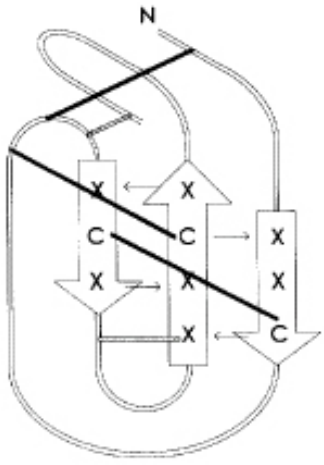
Jamie Schroeder
Biochemistry 218 Final Project
Autumn 2001

Characterization of Hypervariability in Conotoxins

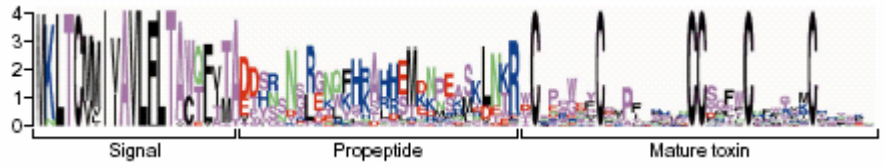
I. Introduction

Hypervariability occurs in genes that are responsible for encoding proteins responsible for direct interface between species, such as peptide toxins and various immune molecules. Variability sequence data has been studied extensively for over thirty years in the immunoglobin gene family and techniques for its analysis were pioneered in part by Wu and Kabat (Johnson et al, 2000). They defined a quantitative measure for variability and aligned sequence data to form what are known as Wu-Kabat plots. This type of plot was used in determining the complementarity-determining-region (CDR) of immunoglobin (Wu et al, 1970), and could be used to characterize similar locations in other hypervariable genes such as conotoxins.

Conotoxins encompass a vast array of small peptides which are used by piscivorous marine cone snails to immobilize and capture prey. These cone snails are gastropods, members of genus conus of the conidae family, and all 500 species are fish predators (Kohn et al, 1956). Each species has a repertoire of 50-200 different short peptide fragments that they use in their venom. The small venom peptides (10-35 amino acids in length) are processed from larger preproprotein precursors (ca. 100 amino acids) and function by inhibiting ion channels to induce excitotoxic shock and neuromuscular paralysis in prey (Terlau et al, 1996). This is occurs by blockage or flooding of various ion channels in order to either block synaptic transmission or overwhelm the nervous system (England et al, 1998; Rigby et al, 1999; Shon et al, 1998). The high specificity of these toxins gives them possible therapeutic value as antinociceptive agents or as components of anticonvulsant drugs (Bowersox et al, 1998).



The sequences of mature conotoxin peptides are interspersed with conserved cysteine residues that form disulfide bonds which provide structural integrity to the highly variable loops in between these residues. The motif of this spacing is known as the inhibitor cysteine knot (ICK) and has been described by the consensus sequence $-CX_{3-7}CX_{3-6}CX_{0-5}CX_{1-4}CX_{4-13}C-$ (Norton et al, 1998). Spacing of these cysteines into this type of motif is important for proper folding of the toxins. The gene family is very rapidly diversifying and evolving, as had been demonstrated by a comparison showing an incredible amount of variability specific to the toxin encoding regions but not as much in the preproprotein region (Duda et al, 1999). There have been recent studies on this type of hypervariability done on alignments across several species of conus' (Conticello et al, 2000; 2001). Methods that were used to characterize and visualize this diversity include sequence logo plots as have been more recently described (Schneider et al, 1990). The sequence logo plot was helpful in identifying conserved residues, but cannot indicate the degree of the variability in the variable regions.



As variability was defined quantitatively by Wu, it is the ratio of the **number of different residues** found across the aligned sequences at a given position to the **frequency of the most common residue** at that position. This formula is applied to each residue in a sequence alignment to obtain a plot visualizing the areas of high and low variability. This method offers the most useful portrait of hypervariability. In this study a script was written using Microsoft Excel to generate Wu-Kabat plots from a multiple sequence alignments. The resulting plots were used to critically analyze the methods in bioinformatics that are involved with obtaining and aligning the sequences, and to quantitatively characterize the hypervariability of the conotoxin gene family.

II. Methods

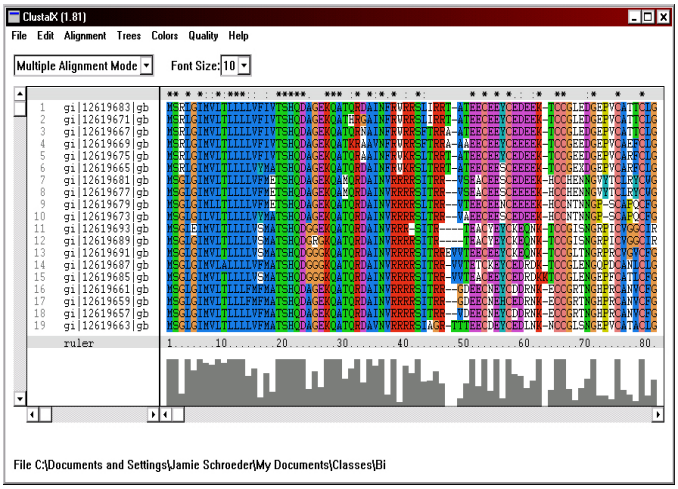
- Sequences were selected by NCBI Taxonomy Browser for genus conus (http://www.ncbi.nlm.nih.gov/htbin-post/Taxonomy/wgetorg?id=6490).
- Sequences were retrieved using NCBI Entrez (http://www.ncbi.nlm.nih.gov/Entrez/).
- Sequences were aligned using CLUSTAL X (1.8) for the PC.
- Wu-Kabat plots made from alignments (Excel Template Protein_Variability.xlt).
- Alignments were used to generate both Hidden Markov Models and to obtain characteristic sequences to be used to find related proteins by both heuristic and dynamic methods. (HMMER2.0)
- Swissprot was searched using HMM iterated search, and by PSI-Blast search with a characteristic sequence from the alignment on DeCypher (http://decypher.stanford.edu/).
- The relevant and significant sequences found in Swissprot were added to their respective alignments and the variability plots were updated with the new larger alignment sets.

III. Results

When selecting sequences for alignment and Wu-Kabat plots, it was important that there were a substantial numbers of sequences that were well aligned to get an accurate portrait of the variability of those genes in question. The NCBI Taxonomy browser allows one to see how many sequences are available for each species. The species that were selected to undergo this type of variability analysis for this study had at least 50 protein sequences, usually of which most were immature toxins. Those selected were Conus arenatus (CA), Conus ebraeus (CE), Conus pennaceus (CP), Conus tessulatus (CT), Conus textile (CX), and Conus ventricosus (CV). The sequences of these species also had eight Popset listings for protein collections used in making alignments for the Conticello, 2001 hypervariability study. These fourteen groups of sequences were the starting point from which alignment and analysis took place

All the protein sequences from each of these groups were downloaded in FASTA and opened using ClustalW. Since the six collections based on species contained all types of proteins that had been sequenced and cataloged, and not just toxins, only those sequences that were had the ICK motif were kept. These sequences were presumably toxins and when they were aligned, the alignment was refined manually to ensure cysteines were matched. Manual refinement of these sequences involved selecting those regions surrounding a prospective cysteine alignment and doing local alignments weighted for the cysteine. ClustalW was able to align all the ICK

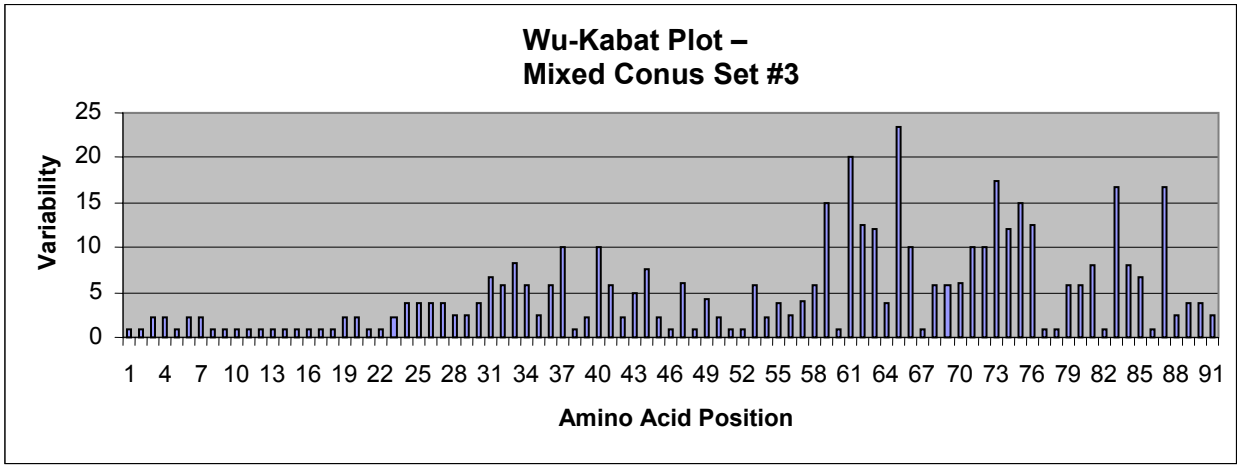
motif cysteines in each of the six species' alignment. Similarly, the collections designated by popset were aligned using the same method of ClustalW with manual refinement.



These alignments were saved as 'Plain sequences' and then designated to have a .csv extension (comma-separated-values format) so they would be opened by Microsoft Excel. This opened the aligned sequences each as strings contained in a single column, each sequence with their own row. The excel scripts that were written parsed each character from each string such that each row would represent a sequence and each column would be a single position leaving one amino acid in each column. This was accomplished using the string-parsing function **=MID($A(2),(B)$1,1)** which queries the cell containing the string (column 1) and pulls out the number for its position. And this was extrapolated to the entire grid. The most common amino acid was determined using **=CHAR(MODE(CODE(B2:B11)))** where char and code are used since mode takes numbers for inputs. The frequency of that AA was determined by the ratio of the number of most common to the total number **=COUNTIF**(B2:B11,B12)/**COUNTA**(B2:B11). The number of different AA's at each position was calculated by counting each letter that was present in each column using the formula

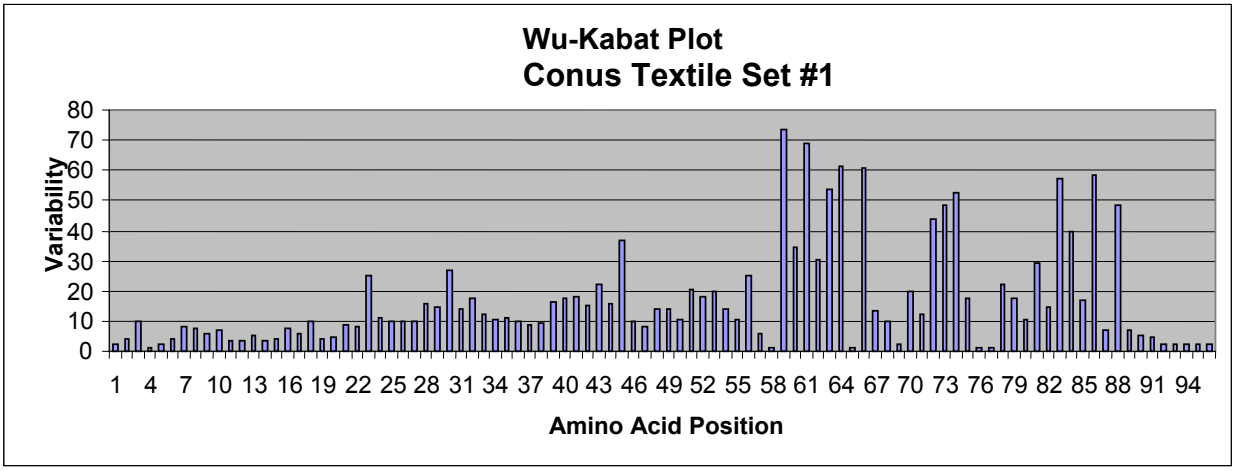
=SUM(IF(COUNTIF(B2:B12,"a")>0,1,0),IF(COUNTIF(B2:B12,"b")>0,1,0),IF(COUNTIF(B2:B12,"c")>0,1,0),IF(COUNTIF(B2:B12,"d")>0,1,0),IF(COUNTIF(B2:B12,"e")>0,1,0),~etc…~**,(COUNTIF(B2:B12,"z")>0,1,0))**

