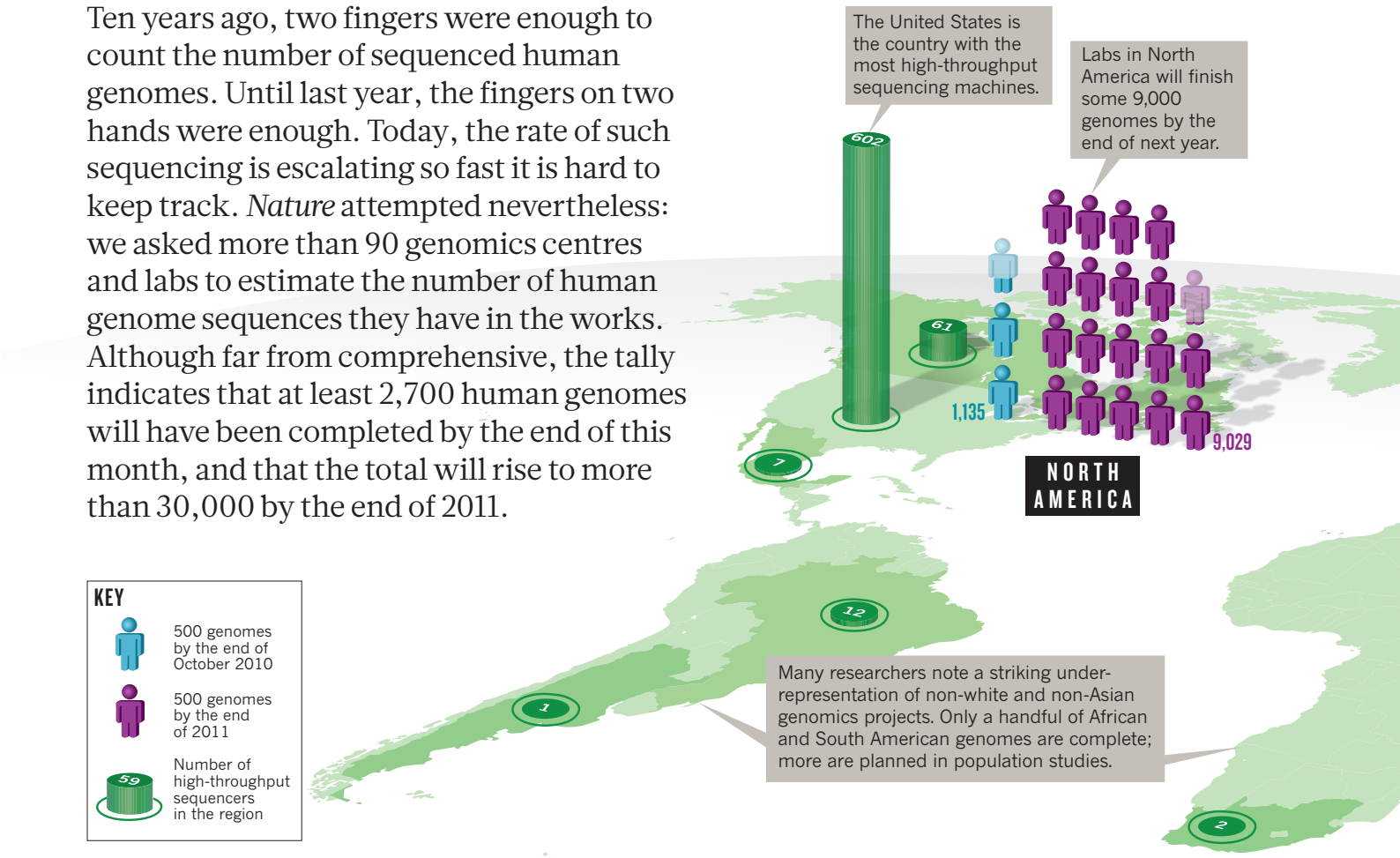
Genomes by the thousand

Ten years ago, two fingers were enough to count the number of sequenced human genomes. Until last year, the fingers on two hands were enough. Today, the rate of such sequencing is escalating so fast it is hard to keep track. *Nature* attempted nevertheless: we asked more than 90 genomics centres and labs to estimate the number of human genome sequences they have in the works. Although far from comprehensive, the tally indicates that at least 2,700 human genomes will have been completed by the end of this month, and that the total will rise to more than 30,000 by the end of 2011.

