

S1. Alignment results

- Data reported in the spreadsheet were generated with the following command:

```
diamond blastp --db reference_database.dmnd --query query_file.fasta --  
out reference_query_diamond.tsv --outfmt 6 qseqid sseqid pident length  
mismatch gapopen qstart qend sstart send evaluate bitscore qlen qseq stitle  
slen sseq staxids sscinames sskingdoms --threads 6 --log
```

- Spreadsheet tabs are named according to the following format:

reference_query_diamond

- Spreadsheet column are named as follow:

```
qseqid sseqid pident length mismatch gapopen qstart qend sstart send  
evaluate bitscore qlen qseq stitle slen sseq staxids sscinames sskingdoms
```

Detailed description is available at the software distribution site:

<https://github.com/bbuchfink/diamond>

<https://github.com/bbuchfink/diamond/wiki/3.-Command-line-options>

S2. Additional plots.

- Figure S2.1. Alignment of the AMPs in the APD3 database: hits distribution against APD3 database, size of the point is proportional to the hit count.
- Figure S2.2. Alignment of the DBAASP against APD3 database: size of the point is proportional to the hit count.
- Figure S2.3. Alignment of the AMPs in the DBAASP database: hits distribution against DBAASP database, size of the point is proportional to the hit count.
- Figure S2.4. Histogram of the AMPs sequences percent identity: content of the DBAASP database compared against itself.