Ann. Zool. Fennici 19: 203-207. 1982

Thin layer isoelectric focusing as a tool for higher category systematics

Håkan Tegelström, Per-Ivan Wyöni & Hans Ryttman

Tegelström, H., Wyöni, P-I. & Ryttman, H. 1982: Thin layer isoelectric focusing as a tool for higher category systematics. — Ann. Zool. Fennici 19: 203-207.

Taxonomic problems cover a wide range of genetic divergence. At the population level and up to closely related genera, standard gel electrophoresis has proved to be a valuable taxonomic tool. Higher category systematics, however, is often difficult to analyse using biochemical methods.

Thin layer isoelectric focusing of tissue proteins has been suggested as a taxonomic tool. We have analyzed 21 species from two orders of birds using heart and breast muscle proteins. Between 80 and 110 protein bands can be identified and the relative frequency of identical bands between species can be calculated.

The low degree of intraspecies variation in tissue proteins demonstrated by electrofocusing is an advantage for higher category classification and makes it possible to use only a few individuals. It has been shown that the relative frequency of identical protein bands between different taxonomical levels is in agreement with a classification based on anatomical and morphological characters.

Håkan Tegelström, Department of Genetics, University of Uppsala, Box 7003, S-750 07 Uppsala, Sweden.

1. Introduction

Since the distinctions between taxa are by definition inherited, it follows that systematics must be concerned with the nature and extent of genetic differences between groups. It is therefore not surprising that a considerable amount of work has been devoted to the use of biochemical methods in taxonomy.

Taxonomic problems cover a great range of genetic divergence. On the one hand we have the population level, with minor differences in genetic material, and on the other we have higher category levels (such as families and orders) in which the genetic complement is much more divergent. Unfortunately there is no biochemical method which can be used at all taxonomic levels. We have to choose the method according to each specific problem.

Starch gel electrophoresis was developed in the 1960s and revealed a surprising amount of intraspecies variation in enzymatic proteins. Allele frequencies can be estimated and used for calculations of genetic distance between taxa. This technique has been used extensively during the last decade for taxonomic purposes at the levels up to closely related genera. For higher category classification several biochemical methods have been applied, among which the most common are amino acid sequence determination, immunological calculation of genetic distance and also various chromosomal methods, including DNA-hybridization.

Taxonomic data from chemical analysis suggest classifications very similar to those based on anatomical and morphological data. In recent years confidence in the hypothesis of chemotaxonomic congruence has decreased as counter examples have appeared (Cronquist 1980, Harris & Bisby 1980, Joysey 1981). If such incongruence proves to be a frequent occurrence we shall no longer be able to rely on classifications based on a single method. Should morphological and anatomical approaches therefore be the methods of choice, or should we include chemosystematic methods?

All biochemical methods for higher category systematics present technical difficulties. Generally they are also time-consuming, which limits their taxonomic use to specialists. In this context we want to call attention to a biochemical method particularly suitable for higher category classification. It is based on many characters and can probably be used for an objective definition of different taxonomic levels. Thin layer isoelectric focusing is an easy and rapid technique, and only a few specimens of each species are needed for the construction of a phenogram.

The method has already occasionally been used taxonomically (Sarich 1977, Ferguson 1980), but

neither the power of the method has been analysed, nor its possible applications evaluated. Our aim in this article is to give a brief account of our, so far limited, experience of the method for higher category classification.

2. Methods

Isoelectric focusing (IEF) is a high resolution technique for the separation of proteins, based on differences in their isoelectric points. Its high resolution is due to a built-in concentration effect. This effect counteracts the diffusion that occurs in ordinary electrophoresis. Electrofocusing thus gives much better resolution between proteins with the same or similar mobility under standard electrophoresis. The method also allows detection of small amounts of protein.

2.1. Treatment of tissues

It is important that the tissues used for comparison are treated carefully, so as not to induce artificial differences. Our experience is that some tissues are more suitable for investigation than others. Heart muscle, for example, is less liable to methodological induction of protein variation than liver. It is preferable to keep the organisms alive as long as possible and to prepare the homogenates just before electrofocusing. The homogenates should be stored at —18°C, but not longer than three weeks because the proteins decompose.

