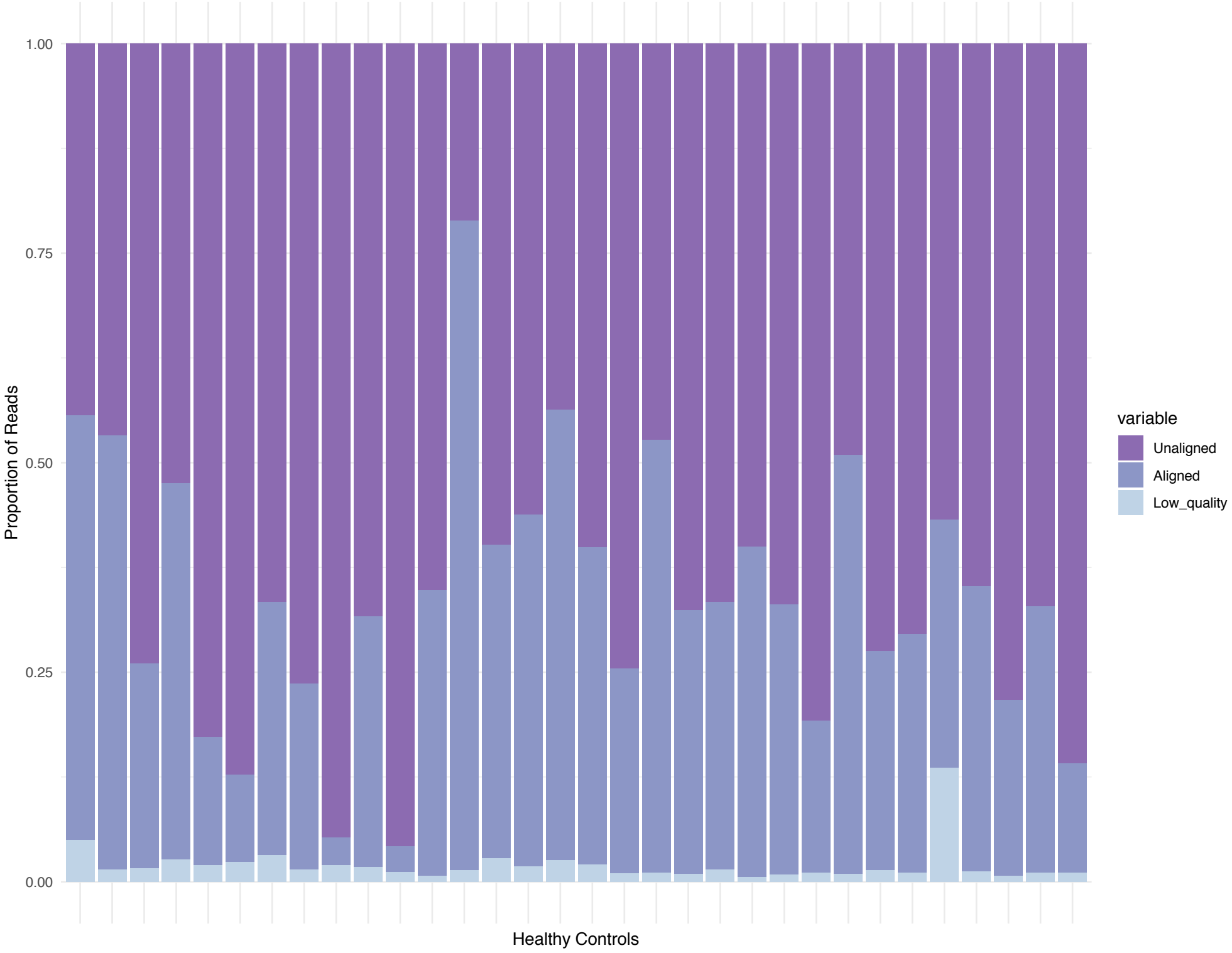
Identifying missing microbial taxa

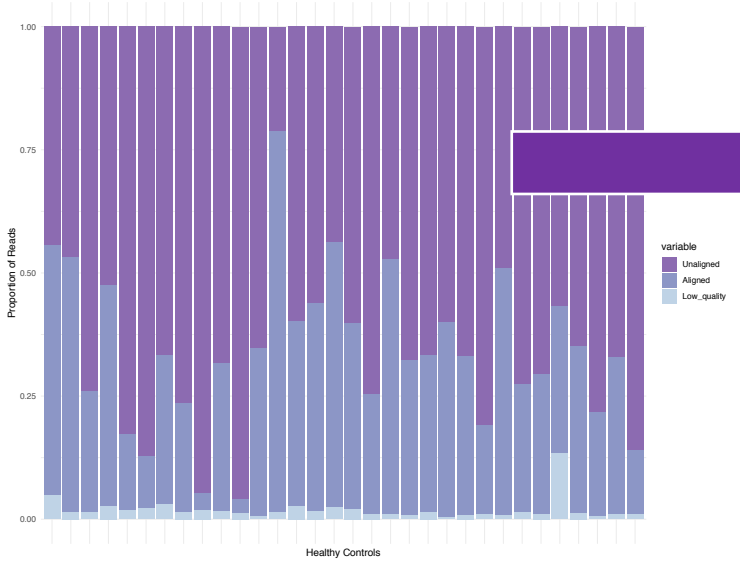

1.00
0.75
0.50
0.25
0.00
Proportion of Reads
Healthy Controls
variable
Unaligned
Aligned
Low_quality

