related homologs, variable regions can be highlighted as such by replacing letters identical to the reference with periods.

Summary plots: conservation, consensus and quantitative annotation. Annotation is important for navigation, in both flat diagrams and interactive systems, because it guides the eye toward 'important' regions of an alignment. MSA workbenches and most of the editing and analysis tools reviewed here allow the user to interactively annotate alignments (see below). However, practically all MSA visualizations include some form of automatically generated annotation, such as consensus lines and alignment quality plots, displayed either above or below the alignment. Consensus annotation has its roots in the textual alignment files generated by MSA programs, but a variety of plots are now provided by modern tools (**Fig. 3**). Quality and consensus plots are calculated from each column's symbol frequency distribution using one of the many measures available[14,15]. Alternatively, sequence logos[16,17] provide a user-friendly indication of the dominant symbols at each position of the alignment. As in shading, described above, annotation can result from other kinds of calculations. For example, PFAAT[18], MacVector[19], VectorNTI[3] and Geneious are able to compute and plot averaged physicochemical quantities such as isoelectric point, and STRAP[20] supports extension of the program to allow complete customization.

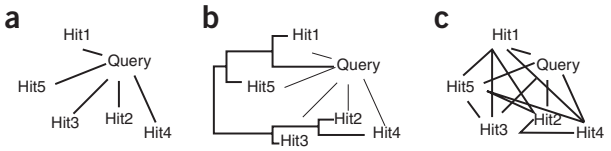

Figure 1 | Alignment topologies. Consistency graphs demonstrating complexity of different types of alignment algorithm. Nodes represent the query and each hit from the results of a sequence search, and edges indicate the mapping between a pair of sequences that the alignment algorithm optimizes. (**a**) MSA constructed directly from pairwise database search results. (**b**) MSA constructed using a guide tree, in which closely related sequences and then groups of sequences are optimally aligned. (**c**) MSA from consistency-based algorithms such as T-Coffee, in which all sequences are optimally aligned with one another.

Alignment editing, analysis and annotation

Integrated systems to support the editing and analysis of sequences have become possible with increased computing power and the ubiquity of internet connectivity. Most of the tools for MSA visualization mentioned here provide alignment coloring, shading and automated annotation facilities, as described above. However, 'editing and analysis' tools and most of the MSA workbench tools also allow alignments to be interactively edited,

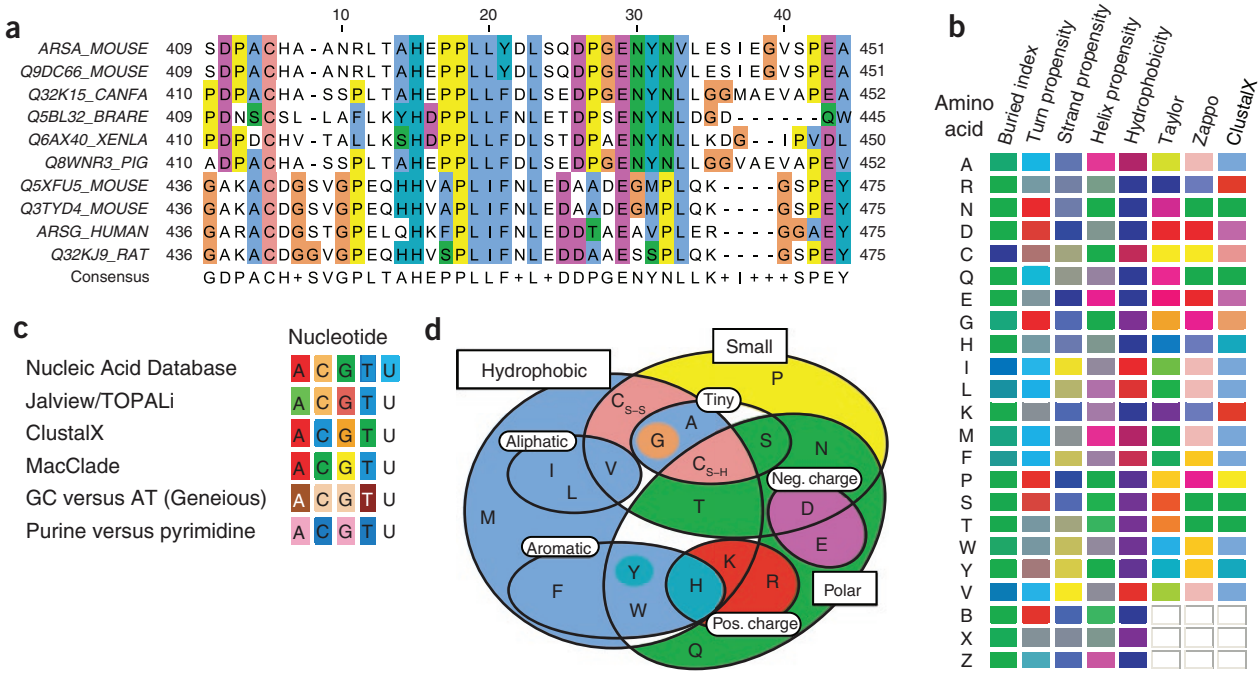

Figure 2 | Multiple alignment visualization. (**a**) A protein sequence alignment diagram rendered with Jalview[25]. Aligned sequences are arranged in rows and placed into a single reference frame, where each aligned position occupies a column in a table. Dashes indicate gaps. The label on the left-hand side of each sequence gives its Uniprot[53] entry name, start and end positions are shown at each end of its row and tick marks at the top allow a particular aligned column or sequence position to be read off. The consensus row at the bottom shows the most frequent residue at each column or a '+' if two or more residues are equally abundant. Residues in the alignment are colored according to the ClustalX[4] shading model: a color is only applied when that residue's abundance in the column is above a residue-specific threshold, highlighting potentially important residues (for example, proline and glycine) or patterns of conservation. (**b**) Examples of amino acid color schemes. Schemes are either quantitative, reflecting empirical or statistical properties of amino acids; or qualitative, reflecting an assignment according to physicochemical attributes. Zappo is a qualitative scheme developed by M. Clamp (personal communication); B, X and Z are amino acid ambiguity codes: B is aspartate or asparagine; Z is glutamate or glutamine; X is an unknown (or 'other'). (**c**) Examples of nucleotide color schemes used by the Nucleic Acid Database[54] and a selection of visualization tools. (**d**) Venn diagram after Taylor[55] showing the amino acids grouped according to their physicochemical properties. Coloring of each group (or amino acid label) is according to ClustalX, demonstrating the correspondence between color and physicochemical properties. Pos., positive; neg., negative.