The thousand-dollar genome

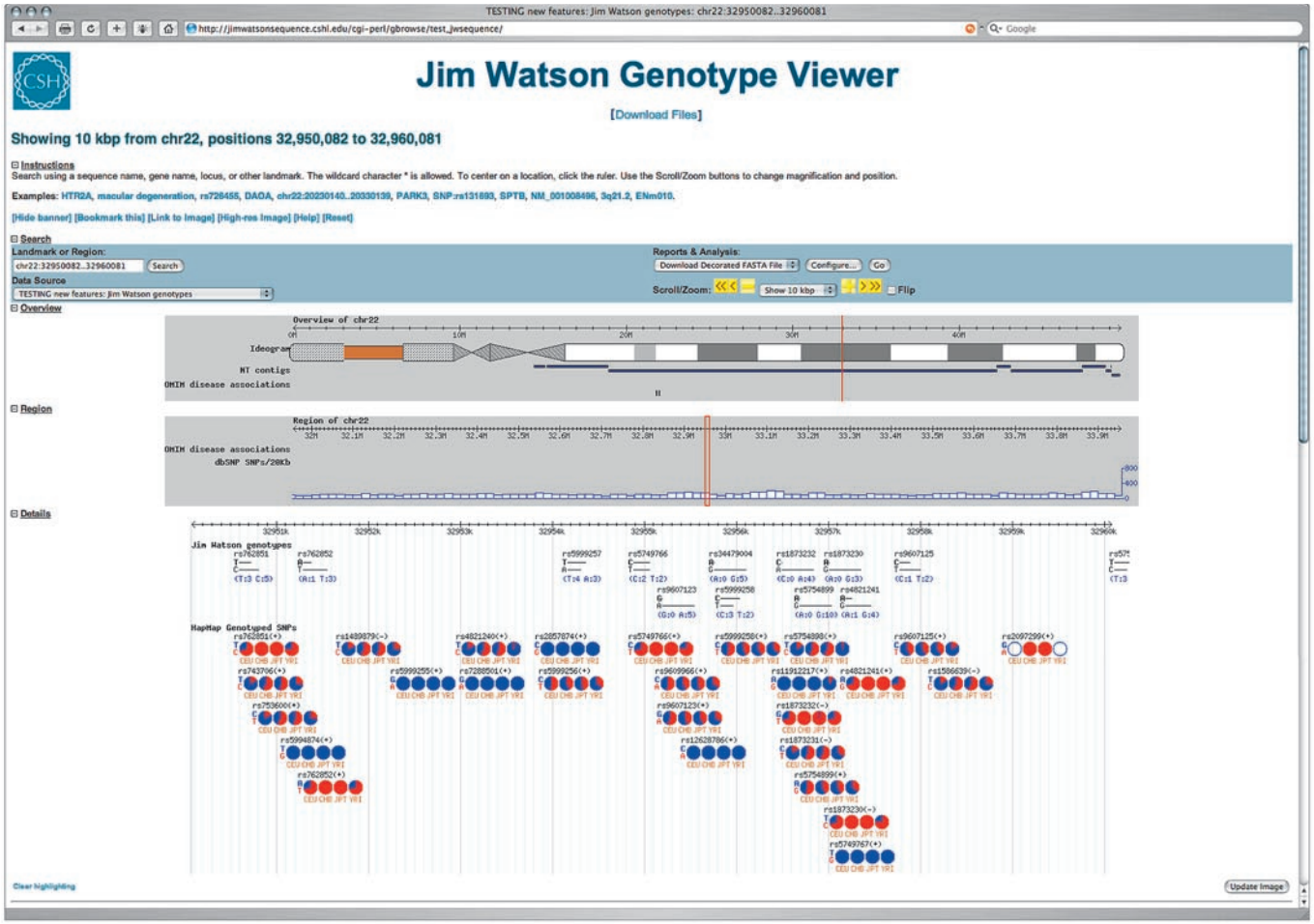
Genetic brinkmanship or personalized medicine?

