The Kusunda people of central Nepal have long been regarded as a relic tribe of South Asia. They are, or were until recently, seminomadic hunter-gatherers, living in jungles and forests, with a language that shows no similarities to surrounding languages. They are often described as shorter and darker than neighboring tribes. Our research indicates that the Kusunda language is a member of the Indo-Pacific family. This is a surprising finding inasmuch as the Indo-Pacific family is located on New Guinea and surrounding islands. The possibility that Kusunda is a remnant of the migration that led to the initial peopling of New Guinea and Australia warrants additional investigation from both a linguistic and genetic perspective.

The Kusunda people of central Nepal are one of the few “relic” tribes found on the Indian subcontinent (the Nathali of India and the Veddas of Sri Lanka are two others). They first appeared in native ethnographic literature in 1848, when they were described by Hodgson as follows: “Amid the dense forests of the central region of Népál, to the westward of the great valley, dwell, in scanty numbers and nearly in a state of nature, two broken tribes having no apparent affinity with the civilized races of that country, and seeming like the fragments of an earlier population” (1). The Kusunda were one of these “broken tribes”; the Chepang were the other. Hodgson went on to show, however, that the Chepang were, on linguistic grounds, closely related to the Lhopa of Bhutan and must be presumed to have split off from this group and moved west at some time in the past. Hodgson had been unable to obtain any data on the Kusunda language, so nothing could be said of their possible affinity with other groups. Nine years later Hodgson published an article that contained the first linguistic data on the Kusunda language (2) as well as data on other Nepalese languages, but he offered no specific discussion of Kusunda even though his data showed quite clearly that the Kusunda language bore virtually no resemblance to any of the other languages he examined. No additional information on Kusunda appeared for more than a century until Reinhard and Toba (3) offered a brief description of the language, which provided some additional data. The final source on Kusunda appeared in an article by Reinhard in 1976 (4), but there is very little additional information that is not already found in the article by Reinhard and Toba (3).

Although Hodgson had predicted in 1848 the demise of the Kusunda in a few generations, a few Kusunda have managed to survive to the present day. Until recently they were seminomadic hunter-gatherers living in jungles and forests, and indeed their name for themselves is “people of the forest.” They are often described as short in stature and having a darker skin color than neighboring tribes. Today the few remaining Kusunda have intermarried with neighboring tribes and drifted apart, and the language has been moribund for decades, although a few elderly speakers with some knowledge of the language still survive.

The Kusunda language is a linguistic isolate, with no clear genetic connections to any other language or language family (4, 5). Curiously, however, it has often been misclassified as a Tibeto-Burman language for purely accidental reasons. Hodgson’s original description of the Kusunda language (2) also includes descriptions of various Tibeto-Burman (and its smaller neighbors), and the lesser Sundas (Timor, Alor, and Pantar). East of New Guinea, Indo-Pacific languages survive on New Britain, New Ireland, the Solomon Islands, Rossel Island, and the Santa Cruz Islands. They also were spoken in Tasmania until 1876. The distribution of Kusunda and the Indo-Pacific family is shown in Fig. 1. Although it is not possible with present evidence to demonstrate conclusively the direction of the migration that separated Kusunda from the other Indo-Pacific languages, it would seem at least plausible that Kusunda is a remnant of the original migration to New Guinea and Australia rather than a backtracking to Nepal from the region in which other Indo-Pacific languages are spoken currently.

Recently, two molecular genetic studies (11, 12) have found that the Andamanese belong to mtDNA haplogroup M, which is found also in East Asia and South Asia and has been interpreted as “a genetic indicator of the migration of modern Homo sapiens from eastern Africa toward Southeast Asia, Australia, and Oceania” (11). In addition, the Andamanese belong to the Asia-specific Y chromosome haplogroup D. Thangaraj et al. (11) conclude that “the presence of a hitherto unidentified subset of the mtDNA Asian haplogroup M, and the Asian-specific Y chromosome D, is consistent with the view that the Andamanese are the descendants of Paleolithic peoples who might have been widely dispersed in Asia in the past.” If molecular genetic evidence can be obtained from the few remaining Kusunda, it will be interesting to determine whether it supports the conclusions we have arrived at on the basis of their language.

**Grammatical Evidence**

Linguistic evidence on Kusunda is sparse, limited to just three sources (2–4), and there are some discrepancies between Hodgson’s 19th-century data and the late 20th-century recordings of Reinhard and Toba (3, 4). For example, Hodgson, using a simple English orthography, represents the Kusunda affricates as ch and j, indicating that he heard them as palatal: [c] and [j]. Reinhard and Toba, however, represent the affricates as [ts] and [dz] and state explicitly that they are alveolar, not palatal. In this article,
the source of each Kusunda form is identified as follows. Words from Reinhard and Toba (3) are taken as the default; words from Hodgson (2) are followed by (H); and words from Reinhard (4) are followed by (R). Sources for the other Indo-Pacific languages mentioned in this article are given in Supporting Appendix 1, which is published as supporting information on the PNAS web site.

