Phenetics, cladistics, and the search for the Alaskan ancestors of the Paleoindians: a reassessment of relationships among the Clovis, Nenana, and Denali archaeological complexes

Briggs Buchanan*, Mark Collard

Laboratory of Human Evolutionary Studies, Department of Archaeology, Simon Fraser University, 9635-8888 University Drive, Burnaby, BC V5A 1S6, Canada

Received 10 August 2007; received in revised form 13 November 2007; accepted 15 November 2007

Abstract

Clovis, Nenana, and Denali are the earliest well-documented archaeological complexes in the New World. Clovis has been found at sites throughout the lower 48 states of the USA and in parts of Canada but not in Alaska, whereas Nenana and Denali have been found in Alaska but not elsewhere in North America. In 1991, Goebel and colleagues reported phenetic analyses of tool types from Clovis, Nenana, and Denali. They found that Nenana clustered more closely with Clovis than with Denali, and concluded from this that Nenana was likely ancestral to Clovis. Goebel et al.’s study has been cited frequently since it was published, and their conclusions have been widely accepted. Here we explain why analyses of the type performed by Goebel et al. are problematic and also demonstrate empirically that their results are dependent on the algorithm and distance measure used. We then reanalyze their dataset with a technique from biology called cladistics. Cladistics has replaced phenetics as the method of choice for reconstructing evolutionary relationships in biology, because it is more objective. The results of this analysis differ from those obtained by Goebel and colleagues. The results strongly suggest that Denali and Clovis are in fact more closely related to each other than either is to Nenana.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: Cladistics; Phenetics; Tool assemblages; Clovis; Denali; Nenana

1. Introduction

The Clovis archaeological complex dates to the Late Pleistocene, and represents the first well-documented evidence of human occupation in the contiguous United States and Canada. It is widely accepted that humans initially colonized North America via Beringia, the landmass exposed between northeast Asia and Alaska during glacial periods. This places Alaska as the entry point into the New World for the First Americans. A few researchers favor an entry point along the eastern coast of North America (Bradley and Stanford, 2004; Stanford and Bradley, 2002), but as a number of researchers have pointed out there are significant problems with this hypothesis (Buchanan and Collard, 2007; Straus, 2000; Straus et al., 2005). With Alaska being the most likely entry point into the New World it is reasonable to assume that sites left by the ancestors of Clovis should be found there. However, the quest to locate the ancestors of Clovis in Alaska has proved challenging (Bever, 2006; Dixon, 1993; Goebel et al., 2003; Hamilton and Goebel, 1999; Meltzer, 2001; Yesner and Pearson, 2002). There are a number of sites in Alaska that date to the Late Pleistocene, but few of these sites have associated dates that are significantly older than the dates for Clovis in the contiguous United States or Canada (Bever, 2006).

The search for a Clovis progenitor in Alaska has focused on two archaeological complexes, Nenana and Denali. The latter is sometimes also referred to as the Paleoarctic archaeological complex. A number of reviews of the Nenana and Denali complexes are available (e.g., Bever, 2001a,b; Dumond, 2001;...
Hamilton and Goebel, 1999; Powers and Hoffecker, 1989; Yesner, 1996, 2001; West, 1996). Therefore, we will not describe them in detail here. Briefly, Nenana assemblages are characterized by a flake and macroblade technology. In contrast to Nenana, the Denali complex has evidence of both microblade technology and use of the burin technique (West, 1967). Neither complex includes fluted points, which are the most distinctive feature of Clovis.

Both Nenana and Denali include assemblages that have been dated to the Late Pleistocene, and therefore both are potential ancestors of Clovis. Nenana has until recently been considered the oldest cultural horizon in Alaska with assemblages recovered stratigraphically below Denali assemblages in cases where the two complexes have been recovered from the same site (e.g., at the Dry Creek and Moose Creek sites in the Nenana Valley and the Owl Ridge site in the Teklanika Valley) (Bever, 2006; Hamilton and Goebel, 1999). However, an assemblage containing microblades recovered from the lowest level of the Swan Point site (Holmes, 1998; Holmes et al., 1996) in the Tanana Valley is now considered the oldest dated assemblage in Alaska (Bever, 2006). Thus, the temporal relationship between Nenana and Denali is uncertain. Accordingly, using age to indicate which of them is most likely ancestral to Clovis is problematic.

Currently, it appears that the majority of researchers consider Nenana to be more closely affiliated with Clovis than Denali. This consensus seems to be based primarily on the results of a study reported by Goebel et al. (1991). Goebel et al.’s (1991) study has been cited on many occasions. Some of the researchers who have cited the study disagree with its conclusions (e.g., Bradley and Stanford, 2004; Dumond, 2001; Meltzer, 2004; Stanford and Bradley, 2002), but the majority appear to accept them (e.g., Bever, 2001a,b, 2006; Boldurian and Cotter, 1999; Dixon, 2001; Fiedel, 2000, 2006; Goebel, 1999, 2004; Haynes, 2002, 2005; Hoffecker, 2001; Holmes, 2001; Pearson, 1999, 2001; Roosevelt et al., 2002; Straus et al., 2005; Yesner et al., 2004). In addition, Goebel et al.’s (1991) conclusions have been incorporated into a number of subsequent influential studies (e.g., Hoffecker et al., 1993).

