# ALLOZYME ELECTROPHORESIS 

# A Handbook for Animal Systematics and Population Studies 

B.J. RICHARDSON<br>Bureau of Flora and Fauna<br>Canberra<br>P.R. BAVERSTOCK and M. ADAMS<br>South Australian Museum<br>Adelaide



## ACADEMIC PRESS

(Harcourt Brace Jovanovich, Publishers)
Sydney Orlando San Diego New York Austin
London Montreal Tokyo Toronto

# ACADEMIC PRESS AUSTRALIA 

Centrecourt, 25-27 Paul Street North
North Ryde, N.S.W. 2113
United States Edition published by
ACADEMIC PRESS INC.
Orlando, Florida 32887

## United Kingdom Edition published by

ACADEMIC PRESS, INC. (LONDON) LTD.
24/28 Oval Road, London NW1 7DX
Copyright © 1986 by
ACADEMIC PRESS AUSTRALIA
All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the publisher.

Printed in Australia
National Library of Australia Cataloguing-in-Publication Data

Richardson, B. J. (Barry John), date- .
Allozyme electrophoresis.
Bibliography.
Includes index.
ISBN 0125878400.

1. Enzymes - Analysis - Technique. 2. Electrophoresis - Technique. I. Baverstock, P. R. (Peter Raymond), date- . II. Adams, M. (Mark), date- . III.
Title.
591.19'25