363,525,943 reads BLASTn to HTGs, viral, EST_others, and nt databases

Retain top hit from each sequence >50 coverage and >90% identity

18% of unaligned reads aligned to at least one genome in a public database

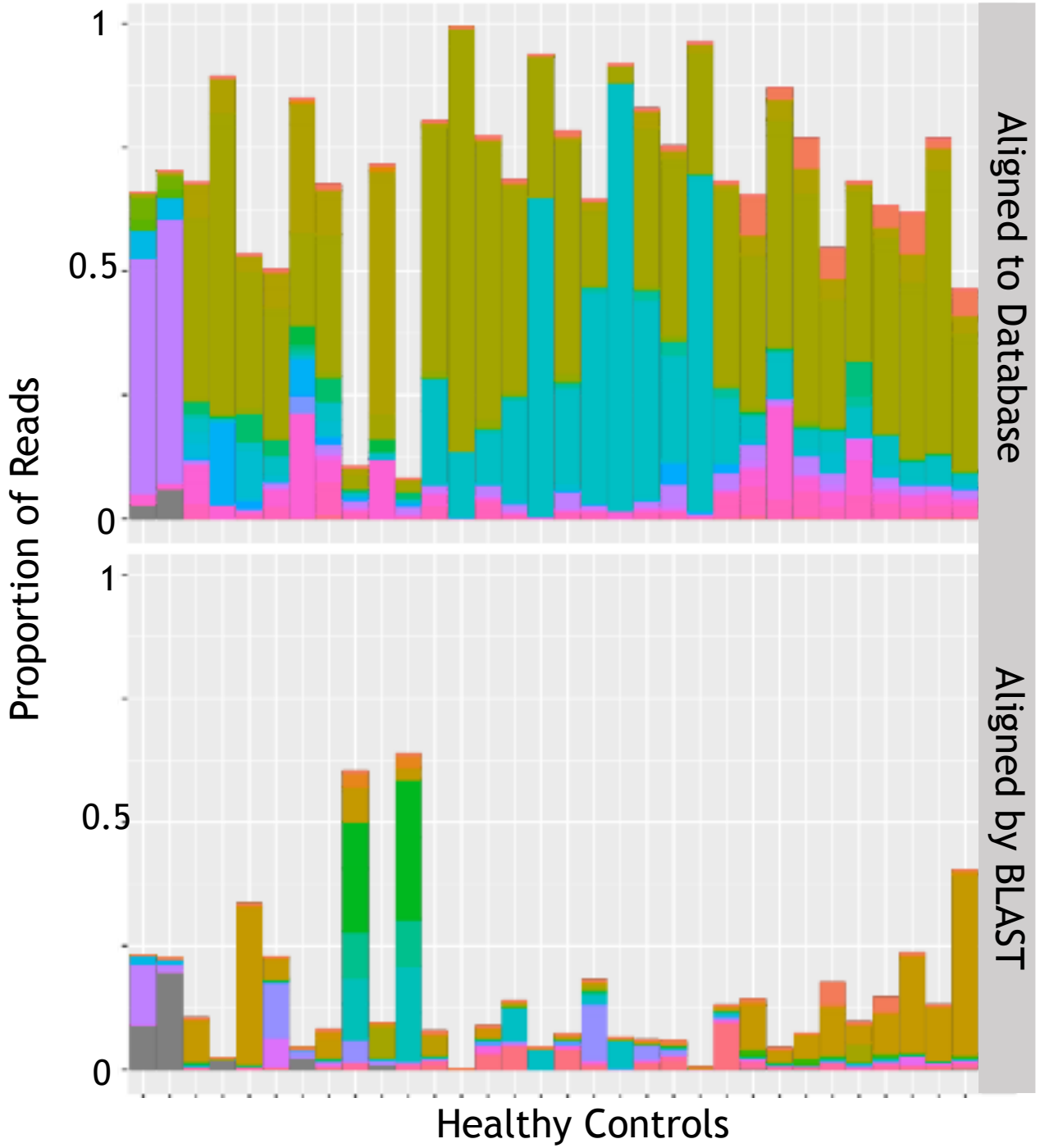
Identifying missing microbial taxa



363,525,943 reads BLASTn to HTGs, viral, EST_others, and nt databases

Retain top hit from each sequence >50 coverage and >90% identity

18% of unaligned reads aligned to at least one genome in a public database

Identifying missing microbial taxa

- *Malassezia*, *Cutibacterium*, *Acinetobacter*
- *Pinus*, *Canis*, *Triticum*