

# Comparative analysis of avian leukosis virus-related endogenous viral genes in experimental strains of the domestic chicken

M Tixier-Boichard<sup>1</sup>, L Durand<sup>1</sup>, M Morisson<sup>1</sup>,  
F Ricard<sup>2</sup>, G Coquerelle<sup>2</sup>

<sup>1</sup> *Institut National de la Recherche Agronomique, Laboratoire de Génétique Factorielle, Domaine de Vilvert, 78952 Jouy-en-Josas Cedex;*

<sup>2</sup> *Institut National de la Recherche Agronomique, Station de Recherches Avicoles, Domaine de l'Orfrasière, 37380 Nouzilly, France*

**Summary** – DNA sequences related to avian leukosis virus (*ev* genes) were identified in the genome of chickens from 6 experimental strains including 3 randombred lines (Wyandotte, Fayoumi and White Leghorn), 2 divergently selected Rhode Island Red lines and a synthetic broiler line. Digestion of genomic DNA with *SacI* and *BamHI* enzymes and hybridization with either RAV-2 sequences or an LTR-*gag* probe revealed 66 different *SacI* bands, 53 of which could be associated with particular *BamHI* bands such that the *ev* loci could be defined at the molecular level. The distribution patterns were very strain-specific. Average band frequency and heterozygosity level could be correlated with the genetic history of the lines and their inbreeding level which varied from 0.04 to 0.28. Within-family segregation did not show any allelic relationship between retroviral insertions that were considered to be 2 genes. The Fayoumi line showed the largest variation in *ev* number per bird and contained a few birds free of any endogenous viral sequences, designated *ev0*. The broiler line showed the highest *ev* number. Very few *ev* loci were shared by different lines; only *ev6* was found in all strains. Comparison with the established classification in White Leghorns was not straightforward and further analysis of DNA polymorphism and viral expression is needed.

chicken / DNA polymorphism / genetic variability / endogenous retrovirus

**Résumé** – Analyse comparative des gènes viraux endogènes apparentés aux virus leucosiques aviaires dans des lignées expérimentales de poule domestique. Les gènes viraux endogènes (*ev*) apparentés aux virus des leucoses et sarcomes aviaires ont été identifiés sur un échantillon de reproducteurs et de leurs descendants issus de 6 populations expérimentales de poule qui différaient par leur origine génétique (Leghorn blanche, Wyandotte, Fayoumi, 2 lignées de Rhode Island Red, une lignée synthétique de poule de chair notée Y11), leur niveau moyen de consanguinité (estimations de 0,04 à 0,28) et la sélection subie. La digestion de l'ADN génomique par les enzymes *SacI* ou *BamHI* suivie

de l'hybridation avec les sondes RAV-2 ou LTR-gag dérivées du virus du sarcome de Rous a révélé 66 fragments avec SacI dont 53 ont pu être associés avec un fragment BamHI afin de caractériser les gènes *ev* au niveau moléculaire. Chaque population a présenté une distribution qui lui est propre. Les différences de fréquences géniques et d'hétérozygotie sont assez bien reliées aux taux de consanguinité des populations. L'étude de la ségrégation familiale n'a montré aucune relation d'allélisme entre ce que nous avons considéré comme 2 gènes *ev*. Le nombre de gènes *ev* par animal était particulièrement variable dans la souche Fayoumi qui présentait quelques animaux sans copies virales endogènes (notés *ev0*). La souche Y11 a montré le plus grand nombre d'insertions virales. Les populations d'origine génétique différente ont très peu de gènes *ev* en commun ce qui n'a pas permis le calcul d'une distance génétique; seul le gène *ev6* est partagé par toutes, fait remarquable dont la signification est discutée. Parmi les 53 loci identifiés avec 2 enzymes de restriction, la majorité apparaît différente des gènes *ev* déjà décrits dans la souche Leghorn blanche actuellement utilisée comme référence pour la nomenclature internationale.

poulet / polymorphisme d'ADN / variabilité génétique / rétrovirus endogène

## INTRODUCTION

Endogenous viral genes (*ev* genes) are DNA sequences of a retroviral origin that have been integrated into the chicken genome and are transmitted in a Mendelian manner. One group of *ev* genes is closely related to the avian leukosis virus (ALV). These endogenous retroviruses have been identified in the red junglefowl (Frisby *et al*, 1979), which is considered to be the ancestor of the domestic chicken. They differ in their insertion site and their proviral structure. Some are complete with viral sequences *gag*, *pol* and *env* flanked by long terminal repeats (LTR) and others are partially deleted (Rovigatti and Astrin, 1983; Crittenden, 1991). Some are transcriptionally active and are responsible for the production of viral proteins (Rovigatti and Astrin, 1983).

The polymorphism of *ev* genes was first studied in White Leghorn chickens (Astrin, 1978). A classification of 22 different *ev* genes has been established from the study of restriction fragment length together with the determination of viral phenotypic expression (Smith, 1987; Crittenden, 1991). Subsequent studies identified further *ev* genes in this breed (Kuhnlein *et al*, 1989a). Initial observations of chickens from breeds other than the White Leghorn suggested that their *ev* genes are quite different and also very numerous (Rovigatti and Astrin, 1983). Indeed, a very high level of polymorphism has been described for *ev* genes in commercial layers and broilers (Aarts *et al*, 1991; Boulliou *et al*, 1991; Iraqi *et al*, 1991; Sabour *et al*, 1992), and new *ev* genes were also found in other experimental strains (Gudkov *et al*, 1981; Ronfort *et al*, 1991; Lamont *et al*, 1992). Independent studies have proved to be very difficult to compare for technical reasons as well as because of differences in animal material.

