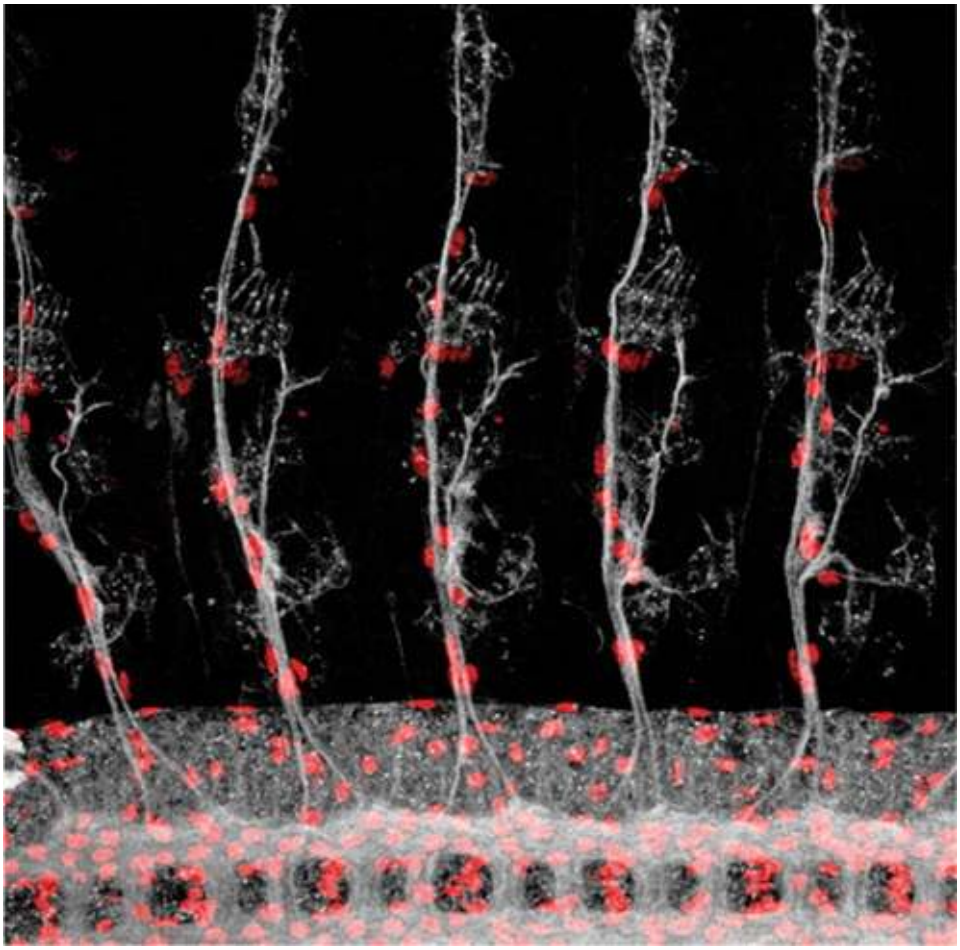
Life’s Blueprint



Trees of embryonic neurons

Life's Blueprint

The Science and Art of Embryo Creation

Benny Shilo

Yale UNIVERSITY PRESS NEW HAVEN AND LONDON

Published with assistance from the Alfred P. Sloan Foundation Public Understanding of Science and Technology Program and from the foundation established in memory of Amasa Stone Mather of the Class of 1907, Yale College.

Printed in China.

Library of Congress Cataloging-in-Publication Data
Shilo, Benny, 1951– author.
Life's blueprint : the science and art of embryo creation / Benny Shilo.
pp ; cm.
Includes bibliographical references and index.
ISBN 978-0-300-19663-4 (clothbound : alk. paper)
I. Alfred P. Sloan Foundation. Public Understanding of Science and Technology Program, sponsoring body. II. Title.
[DNLM: 1. Embryonic Development. 2. Embryonic Structures—cytology. QS 604]
QM601
612.6′4—dc23 2014014622

A catalogue record for this book is available from the British Library.

This paper meets the requirements of aNSI/NISo Z39.48–1992 (Permanence of Paper).

10 9 8 7 6 5 4 3 2 1

Frontispiece: Trees of embryonic neurons. The sensory neurons of the peripheral nervous system of the fruit fly embryo connect to the central nervous system, seen as a ladder at the bottom. They transmit information from external sense organs and internal muscle stretch receptors. The structure of the neurons is repeated accurately at every segment of the embryo. Some of the neural cell nuclei are marked in red, and nerve axons in gray (M. Silies and C. Klämbt, Münster University)

To my parents,

Prof. Moshe Shilo and Dr. Mira Shilo,

who embedded in me the passion for science

and introduced me to the wonders of biology

Contents

Preface

Acknowledgments

1. Introduction: Exploring the Principles of Embryo Creation
2. It’s All in the Genes
3. How an Embryonic Cell “Decides” Its Future Destiny
4. How Cells Talk and Listen to Each Other
5. How Do Simple Modules Lead to Complex Patterns?
6. How Can a Single Substance Generate Multiple Responses?
7. How Do Patterns Evolve Rapidly?
8. How Are Cells Programmed to Follow Predictable Routes to Specification?
9. Ensuring That Embryo Development Is on the Right Track
10. Shaping the Tissues
11. Stem Cells
12. What’s Next?

Glossary of Scientific Terms

Glossary of Scientific Images

Recommended Reading and Web Resources

Index

Preface

During fertilization, the genetic contents of the egg and sperm, each carrying half of the genetic material of the mother and father, respectively, are united in a single cell. This unicellular embryo will give rise to a complete multicellular organism, with its enormous complexity of tissues and multiple types of cells that will appear different from one another and carry out diverse and distinct roles. All the information for this elaborate process is carried within the DNA of that initial single ce that was formed on fertilization, and all cells of the embryo will contain exactly the same genetic information (fig. 1).

