

Are you *on the*
[AAAS.ORG](#) | [FEEDBACK](#) | [HELP](#) | [LIBRARIANS](#)

Daily News

Enter Search Term

ADVANCED

[ALERTS](#) | [ACCESS RIGHTS](#) | [MY ACCOUNT](#) | [SIGN IN](#)
[News Home](#) | [ScienceNOW](#) | [ScienceInsider](#) | [Premium Content from Science](#) | [About Science News](#)
[Home](#) > [News](#) > [ScienceNOW](#) > [September 2012](#) > [Human Genome Is Much More Than Just Genes](#)
**Science** **NOW** UP TO THE MINUTE NEWS FROM SCIENCE

## Human Genome Is Much More Than Just Genes

 by [Elizabeth Pennisi](#) on 5 September 2012, 1:00 PM | [8 Comments](#)
[Email](#) | [Print](#)

2

[More](#)
[PREVIOUS ARTICLE](#)
[NEXT ARTICLE](#)

The human genome—the sum total of hereditary information in a person—contains a lot more than the protein-coding genes teenagers learn about in school, a massive international project has found. When researchers decided to sequence the human genome in the late 1990s, they were focused on finding those traditional genes so as to identify all the proteins necessary for life. Each gene was thought to be a discrete piece of DNA; the order of its DNA bases—the well-known "letter" molecules that are the building blocks of DNA—were thought to code for a particular protein. But scientists deciphering the human genome found, to their surprise, that these protein-coding genes took up less than 3% of the genome. In between were billions of other bases that seemed to have no purpose.

Now a U.S.-funded project, called the Encyclopedia of DNA Elements (ENCODE), has found that many of these bases do, nevertheless, play a role in human biology: They help determine when a gene is turned on or off, for example. This regulation is what makes one cell a kidney cell, for instance, and another a brain cell. "There's a lot more to the genome than genes," says Mark Gerstein, a bioinformatician at Yale University.

The insights from this project are helping researchers understand the links between genetics and disease. "We are informing disease studies in a way that would be very hard to do otherwise," says Ewan Birney, a bioinformatician at the European Bioinformatics Institute in Hinxton, U.K., who led the ENCODE analysis.

As part of ENCODE, 32 institutions did computer analyses, biochemical tests, and sequencing studies on 147 cell types—six fairly extensively—to find out what each of the genome's 3 billion bases does. [About 80% of the genome is biochemically active](#), ENCODE's 442 researchers report today in *Nature*. Some of these DNA bases serve as landing spots for proteins that influence gene activity. Others are converted into strands of RNA that perform functions themselves, such as gene regulation. (RNA is typically thought of as the intermediary messenger molecule that helps make proteins, but ENCODE showed that much of RNA is an end product and is not used to make proteins.) And many bases are simply places where chemical modifications serve to silence stretches of our chromosomes.

[Live Chat: Sports-Related Head Injuries](#) Thursday 3 p.m. EDT

ScienceNOW. ISSN 1947-8062

ENCODE's results are changing how scientists think about genes. It found about 76% of the genome's DNA is transcribed into RNA of one sort or another, way more than researchers had originally expected. That DNA includes slightly less than 21,000 protein-coding genes (some researchers once estimated we had more than 100,000 such genes); "genes" for 8800 small RNA molecules and 9600 long noncoding RNA molecules, each of which is at least 200 bases long; and 11,224 stretches of DNA that are classified as pseudogenes, "dead" genes now known to really be active in some cell types or individuals. In addition, efforts to define the beginning end, and coding regions of these genes revealed that genes can overlap and have multiple beginnings and ends.

The project uncovered 4 million spots in our DNA that act as switches controlling gene activity. Those switches can

ADVERTISEMENT

Register for upcoming webinars.

**Access Now!**

ADVERTISEMENT

**GRADUATE PROGRAMS**

Search by university, topic, or location

[click here](#)

Science Careers  
from the journal Science AAAS

be both near and far from the gene they regulate and act in different combinations in different cell types to give each cell type a unique genomic identity. In addition, at least some of the RNA strands produced by the genome also help to control how much protein results from a particular gene's activity. Thus, the regulation of a gene is proving much more complex than expected.

