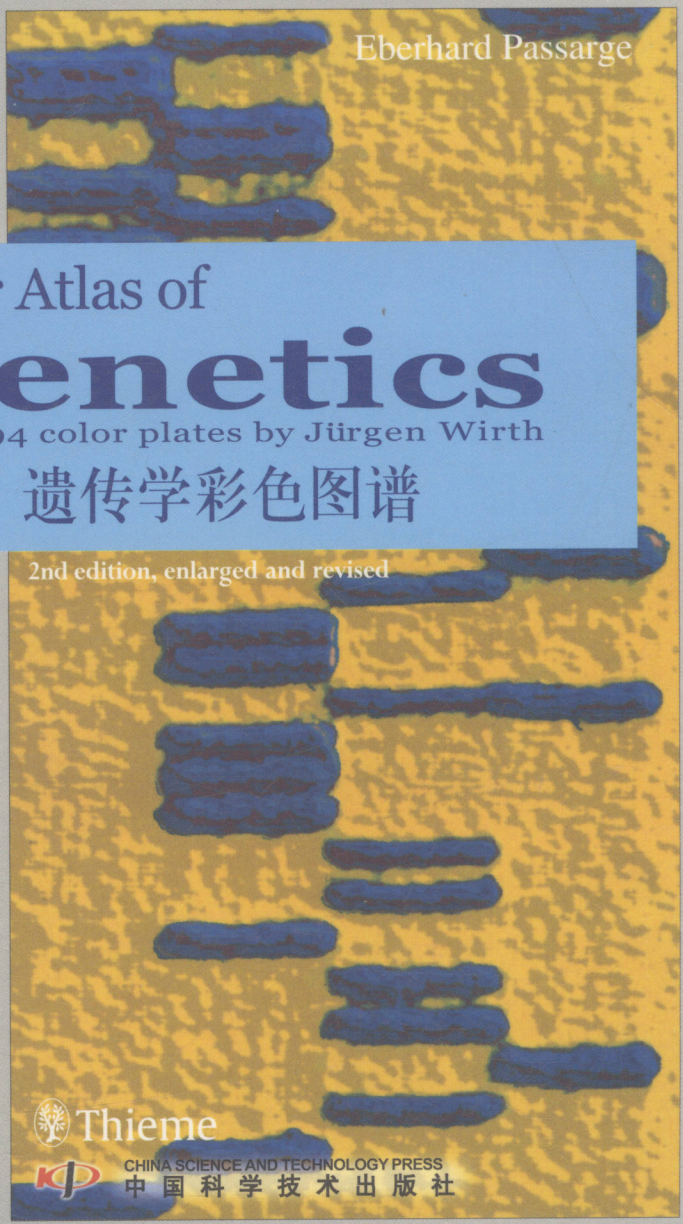
THIEME FLEXIBOOK



Eberhard Passarge

Color Atlas of Genetics

With 194 color plates by Jürgen Wirth

遗传学彩色图谱

2nd edition, enlarged and revised

Thieme

CHINA SCIENCE AND TECHNOLOGY PRESS
中国科学技术出版社

basic sciences

Color Atlas of Genetics

Eberhard Passarge, M.D.
Professor of Human Genetics
Institute of Human Genetics
University of Essen
Essen, Germany

