

1. PSI-BLAST first iteration is the same as a blastp search. Past that point, PSI-BLAST will retrieve the best hits and generate multiple alignment for those. A PSSM will be generated based on the alignment of these hits, with more conserved regions having more weight in the matrix. This way, PSI-BLAST next iteration will “value” the conserved regions and their specific amino acids higher than non-specific regions, bringing more hits with the same motif to your attention. PSI-BLAST may then repeat this process until no new hits are obtained or sooner if one would choose so.

2. PSSM is position-specific scoring matrix, and is used to store information on conserved motifs of sequences. The matrix contains information about positions of elements and their frequency among the chosen sequences. It then stores this information as log odds ratio of elements at this position and frequency– compared to if the sequence was unrelated/random. It can then be used to look for specific motifs in sequences.

3. If PSI-BLAST search converges, it means it can not find any new hits in the next iteration, and PSSM would not change if last query sequences are integrated.

4. <https://www.ncbi.nlm.nih.gov/protein/P08487.1?report=fasta>

5. <https://www.ncbi.nlm.nih.gov/protein/Q95M30.3?report=fasta>

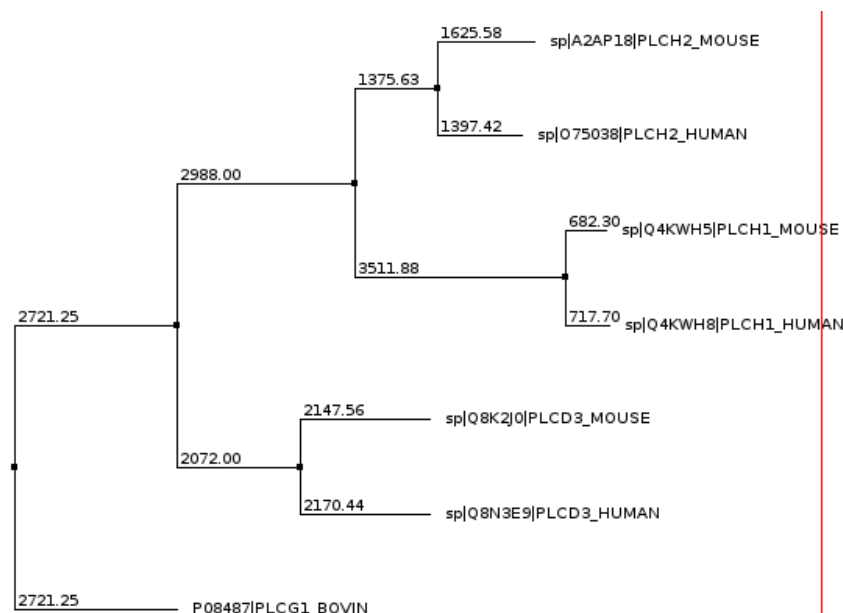
Considering “most similar” as highest value hit according to consensus on similarity, E-value and sequence length (some shorter sequences were more identical)

