Introduction to Jalview

a guided tutorial

EMBL-EBI, Jan 21st, 2013.

Jim Procter & Geoff Barton
College of Life Sciences
University of Dundee, UK.
jprocter@dundee.ac.uk

Day 2 - Tuesday

09:00-10:30 Advanced Jalview and Jalview Hackathon

10:30-11:00 Coffee/Tea Break

11:00-12:30 Advanced Jalview and Jalview Hackathon continued

12:30-13:30 Lunch

13:30-15:00 Jalview Clinic and Jalview Hackathon
15:00-15:30 Coffee/Tea Break
15:30-15:50 Jalview Clinic and Jalview Hackathon continued

15:50-16:15 Wrap-up

16:15 Departure

Course materials

Everything is (or will) be online at

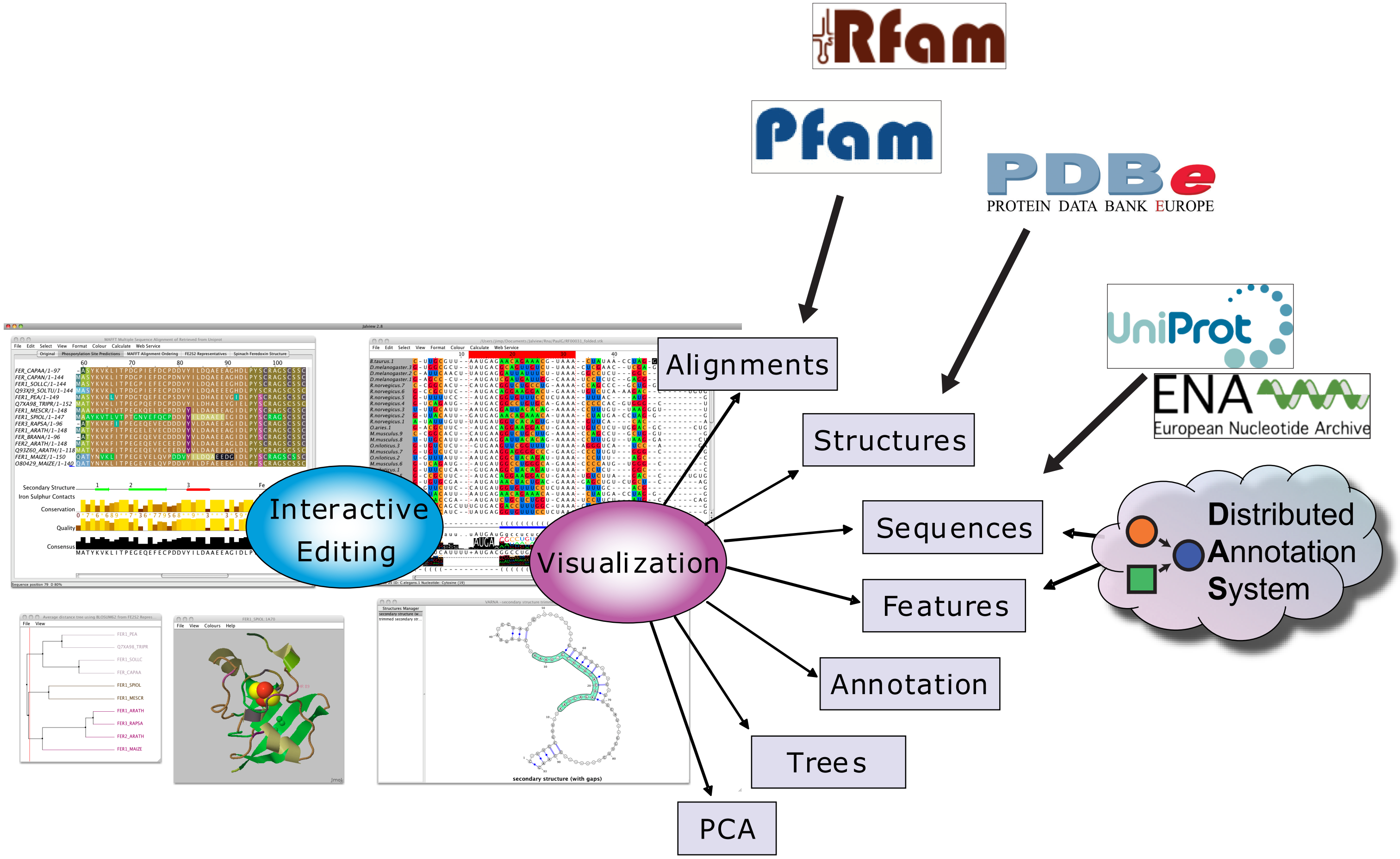
http://www.jalview.org/tutorial/training-materials/2013/EBI/Jan/

