These files are the ones that include judgments of cognacy for individual items. They pertain to the 11 language groups identified by checkmarks in Table 2 and presented in Table 4. For the other groups, the sources provide cognate percentages averaged over items.

The first line of each file is specific to ASJP software. Col. 6 gives the maximum number of synonyms read for each item, cols. 11-12 give the minimum number of items that must be attested in a wordlist for that wordlist to be analyzed by the software, cols. 15-18 give a date that can specify which lists are read, col. 24 is always 3 for the present analyses, and the 1 in col. 30 excludes identified loanwords from the analyses.

The next segment of the file lists the items for which cognates are identified. The first line of this segment gives the format for reading the item numbers in the immediately following list. Depending on the file, the list itself is either the Swadesh 100-item list, or the 40 most stable of the Swadesh items (which are used in most ASJP wordlists), or a subset of one of these. The English names of the items are less important than the preceding numbers, which are used to identify the items in the wordlists. If the number for an item is 0, the item is lacking in the ASJP wordlists even though cognates are identified for the item. The number -1 signals the end of the list.

The next segment lists the languages for which cognates are identified. The name of each language is exactly the same as the name of the corresponding ASJP wordlist. A blank line signals the end of this list.

The next segment contains the judgments of cognacy. The first line gives the name of the language group and the source of the data. The next line gives the format for reading the data. For all groups except Athapaskan, the data are presented in a matrix. Each row is an item, in exactly the same order as in the preceding list of items. Each column is a language, in exactly the same order as in the preceding list of languages (some columns are omitted, particularly the first, as specified in the format). Words judged to be cognate have the same positive number in a given row. Unattested items have 0. Loans from the same source have the same negative number in a given row. The letter X (upper case) in col. 1 signals the end of the matrix. For Athapaskan, each row refers to a pair of languages, which are identified by initials in the first two columns (if the first column is blank, the language is the same as in the previous row). The rest of the row gives item numbers for cognates between the languages in the row. A blank line signals the end of the segment for Athapaskan.

The next segment consists of the ASJPCode symbols that are used to transcribe the wordlists. The last segment consists of the ASJP wordlists themselves. These segments are described on the ASJP website (http://email.eva.mpg.de/~wichmann/ASJPDatabaseDescription.pdf).