The clock-watcher

Biomathematician **Steve Horvath** *has discovered a strikingly accurate way to measure human ageing through epigenetic signatures.*

BY W. WAYT GIBBS

As a teenager in Germany, Steve Horvath, his identical twin Markus and their friend Jörg Zimmermann formed 'the Gilgamesh project', which involved regular meetings where the three discussed mathematics, physics and philosophy. The inspiration for the name, Horvath says, was the ancient Sumerian epic in which a king of Uruk searches for a plant that can restore youth. Fittingly, talk at the meetings often turned to ideas for how science might extend lifespan.

At their final meeting in 1989, the trio made a solemn pact: to dedicate their careers to pursuing science that could prolong healthy human life. Jörg set his eye on computer science and artificial intelligence, Markus on biochemistry and genetics, and Steve says that he "planned to use mathematical modelling and gene networks to understand how to extend life". Jörg did end up working in artificial intelligence, as a computer scientist at the University of Bonn in Germany, but "Markus fell off the wagon", his brother says, "and became a psychiatrist".

Steve, now a human geneticist and biostatistician at the University of California, Los Angeles (UCLA), says that he finally feels poised to make good on the promise. Through a hard-fought project that involved years of solo work, multiple rejections by editors and reviewers and battling through the loss of a child, he has gathered and analysed data on more than 13,000 human tissue samples[1]. The result is a cellular biological clock that has impressed researchers with its accuracy, how easy it is to read and the fact that it ticks at the same rate in many parts of the body — with some intriguing exceptions that might provide clues to the nature of ageing and its maladies.

Horvath's clock emerges from epigenetics, the study of chemical and structural modifications made to the genome that do not alter the DNA sequence but that are passed along as cells divide and can influence how genes are expressed. As cells age, the pattern of epigenetic alterations shifts, and some of the changes seem to mark time. To determine a person's age, Horvath explores data for hundreds of far-flung positions on DNA from a sample of cells and notes how often those positions are methylated — that is, have a methyl group attached.

He has discovered an algorithm, based on the methylation status of a set of these genomic positions, that provides a remarkably accurate age estimate — not of the cells, but of the person the cells inhabit. White blood cells, for example, which may be just a few days or weeks old, will carry the signature of the 50-year-old donor they came from, plus or minus a few years. The same is true for DNA extracted from a cheek swab, the brain, the colon and numerous other organs. This sets the method apart from tests that rely on biomarkers of age that work in only one or two tissues, including the gold-standard dating procedure, aspartic acid racemization, which analyses proteins that are locked away for a lifetime in tooth or bone.

"I wanted to develop a method that would work in many or most tissues. It was a very risky project," Horvath says. But now the gamble seems to be paying off. By the time his findings were finally published last year[1], the clock's median error was 3.6 years, meaning that it could guess the age of half the donors to within 43 months for a broad selection of tissues. That accuracy improves to 2.7 years for saliva alone, 1.9 years for certain types of white blood cell and 1.5 years for the brain cortex. The clock shows stem cells removed from embryos to be extremely young and the brains of centenarians to be about 100.

"Such tight correlations suggest there is something seemingly immutable going on in cells," says Elizabeth Blackburn of the University of California, San Francisco, who won a Nobel prize for her research on telomeres — caps on the ends of chromosomes that shorten with age. It could be a clue to undiscovered biology, she suggests. And there may be medical implications in cases in which epigenetic estimates do not match a person's birth certificate.

In the months since Horvath's paper appeared, other researchers have replicated and extended the results. The study has stirred up excitement about potential applications, but also debate about the underlying biology at work.

"It's something new," says Peter Visscher, chair of quantitative genetics at the University of Queensland in Australia. "If he's right that there is something like an inherently epigenetic clock at work in ageing, that is very interesting. It must be important."

CLOCKING ON

Horvath kept his vow to the Gilgamesh project by supplementing his PhD in mathematics with a doctorate in biostatistics, which led to a position in the genetics department at UCLA in 2000. After receiving tenure in 2006, he began to focus on ageing by searching for shifts that occur in gene activity over the course of life. A doctoral student took the lead, feeding gene-transcription data through statistical filters in the hope of turning up a robust biomarker for age. But after more than a year, Horvath and the student had found no strong clues. If any such signal exists in gene-transcript data, they concluded, it is hopelessly swamped by the noisy variations from organ to organ and person to person. With little to show for their work, "I decided to keep quixotic projects like this away from students and postdocs," Horvath says. "It didn't seem fair to risk their careers."

