A Decipherment of Epi-Olmec Hieroglyphic Writing

John S. Justeson and Terrence Kaufman

The decipherment of part of the epi-Olmec script of ancient Mexico, which yields the earliest currently readable texts in Mesoamerica, has been achieved over the last 2 years. This was made possible by the discovery of a stela with a long inscription at La Mojarra, Veracruz, Mexico, in 1986. This decipherment is based on both a reconstruction of the early stages of languages spoken in the region and semantic clues provided by comparison with cultural practices and other script traditions of early southern Mesoamerica. Summarized here is the current state of the phonetic decipherment, the methods used for the decipherment, and results concerning the epi-Olmec language and script. The language identified in the inscriptions is pre–proto-Zoquean, the ancestor of four languages now spoken in the states of Veracruz, Tabasco, Chiapas, and Oaxaca. The decipherment contributes to knowledge of early Mixe-Zoquean language history. The script is more closely related to Mayan hieroglyphic writing than to other early Mesoamerican scripts, and this relation is closer than previously recognized. .

A distinctive script tradition flourished from about 150 B.C. to A.D. 450 in southern Mexico in the former heartland of Olmec civilization (Fig. 1). We refer to this tradition as epi-Olmec because the archeological cultures of the area descended from that of the Olmecs (1200 B.C. to 500 B.C.). The script may itself descend from an Olmec hieroglyphic system, but too little of the Olmec script has been recovered to confirm or disprove a connection. We outline our partial decipherment of this system and the evidence for this decipherment. The keys to our decipherment were (i) the discovery of a lengthy text in the script; (ii) the assumption that the texts were in Mixe-Zoquean languages; (iii) our analysis of the grammatical structures of available texts; (iv) an account of these structures in terms of the previously reconstructed grammar of proto–Mixe-Zoquean; and (v) clues to word meaning, from calendrical constraints and from comparison with similar Mayan signs, that enabled us to correlate spelled-out words with reconstructed proto-Zoquean and proto–Mixe-Zoquean vocabulary.

The Decipherment

Decipherment is a process of accounting for the patterns of sign use in texts. We summarize our account of epi-Olmec patterns in terms of the language represented in the texts, the principles by which this language was represented in the script, our readings

J. S. Justeson is in the Department of Anthropology, State University of New York, Albany, NY 12222, and is a visiting scientist at the IBM T. J. Watson Research Center. T. Kaufman is in the Departments of Anthropology and Linguistics, University of Pittsburgh, Pittsburgh, PA 15260. Name order is alphabetical; authorship is equal. Correspondence may be addressed to either author.

and interpretations for words and signs, and the content of the texts. The texts whose condition permits analysis are La Mojarra Stela 1 (A.D. 159; Fig. 2) and the Tuxtla Statuette (A.D. 162; Fig. 3); our decipherment makes them the earliest currently readable texts in Mesoamerica. The stela depicts an epi-Olmec warrior-king. Its text, as we read it, provides a lengthy description of his rise to kingship through several years of warfare and ritual activity. It is unusual in its emphasis on the roles of his supporters. The statuette describes the ritual activities of a probable shaman, and its text ends with him calling up his animal spirit com-
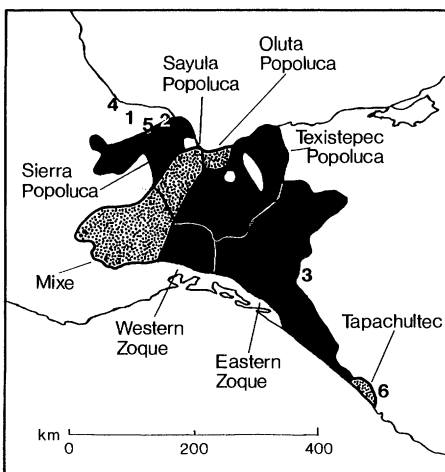

Fig. 1. Sources of epi-Olmec inscriptions and the distribution of Mixe-Zoquean languages of southern Mexico at the time of the Spanish invasion, approximately A.D. 1519 to 1525. Areas of Zoquean speech are indicated in black; areas of Mixean speech are stippled. Sites from which epi-Olmec and related texts have been recovered are: 1, La Mojarra; 2, San Andrés Tuxtla; 3, Chiapa de Corzo; 4, Cerro de las Mesas; 5, Tres Zapotes; and 6, El Sitio.

panion. The statuette appears to depict this or his impersonation of this companion, for it shows a human being in the guise of an animal, wearing a duckbill mask and a cape of bird wings and claws. Although they are from different sites, these two texts are closely related. The statuette postdates the stela by less than 3 years. Its shaman is apparently mentioned twice on the stela in association with the accession rituals of the La Mojarra ruler. The rites described on the statuette were performed on the sixth anniversary of the ruler's last recorded battle before accession.

The language represented in these texts is an early form of Zoquean. Zoquean is a branch of the Mixe-Zoquean language family (Fig. 4), and seven Mixe-Zoquean languages exist today, spoken by 100,000 to 140,000 native Americans. The four modern Zoquean languages descend from proto-Zoquean, spoken about 1400 years ago. We identify epi-Olmec texts as pre–proto-Zoquean (ancestral to proto-Zoquean) by relating textual evidence to Kaufman's (1) reconstruction of Mixe-Zoquean language history. This reconstruction is based on a comparison of the modern Mixe-Zoquean languages, whose attestation in European script goes back to the 17th century.

Proto-Zoquean, proto-Mixean, and proto–Mixe-Zoquean contained 11 consonants (p, t, k, tz, $?$, s, j, m, n, w, y) and 6 vowels (i, e, $ʉ$, a, u, o); Mixean, but not Zoquean, had both long and short vowels (2). (We introduce reconstructed grammatical characteristics as they become relevant to the discussion.) Epi-Olmec texts were written in columns, normally read from left to right (3), and signs usually occupy their column's width. Whenever signs abut, they form together a full word or phrase. More often, successive signs, even within words, are separated by space; words can even be split across columns.

Most epi-Olmec signs are syllabograms, signs that represent a syllable, or logograms, signs that represent words or morphemes (4). All or most epi-Olmec syllabograms represent consonant-vowel (CV) syllables—that is, a consonant and following vowel. Those we can read with assurance are given in Fig. 5A and cover 21 of 66 possible syllabic values. Worldwide, all syllabaries adequately spell the CV sequences of spoken words with the use of the corresponding CV sign, as when ja-ma spells *jama 'day' (Fig. 6L). With the use of only CV syllabograms, however, the pronuncia-

tion of the word and the straightforward reading of the sign sequence necessarily conflict whenever a consonant in the language precedes a consonant or ends a word. This conflict is resolved in one of two ways. (i) The consonant may not be spelled, as in *we-pa* for *wej-pa 'he shouts' (Fig. 6G). (ii) If the consonant is spelled, the vowel of the CV sign is not pronounced, as in *ʔi-ki-p(i)-wʉ* for *ʔi-kip-wʉ 'he fought him/them' (Fig. 6A; unpronounced vowels are transliterated in parentheses). In epi-Olmec writing, all consonants are written when they immediately precede a vowel, and this is the only context in which "weak" consonants (*ʔ, j, w,* or *y*) are written. We have too little evidence to determine whether "strong" consonants may sometimes not be written. The vowel in a CV spelling for a nonprevocalic con-

sonant is the last preceding vowel, which is called the vowel synharmony convention. When a word ending in *ʔV* (word-final) immediately precedes a word that begins with *ʔV* (word-initial) and the vowels are the same, then the word-initial *ʔV* may perhaps be omitted in spelling.

