480  
Ser Cys Gly Arg Lys Gly Ser Gly Trp Val Gln Cys Ser Lys Gln Pro  
                         485                      490                      495  
Gln Gln Gly Arg Ala Val Trp Leu Pro Pro Leu Cys Gln His Leu Trp  
                         500                      505                      510  
Ser Ser Ser Ala Arg Gly Pro Gly Gly Arg His Gln  
                         515                      520                      524 .

10

【 0 0 5 1 】

配列番号 : 3

配列の長さ : 6 7 0

配列の型 : アミノ酸

配列の種類 : タンパク質

トポロジー : 直鎖状

20

起源

生物名 : ホモサピエンス

直接の起源

ライブラリー名 : ヒト胎児脳 c D N A ライブラリー

## 配列

Met Cys Trp Leu Ser His Gln Leu Leu Ser Ser Gln Tyr Val Glu Arg  
 1 5 10 15  
 His Phe Ser Arg Glu Gly Thr Thr Gln His Ser Thr Gly Ala Gly Asp  
 20 25 30  
 His Cys Tyr Tyr Gln Gly Lys Leu Arg Gly Asn Pro His Ser Phe Ala  
 35 40 45  
 Ala Leu Ser Thr Cys Gln Gly Leu His Gly Val Phe Ser Asp Gly Asn  
 50 55 60  
 Leu Thr Tyr Ile Val Glu Pro Gln Glu Val Ala Gly Pro Trp Gly Ala  
 65 70 75 80  
 Pro Gln Gly Pro Leu Pro His Leu Ile Tyr Arg Thr Pro Leu Leu Pro  
 85 90 95  
 Asp Pro Leu Gly Cys Arg Glu Pro Gly Cys Leu Phe Ala Val Pro Ala  
 100 105 110  
 Gln Ser Ala Pro Pro Asn Arg Pro Arg Leu Arg Arg Lys Arg Gln Val  
 115 120 125  
 Arg Arg Gly His Pro Thr Val His Ser Glu Thr Lys Tyr Val Glu Leu  
 130 135 140  
 Ile Val Ile Asn Asp His Gln Leu Phe Glu Gln Met Arg Gln Ser Val  
 145 150 155 160  
 Val Leu Thr Ser Asn Phe Ala Lys Ser Val Val Asn Leu Ala Asp Val  
 165 170 175  
 Ile Tyr Lys Glu Gln Leu Asn Thr Arg Ile Val Leu Val Ala Met Glu  
 180 185 190  
 Thr Trp Ala Asp Gly Asp Lys Ile Gln Val Gln Asp Asp Leu Leu Glu  
 195 200 205  
 Thr Leu Ala Arg Leu Met Val Tyr Arg Arg Glu Gly Leu Pro Glu Pro  
 210 215 220

10

20

30

40

Ser Asn Ala Thr His Leu Phe Ser Gly Arg Thr Phe Gln Ser Thr Ser  
 225 230 235 240  
 Ser Gly Ala Ala Tyr Val Gly Gly Ile Cys Ser Leu Ser His Gly Gly  
 245 250 255  
 Gly Val Asn Glu Tyr Gly Asn Met Gly Ala Met Ala Val Thr Leu Ala  
 260 265 270  
 Gln Thr Leu Gly Gln Asn Leu Gly Met Met Trp Asn Lys His Arg Ser  
 275 280 285  
 Ser Ala Gly Asp Cys Lys Cys Pro Asp Ile Trp Leu Gly Cys Ile Met  
 290 295 300  
 Glu Asp Thr Gly Phe Tyr Leu Pro Arg Lys Phe Ser Arg Cys Ser Ile  
 305 310 315 320  
 Asp Glu Tyr Asn Gln Phe Leu Gln Glu Gly Gly Gly Ser Cys Leu Phe  
 325 330 335  
 Asn Lys Pro Leu Lys Leu Leu Asp Pro Pro Glu Cys Gly Asn Gly Phe  
 340 345 350  
 Val Glu Ala Gly Glu Glu Cys Asp Cys Gly Ser Val Gln Glu Cys Ser  
 355 360 365  
 Arg Ala Gly Gly Asn Cys Cys Lys Lys Cys Thr Leu Thr His Asp Ala  
 370 375 380  
 Met Cys Ser Asp Gly Leu Cys Cys Arg Arg Cys Lys Tyr Glu Pro Arg  
 385 390 395 400  
 Gly Val Ser Cys Arg Glu Ala Val Asn Glu Cys Asp Ile Ala Glu Thr  
 405 410 415  
 Cys Thr Gly Asp Ser Ser Gln Cys Pro Pro Asn Leu His Lys Leu Asp  
 420 425 430  
 Gly Tyr Tyr Cys Asp His Glu Gln Gly Arg Cys Tyr Gly Gly Arg Cys  
 435 440 445  
 Lys Thr Arg Asp Arg Gln Cys Gln Val Leu Trp Gly His Ala Ala Ala

10

20

30

40

450                      455                      460  
 Asp Arg Phe Cys Tyr Glu Lys Leu Asn Val Glu Gly Thr Glu Arg Gly  
 465                      470                      475                      480  
 Ser Cys Gly Arg Lys Gly Ser Gly Trp Val Gln Cys Ser Lys Gln Asp  
                     485                      490                      495  
 Val Leu Cys Gly Phe Leu Leu Cys Val Asn Ile Ser Gly Ala Pro Arg  
                     500                      505                      510  
 Leu Gly Asp Leu Val Gly Asp Ile Ser Ser Val Thr Phe Tyr His Gln  
                     515                      520                      525  
 Gly Lys Glu Leu Asp Cys Arg Gly Gly His Val Gln Leu Ala Asp Gly  
                     530                      535                      540  
 Ser Asp Leu Ser Tyr Val Glu Asp Gly Thr Ala Cys Gly Pro Asn Met  
 545                      550                      555                      560  
 Leu Cys Leu Asp His Arg Cys Leu Pro Ala Ser Ala Phe Asn Phe Ser  
                     565                      570                      575  
 Thr Cys Pro Gly Ser Gly Glu Arg Arg Ile Cys Ser His His Gly Val  
                     580                      585                      590  
 Cys Ser Asn Glu Gly Lys Cys Ile Cys Gln Pro Asp Trp Thr Gly Lys  
                     595                      600                      605  
 Asp Cys Ser Ile His Asn Pro Leu Pro Thr Ser Pro Pro Thr Gly Glu  
                     610                      615                      620  
 Thr Glu Arg Tyr Lys Gly Pro Ser Gly Thr Asn Ile Ile Ile Gly Ser  
 625                      630                      635                      640  
 Ile Ala Gly Ala Val Leu Val Ala Ala Ile Val Leu Gly Gly Thr Gly  
                     645                      650                      655  
 Trp Gly Phe Lys Asn Ile Arg Arg Gly Arg Ser Gly Gly Ala  
                     660                      665                      670 。

【 0 0 5 2 】

配列番号： 4

配列の長さ： 7 6 9

配列の型：アミノ酸

配列の種類：タンパク質

トポロジー：直鎖状

起源

生物名：ホモサピエンス

直接の起源

ライブラリー名：ヒト胎児脳 c D N A ライブラリー

10

20

30

40

50



## 配列

Met Arg Leu Leu Arg Arg Trp Ala Phe Ala Ala Leu Leu Leu Ser Leu  
 1 5 10 15  
 Leu Pro Thr Pro Gly Leu Gly Thr Gln Gly Pro Ala Gly Ala Leu Arg  
 20 25 30  
 Trp Gly Gly Leu Pro Gln Leu Gly Gly Pro Gly Ala Pro Glu Val Thr  
 35 40 45  
 Glu Pro Ser Arg Leu Val Arg Glu Ser Ser Gly Gly Glu Val Arg Lys  
 50 55 60  
 Gln Gln Leu Asp Thr Arg Val Arg Gln Glu Pro Pro Gly Gly Pro Pro  
 65 70 75 80  
 Val His Leu Ala Gln Val Ser Phe Val Ile Pro Ala Phe Asn Ser Asn  
 85 90 95  
 Phe Thr Leu Asp Leu Glu Leu Asn His His Leu Leu Ser Ser Gln Tyr  
 100 105 110  
 Val Glu Arg His Phe Ser Arg Glu Gly Thr Thr Gln His Ser Thr Gly  
 115 120 125  
 Ala Gly Asp His Cys Tyr Tyr Gln Gly Lys Leu Arg Gly Asn Pro His  
 130 135 140  
 Ser Phe Ala Ala Leu Ser Thr Cys Gln Gly Leu His Gly Val Phe Ser  
 145 150 155 160  
 Asp Gly Asn Leu Thr Tyr Ile Val Glu Pro Gln Glu Val Ala Gly Pro  
 165 170 175  
 Trp Gly Ala Pro Gln Gly Pro Leu Pro His Leu Ile Tyr Arg Thr Pro  
 180 185 190  
 Leu Leu Pro Asp Pro Leu Gly Cys Arg Glu Pro Gly Cys Leu Phe Ala  
 195 200 205  
 Val Pro Ala Gln Ser Ala Pro Pro Asn Arg Pro Arg Leu Arg Arg Lys  
 210 215 220

10

20

30

40

Arg Gln Val Arg Arg Gly His Pro Thr Val His Ser Glu Thr Lys Tyr  
 225 230 235 240  
 Val Glu Leu Ile Val Ile Asn Asp His Gln Leu Phe Glu Gln Met Arg  
 245 250 255  
 Gln Ser Val Val Leu Thr Ser Asn Phe Ala Lys Ser Val Val Asn Leu  
 260 265 270  
 Ala Asp Val Ile Tyr Lys Glu Gln Leu Asn Thr Arg Ile Val Leu Val  
 275 280 285  
 Ala Met Glu Thr Trp Ala Asp Gly Asp Lys Ile Gln Val Gln Asp Asp  
 290 295 300  
 Leu Leu Glu Thr Leu Ala Arg Leu Met Val Tyr Arg Arg Glu Gly Leu  
 305 310 315 320  
 Pro Glu Pro Ser Asn Ala Thr His Leu Phe Ser Gly Arg Thr Phe Gln  
 325 330 335  
 Ser Thr Ser Ser Gly Ala Ala Tyr Val Gly Gly Ile Cys Ser Leu Ser  
 340 345 350  
 His Gly Gly Gly Val Asn Glu Tyr Gly Asn Met Gly Ala Met Ala Val  
 355 360 365  
 Thr Leu Ala Gln Thr Leu Gly Gln Asn Leu Gly Met Met Trp Asn Lys  
 370 375 380  
 His Arg Ser Ser Ala Gly Asp Cys Lys Cys Pro Asp Ile Trp Leu Gly  
 385 390 395 400  
 Cys Ile Met Glu Asp Thr Gly Phe Tyr Leu Pro Arg Lys Phe Ser Arg  
 405 410 415  
 Cys Ser Ile Asp Glu Tyr Asn Gln Phe Leu Gln Glu Gly Gly Gly Ser  
 420 425 430  
 Cys Leu Phe Asn Lys Pro Leu Lys Leu Leu Asp Pro Pro Glu Cys Gly  
 435 440 445  
 Asn Gly Phe Val Glu Ala Gly Glu Glu Cys Asp Cys Gly Ser Val Gln

10

20

30

40

450                      455                      460  
 Glu Cys Ser Arg Ala Gly Gly Asn Cys Cys Lys Lys Cys Thr Leu Thr  
 465                      470                      475                      480  
 His Asp Ala Met Cys Ser Asp Gly Leu Cys Cys Arg Arg Cys Lys Tyr  
                     485                      490                      495  
 Glu Pro Arg Gly Val Ser Cys Arg Glu Ala Val Asn Glu Cys Asp Ile  
                     500                      505                      510  
 Ala Glu Thr Cys Thr Gly Asp Ser Ser Gln Cys Pro Pro Asn Leu His  
                     515                      520                      525  
 Lys Leu Asp Gly Tyr Tyr Cys Asp His Glu Gln Gly Arg Cys Tyr Gly  
                     530                      535                      540  
 Gly Arg Cys Lys Thr Arg Asp Arg Gln Cys Gln Val Leu Trp Gly His  
 545                      550                      555                      560  
 Ala Ala Ala Asp Arg Phe Cys Tyr Glu Lys Leu Asn Val Glu Gly Thr  
                     565                      570                      575  
 Glu Arg Gly Ser Cys Gly Arg Lys Gly Ser Gly Trp Val Gln Cys Ser  
                     580                      585                      590  
 Lys Gln Asp Val Leu Cys Gly Phe Leu Leu Cys Val Asn Ile Ser Gly  
                     595                      600                      605  
 Ala Pro Arg Leu Gly Asp Leu Val Gly Asp Ile Ser Ser Val Thr Phe  
                     610                      615                      620  
 Tyr His Gln Gly Lys Glu Leu Asp Cys Arg Gly Gly His Val Gln Leu  
 625                      630                      635                      640  
 Ala Asp Gly Ser Asp Leu Ser Tyr Val Glu Asp Gly Thr Ala Cys Gly  
                     645                      650                      655  
 Pro Asn Met Leu Cys Leu Asp His Arg Cys Leu Pro Ala Ser Ala Phe  
                     660                      665                      670  
 Asn Phe Ser Thr Cys Pro Gly Ser Gly Glu Arg Arg Ile Cys Ser His  
                     675                      680                      685