The long-standing human desire to understand the instructions and rules of forming an embryo an to unravel "life's blueprint" stems from many sources. First and foremost, we want to understand hov we as humans, and all other organisms, build our complex bodies. How do genes shape animals? Wha are the instructions? How are they carried out in such a reliable and reproducible fashion? Can we us this knowledge to better understand pathological situations in which the genetic program goes awry? And, thinking to the future, can we use the information gained in order to instruct cells to generate tissues and organs that could be used as "spare parts"?

As we set out to explore these fundamental questions, more specific and operational issues come t mind. Are cells preprogrammed by the genome to generate distinct cell types, or do cells have a choice of assuming different fates? How do they decide which fate to choose from the arsenal of tissues they are capable of forming? Does every species have its own rules and logic of embryonic development, or, alternatively, can we identify common themes among disparate species? The scientific discoveries of the past three decades have revolutionized our knowledge and understanding of embryonic development in all multicellular organisms. This book conveys these striking findings and the emerging rules of executing development.

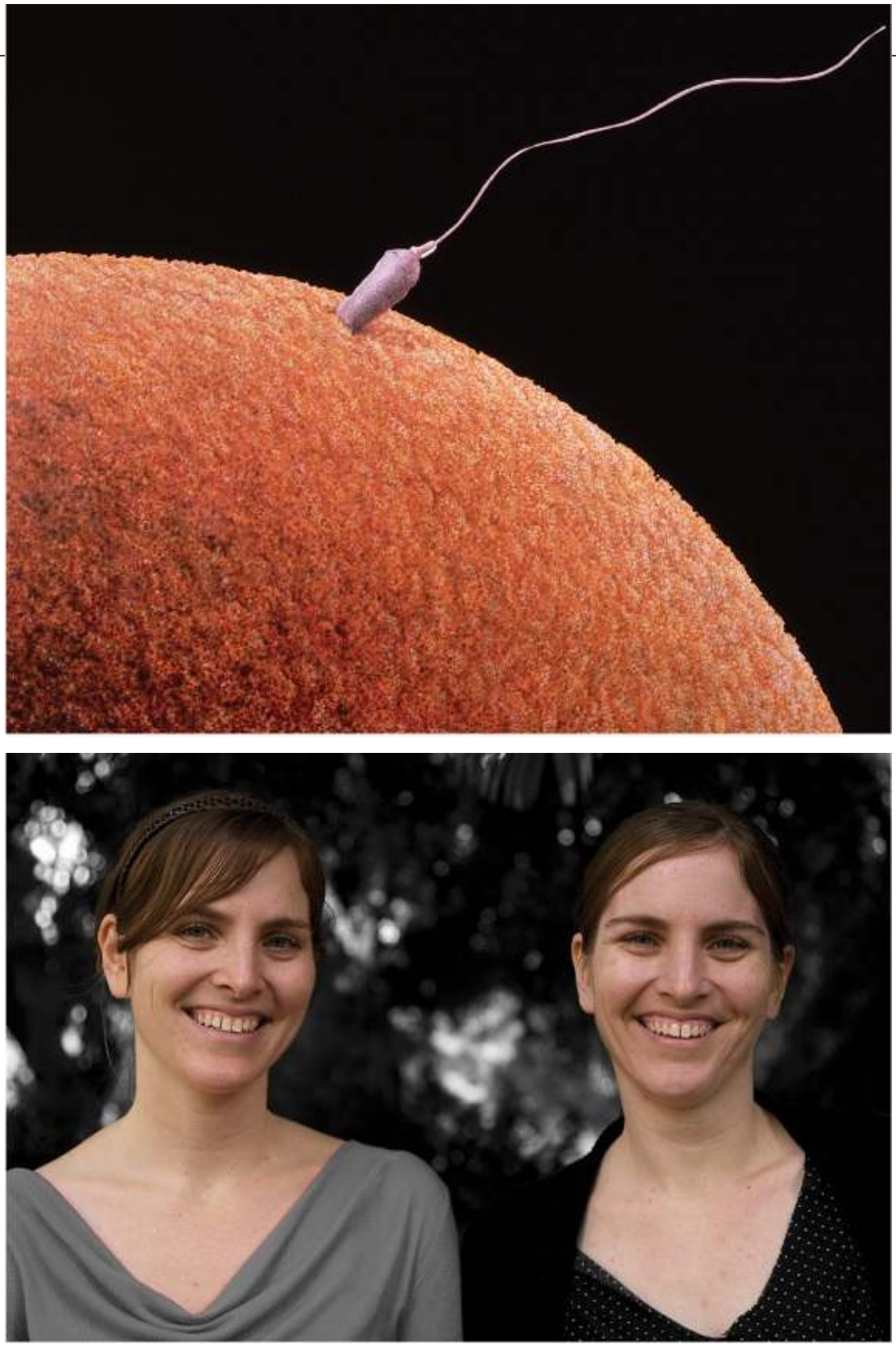


Figure 1. All cells in our body carry identical genetic information
Upon fertilization, the nuclei of the sperm and egg join, and the first cell of the embryo is formed. Throughout the numerous divisions that will follow, all descendants will contain the same genetic legacy. This information dictates the shape of the embryo, a indicated by the fact that identical twins carrying the same genetic material are so similar.

Top: Scanning electron micrograph of fertilized human egg (F. Leroy/Science Source); *bottom:* Identical twins (B. Shilo, Rehovot, Israel)

The analysis of embryonic development relies strongly on microscopic images that we obtain and analyze. These images include fluorescent multicolor depictions of different embryonic tissues,

scanning electron microscope images, and movies of live tissues and embryos showing dynamic processes in real time. The images are not only highly informative but often also aesthetically striking. Their visual beauty is often central in attracting many researchers to the field.

For me, these visual scientific stimulations are combined with another personal passion. Since my high school years, I have been interested in photography and how we view and depict our surroundings. In my early work, I felt more comfortable photographing still objects and how they interact with light. In the past fifteen years, as a result of frequent trips to foreign destinations, including the countries of the Far East and India, and possibly also as a consequence of a more matur view of the world, I have concentrated on photographing people. I try to portray universal aspects tha unite people of all races and ages, as well as the unique features of individuals in their surroundings.