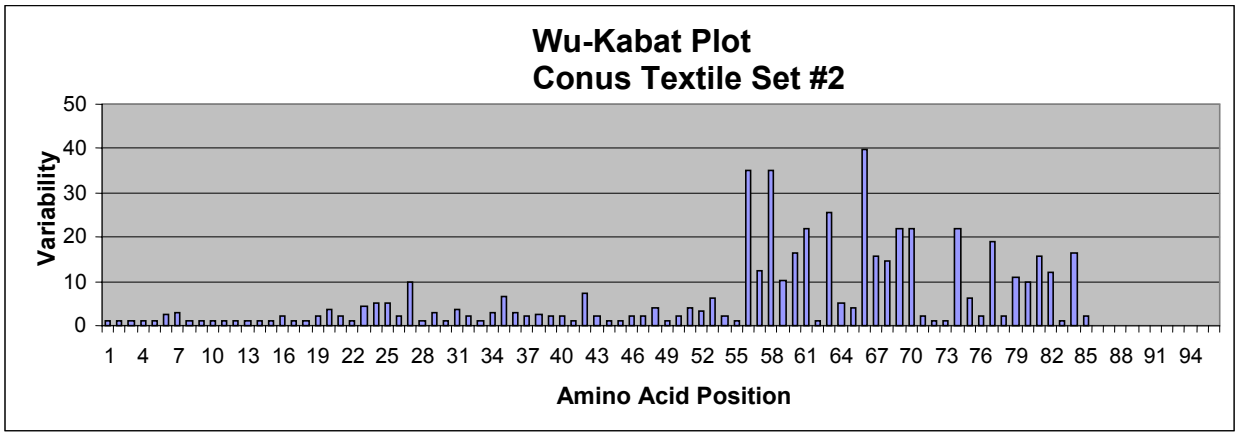
Finally, variability is calculated as the ratio of the preceding two values, the number of different residues to the frequency of the most common residue. This quantity of variability would be plotted by position, yielding a graph that looked something like the following figure, generated from one of the Popset protein collection alignments.



The peak variability values obtained in this graph are misleadingly low compared to a typical Wu-Kabat plot. This is attributed to the smaller dataset used to make the alignment for this plot. It is still clear from this plot that by far the greatest variability occurs in the toxin region (~58~88) and has the conserved ICK motif through the alignment, with those cysteines having a variability of one (complete conservation through all alignments). The first 58 bases show very low variability, indicating possible presence of selection pressure to keep the preproprotein region of the immature peptide conserved.

Similar results were obtained from alignments within a single species, as far as showing the same characteristic pattern of lower variability up until the first conserved cysteine, followed by high variability in the intra-cysteine loops. The following two graphs are from related alignments from the same species, the first from a larger alignment set, and the second a subset of the first. The first showed good contrast in variability between the prepro protein region and toxin region, but had values of variability in parts of the prepro region that were high relative to other alignments. A percentage of the sequences that had divergent prepro regions were removed from the alignment, and the resulting plot showed both better contrast between prepro and toxin regions and lower absolute values of variability in the prepro region. The rationale for removing some sequences would be that the genes aligned to each other in the first set may have been different subsets of the same family, possibly with different targets, even though they shared the same motif structure common to that species' toxins.





In order to increase the significance of each of the Wu-Kabat plots, more sequences from each group needed to be obtained to get more accurate pictures of variability especially compared with previously characterized immunoglobins. Also, if protein databases were searched taking into account the conserved preproprotein region of the conopeptide along with the variable nature of the toxin segment, the quality of the searching method could be gauged by the degree to which they improved on the data in the Wu-Kabat plot. This improvement would most likely involve a reduction of the variability noise in the more conserved regions of the peptide, and additional sequence data might clarify or increase variability peaks in the hypervariable region.

Therefore when all the groups of proteins had been plotted, other proteins of sequence homology were obtained using the groups' alignments by both heuristic and dynamic methods. From each alignment, a Hidden Markov Model was generated using HMMER2.0 algorithm. The

model was used to search Swissprot using DeCypher. The any proteins that were found were added back into the alignment with which they were found, and then the variability plot updated. The same procedure was used on proteins that were found by taking a characteristic sequence from each alignment's conserved region with an ICK type motif from the variable region and using it to performing a PSIBlast search on Swissprot with DeCypher.

An example of these two types of searches and the effects on hypervariability plots is illustrated below using mixed conus alignment #3. Many of the alignments however had no matches in Swissprot.

- **HMM search**

```
HMMER2.0
NAME  query\CGI_29c5dws5230ec_H
DESC
LENG  87
ALPH  Amino
RF    no
CS    no
MAP   yes
COM   c:\hmmbuild.exe ./output/CGI_29c5ews44cd48.out.tmp
NSEQ  8
DATE  Wed Dec 05 14:43:40 2001
CKSUM 6444
XT      -8455     -4  -1000  -1000  -8455     -4  -8455     -4
NULT       -4  -8455
NULE      595  -1558     85    338   -294    453  -1158    197    249    902  -1085   -142    -21   -313     45    531
HMM         A      C      D      E      F      G      H      I      K      L      M      N      P      Q      R      S
          m->m   m->i   m->d   i->m   i->i   d->m   d->d   b->m   m->e
          -415      *  -2000
      1  -1736  -1678  -2975  -2821   -903  -2619  -2137   -342  -2336     24   4777  -2581  -2886  -2281  -2280  -2241
      -   -149   -500    233     43   -381    399    106   -626    210   -466   -720    275    394     45     96    359
      -    -21  -6672  -7714   -894  -1115   -701  -1378   -415      *
      2  -1892  -2397  -1701  -1300  -3050  -2216   -948  -2829   3583  -2696  -2116  -1348  -2450   -619     78  -1865
      -   -149   -500    233     43   -381    399    106   -626    210   -466   -720    275    394     45     96    359
      -    -21  -6672  -7714   -894  -1115   -701  -1378      *      *
      3  -1658  -1349  -3713  -3232   -402  -3390  -2426    705  -2896   2529    705  -3027  -3187  -2430  -2772  -2626
      -   -149   -500    233     43   -381    399    106   -626    210   -466   -720    275    394     45     96    359
      -    -21  -6672  -7714   -894  -1115   -701  -1378      *      *
      4   -436   -748  -1785  -1386   -969  -1557  -1069   -149  -1062   -437   2071  -1211  -1914  -1007  -1218   -768
      -   -149   -500    233     43   -381    399    106   -626    210   -466   -720    275    394     45     96    359
      -    -21  -6672  -7714   -894  -1115   -701  -1378      *      *
      5  -1667   5612  -3405  -3570  -2842  -2113  -2800  -2422  -3305  -2879  -2567  -2846  -2715  -3191  -3057  -1987
      -   -149   -500    233     43   -381    399    106   -626    210   -466   -720    275    394     45     96    359
      -    -21  -6672  -7714   -894  -1115   -701  -1378      *      *
      6  -1234  -1069  -3156  -2686   -523  -2790  -1966    702  -2319    531   4044  -2438  -2805  -2041  -2288  -1999
```

  - First Iteration HMM results-

RANK SCORES QF TARGET LOCUS NAME ACCESSION# TF TARGET P_SCORE DESCRIPTION

1 146.19 1 CXDA_CONTE 1 SWISSPRO 9.8e-045 P18511
conus textile (cloth-of-gold cone). delta-con

2 141.50 1 CXK2_CONTE 1 SWISSPRO 2.5e-043 P18513
conus textile (cloth-of-gold cone). conotoxin

3 140.70 1 CXK1_CONTE 1 SWISSPRO 4.4e-043 P18512
conus textile (cloth-of-gold cone). Conotoxin

- **PSIBlast Search**
  - Characteristic sequence –

mkltcmmivavlfltawtfvtaddsrngleylfpkahyemnpeasklnkkdcxxxxxxcxxxxxxxxccxxxcxxxcx

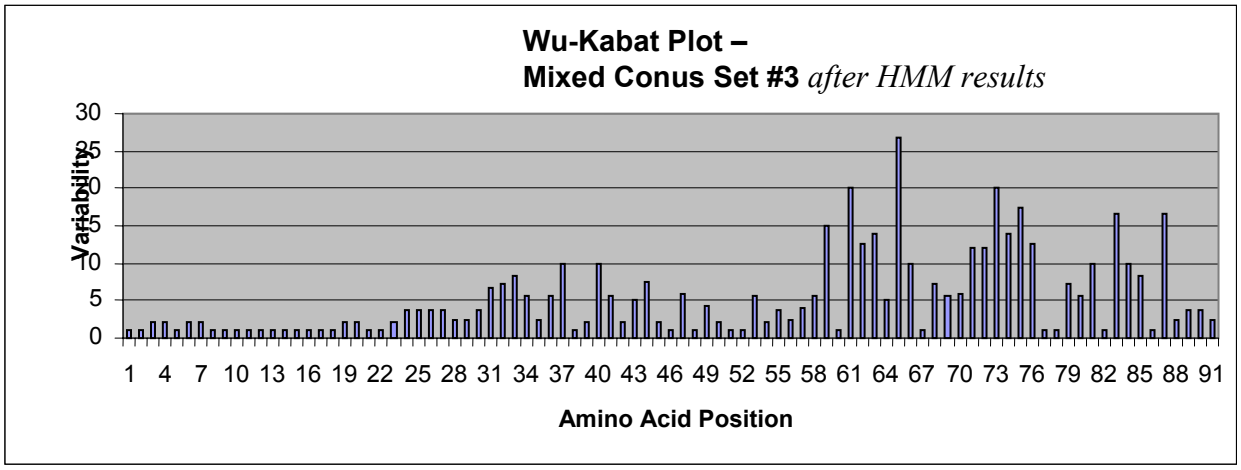
  - Results –

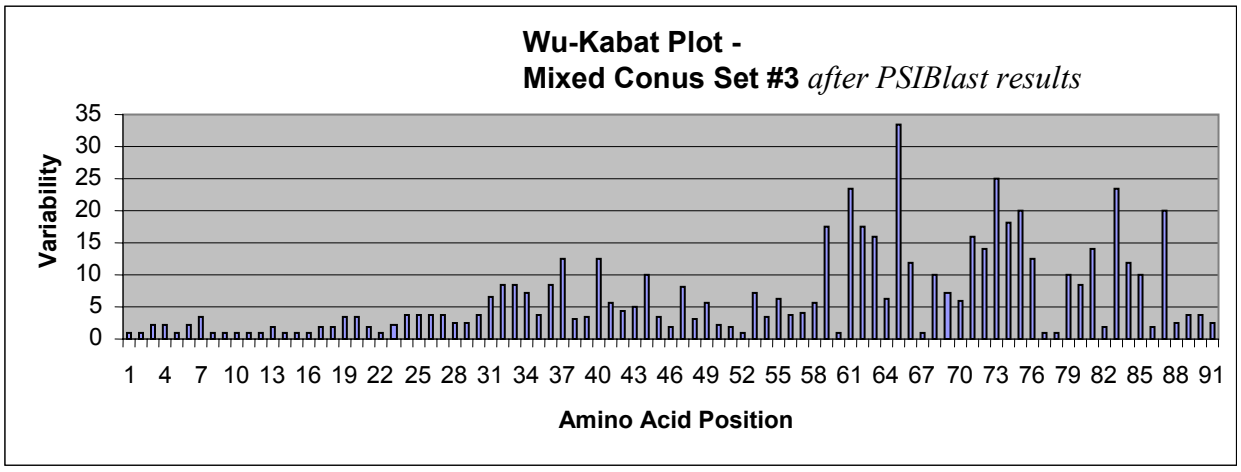
RANK Sequences producing significant alignments: (bits) Value

1 CXDA_CONTE P18511 conus textile (cloth-of-gold cone). delta-con... 110 8e-025
2 CXK2_CONTE P18513 conus textile (cloth-of-gold cone). conotoxin... 103 1e-022
3 CXK1_CONTE P18512 conus textile (cloth-of-gold cone). conotoxin... 103 1e-022
4 CXMB_CONMR Q26443 conus marmoreus (marble cone). mu-o-conotoxin... 95 2e-020
5 CXO6_CONGE P01522 conus geographus (geography cone). omega-cono... 48 5e-006
6 CXOB_CONPE P56713 conus pennaceus. omega-conotoxin pnvib. 5/2000 56 2e-008

Only five alignments returned and HMM results in one or more iterations, and only six yielded PSIBlast results with expectation values less than .01. When these new sequences were added to sequences previously aligned and then realigned three showed a noticeable change in peak variability, while it was the HMM results that tended to show more increase in contrast between peak variability in the toxin region and variability noise in the preproprotein region, when these homologous proteins were added to the existing alignment and plotted. Among those

plots which showed a trend of increasing variability contrast when homologous Swissprot proteins were added was the mixed conus set #3. When proteins found with HMM were added, there was in increase in variability only in the toxin region and no more variability noise in the preproprotein regions. When results from the PSI Blast were added, there was a larger increase in perceived variability in the toxin region, but accompanied by a modest increase in variability noise in the preproprotein region, giving a net effect of a slight increase in overall variability contrast, but not as much as from the HMM sequence addition. However, if a different threshold had been specified for the PSIBlast search, the results of the HMM could have been emulated.