10

20

30

40

His Gly Val Cys Ser Asn Glu Gly Lys Cys Ile Cys Gln Pro Asp Trp  
 690 695 700  
 Thr Gly Lys Asp Cys Ser Ile His Asn Pro Leu Pro Thr Ser Pro Pro  
 705 710 715 720  
 Thr Gly Glu Thr Glu Arg Tyr Lys Gly Pro Ser Gly Thr Asn Ile Ile  
 725 730 735  
 Ile Gly Ser Ile Ala Gly Ala Val Leu Val Ala Ala Ile Val Leu Gly  
 740 745 750  
 Gly Thr Gly Trp Gly Phe Lys Asn Ile Arg Arg Gly Arg Ser Gly Gly  
 755 760 765

Ala

769。

【 0 0 5 3 】

配列番号：5

配列の長さ：1 4 6 4

配列の型：核酸

鎖の数：二本鎖

トポロジー：直鎖状

配列の種類：cDNA to mRNA

起源

生物名：ホモサピエンス

直接の起源

ライブラリー名：ヒト胎児脳 cDNA ライブラリー

配列の特徴

特徴を表す記号：C D S

存在位置：1 . . 1 4 6 4

特徴を決定した方法：E

10

20

30

## 配列

CTC CTC TCC TCG CAA TAC GTG GAG CGC CAC TTC AGC CGG GAG GGG ACA	48	
Leu Leu Ser Ser Gln Tyr Val Glu Arg His Phe Ser Arg Glu Gly Thr		
1 5 10 15		
ACC CAG CAC AGC ACC GGG GCT GGA GAC CAC TGC TAC TAC CAG GGG AAG	96	
Thr Gln His Ser Thr Gly Ala Gly Asp His Cys Tyr Tyr Gln Gly Lys		
20 25 30		10
CTC CGG GGG AAC CCG CAC TCC TTC GCC GCC CTC TCC ACC TGC CAG GGG	144	
Leu Arg Gly Asn Pro His Ser Phe Ala Ala Leu Ser Thr Cys Gln Gly		
35 40 45		
CTG CAT GGG GTC TTC TCT GAT GGG AAC TTG ACT TAC ATC GTG GAG CCC	192	
Leu His Gly Val Phe Ser Asp Gly Asn Leu Thr Tyr Ile Val Glu Pro		
50 55 60		
CAA GAG GTG GCT GGA CCT TGG GGA GCC CCT CAG GGA CCC CTT CCC CAC	240	20
Gln Glu Val Ala Gly Pro Trp Gly Ala Pro Gln Gly Pro Leu Pro His		
65 70 75 80		
CTC ATT TAC CGG ACC CCT CTC CTC CCA GAT CCC CTC GGA TGC AGG GAA	288	
Leu Ile Tyr Arg Thr Pro Leu Leu Pro Asp Pro Leu Gly Cys Arg Glu		
85 90 95		
CCA GGC TGC CTG TTT GCT GTG CCT GCC CAG TCG GCT CCT CCA AAC CGG	336	
Pro Gly Cys Leu Phe Ala Val Pro Ala Gln Ser Ala Pro Pro Asn Arg		30
100 105 110		
CCG AGG CTG AGA AGG AAA AGG CAG GTC CGC CGG GGC CAC CCT ACA GTG	384	
Pro Arg Leu Arg Arg Lys Arg Gln Val Arg Arg Gly His Pro Thr Val		
115 120 125		
CAC AGT GAA ACC AAG TAT GTG GAG CTA ATT GTG ATC AAC GAC CAC CAG	432	
His Ser Glu Thr Lys Tyr Val Glu Leu Ile Val Ile Asn Asp His Gln		
130 135 140		40
CTG TTC GAG CAG ATG CGA CAG TCG GTG GTC CTC ACC AGC AAC TTT GCC	480	

Leu Phe Glu Gln Met Arg Gln Ser Val Val Leu Thr Ser Asn Phe Ala		
145	150	155
AAG TCC GTG GTG AAC CTG GCC GAT GTG ATA TAC AAG GAG CAG CTC AAC	528	
Lys Ser Val Val Asn Leu Ala Asp Val Ile Tyr Lys Glu Gln Leu Asn		
165	170	175
ACT CGC ATC GTC CTG GTT GCC ATG GAA ACA TGG GCA GAT GGG GAC AAG	576	
Thr Arg Ile Val Leu Val Ala Met Glu Thr Trp Ala Asp Gly Asp Lys		10
180	185	190
ATC CAG GTG CAG GAT GAC CTC CTG GAG ACC CTG GCC CGG CTC ATG GTC	624	
Ile Gln Val Gln Asp Asp Leu Leu Glu Thr Leu Ala Arg Leu Met Val		
195	200	205
TAC CGA CGG GAG GGT CTG CCT GAG CCC AGT AAT GCC ACC CAC CTC TTC	672	
Tyr Arg Arg Glu Gly Leu Pro Glu Pro Ser Asn Ala Thr His Leu Phe		
210	215	220
TCG GGC AGG ACC TTC CAG AGC ACG AGC AGC GGG GCA GCC TAC GTG GGG	720	
Ser Gly Arg Thr Phe Gln Ser Thr Ser Ser Gly Ala Ala Tyr Val Gly		
225	230	235
GGC ATA TGC TCC CTG TCC CAT GGC GGG GGT GTG AAC GAG TAC GGC AAC	768	
Gly Ile Cys Ser Leu Ser His Gly Gly Gly Val Asn Glu Tyr Gly Asn		
245	250	255
ATG GGG GCG ATG GCC GTG ACC CTT GCC CAG ACG CTG GGA CAG AAC CTG	816	30
Met Gly Ala Met Ala Val Thr Leu Ala Gln Thr Leu Gly Gln Asn Leu		
260	265	270
GGC ATG ATG TGG AAC AAA CAC CGG AGC TCG GCA GGG GAC TGC AAG TGT	864	
Gly Met Met Trp Asn Lys His Arg Ser Ser Ala Gly Asp Cys Lys Cys		
275	280	285
CCA GAC ATC TGG CTG GGC TGC ATC ATG GAG GAC ACT GGG TTC TAC CTG	912	
Pro Asp Ile Trp Leu Gly Cys Ile Met Glu Asp Thr Gly Phe Tyr Leu		40
290	295	300

CCC CGC AAG TTC TCT CGC TGC AGC ATC GAC GAG TAC AAC CAG TTT CTG	960	
Pro Arg Lys Phe Ser Arg Cys Ser Ile Asp Glu Tyr Asn Gln Phe Leu		
305 310 315 320		
CAG GAG GGT GGT GGC AGC TGC CTC TTC AAC AAG CCC CTC AAG CTC CTG	1008	
Gln Glu Gly Gly Gly Ser Cys Leu Phe Asn Lys Pro Leu Lys Leu Leu		
325 330 335		
GAC CCC CCA GAG TGC GGG AAC GGC TTC GTG GAG GCA GGG GAG GAG TGC	1056	10
Asp Pro Pro Glu Cys Gly Asn Gly Phe Val Glu Ala Gly Glu Glu Cys		
340 345 350		
GAC TGC GGC TCG GTG CAG GAG TGC AGC CGC GCA GGT GGC AAC TGC TGC	1104	
Asp Cys Gly Ser Val Gln Glu Cys Ser Arg Ala Gly Gly Asn Cys Cys		
355 360 365		
AAG AAA TGC ACC CTG ACT CAC GAC GCC ATG TGC AGC GAC GGG CTC TGC	1152	
Lys Lys Cys Thr Leu Thr His Asp Ala Met Cys Ser Asp Gly Leu Cys		20
370 375 380		
TGT CGC CGC TGC AAG TAC GAA CCA CGG GGT GTG TCC TGC CGA GAG GCC	1200	
Cys Arg Arg Cys Lys Tyr Glu Pro Arg Gly Val Ser Cys Arg Glu Ala		
385 390 395 400		
GTG AAC GAG TGC GAC ATC GCG GAG ACC TGC ACC GGG GAC TCT AGC CAG	1248	
Val Asn Glu Cys Asp Ile Ala Glu Thr Cys Thr Gly Asp Ser Ser Gln		
405 410 415		30
TGC CCG CCT AAC CTG CAC AAG CTG GAC GGT TAC TAC TGT GAC CAT GAG	1296	
Cys Pro Pro Asn Leu His Lys Leu Asp Gly Tyr Tyr Cys Asp His Glu		
420 425 430		
CAG GGC CGC TGC TAC GGA GGT CGC TGC AAA ACC CGG GAC CGG CAG TGC	1344	
Gln Gly Arg Cys Tyr Gly Gly Arg Cys Lys Thr Arg Asp Arg Gln Cys		
435 440 445		
CAG GTT CTT TGG GGC CAT GCG GCT GCT GAT CGC TTC TGC TAC GAG AAG	1392	40
Gln Val Leu Trp Gly His Ala Ala Ala Asp Arg Phe Cys Tyr Glu Lys		

450	455	460	
CTG AAT GTG GAG GGG ACG GAG CGT GGG AGC TGT GGG CGC AAG GGA TCC			1440
Leu Asn Val Glu Gly Thr Glu Arg Gly Ser Cys Gly Arg Lys Gly Ser			
465	470	475	480
GGC TGG GTC CAG TGC AGT AAG CAG			1464
Gly Trp Val Gln Cys Ser Lys Gln			
485	。		

10

【 0 0 5 4 】

配列番号：6

配列の長さ：2 9 2 3

配列の型：核酸

鎖の数：二本鎖

トポロジー：直鎖状

配列の種類：cDNA to mRNA

起源

生物名：ホモサピエンス

20

直接の起源

ライブラリー名：ヒト胎児脳 c D N A ライブラリー

配列の特徴

特徴を表す記号：5 ' U T R

存在位置：1 . . 2 7

特徴を決定した方法：E

特徴を表す記号：3 ' U T R

存在位置：1 6 0 0 . . 2 9 2 3

特徴を決定した方法：E

特徴を表す記号：C D S

30

存在位置：2 8 . . 1 5 9 9

特徴を決定した方法：E



## 配列

GCGTTTACTG GCAAACCGCA TTTGTAA ATG TGC TGG CTG AGC CAC CAA CTC	51	
Met Cys Trp Leu Ser His Gln Leu		
1 5		
CTC TCC TCG CAA TAC GTG GAG CGC CAC TTC AGC CGG GAG GGG ACA ACC	99	
Leu Ser Ser Gln Tyr Val Glu Arg His Phe Ser Arg Glu Gly Thr Thr		
10 15 20		10
CAG CAC AGC ACC GGG GCT GGA GAC CAC TGC TAC TAC CAG GGG AAG CTC	147	
Gln His Ser Thr Gly Ala Gly Asp His Cys Tyr Tyr Gln Gly Lys Leu		
25 30 35 40		
CGG GGG AAC CCG CAC TCC TTC GCC GCC CTC TCC ACC TGC CAG GGG CTG	195	
Arg Gly Asn Pro His Ser Phe Ala Ala Leu Ser Thr Cys Gln Gly Leu		
45 50 55		
CAT GGG GTC TTC TCT GAT GGG AAC TTG ACT TAC ATC GTG GAG CCC CAA	243	20
His Gly Val Phe Ser Asp Gly Asn Leu Thr Tyr Ile Val Glu Pro Gln		
60 65 70		
GAG GTG GCT GGA CCT TGG GGA GCC CCT CAG GGA CCC CTT CCC CAC CTC	291	
Glu Val Ala Gly Pro Trp Gly Ala Pro Gln Gly Pro Leu Pro His Leu		
75 80 85		
ATT TAC CGG ACC CCT CTC CTC CCA GAT CCC CTC GGA TGC AGG GAA CCA	339	
Ile Tyr Arg Thr Pro Leu Leu Pro Asp Pro Leu Gly Cys Arg Glu Pro		30
90 95 100		
GGC TGC CTG TTT GCT GTG CCT GCC CAG TCG GCT CCT CCA AAC CGG CCG	387	
Gly Cys Leu Phe Ala Val Pro Ala Gln Ser Ala Pro Pro Asn Arg Pro		
105 110 115 120		
AGG CTG AGA AGG AAA AGG CAG GTC CGC CGG GGC CAC CCT ACA GTG CAC	435	
Arg Leu Arg Arg Lys Arg Gln Val Arg Arg Gly His Pro Thr Val His		
125 130 135		40
AGT GAA ACC AAG TAT GTG GAG CTA ATT GTG ATC AAC GAC CAC CAG CTG	483	

