Brief Definitive Report

THE α-GLOBIN PSEUDOGENE ON MOUSE CHROMOSOME 17 IS CLOSELY LINKED TO H-2

BY PETER D'EUSTACHIO,* BENNETT FEIN,* JAMES MICHAELSON,‡ AND BENJAMIN A. TAYLOR§

*From the *Department of Biochemistry, New York University Medical Center, New York, New York 10016; ‡Center for Blood Research, Boston, Massachusetts 02117; and §The Jackson Laboratory, Bar Harbor, Maine 04609*

Two functional genes encode the adult form of α-globin in the mouse. These genes are tightly linked on chromosome 11. Two DNA sequences, $\alpha\psi3$ and $\alpha\psi4$, have been described which are closely homologous to adult α-globin genes but encode no detectable protein (1–3). These sequences are not linked to the functional α-globin gene cluster. Analysis of somatic cell hybrids has localized $\alpha\psi3$ to chromosome 15 and $\alpha\psi4$ to chromosome 17 (3, 4). To characterize the $\alpha\psi4$ DNA sequence further, we have now defined DNA polymorphisms associated with it and used the polymorphisms to map the $\alpha\psi4$ sequence to the centromeric side of the major histocompatibility complex *H-2*.

Materials and Methods

Livers from individual mice were Dounce homogenized in 0.15 M NaCl, 0.01 M EDTA, 0.01 M Tris, pH 8 (TNE), and the homogenates centrifuged 5 min at 1,000 *g* to yield crude nuclear pellets. Pellets not used immediately were stored at −70°C. DNA was purified from pellets by a standard sequence of protease digestion and phenol/chloroform extraction, followed by dialysis against TNE and ethanol precipitation (5). DNA was redissolved in sterile 10 mM Tris–1 mM EDTA, pH 8.0, and stored at 4°C. 10-μg samples were digested with 40–50 U of TaqI restriction endonuclease (New England Biolabs, Beverly, MA) for 4–5 h according to the manufacturer's instructions, fractionated by agarose gel electrophoresis (1% agarose in 90 mM Tris, 90 mM boric acid, 1.25 mM EDTA, pH 8.3), and blotted onto nitrocellulose filters (BA85, Schleicher and Schuell, Inc., Keene, NH) (6). Filters were baked 4 h at 80°C under vacuum, incubated 3–4 h at 65°C in a solution composed of 6× SSC, 10× Denhardt's solution, 0.05 M sodium phosphate, pH 7.0, 1% glycine, 0.5 mg/ml sonicated heat-denatured (5 min, 100°C) salmon sperm DNA, and then hybridized with the $\alpha\psi4$ DNA probe for 8–12 h at 65°C in a solution composed of 6× SSC, 2× Denhardt's solution, 0.02 M sodium phosphate, pH 7.0, 10% dextran sulfate (mol wt_{avg} = 500,000; Sigma Chemical Co., St. Louis, MO), 0.1 mg/ml sonicated heat-denatured salmon sperm DNA. (1× SSC = 0.15 M NaCl, 0.015 M sodium citrate, pH 7.0; 1× Denhardt's solution = 0.02% polyvinyl pyrrolidone, 0.02% Ficoll 400, 0.02% bovine serum albumin). The probe consisted of the 2.4 kilobase (kb) EcoRI genomic DNA fragment containing the $\alpha\psi4$ sequence, inserted into the EcoRI site of pBR322 (3). It was radiolabeled by nick-translation (7) in the presence of $\alpha[^{32}P]$-dCTP

This work was supported by National Institutes of Health grants GM32105 (to P. D'Eustachio), CA29863 (to J. Michaelson), and GM 18684 (B. A. Taylor). P. D'Eustachio is a Scholar of the Leukemia Society of America. J. Michaelson is supported by a Research Career Development Award from the National Institutes of Health. The Jackson Laboratory is fully accredited by the American Association for Accreditation of Laboratory Animal Care.

(New England Nuclear, Boston, MA) to a specific activity of $5–10 \times 10^7$ Cherenkov cpm per μg, heat denatured, and added to the hybridization solution at a final concentration of 7 ng/ml. After hybridization, filters were washed in 0.3× SSC − 0.1% sodium dodecyl sulfate at 65°C before autoradiography (5).

Mice were typed for *Upg-1* as described previously (8).