IV. Discussion

The usefulness of these plots is apparent especially in protein engineering. Variability analysis of immunoglobin led to the discovery of complementarity determining region in antibodies. Research in protein engineering showed that it was possible to modulate the specificity of the antibody by changing only the CDR (Jones et al. 1986). Similarly, the hypervariable regions of conotoxins could be complementarity determining, so changing of these residues to those similar to for example the normal substrate of a given molecule could give rise to a toxin that interfered with that molecule's normal function.

Though it seems that changes in the accuracy of variability plots might be a good measure of the effectiveness, the similarity between the findings of both the matrix and motif based searching indicated that there are probably not enough conotoxin sequences in the Swissprot database to gauge their effectiveness. The matrix based search had the advantage of not having to choose where the hypervariable region is expected to be, but had the disadvantage

of biasing the search for proteins in the hypervariable region for residues that were already found in those sequences in the hypervariable region.

Finally, comparing the usefulness of the visualization using a Wu-Kabat plot generated by the script written for this study versus and use of Sequence Logos as described by Schnieder, it is clear that the two methods of visualization complement each other quite well. While peaks of the maxima of the Variability occur where that is the least sequence homology at a residue, the peaks of a sequence logo plot are where the residues are most conserved.

V. References

1) Bowersox, S.S. and Luther, R. (1998). Pharmacotherapeutic potential of omega-conotoxin MVIIA (SNX-111), an N-type neuronal calcium channel blocker found in the venom of Conus magus, Toxicon, 36-11, 1651-1658.
2) Conticello, S. G., Gilad, Y., Avidan, N., Ben-Asher, E., Levy, Z., Fainzilber, M. (2001). Mechanisms for Evolving Hypervariability: The Case of Conopeptides. Mol Biol Evol 18: 120-131
3) Conticello, S. G., Y. Pilpel, G. Glusman, and M. Fainzilber. (2000). Position-specific codon conservation in hypervariable gene families. Trends Genet. 16:57–59.
4) Duda, T.F. and Palumbi, S.R. (1999) Molecular genetics of ecological diversification: duplication and rapid evolution of toxin genes of the venomous gastropod Conus. Proc. Natl. Acad. Sci. U. S. A. 96, 6820–6823
5) Drakopoulou, E., J. Vizzavona, J. Neyton, V. Aniort, F. Bouet, H. Virelizier, A. Menez, and C. Vita. (1998). Consequence of the removal of evolutionary conserved disulfide bridges on the structure and function of charybdotoxin and evidence that particular cysteine spacings govern specific disulfide bond formation. Biochemistry 37:1292–1301.
6) England, L. J., Imperial, J., Jacobsen, R., Craig, A. G., Gulyas, J., Akhtar, M., Rivier, J., Julius, D., Olivera, B. M. (1998). Inactivation of a Serotonin-Gated Ion Channel by a Polypeptide Toxin from Marine Snails. *Science* 281: 575-578
7) Johnson G, Wu TT. (2000) Kabat database and its applications: 30 years after the first variability plot. Nucleic Acids Res. Jan 1;28(1):214-8.
8) Jones PT, Dear PH, Foote J, Neuberger MS, Winter G. (1986) Replacing the complementarity-determining regions in a human antibody with those from a mouse. Nature. 321-6069:522-525.
9) Kohn, A.J. (1956). Piscivorous gastropods of the genus Conus. Proc. Natl. Acad. Sci. USA 42, 168–171.
10) Norton, R.S. and Pallaghy, P.K. (1998). The cysteine knot structure of ion channel toxins and related polypeptides. Toxicon36, 1573–1583
11) Rigby, A. C., E. Lucas-Meunier, D. E. Kalume et al. (13 co-authors). (1999). A conotoxin from Conus textile with unusual posttranslational modifications reduces presynaptic Ca2+ influx. Proc. Natl. Acad. Sci. USA 96:5758–5763.
12) Schneider, T. D., and Stephens, E.M. (1990). Sequence logos: a new way to display consensus sequences. Nucleic Acids Res. 18:6097–6100.
13) Shon, K., Stocker, M., Terlau, H., Stühmer, W., Jacobsen, R., Walker, C., Grilley, M., Watkins, M., Hillyard, D.R., Gray, W.R. and Olivera, B.M., (1998). Conotoxin PVIIA: a peptide inhibiting the Shaker K+ channel. J. Biol. Chem. 273, pp. 33-38.
14) Terlau, H., Shon, K., Grilley, M., Stocker, M., Stuhmer, W., and Olivera, B.M. (1996). Strategy for rapid immobilization of prey by a fish-hunting cone snail. Nature 381, 148–151.

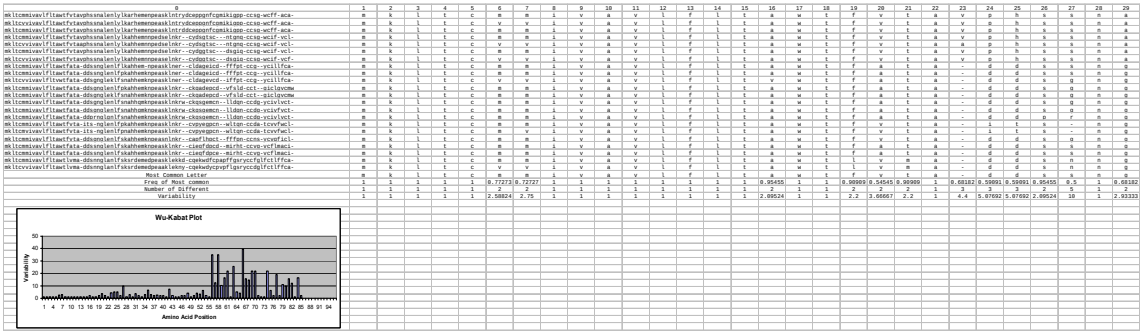
15) Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins. (1997). The CLUSTAL-X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25:4876–4882.
16) Wu,T.T. and Kabat,E.A. (1970) An analysis of the sequences of the variable regions of Bence Jones proteins and myeloma light chains and their implications for antibody complementarity. *J. Exp. Med.*, 132, 211–250

