CRC

# DNA SYSTEMATICS

## Volume III
## Human and Higher Primates

S. K. Dutta
William P. Winter

CRC PRESS

# DNA Systematics

## Volume III: Human and Higher Primates

Editors

**S. K. Dutta, Ph.D.**

Professor
Departments of Botany and
Genetics and Human Genetics
Howard University
Washington, D. C.

**William P. Winter, Ph.D.**

Senior Biochemist
Center for Sickle Cell Disease
Associate Professor
Departments of Genetics and
Human Genetics and Medicine
Howard University
Washington, D. C.


CRC Press, Inc.
Boca Raton, Florida

**Library of Congress Cataloging in Publication Data**
(Revised for vol. 3)

DNA systematics.

Includes bibliographies and index.
Contents: v. 1. Evolution — v. 2. Plants — v. 3. Human and higher primates.
1. Deoxyribonucleic acid. 2. Recombinant DNA. 3. Chemotaxonomy
I. Dutta, S. K. (Sisir K.)
QP624.D19 1986 574.87'3282 85-21285
ISBN 0-8493-5820-5



Direct all inquiries to CRC Press, Inc., 2000 Corporate Blvd., N.W., Boca Raton, Florida, 33431.



International Standard Book Number 0-8493-5820-5 (v.1)
International Standard Book Number 0-8493-5821-3 (v.2)
International Standard Book Number 0-8493-5822-1 (v.3)

Library of Congress Card Number 85-21285
Printed in the United States

## DNA Systematics

### Volume I:
### Evolution

The Role of the Computer in Estimates of DNA Nucleotide Sequence Divergence. Mobile DNA Sequences and Their Possible Role in Evolution. Ribosomal RNA Processing Sites. Analysis of Small RNA Species: Phylogenetic Trends. Mitochondrial DNAs and Phylogenetic Relationships. Evolution of Histone Genes. Phylogeny of Normal and Abnormal Hemoglobin Genes. Transcriptional Regulatory Sequences of Phylogenetic Significance.

### Volume II:
### Plants

Plant DNA: Contents and Systematics. Repeated DNA Sequences and Polyploidy in Cereal Crops. Homology of Nonrepeated DNA Sequences in Phylogeny of Fungal Species. Chloroplast DNA and Phylogenetic Relationships. rDNA: Evolution Over a Billion Years. 23s rRNA-Derived Small Ribosomal RNAs: Their Structure and Evolution with References to Plant Phylogeny. Molecular Analysis of Plant DNA Genomes: Conserved and Diverged DNA Sequences. A Critical Review of Some Terminologies Used for Additional DNA in Plant Chromosomes.

### Volume III:
### Human and Higher Primates

Phylogeny of Hominoids, Based on DNA-DNA Hybridization. The Ribosomal DNA Sequence for the Study of the Evolution of Human, Chimpanzee and Gorilla. A Model for the Evolution of Apolipoprotein Genes from Oligonucleotide Repeats. The Molecular Genetics of the Skin. Mitochondrial DNA: Variation in Humans and Higher Primates. Analysis of DNA Sequences Containing Transcriptional Enhancer Elements. Analysis of the DNA Sequence Homology of Amino terminus of Various Acetylcholine Receptor Subunits. Genetic Polymorphisms of Human Populations in the Tropics: Prospects for DNA Studies. Restriction Fragment Length Polymorphism of DNAs from Chinese Populations.

# PREFACE

In recent years, numerous studies have been performed that use various characteristics of DNA to estimate the diversity or relatedness between both closely and distantly related species. Some of these studies have been concerned with experimental microevolution dealing with the accumulation of relatively small changes within a species, and some with macroevolution involving taxonomic categories and measurements taken over long periods of time or on a geological scale. Most of this explosion of knowledge has been due to the utilization of the powerful methods of recombinant-DNA technology, and a new interdisciplinary science has evolved spontaneously which may be called "DNA Systematics". Included in this science are the characterization of DNA in nuclear and cytoplasmic genomes; DNA:DNA reassociation kinetics of repeated and nonrepeated DNA sequences; thermal stability measurements of heteroduplexes; restriction enzyme pattern analysis of specific nuclear and non-nuclear DNA segments, rDNAs, and mitochondrial and chloroplast DNAs; the rate of evolution of cell organelle genomes vs. nuclear genomes; the implication of gene duplication and gene fusion in evolution and the evolutionary history of specific genes like rDNA genes and hemoglobin genes; evolutionary trends in regulation; and species specificity and DNA sequencing of processing sites of introns and different RNA maturation sites.

