Roberto Macchiarelli & Gerd-Christian Weniger (eds.):
Pleistocene Databases: Acquisition, Storing, Sharing.
Wissenschaftliche Schriften des Neanderthal Museums 4, Mettmann 2011.

THE ANCIENT HUMAN OCCUPATION OF BRITAIN (AHOB) DATABASE

P. David Polly and Chris Stringer

Abstract

The Ancient Human Occupation of Britain (AHOB) project is a multi-institutional programme running since 2001 with the goal of better understanding the geographic, climatic, and environmental contexts in which the earliest humans occupied Britain and northern Europe. The AHOB database serves as a repository for data on which research was carried out and data generated by the project. The data are stored in an SQL-compliant relational database that can be accessed in several formats, including a browsable web interface. Data are structured around sites and their stratigraphic horizons and consist of mammal faunal records, archaeological lithic technologies, age assessments, radiometric dates, stable isotope data, as well as associated photographs, published literature, and location information. The database contains information on 321 sites, most of them British. The database also serves as a research tool in itself for the study of the climates and environments of early humans in Britain.

Keywords: Ancient Human Occupation of Britain, Quaternary, mammals, archaeology, MySQL

Introduction

The Ancient Human Occupation of Britain (AHOB) project is a multi-institutional programme whose goal is to better understand the geographic, climatic, and environmental contexts in which the earliest humans occupied Britain and north-western Europe (Stringer 2006, 2011). The programme, funded by The Leverhulme Trust since 2001, includes paleoanthropologists, archaeologists, geologists, palaeontologists, and geochemists among its core members. Through the last decade, the AHOB project has had three phases: Britain has been a focus of all three phases, with the last two extending to the continent for a more regionally synthetic picture of human habitation in northern Europe. The current and final phase, which involves project members from Britain, the Netherlands, Germany, Belgium, and France, has four overriding research themes: the record of human dispersal at the frontiers of northern Europe, the natural factors controlling human dispersal, population dynamics at the northern limits of their range, and the human responses to marginal environments. The AHOB database is a repository for both data generated by the project and data used by project members as part of their research.

Main Findings of the Project

The overarching goal of the AHOB project has been to better document the record of human occupation and associated faunal changes in Britain and northern Europe from the earliest appearance of humans through the Younger Dryas (Schreve 2004; Stringer 2006; Ashton et al. 2011). Re-evaluation of previously known key sites and the study of new ones has been the core of AHOB's work. Especially important to the enterprise has been fitting the archaeological and palaeontological evidence into the Marine Oxygen Isotope Stage (MIS) chronology that has developed over recent decades.

The Mammal Assemblage Zone (MAZ) system has been key to tying many sites to MIS stages (Currant & Jacobi 2001; Schreve 2001). This biostratigraphic system has established, for example, that the *Hippopotamus-Palaeoloxodon* fauna is restricted to MIS 5, particularly stage 5e and a *Mammuthus-Equus* fauna is typical of parts of MIS 7. The MAZ/MIS chronology has been critical for determining when humans were present and absent in Britain and what the associated environmental context was really like. It has been used to provisionally or more securely place key sites like Norton Subcourse, Pakefield, Happisburgh 1, High Lodge, Hoxne, Grays Thurrock, Banwell Bone Cave, Kent's Cavern, Gough's Cave and Whitemoor Haye into the global Quaternary timescale (Currant & Jacobi 2001; Schreve 2001; Ashton et al. 2004; Lewis et al. 2004; Parfitt et al.

2005; Jacobi et al. 2006; Parfitt et al. 2010).

Among its more important findings, AHOB work has confirmed that humans were absent in Britain during the warm MIS 5 interglacial, following a period of increasing isolation during the preceding cycles (Lewis et al. 2011). Neanderthals returned at the end of MIS 4 and modern humans arrived at the end of MIS 3 (Jacobi & Higham 2011).

As surprising as their absence only 125 kya, AHOB work has also confirmed that humans were present in Britain long before MIS 13. The Boxgrove site had produced what was then some of the earliest evidence of humans in Britain and northern Europe (Stringer et al. 1998). The lack of evidence of humans prior to Boxgrove and Mauer (in Germany) led Roebroeks and van Kolfschoten (1994) to argue that humans colonized Europe only about 500 kya. However, discoveries by AHOB and collaborators of flint artefacts and cutmarks at Pakefield, probably belonging to MIS 17 (ca. 700 kya, Parfitt et al. 2005), and then finds at Happisburgh 3 of artefacts more than 780 kya old (Parfitt et al. 2010), have demonstrated that humans were not only in Europe more than a quarter million years earlier than that, but that they had reached the northern part of the continent.

