



RNA SPLICING—CUTTING AND PASTING GENES

Phillip A. Sharp^{1,2*}

¹Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, United States

²The Koch Institute for Integrative Cancer Research at MIT, Massachusetts Institute of Technology, Cambridge, MA, United States

YOUNG REVIEWERS:



ANASTASIA

AGE: 14



KENZO

AGE: 10



MAXIME

AGE: 10



NOURA

AGE: 15



SAIF

AGE: 15



YOUNIS

AGE: 15

Science in general, and specifically the science of life, offers an intriguing path to walk along. As you deepen your understanding of a particular topic, you become aware of things you did not see at all when you started. Sometimes, this new seeing even drives you to reconsider and redefine very basic concepts that you learned along the way. This is exactly what happened after we discovered a process called RNA splicing. In RNA splicing, pieces of genetic instructions are cut and pasted together to form the final instructions for producing proteins. The discovery of RNA splicing has driven us to rethink what we previously believed about genes, which are the most fundamental units of information in biology. In this article, I will tell you what we discovered about RNA splicing, how it influenced our ideas about genes, and how we now use this knowledge to significantly improve people's lives.

Prof. Phillip Sharp won the Nobel Prize in Physiology or Medicine in 1993, jointly with Prof. Richard Roberts, for their discoveries of split genes.

DNA

Genetic information stored in all cells.

PROTEINS

Tiny biological machines that participate in many crucial functions in the body, such as muscle movement, digestion, and immunity.

TRANSCRIPTION

The process by which DNA is copied into mRNA.

Figure 1

DNA contains the information to make proteins. Living cells use the genetic information in DNA to produce proteins, through two main steps. First, the information in the DNA is copied in a process called transcription, to create a movable mRNA copy of the instructions for a particular protein. The mRNA then travels to the protein-production factory of the cell, where the specific protein is made in a process called translation.

GENE

A piece, or several pieces, of DNA that contain instructions for the production of proteins.

MESSENGER RNA (mRNA)

RNA that is used to produce proteins.

TRANSLATION

The process by which a protein is produced based on instructions from mRNA.

FROM GENES TO PROTEINS—HOW DNA DIRECTS CELL FUNCTIONS

One of the most fundamental processes of life involves reading the genetic information, or “instructions” stored in cells in the form of **DNA**, and turning those instructions into the structures and functions of a living organism (Figure 1). The instructions within the DNA code for the production of **proteins**—tiny biological machines that perform many crucial functions in the body. The creation of proteins from DNA requires a middle step, called **transcription**. In transcription, a set of instructions from the DNA sequence (called a **gene**) is made into a movable copy that can travel to the protein-producing factory in the cell. This copy of the DNA that drives the creation of proteins is called **messenger RNA (mRNA)**. **Translation** is the name of the process by which proteins are produced based on instructions in mRNA.