Goebel et al.’s (1991) study was the first substantive effort to clarify the relationships among the Denali, Nenana, and Clovis complexes. Goebel et al.’s (1991) dataset comprised frequencies of 77 tool types among five assemblages. Two of the assemblages are assigned to Nenana (Dry Creek Component I, Walker Road), two to Clovis (Blackwater Draw, Murray Springs) and one to Denali (Dry Creek Component II). Goebel et al. (1991) investigated the relationships among the assemblages using cumulative percentage curves and hierarchical cluster analysis. Both analyses suggested that Clovis is more similar to Nenana than to Denali, and Goebel et al. (1991) inferred from this that Clovis is descended from Nenana. In discussing this conclusion, Goebel et al. (1991) placed particular emphasis on the lack of microblade technology in both Nenana and Clovis and the presence of such technology in Denali.

Because of the uncertainties regarding the timing of the migration event represented by Nenana and Clovis, Goebel et al. (1991) suggested two possible scenarios. The first was that Nenana and Clovis represent a late entry (12,000–13,000 BP) that occurred when the ice sheets receded and permitted populations to move south. The second was that Nenana and Clovis represent the remnants of a much earlier migration event. Goebel et al. (1991) proposed that this latter event would have occurred after the introduction of blade technology ca. 25,000 BP in north Asia but prior to the widespread use of microblade technology ca. 18,000 BP in that region and most likely during the glacial interval that dates between 22,000 and 25,000 years ago.

Notwithstanding the widespread acceptance of Goebel et al.’s (1991) findings, there are reasons to be skeptical about their validity. Most significantly, both methods they employ are phenetic. Phenetic methods were widely used during the 1960s and 1970s to reconstruct phylogenetic relationships. However, in recent years researchers interested in phylogenetic reconstruction have largely abandoned phenetics (Schuh, 2000; O’Brien and Lyman, 2003). There are two reasons for this. One is that phenetic methods focus on overall similarity, and it has been recognized that the nature of evolution is such that overall similarity can be expected to be a poor guide to phylogenetic relationships. Specifically, phylogeneticists have come to recognize that, because evolution involves descent with modification, only character states that are the result of shared ancestry and derived relative to the ancestral state for the taxa under study have the potential to shed light on phylogenetic relationships. Such character states are referred to as “synapomorphies”, and are contrasted with “symplesiomorphies”, “autapomorphies”, and “homoplasies”. Symplesiomorphies are character states that are inherited from a common ancestor but do not vary among the taxa under study. Autapomorphies are character states that are derived relative to the ancestral state for the study group but only occur in a single taxon. Symplemiosmorphies and autapomorphies are not useful for phylogenetic reconstruction because they do not allow subgroups of taxa to be delineated. Homoplasies are derived character states that are shared by more than one taxon in a study group as a result of processes other than descent from a common ancestor, such as convergence and parallelism. The other reason phylogeneticists no longer rely on phenetics when reconstructing phylogenetic relationships is that the phylogenies it yields are difficult to defend. It has been found that genetics has the potential to produce different results depending on which combination of distance measure and clustering technique is used, but there is no objective means of choosing one such combination over any other (Ridley, 1983; Manly, 2005). Thus, in cases where multiple dendrograms exhibit different relationships, there is no reason to prefer one dendrogram rather than another.

Recently, Meltzer (2004) has highlighted the first of these problems—the failure to identify derived similarities—in connection with Goebel et al.’s (1991) use of phenetics. He criticized Goebel et al.’s (1991) study on the grounds that “the artifact classes that are common to both [Nenana and Clovis] are tool types in use for long periods of prehistory in many different settings, and thus are an unreliable basis for establishing
historical affinities or movements” (Meltzer, 2004: 547). Significantly, however, Meltzer did not provide evidence to support this claim. To date, the implications of the second problem with phenetics for Goebel et al.’s (1991) study has not been discussed or empirically evaluated. Here, we describe a study that was designed to assess the reliability of Goebel et al.’s (1991) findings in light of both problems. We focused on Goebel et al.’s (1991) cluster analysis because it is more rigorous than their comparison of cumulative percentage curves, which does not go beyond simple visual comparison.

2. Materials and methods

Two sets of analyses were carried out. In the first we evaluated the sensitivity of Goebel et al.’s (1991) results to their choice of hierarchical clustering algorithm and distance measure. For these analyses we used Goebel et al.’s (1991) dataset as presented. Goebel et al.’s (1991) dataset comprises tool type frequency data for five assemblages: Dry Creek Component I, Dry Creek Component II, Walker Road, Blackwater Draw and Murray Springs. The Dry Creek I and Walker Road assemblages belong to the Nenana complex, while the Dry Creek II assemblage is part of the Denali complex. The Blackwater Draw and Murray Springs assemblages belong to the Clovis complex. Goebel et al.’s (1991) classification scheme defines 14 general tool categories (retouched flakes, retouched blades, end scrapers, side scrapers, cobble tools, perforators, wedges, notches, bifaces, projectile points, denticulates, pointed tools, burins, and retouched microblades). The general tool categories are subdivided into between two and 11 tool types. A total of 77 different tool types are recognized in the classification scheme.

