Evolution of Antp-class Genes and Differential Expression of Hydra *Hox/paraHox* genes in Anterior Patterning

Dominique Gauchat*, Françoise Mazet*, Cédric Berney*◊, Michèl Schummer°‡, Sylvia Kreger°, Jan Pawlowski*◊ and Brigitte Galliot*•

•brigitte.galliot@zoo.unige.ch

Fig. 7) Conservation of Antp-class HDs from diploblasts to bilaterians:

A) Deduced HD sequences of the hydra *Dlx*, *Cnot* and *CnHex* genes. Percentages on the right indicate identity and similarity rates; dashes indicate identical amino acids.

B) For each Antp-class family, a consensus HD sequence was derived where family-specific residues are written in bold and underlined when conserved from diploblasts to bilaterians. At positions that show variable residues, numbers corresponding to the different residues are given. When only conservative substitutions were noted, the most frequent residue is given in lowercase (conservative substitutions taken into account: E/D, F/Y, I/L/V, K/R, N/Q, S/T). The number of compared sequences is given on the right. Family residues that are specific to deuterostomes (deut), protostomes (prot), diploblasts (dipl) and bilaterians (bilat) are indicated when significant.

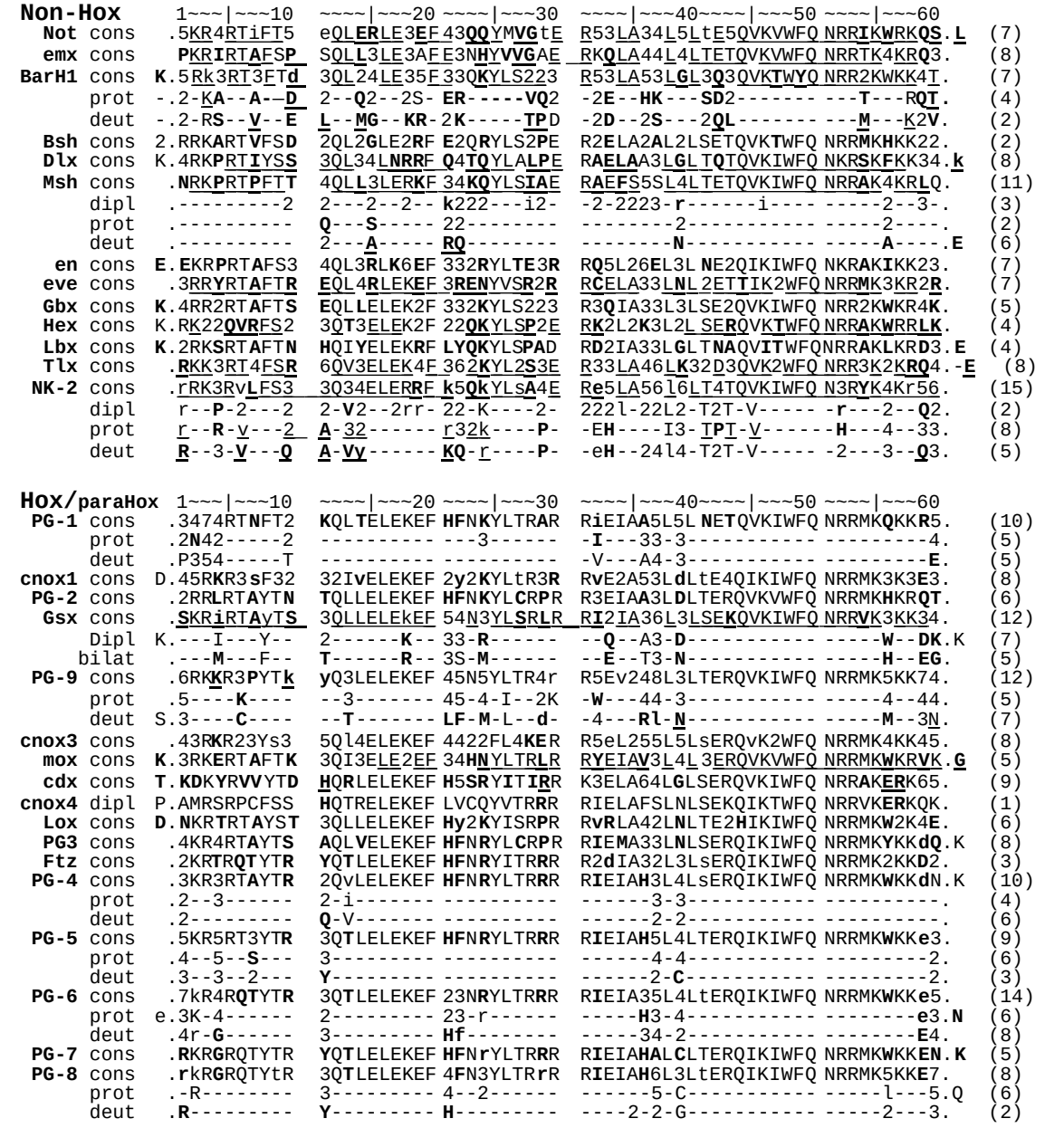
C) From the alignment of the family consensus sequences, an Antp-class consensus sequence was derived where the Antp-class specific residues are in bold and the variable positions indicated with a star. Highly (h) conserved or invariant (i) residues at positions which are common to most HDs classes are labelled above. The signature residues (bold) common either to the non-Hox or to the Hox/paraHox families are indicated underneath. The Prd-class specific residues are given for comparison.

D) Complete alignment of the 240 Antp-class sequences used for the phylogenetic analyses. The sequences were sorted according to their family. For each family, a consensus sequence (cons) and the specific (spec) residues were deduced. Residues are underlined when conserved from diploblasts to bilaterians. The phylum and the species codes are given at the end of the alignment.

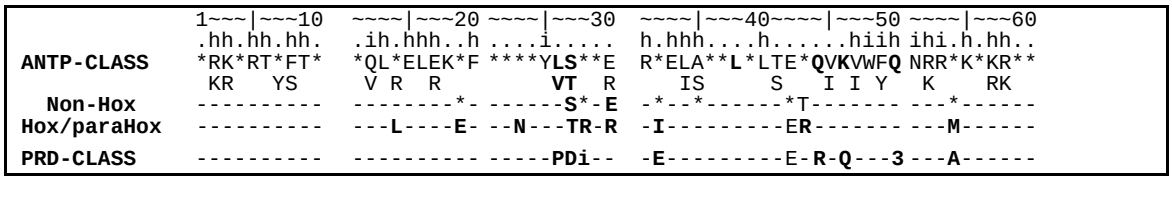
A

```
Prh/Hex      1~~~|~~~10 ~~~~|~~~20 ~~~~|~~~30 ~~~~|~~~40 ~~~~|~~~50 ~~~~|~~~60
  CnHex Hv  S.KKCVQVRFSH SQSTELERVF LVQKYISPYE RKQISRSLDL SERQIKTWFQ NRRAKWRRLK.A  AJ252185
    Prh Mm  K.R-GG-----N D-TV---KK- ET------P- --RLAKM-Q- ----V----- ----------.Q  68.3%78.3%
    Prh Ma  K.R-GG-----N D-TI---KK- ET---L--P- --RLAKM-Q- ----V----- ----------.Q  66.7%78.3%
    Prh Rn  K.R-GG-----N D-TV---KK- ET---L--P- --RLAKM-Q- ----V----- ----------.Q  66.7%78.3%
    Prh Gg  K.R-GG-----N E-TI---KK- ET---L--P- --RLAKL-Q- ----V----- ----------.Q  66.7%78.3%

Not          1~~~|~~~10 ~~~~|~~~20 ~~~~|~~~30 ~~~~|~~~40 ~~~~|~~~50 ~~~~|~~~60
   Cnot Hv  C.NKRRRTIFTI YQLERLECEF QQQQYLVGHE RRYLAKDLGL NEVQVKVWFQ NRRIKWRKSS.L  AJ252184
    Not Sp  R.A--V-----Q E------Q-- DR---M--S- -L---AE-N- S-S------- --------QN.   71.7%73.3%
  Xlhox Xl  K.L--I--V--P E------K-- LK---M--T- -VD--ST-N- T-T------- --------Q-.L  70%  71.7%
  Gnot1 Gg  K.M--V--V--P E------Q-- LK---M--T- -VD--AT-R- T-T------- --------Q-.M  70%  71.7%
  Cnot2 Gg  K.A--V-----S D--A---K-- AR---M--T- -CL--SA-H- T-E------- --------Q-.L  70%  70%
    Flh Br  K.S--M--S--N D--S---K-- AR---M--S- -FL--SA-Q- T-A------- --------Q-.L  68.3%68.3%

Distal-less 1~~~|~~~10 ~~~~|~~~20 ~~~~|~~~30 ~~~~|~~~40 ~~~~|~~~50 ~~~~|~~~60
    Dlx Hv        RTIYSS YQLRELNRRF FKTQYLSLPE KAELAETLGL TPTQVKIWFQ NKRSKFKEVV.R  AJ252183
   Dlx1 Mm  K.IRKP------ L--QA----- QQ----A--- R----AS--- -Q-------- -------KLM.K  76.7%82.1%
    Dll Jc  K.MRKP------ L--QQ----- QR----A--- -----AS--- -Q-------- -R---Y-KMM.K  75%  82.1%
   Dlx3 Ax  K.IRKP------ ---AA-Q--- Q-A---A--- R----AQ--- -Q-------- -R-----KLY.K  75%  82.1%
    Dll Dm  K.MRKP------ L--QQ----- QR----A--- R----AS--- -Q-------- -R---Y-KMM.K  73.2%82.1%
   Dlx3 Mm  K.VRKP------ ---AA-Q--- Q-A---A--- R----AQ--- -Q-------- -R---K-KLY.K  73.2%78.6%
    Dll Bf  K.MRKP----T- F--QQ----- QR----A--- R----AQ--- -Q-------- -R---Y-KLM.K  71.4%83.9%
   Dlx7 Mm  K.LRKP------ L--QH-DQ-- QH----A--- R-Q--AQ--- -Q-------- -----Y-KLL.K  69.6%76.7%
```

