Chromosomal Clustering of Five Defined Endogenous Retrovirus Loci in White Leghorn Chickens

ALLAN TEREBA¹* AND SUSAN M. ASTRIN²

Division of Virology, St. Jude Children's Research Hospital, Memphis, Tennessee 38101,¹ and The Institute for Cancer Research, The Fox Chase Cancer Institute, Philadelphia, Pennsylvania 19111²

Received 26 February 1982/Accepted 3 May 1982

By using in situ hybridization, we localized three endogenous chicken retrovirus loci associated with the $gs^- chf^-$ phenotype to distinct regions on chromosome 1. These three loci, *ev4*, *ev5*, and *ev8*, along with the two other known $gs^- chf^-$ -associated loci in White Leghorn chickens, *ev1* and *ev13*, appear to be evenly distributed along chromosome 1, occurring approximately every 1.8×10^7 base pairs. The presence of all five loci on chromosome 1 represents a highly nonrandom distribution and all but eliminates a mechanism of independent random integration events for the generation of these loci.

In our analysis of the chromosomal distribution of White Leghorn chicken endogenous retrovirus loci, using in situ hybridization (21), we discovered that three loci associated with virus expression (19) and two loci associated with expression of only some viral proteins (17) resided on five different chromosomes. This random distribution was consistent with the presumed random integration of exogenous retroviruses (11, 12, 15) and, along with hybridization studies involving various avian species (9), supported the contention that endogenous retrovirus loci were generated by independent exogenous infections sometime after the speciation of jungle fowl.

Of importance to us in our current study was our initial observation that three unidentified endogenous retrovirus loci associated with non-expression of viral proteins ($gs^- chf^-$) were located on chromosome 1 (21). One of these loci has since been identified as the common *ev1* locus present in almost all White Leghorn chickens. To determine whether the chromosomal positions of this phenotypically similar group of loci reflected an exception to the random chromosomal distribution of other retrovirus loci, we examined all of the known virus loci associated with this phenotype.

Primary fibroblast cultures of SPAFAS Hodel no. 11 chicken embryos were analyzed for the absence of viral protein, using the COFAL test for *gs* antigen (14), and the chicken helper factor assay for viral glycoprotein (22). Embryos which lacked viral protein expression ($gs^- chf^-$) were then analyzed for virus-specific DNA fragments after the cell DNA had been digested with *Sst*I and *Bam*HI restriction endonucleases. As determined from our previous classification of endogenous chicken retrovirus loci (2, 5), all of the

embryos used contained the *ev1* locus. In addition, embryos 251-B1, 11, and 1846 had restriction endonuclease-digested DNA fragments representing *ev4*, *ev5*, and *ev8*, respectively.

To determine on which chromosomes these viral loci were located, metaphase chromosome preparations were made, using fibroblasts from embryos 251-B1, 11, and 1846. These preparations were then hybridized with a probe containing ¹²⁵I-labeled sea urchin DNA that was attached to 35S viral RNA from the Prague strain of Rous sarcoma virus, subgroup C (PR-RSV-C). This RNA would be expected to hybridize to endogenous retrovirus sequences and to the endogenous *c-src* sequence present in all chickens (16). Previous studies (18, 21) have shown that hybridization with this probe provides a good quantitative estimate of the number of retrovirus loci since the number of silver grains developed is dependent on the [¹²⁵I]DNA and not on the size of the hybridized RNA sequences. Our previous work also has shown that the *c-src* gene is nonreiterated and is on the large microchromosome group 10 to 12 (18, 21). Thus, the *c-src* gene provides a good internal quantitative reference in case both viral loci are located on the same chromosome.

Table 1 shows the grain distribution over each chromosome from the three embryos. In every case, after subtracting the background grains, only two chromosome groups, 1 and 10 to 12, contained a statistically significant number of grains above background ($P \leq 0.0003$), assuming a Poisson distribution of background grains. Since the Rous sarcoma virus RNA-containing probe would hybridize to the *c-src* gene and both viral loci, some clustering of the viral loci was indicated.

