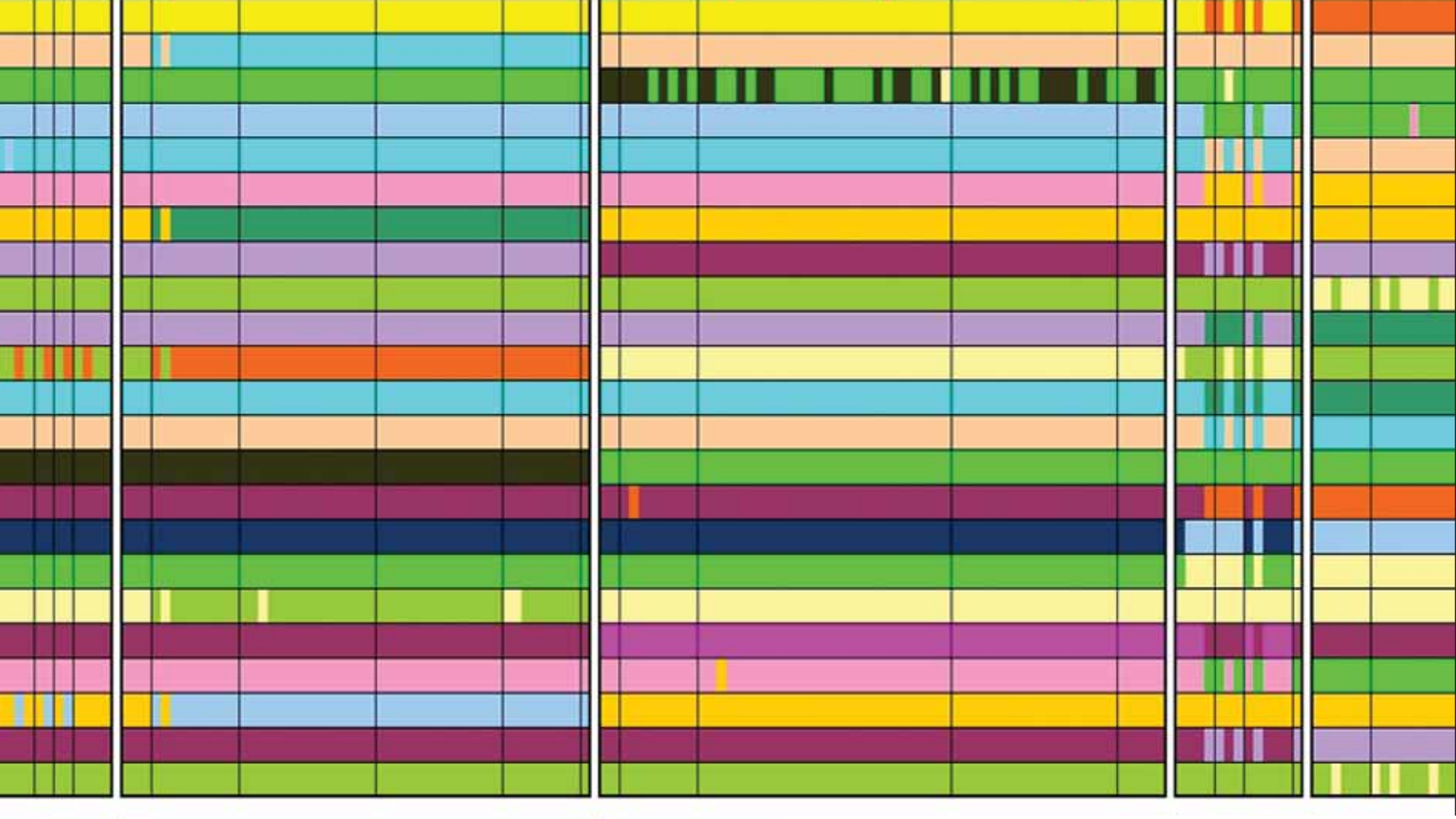
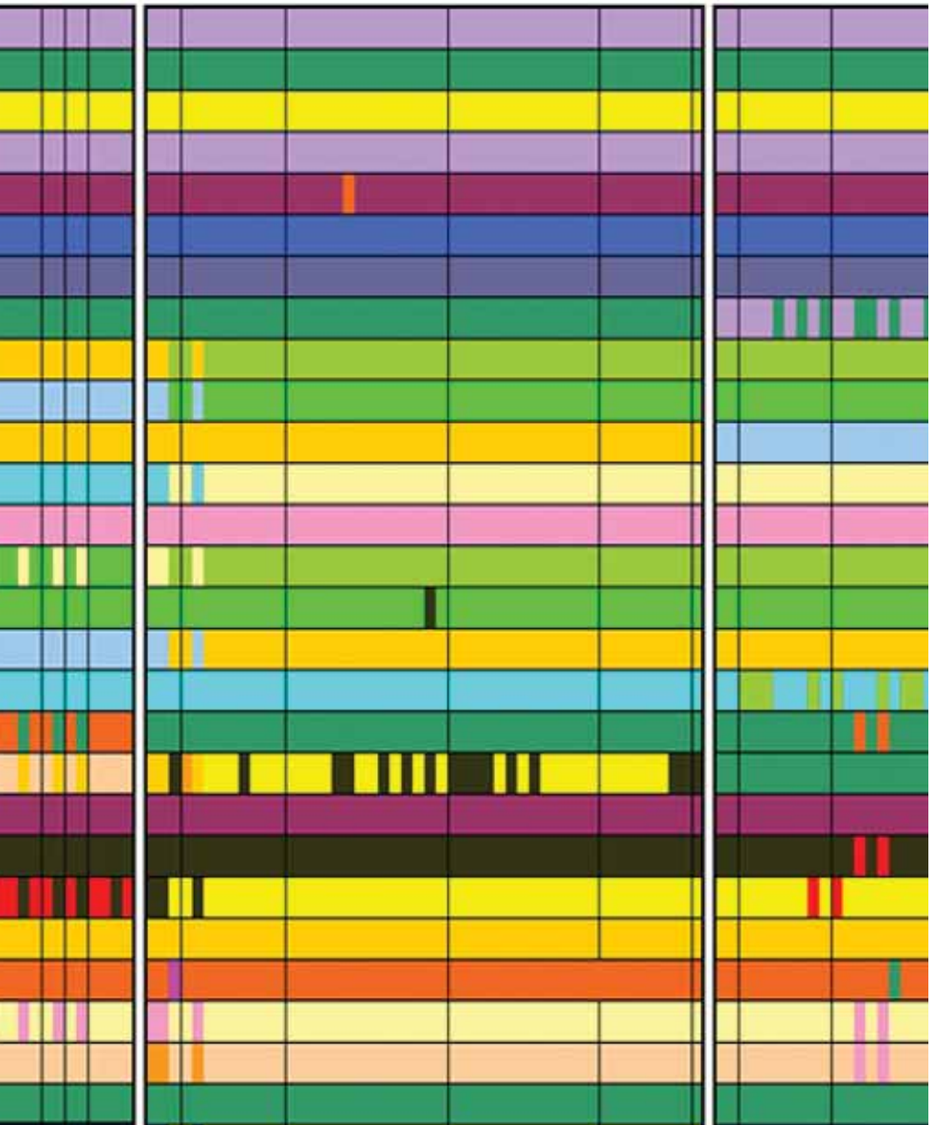
