Sex Chromosome Loss Induced by the “Sex-Ratio” Trait in
Drosophila pseudoobscura Males
G. Cobbs

81

Canalization at Two Extra Scutellar Bristles: Artificial Selection,
Chromosome Analysis, and Effect of Temperature
R. Piñeiro

85

High-Resolution G-Banding of Sheep Chromosomes
H.M. Kaftanovskaya and O.L. Serov

92

Response to Selection for Increased Hybridization Between
Drosophila melanogaster Females and D. simulans Males
J.I. Izquierdo, M.C. Carracedo, R. Piñeiro, and P. Casares

100

Heritability of 90-day Body Weight in Domestic Rabbits from
Tropical Ghana, West Africa
S.D. Lukefahr, J.K.A. Atakora, and E.M. Opoku

105

Use of a Ds Chromosome-Breaking Element to Examine Maize Vp5
Expression
E.T. Wurtzel

109

Mutator Transposable Elements That Occur in Clusters in the Maize
Genome
S.C. Ingels, J.L. Bennetzen, S.H. Hulbert, M. Qin, and A.H. Ellingboe

114

Covariance Between Haploid-Species Hybrid and Tuberosum x
Haploid-Species Hybrid In 4x-2x Crosses of Solanum tuberosum L.
K.G. Haynes

119

Extensive Conservation of Linkage Relationships Between Pea and
Lentil Genetic Maps
N.F. Weeden, F.J. Muehlbauer, and G. Ladizinsky

123

The RSS System of Unidirectional Cross-Incompatibility in Maize: I.
Genetics
A. Rashid and P.A. Peterson

130

Hybrid Weakness in Wild Phaseolus vulgaris L.
E.M.K. Koinange and P. Gepts

135

Brief Communications

Inheritance of Isozyme Phenotypes of the Native Paulownia spp. in
Taiwan
R. Finkeldey

140
REAP: An Integrated Environment for the Manipulation and Phylogenetic Analysis of Restriction Data

D. McElroy, P. Moran, E. Bermingham, and I. Kornfield

Over the past decade, restriction enzyme analysis has become a mainstay for population and evolutionary biologists interested in the phylogenetic relationships of closely related taxa. Typically, one sets out to survey restriction variation by maximizing either (1) the number of individuals in the sample or (2) the number of restriction enzymes employed, which results in a binary matrix of considerable size. Several extremely powerful software packages, such as PAUP (Swofford 1985), PHYLIP (Felsenstein 1988), and NTSYS (Rohlf 1990), as well as any number of more specific programs, exist for the clustering of taxa via a diversity of algorithms; however, generating appropriate input files, each with program-specific format characteristics, is often a labor-intensive task.

The Restriction Enzyme Analysis Package (REAP) is designed to alleviate some of the difficulties inherent in restriction data manipulation, as well as to carry out some common phenetic analyses. The user creates a file of composite haplotypes and a corresponding file of restriction enzyme profiles; from this, REAP will generate a binary matrix, remove uninformative characters or Operational Taxonomic Units (OTUs), and compute estimates of evolutionary distance (\(d \pm SE\)) for site or fragment data. In addition, there are programs to estimate haplotype and nucleotide diversity within populations and nucleotide divergence among populations, to assess geographic heterogeneity in haplotype frequency distributions through Monte Carlo simulation, and to estimate genetic distance from DNA sequence data.

Each of the nine programs can run independently, as part of a batch process, or as a module in the integrated environment. All programs, with the exception of DA and MONTE (see below), can handle an unlimited number of OTUs and a maximum of 30,000 characters per OTU.

1. GENERATE produces a binary character state matrix from two user-created input files. A file of OTUs with their composite haplotypes is compared to a file containing the binary representation of restriction phenotypes produced by each enzyme. In the output, restriction phenotype designations are replaced with the appropriate binary code, thus generating a rectangular character state matrix. GENERATE can be used to create valid PAUP, PHYLIP, or D input files, supplying program-specific format characteristics for each.

2. REDUCE removes uninformative characters (character state 0 for all OTUs) from a binary matrix; optionally, monomorphic characters of state 1 and autapomorphic characters can also be eliminated. Because uninformative characters can be easily removed from the matrix, the user does not have to tailor the file of binary codes in order to eliminate OTUs during analysis. REDUCE is designed to operate on GENERATE-produced files, but it will work on any discrete character data matrix designed for PAUP, PHYLIP, or D.

3. GROUP identifies those OTUs in a binary matrix (or its composite haplotype precursor) that have identical restriction phenotypes and collapses them into a single OTU. This feature simplifies the data matrix and should increase processing speed during clustering, as nodes of zero branch length are eliminated. Any number of different groups of haplotypes can be created; members of each group (OTU names) are retained in comment lines. Again, this program will operate on any valid PAUP, PHYLIP, or D binary file.