Second edition, enlarged and revised

With 194 color plates by Jürgen Wirth

Thieme
Stuttgart · New York

图书在版编目（CIP）数据

遗传学彩色图谱/（德）帕萨尔格(Passarge, E.)著. —北京：中国科学技术出版社，2006.1

ISBN 7-5046-4222-3

Ⅰ. 遗... Ⅱ.①帕... ②E... Ⅲ. 遗传学—图谱 Ⅳ. Q3-64

中国版本图书馆 CIP 数据核字（2005）第 135574 号

著作权合同登记号 北京市版权局图字：01-2005-5210 号

策划编辑 肖 叶 单 亭

责任编辑 郭 璟

责任印制 安利平

法律顾问 宋润君

中国科学技术出版社出版

北京市海淀区中关村南大街 16 号 邮政编码：100081

http://www.kjpbooks.com.cn

电话：010-62103210 传真：010-62183872

科学普及出版社发行部发行

北京汇文商业艺术制作有限公司 · 制版

上海盛杰印刷厂印刷

开本：850×1165 毫米 1/32 印张：15 字数：358 千字

2006 年 1 月第 1 版 2006 年 1 月第 1 次印刷

印数：1—2000 册 定价：68.00 元

Color Atlas of Genetics
2nd edition

To my wife, Mary

Preface

Knowledge about genes (genetics) and genomes (genomics) of different organisms continues to advance at a brisk pace. All manifestations of life are determined by genes and their interactions with the environment. A genetic component contributes to the cause of nearly every human disease. More than a thousand diseases result from alterations in single known genes.

Classical genetics, developed during the first half of the last century, and molecular genetics, developed during the second half, have merged into a fascinating scientific endeavor. This has provided both a theoretical foundation and a broad repertoire of methods to explore cellular mechanisms and to understand normal processes and diseases at the molecular level.

Deciphering the genomes of many different organisms, including bacteria and plants, by determining the sequence of the individual building blocks—the nucleotide bases of deoxyribonucleic acid (DNA)—will augment our understanding of normal and abnormal functions. The new knowledge holds promise for the design of pharmaceutical compounds aimed at individual requirements. This will pave the way to new approaches to therapy and prevention. Insights are gained into how organisms are related by evolution.

Students in biology and medicine face an enormous task when attempting to acquire the new knowledge and to interpret it within a conceptual framework. Many good textbooks are available (see General References, p. 421). This Color Atlas differs from standard textbooks by using a visual approach to convey important concepts and facts in genetics. It is based on carefully designed color plates, each accompanied by a corresponding explanatory text on the opposite page.

In 1594 Mercator first used the term "atlas" for a collection of maps. Although maps of genes are highly important in genetics, the term atlas in the context of this book refers to illustrations in general. Here they provide the basis for an introduction, hopefully stimulating interest in an exciting field of study.

This second edition has been extensively revised, rewritten, updated, and expanded. A new section on genomics (Part II) has been added. Twenty new plates deal with a variety of topics such as the molecular bases of genetics, regulation and expression of genes, genomic imprinting, mutations, chromosomes, genes predisposing to cancer, ion channel diseases, hearing and deafness, a brief guide to genetic diagnosis, human evolution, and many others. The Chronology of Important Advances in Genetics and the Definitions of Genetic Terms have been updated. As in the first edition, references are included for further reading. Here and in the list of general references, the reader will find access to more detailed information than can be presented in the limited space available. Websites for further information are included.

A single-author book cannot provide all the details on which scientific knowledge is based. However, it can present an individual perspective suitable as an introduction. In making the difficult decisions about which material to include and which to leave out, I have relied on 25 years' experience of teaching medical students at preclinical and clinical levels. I have attempted to emphasize the intersection of theoretical fundaments and the medical aspects of genetics, taking a broad viewpoint based on the evolution of living organisms.

All the color plates were produced as computer graphics by Jürgen Wirth, Professor of Visual Communication at the Faculty of Design, University of Applied Sciences, Darmstadt. He created the plates from hand drawings, sketches, photographs, and photocopies assembled by the author. I am deeply indebted to Professor Jürgen Wirth for his most skilful work, the pleasant cooperation, and his patience with all of the author's requests. Without him this book would not have been possible.

Essen, November 2000 *E. Passarge*

Acknowledgements

In updating, revising, and rewriting this second edition, I received invaluable help from many colleagues who generously provided information and advice, photographic material, and other useful suggestions in their areas of expertise: Hans Esche, Essen; Ulrich Langenbeck, Frankfurt; Clemens Müller-Reible, Würzburg; Maximilian Muenke, Bethesda, Maryland; Stefan Mundlos, Berlin; Alfred Pühler, Bielefeld; Gudrun Rappold, Heidelberg; Helga Rehder, Marburg; Hans Hilger Ropers, Berlin; Gerd Scherer, Freiburg; Evelyn Schröck, Bethesda, Maryland; Eric Schulze-Bahr, Münster; Michael Speicher, München; Manfred Stuhrmann-Spangenberg, Hannover; Gerd Utermann, Innsbruck; and Douglas C Wallace and Marie Lott, Atlanta.