To visually present the emerging concepts of embryonic development in this book, I have assembled scientific images of diverse organisms that display the central concepts of embryogenesis from the laboratories of colleagues and from my own lab. Each of these images is paired with an image from our "macro" world taken by me, highlighting a similar theme and reflecting my personal metaphors. In seeing these juxtaposed images, I hope that readers will be able to understand each paradigm in an intuitive way, letting our everyday experiences and recollections resonate with and inform our comprehension of the scientific image.

It is interesting to note that the inception of the word *cell* as a biological term was formulated by Robert Hooke in 1665 and was based on a visual analogy to the macro world. Hooke looked at the ba of a cork oak using a microscope he developed. He saw the regular shapes of the nonliving cell walls and noted that they resemble the cells in which monks live. I envisage that through the pictorial analogies, the reader will "feel" like a cell within the developing embryo (fig. 2).

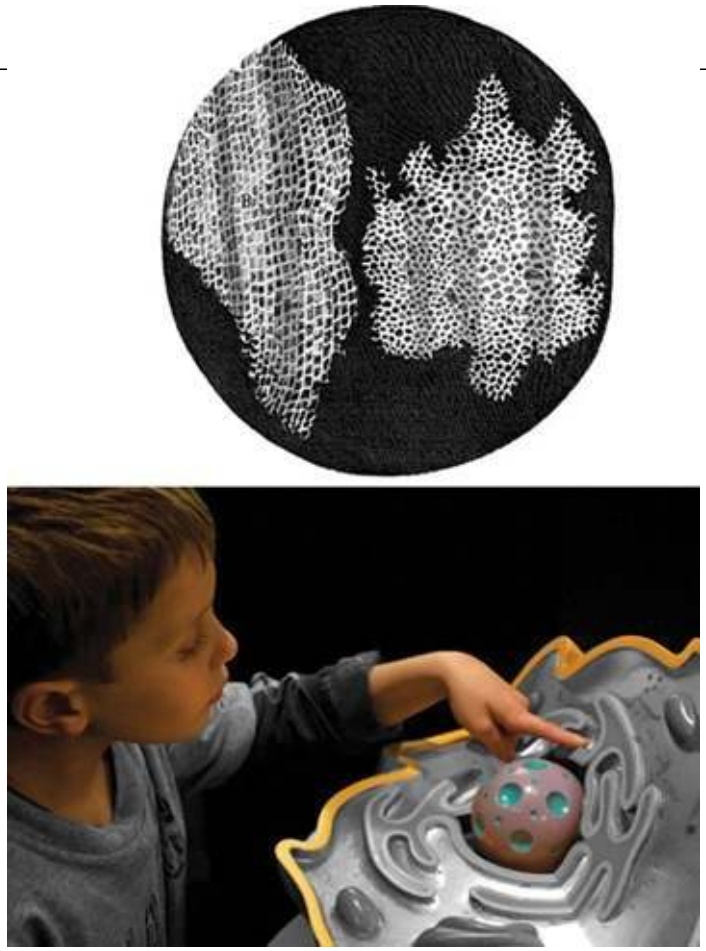


Figure 2. Analogies between cells and the human world
From the inception of the word by Robert Hooke in 1665, the cell has been defined by analogy to the human world. Hooke coined the word *cell* because he felt that the structures he saw in his microscope resembled monks' cells. In this book I will be making analogies between the environment of cells and the human world. I would like the viewer to "feel" like a cell as it participates in forming the complex structure of the embryo by relating this process to visual and conceptual similarities in the human world.

Top: Drawing of cork oak bark by Robert Hooke in *Micrographia; bottom:* Child and cell (B. Shilo, Exploratorium, San Francisco)

The goal of this book—to present a coherent and simplified description of the complicated proces of embryonic development—has forced me to define to myself the true essence of each aspect of life blueprint. The juxta-position of metaphors from the human world with these biological processes has also required a continuous examination of each concept in order to crystallize its essence and employ the most appropriate analogy. Finally, my efforts to present a global view of embryonic development have mandated a critical evaluation of the relative importance of each concept from a broader perspective. As a scientist, my work involves delving deeper and deeper into defined biological problems. Stepping back and viewing the entire scope of embryogenesis in a global manner has been both refreshing and illuminating.

Acknowledgments

The exchanges with so many who contributed their unique expertise and points of view have made th creation of this book a pure joy. During my career I have been inspired by my father, Moshe Shilo, who was a pioneer in microbial ecology; my PhD mentor, Giora Simchen; my postdoctoral adviser, Robert A. Weinberg; and the numerous colleagues with whom I have interacted throughout the years. Especially vital have been the exchanges with the students and colleagues in my lab, who also contributed many of the scientific images displayed in this book. I would also like to thank my international scientific colleagues for generously providing scientific images.

I am truly thankful to the Radcliffe Institute for Advanced Study for hosting me as a fellow during the academic year of 2011/2012 and providing the supportive and multidisciplinary environment that fostered the emergence of this book. The late Lindy Hess at the Radcliffe Institute guided me with he expert and personal touch through the initial and critical stages of daring to write my first book; I am deeply saddened that she will not see the finished product. Dr. Judith Vichniac, director of Radcliffe fellowship programs, provided friendship and support.

I am indebted to Michal Rovner, who inspired me to use black and white with a special combination of color in the images to highlight points of interest; my son, Avner Shilo, who contributed continuous advice in image manipulation; my daughter, Anat Shilo, for printing the images; my grandchil dren, Tom and Noa Dubnov, for serving as models in photographs; and my wif Varda Shilo, who provided the initial trigger for the project, as well as continuous inspiration and discussions as it progressed. Nancy Lynch offered thoughtful comments, and Chani Sacharen edited the text insightfully. At Yale University Press, the support of Jean Thomson Black, who supervised th book from the initial proposal stages to the final version, was crucial. I also thank Laura Dooley for the final editing and Samantha Ostrowski for dedicated administrative assistance. Finally, I am indebted to the Alfred P. Sloan Foundation Public Understanding of Science and Technology Progran for support in publishing this book.