Some signs perhaps represent syllables closed by a consonant (Fig. 5C). One likely example, MS114 (5), may represent different morphemes that have the same pronunciation, *tuk 'to happen' and *tuk 'to harvest'. Others represent not only morphemes but also syllables or phoneme sequences within morphemes. One possible example of this is MS130, used for the syllable *tzaʔ* in the root *matzaʔ 'star' (6). Less securely, MS53 (in 13-MS53-*ma*) is perhaps used for the sequence *jam* in *jama 'day' and to spell *jam 'gray, ashes'; MS54 is perhaps used for

*kak 'to exchange', which relates to succession in office, and for the syllable *kak* in the root *kakpe 'scorpion' (as a constellation name).

Grammatical affixes are spelled by syllabograms. The most common grammatical affixes consist of a CV sequence and are spelled by a CV sign—that is, *ʔi* for the pronominal prefix **ʔi-* and *wʉ* for the verbal suffix **-wʉ* in *ʔi-ki-p(i)-wʉ* for *ʔi-kip-wʉ 'he fought him/them'. Affixes that consist of a single vowel are spelled by a CV sign that conveys that vowel and the consonant preceding it—that is, *ʔaw*-RULE-*mʉ* for **ʔaw=kiʔm-ʉ* 'rulership, authority' (Fig. 6K). Grammatical affixes ending in a weak consonant can be spelled by a CV sign—for example, **-jay* 'indirective' by *ja* (7), **-kʉʔ* 'inceptive' by *kʉ* (8), and **-kuy* 'instrumental' by *ku* (9). Grammatical affixes that end in strong consonants were apparently spelled syllabically in the same way as a root of the same pronunciation—that is, *tʉ-n(ʉ)* spells the pronominal prefix **tʉn-*.

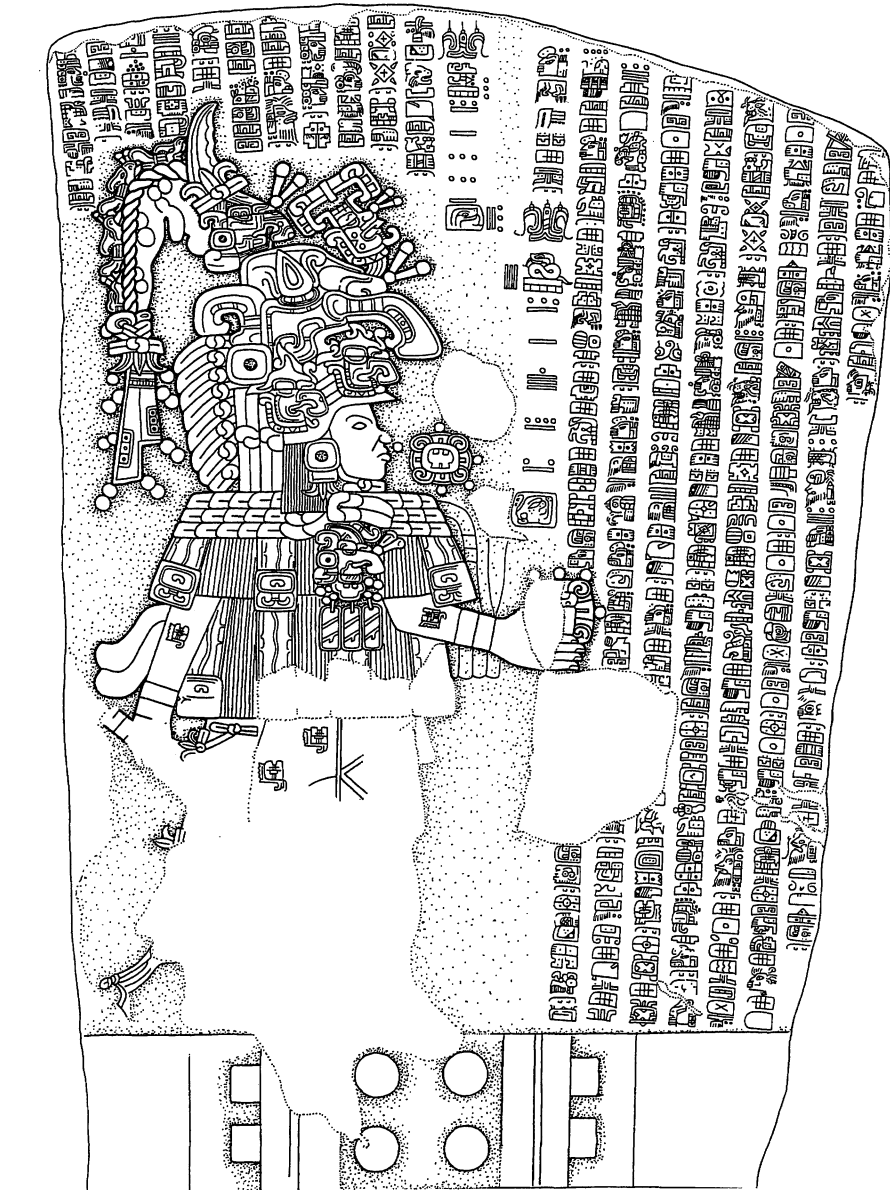Some epi-Olmec logograms represent an uninflected root or stem (*JAMA* for *jama
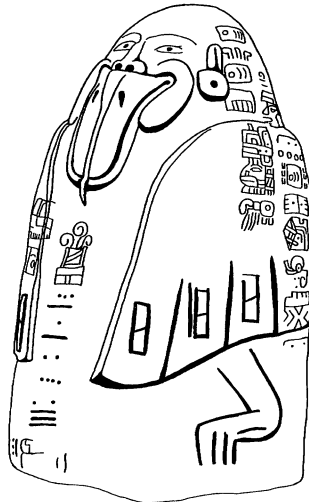

Fig. 3. The Tuxtla Statuette.
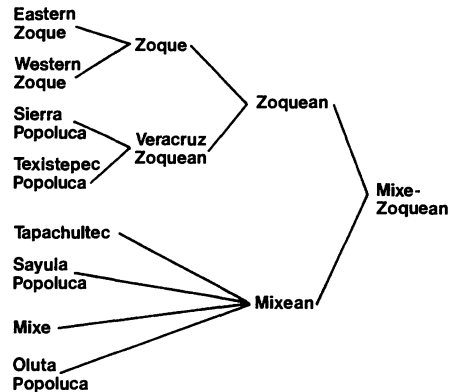

Fig. 4. The Mixe-Zoquean language family.

Fig. 2. Drawing of La Mojarra Stela 1.

'day; animal spirit companion' (10) (Fig. 6N). Some can designate a morpheme within a word; for example, RULE designates *ki?m 'to ascend' in ?aw-RULE-mʉ for *?aw=ki?m-ʉ 'rulership, authority'. In some cases, a sign that spells a verb stem is used with the phonetic value of the verbal noun that corresponds to the verb stem— that is, JU?TZ 'to pierce' is also read ju?tzi 'when'; KI?M 'to accede' is also read ki?m-ʉ 'accession'. Syllabograms can be preposed or postposed (or both) to a logogram either to indicate the initial or final portion of its pronunciation, as in ma-STAR-tza? for *matza? 'star, planet' (Fig. 6M), or to spell grammatical affixes.

