biotech and life sciences worldwide as part of a larger project to map human technological development (Fig 1). "Maps can help answer or at least address a number of key questions," said Hoffman. "Where around the world is biotech really happening or beginning to happen? How important is economic geography now that the life sciences are spilling out of university, governmental and industrial laboratories into regional economies? Where is the search for 'biotech gold' breaking down national boundaries and political jurisdictions, with different regions and different countries banding together to compete in the global economy?"

Indeed, the biotech-clustering craze has already created another level of complexity: meta-clusters. These are transnational networks of clusters operating in geographically connected countries. Europe, in particular, is a fertile place for such initiatives. The latest is EuroBioCluster South, which was announced at the 2005 BioVision conference held in April 2005 in Lyon, France. The project will combine bioclusters from an area stretching from Heidelberg, Germany, to Barcelona, Spain, "to stimulate a supra-regional dynamic of scientific and technological growth and international outreach," according to an official press release (Grand Lyon, 2005). The EuroBioCluster's potential is expected to exceed that of Seattle, San Diego and San Francisco combined, and it will have "the largest concentration of life-science enterprises in the world, with the presence of both major groups and start-ups, research centres, universities and institutions."

ScanBalt is another main cross-border effort in biotech and life-sciences development, encompassing 11 countries and 85 million people in Northern Europe with 60 universities and 870 biotech-related companies. "We do not consider ScanBalt to be a meta-cluster but a meta-region, which is an important difference," clarified Peter Frank, General Secretary of ScanBalt. "Clusters have a certain dynamic connected to the fact that activities within a cluster are characterized by geographical proximity and a regional background, while a meta-region like ScanBalt BioRegion has a different and complementary dynamic." A basic role of ScanBalt is to coordinate between regional

and national networks, promote the establishment of clusters and increase transparency and visibility of competencies in the region, explained Frank.

Although the verdict is still out on the effectiveness of science parks, incubators and mega-clusters in boosting local development, they will be in fashion for some time. "Businesses—from high-tech companies to developers—that get money from these schemes are happy with their windfalls. Politicians are happy to hand out pork while looking like visionaries," wrote Wallsten (2004b). However, "the obsession with becoming the next biotech hub will fade in time, just as dotcom dreams did." But if a good-practice model prevails for the establishment of research parks and clusters, if it is based on a thorough evaluation of local dynamics, research potential and entrepreneurial environment, and if strict policies are enforced to select only technologically sound tenants and spin-offs, then there is space for reasonable optimism that these investments can spur high-tech and biotech progress, at least in some cases.

REFERENCES

- Critical I (2005) *Biotechnology in Europe: 2005 Comparative Study*. Lyon, France: BioVision
Cyranski D (2005) Malaysian biotechnology: the valley of ghosts. *Nature* **436**: 620–621
DeVol R, Wong P, Ki J, Bedroussian A, Koepf R (2004) *America's Biotech and Life Science*

Clusters: San Diego's Position and Economic Contributions. Santa Monica, CA, USA: Milken Institute

- Grand Lyon (2005) EuroBioCluster. Press release, 14 Apr. www.lyon-business.org
House of Commons Trade and Industry Committee (2003) *UK Biotechnology Industry*. Twelfth Report of Session 2002–2003. London, UK: House of Commons
Imperial College (2005) Spin-out News, 30 Nov. www.imperialinnovations.co.uk
Jia H, Jayaraman KS, Orellana C (2003) Biotech parks proliferate, despite concerns over sustainability. *Nat Biotechnol* **21**: 1259–1260
Ketels CHM (2003) *The Development of the Cluster Concept—Present Experiences and Further Developments*. Düsseldorf, Germany: Ministry of Economics, Northrhine-Westphalia
Porter ME (1990) *The Competitive Advantage of Nations*. New York, NY, USA: Free Press
Sansom C (2004) EU expansion: enlarged horizons or false dawn? *Nat Biotechnol* **22**: 501–504
Stone R (2005) Central Asia. Visions of a biotech empire on the Kazakh Steppe. *Science* **308**: 1861
The Scripps Research Institute (2005) Scripps Research breaks ground for construction of Scripps Florida campus. *The Scripps Research Institute News & Views*, 26 Sep www.scripps.edu/newsandviews/
Wallsten S (2004a) Do science parks generate regional economic growth? An empirical analysis of their effects on job growth and venture capital. Working Paper 04–04. Washington, DC, USA: AEI-Brookings Joint Center for Regulatory Studies
Wallsten S (2004b) High-tech cluster bombs. *Nature* **428**: 121–122