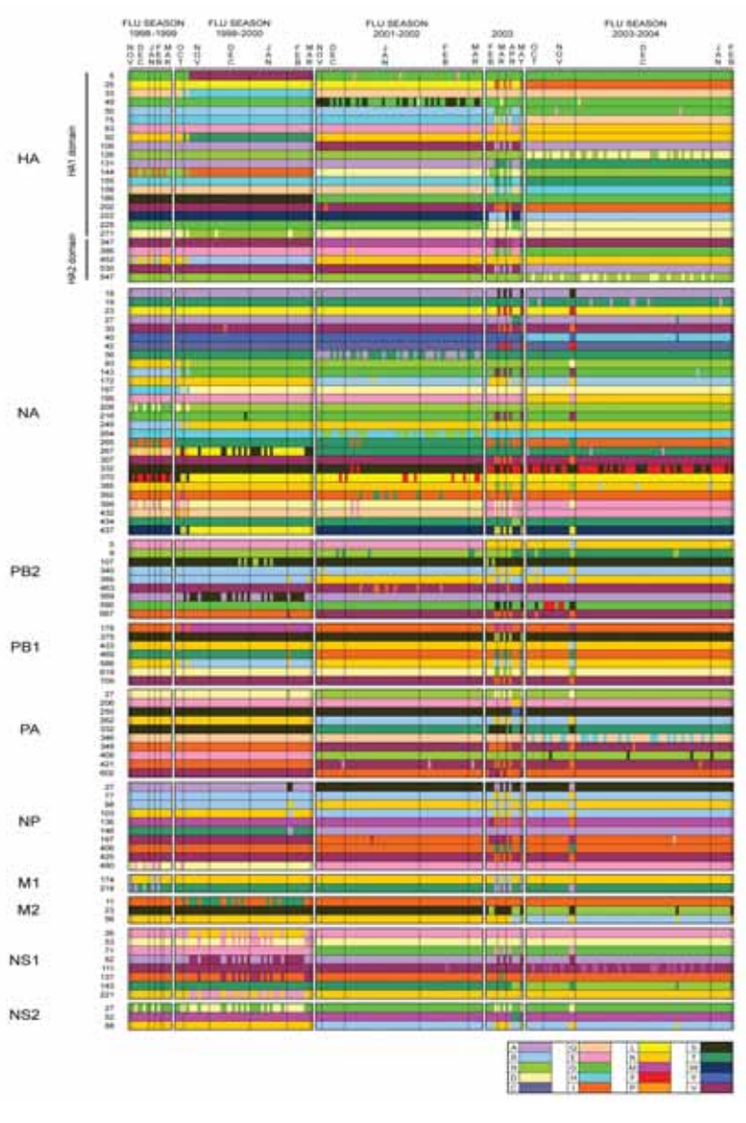





