Theor Appl Genet (1989) 77:271–274

Genetic variability between two breeds based on restriction fragment length polymorphisms (RFLPs) of major histocompatibility complex class I genes in the pig

Y.C. Jung [1], M.F. Rothschild [1] *, M.P. Flanagan [2], E. Pollak [3] and C.M. Warner [2]

Department of Animal Science [1], Department of Biochemistry and Biophysics [2] and Department of Statistics [3], Iowa State University, Ames, IA 50011, USA

Received July 11, 1988; Accepted September 8, 1988
Communicated by L.D. Van Vleck

Summary. Restriction fragment length polymorphism analyses of SLA class I genes were performed on 55 Duroc and 24 Hampshire boars from the 1986-87 national performance tests of each breed. Few boars were inbred. Southern blotting and hybridization procedures were performed on genomic DNA isolated from white blood cells by using Pvu II, Bam HI, and Eco RI endonucleases and a swine MHC class I probe. Genetic variability within and between the two breeds was estimated in terms of nucleotide diversity, by using a mathematical analysis based on the different RFLP patterns. The nucleotide diversity calculated within each breed was less than that between the two breeds. The results from the nucleotide diversity analysis suggested that genetic variability was greater in the Duroc breed than in the Hampshire breed. A relatively high level of genetic variability was shown in the class I major histocompatibility complex genes in the pig.

Key words: RFLP (restriction fragment length polymorphism) – MHC (major histocompatiability complex) – nucleotide diversity – swine

Introduction

The pig major histocompatibility complex (MHC) or swine leukocyte antigen (SLA) complex is comprised of at least three class I, two class II, and three class III loci (Vaiman et al. 1978, 1979; Kirszenbaum et al. 1985; Lie et al. 1987). The total span of the SLA region ranges approximately from 0.8 cM – 1.0 cM (Singer et al. 1983; Vaiman et al. 1988). DNA sequence analysis of the heavy chain of an SLA class I gene from a miniature pig (SLA^d) has revealed that it consists of eight exons spanning 3.6 kb (Satz et al. 1985). The class I gene products, SLA-A, -B, and -C antigens, each consist of a polymorphic heavy chain of approximately 44,000 daltons with a non-polymorphic light chain, β_2-microglobulin, of 12,000 daltons (Chardon et al. 1978; Lunney and Sachs 1978). The number of genes in the SLA class I region has been estimated to be between 6 and 15 (Singer et al. 1982; Chardon et al. 1985; Ehrlich and Singer 1986; Lunney et al. 1986).

Measurement of the variability of different MHC types has been accomplished classically through serological or cellular methods. Recent developments in recombinant DNA techniques has enabled examination of the genetic variation among MHC genes beyond the protein level, that is, in terms of DNA differences based on restriction fragment length polymorphism (RFLP) analysis. Upholt (1977) first attempted to estimate divergence of DNA sequences by analysis of the fraction of conserved fragments. Nei and Li (1979) and Nei and Tajima (1981) justified and refined the models. The average number of nucleotide differences per site between two randomly chosen DNA sequences was termed "nucleotide diversity" and was denoted by π to express the degree of polymorphism in a population at the nucleotide level (Nei and Li 1979).

The purpose of this paper is to report for the first time a mathematical analysis of nucleotide diversity and genetic variability in class I MHC genes in two pig breeds, Durocs and Hampshires, by using RFLP analysis.

Materials and methods

Description of animals

The Duroc breed of swine, which is red in color, originated in the northeastern section of the United States around 1860. The

* To whom correspondence should be addressed

Hampshire breed, which is black in color with a white belt encircling its belly, originated in Kentucky about 1835. The Duroc breed is the more populous of the two breeds and is believed to be genetically more diverse.