Ser	Glu	Thr	Lys	Tyr	Val	Glu	Leu	Ile	Val	Ile	Asn	Asp	His	Gln	Leu		
			140					145					150				
TTC	GAG	CAG	ATG	CGA	CAG	TCG	GTG	GTC	CTC	ACC	AGC	AAC	TTT	GCC	AAG	531	
Phe	Glu	Gln	Met	Arg	Gln	Ser	Val	Val	Leu	Thr	Ser	Asn	Phe	Ala	Lys		
			155					160					165				
TCC	GTG	GTG	AAC	CTG	GCC	GAT	GTG	ATA	TAC	AAG	GAG	CAG	CTC	AAC	ACT	579	
Ser	Val	Val	Asn	Leu	Ala	Asp	Val	Ile	Tyr	Lys	Glu	Gln	Leu	Asn	Thr		10
			170					175					180				
CGC	ATC	GTC	CTG	GTT	GCC	ATG	GAA	ACA	TGG	GCA	GAT	GGG	GAC	AAG	ATC	627	
Arg	Ile	Val	Leu	Val	Ala	Met	Glu	Thr	Trp	Ala	Asp	Gly	Asp	Lys	Ile		
185					190					195					200		
CAG	GTG	CAG	GAT	GAC	CTC	CTG	GAG	ACC	CTG	GCC	CGG	CTC	ATG	GTC	TAC	675	
Gln	Val	Gln	Asp	Asp	Leu	Leu	Glu	Thr	Leu	Ala	Arg	Leu	Met	Val	Tyr		
				205						210					215		20
CGA	CGG	GAG	GGT	CTG	CCT	GAG	CCC	AGT	AAT	GCC	ACC	CAC	CTC	TTC	TCG	723	
Arg	Arg	Glu	Gly	Leu	Pro	Glu	Pro	Ser	Asn	Ala	Thr	His	Leu	Phe	Ser		
			220						225						230		
GGC	AGG	ACC	TTC	CAG	AGC	ACG	AGC	AGC	GGG	GCA	GCC	TAC	GTG	GGG	GGC	771	
Gly	Arg	Thr	Phe	Gln	Ser	Thr	Ser	Ser	Gly	Ala	Ala	Tyr	Val	Gly	Gly		
			235						240						245		
ATA	TGC	TCC	CTG	TCC	CAT	GGC	GGG	GGT	GTG	AAC	GAG	TAC	GGC	AAC	ATG	819	30
Ile	Cys	Ser	Leu	Ser	His	Gly	Gly	Gly	Val	Asn	Glu	Tyr	Gly	Asn	Met		
			250						255						260		
GGG	GCG	ATG	GCC	GTG	ACC	CTT	GCC	CAG	ACG	CTG	GGA	CAG	AAC	CTG	GGC	867	
Gly	Ala	Met	Ala	Val	Thr	Leu	Ala	Gln	Thr	Leu	Gly	Gln	Asn	Leu	Gly		
265					270										280		
ATG	ATG	TGG	AAC	AAA	CAC	CGG	AGC	TCG	GCA	GGG	GAC	TGC	AAG	TGT	CCA	915	
Met	Met	Trp	Asn	Lys	His	Arg	Ser	Ser	Ala	Gly	Asp	Cys	Lys	Cys	Pro		40
					285										295		
									290								

GAC ATC TGG CTG GGC TGC ATC ATG GAG GAC ACT GGG TTC TAC CTG CCC	963	
Asp Ile Trp Leu Gly Cys Ile Met Glu Asp Thr Gly Phe Tyr Leu Pro		
300 305 310		
CGC AAG TTC TCT CGC TGC AGC ATC GAC GAG TAC AAC CAG TTT CTG CAG	1011	
Arg Lys Phe Ser Arg Cys Ser Ile Asp Glu Tyr Asn Gln Phe Leu Gln		
315 320 325		
GAG GGT GGT GGC AGC TGC CTC TTC AAC AAG CCC CTC AAG CTC CTG GAC	1059	10
Glu Gly Gly Gly Ser Cys Leu Phe Asn Lys Pro Leu Lys Leu Leu Asp		
330 335 340		
CCC CCA GAG TGC GGG AAC GGC TTC GTG GAG GCA GGG GAG GAG TGC GAC	1107	
Pro Pro Glu Cys Gly Asn Gly Phe Val Glu Ala Gly Glu Glu Cys Asp		
345 350 355 360		
TGC GGC TCG GTG CAG GAG TGC AGC CGC GCA GGT GGC AAC TGC TGC AAG	1155	
Cys Gly Ser Val Gln Glu Cys Ser Arg Ala Gly Gly Asn Cys Cys Lys		20
365 370 375		
AAA TGC ACC CTG ACT CAC GAC GCC ATG TGC AGC GAC GGG CTC TGC TGT	1203	
Lys Cys Thr Leu Thr His Asp Ala Met Cys Ser Asp Gly Leu Cys Cys		
380 385 390		
CGC CGC TGC AAG TAC GAA CCA CGG GGT GTG TCc TGC CGA GAG GCC GTG	1251	
Arg Arg Cys Lys Tyr Glu Pro Arg Gly Val Ser Cys Arg Glu Ala Val		
395 400 405		30
AAC GAG TGC GAC ATC GCG GAG ACC TGC ACC GGG GAC TCT AGC CAG TGC	1299	
Asn Glu Cys Asp Ile Ala Glu Thr Cys Thr Gly Asp Ser Ser Gln Cys		
410 415 420		
CCG CCT AAC CTG CAC AAG CTG GAC GGT TAC TAC TGT GAC CAT GAG CAG	1347	
Pro Pro Asn Leu His Lys Leu Asp Gly Tyr Tyr Cys Asp His Glu Gln		
425 430 435 440		
GGC CGC TGC TAC GGA GGT CGC TGC AAA ACC CGG GAC CGG CAG TGC CAG	1395	40
Gly Arg Cys Tyr Gly Gly Arg Cys Lys Thr Arg Asp Arg Gln Cys Gln		

445	450	455		
GTT CTT TGG GGC CAT GCG GCT GCT GAT CGC TTC TGC TAC GAG AAG CTG	1443			
Val Leu Trp Gly His Ala Ala Ala Asp Arg Phe Cys Tyr Glu Lys Leu				
460	465	470		
AAT GTG GAG GGG ACG GAG CGT GGG AGC TGT GGG CGC AAG GGA TCC GGC	1491			
Asn Val Glu Gly Thr Glu Arg Gly Ser Cys Gly Arg Lys Gly Ser Gly				
475	480	485		10
TGG GTC CAG TGC AGT AAG CAG CCC CAA CAG GGA CGT GCT GTG TGG CTT	1539			
Trp Val Gln Cys Ser Lys Gln Pro Gln Gln Gly Arg Ala Val Trp Leu				
490	495	500		
CCT CCT CTG TGT CAA CAT CTC TGG AGC TCC TCG GCT AGG GGA CCT GGT	1587			
Pro Pro Leu Cys Gln His Leu Trp Ser Ser Ser Ala Arg Gly Pro Gly				
505	510	515	520	
GGG AGA CAT CAG TAGTGTCACC TTCTACCACC AGGGCAAGGA GCTGGACTGC	1639			20
Gly Arg His Gln				
AGGGGAGGCC ACGTGCAGCT GCGGACGGC TCTGACCTGA GCTATGTGGA GGATGGCACA	1699			
GCCTGCGGGC CTAACATGTT GTGCCTGGAC CATCGCTGCC TGCCAGCTTC TGCCTTCAAC	1759			
TTCAGCACCT GCCCGGCAG TGGGGAGCGC CGGATTTGCT CCCACCACGG GGTCTGCAGC	1819			
AATGAAGGGA AGTGCATCTG TCAGCCAGAC TGGACAGGCA AAGACTGCAG TATCCATAAC	1879			
CCCCTGCCCA CGTCCCCACC CACGGGGGAG ACGGAGAGAT ATAAAGGTCC CAGCGGCACC	1939			
AACATCATCA TTGGCTCCAT CGCTGGGGCT GTCCTGGTTG CAGCCATCGT CCTGGGCGGC	1999			30
ACGGGCTGGG GATTTAAAAA CATTGCGCGA GGAAGGTCCG GAGGGGCCTA AGTGCCACCC	2059			
TCCTCCCTCC AAGCCTGGCA CCCACCGTCT CGGCCCTGAA CCACGAGGCT GCCCCCATCC	2119			
AGCCACGGAG GGAGGCACCA TGCAAATGTC TTCCAGGTCC AAACCCTTCA ACTCCTGGCT	2179			
CCGCAGGGGT TTGGGTGGGG GCTGTGGCCC TGCCCTTGGC ACCACCAGGG TGGACCAGGC	2239			
CTGGAGGGCA CTTCTCCAC AGTCCCCAC CCACCTCCTG CGGCTCAGCC TTGCACACCC	2299			
ACTGCCCCGT GTGAATGTAG CTTCCACCTC ATGGATTGCC ACAGCTCAAC TCGGGGGCAC	2359			
CTGGAGGGAT GCCCCAGGC AGCCACCAGT GGACCTAGCC TGGATGGCCC CTCCTTGCAA	2419			40
CCAGGCAGCT GAGACCAGGG TCTTATCTCT CTGGGACCTA GGGGGACGGG GCTGACATCT	2479			

ACATTTTTTA AAACGAATC TTAATCGATG AATGTAAACT CGGGGGTGCT GGGGCCAGGG 2539  
 CAGATGTGGG GATGTTTTGA CATTTACAGG AGGCCCCGGA GAAACTGAGG TATGGCCATG 2599  
 CCCTAGACCC TCCCCAAGGA TGACCACACC CGAAGTCCTG TCACTGAGCA CAGTCAGGGG 2659  
 CTGGGCATCC CAGCTTGCCC CCGCTTAGCC CCGCTGAGCT TGGAGGAAGT ATGAGTGCTG 2719  
 ATTCAAACCA AAGCTGCCTG TGCCATGCCC AAGGCCTAGG TTATGGGTAC GGCAACCACA 2779  
 TGTCCCAGAT CGTCTCCAAT TCGAAAACAA CCGTCCTGCT GTCCCTGTCA GGACACATGG 2839  
 ATTTTGGCAG GGCAGGGGGG GGTTCCTAGAA AATATAGGTT CCTATAATAA AATGGCACCT 2899  
 TCCCCCTTTA AAAAAAAAAA AAAA 2923

10

•

【 0 0 5 5 】

配列番号：7

配列の長さ：2 9 1 3

配列の型：核酸

鎖の数：二本鎖

トポロジー：直鎖状

配列の種類：cDNA to mRNA

起源

生物名：ホモサピエンス

直接の起源

ライブラリー名：ヒト胎児脳 cDNA ライブラリー

配列の特徴

特徴を表す記号：5' UTR

存在位置：1 . . 2 7

特徴を決定した方法：E

特徴を表す記号：3' UTR

存在位置：2 0 3 8 . . 2 9 1 3

特徴を決定した方法：E

特徴を表す記号：CDS

存在位置：2 8 . . 2 0 3 7

特徴を決定した方法：E

20

30

## 配列

GCGTTTACTG GCAAACCGCA TTTGTAA ATG TGC TGG CTG AGC CAC CAA CTC	51	
Met Cys Trp Leu Ser His Gln Leu		
1 5		
CTC TCC TCG CAA TAC GTG GAG CGC CAC TTC AGC CGG GAG GGG ACA ACC	99	
Leu Ser Ser Gln Tyr Val Glu Arg His Phe Ser Arg Glu Gly Thr Thr		
10 15 20		10
CAG CAC AGC ACC GGG GCT GGA GAC CAC TGC TAC TAC CAG GGG AAG CTC	147	
Gln His Ser Thr Gly Ala Gly Asp His Cys Tyr Tyr Gln Gly Lys Leu		
25 30 35 40		
CGG GGG AAC CCG CAC TCC TTC GCC GCC CTC TCC ACC TGC CAG GGG CTG	195	
Arg Gly Asn Pro His Ser Phe Ala Ala Leu Ser Thr Cys Gln Gly Leu		
45 50 55		
CAT GGG GTC TTC TCT GAT GGG AAC TTG ACT TAC ATC GTG GAG CCC CAA	243	20
His Gly Val Phe Ser Asp Gly Asn Leu Thr Tyr Ile Val Glu Pro Gln		
60 65 70		
GAG GTG GCT GGA CCT TGG GGA GCC CCT CAG GGA CCC CTT CCC CAC CTC	291	
Glu Val Ala Gly Pro Trp Gly Ala Pro Gln Gly Pro Leu Pro His Leu		
75 80 85		
ATT TAC CGG ACC CCT CTC CTC CCA GAT CCC CTC GGA TGC AGG GAA CCA	339	
Ile Tyr Arg Thr Pro Leu Leu Pro Asp Pro Leu Gly Cys Arg Glu Pro		30
90 95 100		
GGC TGC CTG TTT GCT GTG CCT GCC CAG TCG GCT CCT CCA AAC CGG CCG	387	
Gly Cys Leu Phe Ala Val Pro Ala Gln Ser Ala Pro Pro Asn Arg Pro		
105 110 115 120		
AGG CTG AGA AGG AAA AGG CAG GTC CGC CGG GGC CAC CCT ACA GTG CAC	435	
Arg Leu Arg Arg Lys Arg Gln Val Arg Arg Gly His Pro Thr Val His		
125 130 135		40
AGT GAA ACC AAG TAT GTG GAG CTA ATT GTG ATC AAC GAC CAC CAG CTG	483	