Historically, DNA systematics studies were initiated more than 25 years ago by Ellis Bolton, Roy Britten, David Kohne, Brian McCarthy, and others at Carnegie Institution of Washington, Washington, D.C., and a few other scientists of England, France, and of U.S.S.R. whose work has been reviewed by Belozersky and Antonov during 1972 and 1980 at Moscow University Press, U.S.S.R. Their techniques were mostly based on DNA:DNA hybridization, which is now claimed as the "most favorable of all" methods for revealing family trees, as discussed by Lewin.* Unfortunately, there is no available comprehensive treatise of this vast amount of new knowledge particularly on microevolution based on studies in DNA systematics using new tools of recombinant-DNA technology. In the present work we attempt the first comprehensive review of new information on DNA systematics.

The enormous amount of accumulated information has been reviewed by authors who are active in their respective areas and then organized into three volumes. Volume I is devoted to general topics of DNA systematics with respect to general evolution. Hobish reviews the present state-of-the-art use of computers for storage and retrieval of DNA research data. The role of movable elements in evolution and species formation is reviewed by Georgiev and his associates. It is well established now that mobile DNA sequences provide variability for natural selection and for evolutionary jumps. It makes genomes flexible, and mobile sequences are widespread in living creatures. Studies have been performed on the evolutionary significance of various control mechanisms which regulate speciation and evolution. Studies dealing with the regulation of ribosomal RNA processing sites, regulation of transcription, and analysis of various small RNAs along with their phylogenic significance are reviewed by Crouch and Bachellerie; Huang; and Beljanski and Le Goff, respectively. The examination of mitochondrial genomes from mammals, *Drosophila*, and fungi has produced models of mt-DNA variation and offers a comparative treatment of evolution with nuclear genomes; these investigations are reviewed by Birley and Croft. Two of the most important gene sets were selected for discussion in this volume. One is the histone gene set, most gene sets of which do not have introns, and the other is the hemoglobin gene set, which does have introns. Enormous amounts of information on these gene sets on the evolution of *Xenopus*, avians, rodents, and higher primates including humans are reviewed by Marzluff; and Winter, respectively.

The second volume is devoted primarily to the DNA systematics of plants, although, where

* Lewin, R., DNA reveals surprises in human family tree: the application of DNA-DNA hybridization, *Science*, 226, 1179, 1984.

necessary, reference species other than plants have been included to present a complete story. This volume starts with the classical approach to plant systematics using knowledge obtained from the contents of plant nuclear DNAs; this material is reviewed by Ohri and Khoshoo. Plant species, particularly higher green plants, show polyploidy and have 70 to 75% repeated DNA sequences. Studies made on these repeated and single copy DNA sequences of monocot and dicot plants are reviewed by Mitra and Bhatia; and Antonov, respectively. Appels and Honeycutt; and Troitsky and Bobrova have reviewed extensive studies done on ribosomal RNA genes of plants along with information obtained from other species and have given an extensive analysis of the phylogenic significance of these studies. The DNA systematics of some fungal species are reviewed by Ojha and Dutta. The chloroplast genomes of green plants have provided excellent information of plant systematics; this information is reviewed by Palmer, whose group has done extensive work with chloroplast genomes of various plants. A critical glossary of different terminologies used in plant DNA systematics is given by A. K. Sharma.