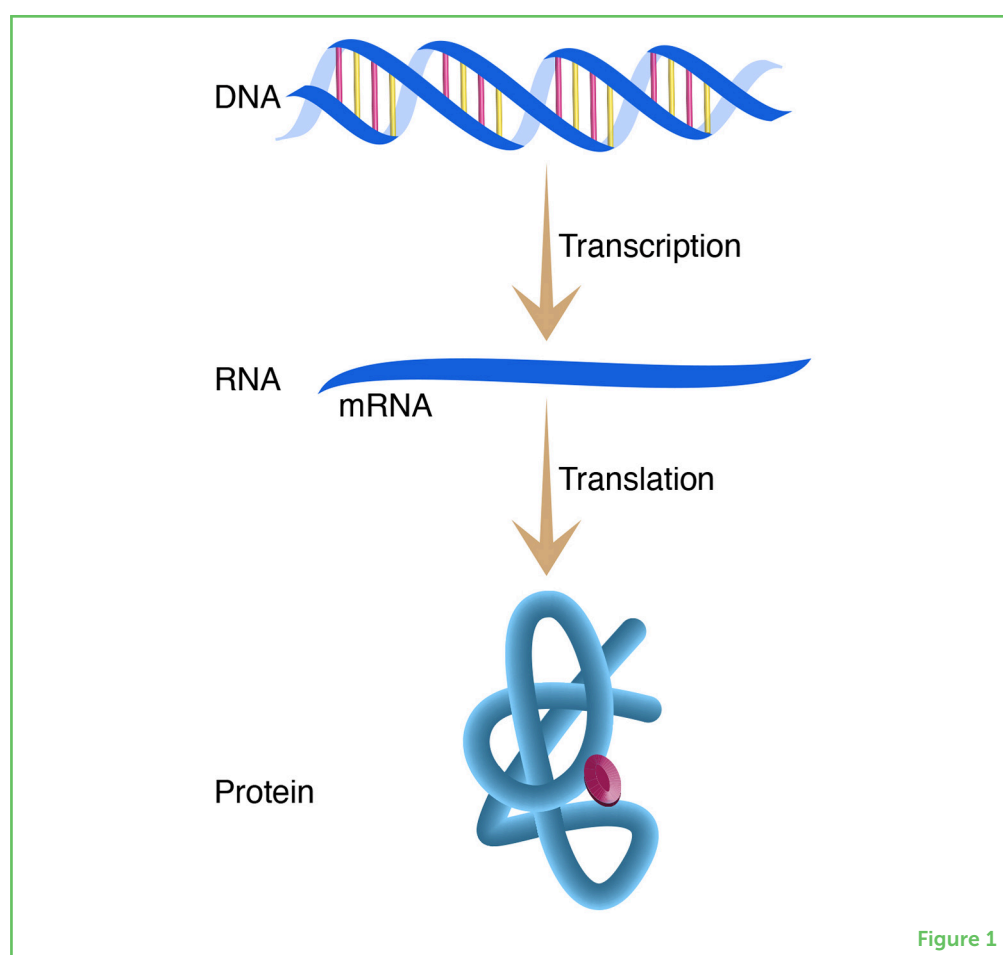


Figure 1

Over the course of evolution, the process of preparing mRNA from DNA became more complex. A completely new intermediate stage emerged between transcription and translation, called *RNA splicing*. In this stage, separate pieces of an mRNA are cut and pasted together to produce the final version of the mRNA. In the next sections, we will explain RNA splicing, how we discovered it, what are its

benefits, and how this knowledge can be used to help people with certain diseases.

FOR RNA SPLICING, NUCLEI MATTER!

The most ancient species that developed on earth had only one cell and no nucleus [1], so its genetic material was not separated by a membrane from the rest of the cell's content. Cells developed nuclei only after about 1–1.5 billion years of evolution. Nucleus-free organisms are called **prokaryotes**, and organisms that have nuclei in their cells are called **eukaryotes**. Bacteria are a well-known example of prokaryotes. In contrast, all mammals, including humans, are eukaryotes.

The process of protein production was first discovered in bacteria. Researchers found that an mRNA copy is made from a continuous piece of DNA (a gene) and is then translated into a specific protein (Figure 2, left) [2]. At that time, most researchers believed that this process was similar in eukaryotic cells. But, some years later, we discovered that an additional step is added before translation and after transcription in eukaryotic cells [3]. In this step, the mRNA produced through transcription is spliced (cut and pasted) to produce the final mRNA that participates in translation. The initial DNA copy is called pre-mRNA, and the final version, after RNA splicing, is called mature mRNA (Figure 2, right).

PROKARYOTE

An organism, such as bacteria, whose cells do not have nuclei.

EUKARYOTE

An organism, such as humans and other mammals, whose cells have nuclei.

Figure 2

Protein production in prokaryotic and eukaryotic cells. **(Left)** In prokaryotic cells like bacteria, which do not have nuclei, mRNA is produced from DNA by transcription and is then directly translated into a specific protein. **(Right)** In eukaryotic cells, a pre-mRNA is made by transcription in the cell's nucleus. Then, the pre-mRNA goes through an additional processing stage called RNA splicing. Only then does the mRNA leave the nucleus and travel to the protein-production factory, where translation occurs (Image adapted from Khan Academy).