Occasionally, spelled-out words are used pictorially, not linguistically. For example, one sign cluster that designates a geographic term (perhaps 'town') consists of WATER alongside a sloped sign (apparently a mountainside) surmounting EARTH (Fig. 7A). Moreover, some titles or offices are designated by a spelled word in the headdress of a profile face sign. For example, logosyllabic (partly logographic and partly syllabic) STAR-tza? (Fig. 7C) *matza? 'star' stands for a warrior title [Fig. 7, D and E; the battles recounted at La Mojarra were timed according to the cycle of Venus (11), as was customarily done among the Mayans], and a knotted cloth (Fig. 7F) around a bundle stands for a title of supreme authority (Fig. 7G; in both epi-Olmec and Mayan writing, the knotted cloth refers to royal accession).

We have epigraphic evidence for both the pronunciation and the meaning of several epi-Olmec logograms (Fig. 8A). For others, we have evidence for meaning but

Fig. 5. (A) The epi-Olmec syllabary. All the numbers next to the signs are Macri-Stark numbers (5). Most of the readings here are based on multiple contexts and highly probative semantic or grammatical constraints (or both); the *po* and *tʉ* readings for MS122 and MS102 are each proposed on the basis of both grammatical and semantic controls, but each occurs in just one context. The functional equivalence of MS38 and MS39 was previously postulated (19). (B) Epi-Olmec syllabograms of uncertain reading. (C) Proposed CVC syllabograms.

not, epigraphically, for pronunciation (Fig. 8B). In some of these cases, we can still be reasonably sure of the Zoquean word represented; one case is the epi-Olmec sign YEAR, given that descendants of proto–Mixe-Zoquean *ʔame are the only native words for 'year' in any Mixe-Zoquean language.

Inference Methods

Phonetic decipherment requires interlocking clues to sign pronunciation. In Old World decipherments, such clues came mainly from proper names, from bilingual texts, or both, with versions in the undeciphered system and in a known language and script. In the decipherment of Egyptian, the Greek text on the Rosetta Stone allowed both Thomas Young (in 1814) and Jean-François Champollion (in 1822) separately to determine semantic correspondences between words spelled in its Egyptian text and known Greek words; comparison with Coptic, the descendant of Egyptian, could then be used to exploit the semantic clues. Proper names were useful because in the Old World they were commonly borrowed roughly as pronounced; their forms as reported in other languages resemble those in the undeciphered script. Thus, when the hieroglyphic spellings for the names of Ptolemy, Cleopatra, and Alexander were recognized through their correspondence with these names in the Greek versions of Egyptian-Greek bilingual inscriptions, the Egyptian signs spelling out these names could be assigned an initial set of approximate phonetic values.

These aids were not available for the decipherment of epi-Olmec writing. There are no known bilingual texts from ancient Mesoamerica. Proper names were usually translated when borrowed, and in any case no epi-Olmec proper names were known to survive (although decipherment has led us to recognize a god name 'Ten Sky', known also in Mayan). The epi-Olmec decipherment therefore required a prior, correct identification of the language being written (pre–proto-Zoquean). It also depended on our assumptions concerning the representational principles of the writing system and on the availability of enough textual data to warrant analysis. In addition, some of the content of the text could be inferred on the basis of practices in other early Mesoamerican writing systems.

It was the discovery of Stela 1 of La Mojarra, Veracruz, Mexico (12; Fig. 2), that made the decipherment of epi-Olmec writing possible. Its lengthy, mostly well preserved text was the first to provide enough data so that multiple lines of evidence could constrain the reading of a large number of signs. Nine other probable texts in the script were already known. One, on

the Tuxtla Statuette (Fig. 3), we partly analyzed grammatically, and it has been quite useful in the decipherment of the script. The other texts are not susceptible to grammatical analysis. The two epi-Olmec texts from Chiapa de Corzo are fragmentary (13), and the four from Cerro de las Mesas (stelas 3, 6, and 8 and the Chapultepec Stone), Tres Zapotes Stela C, and the Alvarado Stela are mostly effaced (14). Related scripts occur on two portable objects, a celt from El Sitio, Guatemala

Fig. 6. Various epi-Olmec spellings. (**A**) ʔ *i-ki-p(i)-wʉ*, spelling ʔ *i-kip-wʉ* 'he fought him/them'. (**B**) SCATTER-*wʉ* ʔ*i*-MS58, spelling 'his MS58 got scattered'. (**C**) *ko-juʔtz-kʉ-wʉ* ʔ*i*-RULE-*si-ʔi*, spelling *ko-juʔtz-kʉ-wʉ* ʔ*i-kiʔm-ʉ=siʔi* 'his royal privates began to get pierced on behalf of others'. (**D**) *20-s(i)*, spelling ʔ*ips* 'twenty'. (**E**) ʔ*i-s(i)*, spelling ʔ*is* 'lo!'. (**F**) *we-ne*, spelling *wen-e* 'part(s)'. (**G**) *we-pa*, spelling *wej-pa* 'he shouts'. (**H**) Probably ʔ*o-tu-pa*, spelling ʔ*otuw-pa* 'he speaks'. (**I**) MS146a-SKY-*pa*, spelling *tzap* 'sky'. (**J**) [...] ʔ*i*-BLOODLETTING PENIS-INSIDE-*mʉ-k(ʉ₂)* TITLE₁ *wʉ*, spelling 'he who is TITLE₁ [underwent] his bloodletting from within the penis'. (**K**) ʔ*aw*-RULE-*mʉ*, spelling ʔ*aw=kiʔm-ʉ* 'rulership, authority'. (**L**) *ja-ma*, spelling *jama* 'day'. (**M**) *ma*-STAR-*tzaʔ*, spelling *matzaʔ* 'star, planet'. (**N**) JAMA, spelling *jama* 'day'. Strings of symbols that have a glyph in common are arranged so that they are adjacent, either vertically or horizontally.

(15), and a ceramic mask of unknown provenience.

The previously known texts had been the basis for limited analysis (16, 17) that indicated that epi-Olmec writing was more closely related to Mayan than to other Mesoamerican scripts, that its numeral and calendrical systems were much the same as Mayan, and that its texts preserved the columnar format that was modified in the Mayan-Izapan area to a more complex system of paired columns. Obvious similarities in form and context between Mayan and epi-Olmec numeral and calendrical signs showed that the script contained logograms. The most important noncalendrical proposals were Ayala's suggestions [as cited in (17)] that MS20 is a clause-ending sign and MS38 a clause-initial sign on the basis of their frequent positions at the end and beginning of the columns.