Andrea Rinaldi

doi:10.1038/sj.embor.7400633

Ancient DNA research goes nuclear

A new technique to extract sequence data from nuclear DNA may reveal exciting new insights into evolution and phylogeny

It is one of the most fascinating questions we face: how did *Homo sapiens*—modern humans—evolve? When did they start using tools, how did they develop language and why did *Homo erectus* and then *Homo sapiens* thrive while other human species, such as the Neanderthals, became extinct? The availability of advanced genetic technologies, most notably the polymerase chain reaction (PCR), originally held great promise to

answer these questions; in theory, sequencing the DNA from fossils could paint a picture of the molecular evolution not only of humans but of other species as well. In practice most of these expectations have not been realized, as the analysis of ancient DNA is anything but easy. The very small amount of DNA in fossil samples, the decay of the molecules over time and contamination with DNA from other organisms have proven to be considerable hurdles. As a

... the biotech-clustering craze has already created another level of complexity: meta-clusters

result, most knowledge about molecular evolution comes from the analysis of mitochondrial and plastid DNA, simply because it is more abundant and easier to analyse.

But the analysis of ancient DNA is about to enter a new era. Two recently developed techniques—multiplex PCR and a new genomic sequencing technology—allow the recovery of meaningful sequence data from nuclear rather than just mitochondrial or chloroplast DNA. Many of the field's leading researchers are already excited about the prospects. "I think we will see a number of studies dealing with functional genetics, where you take and amplify sequences, and see phenotypic features that say something about how the animal behaved or looked like," said Eske Willerslev, from the Department of Evolutionary Biology at the University of Copenhagen, Denmark. "We will definitely see much more in the next two years because there's a huge potential out there."

... in theory, sequencing the DNA from fossils could paint a picture of the molecular evolution not only of humans but of other species as well

Although it may never be possible to recreate extinct organisms from their DNA, the new multiplexing technique could enable scientists to expand their study of phenotypes by investigating properties such as skin colour or behavioural traits. Among the many important questions that could be answered is whether the Neanderthals—the last human species to become extinct—were able to speak and, if so, how well. "To do this, various groups hope to target the *FOXP2* gene, which is believed to confer the ability to speak in humans," said Tom Gilbert, an ancient DNA and evolution researcher at the Niels Bohr Institute, University of Copenhagen, Denmark. This could help to resolve the debate over why the Neanderthals became extinct around 30,000 years ago, as one theory claims that modern humans gained the upper hand linguistically.

Because DNA degrades with time even when preserved under ideal conditions, sequence amplification is compromised by deletions or substitutions that either cause the process to fail or result in mistakes that can lead to false

Phenotype reconstruction based on mammoth nuDNA (possibly contaminated)

conclusions. The probability of such errors is proportional to the length of the strand; although longer sequences contain more information, in practice researchers have had to make do with overlapping shorter ones, and then painstakingly piece them together. More importantly, the process is usually constrained by the small amount of template DNA in the sample. For this reason, no DNA sequences longer than about 1,000 base pairs (bp) have been recovered, even from widely studied Pleistocene mammalian species such as mammoths, ground sloths and cave bears.

The breakthrough in ancient DNA sequencing came in 2005, when Michael Hofreiter from the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany, and colleagues described in *Nature* their new multiplexing technique for reconstructing a longer DNA sequence from several small molecules (Krause *et al*, 2005). They demonstrated this by reconstructing the entire mitochondrial genome of the Pleistocene woolly mammoth, *Mammuthus primigenius*, comprising 16,770 bp, from about 200 mg of bone.