Things began to look up in 2011, however. As part of a team led by his UCLA colleague Eric Vilain, Horvath had analysed methylation patterns in DNA extracted from the saliva of 68 adults. The researchers were looking for an epigenetic pattern that correlated with sexual orientation. None turned up, but with the data in hand Horvath and his colleagues decided to see whether they could use it to predict age.

In human DNA, methyl groups most often attach at 'CpG sites' — places where a cytosine precedes a guanine in the DNA. A typical human genome contains more than 28 million such sites. But the microarray technology used to detect methylation samples finds only a fraction of them: older machines pin down just 27,000 sites and newer ones around 485,000.

Horvath got lucky. He found success with a simple statistical model, which looked at how many cells in a drop of saliva have DNA methylated at just two particular CpG sites. The index roughly paralleled participants' ages with a correlation of 0.85, or 85%, and an average accuracy of about five years[2].

While working on a subsequent study, Horvath identified methylation patterns that hewed even more closely to age in very different cell types, such as brain and blood. Suddenly, a goal that he had thought impossible — finding a biomarker for the age of almost every part of the body — seemed attainable.

But it would not be easy. He would have to pull together myriad data sets that included both peoples' ages and their DNA methylation information. Methylation profiles are used for many kinds of medical research — usually in areas other than ageing (see *Nature* **508,** 22; 2014). And because of variations in the way they are collected and processed, they can be tricky to compare. Horvath worried: "How do you make data sets comparable if they were generated by different labs using different protocols?"

Building on work by Andrew Teschendorff at University College London, Horvath devised a way to normalize methylation profiles and put them all on the same footing. Beyond that, his audacious strategy for dealing with some of the uncertainty was to ignore it — and hope that it didn't clobber the accuracy of his model.

It didn't. By early 2012, his algorithm was using 16 CpG sites in the genome, and was returning correlations with chronological age of 96% in nine kinds of tissue. The accuracy was astonishing: median errors were within three years for blood samples and just 18 months for cheek swabs.

But the editors of two journals rejected Horvath's paper. The "tenor of the reviewers was that it was just too good to be true," he says. They suspected that the clock model fit the training data used to build it but that Horvath had insufficient test data to validate it thoroughly.

Humbled but undaunted, Horvath continued collecting data sets and expanding the algorithm. By December 2012, his methylation database spanned 51 types of non-cancerous tissue and cells, plus 20 kinds of cancer. The age estimator had grown to include 353 CpG sites.

He had completed his analyses and was preparing to rewrite his paper from scratch when

BRAD SWONETZ/REDUX/EYEVINE

his pregnant wife's waters broke — more than three months ahead of her due date. For the next 20 days, he barely left the chair beside her hospital bed while she and the medical staff tried to stave off infection and premature delivery.

The stress focused him. "I wrote every hour as if it was the last hour I had for finishing the article," he says. He made good progress, as did his wife and their baby. Towards the end of the third week, with Christmas approaching, "I started to feel really hopeful," he says. But suddenly the baby's heart rate shot up. After an emergency Caesarean section, the baby struggled to breathe. "The doctors made a heroic effort," Horvath says, "but she died in my hands on the day of her delivery. It wasn't until 10 days later that I found enough strength to upload the paper to *Genome Biology*."

The reviews came back in the spring: more disbelief, and another rejection. Horvath didn't blame the reviewers for being sceptical. "Everyone who develops biomarkers knows what to expect: a very strong biomarker gives you a correlation of, say, 0.6 or 0.7." For example, the correlation between age and the length of telomeres is less than 0.5. For Horvath's clock algorithm, that figure is 0.96. He confesses that he had trouble believing it himself until other researchers independently confirmed the tight association.

This time, Horvath refused to take no for an answer. "After reading the reviewers' comments, I spent the next 10 minutes doing three things that one should never do," he says. "First, I went to the fridge and drank three bottles of beer as fast as I could. Second, I went back to the computer and drafted a letter to the editor. Third, I sent it off."

ABOUT TIME

The appeal worked, and after his article[1] was featured in the October 2013 issue of *Genome Biology*, others began downloading the epigenetic-clock program from Horvath's website to test it on their own data. Marco Boks at the University Medical Centre Utrecht in the Netherlands applied it to blood samples collected from 96 Dutch veterans of the war in Afghanistan aged between 18 and 53. The correlation between predicted and actual ages was 99.7%, with a median error measured in months.

At Zymo Research, a biotechnology company in Irvine, California, Wei Guo and Kevin Bryant wondered whether the program would work on a set of urine samples Zymo had collected from 11 men and women aged between 28 and 72. The correlation was 98%, with a standard error of just 2.7 years. "That's amazingly good," Bryant says. "Urine samples weren't even part of the data that Steve used to develop this algorithm."