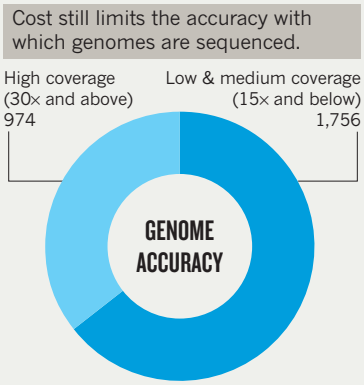


Why scientists want tens of thousands of genomes — and more

To understand populations

Comparing lots of genomes lets researchers identify points at which one genome differs from the next. Costs may be falling, but sequencing and data analysis are still pricey. So most researchers face a trade-off between the number of subjects and the accuracy in the sequences they can afford. For projects examining how populations commonly differ, sequencing a large number of individuals at relatively low accuracy or 'depth of coverage' is enough. About 900 genomes sequenced so far by the 1000 Genomes Project have been read three times on average.

Cost still limits the accuracy with which genomes are sequenced.



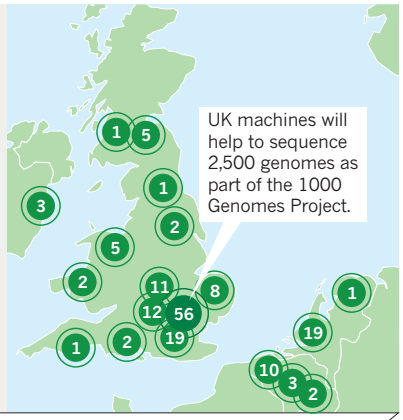
To understand disease

Researchers trying to uncover rare disease-linked mutations — perhaps limited to just one family or an individual — need precision, typically sequencing each genome 30 times on average. Cancer genomes, many sequenced under the auspices of large collaborations, account for a sizeable chunk of high-coverage genome sequences completed to date. Projects scrutinizing people with diabetes, Crohn's disease and other disorders are starting to emerge. Analysing all the genome data is a huge challenge, as is turning genetic discoveries into clinical benefits.

GRAPHICS BY NIK SPENCER; DATA COMPILED BY ALLA KATSNELSON

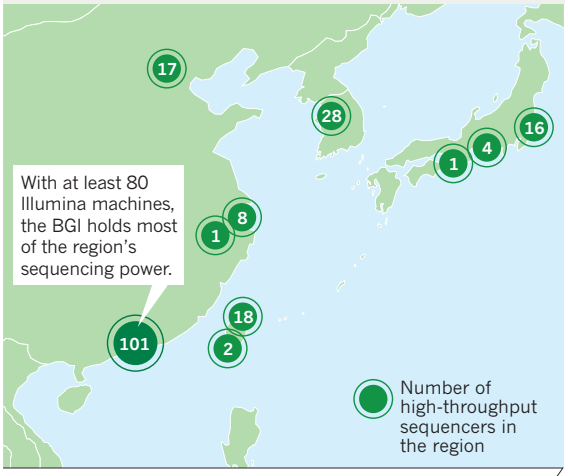
POWER TO THE PEOPLE

The latest sequencing technology is no longer concentrated at a few major centres. In Britain, the Wellcome Trust Sanger Institute in Hinxton houses 38 of the country's high-throughput sequencers, and the rest are scattered over an additional 32 sites. Falling costs mean that a human genome is within the reach of individual labs.