During the fall of 1986, blood samples were collected from 70 Duroc boars, ranging from 20–23 weeks of age, at the National Duroc Performance Test in the Pure Genetics Swine Test Station, located in Ida Grove, Iowa. Similarly, in the spring of 1987, blood samples were collected from 40 Hampshire boars, ranging from 15–25 weeks of age, at the National Hampshire Performance Test in the Farmland Agriservices Inc. Swine Test Station, located in Lisbon, Iowa. Complete results, after DNA isolation, restriction enzyme digestion, and Southern blotting, were obtained on a total of 55 Duroc and 24 Hampshire boars. The 55 Duroc boars were from 40 different litters and from 31 farms. The 24 Hampshire boars were from 17 litters and from 14 different farms.

The sophisticated nature of RFLP analysis makes it difficult to process large numbers of blood samples. The decision to sample boars from the national tests of each breed was made because this enabled collection of blood and data from a single location from a representative sample of each breed. Pedigree analysis of the boars from each breed revealed that they were generally unrelated and were representative of the more popular sires and grandsires of the breed. Therefore, despite the low numbers of boars used, these samples were thought to be sufficient to examine the variability in the MHC class I genes.

DNA isolation and gene detection

Ten ml of blood from each boar was obtained by using sterile method to avoid contamination of the samples. The isolation of genomic DNA, partial digestion by Pvu II, Bam HI, and Eco RI endonucleases, and blotting and hybridization with an SLA class I gene probe were performed as described by Flanagan et al. (1988). The porcine class I gene PD1-A (4.3 kb), described by Singer et al. (1983), was used as the porcine genomic DNA probe. Only class I genes were analyzed because no porcine class II and class III probes were available.

Mathematical analysis of RFLP patterns

Molecular sizes of the restriction fragments were determined by using size markers of Hind III cut lambda DNA restriction fragments. Procedures for quantifying the fragments are in Flanagan et al. (1988). A series of intervals for molecular size was established to allocate all the restriction fragments. The differences between intervals were set at 5%-10% to avoid allocations of more than one restriction fragment to an interval. Therefore, the data set of DNA fragments could be converted to a manageable matrix form with dimensions of the number of intervals representing a fragment by the number of animals. The genetic variability within and between breeds was calculated by estimating nucleotide diversity (π) based on restriction fragment length polymorphism patterns. Two methods were used to estimate nucleotide diversity. The first method was to calculate

$$\hat{\pi} \approx 1 - (((\hat{F}^2 + 8\hat{F})^{1/2} - \hat{F})/2)^{1/r} \qquad \text{(Upholt 1977),}$$

where $\hat{F}$ is an expected proportion (estimated) of DNA sequences that remains unchanged, and r is the number of base pairs in the recognition sequence for the restriction enzyme used. The variance was estimated by

$$V(\hat{\pi}) = \hat{\pi}(1 - \hat{\pi})/N \qquad \text{(Upholt 1977),}$$

where N is the number of independent nucleotide positions in the cleavage sites. Nucleotides in common at cleavage sites were counted only once. The second method was to calculate

$$\hat{\pi} = -(2/r)\log_e \hat{P} \qquad \text{(Nei and Li 1979),}$$

where $P = e^{-r\tau t}$ is the probability that an original restriction site remains unchanged by time t when the substitution rate per nucleotide per unit of time is τ, and is estimated by

$$\hat{P} = (\hat{F}(3 - 2\hat{P}_1))^{1/4} \qquad \text{(Nei 1987).}$$

$\hat{P}_1$ is a trial value of $\hat{P}$ and $\hat{P}$ is obtained by iterating until $\hat{P} = \hat{P}_1$ is achieved. The $\hat{F}^{1/4}$ value was used as the first trial value of $\hat{P}_1$. The proportion F was estimated by

$$\hat{F} = n_{ij}/\tilde{n} \qquad \text{(Nei and Li 1979)}$$

for the nucleotide diversity within a population, where n_{ij} is the number of restriction fragments shared between ith and jth RFLP patterns, and $\tilde{n}$ is the average number of restriction fragments of all the members of a population. The estimated fraction of DNA fragments that remain the same between the two breeds was estimated by

$$\hat{F} = 2n_{xy}/(n_x + n_y) \qquad \text{(Nei and Li 1979),}$$

where n_{xy} is the number of restriction fragments shared between two RFLP patterns of two different breeds and, n_x and n_y are the average numbers of restriction fragments for Duroc and Hampshire breeds, respectively. To estimate total average nucleotide diversity (π), the following equation was used:

$$\hat{\pi} = \frac{\sum_{ij} \pi_{ij}}{n_c} \qquad \text{(Nei et al. 1984),}$$

where π_{ij} is the proportion of different nucleotides between the ith and jth DNA sequence and n_c is the total number of comparisons performed. Net nucleotide differences ($\hat{\pi}_n$, those differences between the two breeds after adjusting for within breed variability), were estimated by the following equation:

$$\hat{\pi}_n = \hat{\pi}_{xy} - (\hat{\pi}_x + \hat{\pi}_y)/2 \qquad \text{(Nei and Li 1979),}$$

where $\hat{\pi}_x$ and $\hat{\pi}_y$ are average nucleotide diversity estimates for each breed and $\hat{\pi}_{xy}$ is the average nucleotide diversity estimate between breeds. The variance of nucleotide diversity was estimated as described by Upholt (1977).

Results

An example of RFLP patterns of some Hampshire pigs are shown for Pvu II in Fig. 1. The same type of results were collected and tabulated for the other restriction enzymes (data not shown). Similar results for Duroc boars have been reported in Flanagan et al. (1988). Visual comparison of the RFLP patterns showed differences between and within the breeds. The calculated estimates of nucleotide diversity, based on all three restriction enzymes used, within the Durocs and within the Hampshires, were the same using both methods (Table 1). Likewise, the calculated estimates of nucleotide diversity between breeds from the two methods were very similar. A final calculation, the net nucleotide differences between two breeds (Nei and Li 1979), was made and gave a value of 0.083. It should be noted that this amount is greater than the variability within breeds (Table 1).

Table 1. Comparison of nucleotide diversities calculated by using three restriction enzymes[a] and two different methods

Method	Upholt (1977)		Nei and Li (1979)	
Breed	Duroc[b]	Hampshire[c]	Duroc	Hampshire
Nucleotide diversity within breeds	0.061 ± 0.023	0.043 ± 0.023	0.061 ± 0.023	0.043 ± 0.023
Nucleotide diversity between breeds	0.138 ± 0.030		0.135 ± 0.029	

[a] Restriction enzymes used were PvuII, BamHI and EcoRI
[b] No. of animals = 55
[c] No. of animals = 24

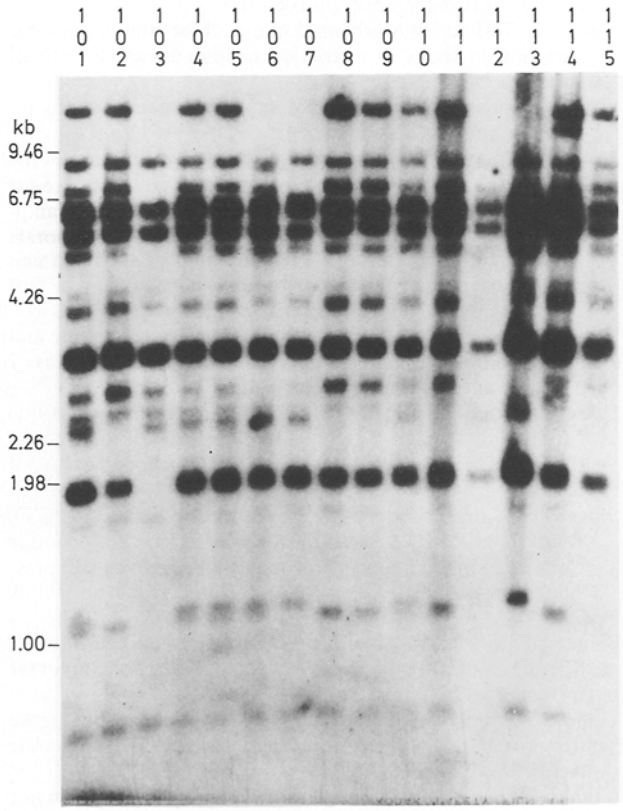

Fig. 1. Example of Southern blot analysis of PvuII digested genomic DNA from Hampshire boars after hybridization with a porcine genomic class I probe, PD1-A. The sizes of the restriction fragments were estimated on the basis of HindIII digested lambda phage DNA fragments