SCOPE OF THE AHOB DATABASE

The AHOB database was instituted during the first phase of the project as a repository for core data on which findings were based. Mammal faunal lists, archaeological industry categorizations, and basic site information dominate the contents of the database, but it also contains radiometric dates, geochemical results, photographs, site plans, bibliographical references and other kinds of data.

To date, the research tables of the database contain basic information about 321

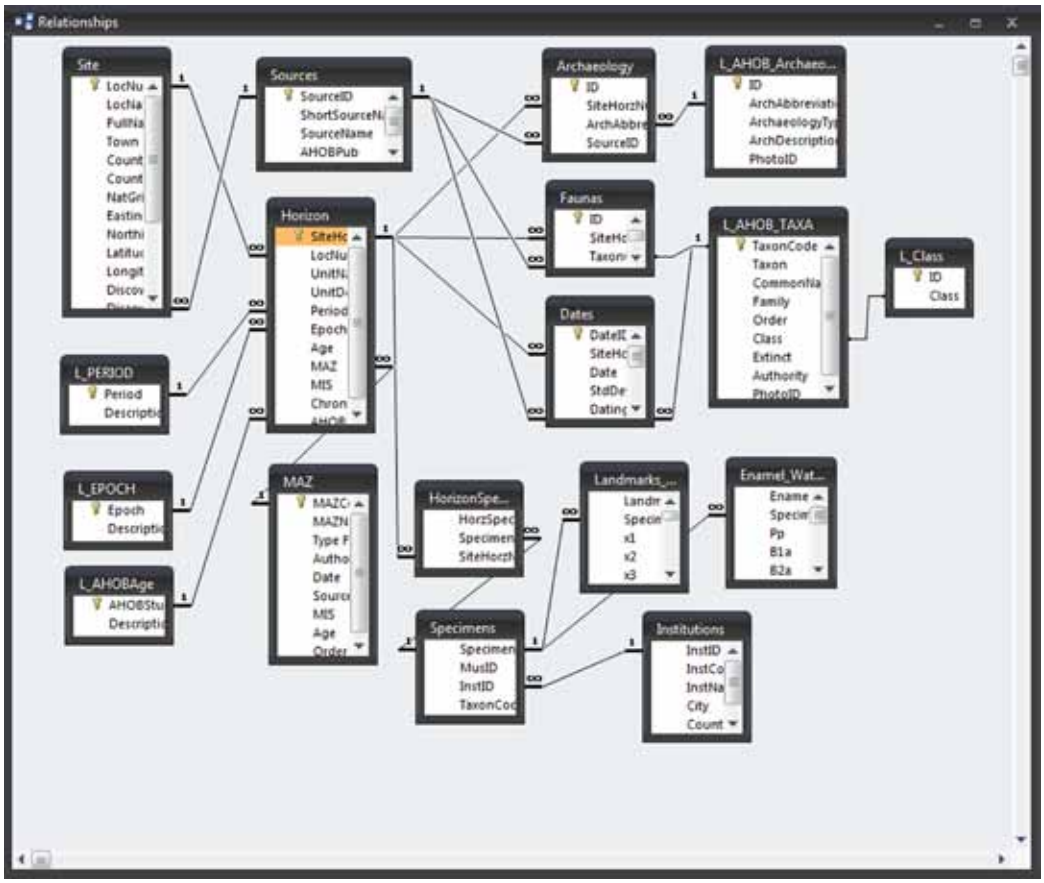

Fig. 1 Data Relationships. Data are organized around horizons, with one or more horizons per site. Age assessments are associated directly with horizons and each horizon may have a fauna or archaeology. The sources table records bibliographic or personal communication citations for the ages, faunas, archaeology, and other information recorded in the database.

sites (most of them in Britain), 1669 species occurrences, 299 lithic categorizations, 146 radiometric dates, 133 isotopic readings, 140 photographs, and 456 bibliographic records. The main source of these data is direct contribution by AHOB researchers, with published literature being a second source.

As the programme matured, the database was extended to include information about the project itself: member details, announcements, minutes of meetings, grant proposal texts, presentations, budgets, and similar material. The administrative tables include contact and biographical details of 76 project researchers, information on 75 project documents, 46 records of the project in the press, and 26 project announcements.

The database thus contains several categories of data, each of which has a different audience. Some data will be of interest to Quaternary researchers, some will be of administrative interest to project members, and some will be of broad interest to the general public. Separate interfaces and user restrictions allow the database to meet all three needs.