The tissues are homogenized with distilled water (1:2) and centrifuged at high speed (~ 30,000 g). The gel used for separation has a pore size that does not permit the penetration of larger molecules and consequently the most efficient centrifugation technique available should be used to remove these particles and proteins. Small samples (like some plant leaves or small insects) can be centrifuged in a haematocrit centrifuge using disposable micropipettes down to 5 μl.

2.2. Isoelectric focusing (IEF)

IEF is performed in 5 % polyacrylamide gels (12 x 12 x 0.075 cm) with Pharmalyte™ (Pharmacia Fine Chemicals) as carrier ampholyte. Gels are prepared according to the instruction manual supplied with the ampholyte, except that Amberlite MB-1 is excluded and only half of the recommended quantity of freshly prepared ammonium persulphate plus 5μl TEMED are added per 15 ml gel solution. Pharmalytes with various pH intervals may be used, but it is best to start with pH 3-10 to find the interval giving the highest number of protein bands. We have usually worked with pH 3-10 and 4-6.5

Samples are applied to pieces of Whatman filter paper (5 × 8 mm) and placed on the gel about one cm from the cathode. About 20 samples can be focused on the gel size used. Gels with pH gradients of 3-10 are run at 10 W constant power for 2500 Vh, followed by 200 Vh at 14 W. The final voltage is 2500 V. Sample applicators are removed after 500 Vh. Gels with pH gradients of 4-6.5 are run at 10 W constant power for 4000 Vh, followed by 200 Vh at 15 W. The final voltage is 3000 V. Sample applicators are removed after 600 Vh.

After electrofocusing, general proteins are stained with Coomassie Brilliant Blue R 250 according to Vesterberg (1972). Gels are destained with a mixture of water, ethanol and glacial acetic acid (20:3:1) with repeated changes until background staining has disappeared, and then treated with destaining solution plus glycerol (1.5 %) for 30 minutes. The gels are dried at room temperature.

2.3. Evaluation

It is only possible to compare the protein bands of two samples placed next to each other on the gel. Therefore all possible combinations of species are run next to each other. It is preferable to use at least two individuals of each species, and to run two independent series with different individuals.

The IEF banding patterns of two samples are compared on a light board equipped with a magnifying lens. Two adjacent patterns (A and B) from the same or different species are compared and each band on the gel is classified by its presence in patterns A and B. The following three values are obtained: Bands present in A only (I), bands present in B only (II) and bands present in both A and B (III). The following ratios are thus possible: III/(I + III) and III/(II + III). The higher of the two ratios is used as the coefficient of similarity, called "IEF identity". This value is the relative frequency of identical bands in the two specimens compared.

All comparisons are made blind, i.e. the observer should have no knowledge of the taxonomic positions of specimens compared. A taxonomic classification based on isoelectric focusing is preferably begun by an examination of the degree of intraspecies variation. About 10 individuals should be used to establish the intraspecies identity. The comparison of IEF banding patterns requires some practice, especially when more distantly related species are involved. The coefficient of similarity may vary slightly when the same two samples on a gel are compared twice. We have used the criterion that two counts of the same comparison on a gel should not give ratios that differ by more than 0.05. We then use the mean value of the two counts. Two trained observers should count all comparisons independently.

3. Results and discussion

Species specific patterns have been found when species have been examined for a high number of proteins. This observation forms the basis for a phenetic classification using the "band counting method" of complex tissue protein patterns after high resolution separation.

One of the advantages in using electrofocusing is the high number of protein bands separated. In birds (Fig. 1) we have used heart muscle tissue, which gives 80-110 bands depending upon the separation characteristics chosen and the protein concentration in the samples. Using head and thorax from *Formica* ants we have identified 60-70 protein bands (Fig. 2). Thus a high number of usable characters can be rapidly identified in a comparison and more reliable estimates of protein similarities will be reached.