Previous studies indicated that *ev1* is located

TABLE 1. Grain distribution over metaphase chromosomes from three characterized embryos hybridized with a probe containing Rous sarcoma virus RNA

Chromosome	251-B1 ^a		11 ^b		1846 ^c	
	Total grains	Grains above background	Total grains	Grains above background	Total grains	Grains above background
1	77	27 ^e	134	60 ^e	121	51 ^e
2	36	1	55	6	52	-2
3	31	3	29	-7	40	-1
4	22	2	30	5	25	-5
Z	16	-1	26	5	25	-1
6	17	1	23	-2	22	0
7-9	25	-1	36	-5	50	11
10-12	31	16 ^e	49	25 ^e	50	26 ^e
Micro	36	-5	55	-1	61	-3

^a From 254 chromosome sets of embryo 251-B1, which contains *ev1* and *ev4* as defined by restriction endonuclease analysis.

^b From 71 chromosome sets of embryo 11, which contains *ev1* and *ev5* as defined by restriction endonuclease analysis.

^c From 179 chromosome sets of embryo 1846, which contains *ev1* and *ev8* as defined by restriction endonuclease analysis.

^d The background was determined as described previously (18). All values have greater than a 1% probability of occurring randomly, assuming a Poisson distribution of background grains over the chromosomes except where noted.

^e Less than a 0.03% probability of occurring randomly.

on chromosome 1 and that the endogenous *c-src* gene is located on chromosome group 10 to 12 (18, 21). To determine which of these chromosome groups also contain *ev4*, *ev5*, and *ev8*, we compared the number of grains above background on chromosome 1 with those on chromosome group 10 to 12. Table 2 summarizes the data in Table 1 and also includes data on *ev1* (embryo 5) and three uncharacterized viral loci present in another embryo (SPAFAS) which was *gs⁻ chf⁻*. Three patterns emerge. Embryo 5 had a ratio of approximately 1 when the grains due to specific hybridization over chromosome 1 were compared with grains over chromosome group 10 to 12. This value agrees with the one viral locus in this embryo detected by restriction endonuclease analysis and localized to chromosome 1 by in situ hybridization (18). Embryos 251-B1, 11, and 1846, which each contain two viral loci, had a ratio closer to 2, indicating that *ev4*, *ev5*, and *ev8* are also on chromosome 1. Finally, the SPAFAS embryo had a ratio of between 3 and 4. In situ hybridization experiments have already shown that this embryo has three loci on chromosome 1 (21). Thus, all of the observed loci appear to be on chromosome 1.

The localization of *ev1*, *ev4*, *ev5*, and *ev8* on chromosome 1 indicated that a mechanism of independent random integration probably was not involved in the positioning of these genes ($P = 10^{-4}$). To determine whether these loci were clustered in one area of chromosome 1, we analyzed the grain distribution along chromosome 1 after normalizing the chromosomes to a uniform length. Figure 1 shows these data and includes results on embryo 5, which contained only *ev1*, and also a *gs⁻ chf⁻* SPAFAS embryo, which had three uncharacterized loci. By using the embryo which contained only *ev1* (Fig. 1A) as a reference, it was easy to distinguish the *ev1* locus on chromosome 1 of each embryo. A peak of grains whose size was consistent with ¹²⁵I emissions (1, 8) was detected in every case at position 0.26 with a standard deviation of ± 0.02 U. Embryo 251-B1, which also contained *ev4* (Fig. 1B), had an additional peak of equal size at position 0.56, which is near the end of the long arm. Embryo 11 (Fig. 1C) had an additional peak at position 0.43. The untyped SPAFAS embryo (Fig. 1E) also contained a peak at this location, and since both embryos originated from the same flock, the untyped bird probably also had the *ev5* locus. Figure 1C also shows minor peaks at 0 and 0.6. These peaks were probably due to random variations in background since they were too narrow for ¹²⁵I emissions, were not statistically highly significant, and could not be accounted for by our restriction endonuclease analysis. Embryo 1846 (Fig. 1D), in addition to the *ev1* locus, had a peak of grains at position -0.16, which is on the middle of the short arm. Finally, the SPAFAS embryo (Fig. 1E), besides the *ev1* and *ev5* loci discussed above, had a peak of grains at position -0.04. Due to its unique

