Evaluation of Genetic Variation Within and Between Different Chicken Lines by DNA Fingerprinting

S. Ponsuksili, K. Wimmers, and P. Horst

A DNA fingerprinting technique was used to assess the extent of genetic variation within and between 12 lines of chickens of different origin. DNA fingerprints were obtained with the restriction endonucleases AluI and HinfI and with the oligonucleotide probes (CA)$_n$, (CAC)$_m$, (GGAT)$_4$, and (GACA)$_4$. Different methods that are based on band sharing and band frequency were tested to estimate genetic variation within and between populations. The genetic variability was significantly different within the populations. It was highest in a commercial broiler line and lowest in a White Leghorn inbred line. All three measures of genetic variation within populations provided results that were highly correlated ($r = 0.89$–$0.97$). A lower degree of correspondence was observed for the estimation of genetic variation between populations by three different methods. This is shown by differences of the phylogenetic trees and by coefficients of correlation between these measures ranging from $r = 0.51$–$0.96$.

Genetic variation, both within and between breeds, is essential for the genetic improvement of domestic animals. Loss of variation will restrict the selection for desirable economic characteristics within current commercial lines. Therefore poultry breeders have an interest in maintaining or increasing genetic variation within and between commercial lines or exotic populations. Molecular markers, such as variable number of tandem repeats (VNTRs) in the form of mini- or microsatellites, allow the assessment of genetic variability among genotypes at the DNA level. Polymorphism at the DNA level is greater than that at the level of gene products and it is detectable independently of environment, age, sex, and tissue. In 1985 multilocus DNA fingerprinting was first developed (Jeffreys et al. 1985) and has been widely used for studies on the genetic structure of natural and domesticated populations. In poultry, DNA fingerprints have been used (1) to characterize individuals or populations genotypically, (2) to study the relative contribution of evolutionary forces to genetic differences between populations, (3) for marker assisted selection, (4) to assist in gene introgression, and (5) to predict hybrid vigor (Hillel et al. 1992). The advantage of DNA fingerprinting is that many widely dispersed VNTR loci are detected simultaneously, even where traditional markers fail to detect variation. But population data on DNA fingerprinting do not provide a complete characterization of the genetic variation in terms of allele frequency distributions, since neither the number of loci nor the locus affiliation of alleles is directly observable. Nevertheless, DNA fingerprinting has been shown to be a cost-effective method to detect hypervariable loci for investigating genetic relationship within and between populations (Dunnington et al. 1994; Gilbert et al. 1990, 1991; Grunder et al. 1994; Jin and Chakraborty 1994; Kuhnlein et al. 1989, 1990; Siegel et al. 1992).

In this study high-yielding, commercially produced chickens whose performance has been changed by selection for high production were compared to local chickens, which still have a great variability in morphological characteristics, and to exotic chickens, which are kept in small flocks in zoos or other institutes. The oligonucleotide probes (CA)$_m$, (CAC)$_m$, (GGAT)$_4$, and (GACA)$_4$, which have been shown to provide informative DNA fingerprints of chickens (Wimmers et al. 1992), were used to assess the present status of genetic variation within and between such different lines or origins. Furthermore, a comparison of different formulas to estimate the level of genetic variation within and between populations was conducted.
Materials and Methods

Chickens
The study was done with 153 individuals from 12 different lines of chickens which included an inbred line, commercial lines, and exotic breeds. These populations are the White Leghorn inbred line ETH77 (IBL, \( n = 13 \)), three commercial, high-yielding strains, namely a broiler male strain (BRO, \( n = 14 \)), a Rhode Island Red layer line (RIR, \( n = 14 \)), and a White Leghorn layer line (LEG, \( n = 12 \)), and eight exotic breeds, namely Bankiva (BAN, \( n = 14 \)), Dandrawi (DAN, \( n = 15 \)), Fayoumi (FAU, \( n = 13 \)), Kadakanath (KAD, \( n = 15 \)), Nunu-kan (Nun, \( n = 14 \)), Silky (SIL, \( n = 15 \)), Taiwan White-darmeat broiler line (TWW, \( n = 7 \)), and Taiwan Brown-darmeat broiler line (TWB, \( n = 7 \)) (Ponsukilii 1995).