The third and final volume of this series, co-edited by S. K. Dutta and William P. Winter, is devoted primarily to the DNA systematics and evolution of man and higher primates. Charles Sibley and Jon Alquist have reviewed our current understanding of the phylogeny of hominoids at the level of DNA while Roy Schmickel and co-authors have examined the question of human-chimpanzee-gorilla relationships from the standpoint of ribosomal genes. The new and exciting applications of mitchondrial DNA as a phylogenetic tool are elegantly treated by Honeycutt and Wheeler. The contributions of specific genes and gene systems have been likewise included. Rajavashisth, Lusis, and Kaplan have described the evolution of apolipoprotein genes, Nicholls has reviewed work by himself and others on the gene systems involved in skin mutations in that system. Roy and Dutta have dealt with acetylcholine receptor subunits. Finally, the study of DNA polymorphisms in human populations in the tropics has been discussed by Sharma while Hsueh has assembled an extensive review of RFLP polymorphisms in Chinese populations. Taken together, these chapters thus comprise a broad overview of the current contributions of DNA systematics to our understanding of phylogeny and evolution in our own and closely related species.

These three volumes are expected to be valuable references, not only to students of evolution but also to others interested in efficient germ plasm resource maintenance and utilization, and fields which are vital for planning plant and animal breeding programs. Knowledge of DNA markers correlating the geographic distribution of genes responsible for heritable diseases such as human sickle cell anemia should be of profound importance to physicians and epidemiologists.

In addition to our contributing authors, who have also helped in reviewing several chapters, several other authors have helped in organizing and improving various chapters. We would like to acknowledge particularly Fransciso Ayala of the University of California, Davis; Igor Dawid, H. Westphal, and A. Schecter of the National Institutes of Health, Bethesda, Maryland; Professor A. K. Sharma of Calcutta University, Calcutta, India; R. L. Pearson, George Mathew, and D. R. Maglott of Howard University, Washington, D.C.; H. James Price, Texas A&M University, College Station, Texas; H. R. Chen, National Biomedical Research Foundation, Georgetown University Medical Center, Washington, D.C.; Bill Hoyer of Georgetown University; E. S. Weinberg of the University of Pennsylvania, Philadelphia; and G. N. Wilson, Pediatric Genetics, The University of Michigan, Ann Arbor.

**S. K. Dutta**
**William P. Winter**
**Co-Editors**

# THE EDITORS

**Sisir K. Dutta, Ph.D.,** is Professor of Molecular Genetics in the Department of Botany and Graduate Professor in the Department of Genetics and Human Genetics at Howard University, Washington, D.C.

Dr. Dutta obtained his B.S. degree from Dacca University, Bangladesh in 1949, and thereafter, served for 6 years as Research Assistant in Genetics and Plant Breeding for the government of West Bengal, India. He received his M.S. and Ph.D. degrees in genetics from Kansas State University, Manhattan, Kansas in 1958 and 1960, respectively. He was a research associate and/or visiting scientist at the University of Chicago, Columbia, and Rockefeller University in New York, the Pasteur Institute in Paris, the National Institutes of Health in Bethesda, Maryland, and Rice University in Houston, Texas. He was Chief Research Officer-cum-Director of the National Pineapple Research Institute of Malaysia from 1960 to 1964, Chairman fo the Division of Natural Sciences, and Chairman of the Biology Department of the Texas Christian University Affiliated College at Hawkins, at Texas Southern University from 1964 to 1967. In 1967 he assumed his present duties at Howard University.

He has been the organizer, chairman, and speaker of several national and international symposia held in the U.S., U.S.S.R., Europe, and Asia as well as international coordinator of a research cooperative on biological nitrogen fixation in rice farming systems. He has been a member of the editorial board of the *East Pakistan Agricultural Journal,* a reviewer and panelist of several government and private agencies. He has been inducted as a personality in America's Hall of Fame for his contribution in molecular genetics, has appeared in *Who's Who in the World, Who's Who in America,* and *Who's Who in Frontier Sciences and Technology.* He is a member of several national and international professional societies, author or co-author of more than 100 papers including monographs and book chapters and editor of five books. He has been a recipient of several research awards for the U.S. National Science Foundation, National Institutes of Health, Department of Energy, Environmental Protection Agency, Research Corporation, Anna Fuller, Fund and several other agencies including the United Nations Development Projects.