CHAPTER 1

Introduction: Exploring the Principles of Embryo Creation

I would like to begin by sharing my personal story and how I embarked on a route that charted my academic career. Along this journey, I was extremely fortunate to be exposed to, and contribute to, th burst of findings that provide us today with a deep and comprehensive understanding of embryonic development. The year was 1981, and I was a twenty-nine-year-old postdoc in the lab of Robert A. Weinberg at the MIT Center for Cancer Research in Cambridge, Massachusetts. During that time, the pioneering work of the lab devised ways to identify and isolate single cancer-causing genes, called oncogenes, from the DNA of human tumors of various origins. I was privileged to be part of the grou that created this breakthrough.

Our discovery was preceded by the work of many labs researching RNA viruses. These viruses, which are called retroviruses, hijack normal genes in the cells they infect and convert them to oncogenes by altering their structure or regulation. Although we recognized the cellular origin of numerous oncogenes, we still did not understand the role they play within healthy cells. The function of these cellular ("normal") oncogenes, before they are hijacked and modified by the retroviruses, must represent central junctions of regulation for the cells. This is why subtle changes in their activit can lead to such deleterious outcomes as cancer. But what are these normal functions?

The answer requires a global view of the entire repertoire of living organisms through the lens of evolution. Instead of viewing all organisms in a "snapshot" as they appear today, we need to consider how this incredibly diverse array of living forms came to be. The appearance of life on earth was a rare process that took place once or at most very few times. A key feature of any living entity is an ability to replicate itself. Through natural selection the most suitable progeny was chosen for each environment. Over the billions of years since life first appeared, life forms have diversified and expanded, giving rise to the vast array of bacteria, plant, and animal species living today.

The process of evolution is dynamic. Hence, many organisms that formerly existed are no longer with us, and we can only glimpse their structure from fossil records. On the other hand, in general, th commonalities shared by different organisms today reflect their shared origin. It is likely that the features common among organisms are generally inherited from a joint ancestor rather than invented multiple times independently. One of the best examples of this shared inheritance is the genetic material, DNA. Short for "deoxyribonucleic acid," DNA is a self-replicating material that is present i the cells of all living organisms. Perhaps there could be other chemical solutions to encode genetic information. But all organisms that have descended from the primordial cell in which the structure o DNA first arose are "stuck" with this design. We can thus draw two general intuitive rules. First, whe a gene is found in a broad range of organisms, it is likely to play a similar role in those different organisms. Second, the wider the range of organisms in which a given gene is found, the more genera or basic its function is likely to be.

That said, in the late 1970s it was not clear whether the genes and biological processes that can be analyzed in simpler organisms, such as the fruit fly or the yeast we use to make bread and wine, bear

any relevance to complex biological mechanisms such as the progression of cancer or the development of the brain. Were these intricate processes introduced late in evolution, together with the emergence of complex organisms? Or were these basic biological processes actually "invented" early on in evolution and exist today in a wide range of multicellular organisms to carry out very bas and common biological processes? If the latter were the case, we realized that we could make great strides in our understanding by studying these processes in the simpler organisms and then extrapolating the knowledge to the mechanisms that may operate in complex organisms like humans.

In the late 1970s and early 1980s, scientists had developed techniques to use the genes hijacked by retroviruses to isolate the corresponding normal genes from the genomes of vertebrate species, including the chicken, mouse, rat, cat, and monkey. But no one had as yet explored whether the presence, and hence possible function, of the cellular genes that can give rise to cancer was specific t these vertebrate organisms or extended more broadly within a wider range of multicellular organisms

My research approach to uncover the normal functions of cancer-causing genes started from a simple notion. If genes similar to oncogenes are also found in the genome of simpler organisms, ther we could decipher their normal function by exploring these simpler organisms and then examining whether the concepts we uncovered also applied to humans. Consider a little boy assembling his first pair of Lego interlocking wheels. Once he grasps the concept, he can see that the wheels of a train or in a watch mechanism, though more complex, employ the same basic principle.

How do you look for oncogenes in lower organisms? Today we know the linear order of the building blocks on the DNA (termed the nucleotide sequence) of the entire genome of hundreds of diverse organisms, ranging from bacteria and worms to human beings. But at that time, although the DNA fragments harboring these genes were accessible, their sequences, let alone their functions, wer not known. My first step involved collecting DNA samples extracted from diverse multicellular organisms. To assemble this collection, I visited labs in Boston and Cambridge where research was under way on such organisms as sea urchins, nematode worms, flies, and even single-cell organisms such as baker's yeast.

I will now present a crash course in molecular biology as I describe how we looked for genes that were similar in sequence to the cancer-causing genes. DNA contains four types of units called nucleotides; these are termed A, C, G, and T. When we speak about the sequence of DNA we refer to the order in which these units, or bases, are linked to one another to form the genetic information of each organism. To test whether "relatives" of a given oncogene were present in these DNAs, we first needed to divide the very long and continuous strand of DNA into fragments. We did this by using an enzyme that cut the DNA each time it recognized a specific stretch of six nucleotides. Although the DNA was prepared from thousands of identical nuclei, the same types of fragments were generated fo every cell. The process was similar to starting a new paragraph in an encyclopedia every time a specific combination of six letters appeared. The same paragraphs would be generated for multiple copies of identical volumes.

We then separated each sample into fragments of discrete sizes by placing them in an agarose gel, which, when combined with an electrical field, separated DNA fragments based on their size and charge. The gel was then overlaid with and blotted onto a special filter paper, and the DNA fragments "stuck" to the filter according to their position on the gel. The filter now contained the entire genetic material of the organisms from which it was extracted. Next came the stage of finding the needle in the haystack—determining whether the DNA of these organisms had sequences related to the mouse or chicken oncogenes.

When you have an electronic version of a book and are looking for a particular sentence you vaguely remember, you ask the computer to find the closest fit to this sentence in the entire book. We

employed this same logic in the molecular procedure we used to look for sequences that are related t the cancer-causing genes. The search was based on the most universal and seminal feature of DNA: i capacity to form a double helix. The DNA bases have the chemical feature of forming pairs: T specifically matches with A, and C pairs with G. The second strand in a DNA double helix contains the pairs that match the first strand (fig. 3).