Our initial working hypothesis—the only one we considered plausible—was that epi-Olmec texts were written in a Mixe-Zoquean language; almost all recent work on epi-Olmec writing makes this assumption
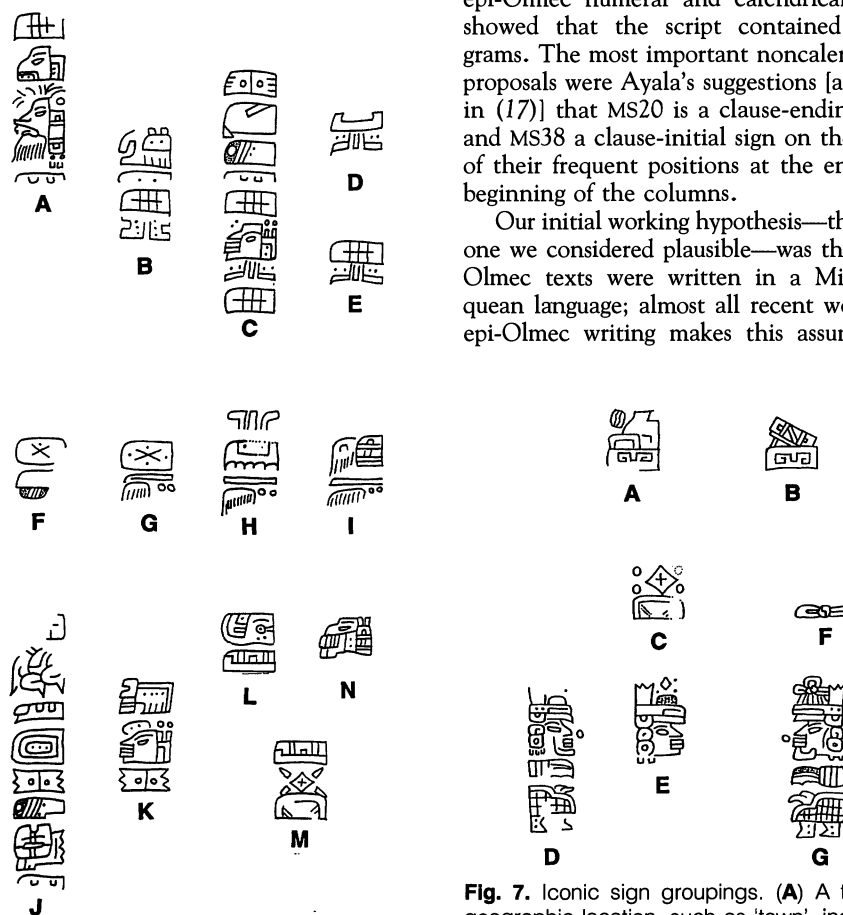

Fig. 7. Iconic sign groupings. (**A**) A type of geographic location, such as 'town', indicated by iconic signs for water, earth, and unidentified elements. (**B**) Sunrise or sunset, indicated by sky and earth signs in the orientation of the Mayan "sun-at-horizon" compound. (**C**) A spelling STAR-*tzaʔ* for **matzaʔ* 'star, planet', which appears in the headdress of the profile face sign in (D) and (E), where it evidently indicates a title or office related to war. (**D**) Title and name of protagonist of La Mojarra Stela 1 text. (**E**) Title of the protagonist of La Mojarra Stela 1 text during military campaigns before accession. (**F**) Knotted cloth, used in Mayan and epi-Olmec texts in reference to accession to rulership evidently through the tying on of authority symbols to the forehead; it appears in the headdress of the profile face sign in (G), where it evidently indicates a title or office of supreme authority. (**G**) Title and name of the protagonist of La Mojarra Stela 1 on the date of his accession. Signs are arranged on the same principle as used in Fig. 6.

(*12, 18, 19*). At the time of the first European records, Mixe-Zoqueans were present in and near the region from which the script is attested (Fig. 1), and speakers of all other languages in the region were intrusive. Further, the correlation of cultural artifacts and practices widely diffused throughout Mesoamerica with vocabulary diffused into other Mesoamerican languages from Mixe-Zoquean indicates the Mixe-Zoquean linguistic identity of the Olmecs (*20*) and thus, presumably, of their descendants.

Both Mixean and Zoquean were, initially, viable candidates for being written in epi-Olmec texts. Speakers of both were involved in the cultural diffusion associated with Olmec and epi-Olmec civilization. Geographically, no Mixean language is known to have been spoken in the region where our texts were found. However, the Mixeans appear to be in each of their present locations because of dislocation, so

we could not assume that they never occupied this region. According to Kaufman's glottochronological estimates (*1*), on the basis of the amount of basic vocabulary shared among various of these languages neither proto-Zoquean nor proto-Mixean split into multiple Zoquean or Mixean languages until about 400 years after our texts were inscribed. We therefore began decipherment entertaining pre–proto-Zoquean and pre–proto-Mixean as viable alternative hypotheses for the language or languages of the epi-Olmec texts. We conclude from the vocabulary of the inscriptions that the language of both texts is pre–proto-Zoquean.

We assumed epi-Olmec writing to be logosyllabic. Logograms for time periods were already known, and our interpretation of Ayala's proposed "clause-final" and "clause-initial" signs as grammatical affixes suggests phonetic usage. We further assumed that simple phonetic signs would

represent at least and perhaps only CV syllables. All syllabaries have CV syllabograms—many have no others—and their CV series are always at least as complete and usually much more complete than other syllable series. All Mixe-Zoquean syllables are consonant-initial, so all syllabograms are consonant-initial as well.

With the use of these assumptions, reconstructed Mixe-Zoquean patterns of grammatical affixation provided us with a systematic framework for phonetic decipherment. Such patterns provide clues to pronunciation because grammatical affixes are almost always spelled by phonetic signs; we first read several epi-Olmec signs by means of their roles as grammatical affixes. The other major sources of evidence are the reconstructed vocabularies of proto-Mixean, proto-Zoquean, and proto–Mixe-Zoquean. We could exploit them because clues to the meanings of some written words

A

	104	?ips	'twenty' this sign is also used for 'moon'
	172	jama	'day; animal spirit companion'
	164	jama	'animal spirit companion'
	23	ju?tz	'to pierce'
	24	ju?tzi	'how, when, as' relative
	171	ki?m	'to accede' this sign is also used for 'rule' and other words derived from ki?m
	57	komi	'boss, lord'
	10	mak	'ten', in deity name Ten Sky
	31V 32	matza?	'star, planet'
	137	poy?a	"month" of 20 days
	50	tokoy	'to get lost or ruined'
	158	tuki	'turtle'
	136/ 144	tzap	'sky'
	150	wik	'to sprinkle'

B

Ritual-related

	91	to perform (ritual)
	22	blood
	132	to let blood
	107	penis
	115	to chop
	147	to cast/scatter
	58	droplets or grains
	99	liquid, drops

Calendric

	72	year (presumably ?ame); drum (presumably kowa)
	104	moon (presumably poy?a)
	89	day count indicator
	106	month name/patron XIII
	110	month name/patron XVI
	156	month name/patron XVIII
	154	day name Deer
	131	day name Earthquake
	155	day name Snake

Kingship

	181- 184	royal title determiner
	129	royal accession
	60	title/office
	61	title/office
	62	title/office
	76 +78	title/office
	51	split conch shell (so-called "knuckle dusters")
	68	symbol-bearing object (possibly kipsi)
	75	hallowed
	83	symbol of status/role
	112	throne

Miscellaneous

	134 +44	geographic location (town?)
	119	mountain
	40- 41	to get pounded
	120	inside (possibly joj)
	33	hide, skin (presumably naka)
	36 +44	sunset

Fig. 8. (**A**) Some epi-Olmec logograms of known reading. All numbers next to the signs refer to Macri-Stark numbers (*5*). (**B**) Some epi-Olmec logograms of known meaning. The iconographic identification of MS33 with hide is due to Winfield Capitaine (*30*).

Table 1. Mixe-Zoquean tense-aspect-mood suffixes. *-ʔi is Zoquean only.

Form	Meaning
*-ʉ ~ *-a	Imperative
*-wʉ	Independent completive (X VERBed)
*-pa	Independent incompletive (X VERBS)
*-ʔi	Optative (X should VERB)
*-ji	Dependent completive (when X VERBed)
*-i ~ *-e	Dependent incompletive (when X VERBS)

can be inferred from patterns of dates, from similar logograms and symbols in the writing and iconography of other early cultures of southern Mesoamerica, and from correspondences between certain syntactic patterns and semantic properties.

