Supplementary material

Geographic distances

Conventional spatial statistics (as used in GIS software) assumes data to be in a Cartesian coordinates system resulting from projecting a spherical surface onto a flat surface. While this approximation is quite accurate for relatively limited areas, it is not appropriate for problems that encompass the whole globe. As we were interested in migrations on a very large scale, we approximated the earth to a sphere and explicitly modeled movement on a curved surface. Even though the earth is a spheroid rather than a sphere, this approach is more precise than any projection of the whole globe onto a Cartesian coordinate system.

Addis Ababa in Ethiopia was chosen as our postulated origin, a location in good agreement with the most recent paleontological evidence [1]. Estimates of distance across land from Addis Ababa to each of the studied populations were obtained using graph theory. Under graph theory, a graph is defined as a set of points ("vertices") that can be connected by "edges". Thus, we can visualise a set of locations on earth as vertices on a graph. To move from one location to another, we can "walk" a path on the graph following the edges that connect the locations of interest. If the graph is non-directional, we can walk along any edge irrespective of which vertex we start from. To simulate a migration across land between two distant locations on earth, we need to break down the route so that no step can cross a body of water. To do this, we would need a large number of vertices regularly spaced on all landmasses, only allowing movement between adjacent vertices. To be able to reach all possible locations on earth, we also need additional vertices to form putative "land-bridges" that might have existed in the past (for example, a land bridge on the Bering straight is needed to allow movement from Eurasia to America). As there are a large number of potential routes for any given migration, we can take the shortest path to be the most parsimonious.

In practice, it is impossible to create a regular grid of a large number of vertices on a spherical surface. However, we created an approximately regular grid on the surface of the globe by laying out 30,000 points on a spiral with a constant angular pitch, moving from one pole of the sphere to the opposite one. The grid of points so obtained had locations that were approximately a hundred kilometres from their four closest neighbours. As discussed above, the points that either fall on land or on one of the putative land bridges can be thought as vertices of a non-directional graph. By restricting movement from one vertex to its eight closest neighbours, we simulated migration on the globe as paths on the graph. The length of an edge connecting two vertices was defined as their great circle distance (i.e. the shortest distance between two points on the surface of a sphere). This provided an exact distance between two locations. The shortest path from Addis Ababa to each population was then obtained using Dijkstra's algorithm, which finds the shortest path from one vertex to all other vertices on a graph. This generated reasonable migration routes (Figure 1) compatible with the ones suggested in [2], with the exception of routes to

America, where the algorithm suggested a rather unlikely scenario with a long stretch at very high latitudes in Asia. To avoid this impossible migration, we forced the route to first reach Eastern Asia before veering to high latitudes to move through the Bering straight.

References

1. White T, Asfaw B, DeGusta D, Gilbert H, Richards G, Suwa G, Howell F: **Pleistocene Homo sapiens from Middle Awash, Ethiopia**. *Nature* 2003; **423**:742-747.
2. Cavalli-Sforza L, Menozzi P, Piazza A: *The history and geography of human genes*. Princeton: Princeton University Press; 1994.