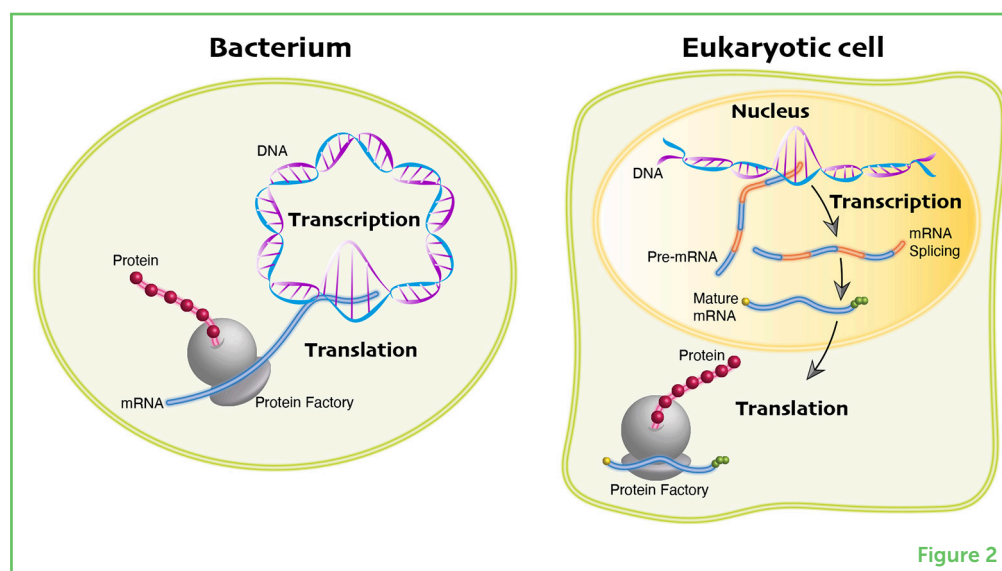


Figure 2

SPLICING—CUTTING AND PASTING RNA

You can think of the RNA splicing process, by which pre-mRNA becomes mature mRNA, like editing a book. Imagine you have a book in which there are stretches of a few words that make sense together,

EXON

A “sense” gene segment that is used in the production of a protein.

INTRON

A “nonsense” gene segment that is removed from pre-mRNA to produce mature mRNA.

ALTERNATIVE SPLICING

Using different combinations of exons to create multiple proteins from the same gene.

Figure 3

Alternative splicing. In alternative splicing, exons are joined together in different combinations, creating different mRNAs, and therefore different proteins with complex functions, from the same gene. This could be the evolutionary advantage that allowed complex organisms like humans to evolve (Image adapted from Wikipedia).

then a lot of gibberish (non-sense) words, and then a few more words that made sense together. To make the book readable, you must edit out the gibberish and join the pieces that make sense. The joined pieces create a sentence that has a function, or makes sense, to the reader. Returning to mRNA, it turns out that pre-mRNA is created from both sense and non-sense DNA sequences. The sense sequences are called **exons**, and the non-sense sequences are called **introns**. When a pre-mRNA molecule is spliced, the sense pieces are brought together and the non-sense pieces are removed and destroyed. This “cutting-and-pasting” process produces the mature mRNA that is used to produce proteins.

But why do eukaryotes have these non-sense introns in the DNA in the first place? Well, scientists are also wondering about that—we do not have a definite answer yet. We know that humans have about 23,000 genes, and almost all of them are made of multiple exons that are spliced together in different combinations in different cells, or at different times in some cells. So this means that multiple proteins (and therefore multiple functions) can be created from the same gene, by selecting which exons are joined together. This is called **alternative splicing** (Figure 3). Some scientists propose that alternative splicing is what allowed the evolution of complex organisms like humans. Alternative splicing allows the DNA to be used efficiently—organisms that use alternative splicing can make complex proteins and perform complex functions using fewer genes.