TABLE 2. Quantitative analysis of grains over chromosome groups 1 and 10 to 12^a

Embryo	Viral ^b loci	Grains above background on following chromosome:		Chromosome 1/ chromosome 10-12 ratio of grains
		1	10-12	
5	<i>ev1</i>	87	71	1.2
251-B1	<i>ev1</i> , <i>ev4</i>	26	16	1.6
11	<i>ev1</i> , <i>ev5</i>	61	25	2.4
1846	<i>ev1</i> , <i>ev8</i>	51	26	2.0
SPAFAS	ND ^c	90	26	3.5

^a In situ hybridization reactions were performed with a PR-RSV-C RNA-containing probe.

^b Determined by restriction endonuclease analysis of virus-specific DNA fragments.

^c ND, Not determined.

FIG. 1. Localization of the $gs^- chf^-$ -associated retrovirus loci on chromosome 1. The distribution of silver grains on chromosome 1 was determined from autoradiograms of chromosomes hybridized in situ with [^{125}I]DNA probes containing PR-RSV-C RNA. The broken lines (---) indicate the calculated level of background silver grains. The length of chromosome 1 is defined as 1 U, and the distance is measured from the centromere, with the long arm having positive values and the short arm having negative values. (A) Embryo 5 containing *ev1*; (B) embryo 251-B1 containing *ev1* and *ev4*; (C) embryo 11 containing *ev1* and *ev5*; (D) embryo 1846 containing *ev1* and *ev8*; (E) uncharacterized SPAFAS embryo; (F) summary of results displayed on chromosome 1.

chromosomal position, it has been defined as a new locus and given the designation *ev13*.

Figure 1F provides a summary of the location of the five loci located on chromosome 1. This analysis clearly demonstrates that all of the $gs^- chf^-$ -associated endogenous retrovirus loci detected so far in White Leghorn chickens reside on chromosome 1. In addition, these five loci are not clustered but occur at regular intervals along

the chromosome. Excluding a gap at 0.11, all of the loci were 0.145 ± 0.025 U apart. This corresponds to approximately 1.8×10^7 base pairs (S. Bloom, personal communication). Recently, we have located *ev6*, a $gs^- chf^-$ -associated locus with a structure very similar to that of *ev4* and *ev5* (10) at position 0.10 on the long arm of chromosome 1 (17), thus confirming this unusual spacing. The significance of this regular spacing is unknown as yet, but a statistical analysis revealed that any regular spacing along the chromosome as observed here would not be due to random occurrence ($P = 6 \times 10^{-5}$).

The reliability of our data with regard to spacing of these loci along chromosome 1 is supported by the reproducibility we obtained in the localization of *ev1* and by animal genetic studies. The *ev1* locus was localized at position 0.26 ± 0.02 on the long arm of chromosome 1 in a total of 14 separate experiments (Fig. 1) (17–21, 23). Dividing the chromosome into fewer or more divisions during analysis reduced the resolution between two close loci when large divisions were used, but the number and mean location of each statistically significant peak remained unchanged.

Animal genetic studies, using three separate chicken crosses where the segregation of *ev4* and *ev5* could be examined, showed seven recombinants in 35 offspring (2–4). This suggests a distance of 20 map units (7 to 33 map units at the 95% confidence level). With a haploid chicken genome complexity of 1.14×10^9 base pairs, a DNA content of 11% of the genome for chromosome 1, and 4.7×10^9 base pairs per map unit (S. Bloom, personal communication), the in situ hybridization data indicate that the distance between *ev4* and *ev5* is $1.6 \times 10^7 \pm 3 \times 10^6$ base pairs or approximately 28 map units. This result is in excellent agreement with the genetic experiments and substantiates the conclusion that *ev4* and *ev5* are physically linked (Fig. 1).