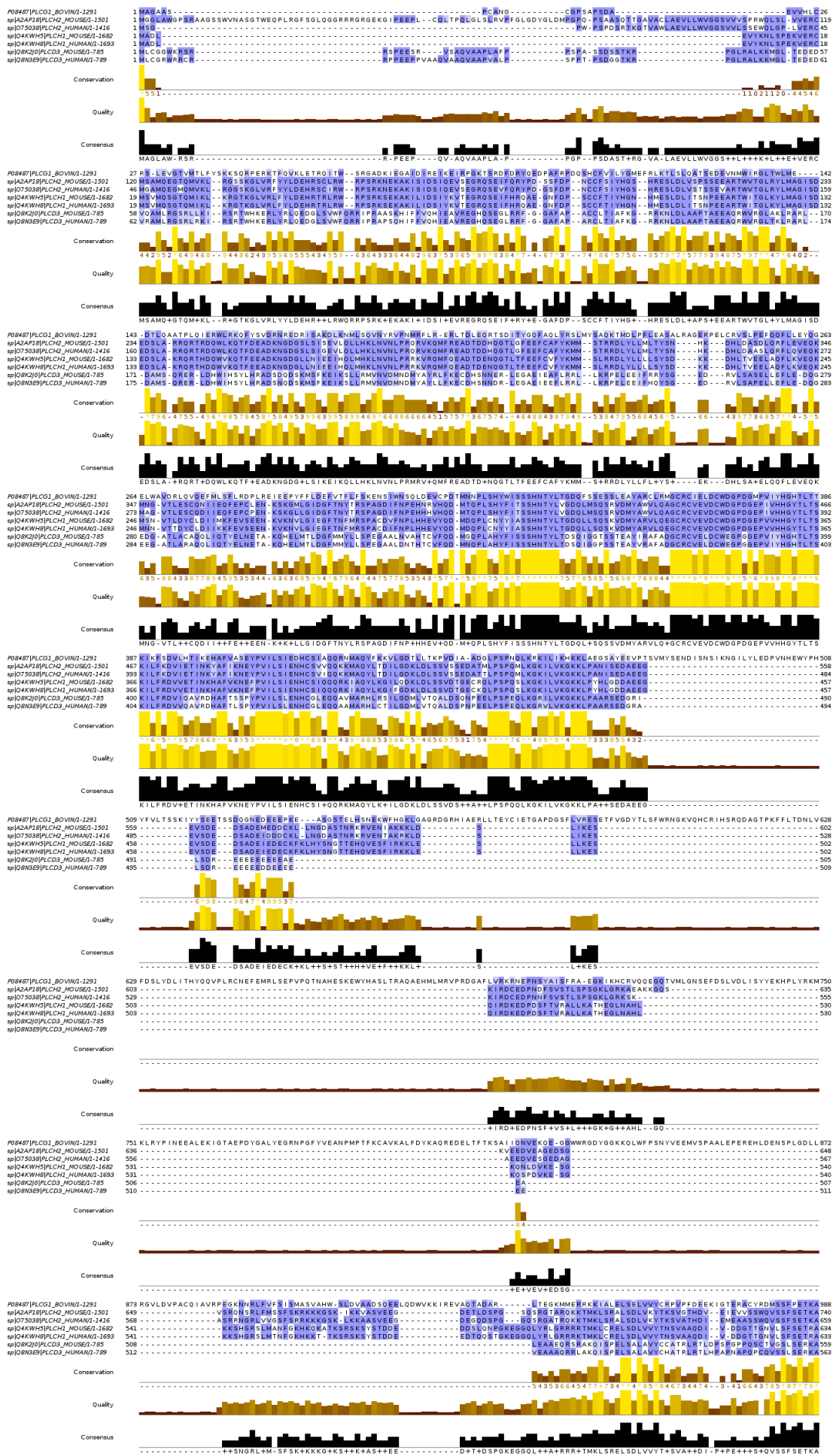
6. MSA allows us to visualise how conserved (and thus important to function) some residues and motifs are. It can also be used to calculate how (relatively) similar or different they are and find points of evolutionary divergence.