The intraspecies variation in investigated birds is low, and the IEF identity for intraspecies comparisons ranges between 0.88 and 1.00 with most values centering around 0.97. Thus only a few protein bands differ between individuals within a species. This low degree of difference between individuals from the same species has two important implications. Firstly it implies that it is not possible to use the method for population studies. Secondly the low degree of variation within a species is an advantage for the application of the method for higher category

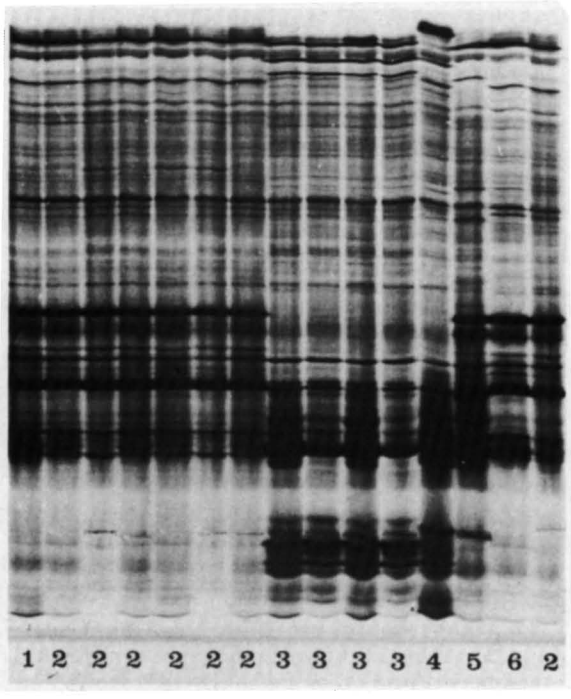

Fig. 1. General protein patterns of heart muscle extracts from several individuals in the order Charadriiformes. Separated by isoelectric focusing in a pH gradient of 3–10 (lowest pH at top of the gel). Samples: 1. *Larus argentatus*, 2. *Larus fuscus*, 3. *Uria aalge*, 4. *Plotus alle*, 5. *Rissa tridactyla*, 6. *Larus marinus*. Note the greater similarity in banding patterns between more closely related species.

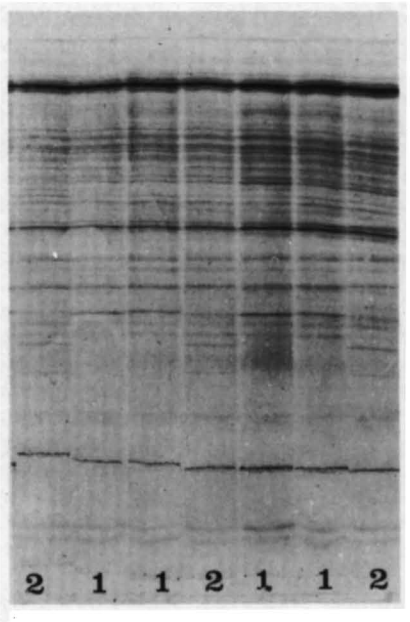

Fig. 2. General protein patterns of head and thorax extracts from different individuals in two species of *Formica* ants. Separated by isoelectric focusing in a pH gradient of 3–10 (lowest pH at top of the gel). Samples: 1. *Formica polyctena*, 2. *Formica rufa*. The two species show identical banding patterns.

classification. This makes it unnecessary to use more than 2 individuals when a comparison is made for higher category classification.

Comparisons between species at different taxonomic levels have been performed for 10 species, representing 3 families, in the order Falconiformes and 11 species, representing 3 families, in the order Charadriiformes. The IEF identity scale is shown in Fig. 3. The figure shows that when two species are compared, the traditional classification based on anatomical and morphological characters is reflected in the IEF identities. For example when two species from different families in the order Falconiformes are compared, the value representing this comparison will be lower than 0.64 but higher than 0.54.