Ser	Glu	Thr	Lys	Tyr	Val	Glu	Leu	Ile	Val	Ile	Asn	Asp	His	Gln	Leu		
			140					145					150				
TTC	GAG	CAG	ATG	CGA	CAG	TCG	GTG	GTC	CTC	ACC	AGC	AAC	TTT	GCC	AAG	531	
Phe	Glu	Gln	Met	Arg	Gln	Ser	Val	Val	Leu	Thr	Ser	Asn	Phe	Ala	Lys		
			155					160					165				
TCC	GTG	GTG	AAC	CTG	GCC	GAT	GTG	ATA	TAC	AAG	GAG	CAG	CTC	AAC	ACT	579	
Ser	Val	Val	Asn	Leu	Ala	Asp	Val	Ile	Tyr	Lys	Glu	Gln	Leu	Asn	Thr		10
			170					175					180				
CGC	ATC	GTC	CTG	GTT	GCC	ATG	GAA	ACA	TGG	GCA	GAT	GGG	GAC	AAG	ATC	627	
Arg	Ile	Val	Leu	Val	Ala	Met	Glu	Thr	Trp	Ala	Asp	Gly	Asp	Lys	Ile		
185					190					195					200		
CAG	GTG	CAG	GAT	GAC	CTC	CTG	GAG	ACC	CTG	GCC	CGG	CTC	ATG	GTC	TAC	675	
Gln	Val	Gln	Asp	Asp	Leu	Leu	Glu	Thr	Leu	Ala	Arg	Leu	Met	Val	Tyr		
				205						210					215		20
CGA	CGG	GAG	GGT	CTG	CCT	GAG	CCC	AGT	AAT	GCC	ACC	CAC	CTC	TTC	TCG	723	
Arg	Arg	Glu	Gly	Leu	Pro	Glu	Pro	Ser	Asn	Ala	Thr	His	Leu	Phe	Ser		
			220						225						230		
GGC	AGG	ACC	TTC	CAG	AGC	ACG	AGC	AGC	GGG	GCA	GCC	TAC	GTG	GGG	GGC	771	
Gly	Arg	Thr	Phe	Gln	Ser	Thr	Ser	Ser	Gly	Ala	Ala	Tyr	Val	Gly	Gly		
			235						240						245		
ATA	TGC	TCC	CTG	TCC	CAT	GGC	GGG	GGT	GTG	AAC	GAG	TAC	GGC	AAC	ATG	819	30
Ile	Cys	Ser	Leu	Ser	His	Gly	Gly	Gly	Val	Asn	Glu	Tyr	Gly	Asn	Met		
			250					255							260		
GGG	GCG	ATG	GCC	GTG	ACC	CTT	GCC	CAG	ACG	CTG	GGA	CAG	AAC	CTG	GGC	867	
Gly	Ala	Met	Ala	Val	Thr	Leu	Ala	Gln	Thr	Leu	Gly	Gln	Asn	Leu	Gly		
265					270					275					280		
ATG	ATG	TGG	AAC	AAA	CAC	CGG	AGC	TCG	GCA	GGG	GAC	TGC	AAG	TGT	CCA	915	
Met	Met	Trp	Asn	Lys	His	Arg	Ser	Ser	Ala	Gly	Asp	Cys	Lys	Cys	Pro		40
				285						290					295		

GAC ATC TGG CTG GGC TGC ATC ATG GAG GAC ACT GGG TTC TAC CTG CCC	963	
Asp Ile Trp Leu Gly Cys Ile Met Glu Asp Thr Gly Phe Tyr Leu Pro		
300 305 310		
CGC AAG TTC TCT CGC TGC AGC ATC GAC GAG TAC AAC CAG TTT CTG CAG	1011	
Arg Lys Phe Ser Arg Cys Ser Ile Asp Glu Tyr Asn Gln Phe Leu Gln		
315 320 325		
GAG GGT GGT GGC AGC TGC CTC TTC AAC AAG CCC CTC AAG CTC CTG GAC	1059	10
Glu Gly Gly Gly Ser Cys Leu Phe Asn Lys Pro Leu Lys Leu Leu Asp		
330 335 340		
CCC CCA GAG TGC GGG AAC GGC TTC GTG GAG GCA GGG GAG GAG TGC GAC	1107	
Pro Pro Glu Cys Gly Asn Gly Phe Val Glu Ala Gly Glu Glu Cys Asp		
345 350 355 360		
TGC GGC TCG GTG CAG GAG TGC AGC CGC GCA GGT GGC AAC TGC TGC AAG	1155	
Cys Gly Ser Val Gln Glu Cys Ser Arg Ala Gly Gly Asn Cys Cys Lys		20
365 370 375		
AAA TGC ACC CTG ACT CAC GAC GCC ATG TGC AGC GAC GGG CTC TGC TGT	1203	
Lys Cys Thr Leu Thr His Asp Ala Met Cys Ser Asp Gly Leu Cys Cys		
380 385 390		
CGC CGC TGC AAG TAC GAA CCA CGG GGT GTG TCc TGC CGA GAG GCC GTG	1251	
Arg Arg Cys Lys Tyr Glu Pro Arg Gly Val Ser Cys Arg Glu Ala Val		
395 400 405		30
AAC GAG TGC GAC ATC GCG GAG ACC TGC ACC GGG GAC TCT AGC CAG TGC	1299	
Asn Glu Cys Asp Ile Ala Glu Thr Cys Thr Gly Asp Ser Ser Gln Cys		
410 415 420		
CCG CCT AAC CTG CAC AAG CTG GAC GGT TAC TAC TGT GAC CAT GAG CAG	1347	
Pro Pro Asn Leu His Lys Leu Asp Gly Tyr Tyr Cys Asp His Glu Gln		
425 430 435 440		
GGC CGC TGC TAC GGA GGT CGC TGC AAA ACC CGG GAC CGG CAG TGC CAG	1395	40
Gly Arg Cys Tyr Gly Gly Arg Cys Lys Thr Arg Asp Arg Gln Cys Gln		



445	450	455		
GTT CTT TGG GGC CAT GCG GCT GCT GAT CGC TTC TGC TAC GAG AAG CTG	1443			
Val Leu Trp Gly His Ala Ala Ala Asp Arg Phe Cys Tyr Glu Lys Leu				
460	465	470		
AAT GTG GAG GGG ACG GAG CGT GGG AGC TGT GGG CGC AAG GGA TCC GGC	1491			
Asn Val Glu Gly Thr Glu Arg Gly Ser Cys Gly Arg Lys Gly Ser Gly				
475	480	485		10
TGG GTC CAG TGC AGT AAG CAG GAC GTG CTG TGT GGC TTC CTC CTC TGT	1539			
Trp Val Gln Cys Ser Lys Gln Asp Val Leu Cys Gly Phe Leu Leu Cys				
490	495	500		
GTC AAC ATC TCT GGA GCT CCT CGG CTA GGG GAC CTG GTG GGA GAC ATC	1587			
Val Asn Ile Ser Gly Ala Pro Arg Leu Gly Asp Leu Val Gly Asp Ile				
505	510	515	520	
AGT AGT GTC ACC TTC TAC CAC CAG GGC AAG GAG CTG GAC TGC AGG GGA	1635			20
Ser Ser Val Thr Phe Tyr His Gln Gly Lys Glu Leu Asp Cys Arg Gly				
525	530	535		
GGC CAC GTG CAG CTG GCG GAC GGC TCT GAC CTG AGC TAT GTG GAG GAT	1683			
Gly His Val Gln Leu Ala Asp Gly Ser Asp Leu Ser Tyr Val Glu Asp				
540	545	550		
GGC ACA GCC TGC GGG CCT AAC ATG TTG TGC CTG GAC CAT CGC TGC CTG	1731			
Gly Thr Ala Cys Gly Pro Asn Met Leu Cys Leu Asp His Arg Cys Leu				30
555	560	565		
CCA GCT TCT GCC TTC AAC TTC AGC ACC TGC CCC GGC AGT GGG GAG CGC	1779			
Pro Ala Ser Ala Phe Asn Phe Ser Thr Cys Pro Gly Ser Gly Glu Arg				
570	575	580		
CGG ATT TGC TCC CAC CAC GGG GTC TGC AGC AAT GAA GGG AAG TGC ATC	1827			
Arg Ile Cys Ser His His Gly Val Cys Ser Asn Glu Gly Lys Cys Ile				
585	590	595	600	40
TGT CAG CCA GAC TGG ACA GGC AAA GAC TGC AGT ATC CAT AAC CCC CTG	1875			

Cys Gln Pro Asp Trp Thr Gly Lys Asp Cys Ser Ile His Asn Pro Leu		
605	610	615
CCC ACG TCC CCA CCC ACG GGG GAG ACG GAG AGA TAT AAA GGT CCC AGC	1923	
Pro Thr Ser Pro Pro Thr Gly Glu Thr Glu Arg Tyr Lys Gly Pro Ser		
620	625	630
GGC ACC AAC ATC ATC ATT GGC TCC ATC GCT GGG GCT GTC CTG GTT GCA	1971	
Gly Thr Asn Ile Ile Ile Gly Ser Ile Ala Gly Ala Val Leu Val Ala		10
635	640	645
GCC TAC GTC CTG GGC GGC ACG GGC TGG GGA TTT AAA AAC ATT CGC CGA	2019	
Ala Ile Val Leu Gly Gly Thr Gly Trp Gly Phe Lys Asn Ile Arg Arg		
650	655	660
GGA AGG TCC GGA GGG GCC TAAGTGCCAC CCTCCTCCCT CCAAGCCTGG	2067	
Gly Arg Ser Gly Gly Ala		
665	670	
CACCCACCGT CTCGGCCCTG AACCACGAGG CTGCCCCCAT CCAGCCACGG AGGGAGGCAC	2127	
CATGCAAATG TCTTCCAGGT CCAAACCCTT CAACTCCTGG CTCCGCAGGG GTTTGGGTGG	2187	
GGGCTGTGGC CCTGCCCTTG GCACCACCAG GGTGGACCAG GCCTGGAGGG CACTTCCTCC	2247	
ACAGTCCCCC ACCCACCTCC TCGGGCTCAG CCTTGACACAC CCACTGCCCC GTGTGAATGT	2307	
AGCTTCCACC TCATGGATTG CCACAGCTCA ACTCGGGGGC ACCTGGAGGG ATGCCCCCAG	2367	
GCAGCCACCA GTGGACCTAG CCTGGATGGC CCCTCCTTGC AACCAGGCAG CTGAGACCAG	2427	
GGTCTTATCT CTCTGGGACC TAGGGGGACG GGGCTGACAT CTACATTTTT TAAAACTGAA	2487	30
TCTTAATCGA TGAATGTAAA CTCGGGGGTG CTGGGGCCAG GGCAGATGTG GGGATGTTTT	2547	
GACATTTACA GGAGGCCCCG GAGAACTGA GGTATGGCCA TGCCCTAGAC CCTCCCCAAG	2607	
GATGACCACA CCCGAAGTCC TGTCACCTGAG CACAGTCAGG GGCTGGGCAT CCCAGCTTGC	2667	
CCCCGCTTAG CCCCCTGAG CTTGGAGGAA GTATGAGTGC TGATTCAAAC CAAAGCTGCC	2727	
TGTGCCATGC CCAAGGCCTA GGTATGGGT ACGGCAACCA CATGTCCCAG ATCGTCTCCA	2787	
ATTCGAAAAC AACCGTCCTG CTGTCCCTGT CAGGACACAT GGATTTTGGC AGGGCGGGGG	2847	
GGGGTTCTAG AAAATATAGG TTCCTATAAT AAAATGGCAC CTTCCCCCTT TAAAAA	2907	40
AAAAA	2913	

•

【 0 0 5 6 】

配列番号： 8

配列の長さ： 3 1 8 3

配列の型： 核酸

鎖の数： 二本鎖

トポロジー：直鎖状

配列の種類：cDNA to mRNA

起源

生物名：ホモサピエンス

直接の起源

ライブラリー名：ヒト脳 cDNA ライブラリー

配列の特徴

特徴を表す記号：3' UTR

存在位置：2308...3183

特徴を決定した方法：E

特徴を表す記号：CDS

存在位置：1...2307

特徴を決定した方法：E

## 配列

ATG AGG CTG CTG CGG CGC TGG GCG TTC GCG GCT CTG CTG CTG TCG CTG	48	
Met Arg Leu Leu Arg Arg Trp Ala Phe Ala Ala Leu Leu Leu Ser Leu		
1 5 10 15		
CTC CCC ACG CCC GGT CTT GGG ACC CAA GGT cct GCT GGA GCT CTG Cga	96	
Leu Pro Thr Pro Gly Leu Gly Thr Gln Gly Pro Ala Gly Ala Leu Arg		
20 25 30		10
TGG GGG GGC TTA CCC CAG CTG GGA GGC CCA GGA GCC CCT GAG GTC ACG	144	
Trp Gly Gly Leu Pro Gln Leu Gly Gly Pro Gly Ala Pro Glu Val Thr		
35 40 45		
GAA CCC AGC CGT CTG GTT AGG GAG AGC TCC GGG GGA GAG GTC CGA AAG	192	
Glu Pro Ser Arg Leu Val Arg Glu Ser Ser Gly Gly Glu Val Arg Lys		
50 55 60		
CAG CAG CTG GAC ACA AGG GTC CGC CAG GAG CCA CCA GGG GGC CCG CCT	240	20
Gln Gln Leu Asp Thr Arg Val Arg Gln Glu Pro Pro Gly Gly Pro Pro		
65 70 75 80		
GTC CAT CTG GCC CAG GTG AGT TTC GTC ATC CCA GCC TTC AAC TCA AAC	288	
Val His Leu Ala Gln Val Ser Phe Val Ile Pro Ala Phe Asn Ser Asn		
85 90 95		
TTC ACC CTG GAC CTG GAG CTG AAC CAC CAC CTC CTC TCC TCG CAA TAC	336	
Phe Thr Leu Asp Leu Glu Leu Asn His His Leu Leu Ser Ser Gln Tyr		30
100 105 110		
GTG GAG CGC CAC TTC AGC CGG GAG GGG ACA ACC CAG CAC AGC ACC GGG	384	
Val Glu Arg His Phe Ser Arg Glu Gly Thr Thr Gln His Ser Thr Gly		
115 120 125		
GCT GGA GAC CAC TGC TAC TAC CAG GGG AAG CTC CGG GGG AAC CCG CAC	432	
Ala Gly Asp His Cys Tyr Tyr Gln Gly Lys Leu Arg Gly Asn Pro His		
130 135 140		40
TCC TTC GCC GCC CTC TCC ACC TGC CAG GGG CTG CAT GGG GTC TTC TCT	480	

