Leafhoppers depend on bacteria for their survival. Some of the bacteria have the smallest genomes ever found. Picture copyright Alex Wild.

THE LOOM: August 23, 2013

And the Genomes Keep Shrinking…

by Carl Zimmer

Here are a few numbers about DNA–some big ones, and then some very small ones.

The human genome contains about 3.2 billion base pairs. Last year, scientists at the University of Leceister printed the sequence out in 130 massive reference-book-sized volumes for a museum exhibit. From start to finish, they would take nearly a century to read.

A typical gene is made up of a few thousand bases. The human genome contains about 21,000 genes that encode proteins. There are other genes in the human genome that encode molecules known as RNA, but how many of those RNA molecules actually do anything useful in the cell is a matter of intense debate. A lot of the human genome is made of neither protein- or RNA-coding genes. Much (maybe most) of it is made up of dead genes and parasite-like stretches of DNA that do little more than making copies of themselves.

As I wrote recently in the *New York Times*, 3.2 billion base pairs and 21,000 genes are not essential requirements for something to stay alive. *E. coli* is doing very well, thank you, with a genome about 4.6 million base pairs. That's .14% the size of our genome. Depending on the strain, the microbe has around 4100 protein-genes. That's about a fifth the number of protein-coding genes that we carry. The high ratio of genes to genome size in *E. coli* is the result of its stripped-down, efficient genetics. Mutations that chop out non-functional DNA spread a lot faster in microbes than in animals.

E. coli, in turn, has proven to be positively gargantuan, genetically speaking, compared to some other species. As scientists explore more of the microbial world, they find species with smaller genomes. In my column for the *Times,* I wrote about the record-holding tiny genome, belonging to a microbe called *Tremblaya.* Its genome is a mere 139,000 base pairs. That's .004% the size of our genome. You could print the entire sequence in a single slim paperback you could slip in your pocket. And in that sleek genome are just 120 protein-coding genes–.6% of our own collection of protein-coding genes.

Whenever I report on such record-breakers, I try to stress that they are only breaking records at that moment. *Tremblaya* has the smallest genome *known*. Or, I should now say, it had the smallest genome known last month.

This month in the journal *Genome Biology and Evolution*, Gordon Bennett and Nancy Moran describe a new record holder, called *Nasuia deltocephalinicola*. It has a genome of just 112,000 base pairs. Imagine taking that slim novella and ripping off the last chapter. Ironically, *Nasuia* packs in *more* genes into its DNA than *Tremblaya*–137 protein-coding genes, Bennett and Moran estimate.

What's really striking about all these current and former record-holders for small genomes is that they all live in a single exotic ecological niche. Without exception, they can be found inside plant-feeding insects. *Tremblaya* lives in mealy bugs, for example, while *Nasuia* lives in a leafhopper (*Macrosteles quadrilineatus*).

Inside those hosts, these microbes carry out chemical reactions on the food that the insects eat. The insects feed on sap and other fluids from plants, which contains few nutrients. But the bacteria can use the compounds floating in the fluid to build amino acids, which the insects can then assemble into proteins.

Leafhoppers, cicadas, sharpshooters, and other related insect species carry related versions of the same stripped-down bacteria. By drawing their evolutionary trees, Bennett and Moran have found that the insects got into a symbiotic relationship with the microbes over 260 million years ago. I've reproduced their tree below for those who want some gory details. The blue lines show *Nasuia* and related lineages of microbes. The insects also acquired another species of bacteria, known as *Sulcia*. Together, these two microbes split the work for millions of years. (In some insects, fungi also jumped into the mix.)

The ancestors of *Nasuia* started out as free-living microbes that had genomes on par with *E. coli*. But once they got inside a host, they were able to lose DNA without paying a price. The insects gave them a stable home, building special organs to shelter them, and they even pass down the bacteria to their offspring. The bacteria cast aside many genes that might otherwise seem essential, such as a number of genes involved in generating energy. All they needed to do was continue to provide a service, by synthesizing some amino acids.

Nasuia holds the record now, but probably not for long. There are many other species of insects left to investigate. Moran had John McCutcheon of the University of Montana have done some back-of-the-envelope calculations to figure out how much smaller the genomes of those symbionts can get. All known insect symbionts share 82 genes in common. It's possible those genes are absolutely required to survive as a symbiont. But a symbiont also needs to provide a benefit to its host, or its host will likely get rid of it. It takes at least 11 genes to synthesize a single amino acid. Those 93 genes, McCutcheon and Moran estimate, could fit in a genome as small as 70,000 base pairs.

It's funny that these bacteria allow us to probe one of the most basic questions about life: how simple life can get and yet still qualify as being alive? While those who make fun of science for a living may consider such research a waste of time, studying these stripped-down organisms is also about as practical as science can get. The leafhoppers that house *Nasuia,* for example, are a nightmare for farmers, causing damage to a wide range of vegetables by spreading fungi and bacteria. Yet they would be helpless if not for their exquisitely simple lodgers. If we can understand how they survive with such tiny genomes, we may be able to stop them from enabling their hosts.