## CONTENTS

General Introduction ..... xi
Part I. Project Planning ..... 1
Introduction to Part I ..... 1
Chapter 1. Genetics ..... 3
1.1 Elementary Genetics ..... 3
1.2 Population Genetics ..... 8
1.2.1 Non-random mating ..... 10
1.2.2 Small population size ..... 11
1.2.3 Differential selection ..... 11
1.3 Clines ..... 12
Chapter 2. Electrophoresis ..... 15
2.1 Protein Structure and Function ..... 15
2.2 Electrophoresis ..... 16
2.3 Staining Systems ..... 20
2.4 Multiple Forms of Enzymes ..... 22
2.4.1 The presence of more than one locus ..... 22
2.4.2 The presence of more than one allele ..... 22
2.4.3 Post-translational modifications ..... 24
2.5 Allelic Variation ..... 26
2.6 Applications of Allozyme Electrophoresis ..... 28
Chapter 3. Population Structures ..... 31
3.1 Alternative Population Models ..... 32
3.2 Differentiating Between Population Models in Simple Situations ..... 35
3.3 More Complex Field Situations ..... 36
3.3.1 Overlapping subpopulations ..... 36
3.3.2 Very small subpopulations ..... 37
3.3.3 Failure to meet isolation-by-distance model assumptions ..... 38
3.4 Sampling Strategies ..... 39
Chapter 4. Systematics ..... 40
4.1 Species ..... 40
4.2 Supraspecific Categories ..... 41
4.3 Electrophoresis in Systematics ..... 46
4.3.1 Delineation of biological species ..... 46
4.3.2 Phylogenetic reconstruction ..... 53
Chapter 5. Statistical Methods ..... 57
5.1 Binomially Distributed Data ..... 57
5.2 Hypothesis Testing ..... 58
5.3 Setting Sample Sizes for Population Studies ..... 60
5.4 Testing for Differences in Allele Frequency ..... 62
5.5 Testing the Assumption that a Single Panmictic Population is Sampled ..... 65
5.6 Testing the Assumption that Alleles at a Single Locus are being Studied ..... 69
5.7 Tests for the Independence of Loci ..... 69
5.8 Measures of Genetic Relatedness for use in Population Studies ..... 70
5.9 Sample Sizes in Taxonomic Studies ..... 74
Chapter 6. Project Planning and Sampling Strategies ..... 76
6.1 Population Studies ..... 76
6.2 Systematic Studies ..... 79
6.2.1 Delineating species boundaries ..... 80
6.2.2 Phylogenetic studies ..... 82
Part II. Collecting the Data ..... 83
Introduction to Part II ..... 83
Chapter 7. Sample Collection, Handling and Preparation ..... 85
7.1 Background ..... 85
7.2 General Principles of Sampling ..... 87
7.3 Tissues ..... 89
7.4 Taking and Processing ..... 91
7.5 Preserving ..... 93
7.6 Transport ..... 94
7.7 Long-term Storage ..... 94
7.8 Preparation of Samples for Electrophoresis ..... 95
7.9 Compromises in Sampling Technique ..... 98
Chapter 8. Electrophoresis Methods ..... 99
8.1 Choice of Electrophoretic Support Medium ..... 99
8.2 General Laboratory Requirements ..... 102
8.2.1 Equipment ..... 102
8.2.2 Consumables ..... 103
8.3 Standard Electrophoretic Procedures ..... 104
8.3.1 Pre-loading ..... 113
8.3.2 Loading ..... 115
8.3.3 The electrophoretic run ..... 118
8.3.4 Staining ..... 119
8.3.5 Scoring ..... 122
8.4 Genetic Interpretation of Zymograms ..... 127
8.4.1 Factors responsible for complex electrophoretic patterns ..... 128
8.4.2 Practical interpretation procedures ..... 136
8.4.3 General principles of interpretation ..... 140
8.4.4 Line-up gels ..... 142
8.4.5 Integration of results across gels ..... 143
Chapter 9. Specific Enzyme Methods for Cellulose Acetate Electrophoresis ..... 145
9.1 The Structure of the Enzyme Sheets ..... 145
9.1.1 General background ..... 145
9.1.2 Standardized nomenclature ..... 146
9.1.3 Consumables ..... 148
9.1.4 The format of the enzyme sheets ..... 148
9.2 The Sheets ..... 161
9.3 Explanatory Notes ..... 219
9.3.1 Note 1: Buffers ..... 219
9.3.2 Note 2: Stabilizers ..... 219
9.3.3 Note 3: Common detection mechanisms ..... 220
9.3.4 Note 4: Isozyme mobility ..... 221
9.3.5 Note 5: Formazan precipitation ..... 221
9.3.6 Note 6: Inhibition of phantom dehydrogerases ..... 221
9.3.7 Note 7: Arsenate ..... 222
9.3.8 Note 8: Post-coupling ..... 222
9.3.9 Note 9: NAD-dependent G6PD ..... 222
9.3.10 Note 10: Fluorescent detection systems ..... 222
9.3.11 Note 11: Cytoplasmic and mitochondrial enzymes ..... 224
9.3.12 Note 12: Multiple locus systems ..... 225
9.3.13 Note 13: Hybridization between the isozyme products of separate loci ..... 225
9.3.14 Note 14: Treatment-related modification ..... 226
9.3.15 Note 15: Complex electrophoretic patterns ..... 226
9.3.16 Note 16: Warping ..... 227
9.3.17 Note 17: Uninterpretable variation ..... 227
9.3.18 Note 18: Atypical heterozygotes ..... 227
9.3.19 Note 19: Stain artifacts ..... 228
Chapter 10. Strategies ..... 229
10.1 Setting Up a New Electrophoresis Unit ..... 229
10.1.1 Background concepts ..... 229
10.1.2 Learning strategies for cellulose acetate ..... 231
