Single mutations can change enzymatic function

Homework #2
Plant Breeding 607
Fall '00

November 8, 2000

1 Background

In a very recent publication [1], researchers demonstrate that single amino acid substitutions can make large differences in the catalytic specificity of certain cytochrome P450 oxygenases. These conclusions underscore the extreme difficulty in predicting biochemical specificity and activity directly from primary sequence. In particular, careful measurements of biochemical function and ingeneous use of "domain swapping" and reciprocal site-directed mutagenesis on show that a single amino acid exchange can lead to the change of catalytic efficiency and specificity of limonene-6-hydroxylase from spearmint to mimic the limonene-3-hydroxylase from peppermint. The converse substitution in limonene-3-hydroxylase does not lead to limonene-6-hydroxylase activity. This lack of reciprocity may be due to differences in the remainder of the proteins.

The cytochrome P450 family show considerable variability in primary structure and biochemical specificity, and include representatives from both prokaryotes and eukaryotes. This structural and phylogenetic diversity can substantially complicate sequence analysis. In this homework, you will be asked to use different parameters and substitution matrices to "tune" a similarity search to detect subsets with increasing degrees of sequence variability.

2 Questions

The amino acid sequences for the hydroxylases from spearmint (SM12) and peppermint (PM2) can be found in section 3. In all cases, use the recommended gap and indel penalties.

1. Use the pairwise BLAST server at `http://www.ncbi.nlm.nih.gov/blast/bl2seq/bl2.html` to align the SM12 and PM2 sequences using the following substitution matrices:

   BLOSUM BLOSUM45, BLOSUM62, and BLOSUM80

   PAM PAM250, PAM160, PAM120, PAM70, and PAM30

   What systematic difference, if any, exist in these pairwise alignments? Why? Discuss trends alignments in terms of entropy and expected scores.

2. Use the standard BLAST server and the same BLOSUM45, BLOSUM62, and BLOSUM80 substitution matrices to identify proteins from the SwissProt and PDB databases that are similar to SM12 and PM2, respectively. What differences do you see in the nature of the alignments among the alignments as one proceeds from strong (BLOSUM80) to moderate (BLOSUM62) to weak (BLOSUM45) alignments? What systematic differences can be found between the alignments detected for PM2 and SM12 for BLOSUM80? What about for the BLOSUM45 alignments? Discuss trends alignments in terms of entropy and expected scores.

3 Amino acid sequences of spearmint and peppermint limonene hydroxylases in FASTA format.

These sequences were obtained[1] from the paper by Schalk and Croteau [1].

```
>gnl|PB607|PM2 peppermint limonene-3-hydroxylase
mellqlwsaliilvvtytisllinqwrkpkpqgkfppgppklplighlhllwgklpqhalasvakeygpvahvqlgevfs
vvlssreatkeamklvdpacanrfesigtrimwydnediifspysehwrqmrkicvsellssrnvrsfgfirqdevsrll
rhlrssagaavdmterietltcsiicraafgsvirdnaelvglvkdalsmasgfeladmfpssklln llcwnksklwrmr
rrvdtileaivdehkfkksgefggediidvlfrmqkdtqikvpittnsikafifdtfsagtetsstttlwvlaelmrnpa
vmakaqaevraalkektnwdvddvqelkymksvvketmrmhppipliprscreecvvngytipnkariminvwsmgrnpl
ywekpdtfwperfdqvskdfmgndfefvpfgagrricpglnfglanvevplaqllyhfdwklaegmkpsdmdmseaeglt
gilknnlllvptpydpss
>gnl|PB607|SM12 spearmint limonene-6-hydroxylase
meldllsaiiilvatyivsllinqwrksksqqnlppsppklpvighlhflwgglpqhvfrsiaqkygpvahvqlgevysv
vlssaeaakqamkvldpnfadrfdgigsrtmwydkddiifspyhdhwrqmrricvtellspknvrsfgyirqeeierlir
llgssggapvdvteevskmscvvvcraafgsvlkdqgslaelvkeslalasgfeladlypsswllnllslnkyrlqrmrr
rldhildgfleehrekksgefggedivdvlfrmqkgsdikipitsncikgfifdtfsagaetssttiswalselmrnpak
makvqaevrealkgktvvdlsevqelkylrsvlketlrlhppfpliprqsreecevngytipaktrifinvwaigrdpqy
wedpdtfrperfdevsrdfmgndfefipfgagrricpglhfglanveiplaqllyhfdwklpqgmtdadldmtetpglsg
pkkknvclvptlyksp
```

References

[1] Michel Schalk and Rodney Croteau. A single amino acid substitution (F363I) converts the regiochemistry of the spearmint (-)-limonene hydroxylase from a C6- to a C3-hydroxylase, *PNAS*, 97(22):11948–11953, 2000.

[1] These sequences were kindly entered by Sarah Schneider.