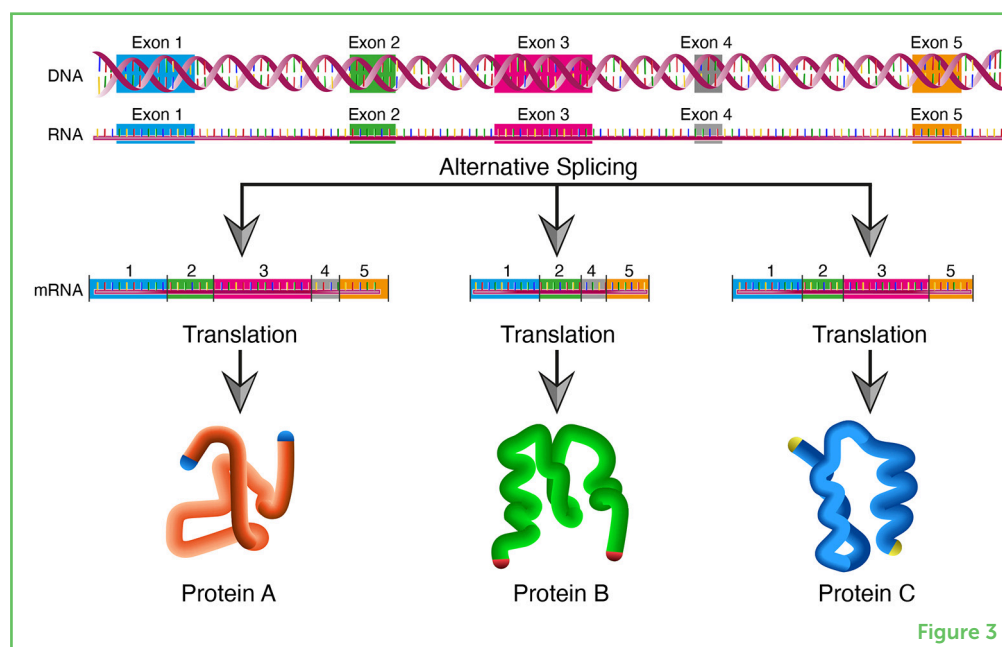


Figure 3

VIRUSES HELPED US DISCOVER RNA SPLICING

To study mRNA splicing in our lab, we used a virus called adenovirus type 2 [3]. This virus has a relatively short DNA, and it uses animal

cells to reproduce its DNA very rapidly. These two facts mean that this adenovirus is relatively easy to use in scientific studies, since it is a simple system that naturally produces many copies of itself.

After choosing adenovirus type 2 as our system of study, we wanted to compare its DNA to its mRNA molecules. If an mRNA exactly matches a specific piece of DNA, that suggests that the mRNA is copied directly from the DNA, without splicing—the way it happens in bacteria. But if the mRNA does *not* match the specific DNA piece, that suggests that splicing is happening to create the mature mRNA.

For our DNA-mRNA comparison, we chose the most common mRNA found in adenovirus two, which codes for a protein called a hexon. When we compared this mRNA with the relevant part of the DNA, we found two areas that did not match between the mRNA and the DNA. One area is called the 3' ("three prime") end of the mRNA strand (Figure 4, green circle), and we already knew that this area is a special "tail" that is added to the mRNA to protect it as it travels through the cell. The other non-matching area was at the other end of the mRNA strand, called the 5' ("five prime") end (Figure 4, blue circle). This surprised us because the 5' end of the mRNA molecule contained a long piece of mRNA that did not match the specific section of DNA, and we did not know where it came from.

Figure 4

Comparing adenovirus DNA and mRNA helped us discover RNA splicing. **(A)** In our study, we used a virus called adenovirus to look for mRNA splicing in eukaryotic cells. When we compared the relevant segment of DNA (thick black line; thin black line is the other DNA strand that separated) with the mRNA (red line) that codes for the hexon protein, we discovered two non-matching areas: the 3' and 5' ends of the mRNA (green and blue circles). The finding that a long piece at the 5' end of the mRNA did not bind to the DNA was surprising, and it suggested that another process was involved in making mRNA. **(B)** A schematic figure showing how a strand of RNA binds to a strand of DNA, when the two strands of DNA are separated in experiments [Image Adapted from Berk [4]].