His current research interest is in the areas of regulation of ribosomal RNA transcription and processing, molecular genetics of neuron-specific genes, and molecular evolution.

**William P. Winter, Ph.D.,** is Senior Biochemist at Howard University Center for Sickle Cell Disease, Associate Professor of Medicine and Associate Professor of Genetics and Human Genetics in the Graduate School of Arts and Sciences of Howard University in Washington, D.C.

Dr. Winter received his B.S. degree in Biochemistry from Pennsylvania State University in 1960 and his M.S. and Ph.D. degrees from the same institution in 1962 and 1965, respectively. He spent the next 4 years in the laboratory of Dr. Hans Neurath at the University of Washington, Seattle, Washington, where he worked on comparative and evolutionary aspects of proteolytic enzymes. From there, Dr. Winter went to the Department of Human Genetics at the University of Michigan, Ann Arbor, Michigan, where he became involved in the structural aspects of human hemoglobin variants and in the study of sickle hemoglobin. He left Michigan in 1977 to move to the Center for Sickle Cell Disease at Howard University where he has continued his research on molecular aspects of variant hemoglobins and on chemical intervention in the sickle cell anemia.

Dr. Winter is a member of the American Chemical Society, the American Society of Hematology, the Sigma Xi, the American Society of Human Genetics, the New York Academy of Science, and the American Association for the Advancement of Science and is a past member of Phi Lambda Upsilon Chemical Honorary Fraternity. He has served as a member of the Advisory Council of the International Hemoglobin Information Center and as a consultant to

NASA and to the Princess Margaret Hospital, Nassau, Bahamas. He is currently a member of the Working Group of the National Committee of Clinical Laboratory Standards and a consultant to the District of Columbia' Office of Child and Maternal Health.

Dr. Winter is the author of numerous publications and abstracts in the areas of comparative protein structure and abnormal human hemoglobins. His previous contributions to CRC Press publications include a contribution to the *Handbook of Biochemistry Selected Data for Molecular Biology;* a chapter on "Chromatography of Hemoglobin" in the *Handbook Series in Clinical Laboratory Science, Section I: Hematology,* Volume IV; the two-volume work entitled *Hemoglobin Variants in Human Populations;* and a chapter on "Phylogeny of Normal and Abnormal Hemoglobin Genes" in Volume I of this series.

# CONTRIBUTORS

**Jon E. Ahlquist, Ph.D.**
Department of Zoological and
Biomedical Sciences
Ohio University
Athens, Ohio

**S. K. Dutta, Ph.D.**
Professor
Departments of Botany and Genetics
and Human Genetics
Howard University
Washington, D.C.

**Iris Laudien Gonzalez, Ph.D.**
Assistant Professor
Department of Pathology
Hahnemann University
Philadelphia, Pennsylvania

**Rodney L. Honeycutt, Ph.D.**
Associate Professor
Department of Wild Life and Fisheries
Sciences
Texas A & M University
College Station, Texas

**Jing Lun Hsueh, Ph.D.**
Institute of Genetics
Fudan University
Shanghai, People's Republic of China

**John S. Kaptein, Ph.D.**
Assistant Research Professor
Comprehensive Cancer Center
University of Southern California
Los Angeles, California

**Aldons J. Lusis, Ph.D.**
Associate Professor
Departments of Medicine and
Microbiology
Molecular Biology Institute
University of California
Los Angeles, California

**E. Maxwell Nicholls, M.D.**
Senior Lecturer
Centre for Biomedical Engineering
University of New South Wales
Kensington, New South Wales, Australia

**Tripathi B. Rajavashisth, Ph.D.**
Assistant Professor
Division of Medical Genetics
UCLA School of Medicine
Torrance, California