We used three different hierarchical clustering algorithms and 15 distance measures to derive dendrograms from the tool frequency data. The clustering algorithms were: (1) the unweighted pair-group average (UPGMA) procedure, which joins clusters based on the average distance between an observation in one cluster and an observation in another cluster; (2) the nearest neighbor or single linkage method, which uses the minimum distance between an observation in one cluster and an observation in another cluster; and (3) Ward’s method, which defines the distance between two clusters as the sum of squared deviations from observations to centroids, and seeks to minimize within-cluster variance. The distance measures were: (1) Bray-Curtis; (2) Chord; (3) Cosine; (4) Dice/Sorensen; (5) Euclidean; (6) Horn’s; (7) Jaccard; (8) Kulczynski; (9) Manhattan; (10) Morisita; (11) Ochiai; (12) Pearson’s correlation; (13) Raup-Crick; (14) Simpson; and (15) Spearman’s rho. Some of these distance measures utilize continuous data directly; others convert continuous data into presence-absence data. The former include Euclidean, Pearson’s correlation, Spearman’s rho, Manhattan, Bray-Curtis, Cosine, Chord, Morisita, and Horn’s. The latter include Dice or Sorensen, Jaccard, Simpson, Kulczynski, Ochiai, and Raup-Crick. Further details of the distance measures can be obtained from Hammer et al. (2001).

A total of 31 analyses were carried out. In each analysis a dendrogram was derived from a different combination of clustering algorithm and distance measure, and the cophenetic correlation coefficient recorded. The cophenetic correlation coefficient is a measure of how faithfully a dendrogram represents the dissimilarities among the taxa (Hammer et al., 2001). UPGMA was employed in 15 analyses, as was the nearest neighbor method. Ward’s method was only used once because it is limited to generating dendrograms from Euclidean distances. All these analyses were carried out in PAST version 1.68 (Hammer et al., 2001).

In the second set of analyses we subjected the tool assemblage data to cladistic analysis. Cladistics is the method of phylogenetic reconstruction that most biologists now rely on instead of phenetics (Schuh, 2000). Recently, cladistics has also started to be used in a number of other disciplines in which phylogenetic relationships are important, including anthropology (e.g., Buchanan and Collard, 2007; Collard et al., 2006; Coward et al., 2008; Foley and Lahr, 1997, 2003; Jordan and Shennan, 2003; Lyckett, 2007; Lyckett et al., 2007; O’Brien et al., 2001; O’Brien and Lyman, 2003; Robson-Brown, 1996; Shennan and Collard, 2005; Tehrani and Collard, 2002), historical linguistics (e.g., Gray and Jordan, 2000; Holden, 2002; Rexová et al., 2003), philology (e.g., Eagleton and Spencer, 2006; Robinson and O’Hara, 1996; Spencer et al., 2004), and business studies (e.g., McCarthy, 2005; Baldwin et al., 2005).

Cladistics is designed to distinguish between the different forms of similarity discussed earlier (i.e. synapomorphies, synapomorphies, autapomorphies, and homoplasies). In its simplest form, cladistic analysis proceeds via four steps. First, a character state data matrix is generated. This shows the states of the characters exhibited by each taxon. Next, the direction of evolutionary change among the states of each character is established. Several methods have been developed to facilitate this, including communality (Eldredge and Cracraft, 1980), ontogenetic analysis (Nelson, 1978), and stratigraphic sequence analysis (Nelson and Platnick, 1981). Currently the favored method is outgroup analysis (Arnold, 1981; Maddison et al., 1984). Outgroup analysis entails examining a close relative of the study group. When a character occurs in two states within the study group, but only one of the states is found in the outgroup, the principle of parsimony is invoked and the state found only in the study group is deemed to be evolutionarily novel with respect to the outgroup state. Having determined the probable direction of change for the character states, the next step in a cladistic analysis is to construct a branching diagram of relationships for each character. This is done by joining the two most derived taxa by two intersecting lines, and then successively connecting each of the other taxa according to how derived they are. Each group of taxa defined by a set of intersecting lines corresponds to a clade, and the diagram is referred to as a cladogram. The final step in a cladistic analysis is to compile an ensemble cladogram from the character cladograms. Ideally, the distribution of the character states among the taxa will be such that all the character cladograms imply relationships among the taxa.
that are congruent with one another. Normally, however, a number of the character cladograms will suggest relationships that are incompatible. This problem is overcome by generating an ensemble cladogram that is consistent with the largest number of characters and therefore requires the smallest number of homoplasies to account for the distribution of character states among the taxa. An example of an ensemble cladogram is shown in Fig. 1.

In order to conduct the cladistic analyses we converted Goebel et al.'s (1991) frequency data into presence-absence data. We used two Asian early Upper Paleolithic assemblages as outgroups. In order to reduce inter-observer bias and maintain consistency in how the tools were classified, the latter were selected from the assemblages analyzed by Goebel (1993). We employed assemblages from the Transbaikal, which is the easternmost of the regions represented in Goebel's (1993) sample. Goebel (1993) examined assemblages from four Transbaikal sites—Masterov Gora, Masterov Kliuch', Tolbaga, and Varvarina Gora. At the time of Goebel's (1993) study, the Masterov Gora and Masterov Kliuch' assemblages were very small. Masterov Gora had seven tools and Masterov Kliuch' 13. The other assemblages were much larger. Tolbaga had 681 tools and Varvarina Gora 156. To avoid introducing an excessive amount of missing data into the analysis, only the two large Transbaikal assemblages were employed as outgroups. Goebel's (1993) classification scheme includes the 77 tool types used by Goebel et al. (1991) plus 24 additional tool types. Three tool types employed by Goebel (1993) were not used in our study because it is clear Goebel et al. (1991) did not recognize them in their earlier analysis. The dataset used in the cladistic analyses is presented in the Appendix.