B



C



D

Antp-class Non-Hox families

```
Not                                                    Cnid. Ech. Vert.
                1~~~|~~~10~~~~|~~~20~~~~|~~~30~~~~|~~~40~~~~|~~~50~~~~|~~~60
   Cnot-Hv  C.NKRRRTIFTIYQLERLECEFQQQQYLVGHERRYLAKDLGLNEVQVKVWFQNRRIKWRKSS.L  AJ252184

    Not-Sp  R.AKRVRTIFTQEQLERLEQEFDRQQYMVGSERLYLAAELNLSESQVKVWFQNRRIKWRKQN.F  AF109903
  Gnot1-Gg  K.MKRVRTVFKPEQLERLEQEFLKQQYMVGTERVDLAATLRLTETQVKVWFQNRRIKWRKQS.M  U20615
  Cnot2-Gg  K.AKRVRTIFTSDQLARLEKEFARQQYMVGTERCLLASALHLTEEQVKVWFQNRRIKWRKQS.L  X98049
    Flh-Br  K.SKRMRTSFTNDQLSRLEKEFARQQYMVGSERFLLASALQLTEAQVKVWFQNRRIKWRKQS.L  L48017
 Xelhox-Xl  K.LKRIRTVFTPEQLERLEKEFLKQQYMVGTERVDLASTLNLTETQVKVWFQNRRIKWRKQS.L  L19566
   Xnot-Xl  K.MKRIRTVFTPEQLEKLEKEFLKQQYMVGTERVDLASTLNLTETQVKVWFQNRRIKWRKQS.L  Z19577

  Not cons   .5KR4RTiFT5eQLERLE3EF43QQYMVGtER53LA34L5LtE5QVKVWFQNRRIKWRKQS.L(7)
  Not-spec    ------I---e--E--------QQ-MVGt-----------------------I-W—QS

Emx (empty-spiracle)                                   Cnid. Arth. Vert.
                1~~~|~~~10~~~~|~~~20~~~~|~~~30~~~~|~~~40~~~~|~~~50~~~~|~~~60
   Ems-Hys  K.RKRHRTAFTPTQLLGLENSFERGHYLVGDERRQLAQFLRLTETQIKVWFQNRRTKWKRQR.N  Y11836

    Ems-Dm  K.PKRIRTAFSPSQLLKLEHAFESNQYVVGAERKALAQNLNLSETQVKVWFQNRRTKHKRMQ.Q  P18488

   Emx1-Hs  ~.PKRIRTAFSPSQLLRLERAFEKNHYVVGAERKQLAGSLSLSETQVKVWFQNRRTKYKRQK.L  Q04741
   Emx1-Mm  ~.PKRIRTAFSPSQLLRLERAFEKNHYVVGAERKQLAGSLSLSETQVKVWFQNRRTKYKRQK.L  Q04742
   Emx2-Br  K.PKRIRTAFSPSQLLRLEHAFEKNHYVVGAERKQLAHSLSLTETQVKVWFQNRRTKFKRQK.L  D32215
   Emx2-Hs  K.PKRIRTAFSPSQLLRLEHAFEKNHYVVGAERKQLAHSLSLTETQVKVWFQNRRTKFKRQK.L  Q04743
   Emx2-Mm  ~.PKRIRTAFSPSQLLRLEHAFEKNHYVVGAERKQLAHSLSLTETQVKVWFQNRRTKFKRQK.L  Q04744
   Emx1-Br  K.PKRIRTAFSPSQLLRLERAFEKNHYVVGAERKQLANGLCLTETQVKVWFQNRRTKHKRQK.L  D32214

  emx cons    PKRIRTAFSPSQLL3LE3AFE3NHYVVGAERKQLA44L4LTETQVKVWFQNRRTK4KRQ3.  (8)
  emx-spec    P--I--A--Ps--L----A-E-NH--VGA--kQ--------------------T----Q-

BarH1/BarH2                                          Cnid. Nema. Arth. Vert.
                1~~~|~~~10~~~~|~~~20~~~~|~~~30~~~~|~~~40~~~~|~~~50~~~~|~~~60
  Cnox3-Cv  K.CRKPRTVFSDLQLMVLEREFNNRKYLSTPQRTNLADRLGLNQTQVKTWYQNRRMKWKKET.F  X64627

  Ceh30-Ce  K.SRKARTIFTDKQLQELENTFEKQKYLSVQDRMDLAHRMGLTDTQVKTWYQNRRTKWKRQA.T  Q22909
   Hm1d-Da  K.QRKARTAFTDHQLQTLEKSFERQKYLSVQERQELAHKLDLSDCQVKTWYQNRRTKWMRQT.A  P22544
  Barh1-Dm  K.QRKARTAFTDHQLQTLEKSFERQKYLSVQERQELAHKLDLSDCQVKTWYQNRRTKWKRQT.A  M73259
  Barh2-Dm  K.QRKARTAFTDHQLQTLEKSFERQKYLSVQDRMELANKLELSDCQVKTWYQNRRTKWKRQT.A  M82884

  Barx1-Mm  K.GRRSRTVFTELQLMGLEKRFEKQKYLSTPDRIDLAESLGLSQLQVKTWYQNRRMKWKKIV.L  Y07960
  Barx2-Hs  K.PRRSRTIFTELQLMGLEKKFQKQKYLSTPDRLDLAQSLGLTQLQVKTWYQNRRMKWKKMV.L  AF031924

BarH1 cons  K.5RK3RT3FTd3QL24LE35F33QKYLS22dR53LA53LGL323QVKTWYQNRR2KWKK43.  (7)
BarH1 deut  -.2-RS--V--EL--MG--KR-2K-----TPD-2D--2S---2QL----------M----2V   (2)
BarH1 prot  -.2-KA--A--D2--Q2--2S-ER-----VQe-2E--HK---SD2----------T---RQT   (4)

Bsh (Brain-specific homeobox)                             Nema. Arth.
                1~~~|~~~10~~~~|~~~20~~~~|~~~30~~~~|~~~40~~~~|~~~50~~~~|~~~60
    Bsh-Dm  R.RRKARTVFSDPQLSGLEKRFEGQRYLSPPERVELATALGLSETQVKTWFQNRRMKHKKQL.A  L06475
 F31e83-Ce  S.RRKARTVFSDQQLQGLERRFESQRYLSTPERIELANALNLSETQVKTWFQNRRMKHKKVV.R  U55856

  Bsh cons  2.RRKARTVFSD2QL2GLE2RFE2QRYLS2PER2ELA2AL2LSETQVKTWFQNRRMKHKK22.   (2)
  Bsh spec   .R--A--V--D----G---R-E-QR----P-------A---------T------M-H----.

Dlx (Distal-less)                                      Cnid. Arth. Chor.
                1~~~|~~~10~~~~|~~~20~~~~|~~~30~~~~|~~~40~~~~|~~~50~~~~|~~~60
    Dll-Hv   .~~~~RTIYSSYQLRELNRRFFKTQYLSLPEKAELAETLGLTPTQVKIWFQNKRSKFKEVV.RA AJ252183

    Dll-Jc  K.MRKPRTIYSSLQLQQLNRRFQRTQYLALPEKAELAASLGLTQTQVKIWFQNRRSKYKKMM.KA AAB32450
    Dll-Dm  K.MRKPRTIYSSLQLQQLNRRFQRTQYLALPERAELAASLGLTQTQVKIWFQNRRSKYKKMM.KA P20009

    Dll-Bf  K.MRKPRTIYTSFQLQQLNRRFQRTQYLALPERAELAAQLGLTQTQVKIWFQNRRSKYKKLM.KQ P53772
  Dlx3-Ame  K.IRKPRTIYSSYQLAALQRRFQKAQYLALPERAELAAQLGLTQTQVKIWFQNRRSKFKKLY.KN Q90229
   Dlx1-Mm  K.IRKPRTIYSSLQLQALNRRFQQTQYLALPERAELAASLGLTQTQVKIWFQNKRSKFKKLM.KQ Q64317
   Dlx7-Mm  K.LRKPRTIYSSLQLQHLDQRFQHTQYLALPERAQLAAQLGLTQTQVKIWFQNKRSKYKKLL.KQ P70436
   Dlx3-Mm  K.VRKPRTIYSSYQLAALQRRFQKAQYLALPERAELAAQLGLTQTQVKIWFQNRRSKFKKLY.KN Q64205

  Dlx cons  K.4RKPRTIYSS3QL34LNRRFQ4TQYLALPERAELAA3LGLTQTQVKIWFQNKRSKFKK34.  (8)
  Dlx spec  K.—--P—-IY-S------N-R-Q-TQ--ALP-R-------G--Q----------S-F--L-

Msx (Muscle-specific homeobox)          Pori. Cnid. Nema. Arth. Echi. Chor.
                1~~~|~~~10~~~~|~~~20~~~~|~~~30~~~~|~~~40~~~~|~~~50~~~~|~~~60
  Prox3-Ef  K.DRKPRTPFTSTQLIALERKFRQQKYLSVAERAEFAEYLKLTETQVKIWFQNRRAKEKRLH.EA L23476
    Msh-Cv  A.NRKPRTPFSVNQLLTLEQKFKRKQYLSISERAELSELLRLTETQIKIWFQNRRAKQKRSK.EA X64629
    Msh-Hv  A.NRKPRTPFSVNQLLALEQKFKRKQYLSISERSELSVMLRLTETQIKIWFQNRRAKQKRRT.EA AJ271008

    Msh-Ce  N.NRKPRTPFSTQQLISLERKFQSKQYLSIAERAEFSASLQLTETQVKIWFQNRRAKSKRLQ.EA Q09604
    Msh-Dm  P.NRKPRTPFTTQQLLSLEKKFREKQYLSIAERAEFSSSLRLTETQVKIWFQNRRAKAKRLQ.EA P13297

    Msx-Sp  T.NRKPRTPFTTSQLLALERKFRQKQYLSIAERAEFSASLNLTETQVKIWFQNRRAKAKRLQ.EA AF042653
   Msx-Hsp  ~.MRKPTTPFTTQQLLALERKFRERQYLSIAERAEFSSSLNLTETQVKIWFQNRRAKSKRLQ.EA U61846
    Msx-Bf  T.NRKPRTPFTTQQLLALERKFRQKQYLSIAERAEFSASLNLTETQVKIWFQNRRAKAKRLQ.EA AJ130766
   Mshb-Br  T.NRKPRTPFSTSQLLSLERKFRQKQYLSIAERAEFSNSLNLTETQVKIWFQNRRAKAKRLQ.EA Q03356
    Msh-Mm  P NRKPRTPFTTQQLLSLEKKFREKQYLSIAERAEFSSSLRLTETQVKIWFQNRRAKAKRLQ.EA P13297
   Msx2-Mm  T.NRKPRTPFTTSQLLALERKFRQKQYLSIAERAEFSSSLNLTETQVKIWFQNRRAKAKRLQ.EA Q03358

  Msh cons   .NRKPRTPFTT4QLL3LERKF34KQYLSIAERAEFS5SL4LTETQVKIWFQNRRAK4KRLQ.EA  (11)
  Msh dipl   .---------22---2--2--k222---i2--2-2223-r-----i----------2--3-.    (3)
  Msh prot   .----------Q---S-----22----------------2----------------2----.    (2)
  Msh deut   .----------2---A-----RQ----------------N----------------A----.    (6)
  Msx spec   .N—-P—-P--t---L----K-r-KQ---IA--A-F-—S-—--------------A—---LQ
```