Data relationships

Data relationships for the research-related data are structured around sites, specifically around individual stratigraphic horizons within sites (Fig. 1). Basic site details, faunal occurrences, archaeological determinations, stratigraphic ages, radiometric dates, isotopic readings, photographs, and publications are all linked to the horizon table in the database structure. General information, including a short description and National Grid coordinates are available for each site. The archaeology of each site is recorded simply as what lithic industry, if any, present at a site (the categories include Lower Palaeolithic, Early Upper Palaeolithic, Late Upper Palaeolithic, Clactonian, Levallois, Mousterian, Human Presence (non-lithic), Cut Marks, Humans Absent (Verified), and Humans Absent (Unverified)). Faunas are recorded in the database as species occurrences associated with a particular site horizon. Currently 217 taxa, mostly mam-

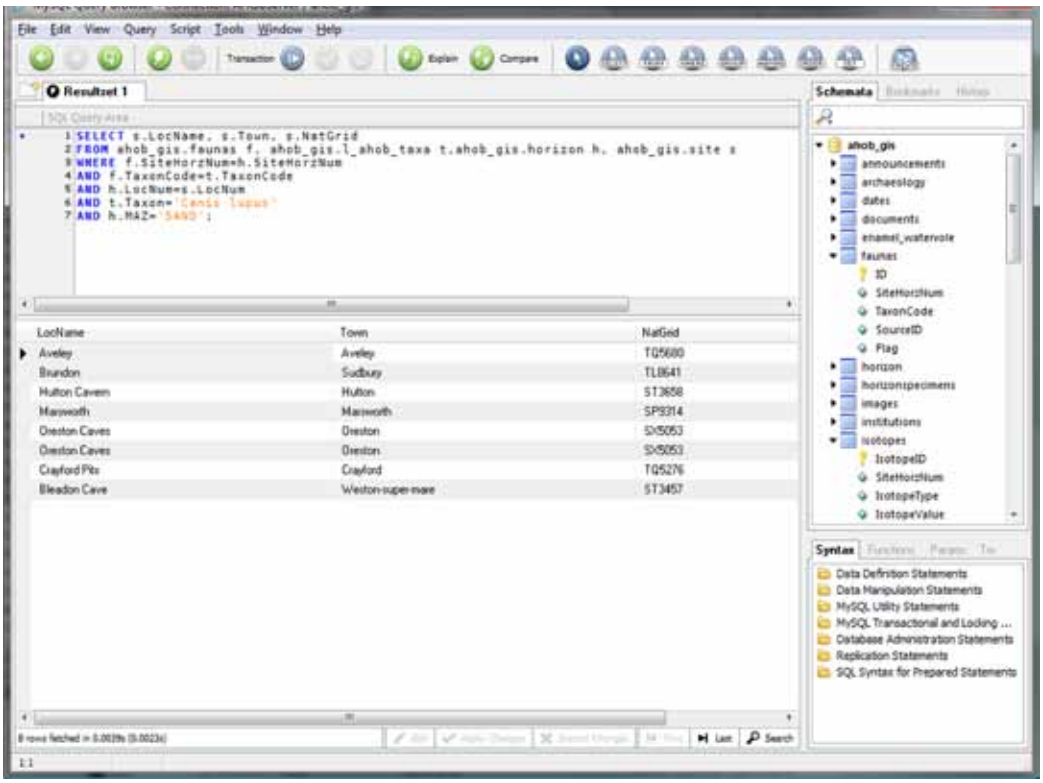

Fig. 2 An example of a raw SQL query made directly to the MySQL server. This particular query joins several tables and then asks for the name, nearest town, and National Grid coordinates for all the sites from the Sandy Lane MAZ that have the species Canis lupus.