In the cladistic analyses we treated each of the 98 tool types as a character, the five North American assemblages as separate taxa in the ingroup, and the two Asian early Upper Paleolithic assemblages as outgroups. We coded the presence of a particular tool type as 1 and its absence as a 0. The character state data matrix was subjected to maximum parsimony analysis in PAUP* 4.0 (Swofford, 1998) using the branch-and-bound search routine, which is guaranteed to find the shortest cladogram. The characters were treated in such a way that a change from 0 to 1 cost the same in terms of number of steps as a change from 1 to 0.

In order to assess the fit between the most parsimonious cladograms and the tool assemblage dataset, three analyses were carried out. The first employed the Permutation Tail Probability (PTP) test. The PTP test was originally proposed as a method of determining whether or not a given dataset contains a statistically significant phylogenetic signal (Archie, 1989; Faith, 1990; Faith and Cranston, 1991). However, following criticism (e.g., Carpenter, 1992; Steel et al., 1993), it is now generally considered to be a heuristic device rather than a formal statistical test (Kitching et al., 1998). In the PTP test, a dataset is reshuffled multiple times and the length of the most parsimonious cladogram computed after each permutation. Thereafter, the length of the most parsimonious cladogram obtained from the unpermuted data is compared to the distribution of lengths of the most parsimonious cladograms and the tool assemblage dataset, three analyses were carried out. The first employed the Permutation Tail Probability (PTP) test. The PTP test was originally proposed as a method of determining whether or not a given dataset contains a statistically significant phylogenetic signal (Archie, 1989; Faith, 1990; Faith and Cranston, 1991). However, following criticism (e.g., Carpenter, 1992; Steel et al., 1993), it is now generally considered to be a heuristic device rather than a formal statistical test (Kitching et al., 1998). In the PTP test, a dataset is reshuffled multiple times and the length of the most parsimonious cladogram computed after each permutation. Thereafter, the length of the most parsimonious cladogram obtained from the unpermuted data is compared to the distribution of lengths of the most parsimonious cladograms and the tool assemblage dataset, three analyses were carried out. The first employed the Permutation Tail Probability (PTP) test. The PTP test was originally proposed as a method of determining whether or not a given dataset contains a statistically significant phylogenetic signal (Archie, 1989; Faith, 1990; Faith and Cranston, 1991). However, following criticism (e.g., Carpenter, 1992; Steel et al., 1993), it is now generally considered to be a heuristic device rather than a formal statistical test (Kitching et al., 1998). In the PTP test, a dataset is reshuffled multiple times and the length of the most parsimonious cladogram computed after each permutation. Thereafter, the length of the most parsimonious cladogram obtained from the unpermuted data is compared to the distribution of lengths of the most parsimonious cladograms and the tool assemblage dataset, three analyses were carried out. The first employed the Permutation Tail Probability (PTP) test. The PTP test was originally proposed as a method of determining whether or not a given dataset contains a statistically significant phylogenetic signal (Archie, 1989; Faith, 1990; Faith and Cranston, 1991). However, following criticism (e.g., Carpenter, 1992; Steel et al., 1993), it is now generally considered to be a heuristic device rather than a formal statistical test (Kitching et al., 1998). In the PTP test, a dataset is reshuffled multiple times and the length of the most parsimonious cladogram computed after each permutation. Thereafter, the length of the most parsimonious cladogram obtained from the unpermuted data is compared to the distribution of lengths of the most parsimonious cladograms and the tool assemblage dataset, three analyses were carried out. The first employed the Permutation Tail Probability (PTP) test. The PTP test was originally proposed as a method of determining whether or not a given dataset contains a statistically significant phylogenetic signal (Archie, 1989; Faith, 1990; Faith and Cranston, 1991). However, following criticism (e.g., Carpenter, 1992; Steel et al., 1993), it is now generally considered to be a heuristic device rather than a formal statistical test (Kitching et al., 1998). In the PTP test, a dataset is reshuffled multiple times and the length of the most parsimonious cladogram computed after each permutation. Thereafter, the length of the most parsimonious cladogram obtained from the unpermuted data is compared to the distribution of lengths of the most parsimonious cladograms and the tool assemblage dataset, three analyses were carried out. The first employed the Permutation Tail Probability (PTP) test. The PTP test was originally proposed as a method of determining whether or not a given dataset contains a statistically significant phylogenetic signal (Archie, 1989; Faith, 1990; Faith and Cranston, 1991). However, following criticism (e.g., Carpenter, 1992; Steel et al., 1993), it is now generally considered to be a heuristic device rather than a formal statistical test (Kitching et al., 1998). In the PTP test, a dataset is reshuffled multiple times and the length of the most parsimonious cladogram computed after each permutation. Thereafter, the length of the most parsimonious cladogram obtained from the unpermuted data is compared to the distribution of lengths of the most parsimonious