The interpretation of the IEF scale presented in Fig. 3 is dependent upon the rate of change in the proteins, i.e. to what extent the proteins we are using in our comparisons behave as molecular clocks. Some proteins change more rapidly than others and what we detect in comparisons of closely related taxa is changes in these rapidly

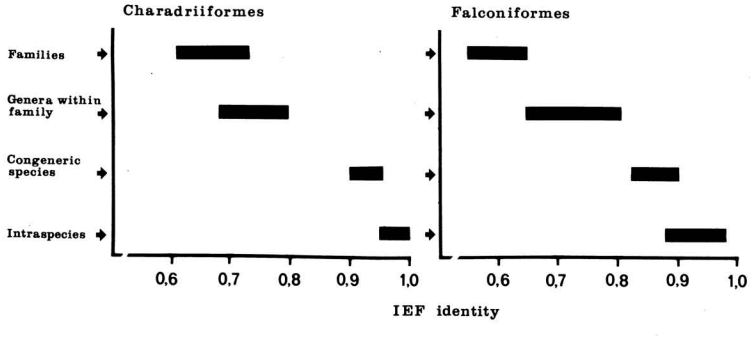

Fig. 3. Range of IEF identity values on different taxonomic levels in the orders Falconiformes and Charadriiformes. Values from comparisons of general protein patterns of heart muscle (Charadriiformes) and heart and breast muscle (Falconiformes) extracts. Separation in a pH gradient of 3–10.

evolving proteins. The more distantly related the organisms are, the more changes will be detected in conservative proteins. What we do not know is the proportion of rapidly and slowly evolving proteins and whether this proportion is different in different groups of organisms. If this proportion is different, the scales of identity between different taxa will be disproportionate. It has been suggested that birds might have a slow protein evolution (Prager et al. 1974) and that the degree of genetic variation (Selander 1976) and genetic distances (Avise et al. 1980) in birds are lower than in mammals. This suggests the need for caution in interpreting protein differences and, for example, transforming them into a time scale. It appears that some proteins evolve at relatively constant rates over long periods of time, whereas major changes in evolutionary rate take place in others. The average rates of evolution of many proteins taken over long periods of time may probably be used as an approximate "molecular clock".

The number of loci represented by the protein patterns after electrofocusing is a crucial point. Under certain conditions the number of bands equals the number of loci studied. Firstly, the resolving power of the separation technique must be high, to avoid the accumulation of several proteins at the same place in the gel. Although the resolving power of electrofocusing is high, several of the bands we use are most probably made up of more than one protein. Even if this is the case, when comparing different individuals from closely related taxonomic groups a change in net charge of a protein will be detected as a new band, despite the fact that the original band was made up from two or more different proteins.

Secondly, all proteins studied should be monomeric proteins giving only one band on the gel. This is not the actual situation, several of the proteins studied being composed of more than one protein chain.

Thirdly, all proteins studied should be monomorphic in the populations studied. This is not the case, but the degree of genetic variation for tissue proteins, and especially membrane bound proteins, is lower than that for soluble enzymes (Edwards & Hopkinson 1980 and Jones 1980).

Fourthly, there should be no post-translational modifications of the proteins giving more than one protein band for each protein. In this case we also include the modification of proteins that may take place when handling and preparing the samples.

The number of monomorphic loci and to some extent post-translational modifications can be estimated by the intraspecies variation, which covers all sources of variation, including that due to degradation of proteins and experimental treatment. In birds and *Formica* ants the intraspecies identity ranges between 0.88 and 1.00 with most values centering around 0.97. Using 80 to 90 protein bands after electrofocusing, the last two points mentioned affect around 2 to 3 bands. As a rough estimate we suggest that, when comparing closely related species, it is possible to detect differences in a number of loci roughly equal to the number of different bands.

We have studied two taxonomically problematic cases in birds: the osprey (*Pandion haliaetus*) and the kittywake (*Rissa tridactyla*). Of all genera in the order Falconiformes, *Pandion* has been the most troublesome. Differing views as to its taxonomic status have been presented, ranging from suborder to subfamily. It is sometimes placed in the Cathartidae family (Compton 1938), or as a subfamily in the family Accipitridae. A close relationship to Accipitridae has been suggested by electrophoretic similarities between the egg white proteins of this and the family Cathartidae (Sibley & Ahlquist 1972).