Results and Discussion

We refer to the $\alpha\psi 4$ DNA sequence as *Hba-a4*: the fourth gene discovered homologous to the adult α-globin polypeptide. This nomenclature conforms to that proposed for mouse genes (9), and provides a neutral and open-ended system for distinguishing the members of dispersed multigene families in mice and humans (10).

Polymorphism at this locus was detected by restriction analysis of genomic DNA from 15 inbred strains of mice. Three patterns were identified in Southern blots of DNA digested with TaqI restriction endonuclease (Fig. 1). BALB/cJ is the prototype strain for *Hba-a4*^{a}; C57BL/6J the prototype for *Hba-a4*^{b}; and SM/J the prototype for *Hba-a4*^{c}. CAST/Ei (inbred *Mus musculus castaneus*) DNA yielded a band pattern indistinguishable from BALB/cJ (*a*), but MOL/Ei (inbred *M. m. molossinus*) DNA yielded a novel pattern, defining a fourth allele (*d*) (not shown). There is no simple correlation between the distributions of *Hba-a4* alleles and *H-2* haplotypes in these mouse strains. Point mutation at a single TaqI recognition site cannot account for the difference between the two most common forms of the gene (e.g. BALB/cJ vs. C57BL/6J). The two forms appear to differ by at least two point mutations, or the insertion of a DNA segment ~1.7 kb in size containing a TaqI recognition site.

To localize the *Hba-a4* gene on chromosome 17, we examined mice from three sets of recombinant inbred (RI) strains whose progenitor strains were distinguishable at this locus (Table I). 37 RI strains were tested. All were homozygous for one of the expected parental forms of *Hba-a4*. No novel forms were detected. 30 strains carried parental allelic combinations at *Hba-a4* and at *H-2* and seven

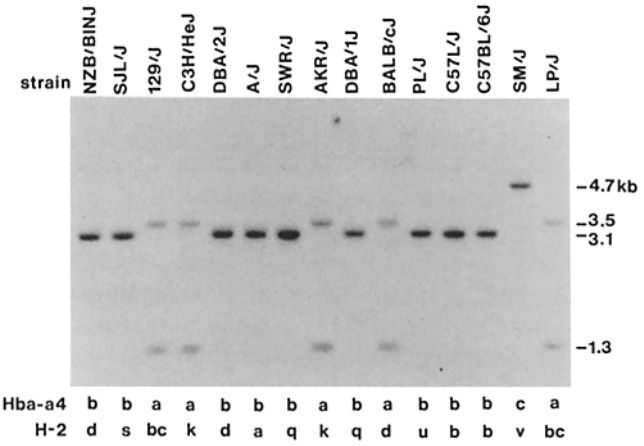

FIGURE 1. DNA polymorphism associated wtih *Hba-a4*. All mice were obtained from The Jackson Laboratory. Sizes of DNA fragments in kilobases (kb) are shown. The letter beneath each lane indicates the allelic form of *Hba-a4* found in the strain.

TABLE I

Distribution of Hba-a4 alleles in recombinant inbred strains

Strain	Locus: *Hba-a4*		*H-2*		*Upg-1*
CXBD	a		d		f
E	b		b		s
G	a	X	b		s
H	a		d		f
I	b		b		s
J	b		b		s
K	b		b		s
BALB/cJ	a		d		f
C57BL/6J	b		b		s
BXH2	a		k	X	s
3	a		k		f
4	b		b		s
6	a		k		f
7	a		k		f
8	b		b		s
9	a	X	b		s
10	b		b		s
11	a	X	b		s
12	a		k		f
14	a		k		f
19	b		b		s
C57BL/6J	b		b		s
C3H/HeJ	a		k		f
AKXL 5	b		b		s
6	b	X	k		f
7	b		b		s
8	a		k		f
9	b		b		s
12	b		b		s
13	a		k		f
14	b	X	k		f
16	b		b		s
17	b		b		s
19	b		b		s
21	a		k		f
24	a	X	b		s
25	b		b		s
28	a		k		f
29	a	X	b		s
37	b		b		s
38	a		k	X	s
AKR/J	a		k		f
C57L/J	b		b		s

Hba-a4 was assayed as in Fig. 1. *H-2* types are from Michaelson (17). Recombinants between *Hba-a4* and *H-2*, and between *H-2* and *Upg-1* are indicated by an "X" in the appropriate interval. All these recombinants are also recombinant between *Hba-a4* and *Upg-1*.

were recombinant, for an estimated recombination frequency of 0.066 ± 0.031 between the two loci (11). There is no indication of heterogeneity of recombination frequency among the three sets of RI strains.

