

**Institute of Mathematical Statistics**  
**LECTURE NOTES—MONOGRAPH SERIES**

# **Statistics in Molecular Biology and Genetics**

Selected Proceedings of a 1997 Joint AMS-IMS-SIAM Summer Conference on Statistics  
in Molecular Biology

**Francoise Seillier-Moiseiwitsch, Editor**

**Volume 33**

**Published by the Institute of Mathematical Statistics  
and the American Mathematical Society**



**Institute of Mathematical Statistics**  
**LECTURE NOTES–MONOGRAPH SERIES**  
**Volume 33**

# **Statistics in Molecular Biology and Genetics**

**Selected Proceedings of a 1997 Joint AMS-IMS-SIAM  
Summer Conference on Statistics in Molecular Biology**

Francoise Seillier-Moiseiwitsch, Editor

**American Mathematical Society**  
Providence, Rhode Island

**Institute of Mathematical Statistics**  
Hayward, California

Institute of Mathematical Statistics

*Lecture Notes-Monograph Series*

Editorial Board

Andrew A. Barbour, Joseph Newton, and David Ruppert (Editor)

The production of the *IMS Lecture Notes-Monograph Series* is managed by the IMS Business Office: Julia A. Norton, IMS Treasurer, and Elyse Gustafson, IMS Business Manager.

This volume was co-published with the American Mathematical Society

Library of Congress Catalog Card Number: 99-076060

International Standard Book Number 0-940600-47-1

Copyright © 1999 Institute of Mathematical Statistics

All rights reserved

Printed in the United States of America

# TABLE OF CONTENTS

Preface .....	v
<i>F. Seillier-Moiseiwitsch</i>	
<b>Genetic Mechanisms</b>	
On a Markov Model for Chromatid Interference .....	1
<i>H. Zhao and T. Speed</i>	
<b>Population Genetics</b>	
Some Statistical Aspects of Cytonuclear Disequilibria .....	21
<i>S. Datta</i>	
Diffusion Process Calculations for Mutant Genes in Nonstationary Populations .....	38
<i>R. Fan and K. Lange</i>	
The Coalescent with Partial Selfing and Balancing Selection: An Application of Structured Coalescent Processes .....	56
<i>M. Nordborg</i>	
<b>Human Genetics</b>	
Statistical Aspects of the Transmission/Disequilibrium Test (TDT) .....	77
<i>W. Ewens</i>	
Estimation of Conditional Multilocus Gene Identity among Relatives .....	95
<i>E. Thompson and S. Heath</i>	
<b>Quantitative Genetics</b>	
A Review of Methods for Identifying QTL's in Experimental Crosses .....	114
<i>K. Broman and T. Speed</i>	
<b>Evolutionary Genetics</b>	
Markov Chain Monte Carlo for the Bayesian Analysis of Evolutionary Trees from Aligned Molecular Sequences .....	143
<i>M. Newton, B. Mau and B. Larget</i>	
Likelihoods on Coalescents: A Monte Carlo Sampling Approach to Inferring Parameters from Population Samples of Molecular Data .....	163
<i>J. Felsenstein, M. Kuhner, J. Yamato and P. Beerli</i>	

Uses of Statistical Parsimony in HIV Analyses .....	186
<i>K. Crandall</i>	
Linear Estimators for the Evolution of Transposable Elements .....	207
<i>P. Joyce, L. Fox, N. Casavant and H. Wichman</i>	
A Conditional Approach to the Detection of Correlated Mutations .....	221
<i>M. Karnoub, F. Seillier-Moiseiwitsch and P.K. Sen</i>	
Correlated Mutations in Protein Sequences: Phylogenetic and Structural Effects .....	236
<i>A. Lapedes, B. Giraud, L. Liu and G. Stormo</i>	

## **Sequence Motifs**

Compound Poisson Approximations for Occurences of Multiple Words ...	257
<i>G. Reinert and S. Schbath</i>	

## **Protein Structure**

Deriving Interatomic Distance Bounds from Chemical Structure .....	276
<i>M. Trosset and G. Phillips</i>	
Protein Fold Class Prediction is a New Field for Statistical Classification and Regression .....	288
<i>L. Edler and J. Grassmann</i>	