### Table 1

<table>
<thead>
<tr>
<th>Distance measure</th>
<th>Dendrogram clusters*</th>
<th>Cophenetic correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bray-Curtis</td>
<td>(DCII(DCI(WR(MS,BWD))))</td>
<td>0.892</td>
</tr>
<tr>
<td>Chord</td>
<td>(DCII(DCI(WR,M,C),(MS,BWD)))</td>
<td>0.726</td>
</tr>
<tr>
<td>cosine</td>
<td>(DCII(WR,DCI),(MS,BWD))</td>
<td>0.702</td>
</tr>
<tr>
<td>Dice</td>
<td>(DCII(DCI),(DCII(MS,BWD)))</td>
<td>0.797</td>
</tr>
<tr>
<td>Euclidean</td>
<td>(DCII(WR,BWD),(MS,DCI))</td>
<td>0.977</td>
</tr>
<tr>
<td>Horn's</td>
<td>(DCII(WR,DCI),(MS,BWD))</td>
<td>0.826</td>
</tr>
<tr>
<td>Jaccard</td>
<td>(DCII(DCI),(DCII(MS,BWD)))</td>
<td>0.826</td>
</tr>
<tr>
<td>Kulczynski</td>
<td>(DCII(DCI),(DCII(MS,BWD)))</td>
<td>0.718</td>
</tr>
<tr>
<td>Manhattan</td>
<td>(DCII(WR),(DCII(MS,BWD)))</td>
<td>0.986</td>
</tr>
<tr>
<td>Morisita</td>
<td>(DCII(WR),(DCII(MS,BWD)))</td>
<td>0.698</td>
</tr>
<tr>
<td>Ochiai</td>
<td>(DCII(WR),(DCII(MS,BWD)))</td>
<td>0.791</td>
</tr>
<tr>
<td>Pearson's correlation</td>
<td>(DCII(WR),(DCII(MS,BWD)))</td>
<td>0.715</td>
</tr>
<tr>
<td>Raup-Crick</td>
<td>(DCII(WR),(DCII(MS,BWD)))</td>
<td>0.679</td>
</tr>
<tr>
<td>Simpson</td>
<td>(DCII(WR),(DCII(MS,BWD)))</td>
<td>0.797</td>
</tr>
<tr>
<td>Spearman's Rho</td>
<td>(DCII(WR),(DCII(MS,BWD)))</td>
<td>0.835</td>
</tr>
</tbody>
</table>

* Assemblage abbreviations: BWD, Blackwater Draw; DCI, Dry Creek I; DCII, Dry Creek II; MS, Murray Springs; WR, Walker Road.
cladograms yielded by the permutations. If the original cladogram is shorter than 95% or more of the cladograms derived from the permutations, then the dataset is considered to contain a phylogenetic signal. The test was carried out in PAUP* 4.0, and the data were reshuffled 10,000 times.

The second analysis employed two measures of goodness of fit, the Consistency Index (CI) and the Retention Index (RI). The CI assesses homoplasy as a fraction of character change in relation to a given cladogram. It ranges between 1.0 and 0.0, with values close to 1 indicating a good fit between the cladogram and the dataset and values close to 0 indicating a poor fit. Sanderson and Donoghue (1989) and Hauser and Boyajian (1997) have shown that there is a significant inverse relationship between the CI and the number of

![Dendrograms](image)

**Fig. 2.** Dendrograms obtained with the UPGMA clustering algorithm and five different distance measures: a) Spearman’s Rho; b) Jaccard; c) Bray-Curtis; d) Manhattan; e) Euclidean.
taxa included in an analysis. This means that it is not possible to simply compare CIs among studies, and that the significance of the CI for a given cladogram has to be assessed relative to the CI that is expected for a cladogram with the same number of taxa (Sanderson and Donoghue, 1989). Accordingly, the CI for the most parsimonious tool assemblage cladograms was compared to the expected CI for a 7-taxa dataset. The expected CI was computed with the aid of a regression equation that Sanderson and Donoghue (1989) derived from a taxonomically diverse set of datasets. The RI measures the number of similarities in a dataset that are retained as homologies in relation to a given cladogram. The RI is insensitive to the presence of derived character states that are present in only a single taxon, or “autapomorphies”. The RI is also insensitive to the number of characters or taxa employed, and therefore can be compared among studies (Hauser and Boyajian, 1997; Sanderson and Donoghue, 1989). RIs for the most parsimonious cladograms were calculated and compared to RIs for 21 biological and 21 cultural datasets reported by Collard et al. (2006).

In the third analysis we employed a technique called bootstrapping in order to evaluate the robustness of support for the branches of the cladogram yielded by the parsimony analysis. In phylogenetics, bootstrapping was originally developed as a way of estimating the statistical likelihood of a given clad being real (Felsenstein, 1985). However, following several recent critiques (e.g., Carpenter, 1992; Kluge and Wolf, 1993), it is now considered by many researchers to be a heuristic tool rather than a statistical test (Kitching et al., 1998; but see Sanderson, 1995). In bootstrapping, a large number of subsets of data (normally between 1000 and 10,000) are randomly sampled with replacement from the character state dataset, with the character state assignments being retained in each sample. Minimum length cladograms are then computed from these subsets of the data, and a list of the clades that comprise the cladograms compiled. Thereafter the percentage of the resampled cladograms in which each clade was found is calculated. Datasets that fit the bifurcating model with little conflicting signal will return high bootstrap support percentages, and vice versa. The bootstrap analysis was carried out in PAUP* 4.0. Ten thousand iterations were conducted, and the consensus cladogram was computed using a confidence region of 50% (cf. Holden, 2002).