- **These slides**
- **Jalview v2.8 Tutorial (v1.4.1)**

- **A copy of the tutorial is also on Penelope**

Getting and working with sequence features and annotation

- Sequence Databases
- Sequence feature sources
  - DAS Sequence feature retrieval
  - GFF and Jalview feature files
- Visualizing features
  - Highlighting annotated regions
  - Shading and reordering based on scores and labels



Rfam

Pfam

PDBe
PROTEIN DATA BANK EUROPE

UniProt

ENA
European Nucleotide Archive

Interactive Editing

Visualization

Alignments

Structures

Sequences

Features

Annotation

Trees

PCA

Distributed **A**nnotation **S**ystem

Jalview and Sequence Databases
Sec 2.9.1 Ex. 29

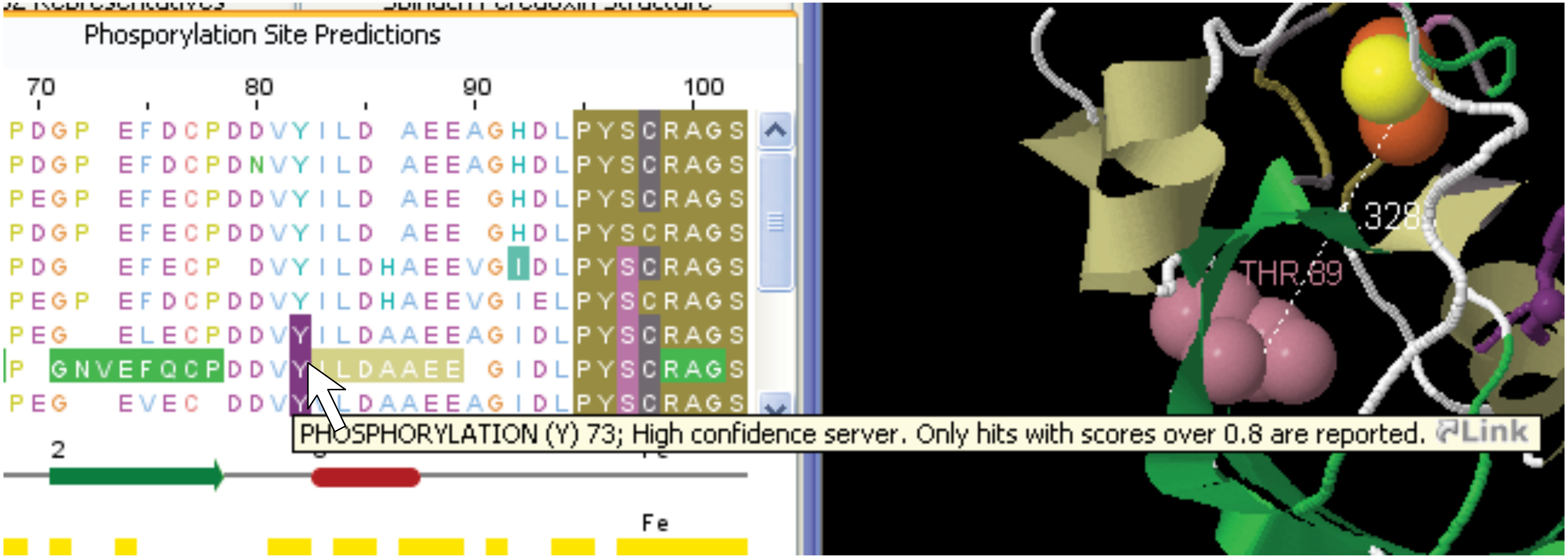
- Can retrieve new sequences or match against existing records using IDs
- Task
  - Recover the Uniprot annotation for the ferredoxin sequences using their IDs
    - **make sure you are using the latest 2.8 Build**

  - Verify retrieval by examining sequence annotation

Jalview and Sequence Databases
Sec 2.9.1 Ex. 29

- Can retrieve new sequences or match against existing records using IDs
- Task
  - Recover the Uniprot annotation for the ferredoxin sequences using their IDs
  - Verify retrieval by examining sequence annotation
- Question
  - What happens if only a subsequence is present in the alignment ?
  - Does database annotation get shared between alignments ?