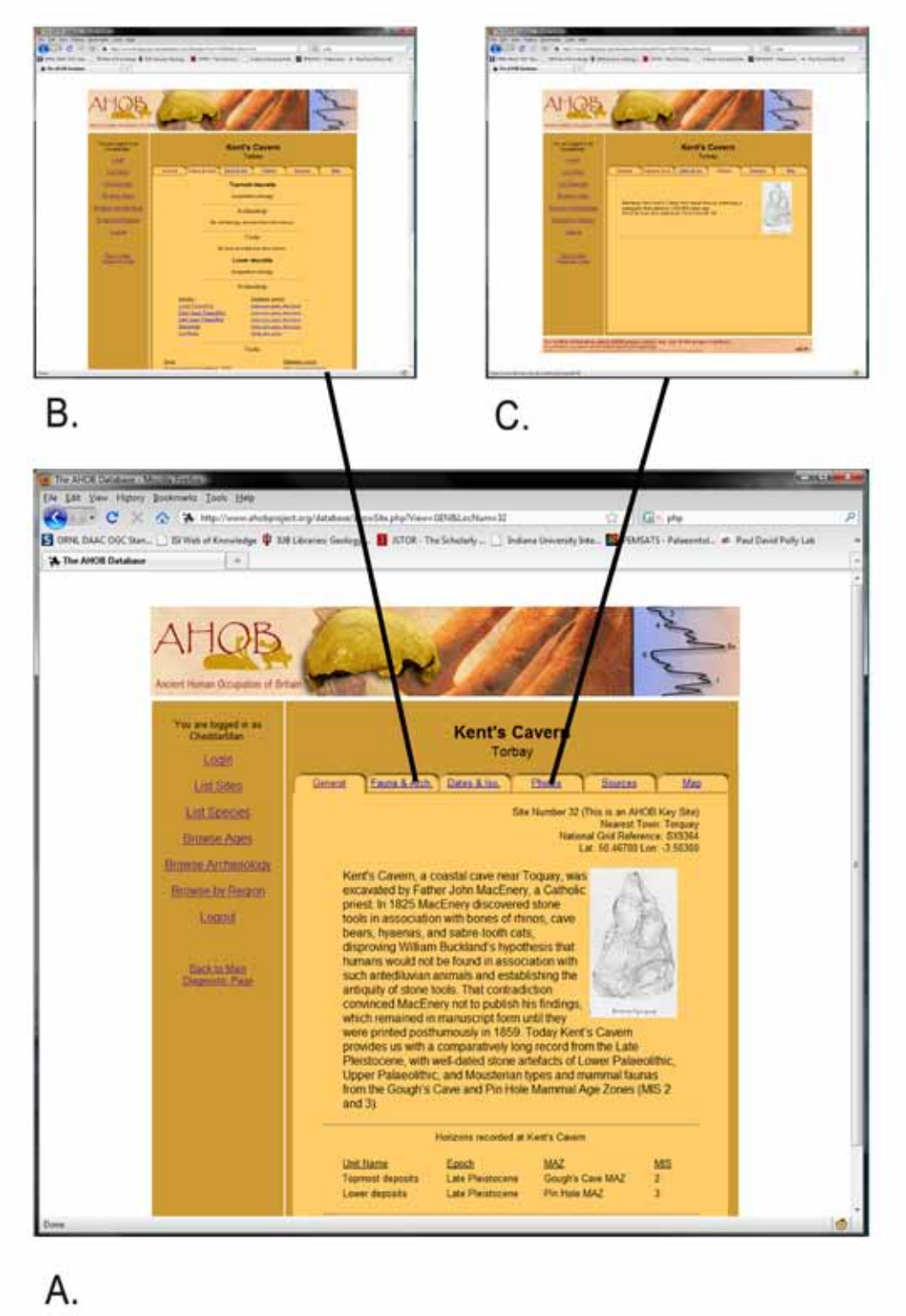

Fig. 3 Web interface. A customised web interface can be used to browse the database by site name, by geographic region, by species, by age, or by lithic industry. The main site page (A.), which displays a short description of the site aimed at a general audience, a photo (if available), and a summary of the horizons at the site and their ages. The tabs across the top of the frame retrieve detailed information. The fauna and archaeology page (B.) reports for each horizon a brief list of lithic industry present, a faunal list (including the taxonomic authority for each species), and the source for each item of information. Industries, species, and sources are all hyperlinked - clicking will retrieve other sites with the same industry, archaeology or source. The photo page (C.) shows a table of photos from the site along with a caption. Clicking on the photo retrieves a high-resolution copy.

mals, are found in the database. Lookup tables contain core taxonomic information about each species (e.g., author and date of the taxon name). Several kinds of stratigraphic age estimates are recorded for each horizon, including mammal age zone for British sites (MAZ) and marine oxygen isotope stage (MIS). A source is recorded for each datum, regardless of whether it is personal communication from a project member or reference to a formal publication. Data are being structured so that the AHOB database will be compatible with cross-database querying on the PaleoAnth Portal (http://www.paleoanthportal.org/), a collaborative project that is still in the organizational stages.