Methods to Produce DNA Fingerprints
DNA was prepared from blood samples by proteinase K digestion and phenol chloroform extraction. The restriction endonucleases Alul or Hinfl were used to digest 10 \( \mu \)g of individual DNA samples or pooled DNA for each banding pattern. The DNA fragments were separated in 0.7% agarose gels (1 vol/cm for 40 h) and transferred to nylon membranes. In order to produce the DNA fingerprint patterns hybridization was performed with the oligonucleotide probes (CAC)\(_5\), (GGAT)\(_4\), (CA)\(_8\), and (GACA)\(_4\) which gels (1 vol/cm for 40 h) and transferred to fragments were separated in 0.7% agarose and the White Leghorn layer line (RIR, \( n = 14 \)), and a White Leghorn layer line (LEG, \( n = 12 \)), and eight exotic breeds, namely Bankiva (BAN, \( n = 14 \)), Dandrawi (DAN, \( n = 15 \)), Fayoumi (FAU, \( n = 13 \)), Kadakanath (KAD, \( n = 15 \)), Nunu-kan (Nun, \( n = 14 \)), Silky (SIL, \( n = 15 \)), Taiwan White-darmeat broiler line (TWW, \( n = 7 \)), and Taiwan Brown-darmeat broiler line (TWB, \( n = 7 \)) (Ponsukilii 1995).

Assessing Genetic Variability Within and Between the Populations
Two types of DNA fingerprints were made. In one case individual DNA samples were used and in the other case DNA mixes were made using equal amounts of DNA from each animal within each line (Figure 1a,b). DNA fingerprints of individual DNA samples were used to determine the band sharing degree and the band frequencies within populations. Pooled DNA from different lines were used to produce DNA fingerprint patterns, which were representative for each of the populations in this study. This approach allowed comparison of the patterns of many populations using only a few DNA fingerprints and determination of the degree of band sharing between populations. The frequency of each band in the representative banding patterns of pooled DNA was taken from individual DNA fingerprints. The DNA fingerprints were repeated several times with different order of the samples to make sure that each banding pattern has been compared with another in close proximity and to check the reproducibility of the banding patterns.

Evaluation of DNA Fingerprints
Bands representing DNA fragments longer than 2 kb were scored by two reviewers. Banding patterns were compared between nearby lanes to classify shared and non-shared bands. Bands were regarded as nonshared if they differed in their position more than the half of the band width and/or if the ratio of the intensities was less than 1.2. Band sharing (BS) was calculated as BS(\%) = (2N\_x\_y/N\_x + N\_y) \times 100\%, where N\_x\_y is the total number of shared bands of the individuals x and y; and N\_x and N\_y are the number of bands in the banding patterns of the individuals x and y (Jeffreys et al. 1985). Based on BS, the probability (\( P \)) that two randomly selected but unrelated individuals show identical banding patterns was calculated as \( P = (1 - 2bs^2 + 2bs^2\times 100\%\), where bs is the average band sharing (BS/100) and N is the average number of bands (Jeffreys and Morton 1987). The number of loci (\( L \)) detected in the DNA fingerprints was estimated based on BS as \( L = N(4 - bs)/[4(2 - bs)] \) (Lynch 1990).

Measures of Genetic Variation Within Populations
The estimation of genetic variation within populations by DNA fingerprinting was based on band sharing or band frequency. All values were averaged over all individuals in each population and over all eight banding patterns produced by different combinations of enzymes and probes. The values are given as arithmetic means (\( x \)) with standard deviations (SD).