Engrailed (en) **Nema. Arth. Echi. Chor.**

```
             1~~~|~~~10~~~~|~~~20~~~~|~~~30~~~~|~~~40~~~~|~~~50~~~~|~~~60
  Ceh16-Ce  E.EKRPRTAFTGDQLDRLKTEFRESRYLTEKRRQELAHELGLNESQIKIWFQNKRAKLKKST.S  P34326
     en-Dm  D.EKRPRTAFSSEQLARLKREFNENRYLTERRRQQLSSELGLNEAQIKIWFQNKRAKIKKST.G  P02836
    inv-Dm  E.DKRPRTAFSGTQLARLKHEFNENRYLTEKRRQQLSGELGLNEAQIKIWFQNKRAKLKKSS.G  P05527
    en-Afr  D EKRPRTAFTAEQLSRLKHEFNENRYLTERRRQDLARELGLHENQIKIWFQNNRAKLKKSS.G  Q05640

     En-Tg  D.EKRPRTAFSASQLQRLKQEFQQSNYLTEQRRRSLAKELTLSESQIKIWFQNKRAKIKKAS.G  P09532
Amphien-Bf  E.EKRPRTAFTSEQLQRLKKEFQENRYLTEQRRQDLARELKLNESQIKIWFQNKRAKIKKAA.G  U82487
     En-Br  E.DKRPRTAFTAEQLQRLKNEFQNNRYLTEQRRQALAQELGLNESQIKIWFQNKRAKIKKAT.G  P31533

   en cons  E.EKRPRTAFS34QL3RLK6EF332RYLTE3RRQ5L26EL3LNE2QIKIWFQNKRAKIKK23.   (7)
   en spec  E.E--P—-A-------R-K---------T--R-Q----E-—-N----------K-A-I----.G
```

Eve **Cnid. Arth. Vert.**

```
             1~~~|~~~10~~~~|~~~20~~~~|~~~30~~~~|~~~40~~~~|~~~50~~~~|~~~60
   Eve-Acf  R.TRRYRTAFTREQLSRLEKEFLRENYVSRTRRSELASMLNLSETTIKIWFQNRRMKAKRRR.M  S36770
Antheve-Nve  .TRRYRTAFTREQLKRLEKEFMRENYVSRTRRCELANALNLSETTIKIWFQNRRMKSKRRR.M  Finnerty, 1998

    Eve-Sa   .IRRYRTAFTREQLARLEKEFYKENYVSRPRRCELASQLNLPESTIKVWFQNRRMKDKRQR.M  Z11845
    Eve-Dm   S.VRRYRTAFTRDQLGRLEKEFYKENYVSRPRRCELAAQLNLPESTIKVWFQNRRMKDKRQR.I  P06602

   Evx2-Br   Q.VRRYRTAFTREQIGRLEKEFYRENYVSRPRRCELAAALNLPETTIKVWFQNRRMKDKRQR.L  X99290
   Evx1-Mm   Q.MRRYRTAFTREQIARLEKEFYRENYVSRPRRCELAAALNLPETTIKVWFQNRRMKDKRQR.L  P23683
   Evx2-Mm   Q.VRRYRTAFTREQIARLEKEFYRENYVSRPRRCELAAALNLPETTIKVWFQNRRMKDKRQR.L  P49749

  eve cons    .3RRYRTAFTREQL4RLEKEF3RENYVSR2RRCELA33LNL2ETTIK2WFQNRRMK3KR2R.  (7)
  eve spec    .---Y—-A—-RE---R—--E—-REN---R-R-C------N----T---------M—----R.
```

Gbx **Nema. Vert.**

```
             1~~~|~~~10~~~~|~~~20~~~~|~~~30~~~~|~~~40~~~~|~~~50~~~~|~~~60
F33d114-Ce   .KMRRPRTAFSSEQLVQLEKQFSDNRYLSRPRRYQLAQQLSLSETQIKIWFQNRRMKNKRCP.S  AF039720

   Gbx2-Hs   .KNRRRRTAFTSEQLLELEKEFHCKKYLSLTERSQIAHALKLSEVQVKIWFQNRRAKWKRVK.A  P52951
   Gbx2-Mm   .KNRRRRTAFTSEQLLELEKEFHCKKYLSLTERSQIAHALKLSEVQVKIWFQNRRAKWKRVK.A  P48031
   Hmx7-Gg   .KSRRRRTAFTSEQLLELEKEFHCKKYLSLTERSQIAHALKLSEVQVKIWFQNRRAKWKRIK.A  P15142
  Hlxb9-Hs   .KCRRPRTAFTSQQLLELEHQFKFNKYLSRPKRFEVATSLMLTETQVKIWFQNRRMKWKRSK.K  P50219

  Gbx cons   .K4RR2RTAFTSEQLLELEK2F332KYLS223R3QIA33L3LSE2QVKIWFQNRR2KWKR4K.   (5)
  Gbx spec    K------A—-SE—-L---------K--------Q-----—----------------W--—K
```

Prh/Hex **Cnid. Vert.**

```
             1~~~|~~~10~~~~|~~~20~~~~|~~~30~~~~|~~~40~~~~|~~~50~~~~|~~~60
   CnHex-Hv  .SKKCVQVRFSHSQSTELERVFLVQKYISPYERKQISRSLDLSERQIKTWFQNRRAKWRRLK.A  AJ252185

     Prh-Hs  .KRKGGQVRFSNDQTIELEKKFETQKYLSPPERKRLAKMLQLSERQVKTWFQNRRAKWRRLK.Q  Q03014
     Prh-Mm  .KRKGGQVRFSNDQTVELEKKFETQKYLSPPERKRLAKMLQLSERQVKTWFQNRRAKWRRLK.Q  P43120
     Prh-Gg  .KRKGGQVRFSNEQTIELEKKFETQKYLSPPERKRLAKLLQLSERQVKTWFQNRRAKWRRLK.Q  Q05502

   Hex cons   KRK22QVRFS23QT3ELEK2F22QKYLSP2ERK2L2K3L2LSERQVKTWFQNRRAKWRRLK.   (4)
   Hex spec   ----—-QVR--—--T---------QK--SP---K---K------R---T------A-W--LK.
```