On May 31 this year, James Watson received the DNA sequence of his full genome in a ceremony at the Baylor College of Medicine in Houston (TX, USA). It marked the end of a two-month project by 454 Life Sciences, a biotech company in Branford (CT, USA) specializing in DNA sequencing, which generated the raw data from a blood sample given by Watson. Baylor College then verified whether the sequence encompassed the entire genome and, on July 6, Watson posted his genome—excluding a portion including the *ApoE* gene, which he had asked to be redacted—online at the GenBank database maintained by the US National Center for Biology Information (Bethesda, MD, USA). Watson, the co-discoverer of the double helical structure of DNA and the father of the Human Genome Project, must have expected to be the first person to make his genome publicly available. He was not. Nine days earlier, on June 27, J. Craig Venter, the inventor of the 'shot-gun' sequencing strategy, who once acknowledged a love–hate relationship with his former boss Watson, had posted his own complete genome online prior to a publication (Levy *et al*, 2007).

There was no official race between Venter and Watson to be the first. The timing of the publications of their genomes was the result of two unrelated projects. But what is the rationale for making the genome data of noted geneticists publicly available in the first place? Is it pure scientific research or genetic exhibitionism? Or has science finally embraced celebrity culture? The media hoopla surrounding the sequencing of Watson's genome has already had some commentators worrying that genome sequencing could become the next must-have for the rich and privileged (Check, 2007). However, beyond the publicity, it is only a matter of time until genome sequencing will be affordable for most people. Once it becomes commonplace, it will generate an enormous quantity of sequence data from a wide range of humans that could benefit biomedical research and drug development. More importantly, a 'thousand-dollar genome' could become an important tool to realize personalized medicine: perfectly tailoring diagnostics and treatments to a patient's genetic make-up.

> **… a 'thousand-dollar genome' could become an important tool to realize personalized medicine…**

Michael Egholm, Vice President of Research and Development for 454 Life Sciences, chuckled when asked whether



sequencing Watson's genome was just a publicity stunt or whether it had practical applications. "It's a fair question," he replied. "It's probably a little bit of everything. The serious answer is that 454 was founded in 2000 with a mission for routine human sequencing […] Some genome sample had to be the first. We had a long discussion at the time on whether we should have an anonymous sample from the HapMap Project, or a well-known individual," he said. "We decided on a well-known individual because we knew it would cause a fair amount of debate, all the consequences would make it much more personal as opposed to an anonymously donated sample." The sequencing itself did not take too long using 454's latest generation of sequencers, which started in earnest in January 2007. "[I]t was quite easy to sequence a human genome. In two one-month periods we generated 20 billion plus bases […] using a handful of machines," Egholm said.

Having their genome sequenced is not something the average person can afford at present. George Church, Professor of Genetics at Harvard Medical School and Director of the Center for Computational Genetics (Boston, MA, USA), and one of the proponents of the Human Genome Project, estimated that it would have cost Venter about US$30 million to sequence his genome. Heather Kowalski, Venter's spokeswoman, declined to comment on the exact costs and maintained that it was not the main consideration in any case. "[T]his research wasn't designed to beat any cost or time estimates, but rather was designed to be a very accurate new reference genome which, by many accounts, it appears to be," she said. "[T]here are many factors that go into the cost since we have a large facility that does a variety of projects at once."

Sequencing Watson's genome with 454's more advanced technology was less expensive. "We did it for well under US$1 million," Egholm said. "We're not disclosing the exact amount. What I can say is [that] by sometime next year the cost is going to go down significantly again."

> **… US$1,000 is a magic number that would attract customers and would put genome sequencing in the same financial bracket as many medical tests**

It seems that the race for the sequencing Holy Grail—the US$1,000 genome for the masses—is on. Egholm commented that 454 Life Sciences expects to be able to provide the US$100,000 genome next year, which would make it more suitable for research applications. "[Y]ou can envision a fairly large study; pharmaceutical companies taking 10 responders and 10 non-responders

and sequencing those and comparing them. You can envisage fairly large public projects," he said.

454, Foundation for Applied Molecular Evolution (Gainesville, FL, USA), Reveo Inc. (Elmsford, NY, USA), and VisiGen Biotechnologies (Foster City, CA, USA) have all announced that they are competing for the US$10 million Archon X Prize, sponsored by the X Prize Foundation (Santa Monica, CA, USA). The prize is promoting development of more efficient sequencing tools. The first team that is able to sequence 100 human genomes in 10 days for not more than US$10,000 per genome will win the prize. Egholm thinks it will take his group five years.

