POPULATION GENETICS

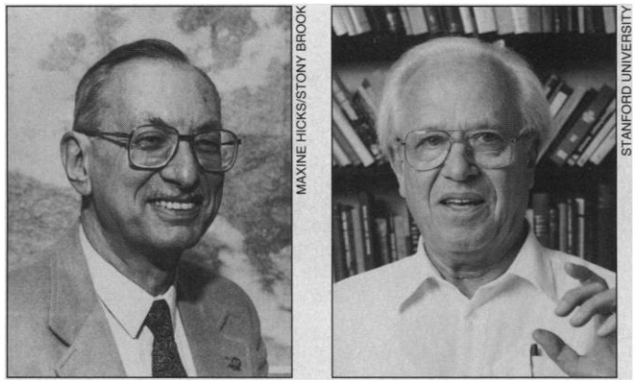
Using Genes to Track Down Indo-European Migrations

For years a debate has been raging about the origins of the Indo-Europeans, the ancestral people whose language lies at the root of most modern European languages: Where did they come from, and when and how did they and their new language come to dominate Europe? Once the debate was waged solely on the territory of archeology and linguistics, by investigators searching for clues in the distribution of artifacts and languages. Now the search has moved to new ground: genetics. But a recent paper and the response it has provoked suggest that the new terrain is just as treacherous as the old.

When Robert Sokal, a population geneticist at the State University of New York at Stony Brook, began his quest a couple of years ago, his hope was that genetic analyses could sort out whether the Indo-Europeans swept into Europe from north of the Black Sea in warrior bands or migrated in from the Middle East as an agricultural people—the two competing theories of the past 5 years. Instead, as he reports in the 15 August *Proceedings of the National Academy of Sciences*, his new data suggest that both theories are wrong. What's more frustrating, he says, is that his analysis doesn't provide a clear alternative to either theory. And to top it off, Sokal's conclusion has been challenged by one of his prominent colleagues, population geneticist Luca Cavalli-Sforza of Stanford.

The standoff Sokal hoped to resolve dates from a 1987 book by Colin Renfrew, an archeologist at Cambridge University. In it Renfrew challenged the prevailing hypothesis, first proposed by archeologist Marija Gimbutas of the University of California, Los Angeles, in the 1970s. She argued that the Indo-European languages were carried into Europe some 6000 years ago by the Kurgan people, a band of warrior tribes. The first to harness the horse for military purposes, the Kurgan people swept out of the Pontic Steppes north of the Black Sea (in what is now eastern Ukraine) and into Europe in three waves beginning in 4500 B.C., imposing both their culture and, Gimbutas maintains, their language on the inhabitants.

Not so, said Renfrew. He argued that the Indo-European languages spread across Europe along with the spread of agriculture, as early farmers searching for more land migrated from Turkey and Asia Minor to southeastern Europe in about 7000 B.C. With that proposal Renfrew embraced the heretical idea, first proposed in 1972 by Cavalli-Sforza and Albert J. Ammerman of the University of Parma, that agriculture was spread by the migration of farmers and not simply by the transmission of new knowledge, as commonly postulated. Cavalli-Sforza's ideas about the spread of agriculture, based on a genetic analysis of modern Europeans, have since gained support, some of it from a reanalysis of the data by Sokal and colleagues. But whether the migrating farmers really were the carriers of the Indo-European tongue remained an open question.



At odds. Sokal *(left)* thinks he has ruled out two scenarios for the Indo-European migration; Cavalli-Sforza thinks they're still viable.

To try to find out, Sokal and Neil Oden and Barbara Thomson of Stony Brook compared the linguistic distances among modern European languages—determined with the help of linguist Merritt Ruhlen, an independent researcher—with the genetic distances among the people speaking them. The genetic distances came from an analysis of gene frequencies from more than 2000 population samples across Europe.

Sokal's group found that genetic and linguistic distances are strongly correlated—that is, the greater the genetic differences between populations, the greater the differences between their languages. As expected, most of the correlation could be explained by geography alone. The reason is simple: People living far apart are likely to develop different dialects and become genetically distinct. But geography is not the whole answer, the group found. When they controlled for its influence, performing a statistical trick that, say, made the distance between Spain and Russia the same as the distance between Spain and Portugal, a "significant partial correlation" between genetic and linguistic distance still remained, says Sokal. Clearly, he says, "something else is going on"—some other process had to be invoked to explain the fact that linguistic distances tend to reflect genetic distances. But what? Could it be explained by the migration of Neolithic farmers or by the Kurgan conquest?

With the help of both Gimbutas and Renfrew, Sokal developed a way to express each hypothesis numerically and then tested each to see if it could account for the remaining correlation. Much to his surprise, says Sokal, "we found absolutely nothing. The genetic evidence does not support either Gimbutas or Renfrew," though he is at a loss to offer an alternative. Even so, the data do offer an intriguing hint about Indo-European origins, he says. They are compatible with the notion that the ancestral Indo-European populations branched or split at different times and then moved into different regions of Europe, where their languages continued to evolve. What is not clear, he says, is whether that branching occurred within Europe or outside, perhaps in Asia.

Sokal hastens to say that the team's negative findings don't cast doubt on the early migrations themselves. He still believes farming was spread by people and not by the passing of knowledge and, similarly, that the Kurgan people moved across Europe. But now, he says, "I seriously doubt these peoples were Indo-Europeans."

Wait a minute, says Cavalli-Sforza, who asserts that "a negative result with this method is not a proof that the hypotheses are wrong." He believes that Sokal's technique, which looks for a linear or proportional relationship between genetic and linguistic distances, may not be sensitive enough to resolve the issue. "My position is that on this basis you can't dismiss either hypothesis," insists Cavalli-Sforza. "It is possible that both Renfrew and Gimbutas are right." Indeed, Cavalli-Sforza speculates, as did Renfrew before him, that when the Neolithic farmers started from Turkey and spread across Europe, they may well have spoken a primitive form of Indo-European. And, he continues, there is good reason to think the Kurgan people may have been descendants of those early farmers. They too would have spoken a variant of Indo-European—albeit a more modern one.

Sokal is not persuaded, noting that Cavalli-Sforza's idea is "interesting but a little difficult to reconcile" with the data. Nor is he convinced that there is any weakness in his technique. He and Cavalli-Sforza will have ample time to discuss data and methods over the next year at Stanford, where Sokal is spending his sabbatical. But given the intensity of the debate on Indo-European origins, 1 year may not be enough.

–Leslie Roberts