In essence, multiplexing is a two-stage PCR. It uses multiple primer pairs in one PCR reaction to target subsequences within

the complete DNA sample. In the case of the woolly mammoth mitochondrial DNA (mtDNA), 46 such primer pairs were chosen that marked overlapping DNA sequence fragments and spanned the entire mtDNA genome. To cut out the overlaps and generate coherent products that could then be amplified in the second PCR stage, the researchers divided the primer pairs into two sets, each comprising alternate pairs. Each of these two sets was amplified in a multiplex PCR, requiring only as much ancient DNA template as would be used normally for a short target sequence.

Among the many important questions that could be answered is whether the Neanderthals... were able to speak and, if so, how well

Having obtained the two amplified sequences, which together spanned the whole mtDNA genome, the samples were divided into 46 parts and used as templates for a secondary PCR to amplify each of the 46 products separately. Although multiplexing does not save time, it allows far greater scaling up from small samples,

Hofreiter pointed out. "You still have to put in a lot of work, but you are not limited by sample amount, because you can now do 10,000 primer pairs in one tube," he commented. "This is a real breakthrough in ancient DNA research," commented Carles Lalueza-Fox, Assistant Professor of Anthropology at the University of Barcelona, Spain.

Only two days after the *Nature* publication, *Science* released a report on the large-scale sequencing of nuclear and mitochondrial DNA from the jawbone of a 27,000-year-old Siberian mammoth (Poinar *et al.*, 2005). The team, headed by Hendrik Poinar, a molecular evolutionary geneticist from McMaster University (Hamilton, ON, Canada) and genome researcher Stephan Schuster from Penn State University (University Park, PA, USA), used a new genome sequencer developed by Stanford University and 454 Life Sciences (Branford, CT, USA). The technology circumvents the need to clone DNA samples before sequencing. Instead, the DNA is broken into small fragments and encapsulated in a lipid bubble, thereby allowing their multiplication in isolation. Using this new tool, the researchers were able to sequence 28 million bp from the ancient fossil, 13 million of which were from the mammoth itself. "The ability to obtain this level of genetic information from extinct species makes it possible to consider detailed analysis of functional genes, and fine scale refinement of mutation rates," the authors commented.

Until now, scientists have usually relied on mtDNA to construct phylogenetic trees. As any sample will typically have 1,000 times more mtDNA than nuclear DNA (nuDNA)—because each animal cell has many mitochondria—this makes it easier to derive meaningful sequence information. However, although nuDNA is less abundant, it seems less prone to degradation and damage over time, so the chance of recovering longer intact strands may actually be better. "DNA damage is lower in nuclear DNA than in mitochondrial DNA, maybe because nuclear DNA is better protected by proteins," said Lalueza-Fox. Combined with the technique of multiplexing, this allows the retrieval of longer sequences.

To some extent, this will put ancient DNA research onto the same footing as

contemporary DNA work. "You can take samples of modern-day DNA, infer back to how that might have evolved, and then test that hypothesis very directly by looking at ancient DNA," explained Ian Barnes, an evolutionary biologist at University College London, UK. Ideally, genes would be identified from modern DNA and then tracked back through time to follow their evolution. But making this feasible on a larger scale requires a library of ancient DNA to facilitate efficient searching and tracking. "Then we can try to bring these genomic studies to a very high scale," Lalueza-Fox said. There are two approaches to building such a library: sequencing whole samples, or systematically selecting genes of interest and amplifying only those. "With that specific approach, if you find something interesting in the genomic library, you can check whether it is real," said Lalueza-Fox.

In some cases, scientists have managed to retrieve nuDNA, usually from relatively 'new' samples around 5,000 years old (Shepherd *et al.*, 2005), but nearly all sequence data obtained so far from fossils has been mitochondrial. However, this has enabled interesting deductions about the phylogenetic positions of some extinct mammals. For example, the giant elk deer (*Megaloceros giganteus*), which stood 2 m tall with antlers spanning 3.5 m and became extinct 8,000 years ago, has as its closest living relative the much smaller fallow deer that roams in similar parts of Europe, according to studies of ancient DNA taken from skeletal remains (Lister *et al.*, 2005; Kuehn *et al.*, 2005).