Church believes that even more valuable breakthroughs will occur much sooner, although he said that full-genome sequencing for US$100,000 would still be too expensive. Instead, he expects a scaled-down sequence for US$1,000 to become commercially available within a year. Although it would include only 1% of the genome, it would cover 90–95% of currently valuable information. "It may be diminishing returns going beyond one percent, but it is a vast improvement over 0.03 percent on current SNP chips," he said. Indeed, that is what the Personal Genome Project (PGP), which Church heads, is about: an affordable sequencing service that will cover all protein-encoding regions of the human genome (Church, 2005).

… the USA in particular could become fertile ground for genome sequencing because of its reliance on private health insurance and the growing popularity of do-it-yourself genetic tests

As Church commented, US$1,000 is a magic number that would attract customers and would put genome sequencing in the same financial bracket as many medical tests. "We're more interested in what you can get for US$1,000," he said. "It's more like the consumer computer model, where you didn't say, 'I don't want to buy a computer until it can beat everybody at chess and do flawless human speech-recognition and all that.' Back in 1977, you said instead, 'I'll take whatever you got for 1,000 bucks.' […] That's the way it's going to be with the genome, too. Rather than wait for the 100-percent genome, people are going to be lining up to get their US$1,000 one-percent genome."

From a marketing perspective, the USA in particular could become fertile ground for genome sequencing because of its reliance on private health insurance and the growing popularity of do-it-yourself genetic tests. "If you talk to people in the United States, you find that maybe they're early adopters in general, and they have a more consumerist approach to healthcare and a greater tendency in general to order screening tests. If you take a simple example of tests like PSA (prostate-specific antigen), I think it's far, far more commonly used in the United States than it is in the United Kingdom," said Stuart Hogarth, a research associate in the Epidemiology for Policy group at the Department of Public Health and Primary Care at Cambridge University, UK.

However, even if a US$1,000 genome sequencing service could be made available, it is not yet clear whether and to what extent it could change both biomedical research and medical practice. Kári Stefánsson, Chairman, CEO and co-founder of deCODE Genetics (Reykjavik, Iceland)—a biopharmaceutical company that uses its genetic, medical and genealogical databases of the Icelandic population to identify disease-related genes for drug and diagnostics development—expects that the US$1,000 whole genome will become available within five years, but he questions how efficiently this information can be mined. "The difference between what you can get out of whole-genome sequencing and genotyping at a very high density is not entirely clear. But it's very likely that the added value of having the entire genome sequenced when it comes to gene hunting is substantial," he said. "The problem I see with this at the very moment is in the computational power needed to pay full attention to all of the data, the lack of data-mining algorithms to take full advantage of the added amount of data. But […] this is a just a short-term problem." In any case, Stefánsson added that he had neither looked at the full personal genomes online, nor would he consider having himself sequenced and posting the results. "I have much better use of my time than to glare at Dr. Watson's genome," he said. "I haven't considered it all. […] Why should I have my genome out there?"

In any case, the next-generation sequencing technology is already changing various research fields. Church commented that the new polymerase colony technology, developed by 454 and Harvard University, will aid various experiments: quantitative biology, association studies, and metagenomics, RNA and chromatin studies; these are already transitioning from DNA chips to sequence-tag counting because the new technology is highly accurate. In the case of association studies, Church also expects that extensive sequencing will replace single-nucleotide polymorphism technology because the latter has difficulties distinguishing between somatic mutations, such as in cancer, individually rare but collectively common mutations and new mutations. Finally, even classical infectious disease research would benefit from applying the new sequencing technologies to metagenomics in order to understand how pathogenic and commensal bacteria vary from person to person and day to day. "This will help define a bio-weather map and make us more aware at a fine-grained geography—down to the resolution of individuals—as infections spread," Church said.