Sequence Features



PHOSPHORYLATION (Y) 73; High confidence server. Only hits with scores over 0.8 are reported.

Sequence Features Section 2.8.1-3 & Ex 27

  - Annotate the whole or part of a sequence
  - Database refs are special case.
- Tasks
  - Visualise, create, modify, import and export features.

Sequence Features
Section 2.6.1-3 & Ex 27

  - Annotate the whole or part of a sequence
  - Database refs are special case.
- Tasks
  - Visualise, create, modify, import and export features.
- Questions
  - What are the different types of file formats available for import and export
  - What services allow you to discover annotation for sequence ?

The Distributed Annotation System
Section 2.9.2, Exercise 30

  - Web servers that jalview can use to discover annotation for a sequence
- Task
  - Browse available DAS sources for protein sequences
  - Retrieve annotation for the ferredoxin alignment.

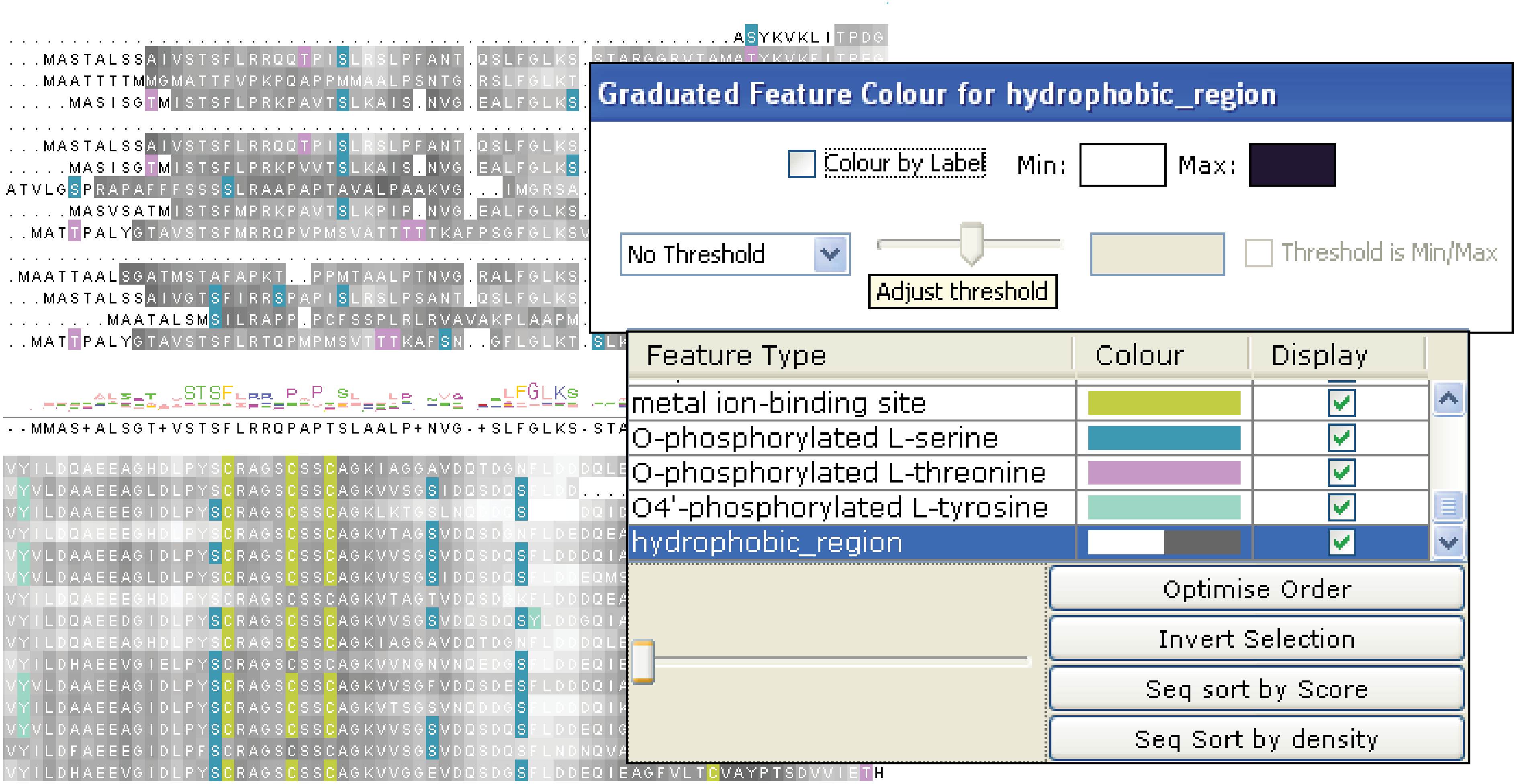
The Distributed Annotation System
Section 2.9.2, Exercise 30

