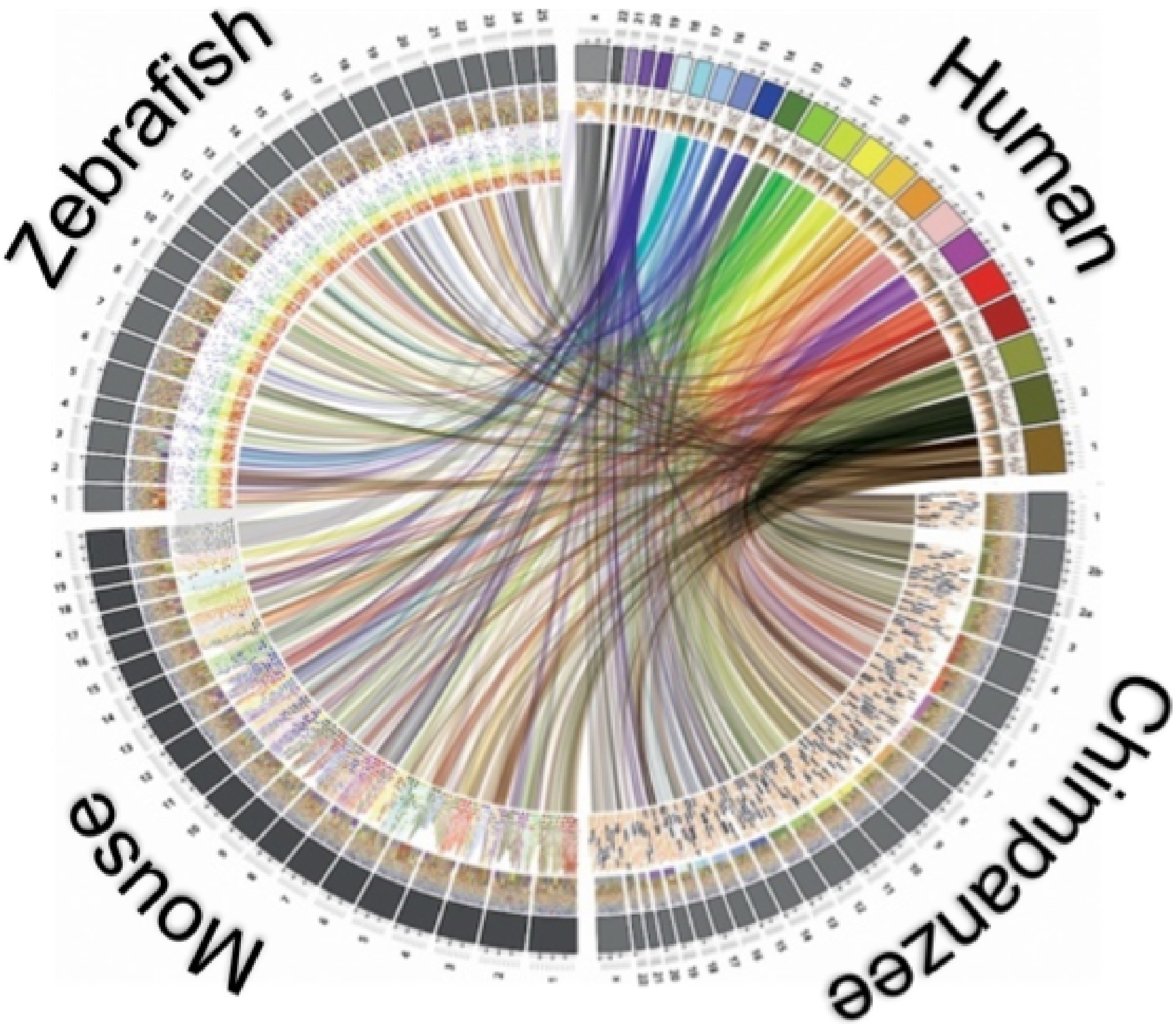
PHYS.ORG

The era of biological cataloging

December 19 2016



Clockwise from top right, the genomes of a human, chimpanzee, mouse and zebrafish are arranged in a circle. Each colored square at the outside of the circle corresponds to a pair of chromosomes, the threadlike packages of DNA stored in the nucleus. Lines connect similar DNA sequences, visually emphasizing just how much DNA humans share with other species. The density of the connections indicates that humans have more in common with chimpanzees than zebrafish. Credit: Martin Krzywinski, Canada's Michel Smith Genome Sciences Center

Have you ever collected coins, cards, toy trains, stuffed animals? Did you feel the need to complete the set? If so, then you may be a completist. A completist will go to great lengths to acquire a complete set of something.