Current data suggest that exogenous avian retrovirus genomes integrate randomly throughout the chicken genome (11, 12, 15). Our data showing the location of all five $gs^- chf^-$ -associated endogenous retrovirus loci on chromosome 1 contrast sharply with this suggested mechanism of integration. Taking into account the relative size of chromosome 1, this concentration of loci would be highly improbable ($P = 1.6 \times 10^{-5}$) if nonselected and left to chance. When all 10 of the retrovirus loci examined so far (17–19) (see above) are considered regardless of associated phenotype, the nonrandom distribution is still equally significant ($P = 10^{-6}$).

One source of selection that may result in a nonrandom distribution of viral loci could involve mating patterns between chickens with few and many viral loci. White Leghorn chick-

ens were originally bred in the United States from a very small number of chickens (7) and have been continually selected on the basis of certain traits (6). If these traits were associated with viral loci on chromosome 1 and were epitomized in birds containing many gs^- chf^- -associated loci, breeding selection could shift a random distribution of viral loci to a nonrandom chromosomal organization. Although possible, it is highly unlikely since all but the adjacent viral loci on chromosome 1 recombine independently of each other. In addition, this explanation requires many assumptions, some of which appear to be incorrect based on the genetic map of chickens. It also does not explain the relatively uniform distribution of the loci along chromosome 1 and the unique 5'-terminal deletion associated with at least three of these chromosome 1-associated loci (10).

Since neither the mechanism of random exogenous viral infection nor selective pressures can reasonably account for the presence of all five viral loci on chromosome 1, an alternative mechanism is likely involved. Of the available alternatives, some mode of duplication and transposition appears to be the most probable. The restriction to chromosome 1, the lack of observable transposition by retroviruses under laboratory conditions (13), and the frequent occurrence of 5' deletions in these loci (10) would also suggest that the transposition was most likely the end result of a duplication of part of chromosome 1 followed by recombination events. This model can be tested by analyzing the sequence homologies between the various endogenous viral loci.

We thank Michael Lai for providing highly purified 35S viral RNA and R. Shah, M. Gryniewicz, D. Keefer, and J. Fang for their expert technical assistance.

This work was supported by grants PCM 78-03253 and PCM 76-09721 from the National Science Foundation, research grant CA28221 from the National Cancer Institute, Cancer Center Support (CORE) grant CA 21765, Biomedical Research Support grant RR 05584 from the National Institutes of Health, grants CA 06927 and RR 05539 from the National Institutes of Health, an appropriation from the Commonwealth of Pennsylvania, and by ALSAC.

