

## Book Reviews

---

*Am. J. Hum. Genet.* 76:190, 2005

*Statistical Methods in Genetic Epidemiology.* By Duncan C. Thomas. New York: Oxford University Press, 2004. \$65.

The Argentine writer Jorge Luis Borges (1899–1986) once wrote, “Nothing is built on stone; all is built on sand, but we must build as if the sand were stone.” The field of genetic epidemiology is no exception. In this constantly mutating field, what was cutting-edge statistical analysis in the 1980s was quickly subverted by the tremendous theoretical and technological advances of the 1990s. However, Thomas’ *Statistical Methods in Genetic Epidemiology* reference and textbook expertly extracts what is perennial from what is temporal and describes it in great detail for the current and future generations of individuals interested in genetic epidemiology.

The book is divided into 12 chapters—all well illustrated and with useful examples from the literature—and a glossary. The first four chapters are good reviews of genetic epidemiology, the basics of human genetics, the principles of Mendelian inheritance, and basic epidemiologic and statistical principles. In chapters 5 through 12, Thomas masterfully describes familial aggregation, segregation analysis, linkage analysis, the principles of population genetics, tests for candidate-gene associations, linkage-disequilibrium mapping, and gene characterization and includes a summary “tying it all together: the genetic epidemiology of colorectal cancer” (p. 339).

Since, in the book’s preface, Thomas states that the text is “a broad overview written at a level that should be accessible to graduate students in epidemiology, biostatistics, and human genetics” (p. vii), I asked a doctoral student in epidemiology, a doctoral student in bioinformatics, and a biostatistics faculty member for their opinions about the book. These three colleagues perform the statistical analysis of various genetic epidemiological studies for the Genetics Program at Boston University School of Medicine. The doctoral student in epidemiology stated that a graduate student in a discipline with limited use of statistical theory may find the derivations and the theory behind some of the methods difficult to follow. However, she pointed out that a major strength of the book is that the explanations are such that the reader could skip the more statistical explanations and still have at least a basic understanding of the concepts being presented. An additional strength of *Statistical Methods in Genetic Epidemiology* mentioned by this doctoral student is that the text presents many of the modern concepts of epidemiology—for instance, when Thomas discusses control selection for association studies. The graduate student in bioinformatics stated that Thomas’ *Statistical Methods* is an insightful and comprehensive reference

book. And then he noted, “If I had read this excellent book when I began my studies in bioinformatics three years ago, I would have a much stronger basis in genetic analysis now.” Finally, the biostatistics faculty member was impressed by the clarity of the topics covered throughout the book: “The reader is able to choose his/her own difficulty level, from the basics to the advanced statistics; there is much for everybody.” She is eager to start using this book in the courses she teaches.

The only weakness that I found in Thomas’ *Statistical Methods* is the absence of exercises at the end of the chapters. Their inclusion would have helped readers solidify the concepts and would have ensured that the statistical methods were well understood. This minor criticism aside, *Statistical Methods in Genetic Epidemiology* is a much welcome volume that all readers of *The American Journal of Human Genetics* should consider having on their bookshelves. It seems likely that this will serve as a classic text for many years to come.

I thank Carolien Panhuysen, Vikki Nolan, and Qianli Ma, for their opinions about the book.

DIEGO F. WYSZYNSKI

*Boston University School of Medicine*  
*Boston, MA*

© 2004 by The American Society of Human Genetics. All rights reserved.  
0002-9297/2005/7601-0022\$15.00

---

*Am. J. Hum. Genet.* 76:190–192, 2005

*Examining the Farming/Language Dispersal Hypothesis.*

Edited by Peter Bellwood and Colin Renfrew. Cambridge: McDonald Institute for Archaeological Research, 2002. \$85.

This book emerged from a conference of scholars from three different disciplines (archaeology, genetics, and comparative linguistics) that was convened by the book’s coeditors. A total of 36 papers written by 43 delegates were circulated before final drafts were prepared, and all the papers focused on the farming/language dispersal hypothesis (FLDH)—namely, that the distributions of some language families resulted from expansions of farming practices from their points of origin. Bellwood introduced the term “triangulation” to refer to the simultaneous focus of evidence from all three disciplines on the hypothesis. The book’s 36 chapters are organized into three

parts, the first of which consists of only two chapters—each written by one of the book’s coeditors—that introduce the rationale and assumptions underlying the FLDH.