Data storage and retrieval

The data are stored in two database formats. Data are stored in a relational Microsoft Access® format that can be distributed for use on individual computers. Data are also stored in MySQL, an open-source, SQL-compliant relational database with client-server architecture (http://www.mysql.com). The advantage of the MySQL server is that it can be queried over the internet and interfaced using the world-wide web. SQL (structured query language) is an international standard for interacting with relational databases. Raw SQL queries can be made directly to the MySQL server, bypassing the customized web interface to provide customised queries for particular problems. For example, one could retrieve the basic information about all the sites in the database by submitting the following SQL query:

```
SELECT * FROM site;
```

Importantly, one can retrieve very specific information using a more complicated query. The following SQL, for example, retrieves the site name, the nearest town and the British National Grid coordinates for all sites where the wolf, *Canis lupus*, occurs during the Sandy Lane mammal age zone:

```
SELECT s.LocName, s.Town, s.NatGrid
FROM ahob_gis.faunas f, ahob_gis.l_ahob_taxa t, ahob_gis.horizon h,
ahob_gis.site s
WHERE f.SiteHorzNum=h.SiteHorzNum
AND f.TaxonCode=t.TaxonCode
AND h.LocNum=s.LocNum
```

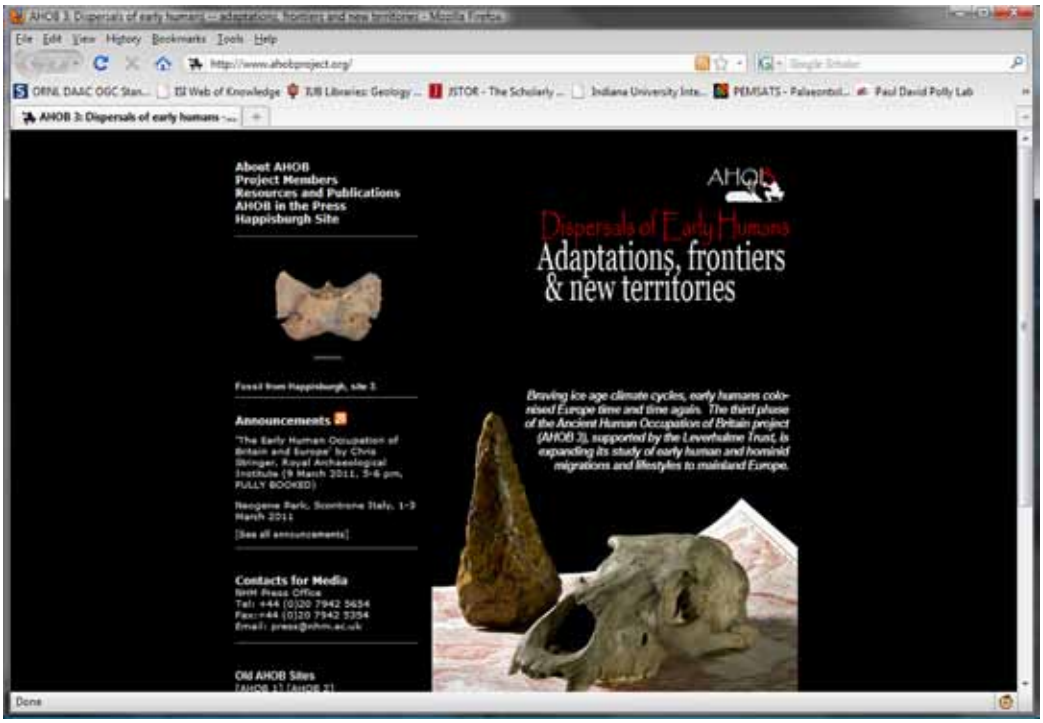

Fig. 4 The public AHOB website at http://www.ahobproject.org/. Many items, such as the image and list of announcements in the left column, are drawn from the AHOB database.