These and other findings appear today in six papers in *Nature*, and 24 in *Genome Research and Genome Biology*. Two additional papers are published today on *Science* online. In a database, [ENCODE has created a map](#) showing the roles of all the different bases. "It's like Google Maps for the human genome," says Elise Feingold, a program director for the National Human Genome Research Institute in Bethesda, Maryland, which funded ENCODE. With Google Maps one can choose various views to see different aspects of the landscape. Likewise, in the ENCODE map, one can zoom in from the chromosome level to the individual bases and switch from looking at whether those bases yield RNA or are places where DNA-regulatory proteins bind, for example.

This catalog "will change the way people think about and actually use the human genome, says John A. Stamatoyannopoulos, an ENCODE researcher at the University of Washington, Seattle.

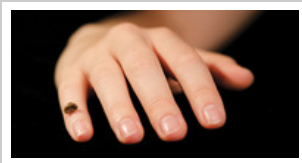
Already he and others are harnessing this information—[much of which is already publicly available](#)—to learn about genetic influences on disease. Many large-scale studies have linked specific base changes to higher or lower risks for disorders ranging from diabetes to arthritis. Now researchers can look to see whether those variants are involved in regulation of some sort and if so, what genes are being regulated. For his study of cancer and epigenetics, "ENCODE data were fundamental," says Mathieu Lupien, a molecular biologist from the University of Toronto in Canada who was not associated with ENCODE.

Follow **ScienceNOW** on [Facebook](#) and [Twitter](#)

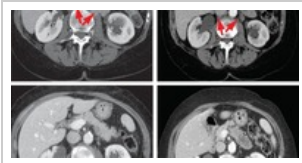
Posted in [Genetics](#)

[Email](#) [Print](#) | [Share](#)  Share **7** retweet [More](#)

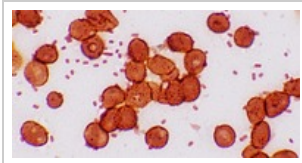
### Related Articles



AUGUST 30, 2012  
**Genome Brings Ancient Girl to Life**




AUGUST 23, 2012  
**Genome Sequencing Clears Up a Cancer Medical Mystery**




AUGUST 22, 2012  
**'Genomic CSI' Helps Contain a Killer**


**8 comments** • **13 reactions**


0 Stars



Discussion  Community  My Disqus 100

 **CZ Yutzman** • 6 hours ago  
We are finally beginning to reverse engineer the programming language of life, the most complex programming in existence.  
1   • Reply • Share ›

 **Preston Garrison** • 4 hours ago  
Who wrote that silly sub-headline? How can the genome have "billions" of control elements when it only has 3 billion nucleotides?  
0   • Reply • Share ›

 **shivaprasad** • 11 hours ago  
The project 'ENCODE' unveiled the function of the non-coding stretch of genetic material. it is really an another window to understand gene expression. construction of panels of all regulatory sequences for an individual gene expression could be 'the end' of genomics research. i wish to wait for the day.  
2   • Reply • Share ›

 **germanfinn** • 15 hours ago  
Is a stretch of DNA in the genome really functional when it is transcribed iust by chance

---

ones in a while. I guess the question is how you define FUNCTION. Just being transcribed would not make it for me!

0 ^ | 1 v • Reply • Share ›

---



**Postdoc** • 16 hours ago

If much of RNA is end product and not used to make proteins, comparison among different techniques such as qRT-PCR, in situ hybridisation, western blots and immunohistochemistry would be more complicated.

0 ^ | 1 v • Reply • Share ›

---



**Lars** • a day ago

Does this mean we might have less or more percentage of dna in common with say for instance chimps?

0 ^ | 1 v • Reply • Share ›

---



**Joe** • 9 hours ago • parent

Doesn't effect how much so much as why. If animals have similar organs etc... they would need similar control structures. What it does not at the moment answer is why so much is common with such different biological structures such as plants and animals.

0 ^ | 1 v • Reply • Share ›

---



**Jon** • 10 hours ago

In a computer program very little of the code is related to performing output. The majority is involved in determining what that output should be.

1 ^ | 2 v • Reply • Share ›

---

#### ALSO ON SCIENCENOW

[What's this?](#) ✕

#### [Semen's Secret Ingredient](#)

0 • 15 comments • 17 days ago

#### [Basic Research Burns Bright in Romney Energy Plan](#)

1 • 9 comments • 14 days ago

#### [DNA: The Ultimate Hard Drive](#)

2 • 11 comments • 21 days ago

---