Pre–proto-Zoquean verbs are required to take one of six tense-aspect-mood suffixes (Table 1); all other verb suffixes are optional. We have identified all six in the La Mojarra text. The independent completive suffix *-wʉ is easily the most common in Mixe-Zoquean, and it occurs frequently in narrative texts. The syllable *-wʉ would be spelled by a CV syllabogram, which would be very frequent at the ends of words and rare elsewhere in any epi-Olmec text. MS20, the most common sign in our texts, is always or almost always word-final—it ends repeating sign sequences and many columns—and its word-final frequency far exceeds that of any other sign. It can hardly be other than a syllabogram for wʉ (24; Fig. 6, A through C).

The syllable *wʉ has one other grammatical use in Mixe-Zoquean: as a relativizing morpheme after nouns (N) and adjectives (A), translating as 'that which is N/A' or 's/he who is N/A'. MS20 is so used when it follows designations for titles or offices in references to the ruler—for example, 'he who is TITLE₁ [underwent] his bloodletting from within the penis'—and such a function accounts for all instances of wʉ after sign sequences we interpret as nonverbs. This twin usage of MS20 not only confirms the reading but is also a signature of Mixe-Zoquean grammar, confirming the hypothesized language identification.

MS143 also appears to be suffixed to independent verbs (Fig. 6, G and H), which suggests the incompletive marker *-pa. A pa reading for MS143 is confirmed by the spelling SKY-pa (Fig. 6I); the only Mixe-Zoquean word for 'sky' (proto–Mixe-Zoquean/proto-Zoquean *tzap) ends in p and has the root vowel a.

MS38 is by far the most common grammatical prefix in epi-Olmec texts, preposed optionally to both verbs and nouns and usually as the first sign in the word. Its only feasible Mixe-Zoquean interpretation

is as the third person ergative (22) pronominal prefix ʔi-, both as the subject (he/she/it/they) of a transitive verb or dependent intransitive verb and as possessor (his/her/its/their) of nouns; it was the second sign we read phonetically. Grammatical alternations between it and MS44 established the latter as also an ergative pronominal prefix, almost certainly first person exclusive and thus *na-. The possessive interpretation of these signs is confirmed in several contexts; for example, na is prefixed to the logogram PENIS and words for body parts usually have possessive prefixes in Mixe-Zoquean. The phonetic reading of ʔi is also confirmed by phonetic spellings discussed below, and the phonetic reading of na is confirmed by its acrophonic origin from a descendant of proto–Mixe-Zoquean *na:s (proto-Zoquean *nas) 'earth', which the sign represents iconographically (23).

A peculiarity of Mixe-Zoquean languages is that a substantial proportion (perhaps 15%) of verbs are bivalent—that is, they occur both as transitive and intransitive verbs without special affixes to transitivize or detransitivize them. As intransitives, they usually have a mediopassive sense ("X got VERBed"). These bivalent patterns do occur at La Mojarra, and they constitute further evidence for the Mixe-Zoquean identity of the texts.

We also recognized several less common affixes. A suffix *-mʉ is required on nouns in locative or ablative functions (*-mʉ 'to, at'; *-mʉ-k 'from'), with a nominal relational location root (for example, 'inside') between them. This almost forces the readings mʉ, kʉ₂, and INSIDE for MS64, MS124, and MS120 in [. . .] ʔi-BLOODLETTING PENIS-INSIDE-mʉ-k(ʉ₂) TITLE₁ wʉ 'he who is TITLE₁ [underwent] his bloodletting from within the penis' (Fig. 6J). The mʉ reading is confirmed by its only other occurrence, in the word ʔaw-RULE-mʉ for ʔaw=kiʔm-ʉ 'authority' (Fig. 6K). Other affixes identified are (i) markers of subordinate status on verbs (in the completive *-ji and in the incompletive *-i); (ii) indirective suffix *-jay and inceptive suffix *-kʉʔ; (iii) possessive and ergative first person inclusive prefix *tʉn-; (iv) suffixes *-e and *-ʉ, which derive nouns from verbs, and *-na, which derives adverbs from verbs; and (v) a plural marker *-taʔm. Recognition of these affixes substantially aided further decipherment. By distinguishing among transitive verbs, intransitive verbs, and nouns, which have differing affixation patterns, we were able to determine much of the grammatical structure of the texts and to circumscribe the meanings of many words.

Patterns of dates provide evidence for the meanings of several signs and sign sequences (11, 24). MS32, for instance, refers to 'star, planet'. Resembling the

Table 2. Epi-Olmec vocabulary descended from proto–Mixe-Zoquean. Uncertain forms are those that have been provisionally identified in the texts with readings from Fig. 5B.

Form	Meaning
	Grammatical morphemes
ʔi-	Third person, possessive and ergative
ko-	'on behalf of others'
na-	First person exclusive, possessive and ergative
tʉn-	First person inclusive, possessive and ergative
-a	Nominalizer
-e	Nominalizer
-i	Incompletive aspect, dependent
-ji	Completive aspect, dependent
-kuy	Instrumental
-kʉʔ	Inceptive
-mʉk	Locative case suffix string, 'from'
-na	Adverbializer
-pa	Incompletive aspect, independent (intransitive)
-taʔm	Plural
-wʉ	Completive aspect, independent
-ʉ	Imperative
-ʉ	Nominalizer
teʔ	Third person pronoun or demonstrative
wʉ	Relativizing enclitic on nouns, adjectives
	Lexical items
ʔis	'to see; lo!'
kij	'to shine'
kiʔm	'to ascend'
kuy	'wood, tree'
kʉʔ	'hand'
mak	'ten'
matzaʔ	'star, planet'
neʔw	'to set stones in order'
nʉks	'to go along'
poyʔa	'moon; 20-day "month"'
saj	'to share out'
tzap	'sky'
wen-e	'part'
	Uncertain grammatical morphemes
-i	Nominalizer
jeʔ	Demonstrative 'that'
yʉʔ	Demonstrative 'this'
	Uncertain lexical items
nas	'to pass by'
naks	'to whip, to beat'
wik	'to sprinkle'

Mayan 'star' logogram, its association with two dates on the La Mojarra text, separated by exactly 9 canonical, synodic Venus years (9 × 584 days), secures the STAR interpretation. Similarly, MS165-63 is a noun that occurs twice in a passage that fails to mention just 2 out of the 4802 days elapsed between successive stated dates, and it has a numeral prefixed in a separate calendrical context. The sequence apparently refers to a period of 1 day. The normal Zoquean word for 'day' is *jama, which suggests a ja-ma reading for the sequence (Fig. 6L). ma for MS63 is confirmed, being prefixed to STAR (Fig. 6M): the only native words for 'star' in Mixe-Zoquean languages descend

Table 3. Epi-Olmec vocabulary specifically associated with the Zoquean branch. All forms given are reconstructible for proto-Zoquean.