Horvath's method has many potential applications. Criminal investigators, for example, might find an epigenetic clock handy for establishing the age of a victim or an assailant by analysing any biological residues left behind. Trey Ideker, chief of the medical-genetics division at the University of California, San Diego, says that his group is working with a forensics lab to test an epigenetic clock that he and his collaborators designed to work specifically on blood, using mathematical methods very similar to Horvath's[3]. Although Ideker's clock is tissue-specific and not quite as accurate, it could be cheaper to use because it is based on fewer CpG sites — 71 rather than 353.

> *"The big question is whether the clock measures a biochemical process that serves a purpose."*

Both Ideker and Horvath expect that the most interesting use of the clock will be to detect 'age acceleration': discrepancies between a person's epigenetic and chronological ages, either overall or in one particular part of their body.

Such discrepancies could be signs that something is awry. In work due to be presented at the November meeting of the Gerontological Society of America, Brian Chen of the US National Heart, Lung, and Blood Institute (NHLBI) in Framingham, Massachusetts, teamed up with Horvath and others to analyse methylation data collected on more than 2,100 men and women aged 40 to 92 as part of the Framingham Heart Study. The researchers concluded that for every five-year increase in age acceleration, the risk of dying from any cause during the study jumped by 15%. Horvath says that unpublished work from two other large studies also finds epigenetic age acceleration to be a substantial risk factor for mortality, even after controlling for chronological age and other well-known risk factors.

Researchers are also comparing the ages of different tissues from the same individual, in the hope of identifying more accurate, less invasive ways to diagnose disease or gauge the risk of future illness. Last year, Ideker and his collaborators reported that the epigenetic ages of breast, kidney, lung and skin cancers were 40% older, on average, than the patients from which they were removed[3]. The picture from Horvath's method is less clear. Some cancers, such as brain tumours, seemed to be decades older, in terms of their methylation, than they should be. But the effect was reversed for some other cancers, such as certain types of endometrial and breast tumours.

Distortions in epigenetic age seem to parallel other diseases more closely. Horvath says that recent work has found that people with HIV who have detectable viral loads appear older, epigenetically, than healthy people or those with HIV who have suppressed the virus. Another study, not yet published, observes that some tissues show significant age acceleration in morbidly obese people, he reports. In the coming months, he will be mining the vast Women's Health Initiative database — which includes thousands of methylation profiles gathered as part of this 20-year, 160,000-person study spearheaded by the NHLBI — for more links.

AN AGE-OLD QUESTION

Medical researchers might be able to use the epigenetic clock to better diagnose and classify illnesses even without really understanding how the biology works. But Horvath hopes that the science won't stop there.

"The big question is whether the clock measures a biochemical process that serves a purpose," he says. His best guess is that the clock corresponds to the function of an epigenomic housekeeping system, which helps to stabilize the genome by maintaining methylation patterns. The more active this mechanism, he proposes, the faster the epigenetic clock ticks.

Because methylation is usually reversible, Wei says, it might be possible to grab the minute hand of the epigenetic clock and retard its incessant progress — an idea that makes Horvath's solemn adolescent vow sound almost attainable. "The greatest hope is that this clock measures the output of a process that really does relate to ageing — even causes ageing," Horvath says.

But some are sceptical. Teschendorff's research has shown that genome-wide patterns of methylation drift gradually as the years slip past[4]. He suspects that some passive process is behind the shift and that it leads to ageing and disease mainly by interfering with the ability of stem cells to differentiate. Ideker agrees that the epigenetic transition from young to old could be mostly random, in which case there may be nothing especially informative about the 353 cogs of Horvath's clock.

Horvath acknowledges that it will take more work to find out whether epigenetic age predicts the onset of disease and decrepitude better than a calendar does. "But the epigenetic clock gives us a new start and a new hope of something that will affect ageing," he says. In that way, Gilgamesh's ancient quest for a way to delay the inevitable lives on. ●

W. Wayt Gibbs *is a freelance science writer in Seattle, Washington.*

1. Horvath, S. *Genome Biol.* **14,** R115 (2013).
2. Bocklandt, S. *et al. PLoS ONE* **6,** e14821 (2011).
3. Hannum, G. *et al. Mol. Cell* **49,** 359–367 (2013).
4. Teschendorff, A. E., West, J. & Beck, S. *Hum. Mol. Genet.* **22,** R7–R15 (2013).