Scientists can also be completists who are inspired to identify and catalog every object in a particular field to further our understanding of it. For example, a comprehensive parts list of the human body—and of other organisms that are important in biomedical research—could aid in the development of novel treatments for diseases in the same way that a parts list for a car enables auto mechanics to build or repair a vehicle.

More than 15 years ago, scientists figured out how to catalog every gene in the human body. In the years since, rapid advances in technology and computational tools have allowed researchers to begin to categorize numerous aspects of the biological world. There's actually a special way to name these collections: Add "ome" to the end of the class of objects being compiled. So, the complete set of genes in the body is called the "genome," and the complete set of proteins is called the "proteome."

Below are three -omes that NIH-funded scientists work with to understand human health.

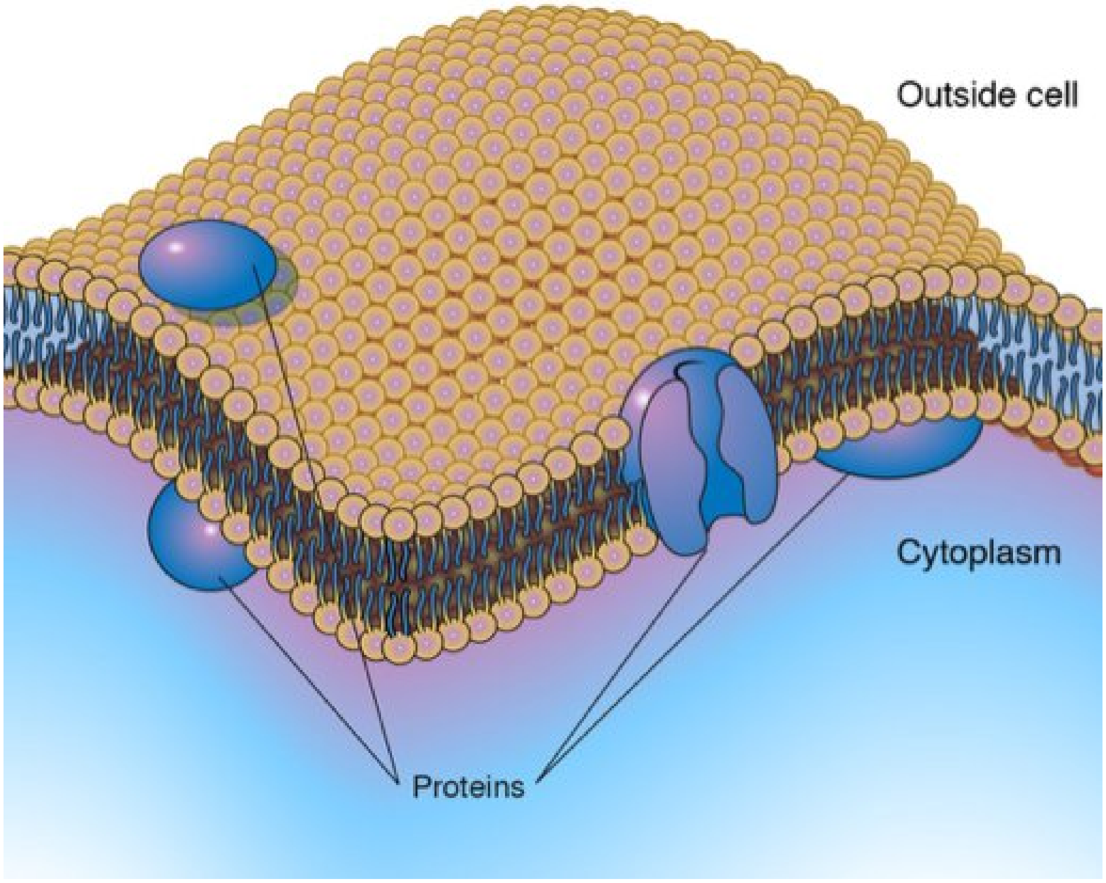
Genome

The genome is the original -ome. In 1976, Belgium scientists identified all 3,569 DNA bases—the As, Cs, Gs and Ts that make up DNA's code—in the genes of bacteriophage MS2, immortalizing this bacteria-infecting virus as possessing the first fully sequenced genome.