At the time, the DNA fragments harboring the oncogenes were isolated and grown in large amoun by procedures of molecular cloning based on their inclusion within the viruses that hijacked them. Th process of molecular cloning is similar, if you will, to photocopying just one page from a thousand-volume encyclopedia over and over again. These are the same proportions between the complexity of the entire mouse or human genome versus a single gene.

Figure 3. The complementary structure of the DNA double helix
The stairway and its matching shadow depict the concept underlying the structure of DNA. The bases in one strand of the helix are complementary to the bases of the adjoining strand. This feature enables not only the pairing that generates the double helix but als the ability to replicate the DNA based on the information contained within each strand and the generation of RNA molecules that copy the DNA sequence. This complementarity of bases was used in order to "fish" from the fly genome the sequences that are complementary to human oncogenes (B. Shilo, Copenhagen)

The process of DNA replication, in which one strand of DNA can be used as a template to generat the complementary strand in the double helix, can be carried out in the test tube by the addition of certain enzymes. If one of the four bases is radioactively labeled, then the entire strand of DNA that i generated will also be labeled. If many copies from a short strand of DNA containing an oncogene ca be marked, they can then be used as a "probe" to explore the vast array of DNA fragments on the filte in a search for DNA stretches that harbor a similar sequence with which they can pair. Even if the oncogene and the DNA on the filter share only a partial identity (that is, they are relatives but not identical twins), this similarity might suffice to allow them to associate with each other.

The filter was then immersed in a radioactively labeled sample of the gene to be tested. After the filters were cleansed of the excess radioactive oncogene probes, the only remaining probe would be bound to the filter at the spots where it found complementary sequences. This binding was then detected by exposing the filter in a sealed cassette to a sheet of film similar to the film used when ou teeth or bones are X-rayed. The radioactive probe excited the film and generated a reaction that is similar to exposure to light. The film was then processed in a darkroom by immersion in solutions of developer and a fixer.

Once I had developed and fixed the film in the dim orange light of the darkroom, I got my first glimpse of the encounter with the labeled oncogenes. There they were, multiple black stripes indicating the presence of similar sequences! Each one reflected the existence of a distinct DNA fragment that paired with the oncogene probe. In the darkroom I could not make out the precise position of the spots on the filter, and hence determine which of the organisms possessed these sequences. But when I later scrutinized the film over a light box, I found that all the spots corresponded to the DNA extracted from the fruit fly *Drosophila melanogaster*.

The implications of this finding were instantly apparent. If the same genes that cause cancer in higher-level vertebrate organisms are also present in fly DNA, this means that they first appeared more than six hundred million years ago, before the evolutionary divergence of vertebrates and invertebrates. Their ancient origin suggests that these genes perform an essential common function i all multicellular organisms. From an experimental standpoint, our ability to manipulate fruit flies, which have been genetically dissected for more than a hundred years, could tell us what functions these same genes carry out in our own human bodies.

For any scientist, there are a few moments in your life when you know you've uncovered an important phenomenon, never before seen or recognized. Even if you don't completely understand th full significance and implications of your discovery, the findings make an immediate impact. These "moments of discovery" are what we scientists live for. I like to think about such moments as analogous in their nature, though not necessarily in scope, to the discovery of the tomb of Tutankhamen on November 26, 1922. The steps leading to the tomb had been cleared, the dignitaries and sponsors of the excavations were assembled outside, and the British archaeologist Howard Carte opened a small hole in the mud wall containing the original royal stamps. Releasing a gust of air that had been sealed up for more than three millennia, he inserted a candle through the hole. Gradually, as his eyes grew accustomed to the dim light, Carter was able to pick out "strange animals, statues and gold, everywhere the glint of gold." His eye was the first to see what would later become a cornerstor in our perception of human civilization.

In the summer of 1981, after having discovered the normal forms of vertebrate cancer-causing genes in fruit flies, I established my lab at the Weizmann Institute in Israel. I spent the next thirty years investigating the role of these genes during the normal life cycle of the fly. This approach can b regarded as "reverse genetics" in the sense that one starts with a known gene and tries to establish its role by generating mutations in this gene and investigating the consequences. In such an approach, yo do not know beforehand what biological functions will be unraveled; you simply "trust your gene," analyze it with the toolkit of technologies at your disposal, and hope to uncover the role it plays in central biological processes. Following this approach, I started my journey with genes that cause cancer in vertebrates and ended up studying the central mechanisms by which cells communicate in forming the fruit fly embryo.

A complementary approach, at that time more common, is to carry out "forward genetics." In othe words, one defines the biological process of interest and seeks mutations in genes that will disrupt th process. In this approach, the biological process being interrogated is well defined, but the nature of

the genes that will be uncovered is not known. The Nobel Prize laureates Christiane Nüsslein-Volhar and Eric Wieschaus used forward genetics to uncover the genes that underlie embryonic patterning in the fruit fly embryo in the late 1970s. Their work revealed most of the genes needed to pattern the fruit fly embryo and triggered a burst of research activity to isolate these genes and identify their molecular nature. When the genes were eventually analyzed at the molecular level, many turned out t correspond to the same genes studied in my reverse genetics approach that had its origin in vertebrat cancer-causing genes. By the late 1980s, the two approaches, which had begun at opposite starting points and used different methodologies, converged.

In the years since then, scores of labs showed that the genes uncovered by these two approaches encoded elements in the communication signals by which cells "talk" to each other during embryonic development to generate the elaborate pattern of the embryo. The presence of these genes in all multicellular organisms indicated that such communication processes might be universal. In fact, these genes embody the essence of multicellularity. The elaborate process of embryonic developmen is predictable and nonchaotic because the behavior of cells is guided by distinct "rules" that are embedded within the DNA. And when these rules are followed correctly, a pattern emerges.

Unraveling the mechanisms by which cells communicate at the molecular level and the discovery that these modules are common to all multicellular organisms was a breakthrough in our knowledge life's blueprint. It also converted modern biology into a more universal discipline, since we now kno that seemingly disparate processes in different organisms are actually controlled by the same genes and cell communication pathways.