VI. Appendix – variability plots generated

Script template

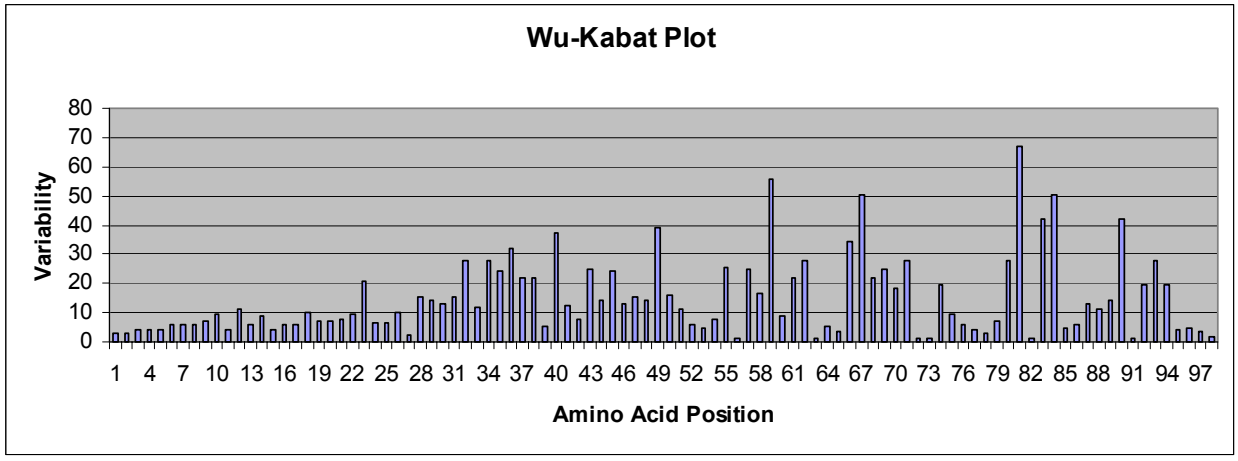
C:\Documents and Settings\Jamie Schroe

Protein_ Variability.xlt

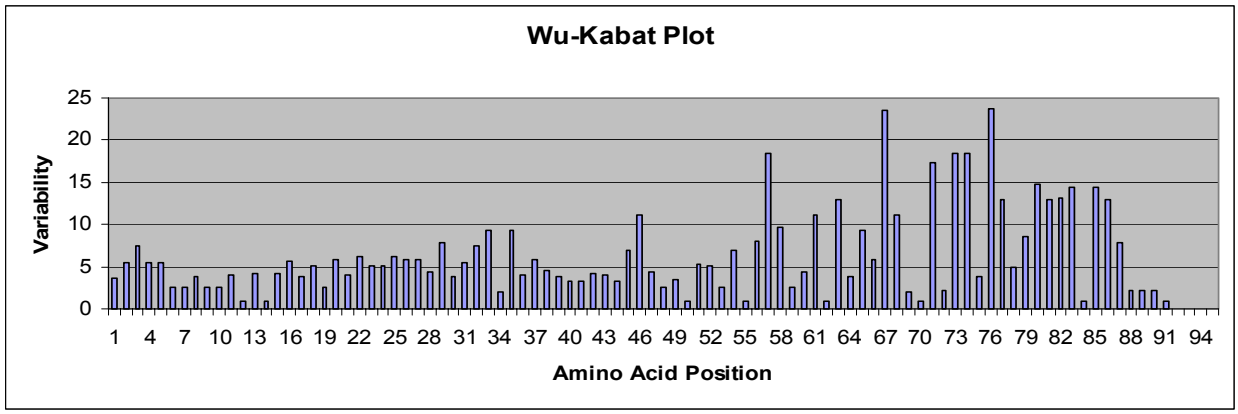


- **Species:**

Conus arenatus CA

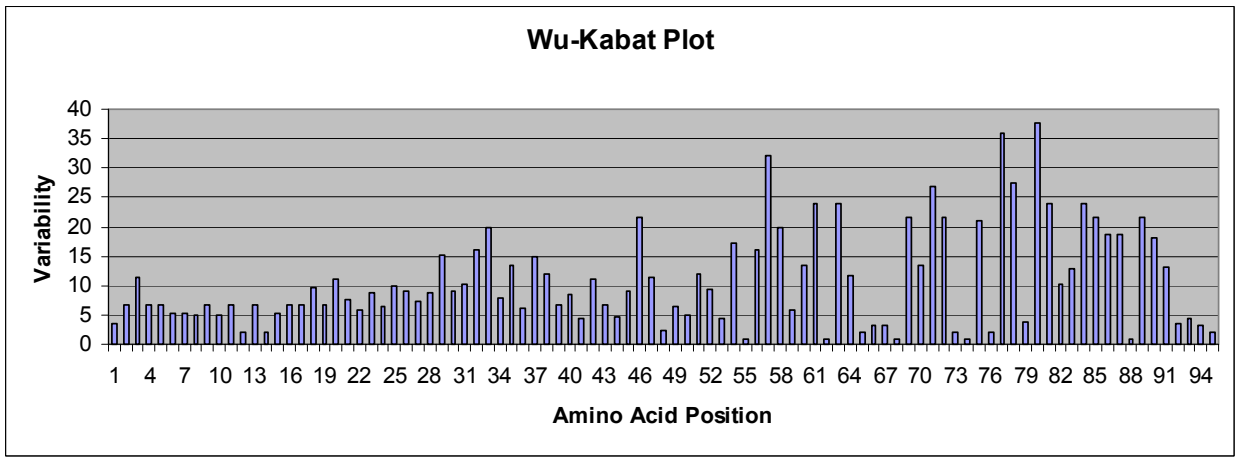



```
---mkltcvliiavlflivcqlntad-dsrdkqeyravrlrdairnsrgs---rscgnlgesc--sah--rccpg----lmcmg--easicipy----
---mkltcvliiamlflivcqlntad-dstdkqeyravklrdamrnfkgsk--rncgeqgegc--atr--pccag----lscvgsrpgglcqyd----
---mkltcvliiavlcltvcqlitad-ylrdkqkyrsvrlrdgmlnfkgs---rqcadlgeec--ytr--fccpg----lrckd-lqvptclla----
---mkltcvliiavlfltacqlitad-ysrdkqeyravrlrdamrysrvr---rqcadlgeec--htr--fccpg----lrced-lqvptclma----
---mkltcvlivavlxltacqliaad-dsrdlkrfsrrkmrdgmlntknmk--rqclpplslc--tmdddeccd------dcxl----flclvts---
---mkltcvlivavlfltacqliaad-dsrdlkrfsrrkmrdgmlntkntk--rqclpplslc--tmdddeccd------dcil----flclvts---
---mkltcvlivavlfltacqliaad-dsrdlkrfsrrkmrdgmlntknte--egclpplslc--tmaddecch------dcil----flclvsp---
---mkltcvlivavlfltacqliaad-dsrdlkrfsrrnmrdgmlntkntk--rqclpplslc--nmadddccn------dcvl----flcsyy----
---mkltcvlivavlfltacqliaad-dsrdlqkfprrkmrdgmlntkntk--rqclpplhwc--nmvddecch------fcvl----lacv------
---mkltcvliiavlfltacqlttgetysrgeqkdhalrstd--knsklt---rqcspnggsc--srhy-hccs-----lwcnk--dsgvcvatsyp-
---mkltcvliiavlfltacqltt------geqkdhaqrsad--rnsklt---rqctpvggsc--srhy-hccs-----lycnk--nigqclatsyp-
---mkltcvliiavlfltacqltt------geqkdhalrstd--knsklt---rqctpvggyc--srhh-hccs-----nhcik--sigrcvah----
---mkltcvliiavlfltacqlttgetysrgeqkdhalrstd--knsklt---rqctpvggyc--fdhh-hccs-----nhcik--sigrcvah----
---mkltcvliiavlfltacqltt------geqkdhalrstd--knsklt---rqctpvggyc--srhy-hccs-----nhcik--sigrcvah----
msgmkltcviivvalfltach-------akdkqehpavrgsddmqdsedlklakkctvdsdfc--dpdnhdccs-----grcideggsgvcaivpvln
---mkltcvlivallfltacqlttad-dsrdkqedplvrshrkmqksedpkmaercsnfgsdc--ipathdccs-----gecfgfedmglct------
---mkltcmmivavlfltawtsvtav-ntrgelenlflrashemnseaskldkkvcvdggtfcgfpkiggpccs-----gwcif-----vcl------
---mkltcmmivavlfltawtsvtav-ntrgelenlflrashemnseaskldkkvcvdggtfcgfpkiggpccs-----gwcif-----vcl------
---mkvtcmmivavlfltawtfvtad-dsisaledlfakadkmenseasplnerdcralgeycglpyvhnsccs-----qlcgf-----icvpesp--
msglgimllllllvfmetshqdage-kqatqrdainvrrrrsltrr-vte---eceen---ce--eeekhccntnngp-scap-----qcfg-----
msglgilvltllllvymatshqdage-kqatqrdainvrrrrsltrr-vae---ecees---ce--deekhccntnngp-scap-----qcfg-----
msglgimvltllllvfmetshqdage-kqamqrdainvrrrrsitrr-vse---acees---ce--eeekhcchenngvytclr-----ycwg-----
msrlgimvltllllvfivtshqdage-kqatqraainfrwkrsltrrtate---eceey---ce--deektccgeedgepvcar-----fclg-----
msrlgimvltllllvymatshqdage-kqatqrdainfrwkrsltrrtate---ecees---ce--eeektccgexdgepvcar-----fclg-----
msrlgimvltllllvfivtshqdage-kqathrgainfrwrrslirrtate---eceey---ce--deektccgledgepvcat-----tclg-----
msrlgimvltllllvfivtshqdage-kqatqrnainfrwrrsftrraate---eceey---ce--deektccgledgepvcat-----tclg-----
msrlgimvltllllvfivtshqdage-kqatkraavnfrwrrsftrraaae---eceey---ce--eeektccgeedgepvcae-----fclg-----
-----------------------------------------rrsltrr-vpe---ecees---ce--eeektccglengqpfcsr-----icwg-----
```




```
mkltcvliiavlflivcqlntad-dsrdkqeyravrlrdairnsrgs-rscgnlgescsah--rccpglmcmg--easicipy----
mkltcvliiamlflivcqlntad-dstdkqeyravklrdamrnfkgskrncgeqgegcatr--pccaglscvgsrpgglcqyd----
mkltcvliiavlcltvcqlitad-ylrdkqkyrsvrlrdgmlnfkgs-rqcadlgeecytr--fccpglrckd-lqvptclla----
mkltcvliiavlfltacqlitad-ysrdkqeyravrlrdamrysrvr-rqcadlgeechtr--fccpglrced-lqvptclma----
mkltcvlivavlxltacqliaad-dsrdlkrfsrrkmrdgmlntknmkrqclpplslctmdddecc--ddcxl----flclvts---
mkltcvlivavlfltacqliaad-dsrdlkrfsrrkmrdgmlntkntkrqclpplslctmdddecc--ddcil----flclvts---
mkltcvlivavlfltacqliaad-dsrdlkrfsrrkmrdgmlntknteegclpplslctmaddecc--hdcil----flclvsp---
mkltcvlivavlfltacqliaad-dsrdlkrfsrrnmrdgmlntkntkrqclpplslcnmadddcc--ndcvl----flcsyy----
mkltcvlivavlfltacqliaad-dsrdlqkfprrkmrdgmlntkntkrqclpplhwcnmvddecc--hfcvl----lacv------
mkltcvliiavlfltacqlttgetysrgeqkdhalrstd--knsklt-rqcspnggscsrhy-hccs-lwcnk--dsgvcvatsyp-
mkltcvliiavlfltacqltt------geqkdhaqrsad--rnsklt-rqctpvggscsrhy-hccs-lycnk--nigqclatsyp-
mkltcvliiavlfltacqltt------geqkdhalrstd--knsklt-rqctpvggycsrhh-hccs-nhcik--sigrcvah----
mkltcvliiavlfltacqlttgetysrgeqkdhalrstd--knsklt-rqctpvggycfdhh-hccs-nhcik--sigrcvah----
mkltcvliiavlfltacqltt------geqkdhalrstd--knsklt-rqctpvggycsrhy-hccs-nhcik--sigrcvah----
```

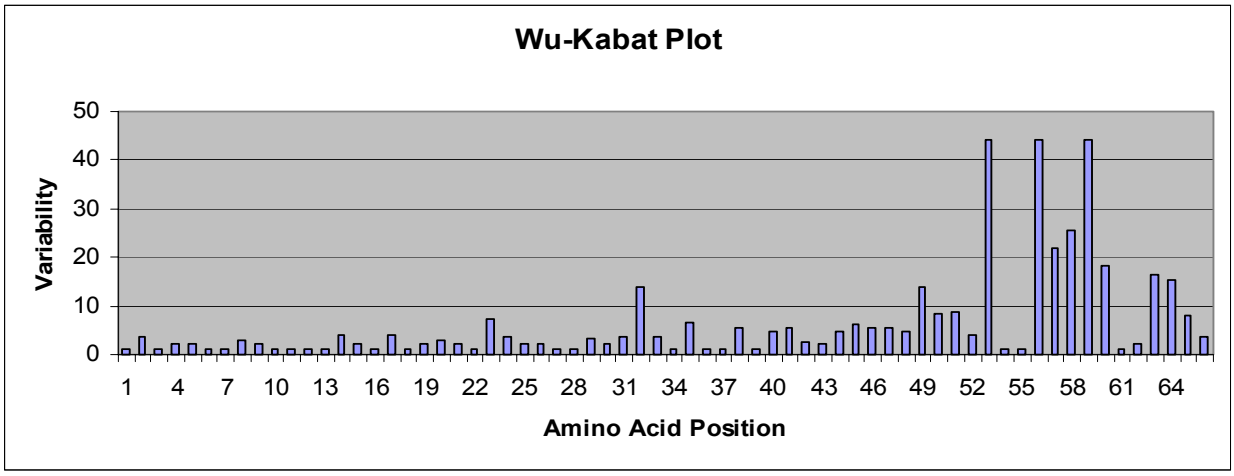
Conus ebraeus CE




```
msglgimlltllllvfmetshqdagekqatqrdainvrr--rrsltrr--vteeceen---ceeeekhccn-tnngp-scapqcfg-----
msglgilvltllllvymatshqdagekqatqrdainvrr--rrsltrr--vaeecees---cedeekhccn-tnngp-scapqcfg-----
msglgimvltllllvfmetshqdagekqamqrdainvrr--rrsitrr--vseacees---ceeeekhcch-enngvytclrycwg-----
msrlgimvltllllvfivtshqdagekqathrgainfrw--rrslirrt-ateeceey---cedeektccg-ledgepvcattclg-----
msrlgimvltllllvfivtshqdagekqatqrnainfrw--rrsftrra-ateeceey---cedeektccg-ledgepvcattclg-----
msrlgimvltllllvfivtshqdagekqatqraainfrw--krsltrrt-ateeceey---cedeektccg-eedgepvcarfclg-----
msrlgimvltllllvfivtshqdagekqatkraavnfrw--rrsftrra-aaeeceey---ceeeektccg-eedgepvcaefclg-----
msrlgimvltllllvymatshqdagekqatqrdainfrw--krsltrrt-ateecees---ceeeektccg-exdgepvcarfclg-----
mkltcvliiavlflivcqlntad-dsrdkqeyravrlrdairnsrgs-----rscgnlgescs--ahrccpglmcmg--easicipy----
mkltcvliiamlflivcqlntad-dstdkqeyravklrdamrnfkgs---k-rncgeqgegca--trpccaglscvgsrpgglcqyd----
mkltcvliiavlcltvcqlitad-ylrdkqkyrsvrlrdgmlnfkgs-----rqcadlgeecy--trfccpglrckd-lqvptclla----
mkltcvliiavlfltacqlitad-ysrdkqeyravrlrdamrysrvr-----rqcadlgeech--trfccpglrced-lqvptclma----
mkltcvliiavlfltacqlttgetysrgeqkdhalrstd--knsklt-----rqcspnggscs-rhyhccs-lwcnk--dsgvcvatsyp-
mkltcvliiavlfltacqltt------geqkdhaqrsad--rnsklt-----rqctpvggscs-rhyhccs-lycnk--nigqclatsyp-
mkltcvliiavlfltacqltt------geqkdhalrstd--knsklt-----rqctpvggycs-rhyhccs-nhcik--sigrcvah----
mkltcvliiavlfltacqltt------geqkdhalrstd--knsklt-----rqctpvggycs-rhhhccs-nhcik--sigrcvah----
mkltcvliiavlfltacqlttgetysrgeqkdhalrstd--knsklt-----rqctpvggycf-dhhhccs-nhcik--sigrcvah----
-----vliiavlfltacqlttaetysrgrqkhrarrstd--knskwt-----rectrsggacn-shtqccd-dfcst--atstci------
-----vliiavlfltacqlttaetysrgrqkhrarrstd--knskwt-----rectrsggacn-shtqccd-dfcst--atstci------
-----vliiavlfltacqltiaetysrgrqkhrarrstd--knskwt-----rectrsggacn-shtqccd-dfcst--atstci------
-----vliiavlfltacqlttaetysrgkqkhrarrstd--knskwt-----rectrsggacn-shtqccd-dfcst--atstci------
-----vliiavlfltacqlttaetysrgrqkhrarrstd--knskwt-----rectrsggacn-shtqccd-dfcst--atstct------
-----vliiavlfltacqlttaetysrgrqkhrarrstd--knskwt-----recthsggacn-shtqccd-dfcst--atstci------
-----vliiavlfltacqlttaetysrgrqkhrarrstd--knskwt-----rectrsggacn-shtqccd-dfcdt--atrtcv------
-----vliiavlfltacqlttaetysrgrqkhrarrstd--knskwt-----recthsggacn-shdqccn-afcdt--atrtcv------
-----vliiavlfltacqlttaetysrgrqkhrarrstd--knskwt-----recthsggacn-shdqccn-afcdt--atrtcv------
-----vliiavlfltacqlttaetysrgrqkhrarrstd--knskwt-----recthsggacn-shdqccn-afcdt--atrtcv------
-----vliiavlfltacqlttaetysrgrqkhrarrstd--knskwt-----recthsggacn-shdqccn-afcdt--atrtcv------
```

```
-----vliiavlfltacqlttaetysrgrqkhrarrstd--knskwt-----recthsggacn-shdqrcn-afcdt--atrtcv------
-----vliiavllltacqlttaetysrgrqkhrarrstd--knskwt-----recthsggacn-shdqccn-afcdt--atrtcv------
-----vliiavlfltacqlttaetysrgrqkhrarrstd--knskwt-----recahsggacn-shdqccn-afcdt--atrtcv------
-----vliiavlfltacqlttaetysrgrqkhrarrstd--knpkwt-----recthsggacn-shdqccn-afcdt--atracv------
-----vliiavlfltacqlttaetysrgrqkhrarrstd--knskwt-----recthsggacn-shdqccn-tfcdt--atrtcv------
-----vliiavlfltacqlttaetysrgrqkhrarrstd--knskwt-----recthsggacn-shnqccn-afcdt--atrtcv------
-----vliiavlfltacqlttaetysrgrqkhrarrstd--knskwt-----rectrsggacn-shdqccn-afcdt--atrtcv------
-----vliiavlfltacqlttaetysrgrqkhrarrstd--knskwt-----rectrsggacn-shtqccn-afcdt--atrtcv------
```