Work with mtDNA has provided further evidence that the Neanderthals were a separate species from modern humans. But more information from the analysis of nuDNA is needed to resolve other questions over the migration of modern humans during the past 10,000 years. This debate came to a head in November 2005 with a publication that explained how farming spread so rapidly across Europe 6,000 years ago (Haak *et al.*, 2005). The paper presented evidence from the ancient mtDNA of the Neolithic LBK culture (from the German "Linearbandkeramik", or Linear Band Pottery), believed to be the first farmers in Europe 7,000 years ago, by analysing the occurrence of one particular mtDNA haplotype—N1a.

In one respect, this study, led by Wolfgang Haak from the Institute of Anthropology at the University of Mainz, Germany, appears to have been successful in convincing most neutral ancient-DNA researchers, and even some sceptics, that genuine sequences from endogenous DNA were recovered—in other words, that the common problem of contamination had been solved. "The DNA sequences look genuine," said Willemslev, who was not involved in the study, "although I wish they would have done some independent validation just in case."

...more information from the analysis of nuDNA is needed to resolve...questions over the migration of modern humans during the past 10,000 years

But there is little consensus over the conclusions: whereas N1a was common among the Neolithic LBK people, it is rare among modern Europeans. The authors deduced that modern Europeans did not descend from the Neolithics, but instead from the hunter-gatherers who were already in Europe. One possible explanation for this, as proposed in the paper, is that the farming culture spread through the communication and dissemination of ideas, rather than being carried directly by the Neolithic people themselves as they migrated through Europe. As the authors wrote, "this includes the possibility that small pioneer groups carried farming into new areas of Europe, and that once the technique had taken root, the surrounding hunter-gatherers adopted the new culture and then outnumbered the original farmers, diluting their N1a frequency to the low modern value."

According to Willerslev, this is too strong a conclusion to draw from the data. "Firstly they haven't tested the haplotype composition of the hunter-gatherers, and secondly you can get bottlenecks that change the composition of a population." Such a bottleneck might have been an epidemic that wiped out a large portion of a population, thus diminishing the N1a haplotype, although there is no evidence either way at present. To be fair, the authors admit that there could have been a subsequent population replacement in Europe, providing an alternative hypothesis for the dilution of the N1a haplotype.

However, the key point according to Hans-Jürgen Bandelt, Professor of Mathematics at Hamburg University, Germany, and a critic of much ancient human DNA work, is the question of what happened when the Neolithic people entered South-East Europe around 10,000 years ago. "Since much of the Neolithic in Europe took its start from South-East Europe, one needs to distinguish between the mitochondrial types that really entered Europe from Asia at the time and the types that were picked up at an early stage within Europe and then moved on," he said. To answer this question, it is essential to look at a much broader spectrum of the human genome, Bandelt added. It is here that the recent breakthroughs could help, if nuDNA sequences can be obtained.

But this does not solve the common problem of contamination with foreign DNA, which is the source of many controversies over ancient DNA research. As PCR does not discriminate, it will amplify whatever DNA is in a sample, and almost every sample excavated, whether from animal, plant, bacterial or other remains, is contaminated with DNA from other organisms. If the contamination occurred recently, for example from humans within the laboratory, it either out-competes the ancient sequences in PCR reactions, or recombines with the old sequences to yield products that look authentic but are in fact erroneous and misleading.

As Hofreiter pointed out, the extent of this problem depends on the organism being studied. In the case of recently extinct mammals such as the mammoth, the endogenous mtDNA recovered from bone or other fossils can usually be quite readily identified and separated from contamination. The mammoth's own DNA is similar to that of modern African and Asian elephants, but is sufficiently different to identify it as a distinct species.

The situation is different for DNA from human samples over the past 30,000 years. There are no clear structural features that can be used to distinguish it from human DNA contaminations that occurred during extraction or handling of the sample. Unfortunately, human DNA gets everywhere. Even if workers in the laboratory have their own DNA sequenced so that the source of

contamination can be identified, it is impossible to be sure that no other human DNA has found its way in.