Ladybird **Arth. Vert.**

```
             1~~~|~~~10~~~~|~~~20~~~~|~~~30~~~~|~~~40~~~~|~~~50~~~~|~~~60
 Ladyb1-Dm   .KKRKSRTAFTNHQIFELEKRFLYQKYLSPADRDEIAASLGLSNAQVITWFQNRRAKQKRDI.E  Y08821
 Ladyb2-Dm   .KKRKSRTAFTNQQIFELEKRFLYQKYLSPADRDEIAGGLGLSNAQVITWFQNRRAKLKRDM.E  Y08822

   Lbx1-Hs   .KRRKSRTAFTNHQIYELEKRFLYQKYLSPADRDQIAQQLGLTNAQVITWFQNRRAKLKREL.E  P52954
   Lbx1-Mm   .KRRKSRTAFTNHQIYELEKRFLYQKYLSPADRDQIAQQLGLTNAQVITWFQNRRAKLKRDL.E  P52955

  Lbx cons   .K2RKSRTAFTNHQIYELEKRFLYQKYLSPADRD2IA33LGLTNAQVITWFQNRRAKLKRD3.E(4)
  Lbx spec   .K---S—-A—-NH—-Y----R-LYQK---PAD-D-I----G--NA--IT------A-L--D-
```

Tlx/Hox11 **Pori. Nema. Arth. Echi. Vert.**

```
             1~~~|~~~10~~~~|~~~20~~~~|~~~30~~~~|~~~40~~~~|~~~50~~~~|~~~60
   _Emh3-Em  R.RKKARTAFSREQVAELEKKFQDKKYLSSAERGELAEKLKLSDMQVKTWFQNRRMKYKRQS.EE U97664
   Prox2-Ef  ~.RKKARTAFSREQVAELEKKFQNKKYLSSTERGELAEKLKLSDMQVKTWFQNRRMKYKRQS.EE L10985
    Ceh9-Ce  ~.RKKARTTFSGKQVFELEKQFEAKKYLSSSDRSELAKRLDVTETQVKIWFQNRRTKWKKIE.~~ P56407
   Hox11-Dm  K.RKKPRTSFTRIQVAELEKRFHKQKYLASAERAALARGLKMTDAQVKTWFQNRRTKWRRQT.AE Z22959
    Hmx-Stp  K.KKKTRTVFSRSQVFQLESTFEVKRYLSSSERAGLAANLHLTETQVKIWFQNRRNKWKRQM.AA D85079
   Hox11-Hs  K.KKKPRTSFTRLQICELEKRFHRQKYLASAERAALAKALKMTDAQVKTWFQNRRTKWRRQT.AE P31314
   Hox11-Mm  K.KKKPRTSFTRLQICELEKRFHRQKYLASAERAALAKALKMTDAQVKTWFQNRRTKWRRQT.AE P43345
    Tlx3-Gg  K.RKKPRTSFSRVQICELEKRFHRQKYLASAERAALAKSLKMTDAQVKTWFQNRRTKWRRQT.AE AF071875

  Tlx cons    .RKK3RT4FSR6QV3ELEK4F362KYL2S3ER33LA46LK32D3QVK2WFQNRR3K2KRQ4.-E(8)
Tlx-specif    .R--P-----R—------------K---S----------K-------T------T-W--Qs.-E
```

```
NK-2                                          Pori. Cnid. Plat. Nema. Arth. Uro. Chor.
              1~~~|~~~10~~~~|~~~20~~~~|~~~30~~~~|~~~40~~~~|~~~50~~~~|~~~60
  Prox1-Ef  K.KRRPRALFSHAQVYELERRFAVQKYLTAHEQSKLATVLHLTETQVKIWFQNRRYKSKRQQ.I  L10984
    Nk2-Hv  S.RKKPRILFSQSQVMELGKKFKDQKYLSASERDQIANKLNLTPTQVKIWFQNKRYKCKKQT.I  AF012538

    Hbx3-Eg  S.QSKRRVLFNKFQISQLEKRLKQ?RYLTAQERQELAHTIGLTPTQVKIWFQNHAYKMKRLF.H  P42585
    Dth2-Gt  Q.RRKRRILFSQAQIYELERRFKQQKYLSAPEREHLANLINLTPTQVKIWFQNHRYKCKRSQ.K  Q00401
    Dth1-Gt  K.KRKRRVLFSKKQILELERHFRQKKYLSAPEREHLANLIGLSPTQVKIWFQNHRYKMKRAH.H  Q00400
   Ceh22-Ce  K.KRKRRVLFTKAQTYELERRFRSQKYLSAPEREALAMQIRLTPTQVKIWFQNHRYKTKKSH.T  P41936
     Tin-Dm  M.KRKPRVLFSQAQVLELECRFRLKKYLTGAEREIIAQKLNLSATQVKIWFQNRRYKSKRGD.I  P22711
      Nk2-Dm  K.KRKRRVLFTKAQTYELERRFRQQRYLSAPEREHLASLIRLTPTQVKIWFQNHRYKTKRAQ.N  P22808
      Nk4-Dm  M.KRKPRVLFSQAQVLELECRFRLKKYLTGAEREIIAQKLNLSATQVKIWFQNRRYKSKRGD.I  D33976
      Vnd-Dm  K.KRKRRVLFTKAQTYELERRFRQQRYLSAPEREHLASLIRLTPTQVKIWFQNHRYKTKRAQ.N  P22808

     Nkx-Ci  T.RRKPRVLFSQAQVFELERRFKQQRYLSAPEREHLAQMLKLTSTQVKIWFQNRRYKCKRMR.Q  AB012666
    Nk21-Bf  Q.RRKRRVLFSQAQVYELERRFKQQKYLSAPEREHLAQLINLTPTQVKIWFQNHRYKCKRQD.K  AF077840
    Nk25-Mm  R.RRKPRVLFSQAQVYELERRFKQQRYLSAPERDQLASVLKLTSTQVKIWFQNRRYKCKRQR.Q  P42582
   Nkx26-Mm  P.QRKSRVLFSQAQVLALERRFKQQRYLTAPEREHLASALQLTSTQVKIWFQNRRYKSKSQR.Q  P43688
    Ttf1-Rn  P.RRKRRVLFSQAQVYELERRFKQQKYLSAPEREHLASMIHLTPTQVKIWFQNHRYKMKRQA.K  P23441

 NK-2 cons   .rRK3RvLFS33Q34ELERRFk5QkYLsA4ERe5LA56l6LT4TQVKIWFQN3RYK4Kr56. (15)
      dipl    r--P-2---22-V2--2rr-22-K----2-222l-22L2-T2T-V------r---2-rQ2.  (2)
      prot    k--R-v---2A-32------r32k----P--EH----I3-TPT-V------H---4-R33. (8)
      deut    R--3-V---QA-Vv------KQ-r----P--eH--24l4-T2T-V------2---3-RQ3. (5)
 NK-2 spec    R----VL--QA--Y----R-k-Qk---AP--eH--------P-----------Y------.

Orphans                                                  Plat. Neme. Arth.
              1~~~|~~~10~~~~|~~~20~~~~|~~~30~~~~|~~~40~~~~|~~~50~~~~|~~~60
     Nk-Ls  K.PRRARTAFTYEQLVALENKFKTTRYLSVCERLNLALSLNLTETQVKIWFQNRRTKWKKQN.   Y16575

    Bcd-Dm  R.PRRTRTTFTSSQIAELEQHFLQGRYLTAPRLADLSAKLALGTAQVKIWFKNRRRRHKIQS.D  P09081

    Hbx4-Eg  .SRRERTIYTPEQLEAMEEVFGVNRYPDVSMREELASRLGINESKIQVWFKNRRAKLRNLE.   P55813
```