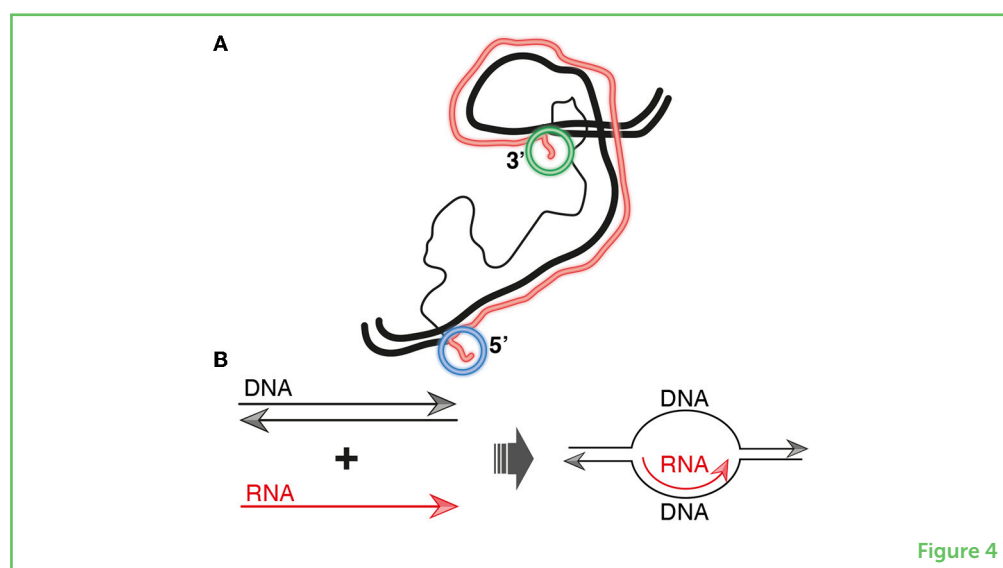


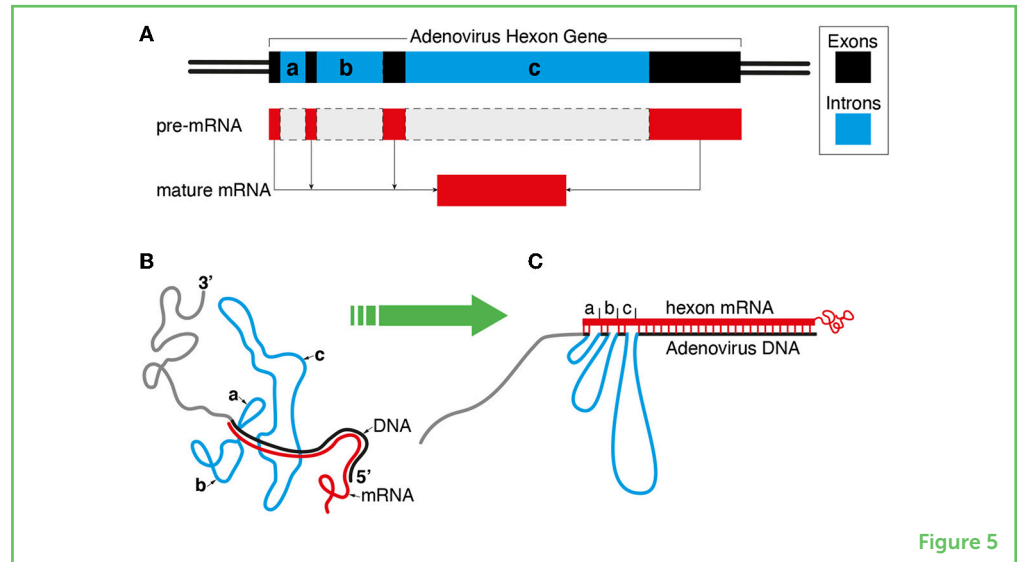
Figure 4

To solve the riddle, we had to find out where the non-matching area at the 5' end of the mRNA came from. After many discussions, we came up with the idea that the non-matched piece could have come from a different place in the DNA. This would mean that there is one continuous piece of DNA that encodes *most* of the hexon mRNA, and a different piece *somewhere else* in the DNA that encodes the 5' end. To check this idea, we compared the hexon mRNA to another section of DNA, located near the section that matched most of the hexon mRNA in our first experiment. What we found was absolutely amazing: the

5' end of the hexon mRNA was attached to the DNA in four *separate segments*, and between those segments there were three DNA loops (Figure 5). This meant that the 5' end was spliced from four separate exons, which were separated by three introns.

Figure 5

Splicing of hexon mRNA. **(A)** Part of the hexon gene (DNA) contains four exons and three introns, from which the 5' end of the hexon mRNA is spliced. The pre-mRNA is a direct copy of the DNA, but the mature mRNA has been spliced to remove introns. **(B)** In our experiment, some parts of the adenovirus DNA were attached to the 5' end of the hexon mRNA (red), while other parts of the DNA were not attached and formed three loops called a, b, and c. The attached parts are exons and the loops are introns. **(C)** The complete spliced hexon mRNA bound to the adenovirus DNA [Image credits: **(A,B)** adapted from Berget et al. [3]; **(C)** Phillip A. Sharp].



This was a truly novel discovery, since no one else had ever suggested that mRNA could be cut and pasted from separate DNA segments. This discovery of RNA splicing changed the way we think about genes!

RETHINKING THE MEANING OF "GENE"

Before we knew about RNA splicing, we thought of a gene as a continuous piece of DNA that encodes a specific protein. But the discovery of RNA splicing taught us that many genes are split genes, meaning that they are combined from different, non-continuous pieces of DNA (exons). Moreover, we know that these pieces can be "mixed and matched" to create not only one, but many different proteins. There are even examples of one specific group of exons coding for tens of thousands of different proteins [5]!