The purpose of this study was to describe the *ev* genes and compare the polymorphism of insertion sites and their segregation pattern between several experimental strains chosen for their diverse genetic origin.

## MATERIAL AND METHODS

### *Strains and sampling procedure*

The strains are listed and described in table I. They cover a wide range of genetic origins and included a White Leghorn (LJ) as a reference strain. The Fayoumi strain (FA) is an old Egyptian breed adapted to hot climate and poor food (Mérat and Bordas, 1982). The experimental Wyandotte strain (X55) was maintained for many generations as part of a conservation programme. The 6 populations studied differed in genetic size and in estimated degree of inbreeding. They were free of leukosis infection and were all fast-feathering. The sampling procedure involved full-sib and half-sib families so as to follow the segregation of *ev* genes and assess the level of heterozygosity. The absence of an *ev* insertion is considered as a null allele which means that homozygous or heterozygous birds cannot be distinguished without family information.

### *Genomic DNA isolation and analysis*

Blood samples were collected on EDTA. After haemolysis, the pellet was incubated overnight in 3 ml buffer containing 50 mM NaCl, 10 mM Tris-HCl (pH 7.5), 10 mM EDTA, 0.4% SDS and 0.1 mg/ml proteinase K. Samples were then extracted with phenol and chloroform-isoamyl alcohol (24:1) and DNA precipitated with isopropanol. The precipitate was washed with 70% ethanol and dissolved in 2–3 ml TE buffer (10 mM Tris-HCl pH 7.5 and 1 mM EDTA). The final DNA concentration was adjusted to around 250  $\mu\text{g}/\text{ml}$ . Seven- $\mu\text{g}$  aliquots of DNA were digested to completion with 50 units of *SacI* or *BamHI* restriction enzymes which were known to yield one major 3' flanking fragment for each *ev* insertion in the genome (fig 1).

Digested DNA samples were electrophoresed in 0.7% agarose at 2 V/cm for 40 h with recirculation of TEA buffer (Maniatis *et al*, 1982). The *Raoul* I ladder (Appligene) was used as a size marker. The DNA was then blotted onto a charged nylon membrane (Biotrace, Gelman Labs) in 0.4 M NaOH. After transfer, the membranes were washed with 2  $\times$  SSPE (Maniatis *et al*, 1982) and prehybridized at 42°C for at least 4 h in 5  $\times$  SSPE, 1% SDS, 50% formamide, 0.1% Denhart, 5% Dextran and 200  $\mu\text{g}/\text{ml}$  salmon sperm DNA. The membranes were probed for hybridization with  $^{32}\text{P}$ -labelled RAV-2 ( $10^7$  counts per minute for a blot of 400  $\text{cm}^2$ ) for 20 h at 42°C. After hybridization, the membranes were washed twice at room temperature with 2  $\times$  SSPE and 0.5% SDS and twice at 65°C with decreasing concentrations of SSPE. The membranes were blotted dry and autoradiographed at -80°C using Kodak Safety Film AR and intensifying screens.

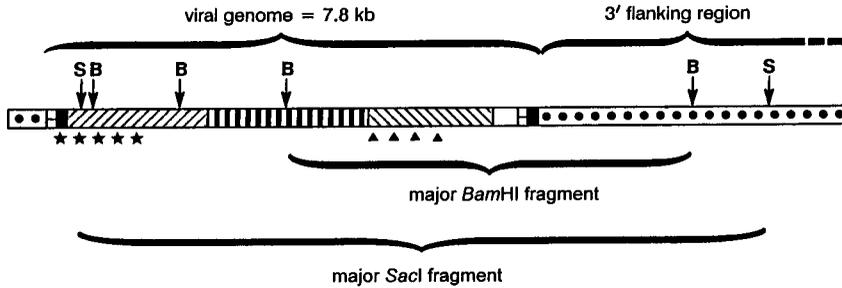
### *Probes*

The plasmid pRAV-2 is a pBR322 derivative that contains cDNA sequences corresponding to genes *gag*, *pol* and *env* and the LTR of the Rous-associated virus 2 and was a generous gift from LB Crittenden (RPRL, East Lansing, MI, USA). The inserted RAV-2 sequences (7.8 kb) were separated from the pBR322 vector by *SaII* digestion in order to avoid excessive labelling of the *Raoul* ladder with pBR322.