Antp-class Hox/paraHox* families

```
PG-1                                                Plat. Neme. Ann. Arth. Chor.
              1~~~|~~~10~~~~|~~~20~~~~|~~~30~~~~|~~~40~~~~|~~~50~~~~|~~~60
  Pnox3-Pn  N.NISVRTNFTTKQLTELEKEFHFNQYLTRARRIEIASKMTLSETQIKIWFQNRRMKQKRRQ.K  L41848
    Lab-Dm  T.NNSVRTNFTTKQLTELEKEFHFNRYLTRARRIEIANTLQLNETQVKIWFQNRRMKQKKRV.K  P10105
   Hox1-Ls  Q.PNTGRTNFTNKQLTELEKEFHFNKYLTRARRIEIAAALGLNETQVKIWFQNRRMKQKKRM.K  Y16570
   Lox7-Ht  S.NNLGRTNFTNKQLTELEKEFHFNKYLTRARRIEIASTLGLNETQVKIWFQNRRMKHKKRL.K  Y10888
   Lab-Nvi  Q.PNMGRTNFTNKQLTELEKEFHFNKYLTRARRIEIAAALGLNETQVKIWFQNRRMKQKKRM.K  AF151663

    Hox1-Bf  G.PNNGRTNFTTKQLTELEKEFHFYKYLTRARRVEIAAALNLNETQVKIWFQNRRMKQKKRE.K  Z35142
Ghoxlab-Gg  Q.PNTIRTNFTTKQLTELEKEFHFNKYLTRARRVEIAATLELNETQVKIWFQNRRMKQKKRE.K  P31259
  Hoxa1-Mm  Q.PNAVRTNFTTKQLTELEKEFHFNKYLTRARRVEIAASLQLNETQVKIWFQNRRMKQKKRE.K  P09022
  Hoxb1-Mm  Q.GGLGRTNFTTRQLTELEKEFHFNKYLSRARRVEIAATLELNETQVKIWFQNRRMKQKKRE.K  P17919
  Hoxd1-Mm  P.PSAIRTNFSNKQLTELEKEFHFNKYLTRARRIEIAANCQLNDTQVKIWFQNRRMKQKKRE.K  Q01822

PG-1 cons    .3474RTNFTTKQLTELEKEFHFNKYLTRARRiEIAA5L5LNETQVKIWFQNRRMKQKKR5.K  (10)
      prot   .2N42-----2-------------3-------I---33-3--------------------4.   (5)
      deut   .P354---------------------------V---A4-3--------------------E.   (5)
 PG-1 spec   .PN*---NF-TK--T------HF-K----A------A----N-T------------Q--R-.

Cnox-1                                                            Cnid.
              1~~~|~~~10~~~~|~~~20~~~~|~~~30~~~~|~~~40~~~~|~~~50~~~~|~~~60
  Antp Acf   .SNKKRFTFTQRQLVELEKEFHFSKYLTRTRRIEIASNLDLTETQIKIWFQNRRMKWKREL.   S36771
Anthox6 Nve P.SQKKRFTFTQRQLVELEKEFHFSKYLTRTRRIEIATTLKLTEMQIKIWFQNRRMKWKGEF.K  AF085282
  Cnox-1 Pc  E.LSKKRVCFSQKQIVELEKEFHFNKYLTRARRVEISHTLDLTEAQIKIWFQNRRMKHKREQ.K  X81455
  Cnox-5 Ed  D.ICKKRVCFTQKQIVELEKEFHYNRYLTRARRVEIAQLLKLTEAQIKIWFQNRRMKQKREQ.D  U41842
  Cnox-1 Cv  E.SFRKRCSFGHRKIIELEREFKYNRYLTRDRRLEFARNLDLSESQIKVWFQNRRMKQKKEQ.T  X64625
  Cnox-1 Hv  D.AFRKRCSFGHRKIIELEKEFKYNKYLSRDRRVEFARNLELSESQIKIWFQNRRMKQKKEQ.T  AJ252181
  Cnox-4 Hm  D.AFRKRCSFGHSKIIELEKEFKYNKYLSRDRRVEFARNLELSESQIKIWFQNRRMKQKKEQ.T  S39067
  Cnox-3 Hv  D.AFRKRCSFGHSKIIELEKEFKYNKYLSRDRRVEFARNLELSESQIKIWFQNRRMKQKKEQ.T  L22787

cnox1 cons  D.45RKR3sF3232IvELEKEF2y2KYLtR3RRvE2A53LdLtE4QIKIWFQNRRMK3K3E3.   (8)
cnox1 spec  D.---K---F--r--v-------Y-K--------------d------------------E-

PG-2                                                      Ann. Arth. Chor.
              1~~~|~~~10~~~~|~~~20~~~~|~~~30~~~~|~~~40~~~~|~~~50~~~~|~~~60
     Pb-Al  ~.PRRLRTAYTNTQLLELEKEFHFNKYLCRPRRIEIAASLDLTERQVKVWFQNRRMKHKRQS.   AF071406
    Pb-Nvi  N.PRRLRTAYTNTQLLELEKEFHFNKYLCRPRRIEIAASLDLTERQVKV                 AF151664
     Pb-Dm  L.PRRLRTAYTNTQLLELEKEFHFNKYLCRPRRIEIAASLDLTERQVKVWFQNRRMKHKRQT.L  P31264

    Hox2-Bf  S.SRRLRTVFTNTQLLELEKEFHYNKYVCKPRRKEIASYLDLNERQVKIWFQNRRMRQKRRD.T  Z35143
  Hoxa2-Mm  G.SRRLRTAYTNTQLLELEKEFHFNKYLCRPRRVEIAALLDLTERQVKVWFQNRRMKHKRQT.Q  P31245
  Hoxb2-Hs  G.ARRLRTAYTNTQLLELEKEFHFNKYLCRPRRVEIAALLDLTERQVKVWFQNRRMKHKRQT.Q  P14652

 PG-2 cons   .3RRLRTAYTNTQLLELEKEFHFNKYLCRPRR3EIAA3LDLTERQVKVWFQNRRMKHKRQT.   (6)
 PG-2 spec   .---L—-A--NT—--------HF-K--C-P---------D----------------H--QT.
```

```
Gsh* / Cnox-2                                          Plac. Cnid. Arth. Chor.
             1~~~|~~~10~~~~|~~~20~~~~|~~~30~~~~|~~~40~~~~|~~~50~~~~|~~~60
 Trox-2 Ta  R.TKRIRTAYTSMQLLELEKEFNSSRYLSRLRRIEIANMLNLSEKQVKIWFQNRRVKWKKDN.K  AF052038
Anthox2 Nve R.SKRIRTAYTSMQLLELEKEFSQNRYLSRLRRIQIAALLDLSEKQVKIWFQNRRVKWKKDK.K  AF085283
 Scox-2 Cx  K.SKRIKTAYTSIQLLELEKEFQNNRYLSRLRRIQIAAMLDLTEKQVKIWFQNRRVKWKKDK.K  AF124592
 Cnox-2 Ed  K.SKRIRTAYTSIQLLELEKEFQNNRYLSRLRRIQIAAMLDLTEKQVKIWFQNRRVKWKKDK.   Kuhn et al. 1996
Cnox-2 Hys  K.VKRIRTAYTSIQLLELEKEFQNNRYLSRLRRIQIAAMLDLTEKQVKIWFQNRRVKWKKDK.K  AF031953
 Cnox-2 Cv  K.SKRIRTAYTSIQLLELEKEFQNNRYLSRLRRIQIAAILDLTEKQVKIWFQNRRVKWKKDK.K  X64626
 Cnox-2 Hv  K.SKRIRTAYTSIQLLELEKEFQNNRYLSRLRRIQIAAILDLTEKQVKIWFQNRRVKWKKDK.K  M62870, AJ277388

    Ind-Dm  S.SKRIRTAFTSTQLLELEREFSHNAYLSRLRRIEIANRLRLSEKQVKIWFQNRRVKQKKGG.S  AF095926

    Gsx-Bf   .SRRMRTAFSSTQLLELEREFASNMYLSRLRRIEIATFLNLSEKQVKIWFQNRRVKHKKEA.   AF052463
   Gsh1-Ol  S.SKRMRTAFTSTQLLELEREFTSNMYLSRLRRIEIATYLNLSEKQVKIWFQNRRVKHKKEG.K  AF035573
   Gsh1-Mm   .SKRMRTAFTSTQLLELEREFASNMYLSRLRRIEIATYLNLSEKQVKIWFQNRRVKHKKEG.K  P31315
   Gsh2-Mm  N.GKRMRTAFTSTQLLELEREFSSNMYLSRLRRIEIATYLNLSEKQVKIWFQNRRVKHKKEG.K  S79041

  Gsx cons   .SKRiRTAyTS3QLLELEKEF54N3YLSRLRRI2IA36L3LSEKQVKIWFQNRRVK3KK34.   (12)
      dipl  K.---I------2------K--33-R--------Q--A3-D----------------W--DK.K  (7)
     tripl   .---M------T------R--3S-M--------E--T3-N----------------H--EG.   (5)
  Gsx spec   .S--i--A--S----------------S-L-------------K----------V------.

PG-9 / PG-10                                       Neme. Ann. Arth. Ech. Uro. Chor.
             1~~~|~~~10~~~~|~~~20~~~~|~~~30~~~~|~~~40~~~~|~~~50~~~~|~~~60
    Hox9-Ls  R.TRKKRKPYTRYQTMVLENEFLTNSYITRQKRWEISCKLHLTERQVKVWFQNRRMKRKKLN.A  Y16574
  Post2-Nvi  R.QRKKRKPYTRYQTMVLENEFMGNSYITRQKRWEISCKLHLSERQVKVWFQNRRMKRKKLN.E  AF151673
  Post1-Nvi  H.MRKKRKPYSKYQIAELEKEYVNNTYITKPKRWELSQRLNLSERQVKIWFQNRRMEKKVT.D  AF151672
    AbdB-Dm  T.VRKKRKPYSKFQTLELEKEFLFNAYVSKQKRWELARNLNLTERQVKIWFQNRRMKNKKNS.Q  P09087

    Hbx4-Tg  S.GRKKRCPYTKFQTLELEKEFLFNMYLTRDRRLEIARLLSLTERQVKIWFQNRRMKMKKQN.R  P10179
    Hox2-Sc  N.GRKKRVPYTKYQLLELEKEFHYNQYLSRERRQEVAKAVSLSDRQVKIWFQNRRMKWKKEK.K  U05571
    Hox9-Bf  S.SRKKRCPYTRFQTLELEKEFLYNMYLTRERRYEISQHVNLTERQVKIWFQNRRMKMKKMS.K  Z35149
   Hox10-Bf  ~.~~KKRCPYTKYQILELEKEFLFNMYVSRERRQEISRHVNLSDRQVKIWFQNRRMKMKRMN.K  Z35150
   Hoxb9-Mm  S.SRKKRCPYTKYQTLELEKEFLFNMYLTRDRRHEVARLLNLSERQVKIWFQNRRMKMKKMN.K  P20615
   Hoxc9-Mm  S.TRKKRCPYTKYQTLELEKEFLFNMYLTRDRRYEVARVLNLTERQVKIWFQNRRMKMKKMN.K  P09633
   Hoxd9-Mm  S.TRKKRCPYTKYQTLELEKEFLFNMYLTRDRRYEVARILNLTERQVKIWFQNRRMKMKKMS.K  P28357
   Hoxa9-Mm  S.TRKKRCPYTKHQTLELEKEFLFNMYLTRDRRYEVARLLNLTERQVKIWFQNRRMKMKKIN.K  P09631

  PG-9 cons   .6RKKR3PYTKYQ3LELEKEF45N5YLTR4rR5Ev248L3LTERQVKIWFQNRRMK5KK74.   (12)
       prot   .5----K------3--------45-4-I--3K-W---44-3----------------4--44.   (5)
       deut  S.3----C------T--------LF-M-L--d--4---Rl-N----------------M--32.   (7)
  PG-9 spec    ------P--ky-T-------L-----------------N--------------------N

Cnox-3                                                                 Cnid.
             1~~~|~~~10~~~~|~~~20~~~~|~~~30~~~~|~~~40~~~~|~~~50~~~~|~~~60
 Scox-1 Cx  S.RHRKRMTYSRNQILELEKEFLYSRYLTKERRKDLSDSLRLTERQIKIWFQNRRTKSKKER.K  AF124591
 Scox-4 Cx  C.KRRKRMTYTSSQTLELEKEFLYNRYLAKDRRKELSQMLCLSERQVKIWFQNRRVKSKRGI.N  AF124594
 Cnox-3 Ed   .SKRKRTSYSRRQIFELENEFNRSRYITREKRIELSMILNLTERQVKTWFQNRRMKTKKEK.T  U41840
 Cnox-1 Ed   .SKRKRMTYSRKQLLELEKEFHLSHFLKKERRVDLAKQLNLSERQIKIWFQNRRMKFKKEN.~  Kuhn et al. 1996
 Cnox-1 Hm  H.SKRKRMSYSKFQLHELEKEFSFNHFLRKERRTELAKLLKFSDRQIKIWFQNRRMKFKKEI.N  Z22638
 Cnox-3 Hv  H.SKRKRMTYSKFQLHELEKEFSFNHFLRKERRTELAKLLKFSDRQIKIWFQNRRMKFKKEI.N  AJ252182
Anthox1 Nv  E.GKRKRTAYTRKQLLELEKEFHFNHFLTKERRSEMASQLNLTERQVKIWFQNRRMLWKKCN.A  Finnerty 1998
 Scox-3 Cx  E.RRRKRTAYTRKQLVELEKEFHFNHFLTRERRLQLASSLSLSERQIKIWFQNRRMKWKKRG.T  AF124593

**cnox3** cons   .43R**K**R23Ys**R**5Ql4ELEKEF4422FL4K**E**RR5eL255L5LsE**R**QvK2WFQNRR**M**K4KK45.   (7)
cnox3 spec   .---K-----R-----------------E-------------------------------.

Mox* / Cnox-5                                          Cnid. Plat. Ann. Vert.
             1~~~|~~~10~~~~|~~~20~~~~|~~~30~~~~|~~~40~~~~|~~~50~~~~|~~~60
  Cnox5-Hm   .~~~~~~~~~~~~~~~LENEFVRNNYLTRLRRYEIAVSLSLSERQVKVWFQNRRMKWKRVK.G  Z22640
Dutarh2-Gt  K.QRKERTAFSKGQILELEKEFAVHNYLTRLRRYELAVALNLNERQIKVWFQNRRMKCKRSK.G  Z32525
  Hrox1-Hr  K.PRKERTAFTKHQIQELEKEFNVHNYLTRLRRYELAVSLELSERQVKVWFQNRRMKWKRVK.G  S38380
   Mox1-Mm  K.ARKERTAFTKEQLRELEAEFAHHNYLTRLRRYEIAVNLDLSERQVKVWFQNRRMKWKRVK.G  P32442
   Mox2-Mm  K.PRKERTAFTKEQIRELEAEFAHHNYLTRLRRYEIAVNLDLTERQVKVWFQNRRMKWKRVK.G  P32443

  Mox cons  K.3RKERTAFTK3QI3ELE2EF34HNYLTRLRRYEIAV3L4L3ERQVKVWFQNRRMKWKRVK.G  (5)
  Mox spec  K.---E--AF-K------------HN----L--Y---V------------------W--VK.G

Caudal (Cdx)                                                 Neme. Arth. Chor.
             1~~~|~~~10~~~~|~~~20~~~~|~~~30~~~~|~~~40~~~~|~~~50~~~~|~~~60
    Cdx-Ls   .PDKYRVVYTDYQRLELEKEFHYSRYITMNRKAELAKSLDLTERQIKIWFQNRRAKERKIN.K  P81193
    Cad-Bm  T.KDKYRVVYSDHQRLELEKEFHYSRYITIRRKAELAVSLGLSERQVKIWFQNRRAKERKQV.K  D16683
    Cad-Dm  T.KDKYRVVYTDFQRLELEKEYCTSRYITIRRKSELAQTLSLSERQVKIWFQNRRAKERTSN.K  P09085

    Cdx-Bf   .KDKYRVVYSDHQRLELEKEFYSNKYITIKRKVQLANELGLSERQVKIWFQNRRAKQRKMA.   Af052465
   Cdx1-Cc  T.KEKYRVVYTDHQRLELEKEFHFNRYITIRRKSELAVNLGLSERQVKIWFQNRRAKERKLI.K  X80668
   CdxC-Gg  T.KDKYRVVYTDHQRLELEKEFHYSRYITIRRKAELASSLGLSERQVKIWFQNRRAKERKIN.K  X80614
   Cdx4-Mm  T.KEKYRVVYTDHQRLELEKEFHCNRYITIRRKSELAVNLGLSERQVKIWFQNRRAKERKMI.K  Q07424
   Cdx1-Mm  T.KDKYRVVYTDHQRLELEKEFHYSRYITIRRKSELAANLGLTERQVKIWFQNRRAKERKVN.K  P18111
   Cdx2-Mm  T.KDKYRVVYTDHQRLELEKEFHFSRYITIRRKSELAATLGLSERQVKIWFQNRRAKERKIK.K  P43241

  cdx cons  T.KDKYRVVYTDHQRLELEKEFH5SRYITIRRK3ELA64LGLSERQVKIWFQNRRAKERK65.   (9)
  cdx spec  T.KD-Y-VV--DH-R-------H--R----R--S------G--------------A-E----.

Cnox-4                                                                 Cnid. Plat.
             1~~~|~~~10~~~~|~~~20~~~~|~~~30~~~~|~~~40~~~~|~~~50~~~~|~~~60
  Cnox4-Ed  P.AMRSRPCFSSHQTRELEKEFLVCQYVTRRRIELAFSLNLSEKQIKTWFQNRRVKERKQK.K  U41841
  Pnox6-Pn  E.KMRLRTSFHSHQLNILEKEFLVNMYLTRMRRIEMASALELGEKQIKIWFQNRRVRLKKIC.A  L41852

cnox4 cons   .2MR2R22F2SHQ222LEKEFLV22YLTR2RRIE2A22L2L2EKQIK2WFQNRRV222K22.   (2)
cnox4 spec   .-M-----F-SH---------LV--------------------K----------V------.
```