FLU SEASON 1998-1999
FLU SEASON 1999-2000
FLU SEASON 2001-2002
2003
FLU SEASON 2003-2004
HA
HA1 domain
HA2 domain
NA
PB2
PB1
PA
NP
M1
M2
NS1
NS2

SCIENTISTS CHART INFLUENZA'S EVOLUTION, TRACKING HOW A SNIFFLE MIGHT BECOME A SCOURGE | BY CHUCK STARESINIC

EVOLUTION OF FLU

On April 24, a 37-year-old woman in a Sumatran village became ill. It is presumed she developed a fever, cough, and aches. She died 10 days later and was promptly buried, long before any distant authorities became interested in why she died. In the next three weeks, however, six of her relatives died of avian influenza, including her two sons, a sister, a brother, a niece, and a nephew. A second brother also caught the flu. He alone survived.

In this village of a few hundred households, chickens wander freely. Those that appear ill are often eaten. The woman is suspected to have caught the virus from live chickens in the market where she sold fruit and peppers or from chicken manure that she used to fertilize her garden.

At least 252 people have been infected by the H5N1 strain of avian flu in the past few years. More than half died. Fortunately, the virus is not readily transmitted from one person to another. Almost all victims caught the virus directly from birds or from another person who did. The Sumatran case set off alarms for two reasons: It was the largest cluster of cases ever, and it

Influenza diversity is staggering. In this chart, narrow columns represent a total of 207 influenza A viruses collected from 1998 to 2004. Each row represents an altered amino acid position in one viral protein. Amino acids are color coded, so that mutations can be seen as changes in color when scanning from left to right along a row. (The large image is a close-up; the inset shows the entire chart.)

COURTESY NATURE PUBLISHING GROUP, © 2005

marked the first time that the virus was known to have been transmitted from one person to another and then to a third person. It showed that avian flu might be adapting to humans. There is precedent for this.