LITERATURE CITED

1. Ada, G. L., J. H. Humphrey, B. A. Askonas, H. O. McDevitt, and G. J. V. Nossal. 1966. Correlation of grain counts with radioactivity (^{125}I and tritium) in autoradiography. *Exp. Cell Res.* 41:557-572.
2. Astrin, S. 1978. Endogenous viral genes of the White Leghorn chickens: common site of residence and sites associated with specific phenotypes of viral gene expression. *Proc. Natl. Acad. Sci. U.S.A.* 75:5941-5945.
3. Astrin, S. M., L. B. Crittenden, and E. G. Buss. 1979. *ev3*, a structural gene locus for endogenous virus, segregates with gs^+ chf^+ phenotype in matings for line 6₃ chickens. *Virology* 99:1-9.
4. Astrin, S. M., and H. L. Robinson. 1979. *Gs*, an allele of chickens for endogenous avian leukosis virus antigens, segregates with *ev3*, a genetic locus that contains structural genes for virus. *J. Virol.* 32:420-425.
5. Astrin, S. M., H. L. Robinson, L. B. Crittenden, E. G. Buss, J. Wyban, and W. S. Hayward. 1979. Ten genetic loci in the chicken that contain structural genes for endogenous avian leukosis viruses. *Cold Spring Harbor Symp. Quant. Biol.* 44:1105-1109.
6. Bacon, L. D. 1979. Origin of inbred and genetically defined strains of chickens, p. 607. *In* P. L. Altman and D. D. Katz (ed.), *Inbred and genetically defined strains of laboratory animals*. Federation of American Societies for Experimental Biology, Bethesda, Md.
7. Brown, E. 1906. *Races of domestic poultry*, p. 40. Arnold, London.
8. Ertl, H. H., L. E. Feinendegen, and H. J. Heiniger. 1970. Iodine-125, a tracer in cell biology: physical properties and biological aspects. *Phys. Med. Biol.* 15:447-456.
9. Frisby, D. P., R. A. Weiss, M. Roussel, and D. Stehelin. 1979. The distribution of endogenous chicken retrovirus sequences in the DNA of galliform birds does not coincide with avian phylogenetic relationships. *Cell* 17:623-634.
10. Hayward, W. S., S. B. Braverman, and S. M. Astrin. 1979. Transcriptional products and DNA structure of endogenous avian proviruses. *Cold Spring Harbor Symp. Quant. Biol.* 44:1111-1121.
11. Hughes, S. H., P. R. Shank, D. H. Spector, H.-J. Kung, J. M. Bishop, H. E. Varmus, P. K. Vogt, and M. L. Breitman. 1978. Provirus of avian sarcoma virus are terminally redundant co-extensive with unintegrated linear DNA, and integrate at many sites. *Cell* 15:1397-1410.
12. Hughes, S. H., P. K. Vogt, E. Stubblefield, J. M. Bishop, and H. E. Varmus. 1981. Integration of avian sarcoma virus DNA in chicken cells. *Virology* 108:208-221.
13. Mason, W. S., M. Linial, T. W. Hsu, R. N. Eisenman, J. Townsend, G. E. Mark, G. Seal, C. Aldrich, and J. M. Taylor. 1982. Alterations in the genomes of avian sarcoma viruses. *Virology* 117:456-474.
14. Payne, L. N., and R. Chubb. 1968. Studies on the nature and genetic control of an antigen in normal chick embryos which reacts in the COFAL test. *J. Gen. Virol.* 3:379-391.
15. Sabran, J. L., T. W. Hsu, C. Yeater, A. Kaji, W. S. Mason, and J. M. Taylor. 1979. Analysis of integrated avian RNA tumor virus DNA in transformed chicken, duck, and quail fibroblasts. *J. Virol.* 29:170-178.
16. Stehelin, D., H. E. Varmus, J. M. Bishop, and P. K. Vogt. 1976. DNA related to the transforming gene(s) of avian sarcoma viruses is present in normal avian DNA. *Nature (London)* 260:170-173.
17. Tereba, A. 1981. 5'-Terminal deletions are a common feature of endogenous retrovirus loci located on chromosome 1 of White Leghorn chickens. *J. Virol.* 40:920-926.
18. Tereba, A., and S. Astrin. 1980. Chromosomal localization of *ev1*, a frequently occurring endogenous retrovirus locus in White Leghorn chickens, by *in situ* hybridization. *J. Virol.* 35:888-894.
19. Tereba, A., L. B. Crittenden, and S. M. Astrin. 1981. Chromosomal localization of three endogenous retrovirus loci associated with virus production in White Leghorn chickens. *J. Virol.* 39:282-289.
20. Tereba, A., and M. M. C. Lai. 1982. Cell oncogenes are located on the large microchromosomes in chicken cells. *Virology* 116:654-657.
21. Tereba, A., M. M. C. Lai, and K. G. Murli. 1979. Chromosome 1 contains the endogenous RAV-O retrovirus sequences in chicken cells. *Proc. Natl. Acad. Sci. U.S.A.* 76:6486-6490.
22. Weiss, R. A., W. S. Mason, and P. K. Vogt. 1973. Genetic recombinants and heterozygotes derived from endogenous and exogenous avian RNA tumor viruses. *Virology* 52:535-552.
23. Wong, T. C., A. Tereba, P. K. Vogt, and M. M. C. Lai. 1981. Characterization of the transformation-specific sequences of avian erythroblastosis virus in normal vertebrate cells. *Virology* 111:418-426.