Ser Phe Ala Ala Leu Ser Thr Cys Gln Gly Leu His Gly Val Phe Ser		
145	150	155
GAT GGG AAC TTG ACT TAC ATC GTG GAG CCC CAA GAG GTG GCT GGA CCT	528	
Asp Gly Asn Leu Thr Tyr Ile Val Glu Pro Gln Glu Val Ala Gly Pro		
165	170	175
TGG GGA GCC CCT CAG GGA CCC CTT CCC CAC CTC ATT TAC CGG ACC CCT	576	
Trp Gly Ala Pro Gln Gly Pro Leu Pro His Leu Ile Tyr Arg Thr Pro		10
180	185	190
CTC CTC CCA GAT CCC CTC GGA TGC AGG GAA CCA GGC TGC CTG TTT GCT	624	
Leu Leu Pro Asp Pro Leu Gly Cys Arg Glu Pro Gly Cys Leu Phe Ala		
195	200	205
GTG CCT GCC CAG TCG GCT CCT CCA AAC CGG CCG AGG CTG AGA AGG AAA	672	
Val Pro Ala Gln Ser Ala Pro Pro Asn Arg Pro Arg Leu Arg Arg Lys		
210	215	220
AGG CAG GTC CGC CGG GGC CAC CCT ACA GTG CAC AGT GAA ACC AAG TAT	720	
Arg Gln Val Arg Arg Gly His Pro Thr Val His Ser Glu Thr Lys Tyr		
225	230	235
GTG GAG CTA ATT GTG ATC AAC GAC CAC CAG CTG TTC GAG CAG ATG CGA	768	
Val Glu Leu Ile Val Ile Asn Asp His Gln Leu Phe Glu Gln Met Arg		
245	250	255
CAG TCG GTG GTC CTC ACC AGC AAC TTT GCC AAG TCC GTG GTG AAC CTG	816	30
Gln Ser Val Val Leu Thr Ser Asn Phe Ala Lys Ser Val Val Asn Leu		
260	265	270
GCC GAT GTG ATA TAC AAG GAG CAG CTC AAC ACT CGC ATC GTC CTG GTT	864	
Ala Asp Val Ile Tyr Lys Glu Gln Leu Asn Thr Arg Ile Val Leu Val		
275	280	285
GCC ATG GAA ACA TGG GCA GAT GGG GAC AAG ATC CAG GTG CAG GAT GAC	912	
Ala Met Glu Thr Trp Ala Asp Gly Asp Lys Ile Gln Val Gln Asp Asp		40
290	295	300

CTC CTG GAG ACC CTG GCC CGG CTC ATG GTC TAC CGA CGG GAG GGT CTG	960	
Leu Leu Glu Thr Leu Ala Arg Leu Met Val Tyr Arg Arg Glu Gly Leu		
305 310 315 320		
CCT GAG CCC AGT AAT GCC ACC CAC CTC TTC TCG GGC AGG ACC TTC CAG	1008	
Pro Glu Pro Ser Asn Ala Thr His Leu Phe Ser Gly Arg Thr Phe Gln		
325 330 335		
AGC ACG AGC AGC GGG GCA GCC TAC GTG GGG GGC ATA TGC TCC CTG TCC	1056	10
Ser Thr Ser Ser Gly Ala Ala Tyr Val Gly Gly Ile Cys Ser Leu Ser		
340 345 350		
CAT GGC GGG GGT GTG AAC GAG TAC GGC AAC ATG GGG GCG ATG GCC GTG	1104	
His Gly Gly Gly Val Asn Glu Tyr Gly Asn Met Gly Ala Met Ala Val		
355 360 365		
ACC CTT GCC CAG ACG CTG GGA CAG AAC CTG GGC ATG ATG TGG AAC AAA	1152	
Thr Leu Ala Gln Thr Leu Gly Gln Asn Leu Gly Met Met Trp Asn Lys		20
370 375 380		
CAC CGG AGC TCG GCA GGG GAC TGC AAG TGT CCA GAC ATC TGG CTG GGC	1200	
His Arg Ser Ser Ala Gly Asp Cys Lys Cys Pro Asp Ile Trp Leu Gly		
385 390 395 400		
TGC ATC ATG GAG GAC ACT GGG TTC TAC CTG CCC CGC AAG TTC TCT CGC	1248	
Cys Ile Met Glu Asp Thr Gly Phe Tyr Leu Pro Arg Lys Phe Ser Arg		
405 410 415		30
TGC AGC ATC GAC GAG TAC AAC CAG TTT CTG CAG GAG GGT GGT GGC AGC	1296	
Cys Ser Ile Asp Glu Tyr Asn Gln Phe Leu Gln Glu Gly Gly Gly Ser		
420 425 430		
TGC CTC TTC AAC AAG CCC CTC AAG CTC CTG GAC CCC CCA GAG TGC GGG	1344	
Cys Leu Phe Asn Lys Pro Leu Lys Leu Leu Asp Pro Pro Glu Cys Gly		
435 440 445		
AAC GGC TTC GTG GAG GCA GGG GAG GAG TGC GAC TGC GGC TCG GTG CAG	1392	40
Asn Gly Phe Val Glu Ala Gly Glu Glu Cys Asp Cys Gly Ser Val Gln		

450	455	460		
GAG TGC AGC CGC GCA GGT GGC AAC TGC TGC AAG AAA TGC ACC CTG ACT			1440	
Glu Cys Ser Arg Ala Gly Gly Asn Cys Cys Lys Lys Cys Thr Leu Thr				
465	470	475	480	
CAC GAC GCC ATG TGC AGC GAC GGG CTC TGC TGT CGC CGC TGC AAG TAC			1488	
His Asp Ala Met Cys Ser Asp Gly Leu Cys Cys Arg Arg Cys Lys Tyr				
485	490	495		10
GAA CCA CGG GGT GTG TCC TGC CGA GAG GCC GTG AAC GAG TGC GAC ATC			1536	
Glu Pro Arg Gly Val Ser Cys Arg Glu Ala Val Asn Glu Cys Asp Ile				
500	505	510		
GCG GAG ACC TGC ACC GGG GAC TCT AGC CAG TGC CCG CCT AAC CTG CAC			1584	
Ala Glu Thr Cys Thr Gly Asp Ser Ser Gln Cys Pro Pro Asn Leu His				
515	520	525		20
AAG CTG GAC GGT TAC TAC TGT GAC CAT GAG CAG GGC CGC TGC TAC GGA			1632	
Lys Leu Asp Gly Tyr Tyr Cys Asp His Glu Gln Gly Arg Cys Tyr Gly				
530	535	540		
GGT CGC TGC AAA ACC CGG GAC CGG CAG TGC CAG GTT CTT TGG GGC CAT			1680	
Gly Arg Cys Lys Thr Arg Asp Arg Gln Cys Gln Val Leu Trp Gly His				
545	550	555	560	
GCG GCT GCT GAT CGC TTC TGC TAC GAG AAG CTG AAT GTG GAG GGG ACG			1728	
Ala Ala Ala Asp Arg Phe Cys Tyr Glu Lys Leu Asn Val Glu Gly Thr				30
565	570	575		
GAG CGT GGG AGC TGT GGG CGC AAG GGA TCC GGC TGG GTC CAG TGC AGT			1776	
Glu Arg Gly Ser Cys Gly Arg Lys Gly Ser Gly Trp Val Gln Cys Ser				
580	585	590		
AAG CAG GAC GTG CTG TGT GGC TTC CTC CTC TGT GTC AAC ATC TCT GGA			1824	
Lys Gln Asp Val Leu Cys Gly Phe Leu Leu Cys Val Asn Ile Ser Gly				
595	600	605		40
GCT CCT CGG CTA GGG GAC CTG GTG GGA GAC ATC AGT AGT GTC ACC TTC			1872	

Ala Pro Arg Leu Gly Asp Leu Val Gly Asp Ile Ser Ser Val Thr Phe		
610	615	620
TAC CAC CAG GGC AAG GAG CTG GAC TGC AGG GGA GGC CAC GTG CAG CTG	1920	
Tyr His Gln Gly Lys Glu Leu Asp Cys Arg Gly Gly His Val Gln Leu		
625	630	635
GCG GAC GGC TCT GAC CTG AGC TAT GTG GAG GAT GGC ACA GCC TGC GGG	1968	
Ala Asp Gly Ser Asp Leu Ser Tyr Val Glu Asp Gly Thr Ala Cys Gly		10
645	650	655
CCT AAC ATG TTG TGC CTG GAC CAT CGC TGC CTG CCA GCT TCT GCC TTC	2016	
Pro Asn Met Leu Cys Leu Asp His Arg Cys Leu Pro Ala Ser Ala Phe		
660	665	670
AAC TTC AGC ACC TGC CCC GGC AGT GGG GAG CGC CGG ATT TGC TCC CAC	2064	
Asn Phe Ser Thr Cys Pro Gly Ser Gly Glu Arg Arg Ile Cys Ser His		
675	680	685
CAC GGG GTC TGC AGC AAT GAA GGG AAG TGC ATC TGT CAG CCA GAC TGG	2112	
His Gly Val Cys Ser Asn Glu Gly Lys Cys Ile Cys Gln Pro Asp Trp		20
690	695	700
ACA GGC AAA GAC TGC AGT ATC CAT AAC CCC CTG CCC ACG TCC CCA CCC	2160	
Thr Gly Lys Asp Cys Ser Ile His Asn Pro Leu Pro Thr Ser Pro Pro		
705	710	715
ACG GGG GAG ACG GAG AGA TAT AAA GGT CCC AGC GGC ACC AAC ATC ATC	2208	
Thr Gly Glu Thr Glu Arg Tyr Lys Gly Pro Ser Gly Thr Asn Ile Ile		30
725	730	735
ATT GGC TCC ATC GCT GGG GCT GTC CTG GTT GCA GCC ATC GTC CTG GGC	2256	
Ile Gly Ser Ile Ala Gly Ala Val Leu Val Ala Ala Ile Val Leu Gly		
740	745	750
GGC ACG GGC TGG GGA TTT AAA AAC ATT CGC CGA GGA AGG TCC GGA GGC	2304	
Gly Thr Gly Trp Gly Phe Lys Asn Ile Arg Arg Gly Arg Ser Gly Gly		40
755	760	765