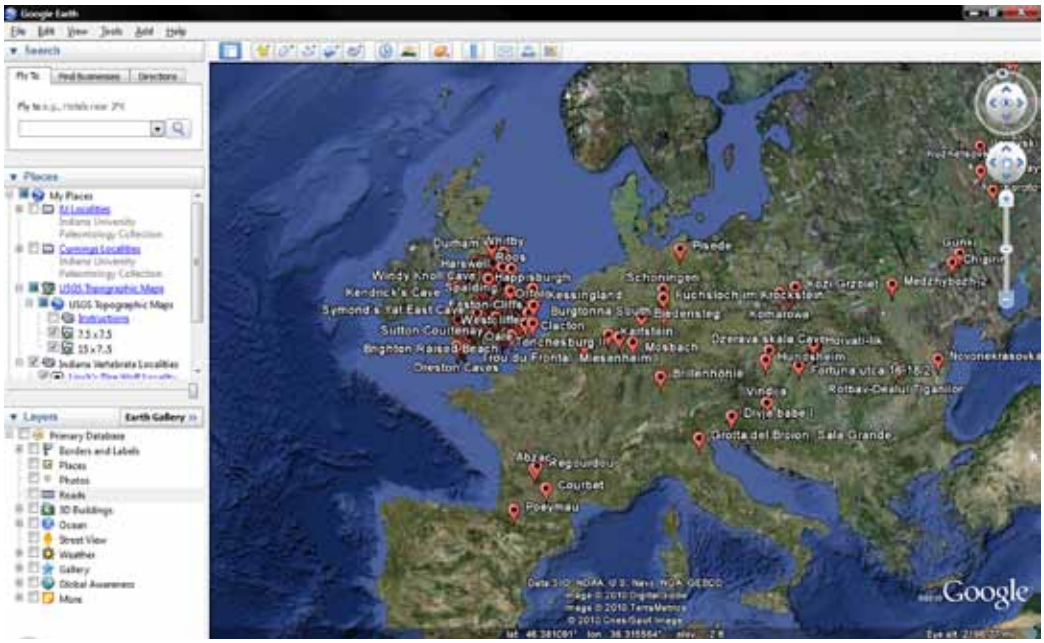

Fig. 5 Display of site locations from the AHOB database in Google Earth shared via a KML file from the web interface to the database.

```
AND t.Taxon='Canis lupus'
AND h.MAZ='SAND';
```

Figure 2 shows this query and its results in the MySQL Query Browser client. Both database engines can be interfaced with standard GIS programs to map query results.

WEB INTERFACE

A customised world-wide web interface has been built to aid access to the data by AHOB researchers and, ultimately, the research public (Fig. 3). Like the relational structure of the database, the web interface is also organized around sites. Users can locate site data by browning lists of sites, lists of species, ages, lithic industries, or geographic regions. Once a particular site is identified, users find general information, fauna and archaeology, dates and isotope readings, photos, bibliographic sources, and an online map by clicking the tabs. Sites are interlinked by clicking. For example, if one browses the fauna of the Aveley site in Greater London, one can click on the name *Bos primigenius*, which retrieves information (and photos, if they are available) about the Aurochs. Another click retrieves a list of all sites with that species in the faunal list. The same principle is functional with archaeology lists, bibliographic sources, and ages.

The web interface is built using PHP hypertext processing scripts (http://www.php.net/) to retrieve data from the MySQL server and present it to the user's internet browser in HTML (hypertext markup language, the native format of the world-wide web). The information presented on any particular page of the interface may come from one or several queries of the database.

NEWSFEEDS, FACEBOOK®, GOOGLE EARTH® AND MORE

The database architecture makes it possible to distribute data to users in several formats, both generic ones such as comma-delimited text files and specialized ones such as FaceBook® news feeds.

Much of the information on the public AHOB website (Fig. 4, http://www.ahob-project.org/) also comes from the database via PHP scripts to the MySQL server. Such items include the randomly selected photograph on the front page, list of project members, and the publication list.

Public announcements about AHOB events, such as lectures and conferences, are

stored in the database and distributed by several means. A list of current announcements scrolls in the sidebar of the public AHOB website, which is generated by a PHP script that queries the database for events whose expiration date has not been reached (Fig. 4). Readers can click on a link that brings all current announcements up on a larger, easier to read page that is also generated by script direct from the database. Another script generates an RSS (Really Simple Syndication) feed, which is an international standard for publishing frequently updated information such as news, headlines or blog entries, in a format that can be used by third-party programs. Users can subscribe to two AHOB RSS feeds, one for announcements and one for press information. Users might display this information in the news portion of their web browser, in an iPhone news app, or in a desktop sidebar app.

AHOB also has a FaceBook® (http://www.facebook.com/) interface to the database. The announcements from the database are piped to FaceBook® using the RSS feed so that users there can read new announcements on the AHOB page or in their own home page when they log in.

Much of the research data presented on pages of the web interface to the AHOB database can be downloaded for use in other programs. Tables of data can be saved as comma-delimited text files that can be imported into programs like Microsoft Excel® or Word®.

The web interface also makes use of embedded Google Earth® maps (http://www.earth.google.com/) to plot site locations. AHOB sites can be plotted amid the rich data available via Google Earth®, including satellite photos and 3D terrain. Users can also download their geographic results as a KML file. KML (keyhole markup language) is an XML-style format for sharing geographic data that can be used in many geographic information system programs, including Google Earth® (Conroy et al. 2008) (Fig. 5).