Renfrew’s chapter explains the concepts of “spread zones”—distributions of closely related languages that are created by language/farming dispersals—and “mosaic zones” (also known as “residual” or “friction” zones)—distributions of many unrelated languages that are associated with ancient colonization. Demographic arguments underlying the “wave of advance” (or demic diffusion) model from which the genetic-research contribution to the FLDH (dubbed “archaeogenetics”) evolved are also presented. The validity and usefulness of these concepts are crucial to the FLDH and become contentious issues in many of the book’s later chapters (for example, Campbell’s chapter presents evidence that challenges the utility of the concept of spread zones). Bellwood’s chapter describes the process through which farming societies are thought to have evolved in situ from Mesolithic foragers in very few locations in the world and how, according to the FLDH, the expansion of farming practices replaced or assimilated those of adjacent hunter-gatherer populations, leading to language spreads.

Part II of the book comprises 7 chapters that focus on the relevant principles and theory of each of the three disciplines that are required for triangulation, and part III includes 27 chapters that provide evidence for or against arguments made in part II of the book, on the basis of studies of four geographic regions: Western Asia/North Africa (6 chapters), Oceania (9 chapters), Mesoamerica/the U.S. Southwest (4 chapters), and Europe (8 chapters). The chapters on comparative studies of languages, archaeology, and genetics are intermingled, but the three disciplines never very effectively merge in any single chapter.

Cavalli-Sforza’s chapter reviews the concept of demic expansion and reminds readers that the resulting gene-frequency gradients (i.e., clines) are caused by varying ratios of both migration and admixture. The studies by Barbujani and Dupanloup and by Chikhi, which simulate geographic gene-frequency clines on the basis of mtDNA and/or Y-chromosome haplotypes, generate the predictable clines in Eurasia, in agreement with Cavalli-Sforza’s earlier studies using conventional markers. Underhill hypothesizes that specific Y-chromosome haplotypes are genetic signatures of the Indo-European, Bantu, middle Yangtze, Austronesian, and Uto-Aztecan expansions, on the basis of samples from his ethnic database. The chapter by Forster and Renfrew reviews the use of haplotype networks to estimate the parameter  $\rho$ . In this chapter, “ $\rho$  dates” are provided for 20 of the world’s lineage clusters, by use of a mutation rate popularized by the senior author (Forster), and most dates are found to converge on several limited time intervals that are argued to reflect major pre-Neolithic demographic events, such as “Out of Africa” and the last glacial maximum. This adds little support for the FLDH in its unmodified form. One should view the inferences drawn from the clustering of  $\rho$  dates about some known prehistoric event with caution and should worry about the unattended potential influences of opportunistic sampling, inaccurate calibration of  $\rho$ , population substructure, and lack of fit between the ages of genes and populations. The chapter by Bandelt et al. takes a “tongue-in-cheek” look at “how data were turned into tales” (p. 100) and provides sage advice to archaeologists and linguists: “Don’t believe everything you read in the papers” (p. 101).

Some of the nongenetics chapters strongly support the FLDH. Jones’s chapter reviews the origins of domesticates that constitute the agricultural/pastoral packages associated with farming/language dispersals. Cattle travel with wheat and barley, and that happy union strengthens the dispersal potential of those grains; pigs serve that function for rice, though less effectively, since pigs cannot pull a plow. Harris’s chapter in part II compares the expansion capacities of early agriculture in five areas of the world, concluding that the evidence supports the FLDH for an Indo-European expansion better than for expansion of any of the other four centers of farming. This is supported by the chapter by Bar-Yosef, which outlines the transformation of the Natufian society in the northern Levant into Neolithic farmers, the chapter by Hassan, which identifies cognate words for food production and cattle keeping in African languages, and the chapter by Militarev, which presents an extensive list of etymologies consistent with the hypothesis that the Natufians spoke proto-Afrasian.

Cohen’s chapter explores the economic and demographic constraints on pre-farming communities that militated against the adoption of farming by hunter/gatherers, and Cohen concludes that some proponents of the FLDH have underestimated these constraints and overestimated the economic advantages provided by farming. His viewpoint draws support from five chapters in part III. Chapters by Hudson and by Barker argue that the Sahara was rapidly colonized in early Holocene times by foragers, not farmers; that farming and domestic-animal keeping developed in North Africa as a desperate attempt to improve life by improving the foraging system; and that Proto-Afroasiatic “farming terms,” cited as evidence of farming practices, refer to the collection of wild plants rather than the cultivation and processing of domesticated species. Zvelebil argues that many foragers were relatively sedentary, whereas many Neolithic farmers incorporated foraging strategies into their economy, which blurs any differences in population-growth potential between the two subsistence strategies. Moreover, the marked cultural continuity between Mesolithic and Neolithic traditions in Europe suggest that most migrations were short-distance ones leading to admixture between adjacent farming and foraging groups that were already somewhat genetically related. Scarre’s review of the origin, development, and expansion of Linearbandkeramik early farming settlements on the Rhine/Meuse delta (c. 7,500 years ago) supports this view. These communities exhibit no evidence of rapid population growth or long-distance migration and were sometimes themselves absorbed by neighboring hunter/gatherer groups. Nor can language rescue the FLDH with regard to Europe. Comrie notes that applying the FLDH to Indo-European languages would require the assumption that Indo-European languages diverged at the almost improbably early time of 9,000 years ago, and he cites evidence that members of all major stocks of Indo-Europe were familiar with farming practices before dispersal. He concludes that Indo-European languages spread as a secondary dispersal, perhaps via elite dominance, with minimal population replacement.