10.2 Enzyme Suitability ..... 235
10.2.1 Presence ..... 235
10.2.2 Staining intensity ..... 236
10.2.3 Electrophoretic resolution ..... 236
10.2.4 Reliability ..... 237
10.2.5 Cost of stain ..... 237
10.2.6 Tissue distribution ..... 237
10.2.7 Number of loci ..... 237
10.2.8 Tissue-specific loci ..... 237
10.2.9 Electrophoretic conservativeness ..... 238
10.2.10 Summary ..... 238
10.3 Laboratory Strategies ..... 238
10.3.1 Routine screen ..... 239
10.3.2 Method modification ..... 239
10.3.3 Further investigation ..... 242
10.3.4 Line-ups ..... 243
10.3.5 Random retyping ..... 243
10.3.6 Reinterpretation ..... 243
10.3.7 General laboratory protocol ..... 243
10.3.8 Investigative work on multiple-locus systems ..... 244
10.3.9 Electromorph splitting ..... 249
10.4 Specific Resource-Saving Strategies ..... 251
10.4.1 Tissue handling ..... 251
10.4.2 Test strips ..... 251
10.4.3 Choice of running buffers ..... 252
10.4.4 Choice of tissues for screening ..... 253
10.4.5 Pre-loading strategies ..... 254
10.4.6 Selection of loading order ..... 254
10.4.7 Mobility controls ..... 256
10.4.8 Origin placements for cellulose acetate electrophoresis ..... 256
10.4.9 Multiple origins ..... 257
10.4.10 Multiple loading ..... 258
10.4.11 Run strategies ..... 258
10.4.12 Staining strategies ..... 260
10.4.13 Staining weak activity enzymes ..... 261
10.4.14 Double stains ..... 261
10.4.15 Freezing gels ..... 264
10.4.16 Trouble-shooting ..... 264
10.4.17 Stain trouble-shooting ..... 268
10.4.18 Large-scale genetic typing exercises ..... 269
10.4.19 Inter-laboratory consultation ..... 270
Part III. Analysing the Data ..... 271
Introduction to Part III ..... 271
Chapter 11. Population Analysis using Electrophoretic Data ..... 273
11.1 Analysis of the Sample Set Data ..... 274
11.1.1 Locus-specific phenomena ..... 277
11.1.2 Population-specific phenomena ..... 279
11.2 Within-Stock and Between-Stock Analysis ..... 280
11.2.1 Comparisons between replicate sample sets ..... 280
11.2.2 Comparisons within stocks ..... 280
11.2.3 Comparisons of different stocks ..... 282
11.2.4 Geographical analysis of allele frequencies ..... 282
11.3 Further Analysis of Geographical Structuring ..... 283
11.3.1 Theoretical relationships between genetic and geographical distance ..... 287
11.3.2 The effects of other factors ..... 288
11.3.3 Mapping the distributions of subpopulations ..... 288
11.3.4 Estimating neighbourhood size ..... 289
11.4 Examples of Analysis ..... 290
11.4.1 The jackass morwong (Cheilodactylus macropterus) ..... 290
11.4.2 The Australian snapper (Chrysophrys auratus) ..... 293
11.4.3 The European rabbit in Australia ..... 295
11.4.4 Brown Trout (Salmo trutta) in Scandinavia ..... 298
11.4.5 Skipjack tuna (Katsuwonus pelamis) in the Bismark Sea ..... 299
Chapter 12. Analysis of Electrophoretic Data in Systematics ..... 303
12.1 Dissection of Species Complexes ..... 303
12.1.1 Sympatric sibling species ..... 303
12.1.2 Allopatric populations ..... 304
12.2 Phylogenetic Reconstruction ..... 309
12.2.1 Phenetic analysis ..... 309
12.2.2 Numerical methods based on distance data ..... 312
12.2.3 The Hennigian approach ..... 316
12.2.4 Reticulate evolution ..... 325
12.2.5 The molecular clock ..... 326
12.3 Integration With Other Data ..... 327
12.3.1 Phylogenetic analysis ..... 327
12.3.2 Species complexes ..... 328
12.4 A Worked Example of Phylogenetic Analysis ..... 330
12.4.1 Distance analysis ..... 331
12.4.2 Hennigian analysis ..... 337
12.4.3 Synthesis ..... 345
12.4.4 Details of computer programs ..... 346
Part IV. Other Types of Project ..... 347
Chapter 13. Other Biological Uses for Allozyme Electrophoresis ..... 349
13.1 Genetic Identification ..... 349
13.1.1 Genetic identification of an individual ..... 349
13.1.2 Species identification ..... 349
13.2 Genetic Quality Control of Laboratory Animal Stocks ..... 351
13.2.1 Background ..... 351
13.2.2 Project planning ..... 351
13.2.3 Interpretation ..... 352
13.3 Cell Culture Contamination ..... 353
13.4 Parental Testing ..... 353
13.5 Relating a Tissue to an Individual ..... 354
13.6 Identifying Cryptic Species ..... 354
13.7 Identifying Non-adult Forms ..... 355
13.8 Identifying Meat Substitutions ..... 355
13.9 Gene Mapping ..... 356
13.10 Origins and Relationships of Domestic Forms ..... 357
13.11 Assessing the Loss of Variability Due to Inbreeding ..... 357
13.12 Applications in Protozoology and Bacteriology ..... 358
13.13 Applications of Allozyme Electrophoresis to Plants ..... 360
13.13.1 Ploidy level ..... 360
13.13.2 Breeding systems ..... 361
13.13.3 Chemical defence systems ..... 361
Appendix I ..... 363
Appendix II ..... 367
Bibliography ..... 389
Index ..... 401