Based on band sharing, the average individual band difference (APDI) was estimated as APDI = 100 - APS, where APS is the average percentage degree of BS.

Based on band frequency, the first value was the additive inverse of mean band frequency called average variation of bands (AVB). It was calculated using the equation \( V = 1 - 1/n(\Sigma v_i) \), where \( v_i \) is the frequency of band i and n is the number of bands scored (Kuhnlein et al. 1989). The second estimator calculated with band frequency was the heterozygosity according to the equation \( H = (2n/2n - 1) \) \((\Sigma S_k/A - (1 - S_k)^{0.5}) - 1\), where n is the number of individuals sampled, \( S_k \) is the frequency of the \( k \)th fragment in the population, and A equals the total number of fragments (Stephens et al. 1992).

Measures of Genetic Variation Between Populations
The genetic variation between the populations was also estimated based on band sharing and band frequency values averaged over all eight banding patterns obtained with different combinations of enzymes and probes.

Based on band sharing between the representative banding patterns the average band difference (APDI) was calculated in the same manner as within-population comparisons.

The second estimator was computed according to the equation \( DL = -\ln(\text{BS}_{xy}/\text{BS}_x\_\text{BS}_y) \), where DL is the genetic distance, BS\(_{xy}\) is the average band sharing between population x and y, and BS\(_x\) and BS\(_y\) are the average band sharing within the populations x and y, respectively (Lynch 1991). Based on average band frequency a third estimator for genetic divergence was calculated by \( DK = -\ln(\bar{f}) \). Here \( f \) is the genetic identity index given by \( f = 1/\Sigma N_i(2v_i^2 + v_i) \), where \( N_i \) is the number of different bands scored in the two populations and \( v_i \) are the frequencies of band i in population x and y (Kuhnlein et al. 1989). The results of these estimations were used to construct genetic trees by UPGMA cluster analyses (unweighted pair-group method using arithmetic averages), a standard algorithm for agglomerative clustering where the distance between two clusters is the average distance between pairs of observations, one in each cluster. By this procedure the input matrix is screened for the pair of objects that are most similar. These objects are merged into a new cluster and the matrix is updated to reflect the deletion of the pair of objects. Then the procedure is repeated with the updated matrix (Rohlfs 1993; SAS 1987).

Statistical Analyses
The statistical analyses were performed using the general linear models (GLM) procedure of the SAS statistical package (SAS 1987). Significance of differences between means was tested by the use of the Duncan range test of the GLM procedure. Coefficients of correlation between the three measures of genetic variation within populations—APDI, AVB, and H—were calculated using the correlation procedure
variables
probes, and restriction enzymes
different strains of chicken, oligonucleotide deviations (SD) of DNA fingerprints for Table 1. Mean values (x, P)

Means with different superscripts differ significantly at P < .05.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Band sharing (%) (x ± SD) (x ± SD)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alu</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BRD</td>
<td>55.2 ± 17.3a</td>
<td>24.2 ± 7.4a</td>
</tr>
<tr>
<td>KAD</td>
<td>55.3 ± 17.9a</td>
<td>45.1 ± 6.4a</td>
</tr>
<tr>
<td>TWB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NUN</td>
<td>58.0 ± 17.4a</td>
<td>47.3 ± 5.8a</td>
</tr>
<tr>
<td>RIR</td>
<td>55.0 ± 17.6a</td>
<td>47.2 ± 5.5a</td>
</tr>
<tr>
<td>LEG</td>
<td>55.9 ± 18.4a</td>
<td>35.0 ± 4.9a</td>
</tr>
<tr>
<td>DAN</td>
<td>56.4 ± 17.0a</td>
<td>45.1 ± 5.7a</td>
</tr>
<tr>
<td>IBL</td>
<td>59.3 ± 18.8a</td>
<td>41.3 ± 8.43a</td>
</tr>
</tbody>
</table>

Table 1. Mean values (x) and standard deviations (SD) of DNA fingerprint data for different strains of chicken, oligonucleotide probes, and restriction enzymes