Over the next two decades, a small handful of additional genomes from other microorganisms followed. The first animal genome was completed in 1998. Just 5 years later, scientists identified all 3.2 billion DNA bases in the human genome, representing the work of more than 1,000 researchers from six countries over a period of 13 years.

As more individuals' genomes have been sequenced, scientists have found that humans share 99.5% of their genome with each other. However, small differences can be quite important. As the cost of sequencing genomes has plummeted from an initial $3 billion to the current $1,000, scientists are sequencing the genomes of individuals as well as those of additional organisms used to investigate biological questions.

And all the effort has started to pay off. Genomics is beginning to reveal many of the basic components of cells and their interactions. Already, researchers are linking the presence of certain genes in the genome to specific diseases. Furthering our understanding of the genome will have a profound impact on the diagnosis and treatment of disease. Also, comparing the genomes of related and disparate species can shed light on how species evolve over time.



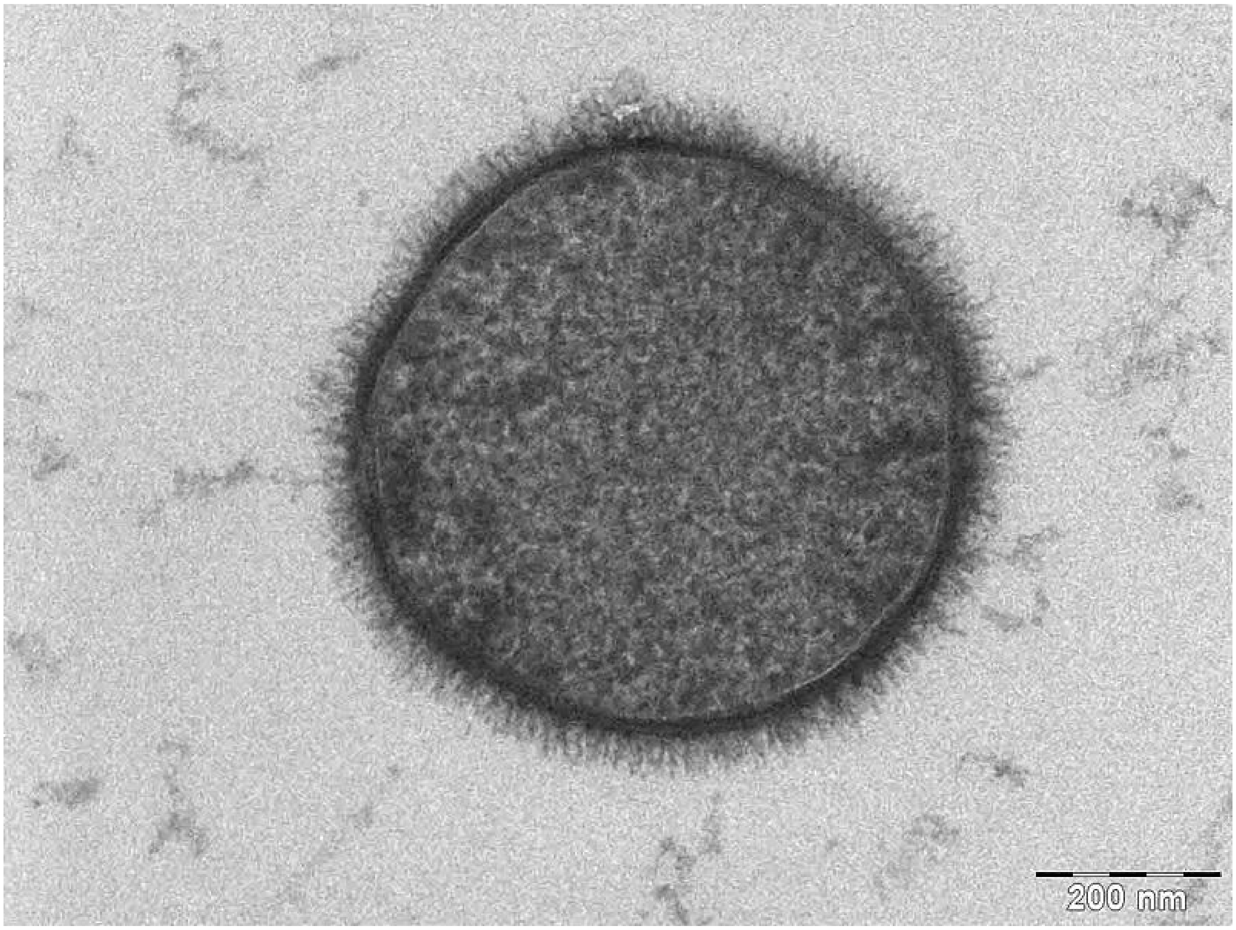
Composed of two layers of lipids (small brown spheres) studded with proteins (bigger purple spheroids), cell membranes form a barrier around cells. Credit: National Institute of General Medical Sciences (NIGMS)