So, if a gene is not one continuous DNA sequence that encodes one protein, what exactly *is* a gene? We still do not know the full answer. As often happens in science, we are faced with incredible complexity, and perhaps we will never have all the answers. But, nevertheless, we always try to expand our knowledge and make the most out of the knowledge that we already have. As you will see in the next section, our knowledge of RNA splicing is being used to help improve the lives of sick people, including children your age.

RNA SPLICING FOR THE GOOD OF HUMANITY

RNA splicing is a fundamental process for all living things, including humans. RNA splicing allows an organism to turn certain genes “on” or “off”, and it therefore helps organisms to carry out many of the complex processes needed for life. As with every important biological process, when something goes wrong, diseases can result. To treat these diseases, we must understand how the damaged biological process works, and then figure out how to fix it.

Some very exciting medical developments are based on RNA splicing. Only a couple of years ago, for the first time in history, we were able to successfully treat children who had a genetic disease called spinal muscular atrophy. This disease is caused by a mutation in the DNA that inactivates the RNA splicing of a particular gene such that an important exon is left out. This mutation causes patients to slowly lose cells in their spinal cord that control their muscles. As the spinal cells and muscles deteriorate, there is progressive loss of muscle function creating severe disabilities and reducing quality of life. Before the discovery of RNA splicing, spinal muscular atrophy was a deadly disease. Now, with our knowledge of RNA splicing, we can manipulate the splicing process so that the missing exon is included in the mature mRNA. This allows a properly functioning protein to be produced and enables individuals with spinal muscular atrophy to develop more normally, and to lead reasonably normal lives.

This is only one example of how our knowledge of RNA splicing can help to improve people’s lives. I have also started two companies that are using knowledge of RNA splicing to develop important medical applications. These companies are called [Biogen](#) and [Alnylam Pharmaceuticals](#), and you can learn more about them by following the links.

It has been extremely exciting to discover RNA splicing, to study the whole splicing process, to think about what it means and how it could be used, and then, 30–40 years later, to see the science turned into a treatment that improves the lives of young children. This is common in science, and it is one of the things that makes doing science, particularly biomedical science, so fulfilling. Not only are scientists discovering deep secrets about biological systems, but occasionally those insights lead to ways of helping others. One of the most gratifying things a person can do in life is to help others, including helping them to understand more about the world, who they are, and their relationship to others. Such understanding is what makes humans unique and interesting.

RECOMMENDATIONS FOR YOUNG MINDS

I want share with you what I learned as I was growing up, going to school, and figuring out what I wanted to do for a career. I found that, if I followed my interests and passions, I was motivated to learn as much as I could about those topics, which allowed me to make important contributions to society. If you try to be too practical in life and do not spend some time dreaming of the future, you might not see enough of the future to fulfill your dreams. However, if you are too day-dreamy, you may never bring yourself to solve the problems you will face as you pursue your goals. So, follow your curiosity and, if it leads you to something you are passionate about, give yourself to it. You do not know what the world will be like 20 or 30 years from now; just do what motivates you to learn and take the time to have fun while you are doing it.

It is also good to remember that life is not a straight path—it often does not advance in predictable steps, but rather in many possible branches. At certain times in your life, you will come to points when you will need to make decisions. One decision will take you one way, and another decision will take you another way. These paths could be quite different from one another, particularly when making decisions that have very long implications, but that is OK. So, when you come to one of those forks in the road, take one! And remember to enjoy it and follow your passion while you are doing so.

ACKNOWLEDGMENTS

I wish to thank [Noa Segev](#) for conducting the interview which served as the basis for this paper and for co-authoring the paper, and Zehava Cohen for the figures.

ADDITIONAL MATERIALS

[RNA Splicing: What is a Gene?—Phillip A. Sharp.](#)

[Nobel Lecture: Split Genes and RNA Splicing—Phillip A. Sharp.](#)

REFERENCES

1. Cooper, G. M. 2000. "The cell: a molecular approach," in *The Origin and Evolution of Cells, 2nd Edn* (Sunderland, MA: Sinauer Associates). Available online at: <https://www.ncbi.nlm.nih.gov/books/NBK9841/>
2. Jacob, F., and Monod, J. 1961. On the regulation of gene activity. *Cold Spring Harb. Symp. Quant. Biol.* 26:193–211. doi: 10.1101/SQB.1961.026.01.024