These RI strains also segregate for *Upg-1*, a biochemically defined locus ~6 cM distal to *H-2* (8). Nine strains were recombinant between *Hba-a4* and *Upg-1*, yielding an estimated recombination frequency of 0.096 ± 0.044 (Table I). These recombination frequencies alone do not allow the three loci to be ordered. The order *Hba-a4* – *H-2* – *Upg-1*, however, requires only single crossover events to account for all observed recombinant genotypes. The order *H-2* – *Hba-a4* –

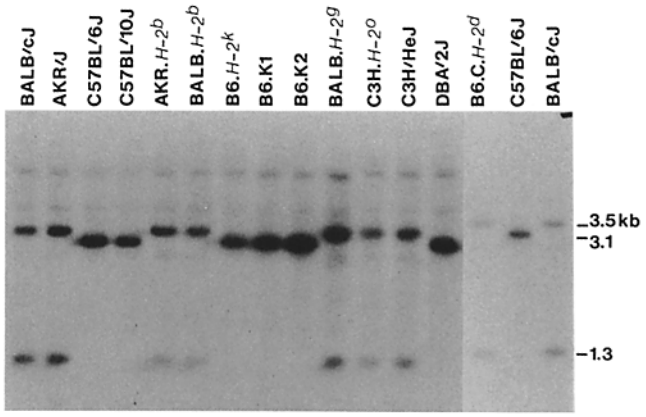

FIGURE 2. *Hba-a4* polymorphism in congenic strains. Mice obtained from The Jackson Laboratory (C3H-*H-2*°, B6.C-*H-2*^{d}) or from E. Boyse, Memorial Sloan-Kettering Institute, New York (all other congenic strains) were assayed as in Fig. 1.

TABLE II

Hba-a4 Alleles in H-2-congenic Mice

Strain	*Hba-a4*	*T* ⬇ *Qglo-1* *H-2K* *H-2D* *Tla* *Upg-1*
AKR.*H-2*^{b}	AKR(*a*)	
BALB.*H-2*^{b}	BALB(*a*)	
B6.*H-2*^{k}	C57BL/6(*b*)	
B6.K1	C57BL/6(*b*)	
B6.K2	C57BL/6(*b*)	
BALB.*H-2*^{g}	BALB(*a*)	
C3H.*H-2*°	C3H(*a*)	
B6.C-*H-2*^{d}	BALB(*a*)	

Each congenic strain is assumed to contain a single contiguous segment of donor chromosome 17 (▬) incorporated by reciprocal crossing over into the recipient strain chromosome (——). Regions of uncertain origin are indicated by dashed lines. *Hba-a4* types are taken from Figure 2. Assignments at other loci are from refs. 9–11. The arrow indicates the apparent site of the *Hba-a4* locus.

Upg-1 would require seven double crossovers (strains CXBG, BXH-9, BXH-11, AKXL-6, AKXL-14, AKXL-24, and AKXL-29), and the order *H-2* – *Upg-1* – *Hba-a4* would require two (strains BXH-2 and AKXL-38), suggesting that *Hba-a4* is in fact proximal to *H-2*.

To confirm the linkage and gene order in an independent set of mice, eight *H-2* congenic strains of mice carrying various segments of chromosome 17 were examined (Fig 2). One, B6.C-*H-2*^{d}, carries a BALB/c-derived segment extending from a point proximal to *Qglo-1* to a point between *H-2D* and *Upg-1*. The remainder of the chromosome is of C57BL/6 origin (12). This strain showed the BALB/c (*a*) form of *Hba-a4*. Seven other strains were of recipient strain *Hba-a4* type, including three whose donor chromosome segment extends distally past *Upg-1* and proximally no further than *H-2D* (12–14) (Table II). Together, these results place *Hba-a4* on chromosome 17, proximal to *H-2*. This conclusion is supported by the observation of tight linkage between *Hba-a4* and *tf* in mouse *t* complex chromosomes (H. S. Fox, L. M. Silver, and G. R. Martin, personal communication).

Hba-a4 promises to be a useful marker for the region proximal to *H-2*. It is

also the first dispersed pseudogene to be placed in a mammalian linkage map. Despite its nomadic origins (3) and extensive present-day polymorphism, the gene shows no signs of instability. Within the limits of our assays, it maps to a single locus and no further variants have arisen in the hundreds of generations separating the recombinant inbred and congenic mice from their inbred strain parents. This behavior contrasts with the apparent instability of the DNA polymorphisms at murine immunoglobulin heavy chain switch regions (15), or at sea urchin histone "orphons" (16).