Within this relatively small and imperfect corpus there is grammatical and lexical evidence pointing toward an Indo-Pacific affinity. The strongest piece of evidence is a pronominal pattern found in the independent pronouns (involving five different parameters) that is widespread in Indo-Pacific and also found in Kusunda in precisely the same form. These five defining features are: (i) a first-person pronoun based on t; (ii) a second-person pronoun based on n or g; (iii) a third-person pronoun based on g or k; (iv) a vowel alternation in the first- and second-person pronouns in which u occurs in subject forms and i in possessive (or oblique) forms; and (v) a possessive suffix -yi found on all three personal pronouns. It is significant that four of these five defining features have to do with the first- and second-person singular pronouns, which are known to be among the most stable elements of language over time (13). Indeed, it is such pronouns that have often been the first evidence for very ancient families such as Eurasiatic and Amerind.

In his original article defining the Indo-Pacific family, Greenberg (10) posited two basic pronominal patterns, n/k “I/you” and t/g “I/you,” and he suggested that the second set originally had a possessive function. However, subsequent research has cast doubt on the antiquity of second-person k, the distribution of which is largely confined to New Guinea itself. In any event, it is the second pattern that Kusunda shares with Indo-Pacific. One finds both ni and nti as the second-person pronoun; Greenberg surmised that nti had been the original form and had changed to ni in some languages as a simple sound change and in others to nti by analogy with the very widespread na “I” of the first pronominal pattern. Greenberg did not notice, however, in his pioneering article either the vowel alternation or the possessive suffix -yi. Table 1 shows the first-, second-, and third-person pronouns for Kusunda and selected Indo-Pacific languages.

In Kusunda the vowel alternation has only been preserved in the second person, having been eliminated through analogy in the first-person form. Furthermore, first-person “t-” has been palatalized to ch-, ts-, or tsch- (R) under the influence of the following i. Such a sound change is extremely common in the world’s languages, and in the present case we can be sure that the original consonant was t-, because t- has been preserved in both the object form tonu “me” and in first-person plural to-ʔi “we” (-ʔi is a plural suffix). In addition to the independent pronouns, the consonantal base also indicates the verbal subject: Kusunda t- “I,” n- “you,” g- “he,” Bea d- “I,” p - “you,” Onge g- “you,” g- “he,” West Makien nV- “I,” nV- “you,” and Brat t- “I,” n- “you.”

The other Indo-Pacific languages in Table 1 have preserved different portions of the original system. It is best preserved in the Andaman Islands (Juwoi, Bo) and North Halmahera (Galela), whereas in Western New Guinea (Seget, Karon Dori),...
New Britain (Kuot), and the Solomon Islands (Savasavo), only the consonants are preserved, in some cases only partially. The final language in Table 1, Bunak, is spoken on Timor and obviously does not preserve either the first-person t or second-person n/n; it does, however, preserve third-person g, and the possessive suffix is attached to all three pronouns, just as in Kuot.

Certainly this unique pronominal pattern shared by Kuot and Indo-Pacific languages cannot be a case of accidental convergence, because the probability that Kuot could have invented this particular pattern independently is vanishingly small. Borrowing is equally unlikely, because there is no evidence that Kuot has ever been in contact with any Indo-Pacific language.

Two other grammatical forms shared by Kuot and Indo-Pacific are demonstrative pronouns based on t and n.

**This.** Kuotan na (H) “this,” n= “that” = Indo-Pacific: Puchikwar ite, Jwuiw ete, Abuii (t)-do, Konda ete, I'ik ide, Biaka te? “this, he;” Kwotmari ita “he,” Timbe ida “this, that,” Selepet eda, Marind iti, Mininabai eti “he,” Humene ida

**Lexical Evidence**

Complementing the grammatical evidence are a number of lexical similarities that also point to an Indo-Pacific affinity. Some of the most convincing are given below. We do not give here all the supporting etymologies or all the supporting forms for each etymology. Rather, we have chosen for each etymology a sample of the forms from different regions of the Indo-Pacific family. The meaning of each form is the same as the head meaning unless specified otherwise.

**Breast.** Kuotan ambu = Indo-Pacific: Sawuy am, Korowai am, Wambon om, Ivori ammug, Gogodala om, Gaima om “milk,” Waia am, Giafo am, Wabuda am, Tebera am, Eki am, Chimb amu, Wahi am, Puri am, Yekora am, Yoda am, Koita am, Nene yama, Morawa am, Arawum ammu, Usu am, Kamba amu, Biyom am, Katiata am, Musak amu, Tani am, Wamembre emi.