3. Berget, S. M., Moore, C., and Sharp, P. A. 1977. Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proc. Natl. Acad. Sci. U.S.A.* 74:3171–5. doi: 10.1073/pnas.74.8.3171
4. Berk, A. J. 2016. Discovery of RNA splicing and genes in pieces. *Proc. Natl. Acad. Sci. U.S.A.* 113:801–5. doi: 10.1073/pnas.1525084113
5. Schmucker, D., Clemens, J. C., Shu, H., Worby, C. A., Xiao, J., Muda, M., et al. 2000. *Drosophila* Dscam is an axon guidance receptor exhibiting extraordinary molecular diversity. *Cell* 101:671–84. doi: 10.1016/S0092-8674(00)80878-8

SUBMITTED: 07 October 2022; **ACCEPTED:** 17 January 2023;
PUBLISHED ONLINE: 31 May 2023.

EDITOR: Robert T. Knight, University of California, Berkeley, United States

SCIENCE MENTORS: Asma Bashir and Alexandra Dimitri

CITATION: Sharp PA (2023) RNA Splicing—Cutting and Pasting Genes. *Front. Young Minds* 11:1063940. doi: 10.3389/frym.2023.1063940

CONFLICT OF INTEREST: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

COPYRIGHT © 2023 Sharp. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

YOUNG REVIEWERS

ANASTASIA, AGE: 14

I am Anastasia, a year 10 student captivated by science! I am very passionate about learning and discovering new opportunities. My dream is to study medicine, as I am very interested in how our body works and how it can be affected. Outside of my studies, I love painting, drawing, and crafting, as well as many other extracurricular clubs. But most of all, I love cooking deserts (I have a sweet tooth)!



KENZO, AGE: 10

Hi I am Kenzo, I am 10 years old, always excited and super friendly. I have a lot of friends and make new friends easily. I like listening to songs and playing video games. I also really like fighting and self defense. My favorite hobby is parkour. At school, I enjoy sports, arts, math, and the breaks. I often express my thoughts and talk while the teacher explains so time goes faster ☺. I love reading comics.



**MAXIME, AGE: 10**

My name is Maxime, I love sports, music, and science, especially space. I am very inquisitive and enjoy problem-solving situations. I just got into coding and I am loving it.

**NOURA, AGE: 15**

I am Noura, I have many interests such as building robots and playing the piano, but science has a special place in my heart and I have always been looking for the latest scientific research and projects.

**SAIF, AGE: 15**

Hi! My name is Saif. I am 15 years old. I have always been passionate about science, particularly biology and chemistry. I love to read science books and biographies about fascinating people such as Steve Jobs. I also participate in programs and competitions that enhance and improve my passion for science, such as winning the best group project at the National Science Fair in 2020. I also enjoy doing experiments with my science teacher.

**YOUNIS, AGE: 15**

Hi! My name is Younis, and I am 15 years old. I have always been attracted to science subjects as they are very fascinating, especially environmental management, geography, geology, and space. And through Frontiers for Young Minds I hope to learn more about the world we live in.

**PHILLIP A. SHARP**

Prof. Phillip Allen Sharp an American professor of molecular biology, Prof. Sharp is a member of the Department of Biology and the Koch Institute for Integrative Cancer Research at the Massachusetts Institute of Technology (MIT), Massachusetts, United States. Prof. Sharp studied chemistry and mathematics at Union College and completed his Ph.D. in chemistry at the University of Illinois at Urbana-Champaign in 1969. After his Ph.D., Prof. Sharp did his postdoctoral training at the California Institute of Technology, and then studied gene expression in human cells at Cold Spring Harbor Laboratory. In 1974, Prof. Sharp went to MIT, where he joined the Center for Cancer Research (now the Koch Institute for Integrative Cancer Research) and conducted experiments that led to the discovery of RNA splicing. During his years at MIT, Prof. Sharp was the director of the Center for Cancer Research (1985–1991), served as head of the biology department from (1991 to 1999), and founded and directed the McGovern Institute for Brain Research (2000–2004). Prof. Sharp won numerous awards and prizes, including the NAS Award in Molecular Biology (1980), the Louisa Gross Horwitz Prize (1998), the Dickson Prize (1991), the Nobel Prize in Physiology or Medicine (1993), and the National Medal of Science (2004). Prof. Sharp started two successful biomedical companies with knowledge of RNA splicing for medical applications. *kids@frontiersin.org