<table>
<thead>
<tr>
<th>Inbred lines</th>
<th>Size of fragments (kb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I and Alu</td>
<td>8</td>
</tr>
<tr>
<td>I and HindIII</td>
<td>6</td>
</tr>
</tbody>
</table>

Results

General Results of the DNA Fingerprints

The banding patterns were absolutely reproducible and contained about 40 bands. Mean numbers of bands calculated for each population, for each restriction enzyme, and for each oligonucleotide probe (COR) of the SAS statistical package. In order to compare the distance matrices based on the three measures of genetic variation between populations—APDI, DL, and DK—matrix correlation coefficients were calculated using the product moment correction (r) and the Mantel test statistic. The computations were performed with the appropriate procedures of the software package NTSYS-pc (Rohlf 1993). This test for matrix correspondence gives the product moment correlation (r) and a statistic test (z) to measure the degree of relationship between two matrices. Rohlf (1993) suggested that the degree of fit can be interpreted as follows: r = 0.9, very good fit; 0.8 ≤ r < 0.9, good fit; 0.7 ≤ r < 0.8, poor fit; r < 0.7, very poor fit.

Genetic Variation Within Populations

Different degrees of genetic variation within the populations were characterized by APDI, AVB, and H. These three measures of genetic variation among populations are given in Table 2.

The inbred line (IBL) had the lowest APDI of only 6.5%. Considering the group of commercial lines, BRD had the highest APDI, whereas RIR and LEG had intermedi-
and IBL had the lowest value of heterozygosity. The correlation coefficient between AVB and heterozygosity was 0.97. The degree of correspondence of the estimators of the genetic relationships between the populations is reflected in the correlation between the results based on the three measurements used. The APD values averaged over all eight combinations of enzymes and probes were correlated with the DL values by $r = 0.60$, which is significant ($P < .01$). The corresponding value for the comparison between APD values and DK values was highly significant with $r = 0.96$ ($P < .01$). The lowest correlation was found between DK and DL with $r = 0.51$.

**Discussion**

The populations used in this study were diverse in their genetic makeup, representing commercial lines, a laboratory inbred line, and a broad range of local populations.

The three examples of commercial lines used in this study represent stocks used in industrialized production of meat and eggs. The population of IBL (ETH77) was founded from different White Leghorn lines and selected for small early egg weight (Somes 1988). Bankiva (BAN) is one of the subspecies of jungle fowl, *G. gallus bankiva* belongs to the group of red jungle fowl, *G. gallus* (Crawford 1990) representing the ancestor of the domestic fowl. Thus it is of special interest to investigate the relationship of this breed to other breeds of different origins and breeding status. The exotic local populations used include a broad variety of populations of different origins, breeding history, and purpose, including those with special morphological and physiological characteristics like black bone/black meat and slow feathering. Many of these may represent important genetic resources. To deal with this question the determination of the genetic distances of these lines is important, since genetic resources are those populations which show the highest genetic differentiation within a species and/or show unique alleles or allelic combinations (FAO 1993). Furthermore these populations represent a useful model case to evaluate the suitability of DNA fingerprints to detect genetic divergence and to measure genetic variability because of the great differences in their origin and breeding history. It should be mentioned however that it is quite problematic to investi-
Genetic Variation in Chicken

Ponsuksili et al. • Genetic Variation in Chicken 21

Figure 2. The average percentage difference (APD), genetic distance, DK, and genetic distance, DL, between different strains of chicken. Mean values (± SD) from eight combinations of two restriction enzymes (Alu, Hinfl) and four probes [(GGAT)₄, (CAC)₅, (GACA)₄, (CA)₈].