3. Results

3.1. Phenetic analyses

The analyses in which the clustering algorithms were combined with different distance measures indicate that Goebel et al.’s (1991) dataset yields different dendrograms depending on the clustering algorithm and distance measure employed. Using UPGMA, the 15 different distance measures produced five different dendrogram topologies (Table 1). The cophenetic correlation coefficients range from 0.679 to 0.986 (Table 1). Examples of the dendrograms are presented in Fig. 2. The dendrogram shown in Fig. 2a was obtained with the Spearman’s Rho distance measure. It shows the Clovis assemblages (Blackwater Draw and Murray Springs) forming one cluster and the Nenana assemblages (Dry Creek I and Walker Road) forming a second. The Denali assemblage (Dry Creek II) is positioned as the sister-group of both the Clovis cluster and the Nenana cluster. This dendrogram has the same topology as the dendrograms yielded by the chord, cosine, Horn’s, Morisita, and Pearson’s correlation distances measures. It also has the same topology as the dendrogram reported by Goebel et al. (1991). The dendrogram shown in Fig. 2b was obtained with the Jaccard distance measure. This dendrogram also suggests that the two Clovis assemblages form a cluster to the exclusion of the other assemblages. However, in contrast to the previous dendrogram, it places the Denali assemblage as the sister taxon of the Clovis assemblages. The Nenana assemblages form a cluster that is the sister taxa of the Clovis and Denali assemblages. This dendrogram has the same topology as the dendrograms yielded by the Dice, Kulczynski, Ochiai, Raup-Crick, and Simpson distances measures. The dendrogram shown in Fig. 2c was obtained with the Bray-Curtis distance measure. In this dendrogram the two Clovis assemblages cluster together, followed by Walker Road, Dry Creek I, and Dry Creek II. The dendrogram shown in Fig. 2d was obtained with the Manhattan distance measure. Like the three preceding dendrograms, it links together the two Clovis assemblages to the exclusion of the Nenana and Denali assemblages. Dry Creek I is linked to this cluster, followed by Walker Road and then Dry Creek II. The dendrogram shown in Fig. 2e was obtained with the Euclidean distance measure. It differs from all the other UPGMA dendrograms in that it does not group together the Clovis assemblages. In this dendrogram, one of the Clovis assemblages, Murray Springs, and one of the Nenana assemblages, Dry Creek I, are grouped together to the exclusion of the other assemblages. Blackwater Draw is then linked to this cluster, followed by Walker Road and Dry Creek II.

Using the nearest neighbor method, the 15 distance measures yielded six different dendrograms. The cophenetic correlation coefficients range from 0.598 to 0.985 (Table 2).

<table>
<thead>
<tr>
<th>Distance measure</th>
<th>Dendrogram clusters</th>
<th>Cophenetic correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bray-Curtis</td>
<td>(DCI(DCI(WR(MS,BWD))))</td>
<td>0.810</td>
</tr>
<tr>
<td>Chord</td>
<td>(DCI(DCI(WR(MS,BWD))))</td>
<td>0.642</td>
</tr>
<tr>
<td>Cosine</td>
<td>(DCI(DCI(WR(MS,BWD))))</td>
<td>0.622</td>
</tr>
<tr>
<td>Dice</td>
<td>(DCI(WR(DCI(MS,BWD))))</td>
<td>0.728</td>
</tr>
<tr>
<td>Euclidean</td>
<td>(DCI(WR(BWD(MS,DCI))))</td>
<td>0.975</td>
</tr>
<tr>
<td>Horn’s</td>
<td>(DCI(WR,DCI)(MS,BWD))</td>
<td>0.757</td>
</tr>
<tr>
<td>Jaccard</td>
<td>(DCI(WR(DCI(MS,BWD))))</td>
<td>0.768</td>
</tr>
<tr>
<td>Kulczynski</td>
<td>((WR,DCI)(DCI(MS,BWD)))</td>
<td>0.614</td>
</tr>
<tr>
<td>Manhattan</td>
<td>(DCI(WR(MS,BWD)))</td>
<td>0.985</td>
</tr>
<tr>
<td>Morisita</td>
<td>(DCI(DCI(WR(MS,BWD))))</td>
<td>0.598</td>
</tr>
<tr>
<td>Ochiai</td>
<td>(I(WR,DCI)(DCI(MS,BWD)))</td>
<td>0.718</td>
</tr>
<tr>
<td>Pearson’s correlation</td>
<td>(DCI(DCI(WR(MS,BWD))))</td>
<td>0.707</td>
</tr>
<tr>
<td>Raup-Crick</td>
<td>((WR,DCI)(DCI(MS,BWD)))</td>
<td>0.666</td>
</tr>
<tr>
<td>Simpson</td>
<td>(DCI(DCI(WR(MS,BWD))))</td>
<td>0.721</td>
</tr>
<tr>
<td>Spearman’s Rho</td>
<td>(DCI(DCI(WR(MS,BWD))))</td>
<td>0.824</td>
</tr>
</tbody>
</table>
Examples of the dendrograms are presented in Fig. 3. The dendrogram shown in Fig. 3a was obtained with the Spearman’s Rho distance measure. It shows the Clovis assemblages forming a cluster to the exclusion of the Nenana and Denali assemblages. Walker Road is linked to this cluster, followed by Dry Creek I, and then Dry Creek II. This dendrogram has the same topology as the dendrograms yielded by the chord, cosine, Morisita, and Pearson’s correlation distances measures. The dendrogram shown in Fig. 3b was obtained with the Simpson distance measure. This dendrogram also suggests that the two Clovis assemblages form a cluster to the exclusion of the other assemblages. However, in contrast to the previous dendrogram, it places the Denali assemblage as the sister taxon of the Clovis assemblages and suggests that the Nenana