Lox* **Ann. Chor.**

```
             1~~~|~~~10~~~~|~~~20~~~~|~~~30~~~~|~~~40~~~~|~~~50~~~~|~~~60
   Lox3-Ht  D.NKRTRTAYSRSQLLELEKEFHFDKYISRPRRVELASSLNLTERHIKIWFQNRRMKWKKME.A  Y09622
 Lox3a-Him  D.NKRTRTAYSRAQLLELEKEFHYDKYISRPRRLELAASLNLTERHIKIWFQNRRMKWKKLE.S  S79684
 Lox3b-Him  D.NKRTRTAYSRAQLLELEKEFHYDKYISRPRRVELAASLNLTECHIKIWFQNRRMKWEKFE.S  S79685
 Lox3c-Him  D.NKRTRTAYSRAQLLELEKEFHYDKYISRPRRLELAASLNLTECHIKIWFQNRRMKWEKFE.S  S79687

   Xlox-Bf  ~.NKRTRTAYTRGQLLELEKEFHFNKYISRPRRIELAAMLNLTERHIKIWFQNRRMKWKKEQ.   AF052464
   Pdx1-Mm  E.NKRTRTAYTRAQLLELEKEFLFNKYISRPRRVELAVMLNLTERHIKIWFQNRRMKWKKEE.D  P52946

  Lox cons  D.NKRTRTAYSR3QLLELEKEFHy2KYISRPRRvELA42LNLTE2HIKIWFQNRRMKW2K4E.   (6)
  Lox spec  D.N--T--A--R--L--------HY-K----P---------N----H-----------W---E
```