  - Web servers that jalview can use to discover annotation for a sequence
- Task
  - Browse available DAS sources for protein sequences
  - Retrieve annotation for the ferredoxin alignment.
- Question
  - What does the ‘optimise order’ button do ?

Working with sequence features Ex 31 and 32 (Sec 2.9.4)

- Task
  - Shading features using labels and scores
  - Sorting alignment using feature scores

Shading, thresholding, colour by label.



Graduated Feature Colour for hydrophobic_region

☐ Colour by Label   Min: ☐   Max: ■

No Threshold ▾   Adjust threshold   ☐ Threshold is Min/Max

Feature Type	Colour	Display
metal ion-binding site	■	☑
O-phosphorylated L-serine	■	☑
O-phosphorylated L-threonine	■	☑
O4'-phosphorylated L-tyrosine	■	☑
hydrophobic_region	■	☑

Working with sequence features Ex 31 and 32 (Sec 2.9.4)

- Task
  - Shading features using labels and scores
  - Sorting alignment using feature scores
- Questions
  - What types of features are best displayed with a 'label' colourscheme ?
  - How would you display only the highest or lowest scoring features ?

Protein secondary structure prediction

Section 2.6 onwards

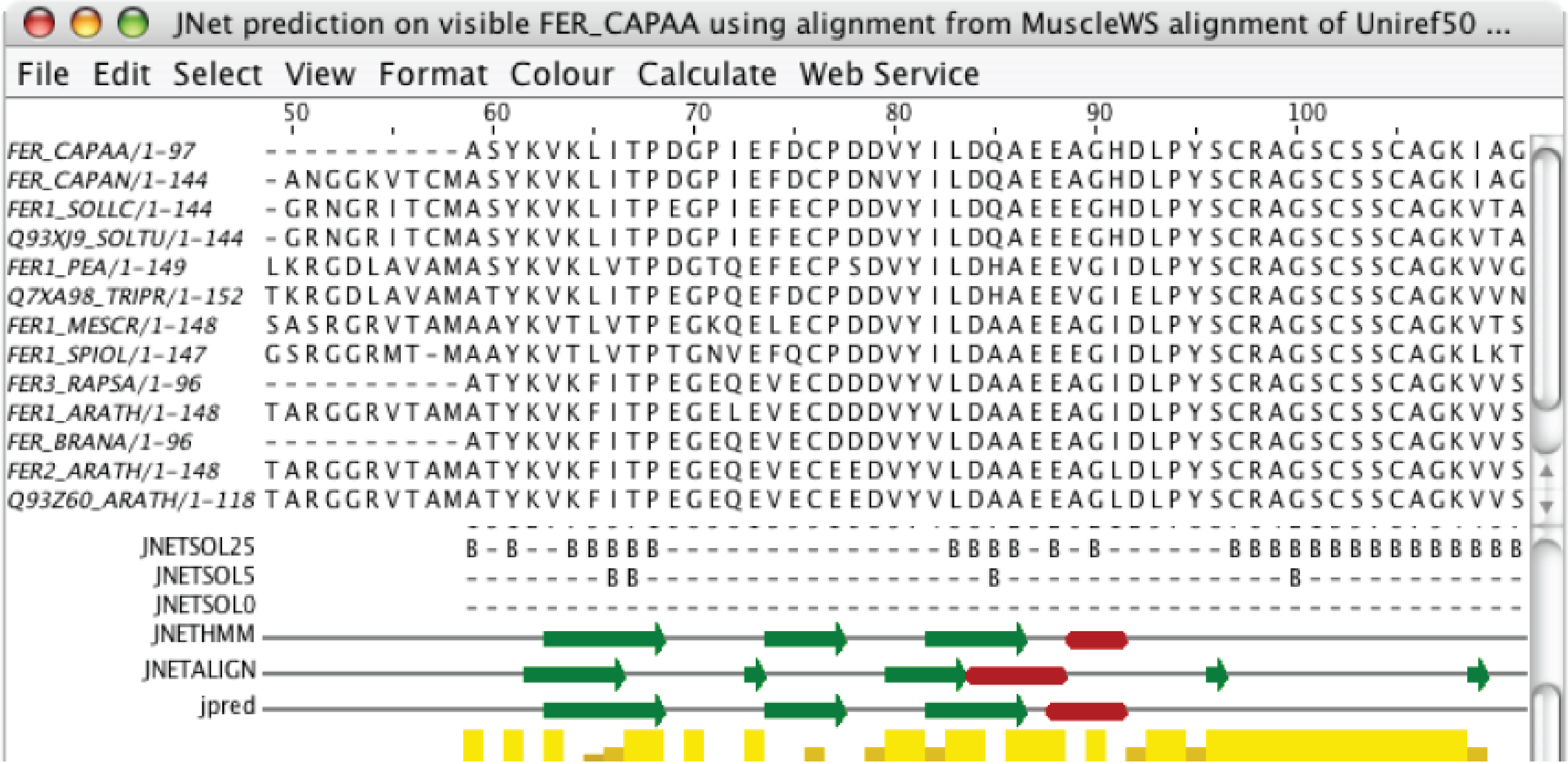
in the Manual

Page 64

Protein Secondary Structure Prediction
Sect 2.6

- Jalview interfaces with the Jpred protein secondary structure predictor
- Prediction is based on
  - Neural net which can recognise helical, coil or beta strand using amino acid patterns
  - Amino acid profile for a sequence
    - Multiple sequence alignment
    - Profile from sequence database search