**Table I.** Main characteristics of the experimental lines sampled for the study of endogenous viral genes.

<i>Line code Breed</i>	<i>Genetic management</i>	<i>No of generations recorded</i>	<i>Effective size<sup>a</sup> N<sub>e</sub></i>	<i>Estimated inbreeding level F<sup>b</sup></i>	<i>Sample Size Parents Progeny</i>
X55	Randombred	19	90	0.10	15 84
R+	Selected for Rhode Island high RFC <sup>c</sup>	16	37	0.28*	11 26
R <sup>-</sup>	Selected for low RFC <sup>c</sup>	16	37	0.25*	11 24
LJ	White Leghorn	15	34	0.20	16 48
FA	Egyptian Fayoumi	11	37	0.14	17 59
Y11	Synthetic Broiler <sup>d</sup> body weight	6	73	0.04	12 58

<sup>a</sup> For LJ, FA, R<sup>-</sup> and R<sup>+</sup>:  $\frac{1}{N_e} = \frac{4}{16N_m} + \frac{4}{16N_f}$ , for Y11:  $\frac{1}{N_e} = \frac{1}{4N_m} + \frac{1}{4N_f} + \frac{1}{4N_f}$ , for X55:  $\frac{1}{N_e} = \frac{1}{16N_m} + \frac{1}{16N_f}$  where  $N_m$  is the number of breeding males and  $N_f$  the number of breeding females; a geometric mean was used in the case where  $N_m$  or  $N_f$  changed across generations. <sup>b</sup>  $F = 1 - (1 - \frac{1}{2N_e})^n$  for the generations recorded; <sup>c</sup> RFC: residual food consumption, divergent selection has been applied from a common base population; <sup>d</sup> Y11 was obtained in 1985 from a cross between 4 commercial parental strains and 2 experimental lines (Ricard, 1989); \* for the R<sup>-</sup> and R<sup>+</sup> selected lines, the true average inbreeding level was calculated using all pedigree information, as described in Bordas *et al*, 1992.



**Fig 1.** Schematic representation of an ALV-related endogenous viral gene inserted in the chicken genome (adapted from Rovigatti and Astrin, 1983). The entire viral genome is identified by the RAV-2 probe; the stars \* indicate the sequence corresponding to the *gag* subprobe and the triangles ▲ the *env* subprobe. ● genomic sequences; ■ LTR DNA sequence; ▨ *gag* DNA sequence; ▩ *pol* DNA sequence; ▭ *env* DNA sequence.

A mixture of 25 ng RAV-2 sequences and 250 pg linearized pBR322 was labelled with a random-primed DNA labelling kit (Amersham) using 50  $\mu$ Ci dC\*TP. A subprobe containing cDNA sequences for LTR and *gag* (figure 1) was obtained by *EcoRI* and *HindIII* digestion of plasmid Spl/LTR-*gag*, a gift from A Boulliou (URA 256 CNRS, Rennes, France). Another subprobe containing cDNA sequences for *env* (fig 1) was obtained by *KpnI* and *SalI* digestion of plasmid Spl/*env*, as described by Boulliou *et al* (1991). Reference genomic DNAs containing either *ev1*, *ev3*, *ev6*, *ev12* or *ev21* were generous gifts from EJ Smith and LB Crittenden (RPRL, East Lansing, MI, USA).

### **Estimation of heterozygosity and nomenclature of *ev* loci**

The banding pattern was analysed to determine the average number of *ev*-related fragments per bird and its variation. The average number of carriers for a given fragment was also determined with its variation. Calculations were done separately for parents and progeny but gave close enough results for the data to be pooled. Mean values obtained for the different strains were compared by Student's *t*-test. Samples from a subset of individuals representing the different lines were run on the same gel to facilitate size comparisons of fragments previously identified on different blots. Samples from families were run on the same blots to simplify the scoring of the bands. Simple rules were followed to determine the genotype of birds: when one offspring was not carrying a gene, the carrier(s) parent(s) was heterozygous; when one parent transmitted a gene to all its progeny (at least 4) it was homozygous. Families in which both parents and sampled progeny carried the same fragment were not informative: this probably led to an underestimation of the number of heterozygous birds.

*Ev* loci that could be defined unambiguously by the association of a *SacI* restriction fragment and a *BamHI* restriction fragment were assigned a letter for

the line code and a number within line, in order of decreasing size of the specific *SacI* fragment (or the specific *Bam*HI fragment when similar *SacI* fragments were obtained).

## RESULTS

### *Distribution pattern*

A total of 66 different *SacI* restriction fragments were identified in the 5 populations: the 10 fragments found in R<sup>-</sup> were also found in R<sup>+</sup> (table II). The distribution pattern appeared to be strain specific (fig 2).

The average number of *SacI* restriction fragments per bird differed between strains, except for FA, LJ and X55, but these 3 lines differed by the variation in number of *ev* per bird and by the total number of *ev*. FA was the only line to contain birds in which no *SacI* fragments could be detected. Lines were classified into 2 groups according to the mean *ev* frequency (table II): group 1 consisted of FA and Y11 and had a low *ev* frequency together with a high rate of heterozygosity among carriers; group 2 contained R<sup>-</sup>, R<sup>+</sup> and X55 and had high *ev* frequency and low heterozygosity; the LJ line was intermediate. FA differed from Y11 by the total number of *ev* genes. In group 2, some *ev* genes were approaching fixation, including the *SacI* fragments of 25 kb, 10.6 kb, 10.2 kb and 9.2 kb in R<sup>+</sup>, and others were rare such as the *SacI* fragments of 8 kb and 7 kb in R<sup>+</sup> (fig 2c). The divergently selected R<sup>-</sup> and R<sup>+</sup> strains differed mainly at 5 *ev* loci: 2 were present only in R<sup>+</sup> and 3 showed significantly different frequencies.