PG-3 **Neme. Ann. Arth. Chor.**

```
             1~~~|~~~10~~~~|~~~20~~~~|~~~30~~~~|~~~40~~~~|~~~50~~~~|~~~60
    Hox3-Ls  P.PKRSRTAYTSAQLVELEKEFHFNRYLCRRRRIEMAALLNLSERQIKIWFQNRRMKYKKDQ.K  Y16571
  Hox3-Nvi   P.SKRARTAYNSAQLVELEKEFHFNRYLCRPRRIEMAALLSLSERQIKIWFQNRRMKYKKDQ.R  AF151665
    Zen-Tc   A.GKRARTAYTSAQLVELEREFHHGKYLSRPRRIQIAENLNLSERQIKIWFQNRRMKHKKEQ.   X97819
  Hox3-Cus   P.AKRARTAYTSAQLVELEKEFHFNRYLCRPRRIEMANLLNLSERQIKIWFQNRRMKYKKEQ.K  AJ005643

   Hox3-Bf   A.GKRARTAYTSAQLVELEKEFHFNRYLCRPRRVEMAAMLNLTERQIKIWFQNRRMKYKKEQ.K  P50901
  Hoxd3-Mm   A.SKRVRTAYTSAQLVELEKEFHFNRYLCRPRRIEMANLLNLTERQIKIWFQNRRMKYKKDQ.K  P09027
  Hoxb3-Mm   A.SKRARTAYTSAQLVELEKEFHFNRYLCRPRRVEMANLLNLSERQIKIWFQNRRMKYKKDQ.K  P09026
  Hoxa3-Mm   S.SKRARTAYTSAQLVELEKEFHFNRYLCRPRRIEMAALLNLTERQIKIWFQNRRMKYKKDQ.K  P02831

 PG-3 cons    .4KR4RTAYTSAQLVELEKEFHFNRYLCRPRRIEMA33LNLSERQIKIWFQNRRMKYKKdQ.K  (8)
 PG-3 spec    .------A--SA-LV------HF-R--C-P----M----N---------------Y--DQ.K
```

Ftz **Arth.**

```
             1~~~|~~~10~~~~|~~~20~~~~|~~~30~~~~|~~~40~~~~|~~~50~~~~|~~~60
     Ftz-Dh  D.SKRTRQTYSRYQTLELEKEFHFNRYITRRRRMDIAHALNLSERQIKIWFQNRRMKSKKDR.T  P48590
     Ftz-Dm  D.SKRTRQTYTRYQTLELEKEFHFNRYITRRRRIDIANALSLSERQIKIWFQNRRMKSKKDR.T  P02835
     Ftz-Tc  G.NKRTRQTYTRYQTLELEKEFHFNKYLTRRRRIEIAESLRLTERQIKIWFQNRRMKAKKDT.K  X97819

  Ftz cons    .2KRTRQTYTRYQTLELEKEFHFNRYITRRRR2dIA32L3LsERQIKIWFQNRRMK2KKD2.   (3)
```

PG-4 **Neme. Ann. Arth. Chor.**

```
             1~~~|~~~10~~~~|~~~20~~~~|~~~30~~~~|~~~40~~~~|~~~50~~~~|~~~60
    Hox4-Ls  E.SKRSRTAYTRHQILELEKEFHFNRYLTRRRRIEIAHALDLSERQIKIWFQNRRMKWKKEH.K  P81192
    Dfd-Nvi  D.SKRTRTAYTRHQVLELEKEFHFNRYLTRRRRIEIAHALCLTERQIKIWFQNRRMKWKKEN.K  Af151666
     Dfd-Af  E.PKRQRTAYTRHQILELEKEFHFNRYLTRRRRIEIAHSLCLSERQIKIWFQNRRMKWKKDN.K  X70078
     Dfd-Dm  ~.PKRQRTAYTRQQILELEKEFHYNRYLTRRRRIEIAHTLVLSERQIKIWFQNRRMKWKKDN.K  P07548

    Hox4-Bf  ~.TKRSRTAYTRQQVLELEKEFHFNRYLTRRRRIEIAHSLGLTERQIKIWFQNRRMKWKKDN.R  Z35144
  Chox1-4Gg  E.PKRSRTAYTRQQVLELEKEFHFNRYLTRRRRIEIAHTLCLSERQIKIWFQNRRMKWKKDH.K  P17277
   Hoxb4-Mm  E.PKRSRTAYTRQQVLELEKEFHYNRYLTRRRRVEIAHALCLSERQIKIWFQNRRMKWKKDN.K  M36654
   Hoxc4-Mm  E.PKRSRTAYTRQQVLELEKEFHYNRYLTRRRRIEIAHSLCLSERQIKIWFQNRRMKWKKDN.K  Q08624
   Hoxa4-Mm  ~.PKRSRTAYTRQQVLELEKEFHFNRYLTRRRRIEIAHTLCLTERQVKIWFQNRRMKWKKDN.K  M27432
   Hoxd4-Mm  ~.PKRSRTAYTRQQVLELEKEFHFNRYLTRRRRIEIAHTLCLSERQIKIWFQNRRMKWKKDN.K  P10628

 PG-4 cons    .3KR3RTAYTR2QvLELEKEFHFNRYLTRRRRIEIAH3L4LsERQIKIWFQNRRMKWKKdN.K  (10)
      prot    .2--3------2-i----------------------3-3--------------------.   (4)
      deut    .2---------Q-V----------------------2-2--------------------.   (6)
```

PG-5 **Neme. Ann. Arth. Echi. Chor.**

```
     Scr-Dm  Q.TKRQRTSYTRYQTLELEKEFHFNRYLTRRRRIEIAHALCLTERQIKIWFQNRRMKWKKEH.K  P09077
    Scr-Nvi  E.SKRTRTSYTRHQTLELEKEFHFNRYLTRRRRIEIAHALNLTERQIKIWFQNRRMKWKKEH.K  Af151667
   Lox1-Him  D.NKRLRTSYSRQQTLELEKEFHFNRYLTKRRRVEIAQMLKLSERQIKIWFQNRRMKWKKEN.S  L07298
   Lox20-Ht  D.NKRTRTSYTRHQTLELEKEFHFNRYLSRRRRIEIAHVLNLSERQIKIWFQNRRMKWKKDH.P  Af006638
    Lox6-Hr  A.DKRARTSYSRYQTLELEKEFHFNRYLNGRRRIEIAHSLGLTERQIKIWFQNRRMKWKKDN.R  Af004386

Hehbox9-He   P.SKRSRTAYTRYQTLELEKEFHFNRYLTRRRRIEIAHALGLTERQIKIWFQNRRMKWKKEH.N  U31563
    Hox5-Ci  P.SKRTRTAYTRYQTLELEKEFHYNRYLTRRRRIEVAHTLCLTERQIKIWFQNRRMKWKKEN.K  AJ002028
    Hox5-Bf  D.NKRTRTAYTRYQTLELEKEFHFNRYLTRRRRIEIAHALCLTERQIKIWFQNRRMKWKKEN.K  Z35145
   Hoxa5-Rn  E.GKRARTCYTRYQTLELEKEFHFNRYLTRRRRIEIAHALCLSERQIKIWFQNRRMKWKKDK.   P52949

 PG-5 cons    .5KR5RT3YTR3QTLELEKEFHFNRYLTRRRRIEIAH5L4LTERQIKIWFQNRRMKWKKe3.   (9)
      prot    .4--5--S---3------------------------4-4--------------------2.   (6)
      deut    .3--3--2---Y------------------------2-C--------------------2.   (3)
```