Conus pennaceus CP

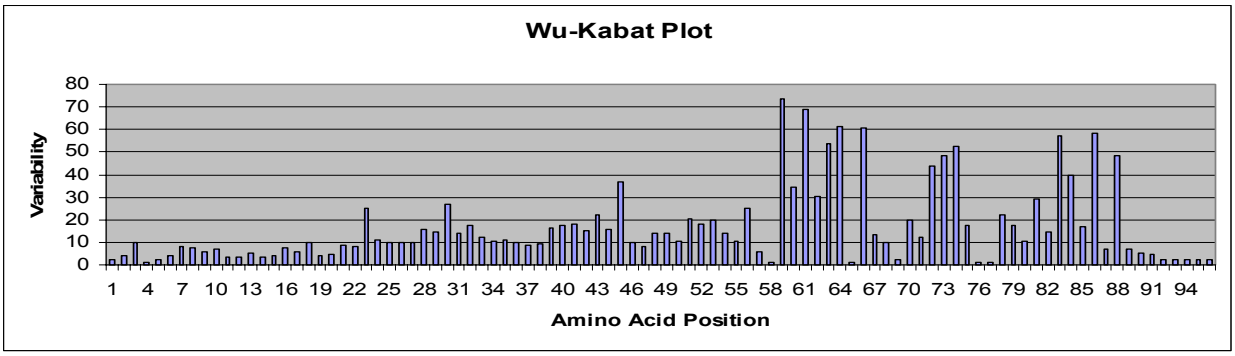



```
mrclpvfvillllliasapsvdarp--ktkddiplvsfqdnakralqilsnkryccyfdyscclylr
mcclpvfvillllliasapsvdalp--ktkddmslasfhdnakrtlqilsnkryccvydyscclswg
mrclpvfvillllliastpsvdalq--ktkddmplasfhdnvkrilqtlsnkrsccptilsccfv--
mrclpvfvillllliasapsvdarp--ktkddiplvsfqdhakrilqtfesrydcckt-fecchwg-
mrclpvfvillllliasapsvdarp--ktkddiplvsfqdhakrilqtfesrydcckt-fecchwg-
mrclpvfvillllliastpsvnarp--ktkd---lasfhdnakrtqhifwskrncciyenwccewi-
mrclpvfvillllltasgpsvdakvhlktkgdgplssfrdnakstlqrlqdkstccgyrm-cvpcg-
mrclpvfivllllivsapgfdarp--kteddvplssfhddlqrtvrtlldirmcclgtsgccpwg-
mhclsvfvillllltasapsvdaqp--kteddvplssfhddlqrtvrtlldirmcclgtsgccpwg-
mrcfpvfiilllllmasapsfdarp--kteddvplssfrdnlkrtlrtlldprrccyetpgccvig-
mrclpvfiilllliasapsfdalp--ktednvplssfhdnlkrtrrihlnireccsdg-wccpag-
```

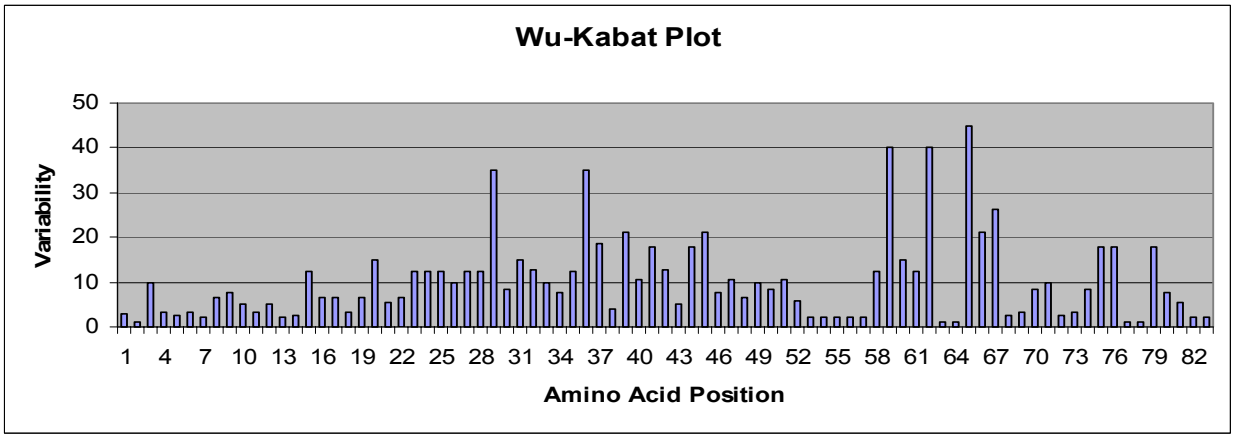
Conus tessulatus CT – See Plot Above

```
------villms-----tqaliqsgv-ekrs-nkikalskrkttaesww---egecy-gwwtscsspeqccsl-nce-niycraw----
-mekltilllaavlvlaqalikkgggekrqkekinflskrkttaesww---egecs-gwsvyctsdpeccsg-ecs-syycelw----
-mekltilllvaavlmstqaliqrgg-akr—rkvnffsirepgaedwr---egnct-pwlggctspeeccpg-nce—tycrawr---
-msglgimvltlllfmfmatshqdagekqatqrdainvrrrrsit-rrg---deecn-eycd—drnkeccgr-tng-hprcanvcfg-
-msglgimvltlllfmfmatshqdagekqatqrdainvrrrrsit-rrg---deecn-ehce—drnkeccgr-tng-hprcanvcfg-
-msglgimvltllllvfmatshqdagekqatqrdainvrrrrsit-rrv---deecn-eycd—drnkeccgr-tng-hprcanvcfg-
-msglgimvltllllvfmatshqdagekqatqrdavnvrrrrsiagrtt---teecd-eyce—dlnkneccgl-sng-epvcataclg-
-mlkmgvvlfvflvlfplatlqldad------qpveryaenkqlvspy----errqi—ilhalgqrq-ccdwqwcdgacdcca-----
-mlkmgvvlfvflvlfplatlqldad------qpveryaenkqlvspy----errqi—ilhalgqrq-ccdwqwcdgacdcca-----
-mlkmgvvlfvflvlfplatlqldad------qpveryaenkqlvspy----errqi—ilhalgqrd-ccvmpwcdgacdccvss---
-mlkmgvvlftflvlfplatlqldad------qpveryaenkqglnpd----esrei—ilsalrqrd-cceqgwcdggcdccq-----
mmsklgvllticlllfpltavpldgd------qhadrpadrmqdiss-----eqhp---lfdpvkr---ccdwpctigcvpcclp----
mmsklgvllticlllfpltavrldgd------qhtdrpadrmqdiat-----eqhp---lfdpvkr---ccdwpctigcvpcclp----
mmsklgvllticlllfpltavsldgd------qpadlpelraqdfap-----ersp---wfdpvrr---ccsqdcrv-cipccpy----
mmsklgvllticlllfpltavqldgd------qpadlpelraqdfap-----ersp---wfdpvrr---ccsqdcrv-cipccpy----
mmsklgvllticlllfpltavqldgd------qpadlpelraqdfap-----ersp---wfdpvrr---ccsqdcrv-cipccpn----
mmsklgvllticlllfpltavqldgd------qpadlpalrtqdiat-----dhsp---wfdpvkr---ccsrycyi-cipccpn----
mmsklgvllticlllfpltavqldgd------qpadlpalrtqdist-----dhsp---wfdpvkr---ccsrycyi-cipccpn----
mmsklgvllticlllfpltvlpmdgd------qpadlpalrtqdiat-----dqsp---wfdpvkr---ccsrycwk-cipccpy----
-------------------pnplerriq----sdliraaledadmktdersvegvi---stikdfavkvccsvslkf----ccpta---
-mlclpvfiilllllaspaapnpletriq----rdliraaledadmktnerflegvi---stikdfagkvccsvsvnf----ccpta---
-mlclpvfiilllllaspaapnplerriq----sdlirtaledadmktpkgvlsgimsnlgtvgnmvggfcctvysg-----ccsek---
-mhclpvlvillllliastpsvdarpktk----ddvplasfhgadnanr—ilrtlw---------nlrgccedkt------ccfig---
-mhclpvlvillllliastpsvdarpktk----ddvplasfhgadnanr—ilrtlw---------nlrgccedkt------ccfig---
-mhclpvpvillllliastpsvdarpktk----ddvppasfhgadnanr—ilrtlw---------nlrgccedkt------ccfig---
-mhclpvlvillllliastpsvdarpktk----ddvppasfhgaddanr—ilqtlw---------nlrgccedkt------ccfig---
-mhclpvlvillllliastpsvdarpnpk----ddvplasfhgavnakr—ylrtlw---------nsrdccdpkep-----ccfig---
mmrclpvfiilllllipsaasa-aqpetk----ddaalasfy—dnakr—tlqrhw---------akslccpedaw-----ccshdegk
-mrclpvfiilllllipsaasv-aqpktk----ddvalasfy—dnakr—tlqrhw---------akslccpedaw-----cc------
```