Fig. 3. Dendrograms obtained with the nearest neighbor clustering algorithm and six different distance measures: a) Spearman’s Rho; b) Simpson; c) Jaccard; d) Manhattan; e) Bray-Curtis; f) Euclidean.
assemblages form a cluster that is the sister taxon of the Clovis and Denali assemblages. This dendrogram has the same topology as the dendrograms yielded by the Kulczynski, Ochiai, and Raup-Crick distances measures. The dendrogram shown in Fig. 3c was obtained with the Jaccard distance measure. In this dendrogram, the two Clovis assemblages form a cluster to the exclusion of the Nenana and Denali assemblages. Dry Creek II is linked to this cluster, followed by Walker Road, and then Dry Creek I. This dendrogram has the same topology as the dendrogram yielded by the Dice distance measure. The dendrogram shown in Fig. 3d was obtained with the Manhattan distance measure. Like the three preceding dendrograms, it links together the two Clovis assemblages to the exclusion of the Nenana and Denali assemblages. Dry Creek I is linked to this cluster, followed by Walker Road, and then Dry Creek II. This dendrogram has the same topology as the dendrogram yielded by Horn's distance measure. The dendrogram shown in Fig. 3e was obtained with the Bray-Curtis distance measure. In this dendrogram, the two Clovis assemblages are grouped together to the exclusion of the other assemblages. Walker Road is linked to this cluster, followed by Dry Creek II, and then Dry Creek I. The dendrogram shown in Fig. 3f was obtained with the Euclidean distance measure. This dendrogram differs from all the other nearest neighbor dendrograms in that it groups together one of the Clovis, Murray Springs, with one of the Nenana assemblages, Dry Creek I, to the exclusion of the other assemblages. Blackwater Draw is linked to this cluster, followed by Walker Road, and then Dry Creek II.

3.2. Cladistic analyses

Of the 98 tool types used in the cladistic analyses, 43 were parsimony uninformative and did not contribute to the construction of the cladogram. The maximum parsimony analysis returned two equally parsimonious cladograms, which are shown in Fig. 5. In one of the cladograms (Fig. 5a), the Denali assemblage, Dry Creek II, is the sister taxon of a clade comprising the two Clovis assemblages, Murray Springs and Blackwater Draw. According to this cladogram, the two Nenana assemblages, Dry Creek I and Walker Road, form a clade that is the sister taxon of the clade consisting of Dry Creek II, Murray Springs, and Blackwater Draw. In the other cladogram (Fig. 5b), the Denali assemblage is again positioned as the sister taxon of a clade comprising the two Clovis assemblages. This cladogram differs from the previously described one in that it suggests that the Nenana assemblages are paraphyletic. The Dry Creek I Nenana assemblage is positioned as the sister taxon of the clade comprising Dry Creek II, Murray Springs, and Blackwater Draw, while the Walker Road Nenana assemblage is positioned as the sister taxon of the clade comprising Dry Creek I, Dry Creek II, Murray Springs, and Blackwater Draw. In both cladograms, the (Dry Creek II, Murray Springs, Blackwater Draw) clade is defined by the gain of elliptical bifaces and the loss of end scraper fragments, steeply keeled end scrapers, double-end scrapers, and wedges. Again, in both

a

Varvarina Gora
Tolbaga
Walker Road
Dry Creek I

Dry Creek II
Murray Springs
Blackwater Draw

b

Varvarina Gora
Tolbaga
Walker Road
Dry Creek I

Dry Creek II
Murray Springs
Blackwater Draw

Fig. 4. Dendrogram obtained with Ward’s clustering algorithm and Euclidean distance.

Fig. 5. Maximum-parsimony cladograms A and B inferred using the branch and bound search algorithm in PAUP* 4.0 of the tool assemblage data (cladogram lengths = 115; RI = 0.52; CI = 0.75).
cladograms, the (Murray Springs and Blackwater Draw) clade is defined by the gain of double straight side scrapers and Clovis projectile points and the loss of transverse scrapers, single concave side scrapers, and retouched blade fragments.

The results of the three goodness-of-fit analyses were consistent. The PTP test indicated that the two most parsimonious cladograms were significantly shorter than any of the 10,000 permuted cladograms ($P = 0.0001$). The CI for the most parsimonious cladograms was 0.75. Using Sanderson and Donoghue’s (1989) regression formula based on biological datasets, the CI of 0.75 derived from the tool type data is not significantly different from the expected value of 0.74 for 7 taxa. The RI associated with the most parsimonious cladograms was 0.52. This is within the range of RIs returned by the datasets examined by Collard et al. (2006). The biological datasets analyzed by Collard et al. (2006) returned RIs between 0.35 and 0.94, whereas the cultural datasets returned RIs between 0.42 and 0.78. Thus, the CI and RI comparisons also suggest that fit between the most parsimonious cladograms and the tool type dataset is good.

The bootstrap analysis yielded the Murray Springs and Blackwater Draw clade 7333 times out of 10,000 replications (Fig. 6; Table 3). The three-assemblage clade comprised of Dry Creek II, Murray Springs, and Blackwater Draw assemblages was supported by 56% of the bootstrap replicates. Examination of the levels of bootstrap support given for clades representing alternative hypotheses suggests they are unlikely.

Most significantly, a clade consistent with Goebel et al.’s (1991) results (i.e. one comprising the Walker Road, Murray Springs, and Blackwater Draw assemblages) was supported by only 9.8% of the bootstrap replicates (Table 3).

4. Discussion and conclusions

The results of the various phenetic analyses demonstrate that the topology of the dendrogram yielded by Goebel et al.’s (1991) dataset is dependent on the combination of clustering algorithm and distance measure employed. Dendrograms showing Denali as most distant to Clovis were produced but so were dendrograms showing Nenana as most distant to Clovis. As noted earlier, in cases where a dataset yields multiple dendrograms when different combinations of clustering algorithm and distance measure are employed, there is no reason to prefer one of the dendrograms to the others. Thus, the results of the first set of analyses suggest that no confidence can be placed in the results of Goebel et al.’s (1991) study.

