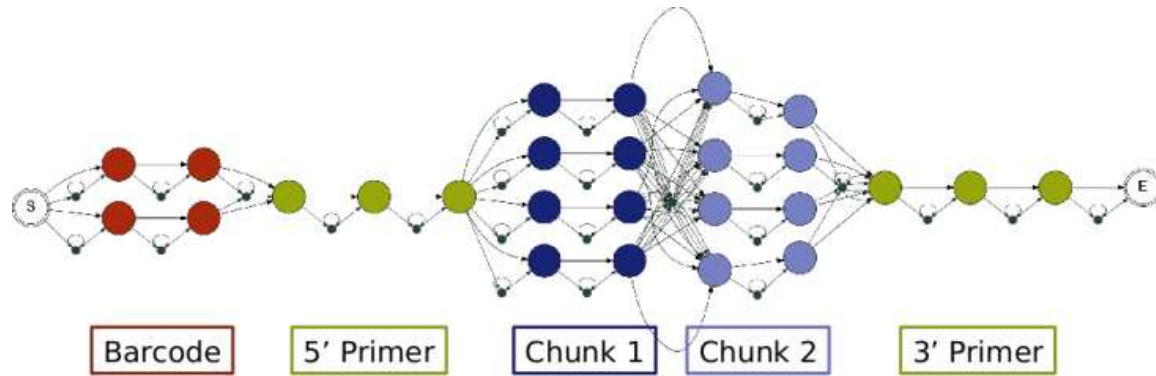
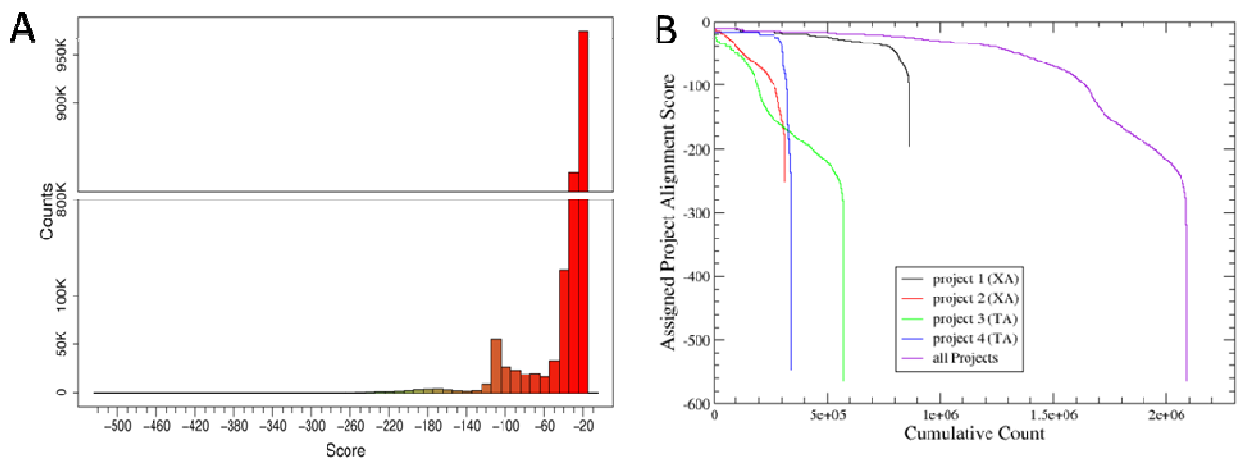




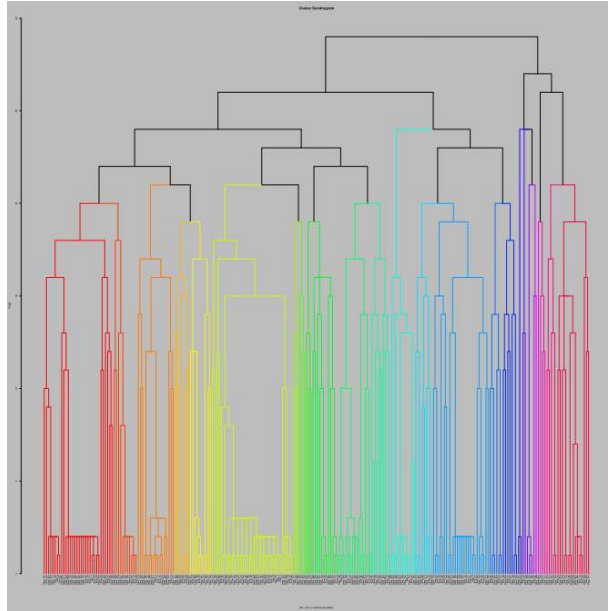
**Figure S1.** Schematic representation of the Markov model topology for a library with only two bar codes, a 3-base 5'-primer, two random fragments with four choices each, and a 3-base 3'-primer. The bar code may be on either end of the 5'- or 3'-primers.