No evidence was obtained for the presence of allelism between restriction fragments. This was particularly clear in a Fayoumi family obtained from an *ev0* dam which scored negative with the RAV-2 probe and was mated to a sire carrying 6 fragments.

### *Structure of ev loci*

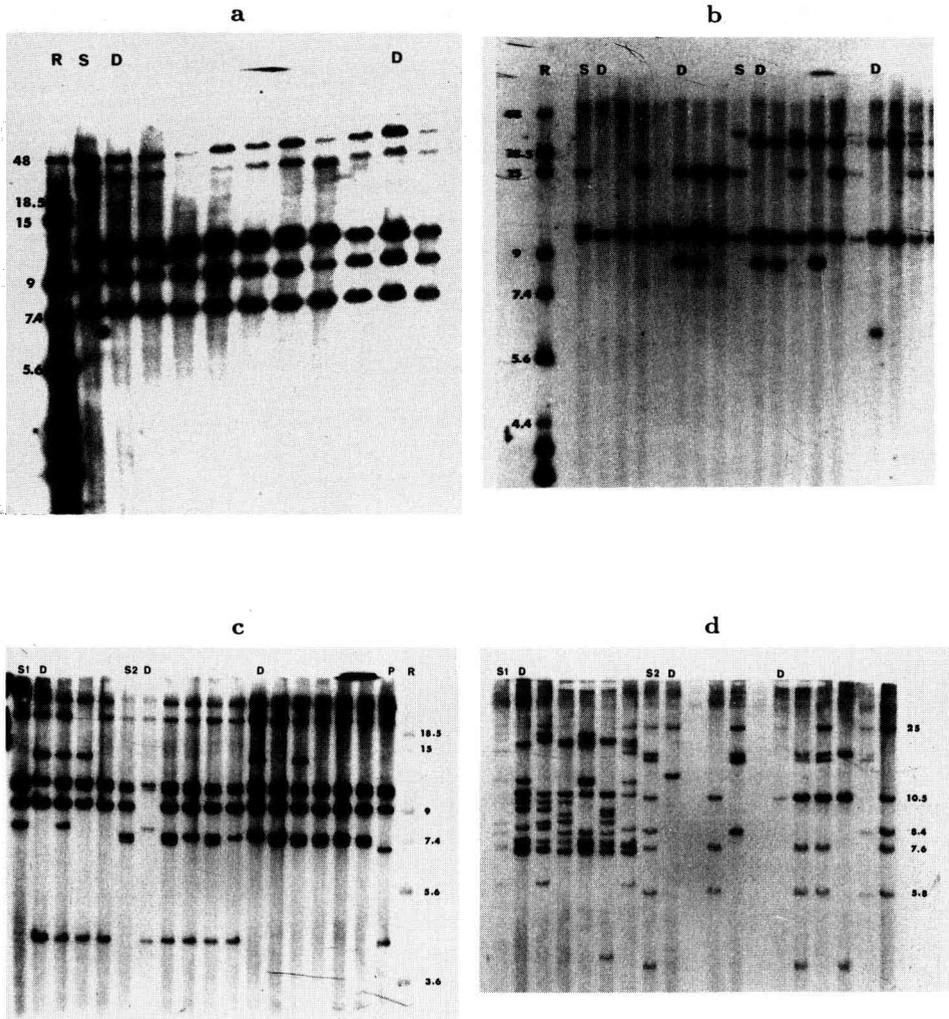
The *SacI* restriction fragments detected were each between 45 kb and 4 kb. Only 53 of the 66 fragments could be associated with a specific *Bam*HI fragment (table III); 19 *SacI* fragments were larger than 11 kb, 20 were between 11 and 8 kb and 11 were smaller than 8 kb. The association of one *SacI* fragment with one *Bam*HI fragment to describe each retroviral insertion was particularly difficult in the Y11 line where many fragments segregated.

Only 2 loci were found in more than one strain: *ev6* was shared by all strains; and *evmt1/evb24* was common to R<sup>-</sup>, R<sup>+</sup> and Y11. Of the 48 different *ev* loci, only 3 appeared to be deleted in the *gag* region, *ev6*, *ev9* and F6. The 9 *ev* loci showing a *SacI* fragment smaller than 7.8 kb were probably deleted in the *pol* region because they hybridized with *gag* and RAV-2 probes but the length of their flanking fragment was not compatible with a complete provirus. The use of *env* subprobe with the *SacI* digests did not reveal any deleted proviruses in the *env* region for the LJ, the Y11 and the R<sup>+</sup> strains. In the Fayoumi strain, one *SacI* fragment of 4.2 kb did not hybridize with the *env* subprobe and could not be associated with a *Bam*HI fragment.

Table II. Distribution pattern of endogenous viral genes (*ev*) identified by the *SacI* restriction fragments hybridizing with RAV-2.