## GENERAL INTRODUCTION

Since the advent of starch gel electrophoresis in 1959, the technique of electrophoresis has been increasingly used to provide useful information in a wide range of biological and medical situations. One major area has been its use as a tool for genetic analysis, using enzymatic proteins as markers of variation in the underlying genes (i.e. allozyme electrophoresis). Such genetic markers are useful for the identification of individuals, population structure analysis, the delineation of species boundaries and phylogenetic reconstruction. The purpose of this book is to address these aspects in a single, convenient handbook for all those intending to use electrophoresis as a tool to answer practical questions in population structure analysis, systematics or specimen identification.

Ferguson (1980) provides useful background information in these areas, but does not give the necessary practical details. The book by Harris and Hopkinson (1976), on the other hand, gives a detailed account of the practical procedures for conducting allozyme electrophoresis, but is directed towards work on humans. Both these books make excellent companion references for this handbook.

Whilst much of this book refers to zoological applications, the principles behind the technique and most of the methodology are applicable to virtually all life forms.

In times of budgeting restraint, considerations of cost in terms of money and, more especially, of time, are of paramount importance. A continuing theme throughout the book is that of maximum return for minimum cost. In electrophoretic work, collecting the samples in the field may be much more expensive than the electrophoresis itself. Therefore particular emphasis is placed upon project planning from the economic point of view without compromising the aims of the project.

Any electrophoretic analysis of population structure, species complexes, or phylogeny goes through three phases; project planning, electrophoresis of the samples, and data analysis. The book includes separate parts dealing with each of these three phases.

Part I considers background concepts. The book assumes no specialized knowledge of either genetics or biochemistry. However, some background knowledge is essential to economical and sensible sampling strategies, to
the process of electrophoresis and to the biological interpretation of the data obtained.

Part II covers the practical aspects of data collection, from the viewpoints of both taking and handling the samples for electrophoresis, and of running, staining and interpreting gels.

Part III discusses the analysis of data with methods and examples being given for the treatment of both population data and systematic data.

Part IV consists of a single chapter which briefly reviews other practical applications of allozyme electrophoresis, such as the identity of cell cultures, checking the 'purity' of inbred strains of laboratory animals, the identification of larval forms, and species identification in mixed fish catches.

A number of support media are suitable for allozyme electrophoresis. We have found that cellulose acetate is a very versatile medium and the easiest to use. Therefore the methods section of this book refers mainly to that medium (and in particular to the 'Cellogel' brand). However, we stress that all of Parts I, III and IV and much of Part II are independent of the support medium used. Therefore the book should prove useful to anyone using allozyme electrophoresis, whatever the support medium.

The authors' task of preparing this work was made easier by the support of many people. We would especially like to thank R. Andrews and S. Donnellan for their sustained effort in forcing us to clarify and clearly describe our ideas. We thank R. Hnatiuk, R. Longmore, P. O'Donahue, J. Orbach, T. Reardon, T. Schwaner and D. Walton, for their comments and suggestions on the manuscript, C. Curtis and A. Gunjko for computing assistance, J. Riede and J. Thurmer for preparing the figures and P. Kidd, G. Downing, V. Ochiltree, W. Riley and E. Tieni for typing assistance. We would also like to thank Grant Walker of Academic Press Australia for his patience and support during the long gestation of this book.