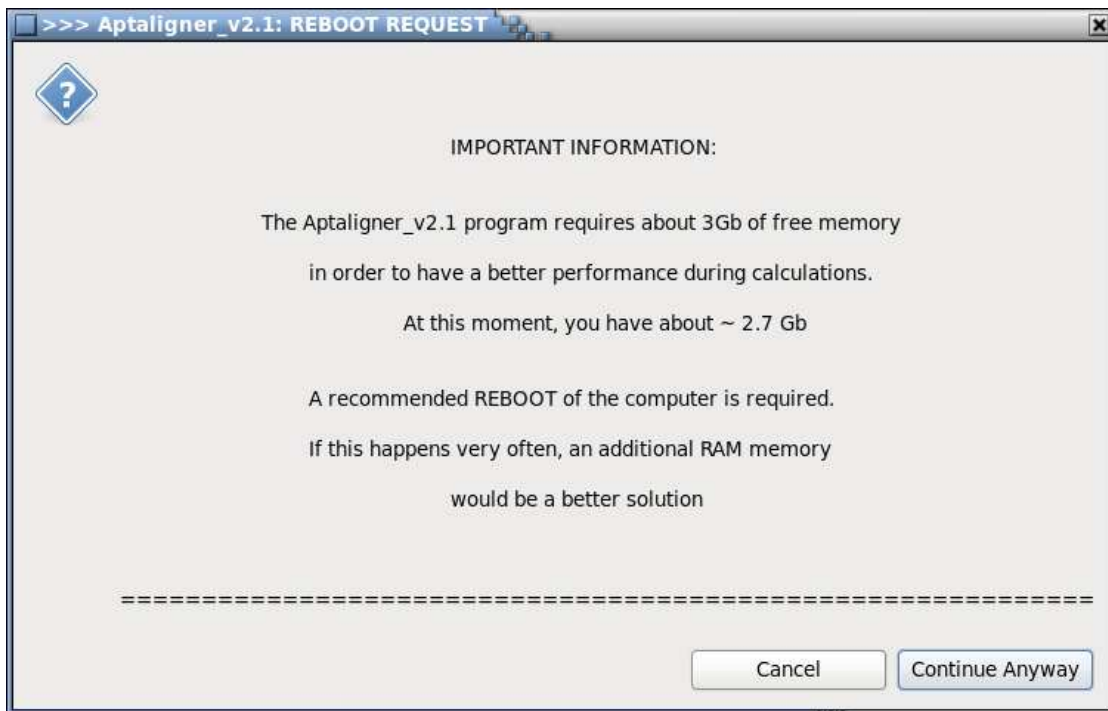
**Figure S2.** Sequence counts by score. A) Histogram showing scores for a single project when noise and cut-off filters were not used. Even when not using length and noise cut-off filters, 75% or more of all sequences in a project have good scores. B) In total, about 1.5 million good sequences (scores of about -60 or higher) were obtained from a data chip containing NGS sequences from four different projects. Projects 1 through 4, respectively, had good sequence scores for approximately 800k, 200k, 200k, and 300k sequences. The sharp drops at the end of each curve represent a very small number of really poor sequences (noise), which are forced into a project when noise and length cut-off filters are not used (older Aptaligner v1.0). XA and TA refer to X-Aptamers and normal aptamers, respectively.