GCC TAAGTGCCAC CCTCCTCCCT CCAAGCCTGG CACCCACCGT CTCGGCCCTG 2357

Ala

AACCACGAGG CTGCCCCCAT CCAGCCACGG AGGGAGGCAC CATGCAAATG TCTTCCAGGT 2417

CCAAACCCTT CAACTCCTGG CTCCGCAGGG GTTTGGGTGG GGGCTGTGGC CCTGCCCTTG 2477

GCACCACCAG GGTGGACCAG GCCTGGAGGG CACTTCCTCC ACAGTCCCCC ACCCACCTCC 2537

TGCGGCTCAG CTTTGACAC CCACTGCCCC GTGTGAATGT AGCTTCCACC TCATGGATTG 2597

CCACAGCTCA ACTCGGGGGC ACCTGGAGGG ATGCCCCCAG GCAGCCACCA GTGGACCTAG 2657

10

CCTGGATGGC CCCTCCTTGC AACCAGGCAG CTGAGACCAG GGTCTTATCT CTCTGGGACC 2717

TAGGGGGACG GGGCTGACAT CTACATTTTT TAAAACTGAA TCTTAATCGA TGAATGTAAA 2777

CTCGGGGGTG CTGGGGCCAG GGCAGATGTG GGGATGTTTT GACATTTACA GGAGGCCCCG 2837

GAGAACTGA GGTATGGCCA TGCCCTAGAC CCTCCCCAAG GATGACCACA CCCGAAGTCC 2897

TGTCACTGAG CACAGTCAGG GGCTGGGCAT CCCAGCTTGC CCCCCTTAG CCCCCTGAG 2957

CTTGGAGGAA GTATGAGTGC TGATTCAAAC CAAAGCTGCC TGTGCCATGC CCAAGGCCTA 3017

GGTTATGGGT ACGGCAACCA CATGTCCCAG ATCGTCTCCA ATTCGAAAAC AACCGTCTG 3077

20

CTGTCCCTGT CAGGACACAT GGATTTTGGC AGGGCGGGGG GGGGTTCTAG AAAATATAGG 3137

TTCCTATAAT AAAATGGCAC CTTCCCCCTT TAAAAA AAAA 3183

•

【 0 0 5 7 】

配列番号 : 9

配列の長さ : 9 2 7 8

配列の型 : 核酸

鎖の数 : 二本鎖

30

トポロジー : 直鎖状

配列の種類 : Genomic DNA

起源

生物名 : ホモサピエンス

直接の起源

ライブラリー名 : ヒトDNAコスミドライブラリー

配列の特徴

特徴を表す記号 : exon 1

存在位置 : 2 8 . . 4 4

特徴を決定した方法 : E

40

特徴を表す記号 : exon 2

存在位置 : 3 0 8 . . 3 7 4

特徴を決定した方法 : E

特徴を表す記号 : exon 3

存在位置 : 9 0 9 . . 9 9 4

特徴を決定した方法 : E

特徴を表す記号 : exon 4

存在位置 : 1 0 8 1 . . 1 1 5 6

特徴を決定した方法 : E

特徴を表す記号 : exon 5

50

存在位置：1 5 9 1 . . 1 6 5 7  
特徴を決定した方法：E  
特徴を表す記号：exon 6  
存在位置：1 7 2 5 . . 1 7 9 2  
特徴を決定した方法：E  
特徴を表す記号：exon 7  
存在位置：2 1 8 2 . . 2 2 5 6  
特徴を決定した方法：E  
特徴を表す記号：exon 8  
存在位置：2 3 3 9 . . 2 4 1 0 10  
特徴を決定した方法：E  
特徴を表す記号：exon 9  
存在位置：2 5 8 8 . . 2 7 5 4  
特徴を決定した方法：E  
特徴を表す記号：exon 10  
存在位置：3 2 4 8 . . 3 3 3 2  
特徴を決定した方法：E  
特徴を表す記号：exon 11  
存在位置：3 4 4 5 . . 3 5 3 5 20  
特徴を決定した方法：E  
特徴を表す記号：exon 12  
存在位置：3 6 4 5 . . 3 6 9 6  
特徴を決定した方法：E  
特徴を表す記号：exon 13  
存在位置：4 0 1 4 . . 4 1 1 3  
特徴を決定した方法：E  
特徴を表す記号：exon 14  
存在位置：4 1 9 6 . . 4 2 6 7  
特徴を決定した方法：E  
特徴を表す記号：exon 15 30  
存在位置：4 3 8 6 . . 4 4 7 8  
特徴を決定した方法：E  
特徴を表す記号：exon 16  
存在位置：4 9 2 0 . . 5 0 0 0  
特徴を決定した方法：E  
特徴を表す記号：exon 17  
存在位置：5 3 4 7 . . 5 3 9 7  
特徴を決定した方法：E  
特徴を表す記号：exon 18  
存在位置：5 5 0 1 . . 5 5 6 4 40  
特徴を決定した方法：E  
特徴を表す記号：exon 19  
存在位置：5 7 6 7 . . 5 8 6 6  
特徴を決定した方法：E  
特徴を表す記号：exon 20  
存在位置：6 0 7 3 . . 6 2 0 2  
特徴を決定した方法：E  
特徴を表す記号：exon 21  
存在位置：6 3 0 0 . . 6 4 6 8  
特徴を決定した方法：E 50

特徴を表す記号 : exon 22

存在位置 : 6 5 5 7 . . 6 6 7 1

特徴を決定した方法 : E

特徴を表す記号 : exon 23

存在位置 : 6 7 5 6 . . 6 8 4 6

特徴を決定した方法 : E

特徴を表す記号 : exon 24

存在位置 : 7 8 2 9 . . 7 8 4 6

特徴を決定した方法 : E

特徴を表す記号 : exon 25

存在位置 : 8 1 6 5 . . 9 0 3 8

特徴を決定した方法 : E

## 配列

GCGTTTACTG GCAAACCGCA TTTGTAA ATG TGC TGG CTG AGC CA NNNNNNNNNN	54	
Met Cys Trp Leu Ser His		
1 5		
NNNNCCAGGT GAGTTTCGTC ATCCAGCCTT CAACTCAAAC TTCACCCTGG ACCTGGAGCT	114	
GAACCAGTGA GNGTGGCCTT GAGCCCAAGA GGAAGGGCAG TGGTGGNNNG GGGGAGACAT	174	
GGCTAGGGCC TGGCTGCTGG GGGTCTGGGG GTTGGGCCTG GCGAGAGGGG ACCTGGGTCC	234	10
TGACCTGAGG CGAGCCTAAA GCCCACCTC ACCTCGCCCG TGACCCCCCT TCCTGCTGCC	294	
CCCTCTGTCT CAG C CAA CTC CTC TCC TCG CAA TAC GTG GAG CGC CAC TTC	344	
Gln Leu Leu Ser Ser Gln Tyr Val Glu Arg His Phe		
10 15		
AGC CGG GAG GGG ACA ACC CAG CAC AGC ACC GTGAGTGCCA CTGCTGGGGA	394	
Ser Arg Glu Gly Thr Thr Gln His Ser Thr		
20 25		20
CCGGGGCCGG GGATGGAAGG GAGGTGCTGT TTCTGTGGTT CTGTGGTCAC AGGTGTAGGG	454	
ACAGGTGGCC ACTGGAGATG GGGTCCTGGG CCTGGCCCCCT CAGCACCTTC CCTCTCTCCC	514	
GACCCAGGAG GCTCTGAGGG TGGACAGTGG GCAGCTTAGT GCATAGGGCC CTGAAGTCCC	574	
CTCACTTGGC CCCAGAGCTC TGACCCCCAG CCAGCCCACG TGGGGCCTAC AGGGACACTC	634	
GTTCCGAGCA GGCTGCCAGG ATCCNNNNNN NNNNNNATAG ATGACGTGAA GGAGGCCCAG	694	
AGGTTCTTAA CCCCAGAGGG CTAGGAACTT GCCCAGGGTG GCACGGCAAA TTAGGAGCAC	754	
CAGCCATCTA GAAACAGGCT CCAGAGCCCC AGGNATACCC AGGGATNGTG GCCACCTGCA	814	30
CACAGGGCAG CTTCAGTGTC CCCC AAAAAG CCTTGAGGCC CATTGGCTGC CCCC GGCTC	874	
ATGCCAGCGT TCTGCTCACT GTTCTGCTCC TTAG GGG GCT GGA GAC CAC TGC TAC	929	
Gly Ala Gly Asp His Cys Tyr		
30 35		
TAC CAG GGG AAG CTC CGG GGG AAC CCG CAC TCC TTC GCC GCC CTC TCC	977	
Tyr Gln Gly Lys Leu Arg Gly Asn Pro His Ser Phe Ala Ala Leu Ser		
40 45 50		40
ACC TGC CAG GGG CTG CA GTGAGTATGG GGAGGGGCCG GGCAGCTGGG	1024	

Thr Cys Gln Gly Leu His

55

AGAAGCCTCT GGCCCAGGCC TGGGGACGGA GGGGAGCTGC GCCTCTCTCT CCACAG T 1081

GGG GTC TTC TCT GAT GGG AAC TTG ACT TAC ATC GTG GAG CCC CAA GAG 1129

Gly Val Phe Ser Asp Gly Asn Leu Thr Tyr Ile Val Glu Pro Gln Glu

60

65

70

GTG GCT GGA CCT TGG GGA GCC CCT CAG GTAAGCCCCA CACAACCCCT 1176

Val Ala Gly Pro Trp Gly Ala Pro Gln

75

80

TGCCATCCTC TCTGGTGGCC CTGCCAAGCT TGTCCCAACA GCTGTTGCTG CCACCTCTTC 1236

CTCCTCCGGC TCCTCCCTCA GTAACCCCAG CCTCACTGCC CTCTTCAGTG ACCCCAGCTC 1396

TGGTTCCCTC CCTCCTGTGC CCCAGCTCCC CCTGTGCCCC CAGCTCCAAT GTCCCATCTG 1356

TCCCATAAGT GACCTCCCAT TGGGCTCCAA TGTCCTTTGC CCCTGTCTCT CAGGGTGCCC 1416

CCAGGTCTTG ACCCGGAAT CTGAGCATCT GGGAGATCAG ATCCGACATG GGAGCTGTGG 1476

CCAGTTCTGG GTCACCCCAG GGTGGGGTGG AGGCGAGGGC TGGATCTGGC CCCCGCCAAG 1536

TGGCCTGGAG CAGGCCAGT TGGCACCCCA AGAACTAATT TCCCCTCATT GCAG GGA 1593

Gly

CCC CTT CCC CAC CTC ATT TAC CGG ACC CCT CTC CTC CCA GAT CCC CTC 1641

Pro Leu Pro His Leu Ile Tyr Arg Thr Pro Leu Leu Pro Asp Pro Leu

85

90

95

GGA TGC AGG GAA CCA G GTAAGGGAGG GGAGGGGGGG TGGGGAGGGG CCNGGCTGTG 1697

Gly Cys Arg Glu Pro Gly

100

CCCCCTCAC CTGCCCCTCC CCGACAG GC TGC CTG TTT GCT GTG CCT GCC CAG 1750

Cys Leu Phe Ala Val Pro Ala Gln

105

110

TCG GCT CCT CCA AAC CGG CCG AGG CTG AGA AGG AAA AGG CAG 1792

Ser Ala Pro Pro Asn Arg Pro Arg Leu Arg Arg Lys Arg Gln

115

120

125

10

20

30

40

GTACGGGGGC CCGCACAGAC CTCGGGCTGC AGAGACCTCG GGCTGCAGAG AGACCTCGGC	1852	
CGTGGCCCAG AGCAGGAGGG CACCCTCATC TATGGCTGGG GCGAAGGAAG GCTCAGATGG	1912	
ATGTGGCTGG GGGCCAGGGA CCGTGTCTGG GAGAAGCCCC CACCCCTTCC CTAATGCTGG	1972	
CATCTACAGA GGCCCCATCC TGGGCAAACC GAGGCTGCCT GCCCTCATTC CAAAGCTGAG	2032	
GAAGGACAGG ACCCTCTGCC AGTGGGGAGC TGGCACTGTC CCTGGCTGGA GTCCAGACCC	2092	
CCCCATCCCC ACGAGTCTG TTCCTGGCTT GGCCATGAGA TCAGTCAGAC ATGGAAGGGA	2152	
CTGATTCCAA GTGCCCACCC ACCCCCCAG GTC CGC CGG GGC CAC CCT ACA GTG	2205	10
Val Arg Arg Gly His Pro Thr Val		
130 135		
CAC AGT GAA ACC AAG TAT GTG GAG CTA ATT GTG ATC AAC GAC CAC CAG	2253	
His Ser Glu Thr Lys Tyr Val Glu Leu Ile Val Ile Asn Asp His Gln		
140 145 150		
CTG GTGAGTGCCA GGGCAGGGAC AGGGCGTGAC ACTGGGAGGC CCCTGAGGAG	2306	
Leu		20
CCTGGCCCTC CTCCCATTCT TCTCTCTCCC AG TTC GAG CAG ATG CGA CAG TCG	2359	
Phe Glu Gln Met Arg Gln Ser		
155		
GTG GTC CTC ACC AGC AAC TTT GCC AAG TCC GTG GTG AAC CTG GCC GAT	2407	
Val Val Leu Thr Ser Asn Phe Ala Lys Ser Val Val Asn Leu Ala Asp		
160 165 170 175		
GTG GTAAGCAGCT CTCCTCCCT CCCTTCCCTC CTCCTCATGC CCCCCACCC	2460	30
Val		
CACCACACAC ATTAGGGGGC ACTGTCAGCC CCTGGCTCCC ACTTCCTGGA GAGAACAGAC	2520	
AGGCCCTCCT CCAGCCCTGG CCCCAACACC CACTCCCACC CTCCAGCCCC CCTCATCTTC	2580	
TCCCCAG ATA TAC AAG GAG CAG CTC AAC ACT CGC ATC GTC CTG GTT GCC	2629	
Ile Tyr Lys Glu Gln Leu Asn Thr Arg Ile Val Leu Val Ala		
180 185 190		
ATG GAA ACA TGG GCA GAT GGG GAC AAG ATC CAG GTG CAG GAT GAC CTC	2677	40
Met Glu Thr Trp Ala Asp Gly Asp Lys Ile Gln Val Gln Asp Asp Leu		