A comparison of general proteins from heart and breast muscle between the osprey and several members of the order Falconiformes shows that the osprey is neither closely related to the Accipitridae family nor to the Falconidae. We have not yet had the opportunity to test the relation between the osprey and members of the family Cathartidae.

The other taxonomic problem investigated is whether the kittywake should be classified as a member of the genus *Larus* or placed in a genus of its own. Sometimes the kittywake, which morphologically resembles the gulls, is placed in the genus *Larus* together with other gulls (see for example Gruson 1976). In comparison with three *Larus* species (*argentatus, fuscus* and *marinus*) and 4 other species in the Charadriiformes order, its IEF identity clearly indicates the placing of the kittywake in a genus other than *Larus* and consequently in a genus of its own — *Rissa.*

The adoption of protein-based classifications of higher categories will probably lead to problems in the ranking of taxa. For example, on protein criteria alone, the current scheme of placental mammal classification is greatly inflated relative to that of frogs (Wilson 1975, Wilson et al. 1977). However, as indicated by the IEF identity scale shown in Fig. 3, the major classification by protein data and traditional classification are in agreement. There is no *a priori* reason to adopt morphological features as being more valuable than those of chemosystematics, but we must be careful not to create with the use of biochemical data more problems than we solve.

References

Avise, J. C., Patton, J. C. & Aquadro, P. F. 1980: Evolutionary genetics of birds: comparative molecular evolution in New World warblers and rodents. — J. Hered. 71: 303-310.

Compton, L. V. 1938: The pterylosis of the Falconiformes with special attention to the taxonomic position of the Osprey. — Univ. Calif. Publ. Zool. 42: 173-212.

Cronquist, A. 1980: Chemistry in plant taxonomy: an assessment of where we stand. — In: Bisby, F. A., Vaughan, J. G. & Wright, C. A. (eds.), Chemosystematics: Principles and Practice. Academic Press, N.Y.: 1-27.

Edwards, E. & Hopkinson, A. 1980: Are abundant proteins less variable? — Nature 284: 511-512.

Ferguson, A. 1980: Biochemical Systematics and Evolution. — Blackie and Sons Ltd.

Gruson, E. S. 1976: Checklist of the birds of the World. — W. Collins Sons & Co Ltd.

Harris, J. A. & Bisby, F. A. 1980: Classification from chemical data. — In: Bisby, F. A., Vaughan, J. G. & Wright, C. A. (eds.), Chemosystematics: Principles and Practice. Academic Press, N.Y.: 305-327.

Jones, J. S. 1980: How much genetic variation? — Nature 288: 10-11.

Joysey, K. A. 1981: Molecular evolution and vertebrate phylogeny in perspective. — Symp. Zool. Soc. Lond. 46: 189-218.

Prager, E. M., Bruch, A. H., Nolan, R. A., Nakanishi, M. & Wilson, A. C. 1974: Slow evolution of transferrin and albumin in birds according to micro-complement fixation analysis. — J. Mol. Evol. 3: 243-262.

Sarich, V. M. 1977: Rates, sample sizes and the neutrality hypothesis for electrophoresis in evolutionary studies. — Nature 265: 24-28.

Selander, R. K. 1976: Genetic variation in natural populations. — In: F. J. Ayala (ed.), Molecular Evolution. Sinauer Ass.: 21-45.

Sibley, C. & Ahlqvist, J. E. 1972: Non-passerine egg-white proteins. — Bull. Peabody Mus. Nat. Hist. 39: 276 pp.

Vesterberg, O. 1972: Isoelectric focusing of proteins in polyacrylamide gels. — Biochim. Biophys. Acta 257: 11-19.

Wilson, A. C. 1975: Relative rates of evolution of organisms and genes. — Stadler. Genet. Symp. 7: 117-134.

Wilson, A. C., Carlson, S. S. & White, T. J. 1977: Biochemical evolution. — Ann. Rev. Biochem. 46: 573-639.

Received 3.III.1982
Printed 10.XII.1982