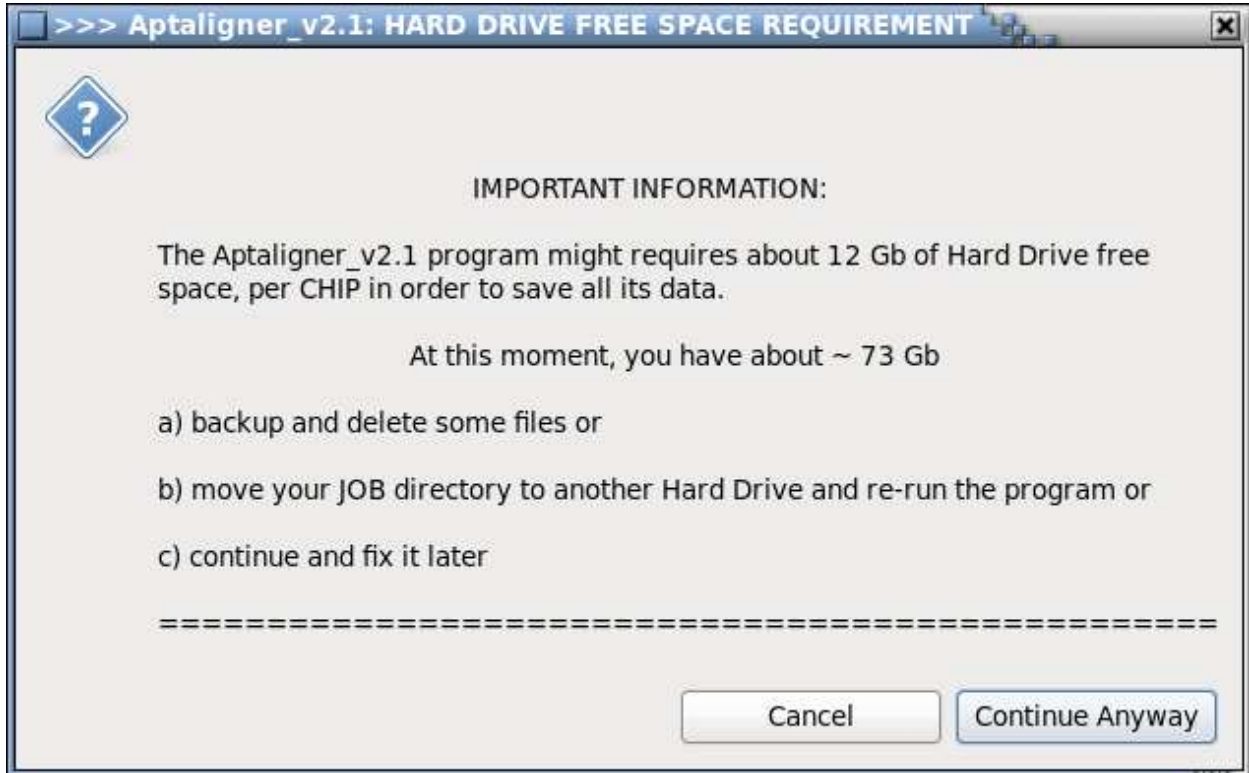
**Figure S3.** Example clustering diagram produced by Aptaligner for the top 250 sequences. Clustering information is also reported in text files for each project.



**Figure S4.** Example warning when RAM *may* be an issue with running Aptaligner successfully. The actual amount of RAM needed will depend on the complexity of the library being analyzed. This warning can sometimes be ignored, but will often lead to failure. When analyzing 5 million sequences, 3+ GB of RAM is recommended.



**Figure S5.** An example GUI comparing needed disk space (up to 12 GB) versus current free disk space. Disk space requirements depend on the number of sequences analyzed and the number of library design files against which they are compared. If the number of free GB are not shown in this box (“you have about ~ GB), and the user continues anyway, the program is highly likely to crash in within a few minutes with a resulting error declaring that the project 1 sequence file “HAS NOT THE EXPECTED FORMAT” – because it is empty.



**Figure S6.** Error resulting from an incorrect Python build. Re-installing python has been shown to fix this error, which appears to result from the C code not compiling correctly during Aptaligner set up. Reducing the RAM to 0.25 GB and CPU to 1 could not reproduce this error when analyzing the “medium test file provided at .

```

#####
"ATTENTION: THE ALIGNMENT INPUT FILE ON: Job_Dir/ p1.Dir/p1_alignments.txt
HAS NOT THE EXPECTED FORMAT
The number of fields expected are 13, but this file has:
THE SCRIPT HAS BEEN STOPPED ...
CHECK THE FILE AND TRY AGAIN
#####

```