Discussion

This paper reports the use of two methods (Upholt 1977; Nei and Li 1979) to calculate nucleotide diversity within and between two breeds of pigs, Duroc and Hampshires, based on data from MHC class I RFLP patterns on Southern blots. Results from computer simulation (Kaplan 1983) have indicated that estimates by both methods are basically identical. Our calculations (Table 1) agree with this observation. Variability within the Duroc breed was greater than that within the Hampshire breed but not significantly so. These results are expected on the basis of the population size of each breed and pedigree information from both breeds. There is also a significant amount of genetic variability between the Duroc and Hampshire breeds as measured by the net nucleotide diversity estimates. Variability between breeds is expected because the breeds have been separated for some time and the breed societies monitor pedigrees carefully to prevent inadvertent transfer of genes between the breeds.

The values for nucleotide diversity seen here for SLA class I genes reflect the characteristically high degree of polymorphism previously found among MHC genes and other gene families. For example, use of results of actual DNA sequences of IgG2a genes that belong to a large gene family showed an estimated divergence of 0.10 (Schreier et al. 1981). The immunoglobulin genes showed considerably greater divergence than other genes, which ranged from 0.002 to 0.019 (Nei et al. 1984). The diversity seen here for the SLA class I genes is of the same magnitude as that for the imunoglobulin genes.

Measurement errors associated with the fragments, however, can occur. This type of error will tend to increase the estimates of nucleotide diversity. For that reason, more than one enzyme should be used to reduce the error. Three enzymes were used for these calculations.

The methods described in this paper are based upon four assumptions: (1) that the four nucleotides are randomly distributed in DNA sequences, (2) that mutations occur with the same probabilities among the four nucleotides, (3) that mutations are caused solely by nucleotide substitution, and (4) that the population is in equilibrium with respect to genetic drift and mutation and without selection. In reality, it has been reported that the distribution of the four nucleotides is not random in mtDNA (Brown 1980), and the rate of nucleotide substitution varies considerably from gene to gene in a genome and from region to region in a gene (Nei 1975). But Nei and Tajima (1983) and Nei et al. (1984) have claimed that the first two assumptions are not necessary to estimate di-

vergence between a pair of DNA sequences, as long as mutations at different nucleotide sites occur independently and π is small. Furthermore, mutation occures in many forms other than single nucleotide substitutions. However, Nei et al. (1984) have indicated that DNA polymorphisms can be studied by the statistical approach they recommend because other mutational events occur rarely. From the animal geneticist's viewpoint, it is assumed that the poulation of a breed is likely not to be in equilibrium, because it is subject to selection. It is doubtful, however, that this selection is focused at the nucleotide level.

The results reported in this paper suggest that calculations of nucleotide diversity are useful for comparison of populations for genetic variability. Such measures are important in detecting variation and in quantifying the level of variability present.

Acknowledgements. The author thank the Pure Genetics Swine Test Station, Ida Grove, Iowa and the Farmland Agriservice Inc. Swine Test Station, Lisbon, Iowa. They also thank the United Duroc Swine Registry and the Hampshire Swine Registry. We thank D. S. Singer for providing the SLA class I probe used in these studies and M. Braet, R. Goodwin, W. R. Lee, I. Pour-El, and N. K. Schwartz for their excellent technical assistance. This work was supported in part by grants from the Iowa High Technology Council, the Iowa Biotechnology Council, and the USDA. Journal Paper No. J-12957 of the Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa. Projects Nos. 1901, 2594 and 2609.