Line code	Average*	Minimum	Maximum	Total <i>ev</i> number	Mean <i>ev</i> frequency	Mean heterozygosity rate among carriers
X55	5.5 <sup>a</sup>	4	7	9 (8)**	0.61 <sup>a</sup>	0.51
R+	6.7 <sup>d</sup>	4	9	12 (11)**	0.56	0.45
R-	5.7 <sup>c</sup>	3	8	10 (9)**	0.57 <sup>a</sup>	0.48
LJ	3.3 <sup>b</sup>	1	5	9 (9)**	0.42 <sup>ab</sup>	0.76
FA	4 <sup>ab</sup>	0	9	13 (12)**	0.26 <sup>b</sup>	0.84
Y11	8.7 <sup>f</sup>	5	12	23 (13)**	0.37 <sup>b</sup>	0.92

\* Estimated values with the same superscript do not differ significantly (Student's *t*-test); \*\* number of *ev* loci defined by both a *SacI* restriction fragment and a *Bam*HI restriction fragment.



**Fig 2.** Autoradiographs showing familial distributions of *SacI* restriction fragments hybridized with the RAV-2 probe, size in kb is indicated at the edges. a) X55 line (from left to right) *Raoul* Marker (lane R), sire 1 (lane S) and its progeny obtained from 2 dams (lanes D), including 7 chicks and 1 chick respectively. b) LJ line, lane R shows the *Raoul* marker, followed (from left to right) by the families of 2 sires (lanes S) mated to 2 different dams (lanes D), giving 3, 2, 5 and 3 progeny respectively. c) (from left to right) the R<sup>-</sup> line is represented by the family of 1 sire (lane S) mated to 1 dam (lane D) with 3 progeny, the R<sup>+</sup> line is then shown with the family of 1 sire (lane S) mated to 2 dams (lanes D) giving 4 and 5 progeny respectively; lane P shows an unrelated chick from line R<sup>+</sup>. d) (from left to right) the Y11 line is represented by the family of 1 sire (lane S1) mated to 1 dam (lane D) with 5 progeny, the FA line is then shown with the family of 1 sire (lane S2) mated to 2 dams (lanes D) giving 4 and 5 progeny.

**Table III.** ALV-related *ev* genes identified by their specific *SacI* and *BamHI* junction fragments.

Fragment <i>SacI</i>	Size (kb)		Hybridize with gag	Line				Similarities with other <i>ev</i> loci	
	<i>BamHI</i>			X55	R <sup>-</sup> R <sup>+</sup>	LJ	FA		Y11
26	4.2		no			L1		<i>ev6</i> <sup>a</sup>	
25	4.2		no	W1			F1	B1	<i>ev6</i>
25	4		no		R1				<i>ev6</i>
23	7.4		?	W2					
22	12		no			L2			<i>ev9</i> <sup>a</sup>
17	7.8		yes					B2	
16.5	5.1		yes				F2		
16	22		yes				F3		<i>evBL4</i> <sup>b</sup>
16	6.8		yes					B3	<i>evR5</i> <sup>c</sup>
15	12.5		yes					B4	
15	7.8		yes				F4		<i>evb11</i> <sup>d</sup>
15	7.4		yes			L3			
14	16		yes	W3					
14	15		yes		R2				
13	25		yes				F5		
12.5	22		yes					B5	
12	13		no				F6		
11.5	8.4		yes	W4					
11.5	4.5		yes					B6	<i>evR8</i> <sup>c</sup>
10.6	25		yes		R3				<i>ev18</i> <sup>a</sup>
10.5	28 or 21		yes	W5					
10.5	4.8		yes				F7		
10.4	5.6		yes			L4			
10.2	10.5		yes		R4				
9.8	15		yes	W6					
9.8	5.2		yes			L5			<i>ev1</i> <sup>a</sup>
9.4	5		yes		R5				
9.2	11.5		yes					B7	
9.2	5.8		yes		R6				<i>evb16</i> <sup>d</sup>
9	17.5		yes	W7					
8.6	18		yes			L6			
8.6	13		yes				F8		
8.5	6.4		yes					B8	
8.4	25		yes					B9	
8.4	7.7		yes			L7			<i>evnewA</i> <sup>e</sup>
8.2	8		yes					B10	<i>evR19</i> <sup>c</sup> <i>evb18</i> <sup>d</sup>
8	14.5		yes		R7				
8	7		yes				F9		
8	6		yes			L8			
7.8	11		yes					B11	
7.8	6.2		yes				F10		
7.6	13		yes		R8				<i>evnewB</i> <sup>e</sup>
7.5	15		yes					B12	<i>evR20</i> <sup>c</sup>
7.4	28 or 21		yes	W8					
7.4	6		yes		R9				
7.4	5.2		yes				F11		
7.2	22		yes		R10				
6.2	7.2		yes			L9			<i>ev3</i> <sup>a</sup>
6	8.4		yes				F12		<i>ev2</i> <sup>a,f</sup>
4.4	3.8		yes		R11			B13	<i>evmt1</i> <sup>g</sup> <i>evb24</i> <sup>d</sup>

<sup>a</sup> Smith, 1987; <sup>b</sup> Ronfort *et al*, 1991; <sup>c</sup> Aarts *et al*, 1992a; <sup>d</sup> Boulliou *et al*, 1991; <sup>e</sup> Kühnlein *et al*, 1989a; <sup>f</sup> Lamont *et al*, 1992; <sup>g</sup> Sabour *et al*, 1992.