In addition, the following colleagues at our Department of Human Genetics, University of Essen Medical School, made helpful suggestions: Beate Albrecht, Karin Buiting, Gabriele Gillessen-Kaesbach, Cornelia Hardt, Bernhard Horsthemke, Frank Kaiser, Dietmar Lohmann, Hermann-Josef Lüdecke, Eva-Christina Prott, Maren Runte, Frank Tschentscher, Dagmar Wieczorek, and Michael Zeschnigk.

I thank my wife, Dr. Mary Fetter Passarge, for her careful reading and numerous helpful suggestions. Liselotte Freimann-Gansert and Astrid Maria Noll transcribed the many versions of the text. I am indebted to Dr. Clifford Bergman, Ms Gabriele Kuhn, Mr Gert Krüger, and their co-workers at Thieme Medical Publishers Stuttgart for their excellent work and cooperative spirit.

About the Author

The author is a medical scientist in human genetics at the University of Essen, Medical Faculty, Germany. He graduated in 1960 from the University of Freiburg with an M.D. degree. He received training in different fields of medicine in Hamburg, Germany, and Worcester, Massachusetts/USA between 1961 and 1963. During a residency in pediatrics at the University of Cincinnati, Children's Medical Center, he worked in human genetics as a student of Josef Warkany (1963-66), followed by a research fellowship in human genetics at the Cornell Medical Center New York with James German (1966-68). Thereafter he established cytogenetics and clinical genetics at the Department of Human Genetics, University of Hamburg (1968 – 1976). In 1976 he became founding chairman of the Department of Human Genetics, University of Essen, from which he will retire in 2001. The author's special research interests are the genetics and the clinical delineation of hereditary disorders, including chromosomal and molecular studies, documented in more than 200 peer-reviewed research articles. He is a former president of the German Society of Human Genetics, secretary-general of the European Society of Human Genetics, and a member of various scientific societies in Europe and the USA. He is a corresponding member of the American College of Medical Genetics. The practice of medical genetics and teaching of human genetics are areas of the author's particular interests.

Table of Contents (Overview)

Table of Contents in Detail

Introduction

Each of the approximately 10^{14} cells of an adult human contains a program with life-sustaining information in its nucleus. This allows an individual to interact with the environment not only through the sensory organs by being able to see, to hear, to taste, to feel heat, cold, and pain, and to communicate, but also to remember and to integrate the input into cognate behavior. It allows the conversion of atmospheric oxygen and ingested food into energy production and regulates the synthesis and transport of biologically important molecules. The immune defense against unwarranted invaders (e.g., viruses, bacteria, fungi) is part of the program. The shape and mobility of bones, muscles, and skin could not be maintained without it. The fate of each cell is determined by the control of cell division and differentiation into different types of cells and tissues, including cell-to-cell interactions and intracellular and extracellular signal transduction. Finally, such different areas as reproduction or the detoxification and excretion of molecules that are not needed depend on this program as well as many other functions of life.

This cellular program is genetically determined. It is transferred from one cell to both daughter cells at each cell division and from one generation to the next through specialized cells, the germ cells (oocytes and spermatozoa). The integrity of the genetic program must be maintained without compromise, yet it should also be adaptable to long-term changes in the environment. This is an enormous task. It is no wonder, therefore, that errors in maintaining and propagating the genetic program occur frequently in all living systems despite the existence of complex systems for damage recognition and repair.

All these biological processes are the result of biochemical reactions performed by biomolecules called proteins. Proteins are involved in the production of almost all molecules (including other proteins) in living cells. Proteins are made up of dozens to several hundreds of amino acids linearly connected to each other as a polypeptide, subsequently to be arranged in a specific three-dimensional structure, often in combination with other polypeptides. Only this latter feature allows biological function. Genetic information is the cell's blueprint to make the proteins that a specific cell typically makes. Most cells do not produce all possible proteins, but a selection depending on the type of cell.

