THE SEQUENCING MACHINE

Faeces, lizards, keyboards, faces — Rob Knight likes to sequence the microbes on anything and everything. Next, he plans to sequence Earth.

BY VIRGINIA GEWIN

Rob Knight wants the spit of a komodo dragon. But he is unsure whether Bintang, a metre-long juvenile of this endangered lizard species, will oblige. Wielding a white cotton swab, Knight cautiously approaches the creature, which is squirming in its keeper's arms at the Denver Zoo in Colorado. With an inquisitive flick of the tongue, Bintang deposits a dab of saliva on the swab.

Knight pops the swab into a sterile plastic tube, opens another tube and goes on to collect samples from the lizard's head and belly. He also swabs the enclosure, which is part of the world's most successful facility for the captive breeding of komodo dragons (*Varanus komodoensis*). The samples are teeming with the bacteria and viruses that live in the reptile's mouth, gut and skin. Back in his lab at the University of Colorado Boulder, Knight will sequence the DNA of these microorganisms — and eventually compare the microbes with those found on komodo dragons in the wild and at other facilities, to try to find out whether and how they affect the animals' survival and why captive females tend to die young.

If there is a link between this species' microbes and its health, Knight is the one to find it. He is a leader in the burgeoning field of microbiome research, which aims to sequence the mass of genes from microbial communities and use computational tools to count and compare species. Knight has helped to reveal differences between the gut microbes in obese and lean people[1]; to show that people's intestinal microbes differ dramatically depending on where in the world they live[2]; and to document the wide differences between the microbes acquired by babies born by Caesarean section and those delivered vaginally[3]. Outside the human body, Knight has probed the microbes that blanket various natural and man-made environments, from freshly fallen snow to computer keyboards and bathroom floors. He does all this at a restless, relentless pace; he co-authored 49 publications in 2011 alone.

Knight is sensitive to the charge that all this is an exercise in microbial surveying, rather than in hard hypothesis-testing. "We don't take on projects if the scientific value isn't clear," he says. "What motivates me, from a pragmatic standpoint, is how understanding the microbial world might help us improve human and environmental health." The microbiota hold clues, he says, to solving major societal problems — preserving endangered species, for example, or treating obesity or malnutrition.

Since 2010, Knight has been involved with the most ambitious effort yet to probe the microbiota: the Earth Microbiome Project, a collaborative effort to sequence and characterize the microbial communities in at least 200,000 environmental samples such as soils and water collected from around the world. He says that the project, which is led by Jack Gilbert, an environmental microbiologist at the US Department of Energy's Argonne National Laboratory in Illinois, will yield a master list of the proteins required to sustain microbial life across the planet; about 500,000 reconstructed microbial genomes;

and the makings of a global-scale model of microbial metabolism.

"Knight and Gilbert literally talk about sampling the entire planet. It is ludicrous and not feasible — yet they are doing it," says Jonathan Eisen, an evolutionary biologist at the University of California, Davis. "Rob is one of the few people who are surfing the exponential curve of increased sequencing power," Eisen adds, "and planning where the wave will take him next."

KEVIN MOLONEY

Rob Knight swabs the microbes from a komodo dragon at Denver Zoo, Colorado.

EVERYTHING OF INTEREST

A tall, lanky 35-year-old, Knight doesn't look the part of a tenured professor. He speaks so rapidly that he can be difficult to follow, and his ideas erupt at a similar rate. The conversation can swerve from using microbes to estimate time of death to his hopes of one day swabbing the International Space Station to reveal its hidden microbial life. "I've never been interested in just one thing," he says.

As a youth growing up in New Zealand, Knight had interests ranging from fossils to chemistry and computers. "I didn't really have this idea that science should be compartmentalized into biology or physics," he says. He was soon dabbling in both at Princeton University in New Jersey, as he worked on a PhD project analysing the evolution of the genetic code. He went on a self-taught computer-programming binge, logging 12–20 hours of screen time per day, writing code, testing programs and working out bugs.

Cathy Lozupone, a technician in the Princeton lab at the time, recalls one night in 2000 when Knight, who was in between apartments, was briefly staying with her. She woke at 2 a.m. to find him trembling with excitement after solving a thorny computational problem. The program he had devised crunched through protein-coding sequences from 600 species spanning all domains of life to show that very simple rules involving mutation and selection could explain a major puzzle about DNA: why different organisms prefer different coding sequences for the same amino acid[4].

Knight moved to Boulder as a postdoc in 2001, and went on to run his own group. He began attending lab meetings with famed microbiologist Norm Pace, whose group was studying microbial diversity by sequencing 16S ribosomal RNA, a stretch of 1,500 nucleotides used for studying evolutionary relationships. Knight saw that mass sequencing of microbes was taking off, but that it was enormously challenging to make sense of so much data. He developed a software tool dubbed UniFrac (unique fraction)[5], which provides a measure of the difference between two microbial communities by constructing evolutionary trees from the sequences in each sample and calculating the fraction of branches that are unique to each.