4. D computes a nucleotide substitution matrix (\(d\) values) from site or fragment data and presents the results in a form suitable for phylogenetic clustering using NTSYS. For a given pair of OTUs, a set of \(d\) values is estimated from the proportion of shared characters in each class of restriction enzyme (\(r\) values = 4, 14/3, 5, 16/3, or 6); an overall weighted estimate of \(d\) is then generated. For site data, \(d\) is computed as per Nei and Tajima (1981) and Nei and Miller (1990, eq. 4), which is suitable for \(d < 0.25\) and agrees well with estimates obtained via the maximum likelihood estimation (Nei and Miller 1990). In the case of fragment data, \(d\) is estimated according to Nei and Li (1979) and Nei (1987, eq. 5.55). Weighting of individual \(d\) follows Nei and Tajima (1983) and is based on the proportion of characters generated by each class of restriction enzyme.

5. DSE is designed to give a more descriptive output for estimates of \(d\) from either site or fragment data and provide standard errors around \(d\) for each pairwise comparison. In addition, the average number of characters generated and nucleotide positions surveyed per OTU by each class of restriction enzyme are reported. Standard errors for site data are computed according to Nei and Tajima (1983) and Nei (1987, eqs. 5.41, 5.44, and 5.51). Standard errors for fragment data are difficult to determine analytically, as the distribution of \(d\) is highly skewed (Nei M, personal communication), although reliable estimates can be achieved numerically via the jacknife (Nei and Miller 1990). Nei and Li (1979) provide no analytic solution for determining standard errors around \(d\) for fragment data; as such, the equation of
Upholt (1977, eq. 6b) is used. Individual SE estimates for both site and fragment data are weighted equivalently to d. The output from DSE is strictly informational and cannot be used as input for clustering.

6. DSIZE is a modification of D that reports d, for a specified class of restriction enzyme. We have noted that estimates of d, from different enzyme classes may vary significantly, so this program is provided as a means of examining the data more closely. Like D, the output file is suitable without modification for NTSYS.

7. DA estimates haplotype and nucleotide diversity within populations and computes the nucleotide divergence among all pairs of populations. The number of populations is limited to 50; a maximum of 100 haplotypes per population is allowed. Haplotype diversity is estimated according to Nei (1987, eqs. 8.4, 8.5, and 8.12). Nucleotide diversity and nucleotide divergence are estimated according to Nei and Tajima (1981) and Nei (1987, eqs. 10.7, 10.19, 10.20, and 10.21). Output consists of (1) a descriptive file of haplotype and nucleotide diversity (±SEs) for all populations and (2) a second file of nucleotide divergence among all pairs of populations. The latter results are presented as a symmetric matrix suitable without modification for NTSYS.

8. MONTE analyzes the extent of geographic heterogeneity among haplotype frequency distributions through a Monte Carlo simulation (Roff and Bentzen 1989). This procedure is designed to minimize the effect of large numbers of empty cells on the validity of chi-square analysis. Up to 50 populations and 200 classes of individuals per population are allowed; the total number of individuals in the sample is limited to 5,000. The extent of heterogeneity in the original matrix is compared to that estimated from repeated randomizations of the data. The output file reports (1) the probability (±SE) of generating by chance alone a χ² value which exceeds that calculated from the original matrix, and (2) average, minimum, and maximum χ² values from the simulation, as well as a cumulative frequency distribution of χ² values.

9. K computes estimates of evolutionary distance from nucleotide sequence data using the two-parameter model of Kimura (1980). For a given pair of OTUs, k (which is equivalent to d) is estimated based on the probabilities of transitions and transversions at each homologous site. Missing bases can be accounted for and are not included in the calculations. Like D, the output is suitable without modification for NTSYS.

REAP was written in Turbo Pascal, Version 5.0 (Borland International 1988) for IBM compatibles. The programs should run without modification on any DOS-based machine. A numeric coprocessor is not required and will be emulated if it is not present. MS-DOS Version 3.0 or greater allows the most flexible use of the REAP Integrated Environment Module. To obtain a copy of this package, along with full documentation and example files, send a formatted diskette (5.25 or 3.5-inch) to D. McElroy.

From the Department of Zoology, the Migratory Fish Research Institute, and the Center for Marine Studies, University of Maine (McElroy, Moran, and Kornfield), and the Smithsonian Tropical Research Institute, Balboa, Panama (Birmingham). Development of this package was supported in part by grants from the National Science Foundation (BSR 85 17040) and NOAA Sea Grant (NA-85 AA-D-SG-020) to L.K. Address reprint requests to D. McElroy, Department of Zoology, University of Maine, 217 Murray Hall, Orono, ME 04469.

The Journal of Heredity 1992;83(2)

References