In the winter of 1918, a doctor in Haskell County, Kan., alerted the U.S. Public Health Service that people in his wide, flat, windblown patch of the Midwest were suffering from "influenza of a severe type." The doctor, Loring Miner, had seen influenza many times before, writes historian John Barry in *The Great Influenza*. But this was different. The disease seemed to race through its victims with great speed. They were wracked by an unusually violent cough and terrible aches. They gasped for air. In the month of February, influenza consumed all of Miner's time and energy. He grew deeply troubled as he crisscrossed the frozen prairie in a horse-drawn buggy. He saw people die that winter who had been in robust health and in the prime of their lives only weeks before. Many experts now believe that an avian flu virus—the H1N1 strain—jumped species in Haskell County.



Elodie Ghedin

Working on viruses in real time is like watching evolution with your finger on the fast-forward button.

When spring arrived, the flu seemed to depart with the wintry weather, even before Miner's alert was published. This is typical of influenza, which is most readily transmitted through dry winter air when people spend long hours together indoors. It's possible that this was the end of the line for Miner's flu—an unusually nasty bug that appeared overnight, raced through much of the population, then fizzled out for lack of new hosts in this sparsely populated area.

But there is another possibility: The virus may have mutated enough to become less lethal and continue circulating. Influenza needs a regular supply of new hosts, because anyone who is infected and recovers becomes immune to that strain. A supply of new hosts was readily available that year. There was a war on, and young men who otherwise might never have left Haskell County traveled hundreds of miles to Camp Funston, where more than 50,000 soldiers lived crowded together like livestock. From there, they would disperse to yet more crowded military bases and, eventually, pack into ships bound for Europe. Many experts believe that soldiers from Camp Funston carried a sleeping demon—a virus that would soon mutate back to its deadly form, circle the globe, and kill millions.

There is one big thing you need to understand about influenza virus. Know this and it all begins to make sense—everything from the pandemic threat of avian flu to why people must be vaccinated annually against more mundane flu strains.

This is what you need to know about influenza virus: It changes. It changes constantly, and it changes quickly. Except sometimes it changes slowly. It evolves in dribs and drabs and then in leaps and bounds. New subtypes appear and disappear with little warning. It changes so much that your immune system won't recognize it from year to year. Sometimes, the vaccine becomes obsolete before flu season is over. Influenza evolves like the sun burns—endlessly and to a degree that's difficult to wrap your mind around.

But let's try. Picture a three-dimensional landscape with mountains, slopes, valleys, and plains. This is the adaptive landscape of influenza A. (Influenza A encompasses the typical flu viruses that become epidemics, the deadly avian flu that concerns us now, and the 1918 strain.) Every possible coordinate on this tumultuous landscape represents a different genetic variant of influenza A. Points that are close together are genetically similar; distant points are genetically varied. The mountains represent fitness in the Darwinian sense—that is, the higher a point is on our landscape, the more successful that particular virus is at generating copies of itself. The lower slopes and valley bottoms are genetic variants that are unlikely to thrive.

The adaptive landscape of influenza A is vast. It's like a relief map of Eurasia, largely unexplored. It has peaks of fitness like the Alps and Himalayas and large, blank regions of mystery that, even today, may as well be marked, "Here be demons," because scientists do not yet understand all of the variations that make a virus lethal, transmissible, drug resistant, or the stuff of pandemics. A virus might have high fitness and remain relatively benign—causing symptoms that eventually subside as the immune system restores order. Others threaten to bring us back to 1918. A small but growing number of scientists are trying to map this world and learn how the viruses in circulation move about the map—how they scale the deadliest peaks, how they leap suddenly from one place to another, and what we can do to be ready when they do.

Some strains of flu virus are always circulating in the human population. Hippocrates is said to have described an influenza epidemic in 412 B.C.E.

When a virus crosses the species barrier and adapts to humans, a pandemic can result. The 1918 Influenza killed between 20 million and 50 million people as it ran its course through 1920; it was probably the deadliest outbreak of disease in human history. The Asian Flu of 1957–58 killed more than 1 million people, and the Hong Kong Flu of 1968–69 killed at least three quarters of a million people. Many experts say we're due for another pandemic.

Scientists could not be certain influenza actually was caused by a virus until the virus was isolated in the 1930s. During the 1918 outbreak, suspicion had largely fallen on an opportunistic bacterium, which now bears the somewhat apocryphal name *Haemophilus influenzae*. (It often turned up in the lungs at autopsy, having caused pneumonia in people weakened by influenza.)