Form	Meaning
Grammatical morphemes	
-ˀi	Optative
Lexical items	
ˀaw=kiˀm-ʉ	'authority, rulership'
jama	'day; animal spirit companion'
juˀtz	'to pierce'
juˀtzi	'how, as, when' (relative)
kip	'to fight'
maw	'to go away'
paki	'hard, strong, powerful'
pak-kuy	'bludgeon'
teˀn	'to stand'
tuki	'water turtle'
wej	'to shout'
Uncertain grammatical morphemes	
-pa	Incompletive aspect, independent (transitive)
Uncertain lexical items	
ˀo	'maguey'
ˀotuw	'to speak'
jetzi	'thus'
kot-e	'container'
masa(n)=wik-i	'holy result of sprinkling', that is, a consecrated prince
miˀks	'to quiver'
pini	'brother-in-law of man'
tuku	'garment, cloth'
tziˀks	'to itch'

from proto–Mixe-Zoquean *ma:tzaˀ (proto-Zoquean *matzaˀ). STAR is followed in each case by a sign we read as tzaˀ; a similar Mayan sign depicts and designates 'stone' (proto–Mixe-Zoquean *tza: ˀ; proto-Zoquean *tzaˀ).

We read MS101 as *si* on the basis of calendrical, grammatical, and other contextual evidence. It occurs immediately after the numeral **20** (proto–Mixe-Zoquean *ˀi:ˀps; proto-Zoquean *ˀips), indicating the final consonant of the word (Fig. 6D). In the word ˀi-s(i) (Fig. 6E), it precedes statements of the form 'in *n* years', which suggests interpretation as a temporal adverb; proto–Mixe-Zoquean *ˀis is a sentence adverb meaning 'lo!'. The fact that both of these cases in which the sign spells a word-final *s* contain the same root vowel *i* suggests the vowel synharmony convention for spelling consonants not followed immediately by vowels and thus that the vowel conveyed by the sign is *i*. This last inference is supported by the apparent use of the sign in part to indicate the (incompletive) dependent status marker *-i (suffixed to verbs whose root vowel is *i*, *ʉ*, or *u* and whose final consonant is *s*). Finally, the spelling *si-ˀi*, apparently for *siˀi* ('backside' in Eastern Zoque but for which we suspect an earlier or additional meaning of "private parts"), appears in a statement: *ko-JUˀTZ-kʉ-wʉ ˀi-RULE-si-ˀi, ko-juˀtz-*

kʉˀ-wʉ ˀi-kiˀm-ʉ=siˀi 'his royal privates began to get pierced on behalf of others' (Fig. 6C). This reading is appropriate contextually in that it almost immediately precedes the statement that the ruler drew blood from his penis in a blood sacrifice ritual.

Discourse pragmatics provides clues to the meanings of some words. The narrative in the La Mojarra text switches at least four times from third person to first person references to the ruler. Such a switch requires a verb of speaking, as in "he says, 'when I chopped it [off], he departed.' " The four switches do follow two verbs, MS25-*pa* and MS97-93-*pa* (Fig. 6, G and H), which we therefore take to be verbs of speaking. MS25 spells the root of the first verb. Every Mixe-Zoquean verb root ends in a consonant and would be spelled syllabically by two CV signs unless the final consonant was not spelled. Among reconstructible Mixe-Zoquean verbs of speaking, only proto–Mixe-Zoquean *wej* 'to shout' ends in a consonant likely not to be written. This suggests the reading *we* for MS25, supported by its use in *we-ne* for *wen-e* 'part(s)' (Fig. 6F), for whose meaning there is contextual support. We have not conclusively identified the second verb, but some evidence supports a reading as *ˀotuw*, with MS97 for *ˀo* and MS93 for *tu*.

Iconic transparency provides a meaning for some signs, though not a specific word, provided the interpretation is verified by another line of evidence. The blood sacrifice verb is indicated by an iconic depiction of a figure in the posture of genital bloodletting, with blood gushing forth; it is verified by an iconic sign PENIS immediately after the BLOODLETTING sign (25; Fig. 6J). MS147 SCATTER, which depicts a casting hand, means 'scattering/sprinkling' in Mayan. In SCATTER-*wʉ ˀi*-MS58 (Fig. 6B), its epi-Olmec meaning is verified by a following possessed noun MS58 that depicts a flowing substance as in Maya iconography, followed by a string of titles belonging to the possessor of the scattered substance, followed by *wʉ* (for 's/he who is . . .'). Grammatically, the phrase is unambiguously interpretable, but we do not know the actual words for the verb, the possessed noun, or the titles involved.

Implications

Our decipherment of epi-Olmec hieroglyphic writing has allowed us to identify the epi-Olmec language, to improve our understanding of Mixe-Zoquean language history, and to demonstrate specific similarities between Mayan and epi-Olmec writing. We have shown that the grammatical structure of epi-Olmec texts is Mixe-Zoquean. The vocabulary of the texts narrows

Table 4. Epi-Olmec vocabulary specifically associated with Eastern (Chiapas) Zoque only. Other branches of Zoquean are less adequately documented, so the forms below may also be proto-Zoquean.

Form	Meaning
Lexical items	
komi	'lord'
siˀi	'backside, privates'
Uncertain lexical item	
tusi	'with hair standing on end'

it more specifically to Zoquean: some epi-Olmec words are attested in Mixe-Zoquean languages generally (Table 2) and many only in Zoquean languages; none is specifically Mixean (Tables 3 and 4). These results are consistent with and supportive of the observations and claims made by Campbell and Kaufman (20), which indicate that Olmecs spoke Mixe-Zoquean languages.

One grammatical feature now limited to Mixean appears in the La Mojarra text, the suffix *-ji for the completive aspect of dependent verbs. It must have formerly existed in Zoquean and been lost before the era of proto-Zoquean, the last common ancestor of modern Zoquean languages; the text must be in a pre–proto-Zoquean language ancestral to all modern Zoquean languages or in a separate, now-extinct branch of Zoquean. This confirms our working hypothesis of either pre–proto-Mixean or pre–proto-Zoquean.

Decipherment has enabled us to improve the reconstruction of some features of Mixe-Zoquean language history. We learned, for example, that in Zoquean the verbal system has been simplified, with a single suffix *-i ~ *-e marking dependent status; this suffix was extended from the incompletive aspect of dependent verbs to the completive, replacing proto–Mixe-Zoquean *-ji. We also verified some features of the reconstruction that deviated from what survived. All the modern languages have a basic word order in which verbs precede both subjects and objects, whether as agents or patients, when they are full nouns or noun phrases (not pronouns). However, there were several indications that this order was adopted under the influence of other Mesoamerican languages, almost all of which have this order, and that agents and objects formerly preceded the verb. In epi-Olmec texts, nominal agents and objects do precede the verb in both transitive and intransitive sentences; however, in intransitive sentences with a mediopassive sense for the verb the subject may instead follow the verb.

A close relationship of the epi-Olmec and Mayan scripts has previously been recognized, largely on the basis of spellings of

calendrical constructs (19, 26) and of the systematic patterns of formal correspondence between them (27). Many of the phonetic values that the two systems share, however, appear to have separate histories. The values of most Mayan CV signs agree acrophonically with the initial consonant and vowel of a Lowland Mayan word for what the sign depicts—for example, hieroglyphic **ka** based on *kay* 'fish', **a** based on *ahk* 'turtle', **na**₃ based on *na?* 'mother', and **ne** based on *ne:h* 'tail'. Similarly, several epi-Olmec CV signs agree with Zoquean vocabulary: epi-Olmec **ki** with **ki?*m 'to ascend' **ku** with **kuy* 'wood, tree', **nʉ** with **nʉ?* 'water', **na** with **nas* 'earth', **ne** with **ne?w* 'to set stones in order', and **po** with **pomo* 'copal incense'.

