

人、小鼠、大鼠三个物种珠蛋白基因家族信息

No 编号	Name 基因名	RefSeq mRNA RefSeq 登录号	BP 长度	CDS 编码区	Gene ID 基因登录号	Protein ID 蛋白质登录号
Hs01	HsHBZ	NM_005332	589	56-484	3050	NP_005323
Hs02	HsHBM	NM_001003938	524	25-450	3042	NP_001003938
Hs03	HsHBA2	NM_000517	622	67-495	3040	NP_000508
Hs04	HsHBA1	NM_000558	627	67-495	3039	NP_000549
Hs05	HsHBQ1	NM_005331	653	154-582	3049	NP_005322
Hs06	HsHBE1	NM_005330	816	254-697	3046	NP_005321
Hs07	HsHBG2	NM_000184	583	54-497	3048	NP_000175
Hs08	HsHBG1	NM_000559	584	54-497	3047	NP_000550
Hs09	HsHBD	NM_000519	774	196-639	3045	NP_000510
Hs10	HsHBB	NM_000518	626	51-494	3043	NP_000509
Hs11	HsMYB	NM_005368	1078	81-454	4151	NP_005359
Hs12	HsCYGB	NM_134268	2166	364-936	114757	NP_599030
Hs13	HsNGB	NM_021257	1885	376-831	58157	NP_067080
Mm01	MmHba-x	NM_010405	630	65-493	15126	NP_034535
Mm02	MsHba-a1	NM_008218	569	36-464	15122	NP_032244
Mm03	MsHbq1a	NM_175000	604	50-487	216635	NP_778165
Mm04	MsHba-a2	NM_001083955	587	33-461	110257	NP_001077424
Mm05	MmHbq1b	NM_001033981	748	50-478	544763	NP_001029153
Mm06	MsHbb-y	NM_008221	619	56-499	15135	NP_032247
Mm07	MmHbb-bh1	NM_008219	610	53-496	15132	NP_032245
Mm08	MmHbb-bh2	NM_001127686	559	95-538	436003	NP_001121158
Mm09	MsHbb-b1	NM_001278161	630	55-498	15129	NP_001265090
Mm10	MsHbb-b2	NM_016956	630	55-498	15130	NP_058652
Mm11	MsMb	NM_001164047	1121	205-699	17189	NP_001157519
Mm12	MsCygb	NM_030206	2331	395-967	114886	NP_084482
Mm13	MmNgb	NM_022414	1630	280-735	64242	NP_071859
Rn01	RnHbz	NM_001172845	589	51-479	287168	NP_001166316
Rn02	RnHba1	NM_013096	556	34-462	25632	NP_037228
Rn03	RnHba2	NM_001007722	545	34-462	360504	NP_001007723
Rn04	RnHbe1	NM_001008890	444	1-444	293267	NP_001008890
Rn05	RnHbe2	NM_001024805	444	1-444	502359	NP_001019976
Rn06	RnHbg1	NM_172093	444	1-444	94164	NP_742090
Rn07	RnHbb	NM_033234	620	48-491	24440	NP_150237
Rn08	RnHbb-b1	NM_198776	659	57-550	361619	NP_942071
Rn09	RnMb	NM_021588	1015	28-492	59108	NP_067599
Rn10	RnCygb	NM_130744	2138	171-743	170520	NP_570100
Rn11	RnNGB	NM_033359	1773	410-865	85382	NP_203523

说明：此表引自“实用生物信息技术课程教学实例”（罗静初，2015，生物技术通报）表3“人、小鼠、大鼠三个物种珠蛋白家族基因信息”。

需要说明的是,小鼠脑红蛋白基因有两个剪接变体,RefSeq 参考序列数据库中 mRNA 序列登录号为 NM\_022414 和 NM\_001294308。NM\_022414 编码区长度 453bp, 编码 151 个氨基酸; NM\_001294308 编码区长度 465bp, 编码 155 个氨基酸。表中只列出其中一个 NM\_022414。小鼠 beta 珠蛋白家族成员 MsHbb-b1 和 MsHbb-b2 为单倍体型 C57BL/6 株系基因组中测得的序列, RefSeq 参考序列数据库中 mRNA 序列登录号为 NM\_001278161 和 NM\_016956。小鼠基因组计划测序样本所用的为融合体 BALB/c 和 129Sv 株系。小鼠基因组信息系统(MGI)中所列小鼠 beta 珠蛋白家族成员基因名为 MsHbb-bs 和 MsHbb-bt, RefSeq mRNA 登录号为 NM\_001201391 和 NM\_008220, 表中未予列出。

此外,大鼠基因组中 $\alpha$ -珠蛋白家族共有 7 个成员,位于 10 号染色体 15.468-5.508Mb 区段,长度约 40KB; 表 8 中只收录已有转录数据的 3 个,即 RnHbz (NM\_013096)、RnHba1 (NM\_013096) 和 RnHba2 (NM\_001007722)。另外 4 个尚无确切证据,也没有确定的基因名,未在表中列出。这 4 个基因中,一个为 $\alpha$ -珠蛋白, RefSeq 数据库中 mRNA 序列登录号为 NM\_001013853, 大鼠基因组数据库 RGD 中暂定基因名为 LOC287167。其它三个为 $\theta$ -珠蛋白,尚无实验证据。大鼠基因组中, $\beta$ -珠蛋白共有 9 个成员,位于 1 号染色体 175.095Mb-175.170 区段,约 75Kb,其中 1 个为假基因,4 个为串联重复排列的 $\alpha$ -珠蛋白,推测由近期发生的基因倍增机制产生。