CHAPTER 2

It's All in the Genes

When we consider all the different cells in our body, we can't help but be amazed by how diverse the are. Each cell type is configured to fulfill its ultimate function in the body. Nerve cells form long extensions that can be up to a meter in length and conduct electrical currents at high speed from one end of the cell to the other. Muscle cells form repeated units of molecular motors that move along tracks to carry out muscle contractions. In the muscles, thousands of cells fuse together to generate giant continuous cells that can extend to a meter in length. Red blood cells play a completely differer role, carrying the hemoglobin protein that helps bind oxygen in the lungs and then releasing oxygen within the target tissues. Where does the body store the instructions for generating this diversity, and how are these instructions transformed to actual cellular structures? The answer lies in our DNA.

The nucleus of each cell harbors DNA, which plays two fundamental roles: it carries the information that allows each cell type to produce the components that build up that cell, and it stores all the information for constructing the remarkable diversity of cell types encompassed in the whole organism. This information is faithfully duplicated during cell division. It is also transmitted from or generation to the next after a sperm fertilizes an egg.

Four building blocks, called bases, make up DNA; these are abbreviated A, C, G, and T (for adenine, cytosine, guanine, and thymidine, respectively). The bases join one another to form extremely long strings. The DNA within a single human cell contains three billion bases, which, if spread out, would form a string a meter long. The genetic material in each cell is divided and packed into units called chromosomes, each carrying a different portion of the genetic material. Every huma cell contains forty-six chromosomes.

The DNA is typically found in the shape of a double helix, meaning that each strand of DNA is paired with its parallel strand. As I mentioned in Chapter 1, the pairing of DNA strands is the essenti underlying feature behind cells' ability to replicate their DNA. There are preferred sets of attraction between the DNA bases: T pairs with A on the opposite strand, and G pairs with C. So when a cell replicates its DNA, the two paired strands separate from each other, and each strand contains the information to form the missing complementary strand. In a way, each strand serves as a "mold." Th complicated process does not take place at once but rather proceeds in a progressive manner, in that the two DNA strands are gradually separated and each of the strands is replicated after the separation In other words, on the template of each strand, free bases pair and assemble to form the matching strand.

The complementary nature of the DNA strands is also the secret behind the cell's ability to conver the information embedded within the DNA into actual functional cellular components. These are the cell's proteins, each carrying out a unique function. Proteins are also composed of repeated units, but the nature of their building blocks is different. How is the information embedded within the DNA use to form proteins, and what is the unit within the very long DNA strands that corresponds to each protein being made?

As stated, the chromosomes are packaging units, and each one contains a large chunk of the genet material. But the universal functional unit within the DNA that carries the information required for making a protein is much smaller, and it is called a gene. It is important to note that although the nucleus in each human cell contains three billion DNA bases, only a small fraction of those bases, about 3 percent, constitute genes. It is estimated that even very complex organisms, such as humans mice, have just under twenty-five thousand genes. Picture the DNA as a desert with very sporadic oases; these sites correspond to the protein-coding regions.

So what is all that "extra stuff" in the DNA, and does it have a purpose? Some parts of it carry instructions on where and when to create proteins. Other parts have more global structural roles, such as separating genes or contributing to the overall function of the chromosome. And some of it represents excess baggage accumulated through evolution—foreign genetic elements that were incorporated into the DNA and remained.

But back to the genes. The two-step process of making proteins, called transcription and translation, involves individual genes on the DNA. First, the information encoded by the DNA that defines a gene is copied into RNA (short for "ribonucleic acid") molecules. The RNA consists of uni that are not only very similar to the DNA units but can also pair with them. The process of copying th information of the gene on the DNA into the RNA (transcription) is therefore very similar to that of DNA replication. In transcription, the gene uses the sequence of DNA bases as a template. A single gene can produce thousands of RNA copies.

I have stated that the DNA is stored in a defined cell compartment called the nucleus. The cellular machinery for making the proteins, however, is located outside the nucleus, in a compartment called the cytoplasm. After an RNA molecule, typically several thousand bases long, is produced in the nucleus, it is exported to the cytoplasm, where it encounters the ribosome and undergoes the second step of making proteins, translation.

The ribosome is a giant molecular machine that catalyzes the translation of the information carrie by the RNA sequence into the language of proteins. The building blocks of proteins are amino acids, of which there are twenty types. To construct proteins, the ribosome translates the linear code of the RNA into a linear code of amino acids; these become firmly associated with each other like the cars a train. The linear order of the amino acids in each protein reflects the information originally carried by the gene on the DNA and converted through an RNA intermediate to the protein.

The resulting proteins are the molecular "machines" that carry out the diverse array of cellular functions, but they cannot carry out their task as linear strings of amino acids. Each protein thus undergoes an elaborate folding process to generate a three-dimensional structure that is dictated by it primary sequence. If you hold a strand of wool and fold it up, you will form a different kind of tangle structure each time. In contrast, every time a protein string of amino acids folds, it eventually creates the same three-dimensional structure, because the "rules" for making the structure are embedded within its linear order. Those rules in turn rely on multiple interactions among the protein's amino acids. As a comparison, think of a creating a seating arrangement at a dinner table in which you try to optimize the arrangement and have each diner sit next to the person he or she likes best. Because so many interactions define the final structure of the protein, it is extremely difficult, if not impossible, to predict the structure the protein will assume based on its sequence, even using the most sophisticated computational tools (fig. 4).

Information embedded in the DNA is thus transmitted and used to generate a protein with a define sequence, using RNA as an intermediary. But how does this process explain the molecular basis for the differences between a nerve cell and a muscle cell? Every cell in the human body carries the sam genetic information that was contributed by the DNA of the sperm and egg upon fertilization. But a

given cell does not use all of that information; rather, it employs only selective parts that are suitable for its biological role. As a comparison, think of an encyclopedia, in which different pages are accessed each time depending on which topics are sought (fig. 5).