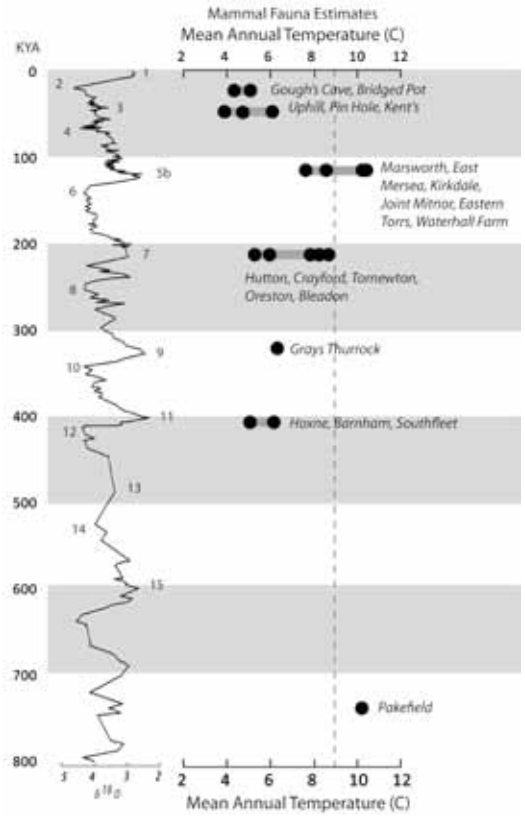

Fig. 6 Palaeotemperature curve estimated from mammalian faunal occurrences in the AHOB database. Black circles indicate the estimate for each site, sites are grouped by marine isotope stage (MIS), site labels refer to the points in each group from left to right. MISs are labelled on the oxygen isotope curve on the left, which is for the northern hemisphere (Raymo and Ruddiman 2004). The broken vertical line shows mean annual temperature today for south-east England outside the Thames valley (Met Office data, 1961-1990 average). (after Polly & Eronen 2011).

Research with the AHOB database

In addition to serving as a repository for data generated by AHOB research, the database is also being used as a research tool in its own right. For example, Polly and Eronen (2011) used the database to study the implications of mammalian palaeofaunas for climate interpretation.

The database was used to identify sites where ten species living today are associated with quite distinct climates. Arctic fox (*Alopex lagopus*), Norway lemming (*Lemmus lemmus*), Musk-ox (*Ovibos moschatus*) and Reindeer (*Rangifer tarandus*) live today in high-latitude cold climates. Spotted hyena (*Crocuta crocuta*), Lion (*Panthera leo*), and Hippo (*Hippopotamus amphibius*) live today in low-latitude warm climates. Water voles (*Arvicola terrestris*), Red deer (*Cervus elaphus*), and Wild boar (*Sus scrofa*) live in a wide range of mid-latitude climates. The Quaternary record has many seem-

ingly surprising mixtures of species, notably co-occurrences of *Panthera leo* or *Crocuta crocuta* with *Rangifer tarandus*, whose geographic ranges do not overlap today.

The AHOB occurrence data were analyzed in light of geographic range data for the ten species and modern global climate data to determine whether such species combinations found in the Quaternary record were consistent with the ranges of climates inhabited by the species today. Twenty-two sites were identified where at least three of the ten species co-occurred. Of those, 82% had pairs of species whose modern ranges do not overlap and 44% had pairs of species whose modern climatic ranges do not overlap, suggesting that many species today, especially carnivores, have a restricted geographic range that is not indicative of their past climatic tolerances.

The same data were used to estimate palaeoclimate parameters for sites ranging from MIS 17 down through MIS 3. The modern climate distributions of the species found in the palaeofaunas were combined using maximum-likelihood to provide the most likely estimate of the palaeoclimate given the mammals that lived there. The results were in good agreement with other palaeoclimate proxies, notably oxygen isotope estimates of palaeotemperature (Fig. 6).

ACKNOWLEDGEMENTS

We thank Nick Ashton, Silvia Bello, Ian Candy, Andy Currant, Tom Higham, Mark Lewis, Simon Lewis, Simon Parfitt, Kirsty Penkman, Richard Preece, Wil Roebroeks, Jim Rose, Mark Ruddy, Danielle Schreve, Mark White and the others who have contributed to the information contained in the AHOB database. This work was facilitated by grants from The Leverhulme Trust to the Ancient Human Occupation of Britain (AHOB) project and by Indiana University. We dedicate this paper to the memory of Roger Jacobi, who contributed tirelessly to the archaeological and faunal information in the database in spite of his good-humoured scepticism of the need for such a database.

P. David Polly
Department of Geological Sciences
Indiana University
1001 E. 10th Street
Bloomington, IN 47405
USA

Chris Stringer
Palaeontology Department
Palaeontology Research Division
Natural History Museum
Cromwell Road
London
United Kingdom

References cited

Ashton, N. M., Lewis, S. G. & Parfitt, S. A. (2004). *Hoxne 2003. Proceedings of the Suffolk Institute of Archaeology and Natural History, XL* (part 4).

Ashton, N. M., Lewis, S. G. & Stringer, C. B. (Eds.) (2011). *The Ancient Human Occupation of Britain*. Elsevier, Amsterdam, 279-304.