UniFrac received a high-profile test when Knight teamed up with Jeff Gordon, a leader in microbiome research at Washington University in St Louis, Missouri, and his postdoc Ruth Ley. The researchers used UniFrac to show that microbial populations in the guts of obese mice differ from those in lean ones[6]. That paper has now been cited more than 500 times, and the program has become a standard analytical tool in the field. "Before UniFrac we had simply been tallying species, but understanding whether the species were related or not gave us deeper insights into the biology of communities," says Ley, now a microbiologist at Cornell University in Ithaca, New York.

> "KNIGHT AND GILBERT LITERALLY TALK ABOUT SAMPLING THE ENTIRE PLANET. IT IS LUDICROUS AND NOT FEASIBLE — YET THEY ARE DOING IT"

UP CLOSE AND PERSONAL

Fascinated by the gut's inhabitants, Knight was soon helping to chart a course for the Human Microbiome Project (HMP) — a US$115-million effort funded by the US National Institutes of Health to sequence the tens of trillions of microbes that live in and on humans — which published its major results in June[7].

He also plunged into data: Knight and his colleagues showed that the human body's bacteria vary dramatically between the gut, skin, mouth, ears and other locations[8]. And by exploring the faeces of mammals ranging from kangaroos to lemurs and zebras, Knight's team showed that microbiota adapt to diet — be it carnivorous, herbivorous or omnivorous — in a similar way across all mammalian lineages[9]. With a rapid succession of high-profile papers, "Rob exploded onto the scene — he went from no profile to being everywhere", says Phil Hugenholtz, a microbial ecologist at the Australian Centre for Ecogenomics in Brisbane.

Knight has continued to develop software for handling the increasingly giant data sets generated by modern sequencing technology. One package, called QIIME (quantitative insights into microbial ecology)[10] uses tools developed by Knight's group and others to simultaneously compare millions of raw sequences, assign species names when possible, build phylogenetic trees and visualize the data in many different ways.

During the HMP, Knight was vocal about his concerns that the tools and data weren't

being shared effectively with the wider community. In one case, he "challenged in a very public way" a rule that was slowing down publication, says Owen White, associate director of the Institute for Genome Sciences in Baltimore, Maryland, who oversees the central data repository for the HMP. The rule was later overturned. White says that Knight was also instrumental in shifting the centre of intellectual power in genomic analysis from larger sequencing centres to smaller labs. Knight says that he wants to "democratize" the data and tools. "To maximize the output of science, it makes more sense to do things that are good for the field overall," he says.

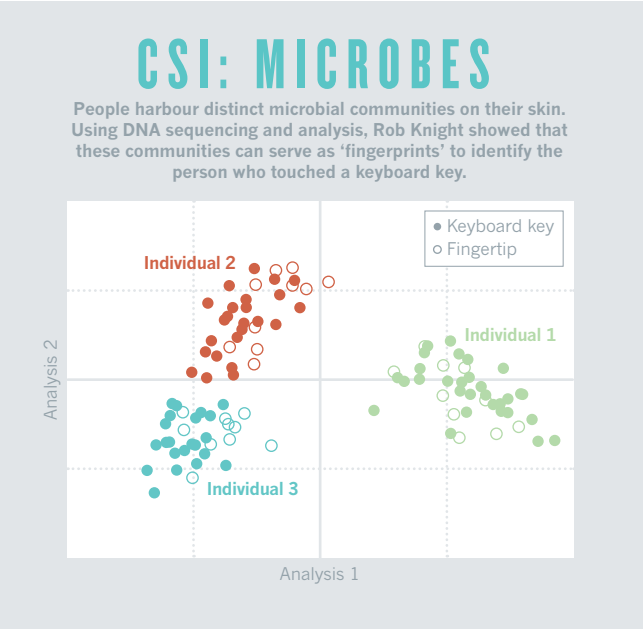
Another of his innovations lies in working out how to link sequence data to the sample's metadata — environmental details such as pH, temperature, salinity and collection time — so that the maximum amount of information can be stored and extracted. "He strongly encouraged people to get organized and put their data in a standard format so that studies could be compared," says Ley.

Knight has ruffled some feathers — perhaps owing to his meteoric rise at an envy-inducing young age, or his unflinching criticism when he thinks something is being done in the wrong way. "He's not always terribly diplomatic," says Ley. Some complain of "superficial, catchy" studies, says one colleague who preferred not to be named. Knight's work certainly can make a splash: in 2010, the popular detective television programme *CSI: Miami* featured a technique developed by Knight and Noah Fierer, a microbial ecologist at Boulder, with which they matched individuals to the unique bacterial fingerprints that they left on computer keyboards[11] (see 'CSI: Microbes'). Knight's team also attracted attention with as-yet unpublished work mapping the microbes that carpet the human face from forehead to lips, and a study documenting how the gut microbiota of a Burmese python (*Python molurus*) shifted from famine to feast mode as the snake digested a rat over three days[12].