The selective use of the information stored in the DNA by a particular cell type means that only some of the genes are used as templates to produce the RNA that will serve as a basis for making the corresponding proteins. The other genes remain silent and are not used within the context of that cell type. This is the essence of how one cell type differs from another. There is an elaborate regulatory process by which genes are selected to create RNA. The regulation determines not only which genes will produce RNA but also how many RNA copies each gene will produce and in which order the genes will be used. Just like a cooking recipe, the combination of ingredients, their amount, and the order in which they are added are critical for the ultimate success.

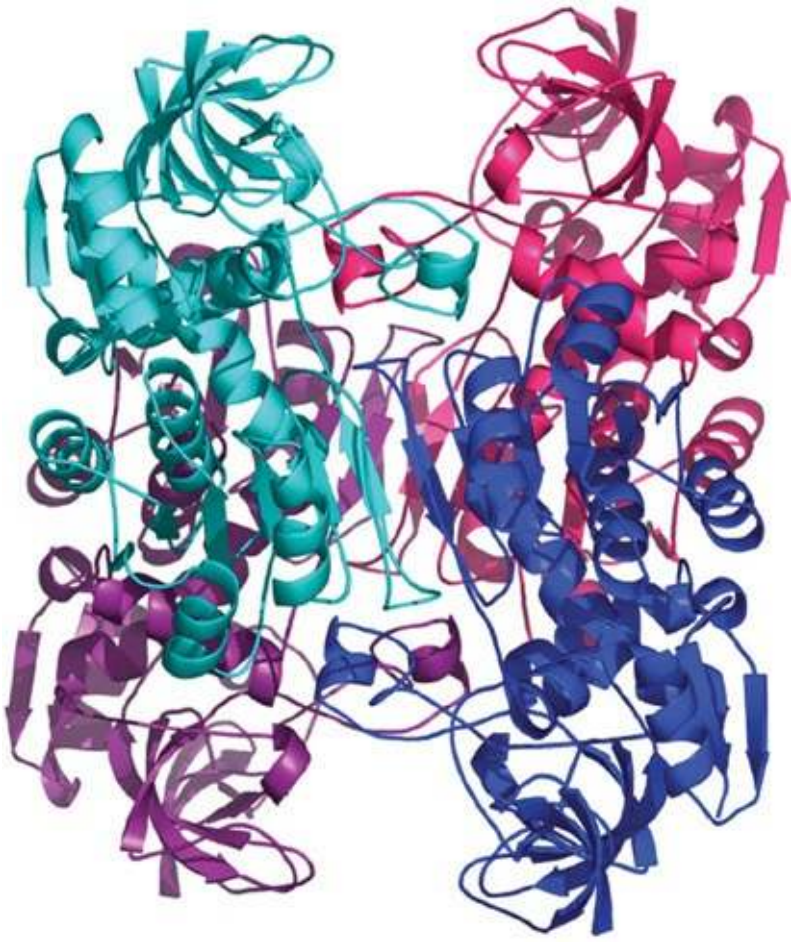


Figure 4. The structure of a protein facilitates its function
The linear information encoded by a gene is transferred from the DNA through an RNA intermediate and is used as a template for production of proteins. Proteins consist of twenty different building blocks termed amino acids. The composition and order of the amino acids (dictated by the sequence of the gene) determine how the protein will reproducibly fold in three-dimensional space. Th final structure of the protein facilitates its function by allowing the protein to pair with proteins or other molecules.

Top: Structure of the bacterial protein alcohol dehydrogenase, which initiates the breakdown of alcohol molecules. Four identical copies of the protein (marked in different colors) associate with one another to generate the higher order structure of the enzyme (Y Burstein, Weizmann Institute); *bottom:* Salt and pepper containers displaying a reproducible and complementary three-dimensional structure (B. Shilo, Mumok, Vienna)

Figure 5. Why cells look different from one another
If all cells carry identical genetic information, why do they look so different? For example, a nerve cell is very distinct from a muscl cell. Each cell type uses a different portion of its genetic material as the template for making RNA and proteins. The distinct protein that each cell type makes reflect and manifest its function.

Top: Scanning electron micrograph of a mouse muscle cell, showing the repeated units called sarcomeres, which carry out muscle contraction (Y. Gruenbaum-Cohen and B. Shilo, Weizmann Institute); *bottom:* Card filing cabinet in which only some of the information is used at any one time (B. Shilo, Schlesinger Library, Radcliffe Institute for Advanced Study, Cambridge, Massachusetts)

How is the elaborate regulation of RNA synthesis controlled? The machinery for producing RNA chains involves a complex of multiple proteins that slide along the DNA and copy the information of the DNA template to RNA. This apparatus is capable of producing RNA from all genes. So why aren' all genes in the genome active all the time? It turns out that the limiting factor in the production of RNA is how the cell marks those specific genes that will be used to produce RNA. By selectively

controlling the association of the copying machinery with some genes and not others, each cell determines the profile of RNA molecules that it will produce.

Now we can define in molecular terms why a nerve cell is so different from a muscle cell. Although they carry identical genetic information on the DNA, each cell type expresses a different se of genes and hence displays a distinct array of proteins that define its structure and function. Only a limited amount of the DNA constitutes the genes and is thus carrying the information for making the proteins. Some of the DNA that is not involved in coding the protein sequence serves to regulate the expression. Gene expression is the process by which the information on the DNA is copied into RNA molecules. The sequences around the gene that are not copied to RNA serve as docking sites for proteins that attach to the DNA. By recruiting a particular set of proteins to a given gene, it becomes possible to attract the universal complex of proteins that constitutes the machinery for making the RNA molecules on the basis of the DNA template. Think of a long queue of people standing in line to buy tickets to a movie. As you reach the line you may recognize someone, and you will be immediately attracted to approach the person you know.