Genetic Variation Within Populations

The three measures—APD, H, and AVB—that were used to study genetic variation within populations were highly correlated. A close relationship between APD and H (r = 0.95) was also found by Stephens et al. (1992) when using cDNA probes in RFLP analyses. The correlation coefficient between H and AVB was the highest (r = 0.97) in our study. This is due to the fact that both measures are based on band frequency. In contrast Zhu et al. (1996) found a higher correlation of APD and H values than of APD and AVD or AVD and H, when using Jeffreys’ probe 33.6 (Jeffreys et al. 1985) and HaeIII to produce DNA fingerprints from turkey. The difference in these findings may result from differences in the evaluation of DNA fingerprints (see below). Our results indicate that all three indices provide estimators that are comparable. Therefore using only one method should be sufficient to estimate the genetic variability within populations.

The comparison of the three measures revealed higher values for APD, than for AVB and H. Furthermore APD showed higher standard deviations. The reason for such differences might be that the band-sharing data were estimated by comparing banding patterns of pairs of individuals. Thus one DNA fingerprint provided 30 band sharing values. These values represent the variation of genetic similarity of individuals. The data presented are means of all pairs of individuals compared and the results of different combinations of enzymes and probes. But one DNA fingerprint provided only one value of AVB and H. This value represents all individuals in the DNA fingerprints, that is, all individuals of a population. The data given here are means of the results of different combinations of enzymes and probes.

Concerning the absolute level of genetic variability within the populations, the values given by APD, and H met to a very high extent our expectations based on the knowledge of the breeding history of the
The Journal of Heredity 1998:89(1)

Figure 3. Dendograms of 12 strains of chicken revealed by UPGMA cluster analysis based on (a) average percent difference, $APD_e$, (b) genetic distance, $DK$, and (c) genetic distance, $DL$.

populations and based on estimations using other DNA markers. The values given by AVB were estimates that were biased slightly downward.

All three methods provided acceptable estimates of the genetic variability within the populations regardless of the level of variability. The relative differences between the values of genetic variability of the populations were similar for all methods. All three methods discriminate between groups of birds which were originated from large outbreeding flocks and/or which were sampled at different locations some generations ago in order to be kept at our institute (BRO, KAD, TWB, TWW, NUN, RIR, and LEG) and groups of chickens which were kept in small flocks for many generations or were inbred (DAN, SIL, BAN, FAU, and IBL). However, of the three tested methods $APD_e$ is the best index to estimate genetic variation within populations. The determination of the degree of band sharing is more reliable than of band frequency, because it is based on the comparison of banding patterns situated in close proximity, whereas the determination of band frequencies requires the comparison of all lanes from one end of the DNA fingerprint to the other. Furthermore, the estimator $APD_e$ is the most sensitive one. It provides the highest number of significant differences of variability within the populations.

Genetic Distance Between Populations
The genetic distance between the 12 populations was also evaluated using three different indices—$APD_e$, $DK$, and $DL$. The $APD_e$ between different strains of chicken used here revealed low variation between different pairwise comparisons. The second method to estimate genetic distance ($DK$) was based on band frequency. Kuhnlein et al. (1989) reported on the genetic distances measured as $DK$ between meat lines and different White Leghorn lines. The highest distance was found between a White Leghorn inbred line and a meat line ($DK = 2.87$). In this study $DK$ between BRO and IBL was 2.66. The relative levels of $DK$ values between meat strains and White Leghorn or within different selected lines of White Leghorn that were found in Kuhnlein’s study and in this study were similar. The few differences that were found may be caused by the use of different combinations of enzymes and probes and the use of individual and representative DNA fingerprints, respectively. The third method to estimate $DL$ is based on band sharing within and between populations. It was derived in order to take into account differences in the degree of variability within the populations. No references were found for this value in studies of chicken. Reports were only found on sympatric brown trout ($Salmo trutta$) populations (Propöhl et al. 1992) and on turkey (Zhu et al. 1996). Our results suggest that by this measure the influence of the genetic variation within populations on the genetic distance between the populations is overestimated. If the band sharing within certain populations of the pairwise comparison is high, the $DL$ between these populations will also be high.