The virus didn't get its due respect in the genomic revolution of the past decade—not until as recently as 2004, anyway. That's when Elodie Ghedin, a molecular biologist working at The Institute for Genomic Research (TIGR), was given a new assignment. TIGR (pronounced *tiger*) is a nonprofit institute in Maryland dedicated to unraveling the genetic code of living organisms, especially important microbes. As a junior investigator at TIGR, Ghedin stood out as one of the few people who had time to take on a big, new project, she says, with a laugh. TIGR had a large grant from the National Institute of Allergy and Infectious Disease for sequencing human pathogens. With a small team of colleagues, Ghedin was tasked with figuring out influenza. She had worked extensively on the genomes of human pathogens, but never that of influenza virus.

Ghedin is now a Pitt assistant professor of medicine. Her first name rhymes with "melody." She is quick to smile. A first-generation Canadian, Ghedin was pleasantly surprised to

find that Pittsburgh reminds her of the Montreal of her childhood, with its working-class, closely built neighborhoods that are by turns beautiful, historic, a bit sad, welcoming, insular, pedestrian, and newly resurgent. She is the child of French immigrants, a product of Montreal's public schools, and a PhD graduate of that city's McGill University. She first came to the United States for a postdoc in molecular parasitology at the National Institutes of Health (NIH). There, she continued working on her thesis subject: *Leishmania donovani*, which causes visceral leishmaniasis in half a million people annually, mainly in Africa and Southeast Asia. (After a second NIH postdoc in parasite genomics, she and several collaborators would publish the complete genomes of three related pathogens, including *Leishmania major*, in *Science* in 2005.)

At the time she was given the influenza job, fewer than 10 samples of flu virus had been fully sequenced (with their genetic code transcribed molecule by molecule). The few sequences available were of viruses chosen because of unusual characteristics or rarity. For a virus as variable as influenza, this is like trying to learn about humanity by looking at the genomes of six people with unusual genetic conditions. Ideally, hundreds or thousands of random, complete genomes would be needed to begin to unlock the secrets of influenza. But sequencing was an expensive and painstaking process. Ghedin's team had to create a new way to do it—faster, cheaper, and in higher volume.

Influenza is simple, in a way. It consists of a spiky ball made of two interlocking proteins. Inside are eight lengths of RNA that comprise the virus' eight genes.

The genomics methods that work so well for sequencing human DNA would not work for an RNA virus like influenza. To sequence the human genome, scientists used what they call the "whole genome shotgun" approach. "You just shatter it," says Ghedin. "Just break it up. Then you clone these pieces and you can get a whole genome [when you reassemble them]. You can't do a whole genome shotgun on RNA. You have to reverse transcribe it into DNA. And when you do that, you are bound to miss pieces." It was the ends of the pieces that were being missed, making it impossible to reconstruct a whole genome. Ghedin's team's solution?

"We just glued the pieces together. … We got rid of the ends. Well, that was inefficient," she admits, laughing.