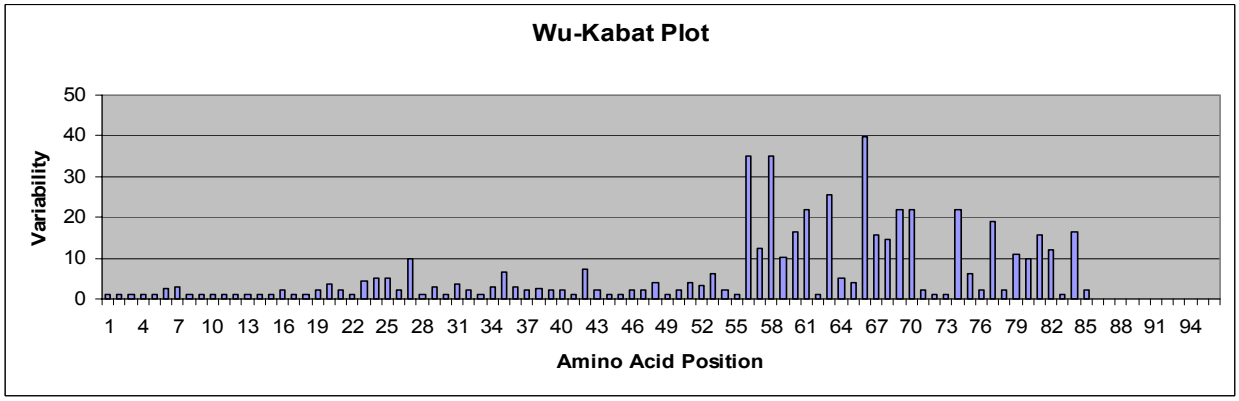
Conus textile CX



Conus ventricosus CV



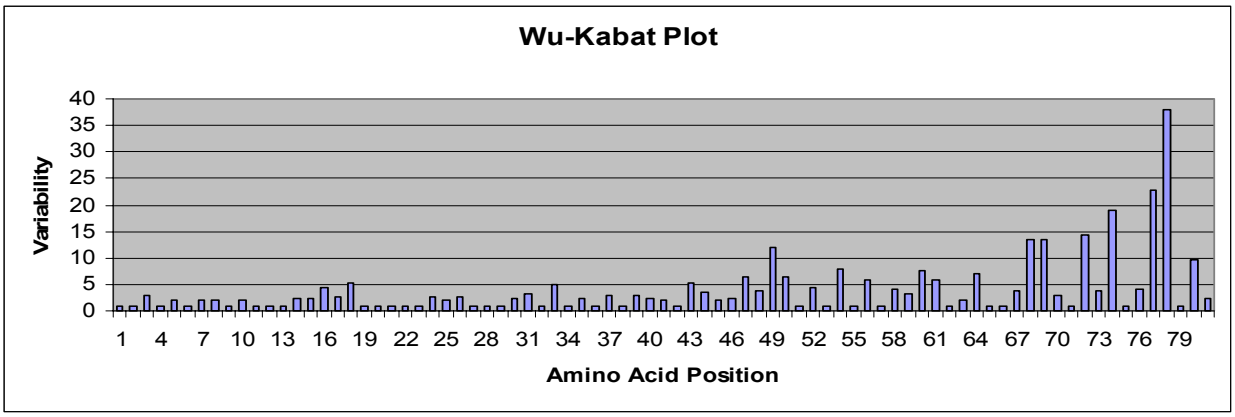

```
-mrclpvfvillllliasapgvdaqpktkydvplasrhdfakktpkr-ls--------kpr-dccrrn-------flcc-----
-mrclpvfvillllliasapgvdaqpktkydvplasrhdfakktpkr-ls--------kpr-dcclrh-------flccv----
-mrclpvfvillllliasapgvd---------------dfakktpkr-ls--------kpr-dccrrn-------flcc-----
-mhclpvfvillllliasapgvdvqpktknfmtraslrdfakktpkr-ls--------klr-gccprs-------flccr----
-mrclpvfvillllliasapgvdvqpktkyyvprasrrdfakktpkr-ls--------klr-gccprs-------flccr----
-mrclpvfiillvlsasapsvdarpkneddvslasfhdnaqrtlqr-ll--------nkr-sccpnn-------pacch----
-mrclpvfvillllliasapsvdarpktkddvplasfhgnaertlln-il--------rdgdnccidk-------qgccpwg--
mmlclpvfiilllllaspaapnplekriqsdliraaledadmktgereilniidsisdvakqiccqit-------vdccvldee
mmlclpvfiilllllaspaapnplekriqsdliraaledadmktdereivniidsisdvakqicceit-------vqccvldee
-mlkmgvmlfiflvlfplatlqldadqpveryaenkrlmspderr-ailhap------rqrgccepdwcdsgcddgcc-----
-mlkmgvvlfiflvlftlatlqldadqpveryaenkrlmspykrr-ailhap------reqeccepqwcdggcda-cc-----
-mlkmgvvlfiflvlftlatlqldadqpveryaenkqllspderrdiilhal------gqrrccdpdwcdagcydgcc-----
mmsklgvvlticllpfpltalpmdgdqpadlpalrtqd-fepers-pwfdp--------vkrccsqd-crv-cip-ccpy---
mmsklgvllticllllfpltavpmdgdqpadlpalrtqd-fepers-pwfdp--------vrrccsqd-csv-cip-ccppp--
mmsklgallticllllfpltalpldgdqpadrpaermqddisseqh-plfdk--------erqcctgs-cln-cwp-ccg----
```




```
-mlkmgvmlfiflvlfplatlqldadqpveryaenkrlmspderr-ailhaprqrgccepdwcdsgcddgcc---
-mlkmgvvlfiflvlftlatlqldadqpveryaenkrlmspykrr-ailhapreqeccepqwcdggcda-cc---
-mlkmgvvlfiflvlftlatlqldadqpveryaenkqllspderrdiilhalgqrrccdpdwcdagcydgcc---
mmsklgvvlticllpfpltalpmdgdqpadlpalrtqd-fepers-pwfdp--vkrccsqd-crv-cip-ccpy-
mmsklgvllticllllfpltavpmdgdqpadlpalrtqd-fepers-pwfdp--vrrccsqd-csv-cip-ccppp
mmsklgallticllllfpltalpldgdqpadrpaermqddisseqh-plfdk--erqcctgs-cln-cwp-ccg--
```

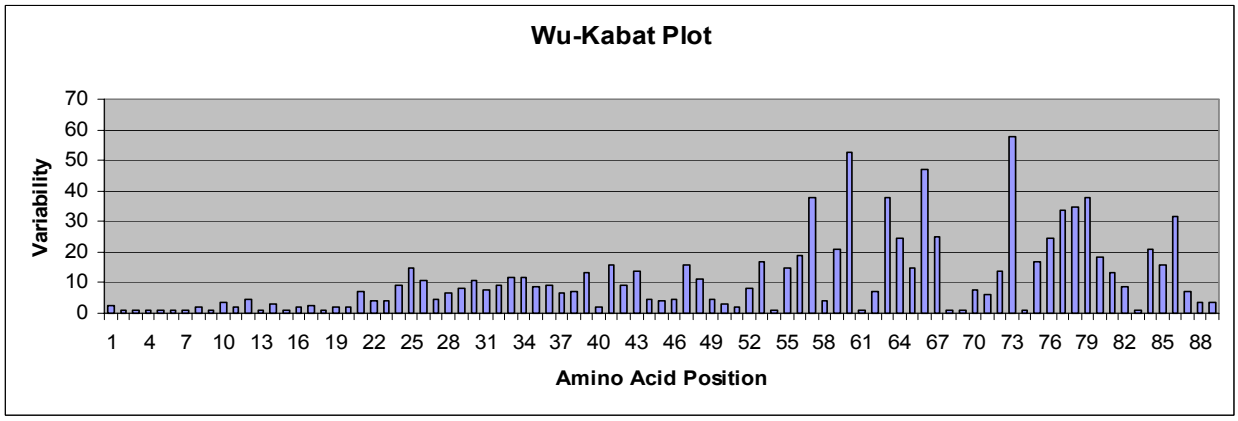
- **Popset Families**

Multiple Conus #1




```
msrlgimvltllllvfivtshqdagekqatqrdainfrwrrslirrt-ateeceeycedeek-tccgledgepvcattclg
msrlgimvltllllvfivtshqdagekqathrgainfrwrrslirrt-ateeceeycedeek-tccgledgepvcattclg
msrlgimvltllllvfivtshqdagekqatqrnainfrwrrsftrra-ateeceeycedeek-tccgledgepvcattclg
msrlgimvltllllvfivtshqdagekqatkraavnfrwrrsftrra-aaeeceeyceeeek-tccgeedgepvcaefclg
msrlgimvltllllvfivtshqdagekqatqraainfrwkrsltrrt-ateeceeycedeek-tccgeedgepvcarfclg
msrlgimvltllllvymatshqdagekqatqrdainfrwkrsltrrt-ateeceesceeeek-tccgexdgepvcarfclg
msglgimvltllllvfmetshqdagekqamqrdainvrrrrsitrr--vseaceescedeek-hcchenngvytclrycwg
msglgimvltllllvfmetshqdagekqamqrdainvrrrrsitrr--vseaceesceeeek-hcchenngvytclrycwg
msglgimlltllllvfmetshqdagekqatqrdainvrrrrsltrr--vteeceenceeeek-hccntnngp-scapqcfg
msglgilvltllllvymatshqdagekqatqrdainvrrrrsltrr--vaeeceescedeek-hccntnngp-scapqcfg
msgleimvltllllvsmatshqdggekqatqrdainvrrr-sitr----teacyeyckeqnk-tccgisngrpicvggcir
msglgimvltllllvsmatshqdgrgkqatqrdainvrrrrsitr----teacyeyckeqnk-tccgisngrpicvggcir
msglgimvltllllvsmatshqdgggkqatqrdainvrrrrsitrrevvteeceeyckeqnk-tccgltngrprcvgvcfg
msglgimvlallllvfmatshqdgggkqatqrdainvrrrrsitrr-vvtetckeycedrdk-tccglengqpdcanlclg
msglgimvltllllvsmatshqdgggkqatqrdainvrrrrsitrr-vvteaceeycedrdkktccglengepfcatlcfg
msglgimvltllllfmfmatshqdagekqatqrdainvrrrrsitrr--gdeecneycddrnk-eccgrtnghprcanvcfg
msglgimvltllllfmfmatshqdagekqatqrdainvrrrrsitrr--gdeecnehcedrnk-eccgrtnghprcanvcfg
msglgimvltllllvfmatshqdagekqatqrdainvrrrrsitrr--vdeecneycddrnk-eccgrtnghprcanvcfg
msglgimvltllllvfmatshqdagekqatqrdavnvrrrrsiagr-ttteecdeycedlnk-nccglsngepvcataclg
```

Multiple Conus #2

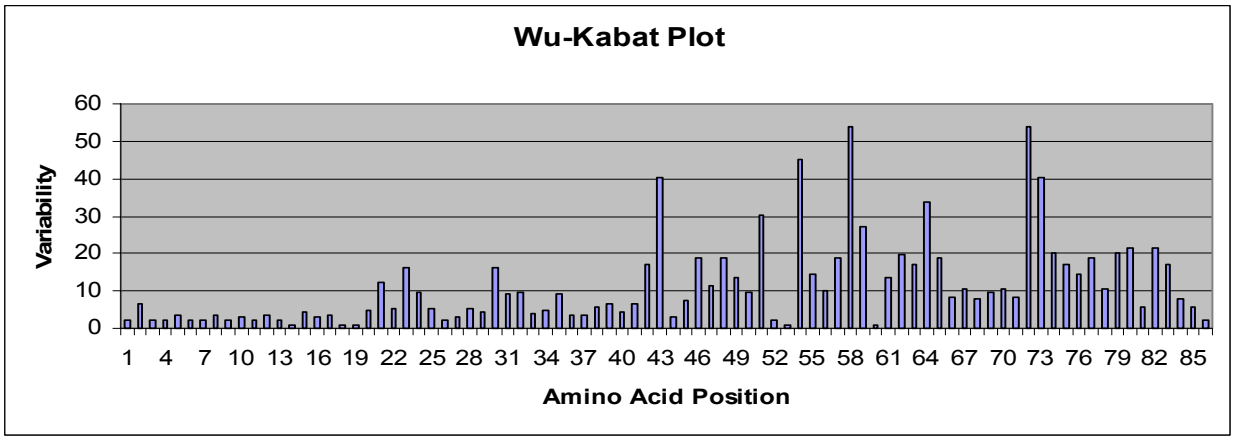



```
-mkltcviivvalfltach-------akdkqehpavrgsddmqdsedlklakkctvdsdfcdpdnhdcc-sgrcideggsgvcaivpvln
-mkltcvlivallfltacqlttadd-srdkqedplvrshrkmqksedpkmaercsnfgsdcipathdcc-sgecfgfedmglct------
-mkltcvliiavlfltacqlttgetysrgeqkdhalrstd--knsklt---rqcspnggsc-srhyhcc-slwc-nkd-sgvcvatsyp-
-mkltcvliiavlfltacqltt------geqkdhaqrsad--rnsklt---rqctpvggsc-srhyhcc-slyc-knnigq-clatsyp-
-mkltcvliiavlfltacqlttgetysrgeqkdhalrstd--knsklt---rqctpvggyc-fdhhhcc-snhc-iksigr-cvah----
-mkltcvliiavlfltacqltt------geqkdhalrstd--knsklt---rqctpvggyc-srhhhcc-snhc-iksigr-cvah----
-mkltcvliiavlfltacqltt------geqkdhalrstd--knsklt---rqctpvggyc-srhyhcc-snhc-iksigr-cvah----
-mkltcvlivavlfltacqlttaasyarserqhpdlgssd--qnsklt---krclasgetc-wrdtscc-sfsctnn--v--cf------
-mkltcvlivavlfltacqlttaasyarserehpdlgssd--qnsklt---krclasgetc-wrdtscc-sfsctnn--v--cf------
-mkltcvlivavlfltacqlttaasyarserqhpdlgssd--qnsklt---krclgsgetc-wldsscc-sfsctnn--v--cf------
-mkltcvliiavlflivcqlntadds-rdkqeyravrlrdairnsrgs---rscgnlgesc-sahr-ccpglmcmgeasi--cipy----
-mkltcvliiamlflivcqlntadds-tdkqeyravklrdamrnfkgsk--rncgeqgegc-atrp-ccaglscvgsrpgglcqyd----
-mkltcvliiavlcltvcqlitadyl-rdkqkyrsvrlrdgmlnfkgs---rqcadlgeec-ytrf-ccpglrc-kdlqvptcl-la---
-mkltcvliiavlfltacqlitadys-rdkqeyravrlrdamrysrvr---rqcadlgeec-htrf-ccpglrced-lqvptcl-ma---
mmkltcvliiavlfltacqlttaet----rdeyravrssdevrnsrst---rdcsgsgygc--kntpccdgltcrgphqgpicl------
mmkltcvliiavlfltacqlttaet----rdeyravrssdevrnsrs------csgsgygc--kntpccagltcrgprqgpicl------
mmkltcvliiavlfltacqlttaet----rdeyravrssdevqnsrst---ddcstag--c--knvpcceglvctgpsqgpvcqpla---
-mkltcvlivavlxltacqliaadds-rdlkrfsrrkmrdgmlntknmk--rqclpplslctmdddecc--ddc-x-lfl--clvts---
-mkltcvlivavlfltacqliaadds-rdlkrfsrrkmrdgmlntkntk--rqclpplslctmdddecc--ddc-i-lfl--clvts---
-mkltcvlivavlfltacqliaadds-rdlkrfsrrkmrdgmlntknte--egclpplslctmaddecc--hdc-i-lfl--clvsp---
-mkltcvlivavlfltacqliaadds-rdlkrfsrrnmrdgmlntkntk--rqclpplslcnmadddcc--ndc-v-lfl--csyy----
-mkltcvlivavlfltacqliaadds-rdlqkfprrkmrdgmlntkntk--rqclpplhwcnmvddecc--hfc-v-lla--cv------
```