Section 2 of part III includes nine papers devoted to Asia and Oceania. The chapter by Kivisild et al. attempts to identify the origins of major mitochondrial and Y-chromosome lineages in India. The paucity of mtDNA lineages originating in the West, as in Europe, suggests only a minor flow of genes from the

presumed origin of the farming/language dispersal. In contrast, however, most Indian Y-chromosomes cluster with those of European and western Asian populations, suggesting, together with the mtDNA evidence, a male-dominated migration from the West, probably from western or central Asia. The authors wisely conclude by cautioning readers to not hinge theories on calibrations of the molecular clock or on inadequate sampling.

The chapter by Paz and the chapter by Pawley cover the most-recent alleged farming/language dispersal—that by speakers of the Proto-Oceanic branch of Austronesia—for which Bellwood has emphasized the role of agriculture and Lapita culture. Linguistic and archaeological data place the origins of the Proto-Oceanic branch in south China, Taiwan, or the Philippines, after which it reached Fiji, via eastern Indonesia, by ~3,000 years ago and then went on to the Bismarck Archipelago. Paz argues that the Philippines was a friction—not a spread—zone, causing the cultural package to break up early and to spread piecemeal at different rates through the islands of Southeast Asia. In contrast, Polynesia was uninhabited (in fact, uninhabitable) when Neolithic farmers first arrived, and extant Polynesians provide a valid portrait of the Neolithic immigrant genome uninfluenced by admixture with indigenous foragers. However, Oppenheimer and Richards show that local mutations in coding genes, mtDNA, and Y chromosomes indicate mixed origins (i.e., Taiwan as well as Indonesia and near Oceania) for Polynesians and suggest that the issue of an expansion of the Austronesian languages has little to do with their origin. The farming-dispersal component of the FLDH with regard to Polynesia fares little better. Although farming was probably necessary for the expansion into Polynesia, it was not sufficient (the outrigger canoe and complex navigational skills were also necessary), nor is there evidence that the farming included rice (rather, it may have included only indigenous Melanesian root and tree crops). Two other chapters, one by Phillipson and one by Hudson, provide the clearest examples of a link between agricultural dispersals and language change: the Bantu expansion from the Cameroon region of Africa eastward and then southward and the spread of Yayoi culture to Japan from the Korean Peninsula. Regrettably, neither author cites genetic evidence for his arguments.

The third region-specific section of part III, the shortest of the four sections, is devoted to Mesoamerica and the U.S.

Southwest. Hill's hypothesis that speakers of Uto-Aztecan introduced maize farming, as well as their language, to the U.S. Southwest from central Mexico reverses conventional wisdom regarding the homeland of the Uto-Aztecan languages, on the basis of rather slim evidence from the Hopi dictionary. She also advances the controversial hypothesis that northern Uto-Aztecs of the Great Basin once farmed but abandoned agriculture in the face of climate changes. Matson argues that the eastern Basketmaker II (BM II) folk derive from the indigenous Archaic tradition and are ancestral to Tanoan (eastern Pueblo) speakers who adopted agriculture from Anasazi immigrants descended from western BM II peoples who are ancestors of contemporary western Pueblo peoples. Matson cites skeletal morphological evidence in support of his hypothesis, which he sees as consistent with the FLDH, but, unfortunately, he ignores crucial evidence from both conventional and DNA markers.

I recommend this book to anyone interested in the origins and movements of populations. Like the controversial work of Greenberg, Turner, and Zegura, "The Settlement of the Americas," and that of Cavalli-Sforza and his colleagues on both the correlations between genes and major language families and the demic expansions of Neolithic farmers from the Near East to Europe, this book provides a paradigm for organizing and testing hypotheses about relationships among genes, material culture, and language. It is tempting to believe Greenberg's argument that "triangulation" has more interpretive power than any one discipline alone. There are many pieces to the farming/language dispersal puzzle that do not fit and that need serious tweaking, but there are many contingencies with which to work. The challenge, as always, will be to keep theory, fact, and fiction each in their proper places.

DAVID GLENN SMITH

*Department of Anthropology  
University of California  
Davis, CA*

© 2004 by The American Society of Human Genetics. All rights reserved.  
0002-9297/2005/7601-0023\$15.00