Summary

DNA sequences homologous to adult α-globin genes are dispersed in the mouse. Two functional genes are tightly linked on chromosome 11. Pseudogenes have been assigned to chromosomes 15 and 17 by analysis of interspecies somatic cell hybrids. We have now further characterized the second of these pseudogenes, *Hba-a4*. The gene is highly polymorphic, with three forms occurring in a panel of 15 inbred strains and a fourth occurring in an inbred strain derived from *M. m. molossinus*. Analysis of *Hba-a4* alleles in CXB, BXH, and AKXL recombinant inbred strains placed *Hba-a4* 6.60 ± 3.14 cM centromeric to *H-2*. Analysis of congenic mouse strains confirmed the linkage and the gene order. *Hba-a4* is the first mammalian dispersed pseudogene to be localized in a linkage map, and should provide a useful marker for the region of chromosome 17 proximal to *H-2*.

We thank Lee Silver for helpful discussions, and Philip Leder for the αψ4 globin probe.

Received for publication 30 November 1983.

References

1. Leder, P., J. N. Hansen, D. Konkel, A. Leder, Y. Nishioka, and C. Talkington. 1980. Mouse globin system: a functional and evolutionary analysis. *Science (Wash. DC).* 209:1336.
2. Vanin, E. F., G. I. Goldberg, P. W. Tucker, and O. Smithies. 1980. A mouse α-globin-related pseudogene lacking intervening sequences. *Nature (Lond.).* 286:222.
3. Leder, A., D. Swan, F. H. Ruddle, P. D'Eustachio, and P. Leder. 1981. Dispersion of α-like globin genes of the mouse to three different chromosomes. *Nature (Lond.).* 293:196.
4. Popp, R. A., P. A. Lalley, J. B. Whitney, and W. F. Anderson. 1981. Mouse α-globin genes and α-globin-like pseudogenes are not syntenic. *Proc. Natl. Acad. Sci. USA.* 78:6362.
5. D'Eustachio, P., A. L. M. Bothwell, T. K. Takaro, D. Baltimore, and F. H. Ruddle. 1981. Chromosomal location of structural genes encoding murine immunoglobulin λ light chains. *J. Exp. Med.* 153:793.
6. Southern, E. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98:503.
7. Rigby, P. W. J., M. Dieckman, C. Rhodes, and P. Berg. 1977. Labelling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. *J. Mol. Biol.* 113:237.
8. Szymura, J. M., B. A. Taylor, and J. Klein. 1982. *Upg-2*: a urinary pepsinogen variant located on chromosome 1 of the mouse. *Biochem. Genet.* 20:1211.

9. Lyon, M. 1981. Nomenclature. *In* The Mouse in Biomedical Research. H. L. Foster, J. D. Small, and J. G. Fox, editors. Academic Press, New York. p. 27.
10. D'Eustachio, P., and F. H. Ruddle. 1983. Somatic cell genetics and gene families. *Science (Wash. DC).* 220:919.
11. Taylor, B. A. 1978. Recombinant inbred strains: use in genetic mapping. *In* Origins of Inbred Mice. H. C. Morse, editor. Academic Press, New York. p. 423.
12. Rinchik, E. M., and D. B. Amos. 1983. Genetic analysis of the *T-H-2* region in non-*t* chromosomes. I. Two new congenic strains isolate *Qglo-1* from *H-2*. *Immunogenetics.* 17:445.
13. Flaherty, L. 1981. Congenic strains. *In* The Mouse in Biomedical Research. H. L. Foster, J. D. Small, and J. G. Fox, editors. Academic Press, New York. p. 215.
14. Klein, J., F. Figueroa, and D. Klein. 1982. *H-2* haplotypes, genes, and antigens: second listing. I. Non-*H-2* loci on chromosome 17. *Immunogenetics.* 16:285.
15. Marcu, K. B., J. Banerji, N. A. Penncavage, R. Lang, and N. Arnheim. 1980. 5′ flanking region of immunoglobulin heavy chain constant region genes displays length heterogeneity in germlines of inbred mouse strains. *Cell.* 22:187.
16. Childs, G., R. Maxson, R. H. Cohn, and L. Kedes. 1981. Orphons: dispersed genetic elements derived from tandemly repetitive genes of eukaryotes. *Cell.* 23:651.
17. Michaelson, J. 1983. Genetics of beta-2 microglobulin in the mouse. *Immunogenetics.* 17:219.