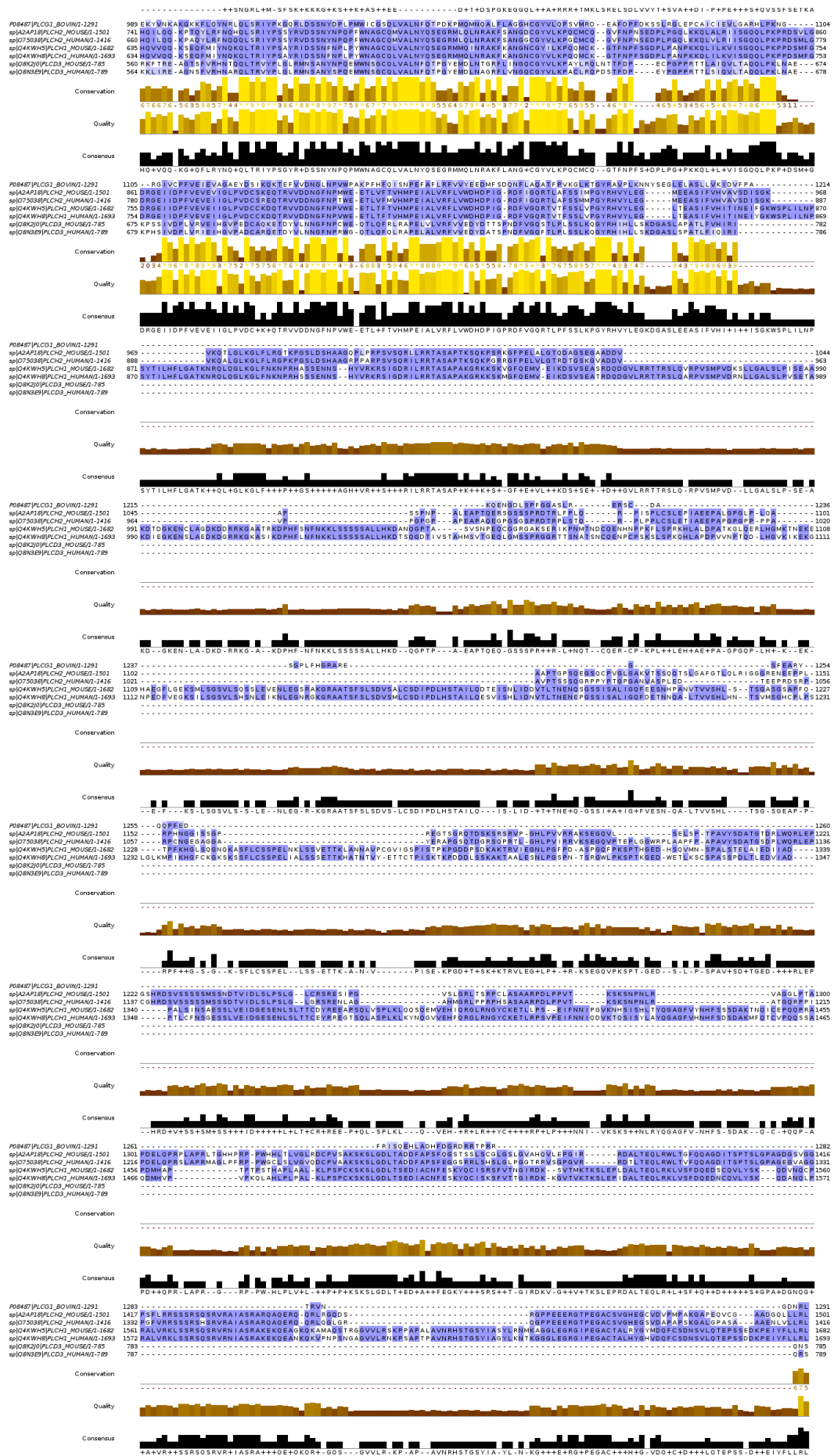
7. Conserved sequences are sequences that have remained homologous throughout evolution. Considering presence of genetic drift and steady introduction, deletion or replacement of residues, these residues or motifs remained similar, or conserved, due to their relevance to function under positive selection.

8. Gaps and non-aligned are not coloured. Residues that have positive BLOSUM score are light blue. Residues that match the consensus residue (most conserved) are dark blue.

9.







10. ClustalX colour scheme:

Residues are grouped depending on their (predicted) chemical properties.