The LJ line contained 5 *ev* genes previously described in White Leghorn, *ev1*, *ev3*, *ev6*, *ev9* and *evnewA* and 4 did not correspond to the usual nomenclature. In the other lines, analogy with *ev* loci previously described was suggested by comparing sizes of individual fragments. In addition to *ev6*, similarities could be found for 4 loci among 12 in the R<sup>-</sup> and R<sup>+</sup> lines, for 3 loci among 13 in the FA line and 5 among 23 in the Y11 line. The similarities found for the Y11 line involved *ev* loci that were identified previously in broiler lines. However, the similarities suggested in R<sup>-</sup> and R<sup>+</sup> involved loci previously identified either in White Leghorns (*ev18/R3*, *evnewB/R8*) or in broilers (*evb16/R6*, *evb24/R11*). The FA line harboured loci similar to those in an inbred Fayoumi, White Leghorn (*ev2/F12*), Brown Leghorn (*evBL4/F3*) and commercial broilers (*evb11/F4*). The X55 line did not share any *ev* insertion other than *ev6* with the other lines. A total of 35 of the *ev* loci identified in *SacI* and *BamHI* restriction digests have not previously been described in any other population.

## DISCUSSION

### *Distribution pattern*

The *ev* gene restriction fragment length patterns were strain-specific. Furthermore, there were only a small number of loci shared by different lines. This confirms the broad genetic diversity of ALV-related *ev* genes. The distribution patterns found in X55, R<sup>+</sup> and R<sup>-</sup> suggested that relatively high *ev* gene frequencies could be a consequence of limited population size and closed pedigree matings for many generations. The lack of fixation of *ev* genes described in some commercial layer populations (Iraqi *et al*, 1991) should not be considered as a general rule. Indeed, there was a large variation in *ev* gene frequency in commercial broilers, suggesting that these loci could often be selectively neutral (Boulliou *et al*, 1991).

The large number of *ev* loci in the Y11 line is consistent with its synthetic origin and previous results obtained with broilers (Boulliou *et al*, 1991; Sabour *et al*, 1992). The Fayoumi line presented a characteristic pattern with a very high degree of polymorphism between birds including some birds with no *ev* insertion; this was different from the pattern found in an inbred Fayoumi (Lamont *et al*, 1992). Birds free of *ev* genes have already been described in a White Leghorn background (Astrin *et al*, 1979; Gavora *et al*, 1989).

The heterozygosity rate (*H*) among carriers of *ev* genes appeared to be inversely correlated with the estimated degree of inbreeding, except for the X55 line which showed low heterozygosity and little inbreeding. However, our scoring method tended to underestimate *H*, particularly when the *ev* frequency was high, which was the case for X55. Inbreeding might also have been underestimated due to incomplete pedigree information because this line was probably kept closed for more than the 19 generations recorded. The use of *ev* loci to measure inbreeding in the domestic fowl has previously been proposed by Bumstead *et al* (1987) who assessed the average heterozygosity of *ev* loci in the entire population using several assumptions, including the Hardy-Weinberg equilibrium for *ev* loci. This method produces high inbreeding values when there is a high proportion of non-carriers, because they are homozygous for the null allele. This was not at all in agreement

with our data where, for instance, lines Y11 and FA showed a high proportion of non-carriers for most *ev* loci but were far from being inbred. Consequently, the method proposed by Bumstead *et al* (1987), did not apply to our data because the assumptions involved did not seem to be fulfilled.

### Characterization of *ev* loci

According to the size of the restriction fragment, the restriction fragment length polymorphism (RFLP) methodology does not necessarily provide a very accurate discrimination between *ev* loci. Furthermore, the number of *SacI* fragments may be an underestimation of the number of *ev* genes because of the presence of comigrating fragments (Iraqi *et al*, 1991). This possibility was supported by the discovery of a slightly larger number of *BamHI* than *SacI* fragments within each line. For instance, additional *BamHI* fragments were found in the X55 line that could not be associated with a *SacI* fragment, although frequent within some families. This also suggests that the true rate of heterozygosity might be higher than the value estimated from the *SacI* banding pattern.

The paired segregation of *SacI* and *BamHI* fragments is necessary but not always sufficient to identify *ev* loci. The analysis of reference *ev* samples helped to recognize certain *ev* genes previously characterized in White Leghorns. Unfortunately, such reference samples are not available for other strains. Loci that have one common restriction fragment could be alleles but allelism has been found to be very rare for ALV-related *ev* genes (Aarts *et al*, 1992a).

Locus-specific probes could be used to compare *ev* insertion sites and heterozygosity levels between strains. Few such probes are currently available because they require cloning of flanking regions from genomic DNA of birds carrying only one *ev* insertion, and such birds are rare. An *ev1*-specific probe showed that the genomic insertion site for *ev1* was not occupied by a viral insert in Fayoumi chickens although they carried the restriction fragments usually associated with an *ev1* occupied site in White Leghorns, *ie* 9.5 kb with *SacI* and 5.2 kb with *BamHI* (Lamont *et al*, 1992). A similar approach revealed a much lower frequency of occupation of the *ev1* site in White Plymouth Rock chickens than in White Leghorns (Aarts *et al*, 1991). The polymerase chain reaction has recently been proposed as a method to identify ALV-related *ev* genes in a locus-specific way (Benkel *et al*, 1992). Since *ev1* was the first locus where this procedure was successful, we have used this approach for a small sample of birds representing the different strains. Preliminary results indicate that the *ev1* insertion site was unoccupied in all the strains except the White Leghorn (data not shown).