Each of the 20 amino acids used by living organisms has a code of three specific chemical structures, the nucleotide bases, that are part of a large molecule, DNA (deoxyribonucleic acid). DNA is a read-only memory of the genetic information system. In contrast to the binary system of strings of ones and zeros used in computers ("bits", which are then combined into "bytes" that are eight binary digits long), the genetic code in the living world uses a quaternary system of four nucleotide bases with chemical names having the initial letters A, C, G, and T (see Part I, Fundamentals). With a quaternary code used in living cells the bytes (called "quytes" by The Economist in a Survey of the Human Genome, July 1, 2000) are shorter: three only, each called a triplet codon. Each linear sequence of amino acids in a protein is encoded by a corresponding sequence of codons in DNA (genetic code). The genetic code is universal and is used by all living cells, including plants and also by viruses. Each unit of genetic information is called a gene. This is the equivalent of a single sentence in a text. In fact, genetic information is highly analogous to a text and is amenable to being stored in computers.

Depending on the organizational complexity of the organism, the number of genes may be small as in viruses and bacteria (10 genes in a small bacteriophage or 4289 genes in *Escherichia coli*), medium (6241 genes in yeast; 13 601 in *Drosophila*, 18 424 in a nematode), or large (about 80 000 in humans and other mammals). Since many proteins are involved in related functions of the same pathway, they and their corresponding genes can be grouped into families of related function. It is estimated that the human genes form about 1000 gene families. Each gene family arose by evolution from one ancestral gene or from a few. The entirety of genes and DNA in each cell of an organism is called the genome. By analogy, the entirety of proteins of an organism is called the proteome. The corresponding fields of study are termed genomics and proteomics, respectively.

Genes are located in chromosomes. These are individual, paired bodies consisting of DNA and special proteins in the cell nucleus. One chromosome of each homologous pair is derived from the mother and the other from the father. Man has 23 pairs. While the number and size of chromosomes in different organisms vary, the total amount of DNA and the total number of

genes are the same for a particular class of organism. Genes are arranged linearly along each chromosome. Each gene has a defined position (gene locus) and an individual structure and function. As a rule, genes in higher organisms are structured into contiguous sections of coding and noncoding sequences called exons (coding) and introns (noncoding), respectively. Genes in multicellular organisms vary with respect to overall size (a few thousand to over a million base pairs), number and size of exons, and regulatory DNA sequences that determine their state of activity, called the expression (most genes in differentiated, specialized cells are permanently turned off). It is remarkable that more than 90% of the total of 3 billion (3×10^9) base pairs of DNA in higher organisms do not carry any coding information (see Part II, Genomics).

The linear text of information contained in the coding sequences of DNA in a gene cannot be read directly. Rather, its total sequence is first transcribed into a structurally related molecule with a corresponding sequence of codons. This molecule is called RNA (ribonucleic acid) because it contains ribose instead of the deoxyribose of DNA. From this molecule the introns (from the noncoding parts) are then removed by special enzymes, and the exons (the coding parts) are spliced together into the final template, called messenger RNA (mRNA). From this molecule the corresponding encoded sequence of amino acids (polypeptide) is read off in a complex cellular machinery (ribosomes) in a process called translation.

Genes with the same, a similar, or a related function in different organisms are the same, similar, or related in certain ways. This is expressed as the degree of structural or functional similarity. The reason for this is evolution. All living organisms are related to each other because their genes are related. In the living world, specialized functions have evolved but once, encoded by the corresponding genes. Therefore, the structures of genes required for fundamental functions are preserved across a wide variety of organisms, for example functions in cell cycle control or in embryonic development and differentiation. Such genes are similar or identical even in organisms quite distantly related, such as yeast, insects, worms, vertebrates, mammals, and even plants. Such genes of fundamental importance do not tolerate changes (mutations), because this would compromise function. As a result, deleterious mutations do not accumulate in any substantial number. Similar or identical genes present in different organisms are referred to as conserved in evolution. All living organisms have elaborate cellular systems that can recognize and eliminate faults in the integrity of DNA and genes (DNA repair). Mechanisms exist to sacrifice a cell by programmed cell death (apoptosis) if the defect cannot be successfully repaired.