Thresholds are set differently for different groups, and detailed threshold explanation is found here:

<http://www.jalview.org/help/html/colourSchemes/clustal.html>

If the residue passes the threshold value, it would be coloured:

K, R = Red (positive charged residues)

E, D = Purple (negative charged residues)

N, Q, S, T = Green (polar residues)

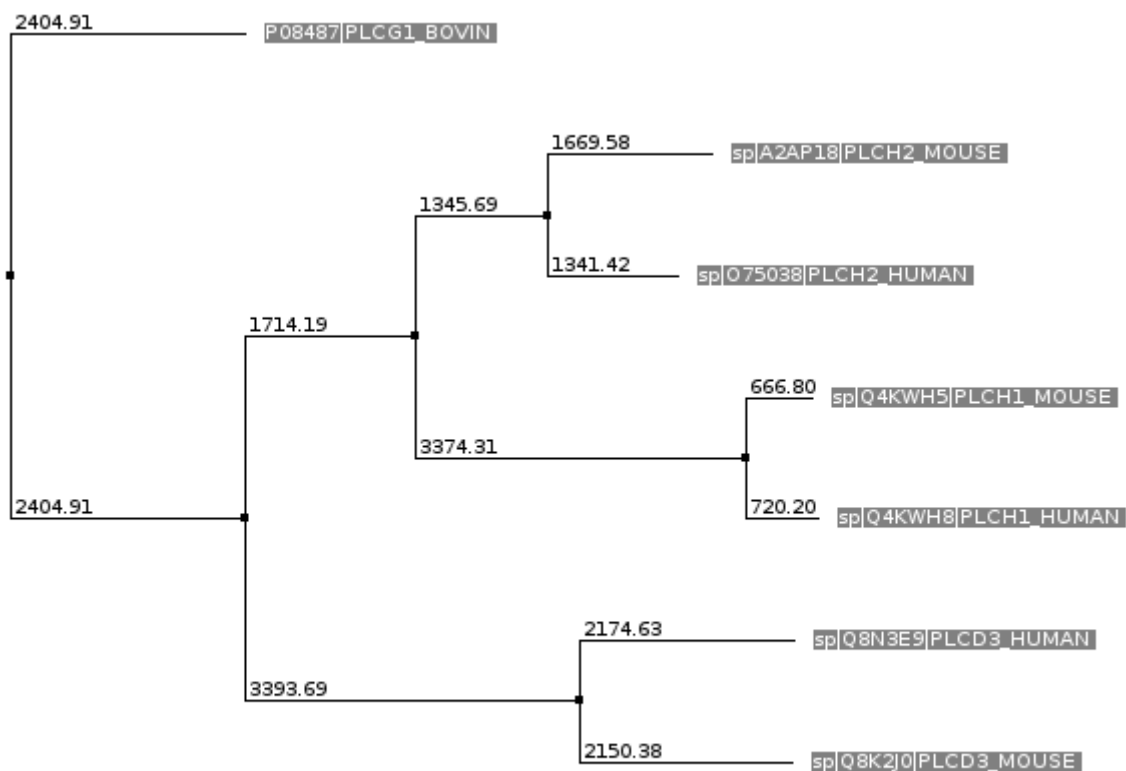
A, I, L, M, F, W, V (Sometimes C) = Blue (hydrophobic residues)

C = Pink or blue (depends on consensus threshold)

G, P = Orange, yellow respectively (>0% threshold, so always coloured)

H, Y = Cyan (aromatic)

11.