There is one problem that DNA studies of population movements within the past 10,000 years do not have: dating of the samples, as radiocarbon dating is considered to be accurate up to 50,000 years. Beyond that, other methods are needed (Hebsgaard *et al*, 2005), which increases the margin for error. Nevertheless, most researchers in the field accept that the oldest DNA sample sequenced so far is from plant material around 400,000 years old (Willerslev *et al*, 2003). From animals, sequences at least 70,000 years old have been obtained—Gilbert and others published a sequence of bison DNA that was at least 64,800 years old (Gilbert *et al*, 2004).

This begs the question of what is the theoretical age limit for recoverable DNA. Various attempts have been made to calculate the theoretical survival rate for DNA at different temperatures, but as Gilbert pointed out, these have been based on simple degradation models that are known to be fairly inaccurate. The generally accepted limit is around 1 million years, certainly no more than 2 million, which makes it impossible to recover, for instance, DNA from dinosaurs. Yet DNA damage occurs in many ways, so it is possible that the absolute limit may be considerably greater, even if a sequence of that age has never been recovered.

The survival prospects of DNA also depend on which part of the organism it came from. In principle, DNA can be recovered from many sources, such as hair, skin and even excrement for animals, along with wood, leaves and roots for plants. In practice, however, bones, teeth and seeds are the best sources of ancient DNA. As Gilbert noted, bones and teeth contain a lot of "mummified" cells, such as osteocytes, that do not degrade or putrefy as quickly as other cells. The oldest samples will always come from permafrost, in locations that have remained continuously frozen from the time the fossil was deposited. Given that only a small proportion of available fossils lie beneath permafrost, there is an obvious desire to recover sequences from other locations. "For cave samples, nothing has been reported older than 100,000 years," Hofreiter noted.

... most researchers in the field accept that the oldest DNA sample sequenced so far is from plant material around 400,000 years old

Although DNA damage limits the amount of accurate sequence data that can be recovered, it does have one advantage: it helps distinguish the endogenous, ancient DNA from contamination. Even so, if contamination of samples can either be significantly reduced or at least differentiated from endogenous DNA, ancient DNA sequences—and the conclusions drawn from their analysis—will be more widely accepted. At the same time, the ability to obtain sequences from nuDNA will provide more interesting and diverse data about both migrations and selective pressures on individual genes.

REFERENCES

- Gilbert MT *et al* (2004) Ancient mitochondrial DNA from hair. *Curr Biol* **14**: R463–R464
- Haak W *et al* (2005) Ancient DNA from the first European farmers in 7500-year-old Neolithic sites. *Science* **310**: 1016–1018
- Hebsgaard MB, Phillips MJ, Willerslev E (2005) Geologically ancient DNA: fact or artefact? *Trends Microbiol* **13**: 212–220
- Krause J, Dear PH, Pollack JL, Slatkin M, Spriggs H, Barnes I, Lister AM, Ebersberger I, Paabo S, Hofreiter M (2005) Multiplex amplification of the mammoth mitochondrial genome and the evolution of *Elephantidae*. *Nature* 18 Dec 2005 [Epub] doi:10.1038/nature.04432
- Kuehn R, Ludt CJ, Schroeder W, Rottmann O (2005) Molecular phylogeny of *Megaloceros giganteus*—the giant deer or just a giant red deer? *Zool Sci* **22**: 1031–1044
- Lister AM, Edwards CJ, Nock DA, Bunce M, van Pijlen IA, Bradley DG, Thomas MG, Barnes I (2005) The phylogenetic position of the 'giant deer' *Megaloceros giganteus*. *Nature* **438**: 850–853
- Poinar HN *et al* (2005) Metagenomics to paleogenomics: large-scale sequencing of mammoth DNA. *Science* 20 Dec 2005 [Epub] doi:10.1126/science.1123360
- Shepherd LD, Millar CD, Ballard G, Ainley DG, Wilson PR, Haynes GD, Baroni C, Lambert DM (2005) Microevolution and mega-icebergs in the Antarctic. *Proc Natl Acad Sci USA* **102**: 16717–16722
- Willerslev E, Hansen AJ, Binladen J, Brand TB, Gilbert MT, Shapiro B, Bunce M, Wiuf C, Gilichinsky DA, Cooper A (2003) Diverse plant and animal genetic records from Holocene and Pleistocene sediments. *Science* **300**: 791–795

Philip Hunter

doi:10.1038/sj.embor.7400634