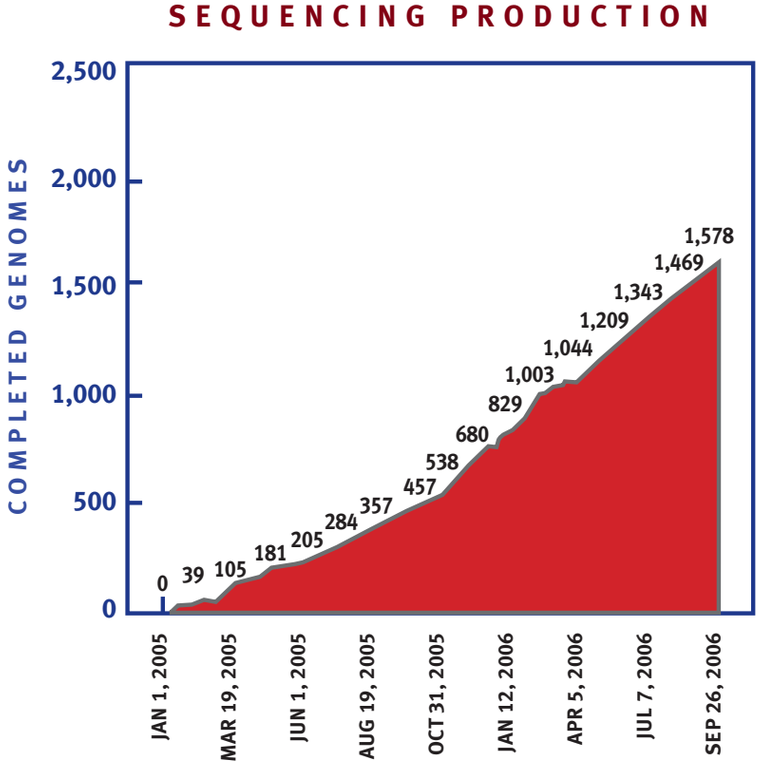
In the end, they settled on a novel version of the whole genome shotgun. They copied the influenza genome into very tiny overlapping pieces—about 250 base-pairs long. By putting the overlapping pieces together, they could reconstruct the entire genome. The volume of work they managed to complete in short order was unprecedented in flu research. They employed robots to inoculate the large number of 96-well plates containing the viral RNA to be duplicated. Then the assembly line went to work.

"We built little teams of people," says Ghedin, who retains an adjunct position at TIGR. "And that's how it works, still. There is a team that does the reverse transcription and the amplification [duplication] of the pieces. Then it's passed on to another team that will do the sequencing. Then another team—the informatic team—will pull all these sequences, clean them up, trim them, and assemble them. Then another team, called the closure team, are the guys that go through and make sure there are no holes in the genome. … The product is then sent to GenBank [the public database of genomic information], which does the annotation. They have a team there that determines whether the coding regions are good. They would come back to the TIGR team and say, 'There was this area that looks a bit strange.' They are the ones that post all the data.

"The brunt of the project was really on the bioinformatics end," says Ghedin. "And this type of project could not have been done 10 years ago, when we didn't have the bioinformatics tools."

"The folks at TIGR have a real institutional experience and understanding of how to wrap things up," says David Lipman, director of the NIH's National Center for Biotechnology Information, which manages GenBank. "And Elodie herself really was very central to solving a number of challenges there."

In less than two years, the team has sequenced more than 1,500 flu viruses in circulation between 1996 and the present. Most have come from upstate New York, where that state's health department conducts exten-



When Ghedin's team set out to sequence the genomes of influenza A viruses in 2004, influenza A had only been sequenced a handful of times. By the time you read this, the Influenza Sequencing Project will have more than 1,600 viruses sequenced. More sequences are added to the public database every week.



sive surveillance of the flu at the Wadsworth Center, near Albany. New sequences are posted as soon as they are completed. Ghedin's team is about to become even more efficient: Here at Pitt, clinicians and infectious disease experts will soon have access to a wildfire-fast technology called Pyrosequencing for partial sequencing of viruses to detect, for example, drug resistance. How fast? It would take about 12 hours to evaluate and compare 96 flu viruses.

Eddie Holmes, an evolutionary biologist, fairly gushes about the influenza project data, and he does so in a British accent that can be difficult to comprehend when he's excited. A professor of biology at Penn State, he does phylogeny, meaning he constructs branching "family trees" showing how organisms evolve over time. Working on viruses in real time, he says, is like watching evolution with your finger on the fast-forward button.

"An RNA virus like flu evolves about a million times faster than humans, okay? So, a phylogeny of 10 years of influenza virus evolution is something like 10 million years of human evolution. That's the difference. And what you see is buckets of diversity. You see lineages being born, dying out, replacing each other—a big sort of melting pot of diversity, [with] lots of various evolutionary processes: selection, drift, extinction, replacement, and reassortment (kind of a sexual thing) going on in single populations."