PG-6 **Plat. Neme. Ann. Arth. Echi. Chor.**

```
             1~~~|~~~10~~~~|~~~20~~~~|~~~30~~~~|~~~40~~~~|~~~50~~~~|~~~60
Dutarh4-Gt   E.YKRCRQAYSRQQTLELEKEFYYNQYLTRRRRIEIANSVCLSERQIKIWFQNRRMKYKKDV.   Z34089
Dutarh3-Gt   D.HKRSRQTYTRYQTLELEKEFHFNKYLTRRRRIEIAHGLSLTERQIKIWFQNRRMKWKKDH.N  Z34091
    Hox6-Ls  E.QKRTRQTYTRYQTLELEKEFHFNRYLTRRRRIEIAHALGLTERQIKIWFQNRRMKWKKEN.N  Y16572
   Lox5-Nvi  E.QKRTRQTYTRYQTLELEKEFHYNRYLTRRRRIEIAHALGLTERQIKIWFQNRRMKWKKEN.N  Af151671
    Lox5-Hr  D.QKRTRQTYTRYQTLELEKEFYSNRYLTRRRRIEIAHSLALSERQIKIWFQNRRMKWKKEN.N  Af004387
    Antp-Dm  E.RKRGRQTYTRYQTLELEKEFHFNRYLTRRRRIEIAHALCLTERQIKIWFQNRRMKWKKEN.K  P02833

    Hbx3-Tg  D.GKRGRQTYTRQQTLELEKEFHFSRYVTRRRRFEIAQSLGLSERQIKIWFQNRRMKWKREH.G  P10178
Pahbox6-Pa   D.RKRGRQTYTRAQTLELEKEFHYNRYLTRKRRIEIAQAVCLSERQIKIWFQNRRMKWKKER.V  X54494
    Hox6-Bf  E.KKRGRQTYTRYQTLELEKEFHFNKYLTRKRRIEIAHLLGLTERQIKIWFQNRRMKWKKEN.K  Z35146
   Hoxy6-Br  ~.~RRGRQIYSRYQTLELEKEFHYNRYLTRRRRIEIANTLCLSERQIKIWFQNRRMKWKKES.N  Y14537
   Hoxb6-Gg  P.ARRGRQTYTRYQTLELEKEFHFNRYLTRRRRIEIAHSLCLTERQIKIWFQNRRMKWKKEN.K  P14839
   Hoxa6-Mm  H.GRRGRQTYTRYQTLELEKEFHFNRYLTRRRRIEIANALCLTERQIKIWFQNRRMKWKKEN.K  P09092
   Hoxb6-Mm  S.GRRGRQTYTRYQTLELEKEFHYNRYLTRRRRIEIAHALCLTERQIKIWFQNRRMKWKKES.K  P09023
   Hoxc6-Mm  D.RRRGRQIYSRYQTLELEKEFHFNKYLTRRRRIEIANALCLTERQIKIWFQNRRMKWKKEL.Q  P10629

 PG-6 cons    .7kR4RQTYTR3QTLELEKEF23NRYLTRRRRIEIA35L4LtERQIKIWFQNRRMKWKKe5.   (14)
      prot   e.3K-4------2---------23-r-----------H3-4------------------e3.N  (6)
      deut    .4r-G------3---------Hf-------------34-2------------------E4.   (8)
```

```
PG-7                                    Plat. Neme. Ann. Arth. Echi. Chor.
   Hox7-Ls  D.RKRGRQTYTRYQTLELEKEFHFNKYLTRRRRIEIAHALCLTERQIKIWFQNRRMKWKKEN.K  Y16573
    Ubx Dm  L.RRRGRQTYTRYQTLELEKEFHTNHYLTRRRRIEMAHALCLTERQIKIWFQNRRMKLKKEI.Q  P02834

   Hox7-Bf  E.RKRGRQTYTRYQTLELEKEFHFNKYLTRRRRIEIAHALCLTERQIKIWFQNRRMKWKKEN.K  Z35147
  Hoxa7-Mm  D.RKRGRQTYTRYQTLELEKEFHFNRYLTRRRRIEIAHALCLTERQIKIWFQNRRMKWKKEH.K  P02830
  Hoxb7-Mm  D.RKRGRQTYTRYQTLELEKEFHYNRYLTRRRRIEIAHTLCLTERQIKIWFQNRRMKWKKEN.K  P09024

 PG-7 cons   .RKRGRQTYTRYQTLELEKEFHFNrYLTRRRRIEIAHALCLTERQIKIWFQNRRMKWKKEN.K  (5)


PG-8                                    Plat. Neme. Ann. Arth. Echi. Chor.
 Pnox1a-Pn  Q.KRRGRQTYSRHQTLELEKEFQFNHYLTRRRRIEIAHNLCLSERQIKIWFQNRRMKLKKER.Q  L41845
 Dthoxf-Gt  Q.KRRGRQTYSRHQTLELEKEFQFNHYLTRRRRIEIAHNLCLSERQIKIWFQNRRMKLKKER.Q  X95412
  Lox2-Him  Q.RRRGRQTYTRYQTLELEKEFKFNRYLTRRRRIELSHTLYLTERQIKIWFQNRRMKEKKEV.Q  P21523
  Lox4-Him  Q.RRRGRQTYSRYQTLELEKEFQFNRYLTRKRRIEIAHCLCLTERQIKIWFQNRRMKVKKEK.Q  S79240
  Lox4-Nvi  E.EERGRQTYSRYQTLELEKEFRFNHYLTRKRRMEIAHVLCLTEHQIKIWFQNRRMKLRKER.L  Af151669
   abdA-Dm  P.RRRGRQTYTRFQTLELEKEFHFNRYLTRRRRIEIAHALCLTERQIKIWFQNRRMKWKKEL.R  P29555

   Hbx1-Tg  G.RKRCRQTYTRYQTLELEKEFHFNRYLTRRRRIELSHLLGLTERQIKIWFQNRRMKYKKES.K  P09080
   Hox8-Bf  E.RRRGRQTYSRYQTLELEKEFHFNKYLTRRRRIEIAHALGLTERQIKIWFQNRRMKLKKEA.A  Z35148

 PG-8 cons   .rkRGRQTYtR3QTLELEKEF4FN3YLTRrRRIEIAH6L3LtERQIKIWFQNRRMK5KKE7.   (8)
      prot   .-R---------3---------4--2------------5-C----------------l---5.Q  (6)
      deut   .R---------Y---------H--k----------2-2-G----------------2---3.   (2)

Orphans                                 Plat. Neme. Arth. Uro.
            1~~~|~~~10~~~~|~~~20~~~~|~~~30~~~~|~~~40~~~~|~~~50~~~~|~~~60
   Hox3-Sp  N.GRKKRVPYTKYQLLELEKEFHYNQYLSRERRQEVAKAVSLSDRQVKIWFQNRRMKWKKEK.K  U05600
    Zen-Dm  K.SKRSRTAFSSLQLIELEREFHLNKYLARTRRIEISQRLALTERQVKIWFQNRRMKLKKST.N  P09090

 Dthoxg-Gt  T.SKRCRSAYTNNQLVELEKEFHYNNYLARGRRTELSKQLLLTERQVKIWFQNRRMKQKKEK.K  X95411
Dutarh1-Gt  R.RVSAAIAYTNNQLVELEKEFHYNNYLARGRRTELSKQLLLTERQVKIWFQNRRMKQKKEK.K  Z34090
  Ceh13-Ce  V.IDENGQNYTRAQTHELEKEFHFNKTATRVNRTEITAALSNNKRQVAIWFQNRRMKQKKEM.K  X17077
   Mab5-Ce  E.SKRTRQTYSRSQTLELEKEFHYHKYLTRKRRQEISETLHLTERQVKIWFQNRRMKHKKEA.K  P10038

  Ceh11-Ce  V.SKKGRQTYQRYQTSVLEAKFQQSSYVSKKQREELRLQTQLTDRQIKIWFQNRRMKAKKEK.Q  P17486

  Lin39-Ce  G.EKRQRTAYTRNQVLELEKEFHTHKYLTRKRRIEVAHSLMLTERQVKIWFQNRRMKHKKEN.   L19639
 Dthoxa-Gt  D.SKRNRTAYTRQQILELEKEFHFNKYLTRKRRNEIAQSLQLSESQVKIWFQNRRMKWKKDH.H  X95414
 Dthoxd-Gt  E.YKRCRQAYSRQQTLELEKEFYYNQYLTRRRRIEIANSVCLSERQIKIWFQNRRMKYKKDV.V  X95417
```

Phylum code

Ann.:	Annelida
Arth.:	Arthropoda
Chor.:	Chordata
Cnid.:	Cnidaria
Echi.:	Echinodermata
Nema.:	Nematoda
Neme.:	Nemerta
Plac.:	Placozoa
Plat.:	Platyhelminthes
Pori.:	Porifera
Pria.:	Priapulides
Uro.:	Urochordata
Vert.:	Vertebrata

Species Code

Porifera: Ef, Em: *Ephydatia fluviatilis*, *E. muelleri*;
Placozoa: Ta: *Trichoplax adhaerens*;
Cnidaria: Af, Am: *Acropora formosa, A. millipora*, Nve: *Nematostella vectensis* (anth.); Cx: *Cassiopeia xamachana* (scyph.); Cv: *Chlorohydra viridissima*; Ed: *Eleutheria dichotoma*; Hl, Hm, Hv: *Hydra littoralis*; *H. magnipapillata, H. vulgaris*; Hys: *Hydractinia symbiolongicarpus*, Pc: *Podocorynae carnea* (hydroz.);
Platyhelminthes: Gt: *Girardia tigrina,* Pn: *Polycelis nigra* (planaria); Eg: *Echinococcus granulosa*;
Nemertea: Ls: *Linneus sanguineus*;
Annelida: Him: *Hirudo medicinalis*, Hsp, Ht: *Helobdella sp., H. triserialis* (leech), Nvi: *Nereis virens*;
Nematoda: Ce: *C. elegans*;
Arthropods: Afr: *Artemia franciscana* (shrimp); *Al: Archegozetes longisetosus* (acari); Bm: *Bombyx mori* (silkworm); Da, Dh, Dm: *Drosophila ananassae*; *D. hydei*; *D. melanogaster* (flies), Jc: *Junonia coenia* (butterfly), *Sa: Schistocerca americana*; Tc: *Tribolium castaneum* (beetle);
Echinodermata: He: *Heliocidaris erythrogramma*, Hp: *Hemicentrotus pulcherrimus*, Pa: *Parechinus angulosus*, Pl: *Purpuratus lividus*, Sp: *Strongylocentrotus purpuratus*, Tg: *Tripneustes gratilla* (sea urchin);
Urochordates: Ci: *Ciona intestinalis,* Hr: *Halocynthia roretzi*, Sc: *Styela clava*; *St: Styela plicata* (ascidia);
Cephalochordates: Bf: *Branchiostoma floridae* (amphioxus);
Vertebrates: Br: *Brachydanio rerio*, Cc: *Cyprinus carpio*, Ol: *Oryzia latipes* (fish), Ame: *Ambystoma mexicanus* (axolotl), Xl: *Xenopus laevis*, Gg: *Gallus gallus* (chick), Ma: *Mesocricetus auratus* (hamster), Mm: *Mus musculus*; Rn: *Rattus norvegicus*; Hs: *Homo sapiens*.