sp|A2AF18|PLCH2_MOUSE1-1501 1 MGLGAWGSRRAAGSSWNVASITWEOGLREFSLOQRRRRSRSEKILTEELQLCLTQLSLRVLFGLDYGLDMFGPQSAASQTTGAVACLAEEVLWVGGLVV-VSRWQLSVVVERCMSAMGEG 128
sp|O75038|PLCH2_HUMAN1-1416 1MSRWWSQDSRTKGTVAWLAEVLWVGGLVV-LSEWQLSRLVERCMSAMGEG 154
sp|O4KWH5|PLCH2_MOUSE2-1682 1MADLEWTKN.....LSREK.....VERCMSVMGSGT 27
sp|O4KWH5|PLCH2_HUMAN1-1693 1MADLEWTKN.....LSREK.....VERCMSVMGSGT 27
sp|O8N39|PLC3_MOUSE1-789 1MLCRRWRRCRRRPFEEPRVAQAQVAQVAALVLPSPPTDGGGKRPGLRALKKMLTEDEVRAMLRRGR 70
sp|O8K20|PLC3_MOUSE1-785 1MLCGGWKSRSSREE.....SRVSAQVAARLAFPPSAASDSSIKRPGLRALKMLTEDEVOAMLRRGR 66
P08487|PLC1_BOVIN1-1291 1MAGAASPCANR.....LGSASDSEAVHLRPLFVGLV 34

Conservation



Quality



Consensus



sp|A2AF18|PLCH2_MOUSE1-1501 129 MVVLKQSSKELVRYFYLYDEH--RCLRWPPPKKNEKAKISIDSIIEVVEGQEGEIPQVYD-GSFDP--NCCFSIYHGSHR--ELDLVSPSEEARVWVQLRPLMAGISDEDSLARQRTDRLWLCQ 251
sp|O75038|PLCH2_HUMAN1-1416 129 MVVLKQSSKELVRYFYLYDEH--RCLRWPPPKKNEKAKISIDSIIEVVEGQEGEIPQVYD-GSFDP--NCCFSIYHGSHR--ELDLVSPSEEARVWVQLRPLMAGISDEDSLARQRTDRLWLCQ 251
sp|O4KWH5|PLCH2_MOUSE2-1682 129 MVVLKQSSKELVRYFYLYDEH--RCLRWPPPKKNEKAKILDSIKYKVEGQEGEIPQVYD-GSFDP--NCCFSIYHGSHR--ELDLVSPSEEARVWVQLRPLMAGISDEDSLARQRTDRLWLCQ 251
sp|O4KWH5|PLCH2_HUMAN1-1693 129 MVVLKQSSKELVRYFYLYDEH--RCLRWPPPKKNEKAKILDSIKYKVEGQEGEIPQVYD-GSFDP--NCCFSIYHGSHR--ELDLVSPSEEARVWVQLRPLMAGISDEDSLARQRTDRLWLCQ 251
sp|O8N39|PLC3_MOUSE1-789 129 LKNTSRWHKELRYLROEDLSVWFORIPPAASSHIFFVOHLEAVREGHSGGRPFEE--GAFAP--ACCLTIAFKERR--KNLDLAPTAEEAQRWVRWLTKEARL--DAMSOERLDHWHSY 130
sp|O8K20|PLC3_MOUSE1-785 129 LKNTSRWHKELRYLROEDLSVWFORIPPAASSHIFFVOHLEAVREGHSGGRPFEE--GAFAP--ACCLTIAFKERR--KNLDLAPTAEEAQRWVRWLTKEARL--DAMSOERLDHWHSY 130
P08487|PLC1_BOVIN1-1291 129 MLRLYSKKSSPPEKRTFQVKLETRQITWRRADKLEAIDQIRKEIPQVYD-RQEDKQDQARRPDQSFEVILTGMEFLKLSLQAALGDEVNWLRLTWMLELQAALPLQIE--WLKQ 160