Lipidome

The lipidome is the collection of all the lipids, or fat molecules, within a cell. Cells use lipids to form a continuous lipid membrane around themselves and to separate their inner organelles from each other. These cellular membranes aren't simply for protection. They're also highly organized and dynamic work zones, seeded with proteins that help regulate the way cells attach to other cells, talk to each other, collect nutrients and grow.

The lipid membranes inside the cell can similarly act as points of contact between cellular compartments, and they're involved in nearly every aspect of cellular physiology and function. Recent experiments have revealed hundreds of distinct types of lipids produced by cells. The lipidome has also been found to be remarkably flexible. It's capable of rapid, large- and small-scale rearrangements in response to different situations, including early development and disease. In the case of development, lipids within the membrane reorganize as a cell grows. In the case of disease, viruses delivering their infectious payloads can slam into and rupture the lipid membrane of human cells, causing localized reshuffling.

Disturbances to the lipid components of cellular membranes are associated with diverse diseases, including cardiovascular disease, autoimmunity, osteoporosis, neurological disorders and cancer. Experiments investigating the lipidome of specific cells with known roles in particular diseases could help researchers identify novel treatments.



Thousands of glycans protrude from the bacterium Bactillus subtillus, forming a unique carbohydrate coat. Credit: Wikimedia Commons, Allonweiner

Glycome

The glycome is the complete set of glycans, also known as carbohydrates or sugars, that cells produce. Many of these glycans are linked to proteins and lipids on cell surfaces, where they can interact with molecules on other cells. Single sugars can also act as signaling molecules inside cells, altering gene editing, protein folding and other cellular functions.

A recent study of 650 different species suggests that about 5% of an organism's DNA codes for the proteins that synthesize, degrade and/or recognize and bind to carbohydrates. Mutations in these genes can result in the dysfunction of many organs, underscoring the importance of carbohydrates to human health. In addition, changes in the patterns of glycans in a person's cells can be an indication of a range of diseases, including cancer, inflammatory bowel disease and cardiovascular disease . One day, scientists may use imaging techniques to rapidly identify a cell's glycome to diagnose specific kinds of cancer, for example.

Cells also use the glycans on their outer surface, commonly referred to as "carbohydrate coats," to recognize one another. Likewise, viruses can recognize and bind to carbohydrate coats. By analyzing the carbohydrate binding properties of the flu virus, researchers have been able to design antiviral drugs that interfere with the virus' ability to infect our cells.

Provided by National Institute of General Medical Sciences (NIGMS)