**Nirmal K. Roy, Ph.D.**
Insititute of Environmental Medicine
NYU Medical Center
New York, New York

**Roy D. Schmickel, M.D.**
Department of Human Genetics
University of Pennsylvania
Philadelphia, Pennsylvania

**Archana Sharma, Ph.D.**
Professor of Genetics
Department of Botany
Center for Advanced Study in Cell and
Chromosome Research
University of Calcutta
Calcutta, India

**Charles G. Sibley, Ph.D.**
Professor of Biology
Tiburon Center
San Francisco State University
Tiburon, California

**Dwight Stambolian, Ph.D.**
Assistant Professor
Department of Human Genetics
University of Pennsylvania
Philadelphia, Pennsylvania

**James Sylvester, Ph.D.**
Associate Professor
Department of Pathology
Hahnemann University
Philadelphia, Pennsylvania

**Mukesh Verma, Ph.D.**
Research Scientist
Department of Biochemistry
George Washington University School
of Medicine
Washington, D.C.

**Ward C. Wheeler, Ph.D.**
Research Fellow
Molecular Biology Institute
University of California
Los Angeles, California

**William P. Winter, Ph.D.**
Senior Biochemist
Center for Sickle Cell Disease
Associate Professor
Departments of Genetics and Human Genetics and Medicine
Howard University
Washington, D.C.

# TABLE OF CONTENTS

## Volume III

# TABLE OF CONTENTS

## Volume III

Chapter 1

# PHYLOGENY OF THE HOMINOIDS, BASED ON DNA-DNA HYBRIDIZATION

**Charles G. Sibley and Jon E. Ahlquist**

## TABLE OF CONTENTS

# I. INTRODUCTION

The living members of the superfamily Hominoidea are man *(Homo sapiens)*, common chimpanzee *(Pan troglodytes)*, pygmy chimpanzee or bonobo *(Pan paniscus)*, gorilla *(Gorilla gorilla)*, orangutan *(Pongo pygmaeus)*, and the nine species of gibbons *(Hylobates)*. The next nearest lineage, or sister group, of the hominoids is the Cercopithecoidea, including the Old World monkeys, baboons, macaques, colobines, etc. The composition of these groups and their sister group relationship are not in dispute, but the branching sequence within the Hominoidea, and the dating of the divergences of the hominoid lineages, are still being debated.

The literature pertaining to these problems is large and we will not attempt to review or evaluate the evidence of hominoid relationships based on morphology, fossils, or behavior. In two recent papers[1,2] we reviewed most of the previous evidence based on "molecular" methods. In this chapter, we present a synopsis of our DNA-DNA hybridization data.

# II. THE HOMINOID BRANCHING PATTERN

In recent years, at least five branching sequences have been proposed for the five hominoid genera, with the cercopithecoids as the outgroup. All authors accept the gibbons as the living descendants of the oldest hominoid branch; thus, the disagreements concern the branching sequence of the great apes (orangutan, gorilla, chimpanzee) and man.

Some studies have placed the orangutan lineage as the sister group of the gorilla and chimpanzees[3] or as forming a clade with man distinct from that containing the African apes.[4,5] However, based on fossils[6-10] and molecular evidence,[1,2,11-20] the consensus is that the orangutan lineage branched next after the gibbon clade, and that the gorilla, chimpanzees, and man form a clade. Therefore, we will consider only the following four branching patterns for the African apes and man in which the taxa are listed in the order of branching from oldest to youngest, as in Figure 1.

1. Gorilla-chimpanzee-human
2. Human-chimpanzee-gorilla
3. Chimpanzee-gorilla-human
4. A trichotomy, from which the lineages of the gorilla, chimpanzees, and man branched simultaneously

An exact trichotomy is unlikely to occur in a phylogeny; thus, when a method produces a trichotomy, it is probable that there are two branches so close together that they are not being resolved. Several molecular methods have produced a trichotomy, including comparisons of the sequences of 2251 bases of the eta locus, one of the five beta-related globin genes.[20]

A chimpanzee-gorilla-human sequence has received little support from any source, and is not indicated by any of the molecular data. We will not discuss it further.