This suggests that some signs that appear to be shared by the epi-Olmec and Mayan scripts had different phonetic values, locally developed on the basis of Mixe-Zoquean and Mayan vocabulary for the same depicted entity or logographic value. A shared sign depicting a shiny stone is used in epi-Olmec writing for the syllable **tza?** or perhaps **tza**, based on proto–Mixe-Zoquean **tza:?* (proto-Zoquean **tza?*), and in Mayan for *tu:n* 'stone; year'. The epi-Olmec RULE logogram is read **ki?m* 'to ascend' in spelling ** ?aw=ki?m-ʉ* 'rulership, authority'. This logogram resembles a Mayan sign of similar form, read *a:ja:w* 'lord, ruler' and *a:jw-a:l* 'rulership'. Epi-Olmec **ki** was evidently derived from RULE acrophonically, on the basis of the root **ki?m*, by the alteration of its mouth (much as the Roman letter G was derived from the letter C by the addition of a horizontal stroke).

Nonetheless, several epi-Olmec phonetic signs may share values as well as forms with Mayan signs, which would indicate a more intimate historical connection. Epi-Olmec **mʉ** is probably the source of the Mayan sign **ma**/GREAT and its values; **ʉ** approximated the pronunciation of Mayan short **a**, and proto–Mixe-Zoquean/proto-Zoquean **mʉj* meant 'great', in the sense of 'big, important' (28). Another possible instance is epi-Olmec **wʉ** corresponding to Mayan **?u**. If valid, this too points to Mayan adoption of a Mixe-Zoquean sign; Mixe-Zoquean **wʉ** might be heard as **?u** by Mayan speakers, but Mayan **?u** would surely have been heard as **?u** by Zoqueans.

These results confirm earlier suggestions that epi-Olmec writing was closely related to Mayan and in fact that Mayan writing was very likely borrowed from a related Mixe-Zoquean script attested at the highland Guatemalan site of Kaminaljuyú (16). Epi-Olmec spelling conventions also closely resemble those of Mayan writing, and we speculate that the structure of Mixe-Zoquean languages provides a better internal basis for the development of these conven-

tions. The epi-Olmec texts seem to be more prosaic, discursive, and explicit than Mayan texts. This is a reflection of the fact that the textual genres used by epi-Olmecs and Mayans may have been different and, even when the same, probably had different standard traits.

Retrospect and Prospect

Our decipherment has laid both the substantive and methodological foundations for the study of the epi-Olmec written language. Methodologically, the epi-Olmec situation provided almost ideal circumstances for decipherment. The approximate identity of the language being written could be inferred, and the grammatical structure and vocabulary of this family at the relevant era were already long known through a well-developed historical reconstruction. The approximate nature of the writing system could be inferred in several ways: (i) partly from general practices in early scripts with as many signs as the epi-Olmec; (ii) more specifically by comparison to practices in Mayan writing, to which epi-Olmec was believed related; and (iii) partly on features of Mixe-Zoquean languages that constrain the structural possibilities of a logosyllabic script. The texts we studied also provided substantial evidence concerning the readings of many signs. Because Mixe-Zoquean languages had only 66 different CV sequences and the script made substantial use of phonetic spelling with signs with CV values, a given CV sign is reused in different contexts often enough to provide a variety of lines of evidence concerning its value. These distinct clues could be exploited if one could link them with specific grammatical or lexical knowledge about the languages. This grammatical linkage was straightforward because the obligatory inflectional morphology in Mixe-Zoquean languages is relatively limited. A lexical linkage depends on clues to the meanings of words; a systematic semantic framework could be derived from the structure of numeral and calendrical constructs that were widely used in southeastern Mesoamerica at that time and whose epi-Olmec forms could be readily recovered on the basis of similar Mayan practices.

The resources for decipherment have yet to be exhausted, and the process of decipherment is continuing (29). But substantial progress will require more archeological and linguistic fieldwork. The recovery of additional inscribed monuments is most critical to provide new contexts for the syllabograms, and renewed examination of the monuments of Tres Zapotes and Cerro de las Mesas may also help. Almost as important is a renewed and systematic study of the Mixe-Zoquean languages. None has

been described adequately for either historical reconstruction or decipherment. Work on Veracruz Zoquean is for now the most crucial, but work on Mixean is needed to recover pre–proto-Zoquean forms inherited from proto–Mixe-Zoquean but lost by proto-Zoquean times and will also prove useful if still-undiscovered texts were written within Mixean-speaking segments of Olmec or epi-Olmec civilization.

REFERENCES AND NOTES

1. The Mixe-Zoquean reconstruction is given primarily in T. Kaufman, *Mixe-Zoque Diachronic Studies* (unpublished monograph, 1963), which Kaufman has modified in some particulars, stimulated by new data.
2. We write Mixe-Zoquean forms in a practical, Spanish-based orthography. Most letters have their usual Spanish pronunciations, but *j* represents [h]. We use *ʉ* to represent a high, central-back unrounded vowel, like the *u* of *put* and *bush*, as pronounced by many Southerners and Westerners, and of *just* as in *just now*. The symbol *?* represents a glottal stop. Phonologically explicit representations of Mixe-Zoquean words are in italics; phonetic transliterations of epi-Olmec signs are in bold italics. In phonologically explicit representations of Mixe-Zoquean words, grammatical affixes are joined to roots or to one another by a hyphen; elements of compound words are joined by =. In transliterations of epi-Olmec sign sequences, signs spelling a single word are joined by hyphens. Reconstructed words, which are labeled by an asterisk, are pre–proto-Zoquean forms unless otherwise stated.
3. In the La Mojarra text, columns are read from the center leftward and then from the center rightward. This variation has parallels in many ancient scripts.
4. A morpheme is a linguistic unit that conveys a meaning and that is not decomposable into further meaningful forms. For example, the word *writing* in English contains the morphemes *write* and *-ing*.
5. The numbers associated with these signs are called Macri-Stark (MS) numbers after the sign transcription numbers given in M. J. Macri and L. M. Stark, *A Sign Catalog of the La Mojarra Script* (Monograph 5, Pre-Columbian Art Research Institute, in press); for instance, in Fig. 5A MS179 is *pi*.
6. For now, we transliterate MS30 as *tza?*, but it may in fact be used simply for *tza*.
7. An indirective is a suffix on transitive verbs that indicates that the object marked on the verb is indirect and not direct.
8. An inceptive is a suffix on verbs that means 'to begin to VERB'.
9. Syllabograms for *CVC* sequences, if they do exist, might also spell such affixes.
10. We indicate logograms in transliteration by rendering them in small capital letters. These transliterations are in bold italics when they specify a Zoquean word they are known to represent and are in roman type when they specify its (usually basic) meaning only. For example, the sign MS104, representing Zoquean *?ips* 'twenty', may be transliterated either *?ips* or TWENTY.
11. J. S. Justeson and T. Kaufman, "A chronological framework for the decipherment of epi-Olmec hieroglyphic writing" (unpublished manuscript, 1992).
12. F. Winfield Capitaine, *La Estela 1 de La Mojarra, Veracruz* (Publ. 16, Research Reports on Ancient Maya Writing, Washington, DC, 1988).
13. G. Lowe, *Estud. Cult. Maya* 2, 185 (1962); C. Navarrete, *An. Antropol.* 8, 69 (1971).
14. M. W. Stirling, *Stone Monuments of Southern Mexico* (Bulletin 138, Bureau of American Ethnology, Washington, DC, 1943), pp. 31–48. Some of these texts may be better preserved than the published drawings and photos suggest.