195	200	205	
CTG GAG ACC CTG GCC CGG CTC ATG GTC TAC CGA CGG GAG GGT CTG CCT			2725
Leu Glu Thr Leu Ala Arg Leu Met Val Tyr Arg Arg Glu Gly Leu Pro			
210	215	220	
GAG CCC AGT AAT GCC ACC CAC CTC TTC TC GTGAGTCCCC CACCCTGCAC			2774
Glu Pro Ser Asn Ala Thr His Leu Phe Ser			
225	230		
CTCCTGCCAG CCTCTGCTAG TTGCTACAGT GCTTGGGATT ACTTAACACC TGCCCTGTGC			2834
TGGCTGCTCC TCTCAGAGTC TGGGGACTGG GCTCACCTTG CACCTGCCAC CTACCCCCAG			2894
CCACATGCAA CAGCTGGGCA TCATCCCCTG AATCTGAGGT TGATGCCCTT GTCTTAGCCC			2954
TGGTGGTCCT CTTCTGCCTC TCACCTCCCC TTAGTTCTGT CTTTCCCTTC AACTGTCCCN			3014
NNNNNNNNN NAGAGTGAAA CTCTGTCTCA AAAGAAAAAN AAAANAAAAG AAGAAAAAAA			3074
AGAACCCAAG GAGCGGGGA AGGGTCTTGC CTGGGGTCAC CAAGGCTGAT GTAAAGGGCC			3134
AGGCTCACCT CCTGAGGAAG GACTCTAGTG TGAGGGGCTC CCAAGGCC CACCACCACC			3194
CGGGGAGCCA CAGGGGAGGG CAGAAGCCAT CCTGACAGCG CACTCCCTTC CAG G GGC			3251
		Gly	
AGG ACC TTC CAG AGC ACG AGC AGC GGG GCA GCC TAC GTG GGG GGC ATA			3299
Arg Thr Phe Gln Ser Thr Ser Ser Gly Ala Ala Tyr Val Gly Gly Ile			
235	240	245	
TGC TCC CTG TCC CAT GGC GGG GGT GTG AAC GAG GTGAGCAGTG			3342
Cys Ser Leu Ser His Gly Gly Gly Val Asn Glu			
250	255	260	
GGGGGACATG GCTGGGGTGG CGGCTGAGGG AAAGGGGCTT AGGGGCACGA CGTGCCTGNT			3402
TGGAAGATGT AGACATCTGT GCCCATCTT CCCCACCCCC AG TAC GGC AAC ATG			3456
		Tyr Gly Asn Met	
GGG GCG ATG GCC GTG ACC CTT GCC CAG ACG CTG GGA CAG AAC CTG GGC			3504
Gly Ala Met Ala Val Thr Leu Ala Gln Thr Leu Gly Gln Asn Leu Gly			
265	270	275	280
ATG ATG TGG AAC AAA CAC CGG AGC TCG GCA G GTATCCTCCC CCAGAGGCCC			3555