The principal debate is between the proponents of branching sequences 1 and 2. Most morphological studies favor 2, for example.[9,10,21-23] Our DNA-DNA hybridization data favor branching sequence 1.

# III. METHODS

The DNA-DNA hybridization technique measures the median base sequence divergence between the reassociated DNA single strands from the same species, or from two different species, by measuring the temperature required to melt the hydrogen bonds between base pairs thereby converting double-stranded DNA into single strands.

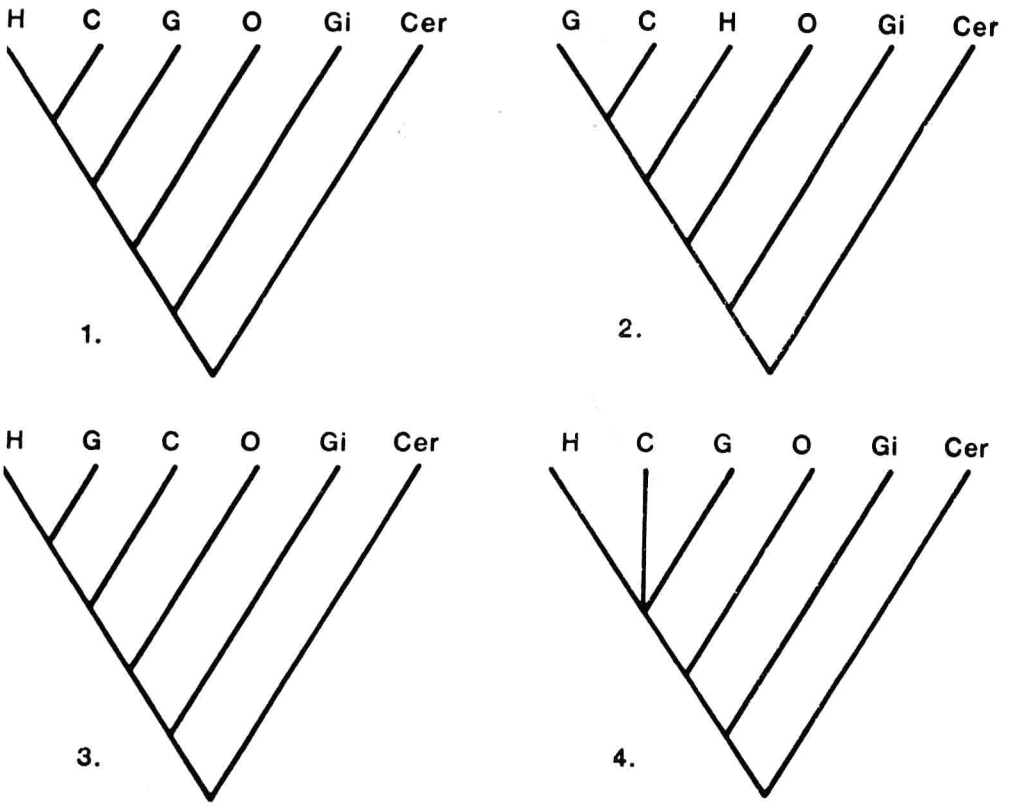


FIGURE 1. The four branching patterns most frequently proposed for the hominoid lineages. H: human, C: chimpanzee, G: gorilla, O: orangutan, Gi: gibbons, Cer: cercopithecidae.

The thermal dissociation curve of the double-stranded DNA molecules formed by the "hydridization" of single strands from two different species is a plot of the different rates of base substitution that have occurred in different sequences (Grula et al.[24]). Those hybrid duplexes composed of homologous sequences that have diverged most rapidly, and hence are most different, will contain the largest percentage of mismatched base pairs and will, therefore, melt at the lowest temperatures. The DNA-DNA hybrids that do not melt until exposed to the highest temperatures are composed of sequences that have evolved slowly and contain few, if any, mismatched bases. Therefore, the median *amount* of divergence is a measure of the median *rate* of divergence. In the absence of evidence to the contrary, we will assume that the frequency distribution of rates of base substitution is normal; thus, the median sequence divergence measures the average rate of divergence.