Conservation



Quality



Consensus



sp|A2AF18|PLCH2_MOUSE1-1501 252 FDEADNGDGLSIEVLLDLHKLNVNLRORVKMPFREADTDHGGTLGFEFCFAFY--MMSTRDLYLMLTYNHNK----HDAASLQRFLEVEKMGVPLESCNIIQEPFECLEN 570
sp|O75038|PLCH2_HUMAN1-1416 252 FDEADNGDGLSIEVLLDLHKLNVNLRORVKMPFREADTDHGGTLGFEFCFAFY--MMSTRDLYLMLTYNHNK----HDAASLQRFLEVEKMGVPLESCNIIQEPFECLEN 570
sp|O4KWH5|PLCH2_MOUSE2-1682 252 FDEADNGDGLSIEVLLDLHKLNVNLRORVKMPFREADTDHGGTLGFEFCFAFY--MMSTRDLYLMLTYNHNK----HDAASLQRFLEVEKMGVPLESCNIIQEPFECLEN 570
sp|O4KWH5|PLCH2_HUMAN1-1693 252 FDEADNGDGLSIEVLLDLHKLNVNLRORVKMPFREADTDHGGTLGFEFCFAFY--MMSTRDLYLMLTYNHNK----HDAASLQRFLEVEKMGVPLESCNIIQEPFECLEN 570
sp|O8N39|PLC3_MOUSE1-789 252 LHRASNDQSKMSKEIKSLRMVYDMDMYAYLLFCCHDSNRL-LEGAEIEFLR--RLKPELEEFIFRYSGEDR--VLSAPELLEFLE-DGGEBAALARACLOTIVELNETAN 907
sp|O8K20|PLC3_MOUSE1-785 252 LHRASNDQSKMSKEIKSLRMVYDMDMYAYLLFCCHDSNRL-LEGAEIEFLR--RLKPELEEFIFRYSGEDR--VLSAPELLEFLE-DGGEBAALARACLOTIVELNETAN 903
P08487|PLC1_BOVIN1-1291 252 FYSVDRFEDRLAKDLKNMLSQVNFVNMFLR-ELLTDLQRTSDITVQPAQLMSLYNARKTMDLFEALALAGERFELCRVSLPEEQFLLEKQELWAVORLQVFMFLFDRFLRE 289

Conservation



Quality



Consensus



sp|A2AF18|PLCH2_MOUSE1-1501 371 KBMLGIDGFNTYRSPAGDIFNPENHNVHQD-MTPQLSHYFISSSHNTYLTGQDL+SQSSVDMYARVLQ+GCRRCVEVDCWGDGPDGEPVHHGYTLTSKILFQDVIEINKAFIKNEYPVILSIEHNC 499
sp|O75038|PLCH2_HUMAN1-1416 371 KBMLGIDGFNTYRSPAGDIFNPENHNVHQD-MTPQLSHYFISSSHNTYLTGQDL+SQSSVDMYARVLQ+GCRRCVEVDCWGDGPDGEPVHHGYTLTSKILFQDVIEINKAFIKNEYPVILSIEHNC 499
sp|O4KWH5|PLCH2_MOUSE2-1682 371 KBMLGIDGFNTYRSPAGDIFNPENHNVHQD-MTPQLSHYFISSSHNTYLTGQDL+SQSSVDMYARVLQ+GCRRCVEVDCWGDGPDGEPVHHGYTLTSKILFQDVIEINKAFIKNEYPVILSIEHNC 499
sp|O4KWH5|PLCH2_HUMAN1-1693 371 KBMLGIDGFNTYRSPAGDIFNPENHNVHQD-MTPQLSHYFISSSHNTYLTGQDL+SQSSVDMYARVLQ+GCRRCVEVDCWGDGPDGEPVHHGYTLTSKILFQDVIEINKAFIKNEYPVILSIEHNC 499
sp|O8N39|PLC3_MOUSE1-789 371 KNLVGLGDFNTYRSPAGDIFNPENHNVHQD-MTPQLSHYFISSSHNTYLTGQDL+SQSSVDMYARVLQ+GCRRCVEVDCWGDGPDGEPVHHGYTLTSKILFQDVIEINKAFIKNEYPVILSIEHNC 499
sp|O8K20|PLC3_MOUSE1-785 371 KNLVGLGDFNTYRSPAGDIFNPENHNVHQD-MTPQLSHYFISSSHNTYLTGQDL+SQSSVDMYARVLQ+GCRRCVEVDCWGDGPDGEPVHHGYTLTSKILFQDVIEINKAFIKNEYPVILSIEHNC 499
P08487|PLC1_BOVIN1-1291 371 HELNLTLDGFMMLYLLREBAALNVAHTCTCFVD-MGDPPLAHYFISSSHNTYLTGQDL+SQSSVDMYARVLQ+GCRRCVEVDCWGDGPDGEPVHHGYTLTSKILFQDVIEINKAFIKNEYPVILSIEHNC 419