Virus sex? Not exactly, but influenza evolves by two main mechanisms, one of which is analogous to sexual reproduction. Scientists sometimes refer to the mechanisms as "drift" and "shift." Drift, aptly named, is a rather passive process. It's the result of random errors in the virus' RNA as it is reproduced inside infected cells. RNA is much more error-prone than DNA. In the case of viruses, this is an advantage, because random errors, or mutations, sometimes help the virus adapt. For example, a change to its surface protein might mean that the immune system doesn't recognize it as flu. Other mutations can change the way the virus makes you sick or reacts to medication.

Shift, or reassortment, is a much more dynamic process, involving two flu viruses contributing a portion of their genes to create one dramatically new virus—sort of like how your parents created you. It occurs when a cell is infected with two different influenza viruses simultaneously. The cell's machinery is hijacked so that it produces many copies of influenza's eight genes. But with two viruses infecting this cell, there are actually 16 influenza genes being copied. They mix and mingle and come together in various new combinations. Then, they are enclosed in a new capsid—the spiky shell of interlocking proteins that completes the construction of the new virus.

"I think they found a whole new way of looking at influenza diversity and evolution that had been missing before," Holmes says. "We found the virus was much more diverse in single populations than we had thought before. We found lots of reassortments."

"It was thought that drift mutations—just the little sequence mutations—were a major component of the way influenza evolves," says Kirsten St. George, director of the clinical virology program in the Wadsworth Center. "But it has become very clear with the analyses of the whole genomic sequence of hundreds of [viruses] that reassortment is constant and frequent in the population. It is a major, major component of the evolutionary path of influenza."

In her office in Pitt's Scaife Hall, Ghedin points to a colorful chart on the wall of her office (see p.18). "My favorite poster," she says. Each color represents a different amino acid, a key component of the structure of the flu virus. The chart shows 207 flu viruses sampled in New York from 1996 through 2004. She points to a narrow column in November 2003, representing just three viruses out of dozens sampled that season. The colors match all of the other viruses in circulation at the time—but only at the very top of the chart, on the gene for the surface protein. At the lower part of the chart, which shows the other seven genes, the colors are wildly different.

"The take-home message of this one is that you can see reassortment occurring," she explains. "And we knew that was occurring, but in this case it was really important, because it involved the surface protein." In this narrow band representing three viruses, a surface protein appears to have been transplanted from one set of viruses to another. The human immune system only recognizes the flu virus by its surface proteins, so swapping one for another can be like donning a disguise. The results can be dramatic.

Reassortment may be the event that triggers the next pandemic. Imagine a farmer in China who is sick with a typical human flu and a deadly avian flu that he picked up from his chickens. He is dying, and his family gathers round him as he lies coughing, gasping, and shedding a new virus that is a combination of a deadly avian flu and a highly transmissible human flu.

In 2005, scientists modeled a theoretical avian flu outbreak. They predicted that the World Health Organization would have, at most, three weeks to trace the contacts of all those exposed and distribute antiviral medications in a protective ring around the outbreak. But in Sumatra, the first WHO official reached the site of the outbreak approximately 19 days after the first woman became sick. Others arrived a week later. Villagers were suspicious of health officials and did not always cooperate. Some of those infected apparently fled the hospital while still shedding virus.

Meanwhile, people moved about the country as people always do. They squeezed aboard crowded buses, visited distant relatives, cared for the sick, and sent their kids off to school. They boarded flights bound for the capital cities of Asia and beyond. Fortunately, the virus did not completely adapt to humans this time, and the chain of infection ended.

In North America, teams of scientists quietly went about the work of translating the peculiar alphabet that is RNA, learning to spell influenza 1,500 different ways.

"Genomics has to really be looked at as a discovery science," says Ghedin.

"You may go in with a preconceived idea, but then you have to be really open to learning something completely different."

In other words, she and her colleagues have every intention of learning practical things about influenza, such as which genes confer drug resistance and which virus is likely to be dominant next year. But the work is not all hypothesis driven. They are simply sequencing all of the random viruses they can get from one region—creating a mountain of data for scientists to mine. The questions they ask and the answers they find may warn us of the next pandemic, or they may ultimately teach us about other RNA viruses like HIV, about the evolution of parasites, or about the properties of our own RNA and DNA.

"Increasingly, we see in biomedical research the interconnectedness of life," says Lipman, "and discoveries in one area can all of a sudden inform searches in another area." ■