We have published descriptions of our methods,[25-27] a synopsis follows.

DNA extracts were obtained from the nuclei of tissue cells, purified to remove proteins, ribonucleic acids, and other non-DNA components, and sheared by sonication into fragments with an average length of 500 bases. The single-stranded fragments of the species to be used as radiolabeled "tracers" were allowed to reassociate to a $C_0t$ of 1000 at 50°C in 0.48 *M* sodium phosphate buffer. ($C_0t$ = the initial concentration of single-stranded DNA in moles per liter, times the length of incubation in seconds.) This permitted most of the rapidly reassociating repeated sequences to form double-stranded molecules while the slowly reassociating single-copy sequences remained single-stranded. The latter were recovered by chromatography on a hydroxyapatite (HAP) column. This process produced a single-copy DNA preparation consisting of one copy per genome of each original single-copy sequence plus *at least* one copy per

genome of each repeated sequence that was kinetically different under these conditions. The single-copy DNA of the "tracer" species was labeled with radioiodine ($^{125}I$).

DNA-DNA hybrids were formed from a mixture composed of one part (=200 ng) of the tracer DNA and 1000 parts (=200 μg) of sheared, whole DNA of the "driver" species. The 1:1000 ratio between tracer and driver sequences increases the probability that the radiolabeled tracer DNA sequences will form heteroduplexes with the driver DNA, rather than homoduplexes by reassociating among themselves. Under these conditions only about 1% of the duplexes that form are between tracer fragments.

The hybrid combinations were heated to 100°C for 5 min to dissociate the double-stranded molecules into single strands, then incubated for 120 hours at 60°C in 0.48 *M* sodium phosphate buffer to permit the single strands to form hybrid duplexes. After incubation the buffer was diluted to 0.12 *M* and the hybrids were bound to HAP columns immersed in a temperature-controlled water bath at 55°C. The temperature was then raised in 2.5°C increments from 55°C to 95°C. At each of 17 temperatures the single-stranded fragments produced by the melting of duplexes were eluted in 20 ml of 0.12 *M* sodium phosphate buffer. The radioactivity in each sample was counted and the counts were used to calculate the $T_{50}H$ values.

We use the $T_{50}H$ statistic as a measure of the genealogical distance between the two taxa composing a DNA-DNA hybrid. $T_{50}H$ is the temperature in degrees Celsius in an ideal, normalized, cumulative frequency distribution function at which 50% of all single-copy DNA sequences are in the duplex form, and 50% have dissociated into single strands. Thus, the delta $T_{50}H$ measures the median sequence divergence between the genomes of the two taxa forming a DNA-DNA hybrid and, as noted above, the median divergence is the result of the median rate of nucleotide substitution.

In the calculation of $T_{50}H$ it is assumed that the single-copy sequences in the genomes of the two species being compared have homologs in the other species, that all single-copy sequences potentially can hybridize with their homologs, and that all degrees of divergence can be detected. For homologous hybrids the $T_{50}H$ and $T_m$ values derived from normalized cumulative distributions are equal. For hybrids between more diverged taxa the percentage of hybridization declines and the thermal stability curve is progressively truncated by the effects of experimental conditions. It is possible to estimate the $T_{50}H$ by making a graphic extrapolation of the portion of the sigmoid melting curve that is most nearly linear. A better procedure is to calculate the best-fitting linear regression to the part of the curve having the most constant slope and to find the temperature corresponding to its intercept with the 50% hybridization level. The most objective procedure is to fit a cumulative distribution function to the data and find its intercept with the temperature axis at the 50% hybridization level.

To convert the delta $T_{50}H$ values into a phylogeny requires a procedure to obtain a hierarchical clustering of taxa. We use the "average linkage" unweighted-pair-group (UPGMA) procedure which begins by averaging the hybrids for each pairwise set of comparisons, then clustering the closest pair or pairs of taxa. The next step links the taxa which have the smallest average distance to any existing cluster. This procedure continues until all taxa are linked. DNA hybridization data are compatible with this method because there is no known reason, other than nucleotide sequence homology, to explain the thermal stabilities of DNA-DNA hybrids. We have also used other methods of clustering that do not assume equal branch lengths.