Multiple Conus #3 – see above for all three plot (incl. protein additions)

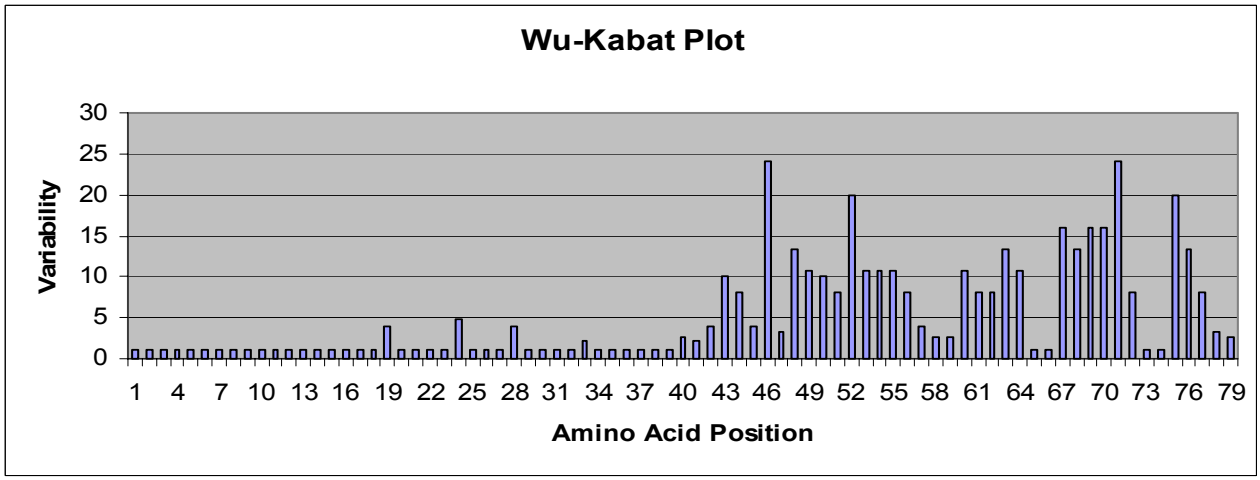
```
mkltcmmivavlfltawtfvtadd------srngleylfpkahyem-npeasklnkkedceaggrfcg-fpkigepccsgwcffvca----
mkltcmmivavlfltawtfvtad-------srngleylfpkahyem-npeasklnkkrdcvagghfcg-fpkiggpccsgwcffvca----
mkltcmkivavlfltawtfvtadd------srngleylfpkahyem-npeasklnkkqdcaaggqfcg-fpkiggpccsgwclgvca----
mkltcmmivavlfltawtfvtadd------srngldylfpkarhem-npkasrdikr--crpggmicg-fpkpgpyccsgwcfvvcl----
mkltcmmivavlfltawtsvtavn------trgelenlflrashem-nseaskldkk-vcvdggtfcg-fpkiggpccsgwcifvcl----
mkvtcmmivavlfltawtfvtadd------sisaledlfakahdkmenseasplner-dcralgeycglpyvhnsrccsqlcgficvpesp
mkltcvmivavlfltawtfvtaddprdgpdtavgwrklfseardemknreasklner-gciedkkycgipfansgvccsylcifvcvpkap
mklmcmmivavlfltawtfvtaddsingpenrriwekllsktrdemknpeasklnkk-ecrqpgefc-fpvvak--ccggtclvici----
```

Multiple Conus #4




```
mekltilllvaavlmstqamfq----eksrkaeinfsetrkl-arnkqkr--cggystyce--vdseccsdncvrs---yctlfg------
mekltilllvaavlmstqamfqgdg-eksrkaeinfsetrkl-arnkqkr--cggystyce--vdseccsdncvrs---yctlfg------
mekltilllvaavlmstqaliqhdg-eksqkakmkfltartlsaktrgvd--cvglssycgpwnnppccswytcdy---yckf--------
-mkltilflvaavlmstqaliqhdg-eksqkakmkfltartlsaktrgvd--cvglssycgpwnnppccswytcdy---yckf--------
-mkltilflvaavlmstqaliqhdg-eksqkakmkfltartlsakkrdvd--cvgwssycgpwnnppccswytcdy---yckl--------
mqkltilllvaavlmstqalirggv-ekrqeakrnffskrkttaeswwege-crtwyapcn--fpsqccs-evcssktgrcltw-------
mekltilllvaavltstqaliqggg-derqkakinflsrs--------drd-crgydapcs--sgapccdwwtcsartnrcf---------
mekltilllvaavlmstqalverag-enhskeninfllkrkraadrgmwge-ckdglttcl--apseccs-edceg---sctmw-------
mekltilllvaavlmstqalverag-enrskenikfllkrkraadrgmwgk-ckdglttcl--apseccs-gnceq---nckmw-------
mekltilllvaavlmstqalpqggg-enrlkenikfllkrktaadrgmwgd-cddwlaact--tpsqcct-evcdg---fcrlwe------
mekltilllvavvlmstqalpqggg-ekrprenirflskrksnaerwregs-ctswlatct--dasqcct-gvcykr-aycalwe------
mekltilllvavvlmstqalpqggg-ekrprenirflskrksnaerwregs-ctswlatct--dasqcct-gvcykr-aycalwe------
mqkltilllvaavlmstqalikggg-ekrpkekikflskrrtnaerwwegd-ctgwldgct--spaecct-avcda--t-cklw-------
mqkltilllvaavlmstqalikggg-ekrpkekirflskrkttaerwwege-crgwsngct--tnsdccs-nncdg--tfcklw-------
mekltilllvaavlmwtqaliq----ekrpkekikflskrkttaeswwege-csgwsvyctqhseccsgectgnycelf-------
mqkltilllvaavlmstqalikggg-ekrpkekikflskrkttaeswwege-csgwsvyctqhseccsgectgnycelf-------
mekltilllvaavlmstqaliq----ekrpkekikflskrksipeswwege-csgwsvhctqhsdccsgectgsycely-------
mekltilllaavlvlaqalikkgggekrqkekinflskrkttaeswwege-csgwsvyctsdpeccsgecssyycelw-------
----------villmstqaliqsgv-ekrs-nkikalskrkttaeswwege-cygwwtscsspeqccslnceniycraw-------
mekltilllvaavlmstqaliqrgg-akr--rkvnffsirepgaedwregn-ctpwlggctspeeccpgnce-tycrawr------
mekltilllvaavllsiqalnq----ekhqrakinllskrkppaerwwrwggcmawfglcskdseccsnscdvtrcelmpfppdw-
mekltilllvaavlmstqalnq----eqhqrakinllskrkppaerwwe---cgiwfsrctkdseccsnscdqtycelmpfppdw-
mekltilllvaavlmstqaqnq----eqrqqakinflskrkpsaerwrrd--ctswfgrctvnseccsnscdqtycelyafpsfga
mqkliilllvaavlmsaqavlq----ekrpkekikflskrktdaekqqkrl-cpdytepcshaheccswncynghctg--------
mqkliilllvaavlmstqavlq----ekrpkekikflskrktdaekqqkrl-cpdytdpcshaheccswncynghctg--------
mqkliilllvaavlmstqalfq----ekrlkekinflskekadaekqqkry-csdqwkscsypheccrwsc-nryca---------
mqkliilllvaavlmstqalfq----ekrpmkkidflskgktdaekqqkrs-csddwqycesptdccswdc-dvvcsg--------
```

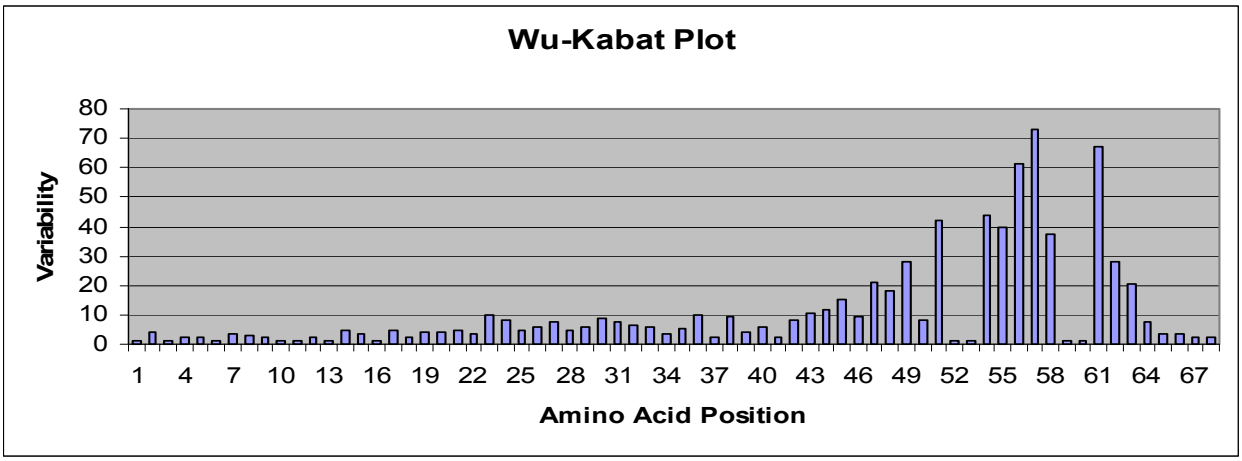
Multiple Conus #5




```
mlclpvfiiillllaspaapnplekriqsdliraaledadmktdereivniidsisd---vakqicceitvq-ccvldee
mlclpvfiiillllaspaapnplekriqsdliraaledadmktgereilniidsisd---vakqiccqitvd-ccvldee
mlclpvfiiillllaspaapnpletriqrdliraaledadmktnerflegvistikd---fagkvccsvsvnfccpta--
```