The results of the cladistic analysis are not consistent with those obtained by Goebel et al. (1991). Using two Asian early Upper Paleolithic assemblages as outgroups, we obtained two equally parsimonious cladograms. The cladograms suggest that, contrary to what Goebel et al. (1991) found, Denali and Clovis are in fact more closely related to each other than either is to Nenana. In both cladograms the Denali assemblage, Dry Creek II, is the sister taxon of the two Clovis assemblages, Murray Springs and Blackwater Draw, to the exclusion of the Nenana assemblages. The good fit between the most parsimonious cladograms and the character state data matrix revealed by the PTP test, the CI and RI comparisons, and the bootstrap analysis suggests that the interassemblage variation in tool assemblages is well explained by the most parsimonious cladograms.

The cladistic analyses suggest that Clovis is not the direct descendent of Nenana as Goebel et al. (1991) contend. Rather, the analyses suggest that Clovis is either descended from Denali, or Clovis and Denali are descended from an as yet unknown ancestor. If the new dates on the Denali assemblage from Swan Point are verified and Denali is established to be significantly older than Clovis (Bever, 2006), this would support the idea that the relationship between Denali and Clovis is
one of ancestor and descendent. However, if Denali is found to be contemporaneous with or younger than Clovis, as some researchers have suggested (e.g., Hamilton and Goebel, 1999), then the cladograms described here suggest that Denali and Clovis were derived from the as yet unknown ancestor. Refinement of the radiocarbon record for the Alaskan complexes or the discovery of a heretofore-unknown complex in Beringia will help to resolve this question. Furthermore, we can predict the relative chronological order of the appearance of the three archaeological complexes based on the cladograms. As researchers refine the radiocarbon record associated with the Nenana, Denali, and Clovis complexes, some Nenana assemblages should turn out to be the oldest, followed by Denali and then Clovis. It is possible that there will be a degree of temporal overlap among the complexes. However, as the chronological framework becomes more precise, the precedence of Nenana first, followed by Denali, and then Clovis should become clearer. It should be noted that the cladogram only makes predictions about the first appearance dates of the complexes; their last appearance dates need not be consistent with the cladogram.

As noted earlier, microblade technology has been treated as a special character in most previous attempts to link Clovis, which lacks microblades, to an ancestral complex and as a result Denali has not been considered by most researchers to be ancestral to Clovis. However, in the cladistic analysis the relationships among the ingroup assemblages are not determined by the presence or absence of microblades. This is because microblade technology in the form of retouched microblade tools occurs only in the Dry Creek II Denali-affiliated assemblage, and uniquely derived characters (autapomorphies) are incapable of shedding light on phylogenetic relationships in the cladistic framework. Instead, it is the gain of elliptical bifaces and the loss of end scraper fragments, steeply keeled end scrapers, double-end scrapers, and wedges that links the Clovis and Denali assemblages, and the gain of double straight side scrapers and Clovis projectile points and the loss of transverse scrapers, single concave side scrapers, and retouched blade fragments that links together the two Clovis assemblages. An important corollary of this concerns the timing of the peopling of the New World. To reiterate, Goebel et al. (1991) proposed late and early entry models to fit their finding that Nenana is ancestral to Clovis and hypothesized that its Asian ancestor should lack microblade technology. The autapomorphic status of the Denali microblade technology means that there is no need to restrict hypotheses based on the presence or absence of microblade technology. It is unnecessary to search exclusively for an ancestor of Clovis that is distinguished by a lack of microblade technology. Accordingly, hypotheses do not need to be limited to time periods when non-microblade assemblages were prevalent. The cladograms derived from our analyses suggest that ancestors without microblades (Nenana and the early Upper Paleoindian assemblages) gave rise to descendants with microblades (Denali) and descendants without microblades (Clovis).

Goebel et al.’s (1991) study did not include a number of lithic assemblages in central Alaska have been dated to the same time period as the Nenana and Denali assemblages analyzed by them (Bever, 2006; Hamilton and Goebel, 1999). Some of these assemblages have been assigned to Nenana (e.g., Chugwater, 1 Moose Creek 1) or Denali (e.g., Chugwater 2, Moose Creek 2 and 3, Panguingue Creek 1 and 2). However, the affiliation of others is still a matter of debate (e.g., Broken Mammoth 3 and 4, Swan Point 6 and 7, Owl Ridge 1 and 2). In addition, research in other regions of Alaska has identified lithic assemblages that may date to the same time period as Nenana and Denali, and therefore may be relevant to both the nature of the relationship between Nenana and Denali, and the origins of Clovis (e.g., Mesa in northern Alaska and Onion Portage in northwestern Alaska). Accordingly, the obvious next step is to apply Goebel et al.’s (1991) classification scheme to these other assemblages, combine the resulting data with Goebel et al.’s (1991) data, and then subject the expanded dataset to cladistic analysis.

Acknowledgements

We wish to thank Michael O’Brien for his comments on an earlier draft of this work. BB gratefully acknowledges financial support from National Science Foundation grant 0502293. MC’s work on this project was supported by the Social Sciences and Humanities Research Council and the Canada Research Chairs Program. Any opinions, findings, and conclusions expressed in this research are those of the authors and do not necessarily reflect the views of the National Science Foundation or any of the other funding bodies.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version, at doi:10.1016/j.jas.2007.11.009.

References


Holmes, C.E., 2001. Tanana valley archaeology circa 14,000 to 9000 B.P. Arctic Anthropology 38, 54–120.