Conroy, G. C., Anemone, R. L., Regenmorter, J. V. & Addison, A. (2008). Google Earth, GIS, and the Great Divide: a new and simple method for sharing paleontological data. *Journal of Human Evolution* 55: 751-755.

Currant, A. P. & Jacobi, R. M. (2001). A formal mammalian biostratigraphy for the Late Pleistocene of Britain. *Quaternary Science Reviews* 20: 1707-1716.

Jacobi, R. M., Higham, T. G. F. & Bronk Ramsey, C. (2006). AMS radiocarbon dating of Middle and Upper Palaeolithic bone in the British Isles: improved reliability using ultrafiltration. *Journal of Quaternary Science* 21: 557-573.

Jacobi, R. M. & Higham, T. G. F. (2011). The British earlier Upper Palaeolithic: settlement and chronology. *In*: N. M. Ashton, S. G. Lewis & C. B. Stringer (Eds.) *The Ancient Human Occupation of Britain*. Elsevier, Amsterdam, 181-222.

Lewis, S. G., Parfitt, S. A., Preece, R. C., Sinclair, J., Coope, G. R., Field, M. H., Maher, B. A., Scaife, R. G. & Whittaker, J. E. (2004). Age and palaeoenvironmental setting of the Pleistocene vertebrate fauna at Norton Subcourse, Norfolk. *In*: D. C. Schreve (Ed.) *The Quaternary Mammals of Southern and Eastern England*. Field Guide. Quaternary Research Association, London, 5-17.

Lewis, S. G., Ashton, N. & Jacobi, R. (2011). Testing human presence during the last interglacial (MIS 5e): a review of the British evidence. *In*: N. M. Ashton, S. G. Lewis & C. B. Stringer (Eds.) *The Ancient Human Occupation of Britain*. Elsevier, Amsterdam, 125-164.

Parfitt, S. A., Barendregt, R. W., Breda, M., Candy, I., Collins, M. J., Coope, G. R., Durbidge, P., Field, M. H., Lee, J. R., Lister, A. M, Mutch, R., Penkman, K. E. H., Preece, R. C., Rose, J., Stringer, C. B., Symmons, R., Whittaker, J. E., Wymer, J. J. & Stuart, A. J. (2005). The earliest record of human activity in northern Europe. *Nature* 438: 1008-1012.

Parfitt, S. A., Ashton, N. M., Lewis, S. G., Abel, R. L, Coope, G. R., Field, M. H., Gale, R., Hoare, P. G., Larkin, N. R., Lewis, M. D., Karloukovski, V., Maher, B. A., Peglar, S. M., Preece, R. C., Whittaker, J. E. & Stringer, B. B. (2010). Early Pleistocene human occupation at the edge of the boreal zone in northwest Europe. *Nature* 466: 229-233.

Polly, P. D. & Eronen, J. T. (2011). Mammal associations in the Pleistocene of Britain: implications of ecological niche modelling and a method for reconstructing palaeoclimate. *In*: N. M. Ashton, S. G. Lewis & C. B. Stringer (Eds.) *The Ancient Human Occupation of Britain*. Elsevier, Amsterdam, 279-304.

Raymo, M. E. & Ruddiman, W. F. (2004). DSDP Site 607 Isotope Data and Age Models, IGBP PAGES/World Data Center for Paleoclimatology, Data Contribution Series #2004-010. NOAA/NGDC Paleoclimatology Program, Boulder, CO, USA.

Roebroeks, W. & van Kolfschoten, T. (1994). The earliest occupation of Europe: a short chronology. *Antiquity* 68: 489-503.

Schreve, D. C. (2001). Differentiation of the British late Middle Pleistocene interglacials: the evidence from mammalian biostratigraphy. *Quaternary Science Reviews* 20: 1693-1705.

Schreve, D. C. (Ed.) (2004). *The Quaternary Mammals of Southern and Eastern England. Field Guide*. Quaternary Research Association, London.

Stringer, C. (2006). *Homo britannicus*. Penguin, London.

Stringer, C. (2011). The changing landscapes of the earliest human occupation of Britain and Europe. *In*: N. M. Ashton, S. G. Lewis & C. B. Stringer (Eds.) *The Ancient Human Occupation of Britain*. Elsevier, Amsterdam, 1-10.

Stringer, C., Trinkaus, E., Roberts, M. B., Parfitt, S. A. & Macphail, R. (1998). The Middle Pleistocene human tibia from Boxgrove. *Journal of Human Evolution* 34: 509-547.

WEBSITES CITED

http://www.ahobproject.org/

http://www.earth.google.com/

http://www.facebook.com/

http://www.mysql.com/

http://www.paleoanthportal.org/

http://www.php.net/