10

20

30

40

Met Met Trp Asn Lys His Arg Ser Ser Ala Gly

285

290

CCGTGTGGCC CAGCAGCTCT GGAACGGGAG GGTGACAGTG GGAGGGGTGG TCCTTGGCCT 3615

CCCTCATATC CGCCTGGCTC ACCCCTCAG GG GAC TGC AAG TGT CCA GAC ATC 3667

Asp Cys Lys Cys Pro Asp Ile

295

TGG CTG GGC TGC ATC ATG GAG GAC ACT GG GTGAGTTCTT GGGGACAACC 3716

10

Trp Leu Gly Cys Ile Met Glu Asp Thr Gly

300

305

GGGGGAAGGT CTTGGGCGAG GGGAGTCTTA GAGCGAGCAT TGTTTGGCAG TCTGGACCAG 3776

GGGNNNNNNN NNNNGAACA CACCTTCCCT TCCAGGCCGG CTTGCGAGTC CCAGGTTCAA 3836

GCGAGGGATG GGAGCGACAA GGGACAAGGC GGAGGATTCT GGTGCAATCC CGGGGCAGAT 3896

CCTCCGCCTC CTCGCGATGG TGACGAAGTC CCCCAGTGTA CCCCCTCCCC AGCCTTGAGA 3956

GGGGTGAGGG TGGGTTGGAG GGGAGCAGCC AGCAGCACCT CCCCTCGCCC TATCCAG G 4014

20

TTC TAC CTG CCC CGC AAG TTC TCT CGC TGC AGC ATC GAC GAG TAC AAC 4062

Phe Tyr Leu Pro Arg Lys Phe Ser Arg Cys Ser Ile Asp Glu Tyr Asn

310

315

320

CAG TTT CTG CAG GAG GGT GGT GGC AGC TGC CTC TTC AAC AAG CCC CTC 4110

Gln Phe Leu Gln Glu Gly Gly Gly Ser Cys Leu Phe Asn Lys Pro Leu

325

330

335

340

AAG GTACCAGCCC CGCGGCGGGG AGCATGGGAG CGGGCCCTGG GCGGGGTCCG 4163

30

Lys

GGCCAGACTC CCGACCTGTC CTCCCGGTCC AG CTC CTG GAC CCC CCA GAG TGC 4216

Leu Leu Asp Pro Pro Glu Cys

345

GGG AAC GGC TTC GTG GAG GCA GGG GAG GAG TGC GAC TGC GGC TCG GTG 4264

Gly Asn Gly Phe Val Glu Ala Gly Glu Glu Cys Asp Cys Gly Ser Val

350

355

360

40

CAG GTGAGCGGTG GTGCGGGCGC CAGGTGGGGA ACCGGGATGC GGGGGTGGGC 4317



Gln

365

ACCAGGGAGC GTCTGAGTGG GAGGATTAGG GCTCGCCCGC CTCCTTCCCC TCCTCCCGCG 4377

TCCCTCAG GAG TGC AGC CGC GCA GGT GGC AAC TGC TGC AAG AAA TGC ACC 4427

Glu Cys Ser Arg Ala Gly Gly Asn Cys Cys Lys Lys Cys Thr

370

375

CTG ACT CAC GAC GCC ATG TGC AGC GAC GGG CTC TGC TGT CGC CGC TGC 4475

Leu Thr His Asp Ala Met Cys Ser Asp Gly Leu Cys Cys Arg Arg Cys

380

385

390

395

AAG GTAAGCAGGA CCGGCCGGGA GCGGGGCCA GGACGCAGGA GGAGCGATTG 4528

Lys

GAGGCCTTCA TATAAGGGGT GGGAGCTAGG GAGGGAAGCG GAGCCTTCGG GGACGAAGGC 4588

CTCTGGGGCA GGGCTTGATG CGAAGACAGC GCCAATGGGA GCAAGGGCGG GCTGAAGGAT 4648

GTTGAAGGCN NNNNNNNNNN NNNCGGACGG GAAGCTCCCA GAATCAAGGA GGGCGGGAAG 4708

GTGGGCGGGC TTGGGGCGGT GCTGAGTGGC CTGGGAGCGA GGTGGGAGC GTTCAAGAGG 4768

TGGTGGGAGC AGGGAAATAA GAACAGGCCT AAACGGGGCC CTGGGGAGCT GGAGGGCCCG 4828

GGGATGTGGG GGTCCAGAGA GCGGGGGGCC TGGGAGGGC AGGGCCGAGG CATCCATCCT 4888

GCCTGACTCG AGGAGCGCGT CTCTTCCCTA G TAC GAA CCA CGG GGT GTG TCC 4940

Tyr Glu Pro Arg Gly Val Ser

400

TGC CGA GAG GCC GTG AAC GAG TGC GAC ATC GCG GAG ACC TGC ACC GGG 4988

Cys Arg Glu Ala Val Asn Glu Cys Asp Ile Ala Glu Thr Cys Thr Gly

405

410

415

GAC TCT AGC CAG GTCCGCCCG CCGCGCGTC TTGTGGAGCC CTGGGCGAGG 5040

Asp Ser Ser Gln

420

CAACCCCTAC CCTTGTCGAT TTGGTTTTCC CGGACGAGTG CTCAGCACTC CCCTCCTCTC 5100

CACAGCTGGC ATCGACCTTC ACTGATCAGA CTGTTTTCTT ATCTGAGAAA GGGGTTCTTC 5160

ATGCTCCTGG CCTTGTTCTT TCAATCATT AACCAGAATG TATCGTCTGG CTGGTATCCC 5220

10

20

30

40

40

## Pro Gln Gln Gly Arg Ala Val Trp

500

CTT CCT CCT CTG TGT CAA CAT CTC TGG AGC TCC TCG GCT AGG GGA CCT 6145

Leu Pro Pro Leu Cys Gln His Leu Trp Ser Ser Ser Ala Arg Gly Pro

505

510

515

GGT GGG AGA CAT CAG TAGTGTACC TTCTACCACC AGGGCAAGGA GCTGGACTGC 6200

Gly Gly Arg His Gln

10

520

AGGTGCTGAC CAGCACCAAA ACTCAGGGAG GGGACCTGGC AGCTGTGCTG GGGGTTAGAA 6260

GATCTGGGGG CTGGAGGCTG GGCTGTGTCA CTTCCCCAGG GGAGGCCACG TGCAGCTGGC 6320

GGACGGCTCT GACCTGAGCT ATGTGGAGGA TGGCACAGCC TCGGGCCTA ACATGTTGTG 6380

CCTGGACCAT CGCTGCCTGC CAGCTTCTGC CTTCAACTTC AGCACCTGCC CCGGCAGTGG 6440

GGAGCGCCGG ATTTGCTCCC ACCACGGGGT GACTGCCTGG AGCCCGGGAT GGCGGGAGAA 6500

GCTTACAAGA GGGGACAGGC CCCTGCTCAC CTCTCCTGGC CCTGCCCTGC CTCTAGGTCT 6560

20

GCAGCAATGA AGGGAAGTGC ATCTGTCAGC CAGACTGGAC AGGCAAAGAC TGCAGTATCC 6620

ATAACCCCT GCCACGTCC CCACCCACGG GGGAGACGGA GAGATATAAA GGTGAGGCTG 6680

GAGCTGGCCG AGGGGGGTCT GTCTGTCCCG CTCTCTATGC CTGTCTTGC CAGCTAAGCC 6740

CTGCCATCCT CCCAGGTCCC AGCGGCACCA ACATCATCAT TGGCTCCATC GCTGGGGCTG 6800

TCCTGGTTGC AGCCATCGTC CTGGGCGGCA CGGGCTGGGG ATTTAAGTAA GAGACACACA 6860

CACCTGTGC CCCCTGGCAT CCTTGAGGGG GGATCAGAAT CCCTACTGGT GGAGCTGAGG 6920

GGGCCCTCCC TGAAAGCCCA ACTGAACCAG AGCTCACACG TCATAGGTCC AAGTAGCCTG 6980

30

CAGGGCTTAA CATTTAGAAA CTAGGAGATT TTAGGCTAGA TGAGGTGCTC ACGCCTGTAA 7040

TCCCAGCACT TTGGGAGGCC AAGGCAGGCG GATCACCTGA GGTGAGGAAT TCAAGACCAG 7100

TCTGGCCAAC ATGGTGAAAC CCGTCTCTAT TAAAAATACA AAAATTAGCC AGCCATGGTG 7160

GTGCACACCT GTAATCCCAG CTAATTGCGA GGCTGAGGCA GAGAATTGCT TGAACCCGGG 7220

AGGTGGAGGT TGCAGTGAGC TGAGATCGCA CCATTGCACT CCAGCCTTGG GTGACAGAGC 7280

AAGACTGCGT CAAAAAAAAA AAAAAAAAAA AAAAAAGGA AAGAAAGAGA GAAAGAAAAG 7340

AAAAGAGAAA AGAAATCAGG AGATTTTACA CTAGCAATTC GGATTTCCAG CTCTGGAAAC 7400

40

ATGAAAAGGT TGAGCCCCAG CGTGCCTCTA AGCATCCCCA AATAGCCACA GAGTGGAGCT 7460

GGGCAGGGGC	CACCCAAGCC	AGGCATGTGT	CCTCCAGTCT	CCAGTTCCCA	CCAGCCTATA	7520	
CTCCTTTGTG	CGTGTCTAAG	TTTGGGGTCC	TTGTGCCTGG	TCTTACCCCC	CTTAATGTGC	7580	
AGAGGGAGGA	ACCCACGGCC	CAAGGTCACA	TGATTGAGTT	AGTAGCAGAG	TCAGAGCTGG	7640	
AACCGGGACG	CATTTTTTGTG	GGTGCCCTGG	GTAATTCTCC	CTGGCCCTTA	CATTAGTGTC	7700	
CAGGCCCCGG	GGACCCCGGC	CCCGCTCTGG	GGCAAGGGGT	CGCATGGCAG	CCAAAGGCCC	7760	
CTCCCTGAGA	GAAGCAAAAG	GTCAGATGTC	TCCTTTTCCT	CTCCCCCTCC	ACCATCCTCC	7820	
CCCTGCAGAA	ACATTCGCCG	AGGAAGGTAC	GACCCGACCC	AGCTGGGGGC	AGTGTGATGC	7880	10
CGGCCACGTC	ATCCCTCCCG	CTGTCTTTGT	CTCCTCCATC	TCATTGCTCA	CCCGCGTTCT	7940	
GTTGATGGGG	TGCGGGGCCG	ATCCCACCCT	GCGTGCCNNN	NNNNNNNNNN	ATCTGTTTTG	8000	
TCTTCCATAT	CACCACTGTC	TGACCTCCCG	CAGATCCCTT	CCCTGGCCAG	CCTGTGACTT	8060	
GCCGCCTGCC	TCCAGGGCCC	AGAACTGAGC	TCCGGGGCCC	TGCTGGGGGG	CTCTCCCCGA	8120	
GGCCCTGCT	CACGTCTCTC	CCTGATGCCC	CCTCTCCGTT	CCAGGTCCGG	AGGGGCCTAA	8180	
GTGCCACCCT	CCTCCCTCCA	AGCCTGGCAC	CCACCGTCTC	GGCCCTGAAC	CACGAGGCTG	8240	
CCCCCATCCA	GCCACGGAGG	GAGGCACCAT	GCAAATGTCT	TCCAGGTCCA	AACCCCTCAA	8300	20
CTCCTGGCTC	CGCAGGGGTT	TGGGTGGGGG	CTGTGGCCCT	GCCCTTGGCA	CCACCAGGGT	8360	
GGACCAGGCC	TGGAGGGCAC	TTCCTCCACA	GTCCCCCACC	CACCTCCTGC	GGCTCAGCCT	8420	
TGCACACCCA	CTGCCCCGTG	TGAATGTAGC	TTCCACCTCA	TGGATTGCCA	CAGCTCAACT	8480	
CGGGGGCACC	TGGAGGGATG	CCCCCAGGCA	GCCACCAGTG	GACCTAGCCT	GGATGGCCCC	8540	
TCCTTGCAAC	CAGGCAGCTG	AGACCAGGGT	CTTATCTCTC	TGGGACCTAG	GGGGACGGGG	8600	
CTGACATCTA	CATTTTTTTAA	AACTGAATCT	TAATCGATGA	ATGTAAACTC	GGGGGTGCTG	8660	
GGGCCAGGGC	AGATGTGGGG	ATGTTTTGAC	ATTTACAGGA	GGCCCCGGAG	AAACTGAGGT	8720	30
ATGGCCATGC	CCTAGACCCT	CCCCAAGGAT	GACCACACCC	GAAGTCCTGT	CACTGAGCAC	8780	
AGTCAGGGGC	TGGGCATCCC	AGCTTGCCCC	CGCTTAGCCC	CGCTGAGCTT	GGAGGAAGTA	8840	
TGAGTGCTGA	TTCAAACCAA	AGCTGCCTGT	GCCATGCCCA	AGGCCTAGGT	TATGGGTACG	8900	
GCAACCACAT	GTCCCAGATC	GTCTCCAATT	CGAAAACAAC	CGTCCTGCTG	TCCCTGTCAG	8960	
GACACATGGA	TTTTGGCAGG	GCGGGGGGGG	GTTCTAGAAA	ATATAGGTTC	CTATAATAAA	9020	
ATGGCACCTT	CCCCCTTTNN	NNNNNNNNNN	NNNGGGATAC	CTCTGAATAT	GGGTATCTGG	9080	
GGCTGGATAT	GGGTGGGACA	TGAGACTTCC	TGTGACCAGC	CACCCTGGCT	CCCAGCTCTC	9140	40
TGTATCCTCC	TGCCCCGCCC	TGGGGGGTGC	CTACCCTGGN	AGAACCAGG	GAGGAGTGGA	9200	
GGCTGCCTCT	GCCTGGGCCT	CCACACAGCA	TCCTGACATA	CGCCACCTGG	GGTGGGGGTG	9260	
GGGAGGCAGG	GCCAGGAG					9278	

【図面の簡単な説明】

【図 1】3 4 2 個のコスミド・クローンの第 1 7 番染色体上の位置を示した図である。クローン名はクローン番号のみで示してある。

【図2】卵巣癌における第17番染色体長腕の部分的欠失を示した図である。黒丸はヘテロ接合性の消失（LOH）、白丸は保持を表している。2カ所の共通欠失領域は傍線で示した。

【図3】乳癌における第17番染色体長腕の部分的欠失を示した図である。黒丸はヘテロ接合性の消失（LOH）、白丸は保持を表している。2カ所の共通欠失領域は傍線で示した。

【図4】第17番染色体長腕21.3領域に位置づけられたマーカーから該遺伝子の単離に至る過程、および乳癌組織で遺伝子再構成の起きている箇所（斜線の矩形）を示した図である。クローン名はクローン番号のみで示してある。

【図5】乳癌における遺伝子再構成のサザンプロットによる検出を示した図であり、Nは正常組織のDNA、Tは癌組織DNAを示す。

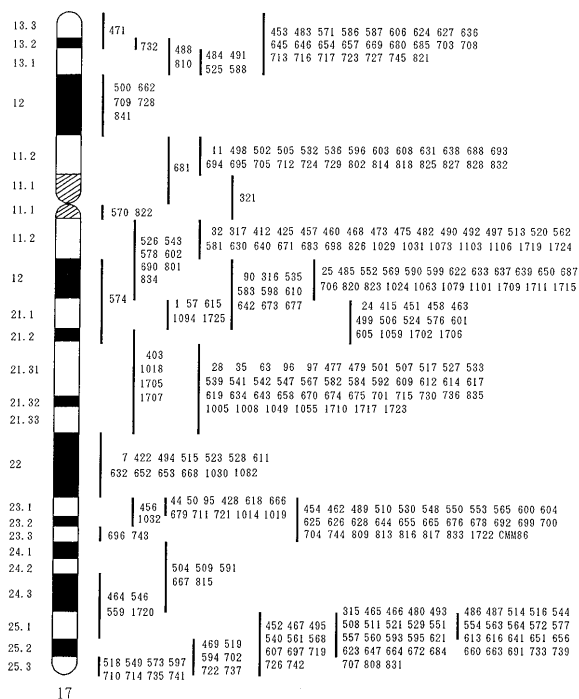
10

【図6】乳癌における遺伝子再構成のサザンプロットによる検出を示した図であり、Nは正常組織のDNA、Tは癌組織DNAを示す。

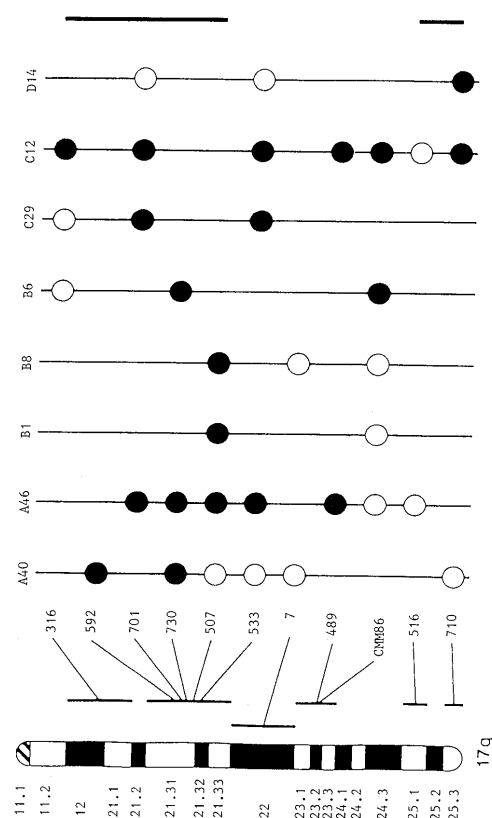
【図7】乳癌における遺伝子再構成のサザンプロットによる検出を示した図であり、Nは正常組織のDNA、Tは癌組織DNAを示す。

【図8】モノクローナル抗体およびウサギポリクローナル抗体でのELISA法によるMDC蛋白質の濃度測定検量線の図である。

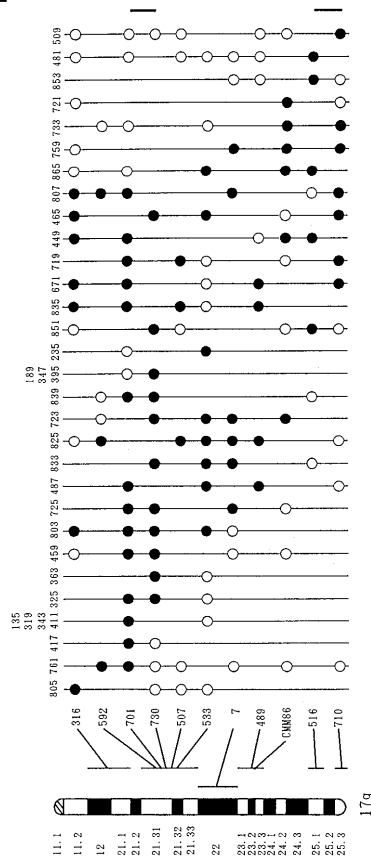
【図1】



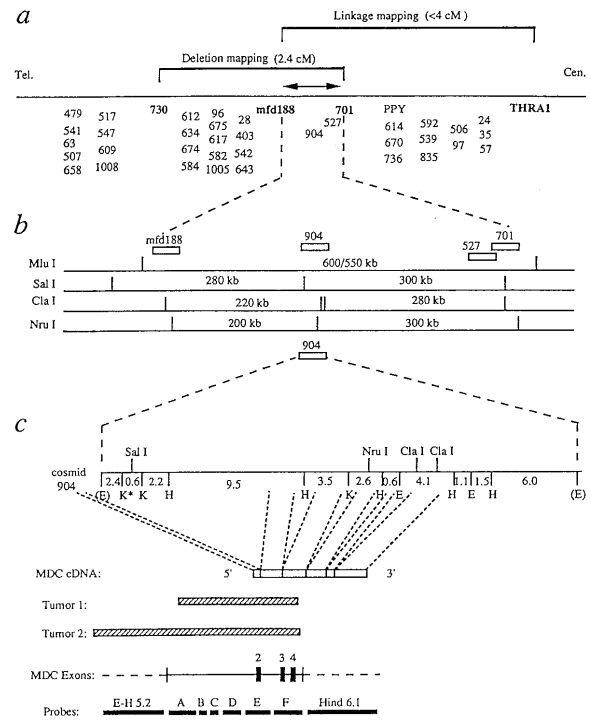
【図2】



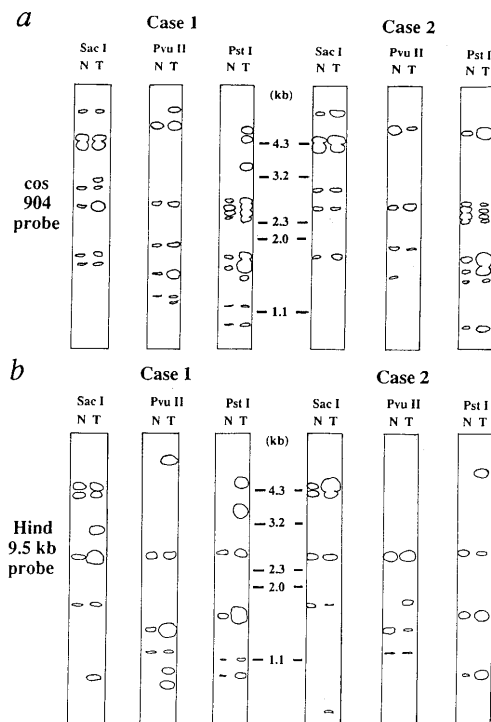
【 図 3 】



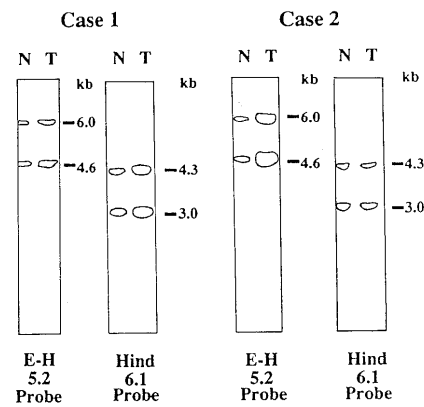
【 図 4 】



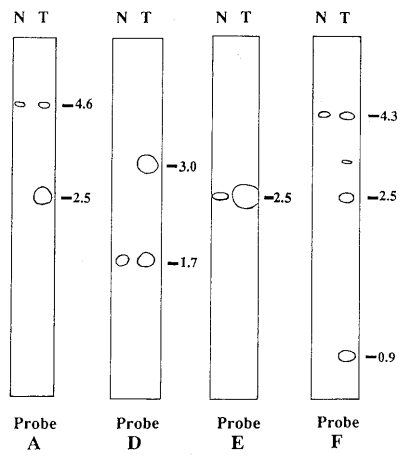
【 図 5 】



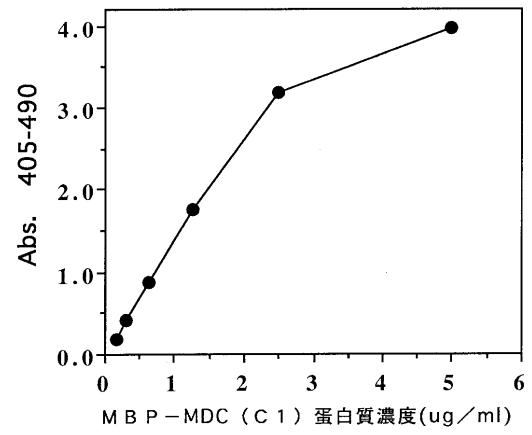
【 図 6 】



【 図 7 】



【 図 8 】



## フロントページの続き

(51) Int.Cl. <sup>7</sup>	F I
C 1 2 P 21/08	C 1 2 Q 1/68 A
C 1 2 Q 1/68	C 1 2 N 15/00 Z N A A

(31)優先権主張番号 特願平6-73328

(32)優先日 平成6年4月12日(1994.4.12)

(33)優先権主張国 日本国(JP)

## 前置審査

(72)発明者 江見 充

東京都北区滝野川 6 - 1 - 1 1 - 1 0 6

審査官 深草 亜子

(56)参考文献 J.Biol.Chem. , 1 9 9 0 年 , Vol.265 , p.16068-16073

Nature Genet. , 1 9 9 3 年 1 0 月 , Vol.5 , p.151-157

(58)調査した分野(Int.Cl.<sup>7</sup> , D B 名)

C12N 15/09

GenBank/EMBL/DDBJ/PIR/Swissprot/GeneSeq