References

Brown WM (1980) Polymorphism in mitochondrial DNA of humans as revealed by restriction endonuclease analysis. Proc Natl Acad Sci USA 77:3605–3609

Chardon P, Vaiman M, Renard C, Arnoux B (1978) Pig histocompatibility antigens and β_2-microglobulin. Transplantation 26:107–112

Chardon P, Vaiman M, Kirszenbaum M, Geffrotin C, Renard C, Cohen D (1985) Restriction fragment length polymorphism of the major histocompatibility complex of pig. Immunogenetics 21:161–171

Ehrlich R, Singer DS (1986) Organization and expression of class I MHC genes in the miniature swine. Fed Proc 4:846

Flanagan MP, Jung YC, Rothschild MF, Warner CM (1988) RFLP analysis of SLA class I genotypes in Duroc swine. Immunogenetics 27:465–469

Kaplan N (1983) Statistical analysis of restriction enzyme map data and nucleotide sequence data. In: Wier BS (ed) Statistical analysis of DNA sequence data. Marcel Dekker, New York, pp 75–106

Kirszenbaum M, Renard C, Geffrotin C, Chardon P, Vaiman M (1985) Evidence for mapping pig C4 gene(s) within the pig major histocompatibility complex (SLA). Anim Blood Groups Biochem Genet 16:65–68

Lie WR, Rothschild MF, Warner CM (1987) Mapping of C2, Bf and C4 genes to the swine major histocompatibility complex (SLA). J Immunol 139:3388–3395

Lunney JK, Sachs DH (1978) Transplantation in miniature swine. IV. Chemical characterization of MSLA and Ia-like antigens. J Immunol 120:607–612

Lunney JK, Pescovitz D, Sachs DH (1986) The swine major histocompatibility complex: Its structure and function. In: Tumbleson ME (ed) Swine in biomedial research vol. 3. Plenum Press, New York London, pp 1821–1836

Nei M (1975) Molecular population genetics and evolution. North-Holland, Amsterdam Oxford, pp 224–230

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York, pp 106–107

Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci USA 76:5269–5273

Nei M, Tajima F (1981) DNA polymorphism detectable by restriction endonucleases. Genetics 97:145–163

Nei M, Tajima F (1983) Maximum likelihood estimation of the number of nucleotide substitutions from restriction sites data. Genetics 105:207–217

Nei M, Tajima F, Gojobori T (1984) Classification and measurement of DNA polymorphism. In: Chakravarti A (ed) Human population genetics: The Pittsburgh Symposium. Van Nostrand Reinhold, New York pp 307–330

Satz ML, Wang L-C, Singer DS, Rudikoff S (1985) Structure and expression of two porcine genome clones encoding class I MHC antigens. J Immunol 135:2167–2175

Schreier PH, Bothwell ALM, Muller-Hill B, Baltimore D (1981) Multiple differences between the nucleic acid sequences of the $IgG2a^a$ and $IgG2a^b$ alleles of the mouse. Proc Natl Acad Sci USA 78:4495–4499

Singer DS, Camerini-Otero RD, Satz Ml, Osborne B, Sachs D, Rudikoff S (1982) Characterization of a porcine genomic clone encoding a major histocompatibility antigen: Expression in mouse L cells. Proc Natl Acad Sci USA 79:1403–1407

Singer DS, Lifshitz R, Ableson L, Nyrjesy P, Rudikoff S (1983) Specific association of repetitive DNA-sequences with major histocompatibility genes. Mol Cell Biol 3:903–913

Upholt WB (1977) Estimation of DNA sequence divergence from comparison of restriction endonuclease digests. Nucleic Acids Res 4:1257–1265

Vaiman M, Hauptman G, Mayer S (1978) Influence of the major histocompatibility complex in the pig (SLA) on serum haemolytic complement levels. J Immunogenet 5:59–65

Vaiman M, Chardon P, Renard Ch (1979) Genetic organization of the pig SLA complex. Studies of nine recombinant and biochemical and lysostrip analysis. Immunogenetics 9:353–361

Vaiman M, Renard Ch, Bourgeaux N (1988) The SLA complex, the major histocompatibility system in swine: its influence on physiological and pathological traits. In: Warner CM, Rothschild MF, Lamont SJ (eds) Proceedings of International Symposium on the Molecular Biology of the Major Histocompatibility Complex of Domestic Animal Species Iowa Sate University Press, Ames, Iowa, pp 23–38