Conservation



Quality



Consensus



sp|A2AF18|PLCH2_MOUSE1-1501 500 VVQKMAQYLTDLGDKLDLSVSEBATMLPSPMMLKGLVKGKLPANISDAEEGEVDEDEADEIDDDCKL-LNGDASNRKRVENIAKKKLDLSLKEKIRDCEDNDFVSTLSRSGKLRKA 628
sp|O75038|PLCH2_HUMAN1-1416 500 VVQKMAQYLTDLGDKLDLSVSEBATMLPSPMMLKGLVKGKLPANISDAEEGEVDEDEADEIDDDCKL-LNGDASNRKRVENIAKKKLDLSLKEKIRDCEDNDFVSTLSRSGKLRKA 628
sp|O4KWH5|PLCH2_MOUSE2-1682 500 VVQKMAQYLTDLGDKLDLSVSEBATMLPSPMMLKGLVKGKLPANISDAEEGEVDEDEADEIDDDCKL-LNGDASNRKRVENIAKKKLDLSLKEKIRDCEDNDFVSTLSRSGKLRKA 628
sp|O4KWH5|PLCH2_HUMAN1-1693 500 VVQKMAQYLTDLGDKLDLSVSEBATMLPSPMMLKGLVKGKLPANISDAEEGEVDEDEADEIDDDCKL-LNGDASNRKRVENIAKKKLDLSLKEKIRDCEDNDFVSTLSRSGKLRKA 628
sp|O8N39|PLC3_MOUSE1-789 500 EDAAMARHCTILGDMVLVTOALDNPFEELPSPOLKGRVLYKGLKPAAR-EDGR--ALDREEEEEEED----- 505
sp|O8K20|PLC3_MOUSE1-785 500 EDAAMARHCTILGDMVLVTOALDNPFEELPSPOLKGRVLYKGLKPAAR-EDGR--ALDREEEEEEED----- 501
P08487|PLC1_BOVIN1-1291 500 ADGRNMAHYKVKVGGDTLLTKVYDIAADR-LPFRNQLRGLILHKLKAFESAYEYVTVSMYSENDLSNISKNLILLEDVYHWEVYHYFLVTSKILVYESTSQDN 529

