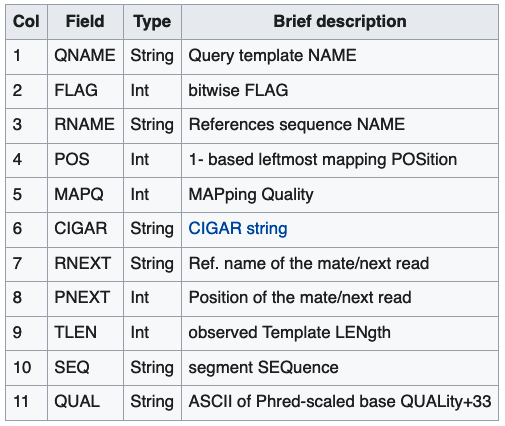
**MIP 280A4: Microbial Sequence Analysis**

**Mapping, part 2, In-class exercise questions**

1. Last week, you downloaded and cleaned-up some reads from an existing NGS dataset. What kind of animal did these reads come from? What was unusual about this particular animal? (2 pts)
2. After mapping reads to the boa constrictor mitochondrial genome, insect the resulting SAM file and answer the following questions about the first mapped read. (1 pt each):

[This Wikipedia page](https://en.wikipedia.org/wiki/SAM_(file_format)#Format) includes a table listing the content of each column in a SAM-formatted file:



* 1. What position in the mtDNA sequence did this first mapped read map to?
  2. What is the mapping quality for this read's mapping?
  3. What does the mapping quality score of this first mapped read indicate?

1. After importing the read mapping information into Geneious, answer the following questions about reads mapped to the boa constrictor mitochondrial genome. (1 pt each)
   1. What is the average coverage depth across the mitochondrial genome?
   2. Is coverage depth even across the mitochondrial genome?
   3. Would you expect coverage to be even across the genome? (Recall that this data is derived from total RNA from liver tissue).
   4. How does differential expression of mitochondrial genes relate to their coverage?
   5. Are there any variants between this snake's mitochondrial genome sequence and the boa constrictor reference sequence?
   6. Is it expected that there are variants between these reads and this reference sequence? Explain your answer.
   7. Can you distinguish true variants from sequencing errors?
   8. In general, how can you distinguish true variants from sequencing errors?
   9. Is it possible that reads that derive from the boa constrictor *nuclear* genome are mapping to this sequence?
   10. How would you prevent nuclear reads from mapping to the mitochondrial genome?
   11. Can you identify mapped read pairs?