SERIOUS BUSINESS

Knight doesn't deny that he is one of two "healthy subjects" involved in generating the largest time series of human microbiota yet collected[13]: he swabbed his palms, mouth and faeces every day for 15 months for the cause. And he relishes the story of an airport official who briefly confiscated the thermos of dry ice he was carrying to transport the samples. Six hours later, in Knight's office, the thermos exploded because the official had cinched down the lid too tightly.

Knight says that all his studies have a serious intent. The time-series data, for example, showed that there was vastly more variation within one person's microbiota than expected. He hopes that the finding will convince clinicians that monitoring people's microbiomes over time could be useful in a clinical trial, allowing investigators to, for example, test whether variations in microbiota correlate with a drug response. "It's the kind of thing where you wouldn't want to ask someone else to go through the trouble if you haven't demonstrated the feasibility yourself," he says.



Knight says he currently has about 70 papers in draft form, involving roughly 50 principal investigators. But that looks set to increase thanks to the Earth Microbiome Project. Knight, Gilbert and Janet Jansson, a microbial ecologist at the Lawrence Berkeley National Laboratory in Berkeley, California, are encouraging microbiologists worldwide to send in samples. They have already received some 60,000, ranging from deep-sea sediments from the Pacific Ocean to owl nests from Alaska. The researchers, funded so far on about $3 million from the US Department of Energy, as well as scraps from private sponsors and their own and their contributors' funds, have sequenced around 15,000 of them. They are making all the data and tools — many of which Knight developed or is developing — openly available.

"The Earth Microbiome Project is not a purely exploratory project," Knight emphasizes. The plan, he says, is to collect a number of hypothesis-driven data sets, such as samples collected in pristine and disturbed areas of the Australia coastline to determine whether the disease-driven decline of algae has knock-on effects for other organisms. By combining data sets, the team should also be able to test broader hypotheses — such as whether the dominant microbes in a sample always have the most important functions.

To that end, Knight continues to collect samples whenever he has the chance; later this month, for example, he will be scraping up microbial 'mats', which are among the most diverse communities known, from hypersaline waters off the Californian coast. The komodo-dragon samples are going into the project, and Knight now has approval to sample these and other species at a further three zoos.

Knight has other ventures on the boil. In one project, he is exploring whether gut microbial communities can influence mental health by, for example, changing the signalling patterns between the gut and brain. "My family has schizophrenia on one side and bipolar disorder on the other," he says, "which has led to my interest in trying to figure out whether we could potentially address those conditions." With Gordon and Lozupone, he is also part of a global network for the study of malnutrition and intestinal diseases, funded by the Bill & Melinda Gates Foundation in Seattle, Washington. Working in Bangladesh, the team hopes to identify the microbes associated with malnutrition, and to explore potential probiotic or other microbial treatments.

Knight struggles to explain how he sustains his eclectic interests and level of intensity. With so much to sequence, over-commitment is an occupational hazard for everyone in microbiome research, he says.

But Gilbert says that the research community faces a challenge in keeping up with Knight. "Unless he burns out in the next 20 years," says Gilbert, "I think we'll look back at him as a pioneer." ■

Virginia Gewin *is a freelance writer based in Portland, Oregon.*

1. Turnbaugh, P. J. *et al. Nature* **457,** 480–484 (2009).
2. Yatsunenko, T. *et al. Nature* http://dx.doi.org/10.1038/nature11053 (2012).
3. Dominguez-Bello, M. G. *et al. Proc. Natl Acad. Sci. USA* **107,** 11971–11975 (2010).
4. Knight, R. D., Freeland, S. J. & Landweber, L. F. *Genome Biol.* **2,** research0010.1–research0010.13 (2001).
5. Lozupone, C. & Knight. R. *Appl. Environ. Microbiol.* **71,** 8228–8232 (2005).
6. Ley, R. E. *et al. Proc. Natl Acad. Sci. USA* **102,** 11070–11075 (2005).
7. The Human Microbiome Project Consortium *Nature* **486,** 207–214 (2012).
8. Costello, E. K. *et al. Science* **326,** 1694–1697 (2009).
9. Muegge, B. D. *et al. Science* **332,** 970–974 (2011).
10. Caporase, J. G. *et al. Nature Meth.* **7,** 335–336 (2010).
11. Fierer, N. *et al. Proc. Natl Acad. Sci. USA* **107,** 6477–6481 (2010).
12. Costello, E. K., Gordon, J. I., Secor, S. M. & Knight, R. *Isme J.* **4,** 1375–1385 (2010).
13. Caporaso, J. G. *et al. Genome Biol.* **12,** R50 (2011).