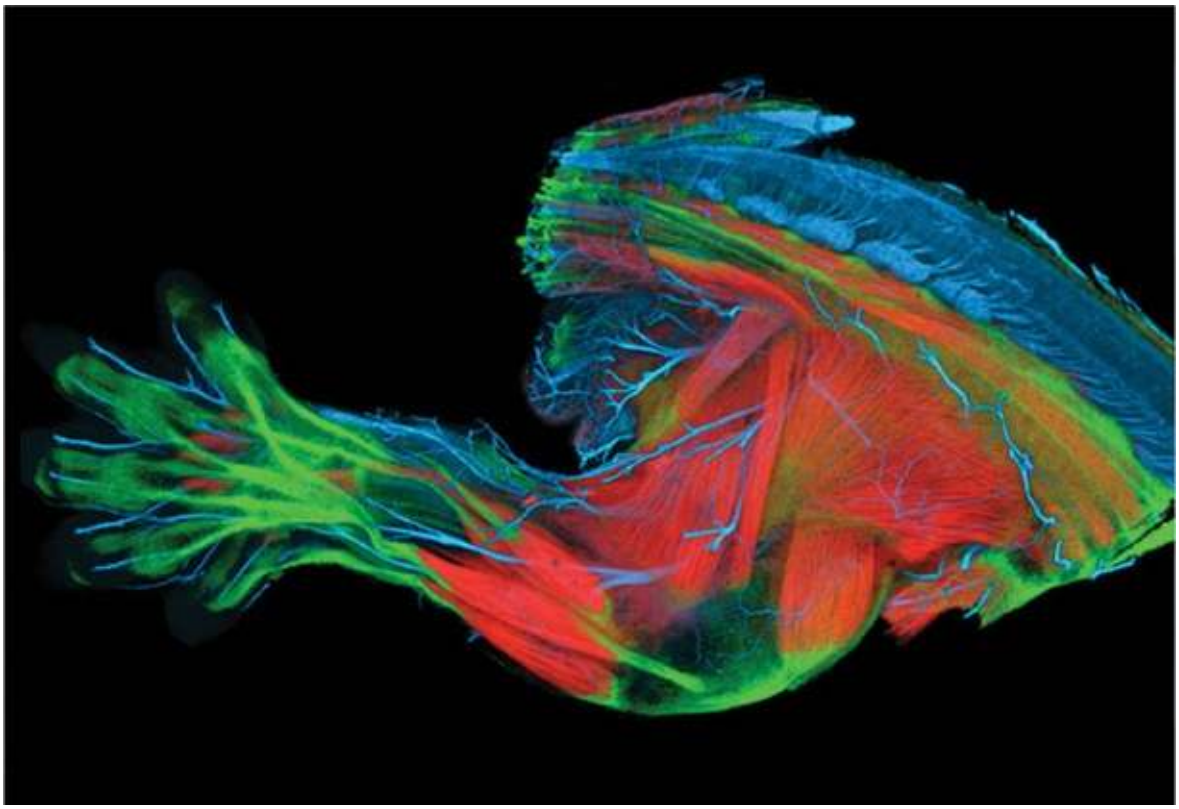


Figure 6. Coordinating the generation of different cell types

Although each cell type adopts a distinct shape and function, this process does not happen in isolation. As cells assume different roles, they must be tightly coordinated with the other cells. This enables the creation of such elaborate structures as the mouse limb, which consists of bones, blood vessels, nerves (blue), muscles (red), and tendons (green) (A. K. Lewis, G. Kardon, Y. Wan, and R. Schweitzer, University of Utah and Shriners Research Center, Portland, Oregon)

How can a population of identical cells that originated from the same fertilized egg generate the intricate structure of the embryo that contains billions of cells and thousands of cell types, each expressing a different set of genes (fig. 6)? In other words, how is the information embedded within the DNA converted into the elaborate structure of the embryo, and how do genes dictate this pattern?

Even at first glance, it is obvious that there is a genetic basis for patterning the embryo. Within a given species, all embryos are similar, at least at the global level. The striking degree of similarity between identical twins, which share identical genetic information, is a powerful demonstration of th genetic basis for patterning. Clearly, most of the pattern that will be generated in the embryo is

embedded within the genes and the DNA sequences that control them. But how does this happen?

Figure 7. Embryonic development is a concert without a conductor
In a concert, the conductor's score shows all the notes that will be played or sung by each musician. Every individual member has t score of all the notes he or she will sing or play. The complete plan is laid out from the beginning. The conductor harbors all the information to coordinate the performance, and the end result is predictable.

It is interesting to compare and contrast this example with embryonic development, where the final result is also reproducible and written down, somehow, in the genome of the cells. But there is no central coordinator who "knows it all." At different stages the cells receive inputs from their neighbors on what to do, but this information relates only to that specific point in time and place, without any clues regarding long-term future actions. Eventually, the local information that is processed individually by each cell gives rise to the elaborate and reproducible shape of the whole embryo, which is infinitely more complex than a Beethoven symphony (B. Shilo, Israel Philharmonic Orchestra and Gary Bertini Israeli Choir)

Broadly speaking, and based on our knowledge of organizations within our society, there are two extreme types of management. According to one model, there is a senior executive who has the plans for the final product. This executive organizes and manages groups of workers who will fulfill their distinct tasks. For example, to build an airplane, one group will construct the wings, another will assemble the engines, and so forth. From the outset, this system must have a detailed plan for the completed product. That plan will include how the workers will obtain the parts to be assembled and the established hierarchy of managers who will coordinate and micromanage the activities of the specific task forces. Continuing with our analogy to human interactions, an alternative method would be to have a group of people who are equivalent to one another. The people would establish a defined outcome, setting some general rules for the dynamics of interactions among group members. Althoug this approach could be an excellent setting for a group of improvisational theater actors, it would not be effective for assembling an airplane.

Is embryonic development more like the assembly of an airplane or the operation of an improvisational theater? An embryo is much more complex than an airplane. But the final plan for th embryo is not laid out from the outset. The parts that will be assembled are not prefabricated, and there is no high-level executive cell who will explicitly instruct each of the other cells which tasks they have to fulfill. In contrast to an orchestra, the elaborate process of embryogenesis progresses

without a designated conductor who oversees the entire succession (fig. 7). The shared and identical information that is contained within the DNA does not specify directly the shape of the final product; rather, it writes the rules that cells must follow to reach the desired end point. These rules are spelled out accurately and clearly enough to generate a reproducible pattern. Our challenge is therefore to decipher these instructions. How does a population of identical cells follow such rules to generate the complex and reproducible structure of an embryo?