15. C. Navarrete, *The Olmec Rock Carvings at Pijijia-pan, Chiapas, Mexico and Other Olmec Pieces from Chiapas and Guatemala* (Paper 35, New World Archaeological Foundation, Provo, UT, 1974).
16. M. D. Coe, in *Origins of Religious Art and Iconography in Preclassic Mesoamerica*, H. B. Nicholson, Ed. (Univ. of California Press, Berkeley, CA, 1976), pp. 107–122; J. Marcus, *Annu. Rev. Anthropol.* 5, 35 (1976); M. Ayala, in *Antropología e Historia de los Mixe-Zoques y Mayas (Homenaje a Frans Blom)*, L. Ochoa and T. A. Lee, Jr., Eds. (Universidad Nacional Autónoma de México, Coyoacán, Mexico, 1983), pp. 175–221; J. S. Justeson, W. M. Norman, L. Campbell, T. Kaufman, *The Foreign Impact on Lowland Mayan Language and Script* (Publ. 53, Middle American Research Institute, New Orleans, LA, 1985); S. Meluzin, in *The Periphery of the Southeastern Maya Realm*, G. Pahl, Ed. (UCLA Latin American Center, Los Angeles, 1987), pp. 67–112.
17. M. Ayala, in (*16*), p. 197; S. Meluzin, in (*16*), p. 69; J. Justeson, report in a University of California at Berkeley seminar on non-Mayan scripts of Mesoamerica (1970), directed by J. A. Graham.
18. G. W. Lowe, in *The Origins of Maya Civilization*, R. E. W. Adams, Ed. (Univ. of New Mexico Press, Albuquerque, NM, 1977), pp. 197–248; J. S. Justeson, *World Archaeol.* 17, 437 (1986); ____ and P. Mathews, *Visbl. Lang.* 24, 88 (1990); B. Stross, ibid., p. 38. However, the text is instead addressed in terms of Mayan vocabulary and grammar by M. D. Coe [as cited in (*16*)] and by L. B. Anderson [*The Writing System of La Mojarra and Associated Monuments* (privately printed, Washington, DC, 1991)].
19. M. Macri, in *Literacies: Writing Systems and Literate Practices*, D. L. Schmidt and J. S. Smith, Eds. (Davis Working Papers in Linguistics, vol. 4, Department of Linguistics, University of California at Davis, CA, 1991), pp. 11–23.
20. L. Campbell and T. Kaufman, *Am. Antiq.* 41, 80 (1976).
21. Based on part of this evidence, the *wu* reading was proposed by T. Kaufman in a working group on La Mojarra Stela 1 at the meeting of the American Anthropological Association, Phoenix, AZ, 16 to 20 November 1988.
22. Ergative is the grammatical category of the agent of a transitive verb and, in Mixe-Zoquean languages, the possessor of a noun.
23. B. Stross, in (*18*), pp. 48–51.
24. The basis for our calendrical framework was presented by J. S. Justeson at the Workshop on La Mojarra Stela 1, University of California at Santa Barbara, CA, April 1989.
25. J. S. Justeson and P. Mathews, in (*18*), pp. 97 and 99; also presented by J. S. Justeson at the workshop in (*24*). The 'bloodletting' meaning of MS132 was identified by Mathews.
26. J. S. Justeson, W. M. Norman, L. Campbell, T. Kaufman, in (*16*), pp. 38–44; L. B. Anderson, in (*18*). For a contrary view, see J. E. S. Thompson, *Archaeology of Southern Mesoamerica: Part 2*, vol. 3 of *Handbook of Middle American Indians* (Univ. of Texas Press, Austin, 1965), p. 651.
27. J. S. Justeson and P. Mathews, in (*18*), p. 114.
28. A Mixe-Zoquean source for this pair of values in Mayan had been postulated by J. S. Justeson before the epi-Olmec sign was known [J. S. Justeson, W. M. Norman, L. Campbell, T. Kaufman, in (*16*), p. 44].
29. J. S. Justeson and T. Kaufman, *The Decipherment of Epi-Olmec Hieroglyphic Writing and Mixe-Zoquean Comparative Linguistics* (Univ. of Oklahoma Press, Norman, OK, in press).
30. F. Winfield Capitaine, in (*12*), p. 16.
31. All who have studied La Mojarra Stela 1 owe a tremendous debt to L. Wagner, G. Stuart, and F. Winfield Capitaine, who were instrumental in effecting its unrestricted dissemination. G. Stuart also produced the drawings of the text and monument that have been the indispensable basis for all subsequent work and provided us with access to his unpublished photographs. We began our joint work on epi-Olmec writing in

Ancient Conserved Regions in New Gene Sequences and the Protein Databases

Philip Green,* David Lipman, LaDeana Hillier, Robert Waterston, David States, Jean-Michel Claverie

Sets of new gene sequences from human, nematode, and yeast were compared with each other and with a set of *Escherichia coli* genes in order to detect ancient evolutionarily conserved regions (ACRs) in the encoded proteins. Nearly all of the ACRs so identified were found to be homologous to sequences in the protein databases. This suggests that currently known proteins may already include representatives of most ACRs and that new sequences not similar to any database sequence are unlikely to contain ACRs. Preliminary analyses indicate that moderately expressed genes may be more likely to contain ACRs than rarely expressed genes. It is estimated that there are fewer than 900 ACRs in all.

Understanding the functions and structures of the array of proteins expressed in living organisms is a fundamental goal of molecular biology. Our hope of attaining this goal stems largely from the unifying theme of shared evolutionary ancestry: related organisms have similar proteins and, within an organism, different proteins of related function are often wholly or partly similar in sequence, reflecting gene duplication and exon shuffling (*1*) during evolution. Such similarities can provide important functional insights, and consequently an important step in characterizing any newly sequenced gene is to compare its encoded protein sequence with the protein sequence databases in order to look for conserved regions shared with known proteins.

The present study uses extensive new sets of gene sequences to address several general questions about conserved regions: how many of these regions exist, what fraction has been discovered, and what proportion and types of proteins contain them. We focus on ancient conserved regions, or ACRs, detected through similarities between proteins from distantly related organisms. Over long evolutionary periods the less constrained portions of the sequences will have significantly diverged; consequently, the regions of

similarity are usually those of greatest structural or functional significance. ACRs often correspond to specific domains (or motifs) present in a variety of proteins, such as zinc finger DNA binding domains (*2*), or to enzyme active sites, but they can also comprise most or all of the sequence of a single highly conserved protein or protein family, such as actins and histones. Conserved regions of all of these types have been extensively cataloged (*3*, *4*). Because the degree of similarity between two related proteins reflects not only the amount of time since their last common ancestor but also their rates of sequence evolution, which can vary greatly for different proteins (*5*), not all proteins need contain ACRs.

The precise definition of an ACR depends on its required age and distribution among organisms and on the method used to detect sequence similarities. The present study involves ACRs that antedate the radiation of the major animal phyla [some 580 to 540 million years ago (*6*)] and that are present in diverse eukaryotes. We detected similarities by using the sequence alignment program BLAST (*7*) with a score cutoff sufficiently high to distinguish confidently true homologies from background in database searches (*8*). Figure 1 shows a representative BLAST alignment at this score level. Typically, a BLAST comparison of two related proteins reveals several (gap-free) aligned segments, separated by unaligned regions; in such cases we considered the entire collection of aligned segments to constitute a single conserved region, provided the segments always tended

P. Green, L. Hillier, and R. Waterston are in the Genetics Department, Washington University Medical School, St. Louis, MO 63110. D. Lipman, D. States, and J.-M. Claverie are at the National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20894.

*To whom correspondence should be addressed.