Unlike the important structures that time has evolutionarily conserved, DNA sequences of no or of limited direct individual importance differ even among individuals of the same species. These individual differences (genetic polymorphism) constitute the genetic basis for the uniqueness of each individual. At least one in 1000 base pairs of human DNA differs among individuals (single nucleotide polymorphism, SNP). In addition, many other forms of DNA polymorphism exist that reflect a high degree of individual genetic diversity.

Individual genetic differences in the efficiency of metabolic pathways are thought to predispose to diseases that result from the interaction of many genes, often in combination with particular environmental influences. They may also protect one individual from an illness to which another is prone. Such individual genetic differences are targets of individual therapies by specifically designed pharmaceutical substances aimed at high efficacy and low risk of side effects (pharmacogenomics). The Human Genome Project should greatly contribute to the development of an individual approach to diagnostics and therapy (genetic medicine).

Human populations of different geographic origins also are related by evolution (see section on human evolution in Part II). They are often mistakenly referred to as races. Modern mankind originated in Africa about 200000 years ago and had migrated to all parts of the world by about 100000 years ago. Owing to regional adaptation to climatic and other conditions, and favored by geographic isolation, different ethnic groups evolved. They are recognizable by literally superficial features, such as color of the skin, eyes, and hair, that betray the low degree of human genetic variation between different populations. Genetically speaking, *Homo sapiens* is one rather homogeneous species of re-

cent origin. Of the total genetic variation, about 85% is interindividual within a given group, only 15% is among different groups (populations). In contrast, chimpanzees from one group in West Africa are genetically more diverse than all humans ever studied. As a result of evolutionary history, humans are well adapted to live peacefully in relatively small groups with a similar cultural and linguistic background. Unfortunately, humans are not yet adapted to global conditions. They tend to react with hostility to groups with a different cultural background in spite of negligible genetic differences. Genetics does not provide any scientific basis for claims that favor discrimination, but it does provide direct evidence for the evolution of life on earth. Genetics is the science concerned with the structure and function of all genes in different organisms (analysis of biological variation). New investigative methods and observations, especially during the last 10 to 20 years, have helped to integrate this field into the mainstream of biology and medicine. Today, it plays a central, unifying role comparable to that of cellular pathology at the beginning of the last century. Genetics is relevant to virtually all medical specialties. Knowledge of basic genetic principles and their application in diagnosis are becoming an essential part of medical education today.

Classical Genetics Between 1900 and 1953

(see chronological table on p. 13)

In 1906, the English biologist William Bateson (1861 – 1926) proposed the term *genetics* for the new biological field devoted to investigating the rules governing heredity and variation. Bateson referred to heredity and variation when comparing the similarities and differences, respectively, of genealogically related organisms, two aspects of the same phenomenon. Bateson clearly recognized the significance of the Mendelian rules, which had been rediscovered in 1900 by Correns, Tschermak, and DeVries.

The Mendelian rules are named for the Augustinian monk Gregor Mendel (1822 – 1884), who conducted crossbreeding experiments on garden peas in his monastery garden in Brünn (Brno, Czech Republic) well over a century ago. In 1866, Mendel wrote that heredity is based on individual factors that are independent of each other (see Brink and Styles, 1965; Mayr, 1982). Transmission of these factors to the next plant generation, i.e., the distribution of different traits among the offspring, occurred in predictable proportions. Each factor was responsible for a certain trait. The term *gene* for such a heritable factor was introduced in 1909 by the Danish biologist Wilhelm Johannsen (1857 – 1927).