Cluster analyses of $APD_e$ and $DK$ (Figure 2a,b) showed the same relationship among strains of all nested sets, but cluster analysis of $DL$ showed different nested sets. As the correlation coefficient between $APD_e$ and $DK$ was high ($r = 0.96$) the cluster analysis produced nearly the same dendrogram. In an analysis of relatedness in geese by DNA fingerprinting, Grunder et al. (1994) also found a high correlation ($r = 0.99$) between $APD_e$ and $DK$ when estimating band frequencies based on the optical density of bands obtained from pooled DNA. Correlation was much lower ($r = 0.76$) when band frequencies were derived from DNA fingerprints of six individual DNA samples. Zhu et al. (1996) reported a correlation between $ADPP$ and $DK$ of $r = 0.84$ in an analysis of individual banding patterns of turkey. They used a computer-assisted method to evaluate the DNA fingerprints where both band sharing and band frequency values were based on the same fragment size calculations. In our approach evaluation of DNA fingerprints
was done by visual inspection, taking into account fragment size and intensity. For band sharing only adjacent lanes were compared, but for band frequency all lanes of a DNA fingerprint were compared. This might have caused some identification mistakes and might have led to differences in the identification of different and common bands when determining band sharing and band frequencies, respectively. In this regard our approach, where representative banding patterns of pooled DNA were compared to obtain band sharing but band frequencies were taken from individual patterns, is similar to the assessment of band frequencies from pooled DNA as done by Grunder et al. (1994) and may reveal biased estimates. Therefore we observed a high correlation between APD, and DK comparable to the results of Grunder et al. (1994) when using pooled DNA.

In comparison to Zhu et al. (1996) we found lower correlation between APD, and DK and DL, respectively. This is due to the fact that lines highly differing in their genetic variability were used here in contrast to the other study.

Based on the background knowledge of history, origin, and morphology of the populations investigated here, one comes to the conclusion that the data of APD, and DK provided more informative values than DL, as the dendrograms based on APD, and DK meet our expectations to a higher degree than that based on DL.

In this study, and in agreement with some other publications (Gilbert et al. 1990; Kuhnlein et al. 1989; Siegel et al. 1992), the usefulness of DNA fingerprints for the purpose of estimating genetic distances has successfully been demonstrated with APD, and DK and to a lesser extent with DL. APD, and DK led to the same grouping of the populations. Using DL led to a dendrogram that was quite different from the first two cluster diagrams, especially the position of the populations SIL and RIR. In general it will not be possible to determine which dendrogram is the best until such time as the entire genomic sequence can be compared (FAO 1993), but the fact that APD, and DK results in the same dendrogram suggests that these cluster diagrams represent better pictures of the genetic relationship between the populations. For the same reasons as described for the assessment of genetic variation within populations, $APD_r$—corresponding to APD—is the most favorable measure to estimate genetic distances between populations. When compared to DK, which also reveals realistic estimations of the genetic distances, the advantage of $APD_r$ is that it can be determined using pooled DNA samples.

To summarize, this study suggests the use of the measures APD, or $APD_r$, both based on band sharing, to estimate genetic variability within and between populations by DNA fingerprinting. As an alternative to the technical or statistical limitations of DNA fingerprint approaches, a new class of genetic markers is available. These markers are based on amplification of genomic DNA by polymerase chain reaction (PCR) with locus-specific primers of the simple repetitive sequence or microsatellite sequence. These sequences are currently of great interest in studies of variation at the level of populations because they are highly polymorphic and well distributed. Therefore further investigations are under way to evaluate the usefulness of monolocus microsatellite markers compared to multilocus microsatellite markers for estimating the genetic variability and genetic distance of these chicken populations.

References


Received March 1, 1996

Accepted January 31, 1997

Corresponding Editor: J. James Bitgood

Ponsuksili et al • Genetic Variation in Chicken 23