We measured the experimental error in our data and found that a single delta $T_{50}H$ value should be assumed to have a possible error of ±1.0. The delta values may be corrected for some sources of error (e.g., variation in fragment size) but we have found it possible to compensate for most of the sources of error by using five or more species and/or replicates for each pairwise comparison so that an average delta $T_{50}H$, its standard error (SE), and standard deviation (SD) can be calculated. The average SD of five or more delta $T_{50}H$ values for the same comparison is 0.35. Because of the incremental summation of values in the average linkage procedure, the older nodes in our phylogenies are often the averages of 20, 30, or more delta $T_{50}H$ measure-

**TABLE 1**
**Reduced and Folded Matrix of Delta $T_{50}H$ Values Produced by DNA-DNA Hybridization**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Pan troglodytes | | | | | | | |
| Pan paniscus | 0.73<br>0.30<br>N=6 | 0.73<br>0.30<br>N=6 | | | | | |
| Homo sapiens | 1.63<br>0.24<br>N=64 | 1.65<br>0.18<br>N=11 | 1.63<br>0.24<br>N=75 | | | | |
| Gorilla gorilla | 2.21<br>0.25<br>N=69 | 2.37<br>0.23<br>N=12 | 2.27<br>0.24<br>N=69 | 2.25<br>0.25<br>N=150 | | | |
| Pongo pygmaeus | 3.56<br>0.32<br>N=20 | 3.56<br>0.25<br>N=13 | 3.62<br>0.33<br>N=24 | 3.54<br>0.34<br>N=22 | 3.57<br>0.32<br>N=79 | | |
| Hylobates lar | 4.82<br>0.40<br>N=18 | 4.87<br>0.52<br>N=6 | 4.76<br>0.34<br>N=17 | 4.72<br>0.30<br>N=24 | 4.83<br>0.34<br>N=20 | 4.79<br>0.35<br>N=85 | |
| Cercopithecidae | 7.30<br>0.41<br>N=22 | 7.01<br>0.38<br>N=11 | 7.36<br>0.32<br>N=22 | 7.18<br>0.41<br>N=28 | 7.43<br>0.41<br>N=15 | 7.05<br>0.40<br>N=8 | 7.25<br>0.40<br>N=106 |
| | Pan trog. | Pan pan. | Homo | Gorilla | Pongo | Hylobat. | Cercopit. |

*Note:* The top value in each set of three is the average delta $T_{50}H$; the middle value is the standard deviation; the bottom value is the number of DNA-DNA hybrids that were averaged. The average, standard deviation, and number of DNA hybrids for each horizontal row are given in the boxes at the ends of the rows.

ments. The accuracy of the average delta $T_{50}H$ for each node is proportional to the number of DNA hybrids and any node based on more than five will usually have a standard error of the mean of ±0.2, or less.

## IV. RESULTS

Table 1 is a reduced and folded matrix based on 501 delta $T_{50}H$ values. Figure 2 is a phylogram derived from an average linkage analysis of the data. This method assumes equal branch lengths. Figure 3 is based on a method that does not assume equal branch lengths and it shows that some of the branches differ in length, with the gibbon *(Hylobates)* branch being the shortest. The different branch lengths may be due to different average genomic rates of nucleotide substitution, to experimental error or both.

It is apparent from the phylograms that the branches are well resolved, except for those of the gorilla lineage vs. the human-chimpanzee lineage. The distance between the averages for the gorilla branch and the human-chimpanzee branch is only 0.7 delta $T_{50}H$ units, and there is an overlap of 0.6 delta units in the ranges of the values for these two nodes.

Statistical tests of the signigficance of the difference between the Gorilla branch and the human-chimpanzee branch have produced controversy. Templeton[28] used "Q" statistics to test our data[1] and concluded that this test supports our phylogeny, but not significantly. Saitou,[29]