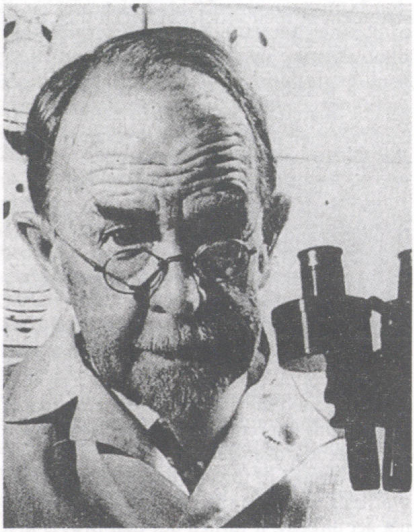
Johann Gregor Mendel

Starting in 1902, Mendelian inheritance was systematically analyzed in animals, plants, and also in man. Some human diseases were recognized as having a hereditary cause. A form of brachydactyly (type A1, McKusick 112500) observed in a large Pennsylvania sibship by W. C. Farabee (PhD thesis, Harvard University, 1902) was the first condition in man to be described as being transmitted by autosomal dominant inheritance (Haws and McKusick, 1963).

In 1909, Archibald Garrod (1857 – 1936), later Regius Professor of Medicine at Oxford University, demonstrated that four congenital metabolic diseases (albinism, alkaptonuria, cystinuria, and pentosuria) are transmitted by autosomal recessive inheritance (Garrod, 1909). Garrod was the first to recognize that there are biochemical differences among individuals that do not lead to illness but that have a genetic

basis. However, the relationship of genetic and biochemical findings revealed by this concept was ahead of its time: the far-reaching significance for the genetic individuality of man was not recognized (Bearn, 1993). Certainly part of the reason was that the nature of genes and how they function was completely unclear. Early genetics was not based on chemistry or on cytology (Dunn, 1965; Sturtevant, 1965). Chromosomes in mitosis (Flemming, 1879) and meiosis (Strasburger, 1888) were observed; the term chromosome was coined by Waldeyer in 1888, but a functional relationship between genes and chromosomes was not considered. An exception was the prescient work of Theodor Boveri (1862 – 1915) about the genetic individuality of chromosomes (in 1902).

Genetics became an independent scientific field in 1910 with the study of the fruit fly *(Drosophila melanogaster)* by Thomas H. Morgan at Columbia University in New York. Subsequent systematic genetic studies on *Drosophila* over many years (Dunn, 1965; Sturtevant, 1965; Whitehouse, 1973) showed that genes are arranged linearly on chromosomes. This led to the chromosome theory of inheritance (Morgan, 1915).



Thomas H. Morgan

The English mathematician Hardy and the German physician Weinberg recognized that Mendelian inheritance accounts for certain regularities in the genetic structure of populations (1908). Their work contributed to the successful introduction of genetic concepts into plant and animal breeding. Although genetics was well established as a biological field by the end of the third decade of last century, knowledge of the physical and chemical nature of genes was sorely lacking. Structure and function remained unknown.

That genes can change and become altered was recognized by DeVries in 1901. He introduced the term mutation. In 1927, H. J. Muller determined the spontaneous mutation rate in *Drosophila* and demonstrated that mutations can be induced by roentgen rays. C. Auerbach and J. M. Robson (1941) and, independently, F. Oehlkers (1943) observed that certain chemical substances also could induce mutations. However, it remained unclear what a mutation actually was, since the physical basis for the transfer of genetic information was not known.

The complete lack of knowledge of the structure and function of genes contributed to misconceptions in the 1920s and 30s about the possibility of eliminating "bad genes" from human populations (eugenics). However, modern genetics has shown that the ill conceived eugenic approach to eliminating human genetic disease is also ineffective.

Thus, incomplete genetic knowledge was applied to human individuals at a time when nothing was known about the structure of genes. Indeed, up to 1949 no essential genetic findings had been gained from studies in man. Quite the opposite holds true today.

Today, it is evident that genetically determined diseases generally cannot be eradicated. No one is free from a genetic burden. Every individual carries about five or six severe genetic defects that are inapparent, but that may show up in offspring.

With the demonstration in the fungus *Neurospora* that one gene is responsible for the formation of one enzyme ("one gene, one enzyme", Beadle and Tatum, 1941), the close relationship of genetics and biochemistry became apparent, quite in agreement with Garrod's concept of inborn errors of metabolism. Systematic studies in microorganisms led to other important advances in the 1940s: genetic recombination was