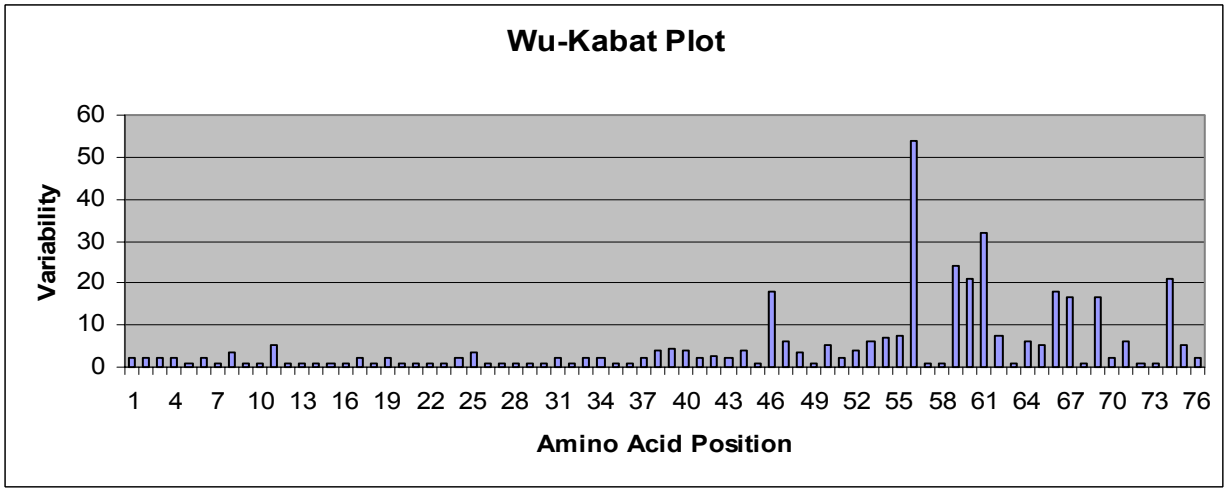
```
mlclpvfiilllllaspaasnplekriqsdliraaledadtkndpr-llglvtga----------ccvlkfsfccgkk--
mlclpvfiilllllaspaasnplekriqsdliraaledadtkndpr-lldyvtga----------ccaglnfvccgkk--
mlclpvfiilllllaspaasnplekriqndliraaledadmendprsiidsvktf----------ccstnlgiccskk--
mlclpvfiilllllaspaasnpletriqsdliraaledadmkneknilssimgslgtignvvgnvccsitks-ccasee-
mlclpvfiilllllaspaapnplerriqsdlirtaledadmktpkgvlsgimsnlgtvgnmvggfcctvysg-ccsek--
```

Multiple Conus #6




```
mrclpvfiilllllipsaas-aaqpetkddaalasfydnakrtlqrhwakslccpedawccshdegk---
mrclpvfiilllllipsaas-vaqpktkddvalasfydnakrtlqrhwakslccpedawcc---------
mhclpvlvilllliastpsvdarpktkddvppasfhgaanrilqtlwnlrgccedkt-ccfig------
mhclpvpvilllliastpsvdarpktkddvppasfhgaanrilrtlwnlrgccedkt-ccfig------
mhclpvlvilllliastpsvdarpktkddvplasfhgaanrilrtlwnlrgccedkt-ccfig------
mhclpvlvilllliastpsvdarpnpkddvplasfhgaakrylrtlwnsrdccdpkepccfig------
mrclpvfvilllliasapsvdarpktkddvplasfhgnaertllnilrddnccidkqgccpwg------
mrclpvfvilllliastpsvnarpktkd---lasfhdnakrtqhifwskrncciyenwccewi------
mrclpvfvilllliasvpsvdaelkakddmpqasfhdnaerdqqkkts--dccfyhn-ccc--------
mrclpvfvilllliasapsvdarpktkddiplvsfqdhakrilqtfesrydccktfe-cchwg------
mrclpvfvilllliastpsvdaqlktkddmslasfhdnvkrilqirtt-eeccpfivgccsr-------
mrclpvfvilllliastpsvdalqktkddmplasfhdnvkrilqtlsnkrsccptilsccfv-------
mrclpvfvilllliastpsvdaraktrddmslasfhddakrilqilqdrsgccvidsnccg--------
mrclpvfvilllliasapsvdarpktkddiplvsfqdnakralqilsnkryccyfdyscclylr-----
mcclpvfvilllliasapsvdalpktkddmslasfhdnakrtlqilsnkryccvydyscclswg-----
mrclpvfvilllliasapgvd---------------dfakktpkrlskprdccrrnflcc---------
mrclpvfvilllliasapgvdaqpktkydvplasrhdfakktpkrlskprdccrrnflcc---------
mrclpvfvilllliasapgvdaqpktkydvplasrhdfakktpkrlskprdcclrhflccv--------
mrclpvfvilllliasapgvdvqpktkyyvprasrrdfakktpkrlsklrgccprsflccr--------
mhclpvfvilllliasapgvdvqpktknfmtraslrdfakktpkrlsklrgccprsflccr--------
mrclpvfvilllliasapsvdaqpktkddipqasfldnakrylqvleskrnccrrqi-ccgrps-----
mrclpvfivllllivsapgfdarpkteddvplssfhddlqrtvrtlldirmcclgtsgccpwg------
mhclsvfvilllltasapsvdaqpkteddvplssfhddlqrtvrtlldirmcclgtsgccpwg------
mrcfpvfiilllllmasapsfdarpkteddvplssfrdnlkrtlrtlldprrccyetpgccvig------
mrcfpvfiilllliasapcfdartktdddvplsplrdnlkrtirtrlnireccedg-wcctaapltgr-
mrcfpvfiilllliasapcfdartktdddvplsslrdnlkrtirtrlnireccedg-wcctaapltgr-
mrclpvfiilllliasapsfdalpktednvplssfhdnlkrtrrihlnireccsdg-wccpag------
mrclpvfiillvlsasapsvdarpkneddvslasfhdnaqrtlqrllnkrsccpnnpacch--------
```

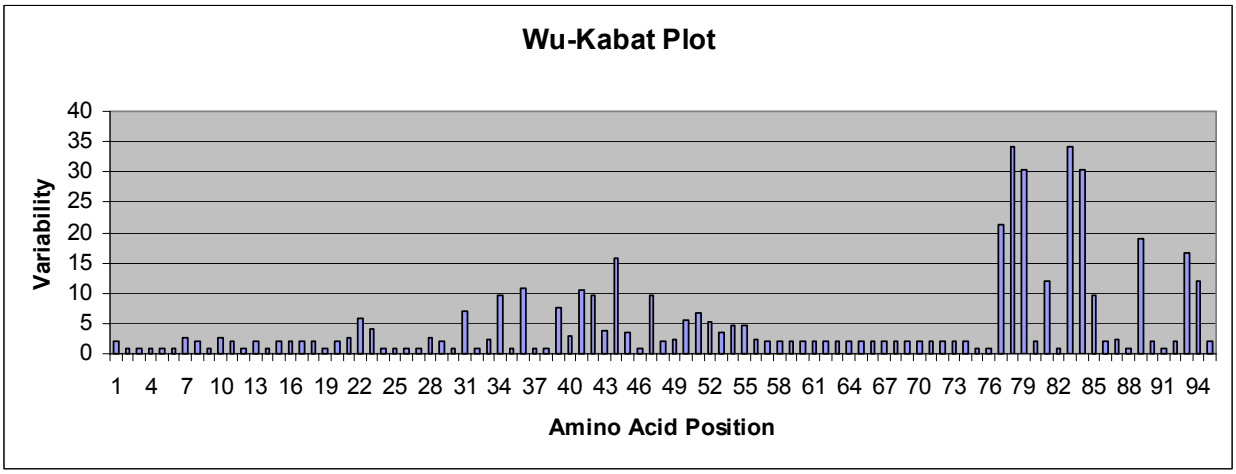
Multiple Conus #7




```
-mlkmgvvlfiflvlftlatlqldadqpveryaenkrlmspykrr-ailhapreqeccepqwcd-ggcd-acc---
-mlkmgvmlfiflvlfplatlqldadqpveryaenkrlmspderr-ailhaprqrgccepdwcd-sgcddgcc---
-mlkmgvvlfiflvlftlatlqldadqpveryaenkqllspderrdiilhalgqrrccdpdwcd-agcydgcc---
-mlkmgvvlfvflvlfplatlqldadqpveryaenkqlvspyerrqiilhalgqrdccvmpwcd-gacd--ccvss
-mlkmgvvlfvflvlfplatlqldadqpveryaenkqlvspyerrqiilhalgqrqccdwqwcd-gacd--cca--
-mlkmgvvlfiflvlfplatlqldadqpveryaenkqllspderreiilhalgtr-ccswdvcdhpsc--tcc-g-
```

```
----mgvvlfiflvlfplatlqldpdqpveryaenkqllnpderrgiilhalgqrvccppesct-drcl--cclg-
-mlkmgvvlfiflvlfplatlqldadqpveryaenkqllntderreiilsalrtrvccpfggch-elcq--cceg-
mmlkmgvvlfiflvlfplatlqldadqpveryaenkqllnpderrgillpalr-kfccdsnwchisdce--ccyg-
-mlkmgvvlftflvlfplatlqldadqpveryaenkqglnpdesreiilsalrqrdcceqgwcd-ggcd--ccq--
-mlkmgvllftflvlfplatlqldadqpveqyagnkqdlnpdekremilpalrqwqcctmawcd-sgcl--cce--
-mlkmevvlftflvlfplstlqletdqpveryvenkqdlnpdesrnfmlpivk--kccta--crmppck--cca--
```

Multiple Conus #8




```
mmsklgallticllllfpltalpldgdqpadrpaermqddisseqhplfd---kerq-----------------cctg--sc-ln--cwpccg--
mmsklgvllticllllfpltavpldgdqpadrpaermqdgissehhpffdsvkkkqq-----------------ccppv-acnmg--cepccg--
mmsklgalliicllllfpltavpldgdqhadrpaerlqddisskhhpmfdavrg--------------------cchps-tchmrkgcsrccs--
-msklgalliicllllfpltavpmdgdqpadrpaermqddisfeqhpmfdatrr--------------------ccnag-fcrf--gctpccy--
mmsklgvllticllllfphtavpldgdqhadqpaerlqddissehhpmlnsirrreqnqfmsftsvklrdsrgerccgpt-acma--gcrpccg--
mmsklgvllitcllllfpltavpldgdqpadqpaerlqddissenhpffdpvkr---------------------ccrl--lc-ls--cnpccg--
mmsklgvlliicllllcpltavpqdgdqpadqpaermqddissehhpffdpvkr---------------------cckygwtcwlg--cspcgc--
mmsklgvllticllllfpltavpldgdqhadrpadrmqd-isseqhplfdpvkr---------------------ccdwp--ctig--cvpcclp-
mmsklgvllticllllfpltavrldgddqhtdrpadrmqd-iateqhplfdpvkr---------------------ccdwp--ctig--cvpcclp-
mmsklgvllticmllfpltalpldgdqpadrpaermqddfiseqhplfnpikr---------------------ccdwp--ctig--cvpcck--
mmsklgallticllllfsltavpldgdqhadqpaqrlqdriptedhplfdpnkr---------------------ccdds-ecsys--cwpccyg-
mmsklgvllticllllfpltavqldgdqpadlpalrtqd-iatdhspwfdpvkr---------------------ccsry--cyi---cipccpn-
mmsklgvllticllllfpltavqldgdqpadlpalrtqd-istdhspwfdpvkr---------------------ccsry--cyi---cipccpn-
mmsklgvllticllllfpltvlpmdgdqpadlpalrtqd-iatdqspwfdpvkr---------------------ccsry--cwk---cipccpy-
mmsklgvllticllllfpltavsldgdqpadlpelraqd-faperspwfdpvrr---------------------ccsqd--crv---cipccpy-
mmsklgvllticllllfpltavqldgdqpadlpelraqd-faperspwfdpvrr---------------------ccsqd--crv---cipccpy-
mmsklgvllticllllfpltavqldgdqpadlpelraqd-faperspwfdpvrr---------------------ccsqd--crv---cipccpn-
mmsklgvvlticllpfpltalpmdgdqpadlpalrtqd-feperspwfdpvkr---------------------ccsqd--crv---cipccpy-
mmsklgvllticllllfpltavpmdgdqpadlpalrtqd-feperspwfdpvrr---------------------ccsqd--csv---cipccppp
```