Characterization of structural deletions could provide more information but it is not easy if several *ev* genes are present in each bird. A combination of several enzymes together with viral subprobes is necessary, as shown by Boulliou *et al*, (1991), but was beyond the scope of our study. For instance, the description of the Fayoumi *ev* locus designated by F1 (table III) included an *EcoRI* digestion which yielded a 6.8 kb fragment compatible with the pattern obtained for *ev6* in White Leghorn (Hughes *et al*, 1981). Furthermore, the finding of a deletion in the *gag* region was useful to confirm the identification of *ev6* in all the strains. Deletions within the *gag* region were found to be rare in a survey of several strains (Aarts

*et al*, 1991) but more frequent in White or Brown Leghorns (Hughes *et al*, 1981; Gudkov *et al*, 1986; Ronfort *et al*, 1991). Further studies will be necessary for a complete description of the new *ev* loci identified here.

### **Interpretation of *ev* loci polymorphism**

Differences between *ev* gene structure and distribution pattern can result from differences in genetic origin or selection procedure or both. The large differences found between randombred lines in this study suggest that selection pressure was not the main cause for the observed variation. However, selection effects can not be completely ruled out because they may have occurred before the start of experimental recording. The number of *ev* loci shared between lines correlates with phylogenetic distance between lines (Aarts *et al*, 1992b). We were unable to calculate phylogenetic distances because the number of shared *ev* loci was too small. The strain-specificity of *ev* loci was consistent with that observed in many other studies and could indicate that new retroviral insertions took place after the creation of specialized chicken breeds. This could be the case for the Egyptian Fayoumi although there were possible homologies with 3 loci from other strains (F3, F4 and F12 in table III). New *ev* loci do not seem to appear very often as shown in the R<sup>-</sup> and R<sup>+</sup> lines which are separated by 16 generations and differed only by 2 out of 12 loci. This difference could be due to genetic drift or selection effects.

The degree of diversity was such that the few *ev* genes that appear to be shared between different genetic origins deserve special attention. The finding of *ev6* in all the lines was remarkable. It must be noticed that *ev6* has also been found in other populations such as the Brown Leghorn (Ronfort *et al*, 1991), Rhode Island Red layers (Iraqi *et al*, 1991) and commercial broilers (Boulliou *et al*, 1991). This could suggest that *ev6* is a very old retroviral insertion. Furthermore, *ev6* is expressed, and is responsible for the production of the endogenous envelope glycoproteins in many tissues. These proteins induce protection from infection with subgroup E viral particles in White Leghorn chickens (Robinson *et al*, 1981). This observation suggested that *ev* genes expressing envelope antigens, such as *ev6*, may confer a selective advantage (Robinson *et al*, 1981), which could explain the ubiquity of *ev6*. However, *ev6* was found to be associated with higher susceptibility to Marek's disease (Kuhnlein *et al*, 1989b) and lower immunoresponsiveness (Lamont *et al*, 1992). The presence of *ev6* in the Fayoumi line at a high frequency (0.56) could be of interest because this strain is said to be resistant to diseases and particularly to Marek's disease. Thus the Fayoumi line may have developed novel mechanisms of resistance.

In conclusion, ALV-related *ev* genes may be of use in discriminating between more or less distantly related strains but cannot provide reliable estimates of genetic distance because of the dissimilarity between *ev* insertions of different strains. However, the comparison of *ev* loci between strains could provide insight into the total number of possible sites for insertion of these retroviruses in the chicken genome. Information from many populations may help to understand the role of these genes in genome evolution. An agreement on gene nomenclature will be necessary for comparison of results from different studies and strains, which will also require exchange of reference DNA samples and probes or sequences. Within a

given population, *ev* loci could be associated with effects on performances related to the expression of the viral genome (Gavora *et al.*, 1991) or to linkage with production trait genes for silent *ev* genes (Kuhnlein *et al.*, 1989b). Thus, endogenous viral genes can be proposed for mapping studies (Crittenden, 1991).

## ACKNOWLEDGMENTS

The advise and technical expertise of JP Furet, C Grohs and C Printz are gratefully acknowledged. This work was supported by grant 88R702 from the French Ministry of Research and Technology and was included in the research programme Prodige-Agrotech launched by INRA.