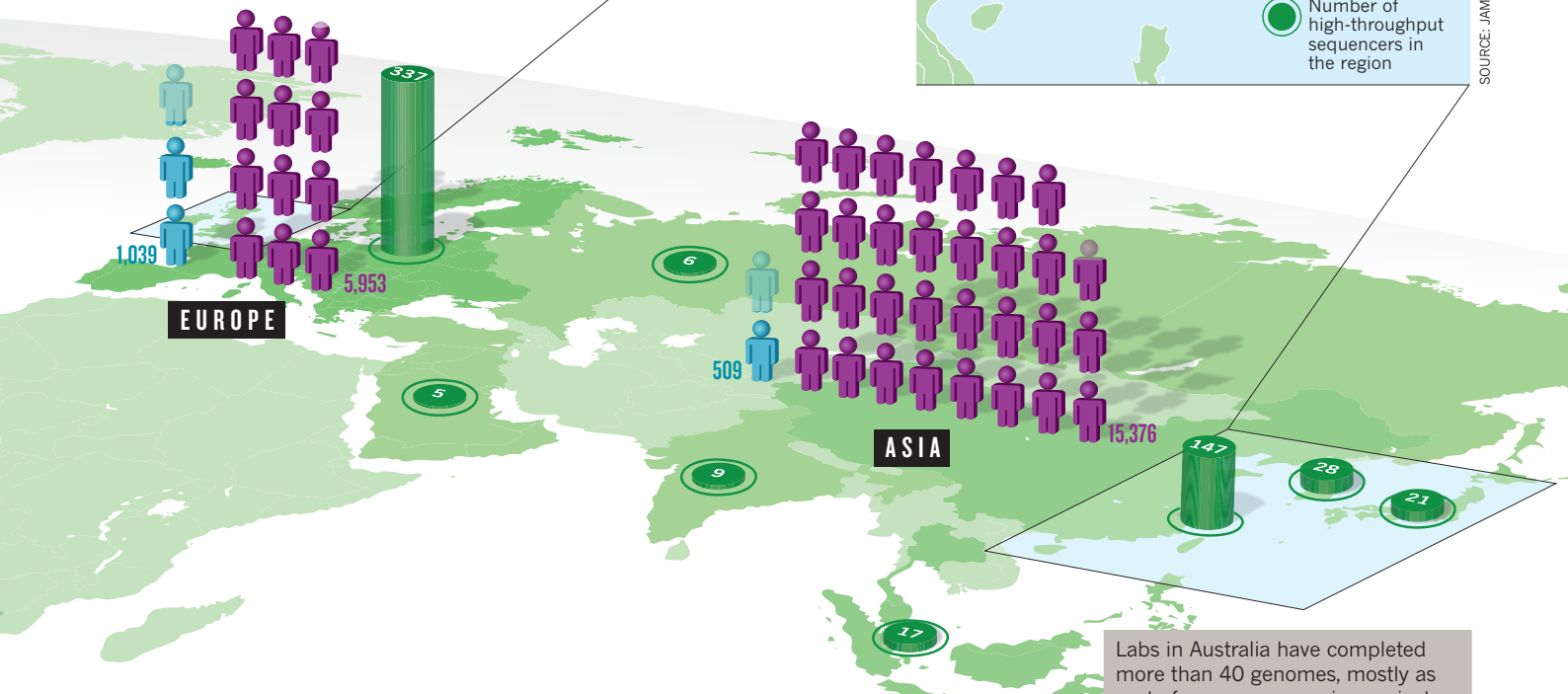


THE RISE OF GENOME FACTORIES

Individual labs may still find it cheaper and easier to outsource a human genome to a power-house 'sequencing service provider'. The BGI in Shenzen, which has global expansion plans, predicts that its machines will have completed some 10,000 to 20,000 human genomes by the end of 2011.

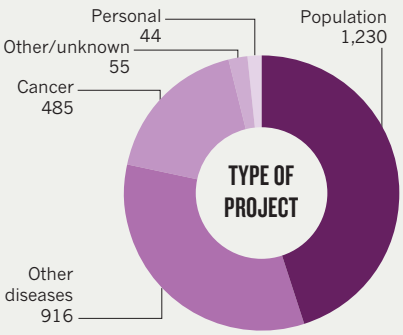


SOURCE: JAMES HADFIELD & NICK LOMAN



Labs in Australia have completed more than 40 genomes, mostly as part of cancer sequencing projects. They plan to finish well over 100 genomes by the end of next year.

Disease-specific projects make up more than half of the complete genomes.



To understand individuals

The rate at which human genomes are being sequenced — at least in mega-projects — will probably slow once researchers have extracted most of the common variation shared by populations and diseases. But individuals are genetically unique. If the cost of a genome sequence becomes trivial and the benefits of knowing one increase (through gene-tailored medicine), then personal genome sequencing will continue to push the genome count up and up.

NATURE.COM
For more on human genomes see:
www.nature.com/humangenome

METHODS: Our survey focused on large, academic projects rather than individual labs; we included complete genome sequences, both high- and low-coverage, and excluded partial (exome) sequences. The list excludes all biotechnology and pharmaceutical companies and most sequencing service providers, which do not disclose their work. The sequencer locations, based on a user-generated map at go.nature.com/b74acy, includes some 60–70% of all machines.