While a huge drop in the costs of sequencing might, in any case, benefit basic research, Hogarth wonders about the medical value of whole-genome sequencing. "Even if we have medically actionable information, do we really need the full genome scan when we could just have individual tests? That's an economic question," he commented. "But the clinical question is, do we have well-validated gene–disease associations for common complex diseases and could that become a clinically useful test? So it's not just gene–disease association or bust, but can we do anything with the information?" He added that companies such as IntegraGen (Evry, France), deCODE and others have already launched, or are developing, diagnostic tests that are based on selected parts of the genome.

Amy McGuire, Assistant Professor of Medicine at the Center for Medical Ethics and Health Policy at Baylor College of Medicine, who discussed the implications with Watson of publicizing his genome sequence with Watson, shares Hogarth's doubts. "If you're doing a specific genetic test to try to determine if you have a *BRCA1* mutation, it is very specific and you're looking for one particular thing, the counseling

is pretty standard," she said. "The problem with the whole genome—or even less than the whole genome, a larger region of genes—is oftentimes you get information that we really don't know what its clinical significance is. There is a lot of noise that probably means nothing."

> **"Despite all the hype about new genetic knowledge, some of it is going to fall flat because the recommendations are going to involve simple lifestyle changes you could do anyway."**

Even if it were possible to extract meaningful information from a genomic sequence, it would not necessarily have much medical utility, McGuire added. "Usually, what it's telling us, if anything, is some very small increase in your risk of developing some complex disorder," she said. "So even [in] the disease association studies that they're doing now—where they're finding associations between different genetic variations and complex diseases—the degree to which it increases your risk of developing the disease goes from one percent to like 1.6 percent."

Paul Burton, Professor of Genetic Epidemiology at the University of Leicester, UK, and a Principal Regional Investor with the UK Biobank, questions the medical value of such weak associations. "I wouldn't want any doctor treating me on the basis of a relative risk of 1.2. There would be far too much uncertainty," he said. "Basically, we're in a situation where almost all these effects that are in common genes are almost certainly going to be small, [which] makes them hard to use as predictors […] There are definitely going to be some conditions which do have big relative risks associated with alleles […]. But we know so little about them at the moment that […] if someone gave me my genome for free, I wouldn't know what to do with it." He therefore forecasts that full-genome sequencing will not become a common medical tool for 40–50 years.

Similarly, Arthur Caplan, Professor of Bioethics and Director of the University of Pennsylvania's Center for Bioethics (Philadelphia, PA, USA), thinks that most genetic information will not necessarily have much value even for preventive medicine. "Despite all the hype about new genetic knowledge, some of it is going to fall flat because the recommendations are going to involve simple lifestyle changes you could do anyway," he said. "I'm a little skeptical that you're going to need to go out and spend money on testing, or get personalized genetic profiling to tell you to wear your seatbelt, so to speak."

However, as Church pointed out, people's genome sequences could potentially reveal a lot of useful medical information, such as the sequences of the *CYP2C9*, *BRCA1/2* and *ApoE* genes, which relate to drug metabolism, the risk of developing breast cancer and the risk of Alzheimer disease, respectively. He also thinks that knowledge of even small relative risks would be helpful. "[I]f given a [relative return] of 1.2 on an investment—or a cancer treatment—most folks will gladly take that 20% edge." Furthermore, running a range of diagnostic tests on someone's whole genome might be more cost-efficient than current practices. "There are many tests already available that genetic counselors and physicians trained in genetics consider reliable and which sum up to much more than $1K, so getting them all at once is cost-effective and permits easy updates without fresh blood draws."

Despite the recent advances in sequencing technology, as long as it is not possible to accurately and cheaply sequence whole human genomes, it will remain unclear how useful a whole-genome sequence will be to an individual. Beyond the media attention given to publishing Watson and Venter's genomes, and the ethical arguments and debates about their utility, it seems likely that once a US$1,000—or a €1,000—test is achieved, new services will become available. As Caplan pointed out, it will just take time because the science and its utility are still in a transitional period.

REFERENCES
Check E (2007) Celebrity genomes alarm researchers. *Nature* **447:** 358–359
Church GM (2005) The personal genome project. *Mol Syst Biol* **1:** 2005.0030. doi:10.1038/msb4100040
Levy S *et al* (2007) The diploid genome sequence of an individual human. *PLoS Biol* **5:** e254

Howard Wolinsky
doi:10.1038/sj.embor.7401070