Conservation



Quality



Consensus



sp|A2AF18|PLCH2_MOUSE1-1501 629 EAKKGQSKVEQVAGEDSVSRONSRLFMSSFSKRKKKSKIK-KVA-VEEGDETLDSPG-----SRSRGTARQKTMKLRALSDLVKYTK--VGTVDVIEVVSQWVSFSETKAHOILQKQVGL 752
sp|O75038|PLCH2_HUMAN1-1416 629 EAKKGQSKVEQVAGEDSVSRONSRLFMSSFSKRKKKSKIK-KAA-VEEGDETLDSPG-----GSRBGTARQKTMKLRALSDLVKYTK--VATHDVIEEASVQWVSFSETKAHOILQKQVGL 752
sp|O4KWH5|PLCH2_MOUSE2-1682 629 EAKKGQSKVEQVAGEDSVSRONSRLFMSSFSKRKKKSKIK-KAA-VEEGDETLDSPG-----GSRBGTARQKTMKLRALSDLVKYTK--VATHDVIEEASVQWVSFSETKAHOILQKQVGL 752
sp|O4KWH5|PLCH2_HUMAN1-1693 629 EAKKGQSKVEQVAGEDSVSRONSRLFMSSFSKRKKKSKIK-KAA-VEEGDETLDSPG-----GSRBGTARQKTMKLRALSDLVKYTK--VATHDVIEEASVQWVSFSETKAHOILQKQVGL 752
sp|O8N39|PLC3_MOUSE1-789 629 EAKKGQSKVEQVAGEDSVSRONSRLFMSSFSKRKKKSKIK-KAA-VEEGDETLDSPG-----GSRBGTARQKTMKLRALSDLVKYTK--VATHDVIEEASVQWVSFSETKAHOILQKQVGL 752
sp|O8K20|PLC3_MOUSE1-785 629 EAKKGQSKVEQVAGEDSVSRONSRLFMSSFSKRKKKSKIK-KAA-VEEGDETLDSPG-----GSRBGTARQKTMKLRALSDLVKYTK--VATHDVIEEASVQWVSFSETKAHOILQKQVGL 752
P08487|PLC1_BOVIN1-1291 629 EAKKGQSKVEQVAGEDSVSRONSRLFMSSFSKRKKKSKIK-KAA-VEEGDETLDSPG-----GSRBGTARQKTMKLRALSDLVKYTK--VATHDVIEEASVQWVSFSETKAHOILQKQVGL 752

Conservation



Quality



Consensus



sp|A2AF18|PLCH2_MOUSE1-1501 753 FPHQHLRLIYPSISVDSNHYNDPFWHAGCQVVALNYSEGRMLDLNRAFSANGCGVYLKPKQCMCGVFNHNEEDPLPQDLKQLALRIISGGQLKPRRDLQDGEIIDPFVEVEIIGLPVDC 882
sp|O75038|PLCH2_HUMAN1-1416 753 FPHQHLRLIYPSISVDSNHYNDPFWHAGCQVVALNYSEGRMLDLNRAFSANGCGVYLKPKQCMCGVFNHNEEDPLPQDLKQLALRIISGGQLKPRRDLQDGEIIDPFVEVEIIGLPVDC 882
sp|O4KWH5|PLCH2_MOUSE2-1682 753 FPHQHLRLIYPSISVDSNHYNDPFWHAGCQVVALNYSEGRMLDLNRAFSANGCGVYLKPKQCMCGVFNHNEEDPLPQDLKQLALRIISGGQLKPRRDLQDGEIIDPFVEVEIIGLPVDC 882
sp|O4KWH5|PLCH2_HUMAN1-1693 753 FPHQHLRLIYPSISVDSNHYNDPFWHAGCQVVALNYSEGRMLDLNRAFSANGCGVYLKPKQCMCGVFNHNEEDPLPQDLKQLALRIISGGQLKPRRDLQDGEIIDPFVEVEIIGLPVDC 882
sp|O8N39|PLC3_MOUSE1-789 753 FPHQHLRLIYPSISVDSNHYNDPFWHAGCQVVALNYSEGRMLDLNRAFSANGCGVYLKPKQCMCGVFNHNEEDPLPQDLKQLALRIISGGQLKPRRDLQDGEIIDPFVEVEIIGLPVDC 882
sp|O8K20|PLC3_MOUSE1-785 753 FPHQHLRLIYPSISVDSNHYNDPFWHAGCQVVALNYSEGRMLDLNRAFSANGCGVYLKPKQCMCGVFNHNEEDPLPQDLKQLALRIISGGQLKPRRDLQDGEIIDPFVEVEIIGLPVDC 882
P08487|PLC1_BOVIN1-1291 753 FPHQHLRLIYPSISVDSNHYNDPFWHAGCQVVALNYSEGRMLDLNRAFSANGCGVYLKPKQCMCGVFNHNEEDPLPQDLKQLALRIISGGQLKPRRDLQDGEIIDPFVEVEIIGLPVDC 882

Conservation



Quality



Consensus