**Table 1. An Indo-Pacific pronominal pattern**

<table>
<thead>
<tr>
<th></th>
<th>Kusunda</th>
<th>Juwoi</th>
<th>Bo</th>
<th>Galela</th>
<th>Seget</th>
<th>Karon Dori</th>
<th>Kuot</th>
<th>Savasavo</th>
<th>Bunak</th>
</tr>
</thead>
<tbody>
<tr>
<td>l</td>
<td>chi (H)</td>
<td>tui</td>
<td>tu-la</td>
<td>to</td>
<td>tet</td>
<td>tuo</td>
<td>-tuo</td>
<td>-tu</td>
<td>ne-tu</td>
</tr>
<tr>
<td>me</td>
<td>chi-yi (H)</td>
<td>tii-ye</td>
<td>tii-e</td>
<td>ji-“me”</td>
<td>nen</td>
<td>nüo</td>
<td>-nüo</td>
<td>-no</td>
<td>n-ie</td>
</tr>
<tr>
<td>you</td>
<td>nu (H)</td>
<td>gii-ye</td>
<td>gii-la</td>
<td>no</td>
<td>gao</td>
<td>go</td>
<td>0-IE</td>
<td>gi</td>
<td>g-ie</td>
</tr>
<tr>
<td>he/she</td>
<td>gida (H)</td>
<td>kiter</td>
<td>kiter</td>
<td>ni-“thee”</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>his/hers</td>
<td>gida-yi (H)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>gidi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


**Father.** Kuotan mam “older brother,” mam (R) “older brother, father’s sister’s older son, mother’s sister’s older son” = Indo-Pacific: Abui mana, Mii -mam, Arandai mane, Eipo pā “mother’s brother,” Dentma mana, Manambu mana “older brother,” Angoram man, Korowai mana “mother’s brother,” Huli mana “grandfather,” Kobon man “brother,” Kate ma-ni?, Kwale mana, Palubu mana, Saep mana, Jihim mana, Bongu mem, Kare momo-, Sihan meme-, Samosa mane-, Wamas mana, Garuh mana, Mugil -mam, Kuot mana, Baining mana, Taulil mana, Baniata mana.

**Fire.** Kuotan stay = Indo-Pacific: Isam ea, Bauzi ai, Gresi aya, Nimboran aya, Taikat aya, Yuni aya, Dera aya, Kwotmari aya?, Busa aya(?) , Amto aya, Urait aya, Yis aya, Seti aya, Wiaki yaye, Hewa aya, Amal aya, Siyage aya, Dibulug a’aya, Eki aya “great grandfather,” Sausi ai- “older sibling (same sex),” Danaru aya “older sibling (same sex),” Uru aya, Baniata ai.

**Give.** Kuotan aí (H), yu-gan, yo-wu “give” (imperative) = Indo-Pacific: Juwoi a, Jarawa a yu, Bale a, Brat e, Hatam -yai “take, give,” Senati ye, Manem ya, Elepi yu, Kamasau nge-

**Dog.** Kuotan agai (H), aqsi = Indo-Pacific: Woisika waggu, Sentani yoku, Grand Valley Dani yekke ~ yege, South Ngalik yege, Aghu yagg, Kaeti aqga, Yelmek agoa, Noraria aqa, Girda yauga, Siroi age.
“give it to me,” Wambon yo-, Riantana ya-, Maklewe -ai-, Gidra aia(o), Northeast Kiawai ai.


River. Kusunda wido‭-‬o “flow (noun)” = Indo-Pacific: Baham weja, Iha wadar, Puragi owoedo, Aikwakai wetai, Siagha wedi, Pisa wadi, Aghu widi, Kombai wodet, South Kati ok-wiri (ok-“water”), Awim waidou “Fly River.”


Tree. Kusunda i (H), yi, ii (R) = Indo-Pacific: Sentani i “fire,” Biaka yei? “fire,” Kwomtari i? “fire,” Rocky Peak yeu “fire,” Siagha yi, Kombai e, Girara e, Gogodala i, Kairi i “tree, fire,” Tumu ii, Kibiri i, Mena bi, Pawanai (n), Kasua ii, Pa i, Angataa i-pati [paiti (class prefix)], Fuyuge i-ye “tree, wood,” Zia ii, Yeletyme yi.


This article was written as part of the Santa Fe Institute’s program on the Evolution of Human Languages, directed by Murray Gell-Mann, and we gratefully acknowledge their support, the support of City University of Hong Kong (Grants 901001 and CERG 9040781 to W.S.-Y.W., Principal Investigator), where part of the article was written, and the invaluable assistance of Ruth Thurman at the Summer Institute of Linguistics office in Papua, New Guinea.