## REFERENCES

- Aarts HJM, van der Hulst-van Arkel MC, Beuving G, Leenstra FR (1991) Variations in endogenous viral (*ev*)-gene patterns in White Leghorn, medium heavy, White Plymouth Rock and Cornish type chickens. *Poult Sci* 70, 1281-1286
- Aarts HJM, van der Hulst-van Arkel MC, Leenstra FR (1992a) Different endogenous viral loci in Cornish and White Plymouth Rock chickens. *Theor Appl Genet* 85, 325-330
- Aarts HJM, van der Hulst-van Arkel MC, Leenstra FR (1992b) Are endogenous viral sequences recently integrated within the chicken genome? In *XIX World's Poultry Congress*, Amsterdam, 20-24 September 1992, Ponsen & Looijen, Wageningen, Vol 1, 509-513
- Astrin SM (1978) Endogenous viral genes of the White Leghorn chicken: common site of residence and sites associated with specific phenotypes of viral gene expression. *Proc Natl Acad Sci USA* 75, 5941-5945
- Astrin SM, Buss EG, Hayward WS (1979) Endogenous viral genes are non-essential in the chicken. *Nature (Lond)* 282, 339-341
- Benkel BF, Mucha J, Gavora JS (1992) A new diagnostic method for the detection of endogenous Rous-associated virus-type provirus in chickens. *Poult Sci* 71, 1520-1526
- Bordas A, Tixier-Boichard M, Mérat P (1992) Direct and correlated responses to divergent selection for residual food intake in Rhode Island Red laying hens. *Br Poult Sci* 33, 741-754
- Boulliou A, Le Pennec JP, Hubert G, Donal R, Smiley M (1991) Restriction fragment length polymorphism analysis of endogenous avian leukosis viral loci: determination of frequencies in commercial broiler lines. *Poult Sci* 70, 1287-1296
- Bumstead N, Messer LI, Greenwood NG (1987) Use of *ev* loci as a measure of inbreeding in domestic fowls. *Br Poult Sci* 28, 717-725
- Crittenden LB (1991) Retroviral elements in the genome of the chicken: implications for poultry genetics and breeding. *Crit Rev Poultry Biol* 3, 73-109
- Frisby DP, Weiss RA, Roussel M, Stehelin D (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

- Gavora JS, Kuhnlein U, Spencer JL (1989) Absence of endogenous viral genes in an inbred line of Leghorn chickens selected for high egg production and Marek's disease resistance. *J Anim Breed Genet* 106, 217-224
- Gavora JS, Kuhnlein U, Crittenden LB, Spencer JL, Sabour MP (1991) Endogenous viral genes: association with reduced egg production rate and egg size in White Leghorns. *Poult Sci* 70, 618-623
- Gudkov AV, Obukh IB, Serov SM, Naroditsky BS (1981) Variety of endogenous proviruses in the genomes of chickens of different breeds. *J Gen Virol* 57, 85-94
- Gudkov AV, Korec E, Chernov MV, Tikhonenko AT, Obukh IB, Hlozanek I (1986) Genetic structure of the endogenous proviruses and expression of the *gag* gene in Brown Leghorn chickens. *Folia Biol (Praha)* 32, 65-72
- Hughes SH, Toyoshima K, Bishop JM, Varmus HE (1981) Organization of the endogenous proviruses of chickens: implications for origin and expression. *Virology* 108, 189-207
- Iraqi F, Soller M, Beckmann JS (1991) Distribution of endogenous viruses in some commercial chicken layer populations. *Poult Sci* 70, 665-679
- Kühnlein U, Gavora JS, Spencer JL, Bernon DE, Sabour M (1989a) Incidence of endogenous viral genes in two strains of White Leghorn chickens selected for egg production and susceptibility or resistance to Marek's disease. *Theor Appl Genet* 77, 26-32.
- Kühnlein U, Sabour M, Gavora JS, Fairfull RW, Bernon DE (1989b) Influence of selection for egg production and Marek's disease resistance on the incidence of endogenous viral genes in White Leghorns. *Poult Sci* 68, 1161-1167
- Lamont SJ, Chen Y, Aarts HJM, van der Hulst-van Arkel MC, Beuving G, Leenstra FR (1992). Endogenous viral genes in 13 highly inbred chicken lines and in lines selected for immune response traits. *Poult Sci* 71, 530-538
- Maniatis T, Fritsch EF, Sambrook J (1982) *Molecular cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- Mérat P, Bordas A (1982) Etude des particularités de la poule Fayoumi I. Performances de ponte en cages individuelles à deux températures. *Ann Génét Sél Anim* 14, 241-244
- Ricard F (1989) Selecting broiler chickens for improved carcass quality. In: *IX European Symposium of the World Poultry Science Association on Poultry Meat*. Stuttgart 22-25 August 1989 (Scholtyssek S, ed), Eugen Ulmer, Stuttgart, 13-16
- Robinson HL, Astrin SM, Senior AM, Salazar FH (1981). Host susceptibility to endogenous viruses: defective, glycoprotein-expressing proviruses interfere with infections. *J Virol* 40, 745-751
- Ronfort C, Afanassieff M, Chebloune Y, Dambrine G, Nigon VM, Verdier G (1991) Identification and structure analysis of endogenous proviral sequences in a Brown Leghorn chicken strain. *Poult Sci* 70, 2161-2175
- Rovigatti UG, Astrin SM (1983) Avian endogenous viral genes. *Curr Top Microbio Immunol* 103, 1-21
- Sabour MP, Chambers JR, Grunder AA, Kühnlein U, Gavora JS (1992) Endogenous viral gene distribution in populations of meat-type chickens. *Poult Sci* 71, 1259-1270
- Smith EJ (1987) Endogenous avian leukemia viruses. In: *Avian Leukosis* (de Boer GF, ed). Nijhof, Boston, 101-120