      - **P**osition **S**pecific **S**ubstitution **M**atrix

Protein Secondary Structure Prediction
Sect 2.6



```
                      50        60        70        80        90        100
FER_CAPAA/1-97      ----------ASYKVKLITPDGPIEFDCPDDVYILDQAEEAGHDLPYSCRAGSCSSCAGKIAG
FER_CAPAN/1-144     -ANGGKVTCMASYKVKLITPDGPIEFDCPDNVYILDQAEEAGHDLPYSCRAGSCSSCAGKIAG
FER1_SOLLC/1-144    -GRNGRITCMASYKVKLITPEGPIEFECPDDVYILDQAEEEGHDLPYSCRAGSCSSCAGKVTA
Q93XJ9_SOLTU/1-144  -GRNGRITCMASYKVKLITPDGPIEFECPDDVYILDQAEEEGHDLPYSCRAGSCSSCAGKVTA
FER1_PEA/1-149      LKRGDLAVAMASYKVKLVTPDGTQEFECPSDVYILDHAEEVGIDLPYSCRAGSCSSCAGKVVG
Q7XA98_TRIPR/1-152  TKRGDLAVAMATYKVKLITPEGPQEFDCPDDVYILDHAEEVGIELPYSCRAGSCSSCAGKVVN
FER1_MESCR/1-148    SASRGRVTAMAAYKVTLVTPEGKQELECPDDVYILDAAEEAGIDLPYSCRAGSCSSCAGKVTS
FER1_SPIOL/1-147    GSRGGRMT-MAAYKVTLVTPTGNVEFQCPDDVYILDAAEEEGIDLPYSCRAGSCSSCAGKLKT
FER3_RAPSA/1-96     ----------ATYKVKFITPEGEQEVECDDDVYVLDAAEEAGIDLPYSCRAGSCSSCAGKVVS
FER1_ARATH/1-148    TARGGRVTAMATYKVKFITPEGELEVECDDDVYVLDAAEEAGIDLPYSCRAGSCSSCAGKVVS
FER_BRANA/1-96      ----------ATYKVKFITPEGEQEVECDDDVYVLDAAEEAGIDLPYSCRAGSCSSCAGKVVS
FER2_ARATH/1-148    TARGGRVTAMATYKVKFITPEGEQEVECEEDVYVLDAAEEAGLDLPYSCRAGSCSSCAGKVVS
Q93Z60_ARATH/1-118  TARGGRVTAMATYKVKFITPEGEQEVECEEDVYVLDAAEEAGLDLPYSCRAGSCSSCAGKVVS

         JNETSOL25            B-B--BBBBB--------------BBBB-B-B------BBBBBBBBBBBBBBB
         JNETSOL5             -------BB-----------------B--------------B-----------
         JNETSOL0             -----------------------------------------------------
         JNETHMM
         JNETALIGN
         jpred
```

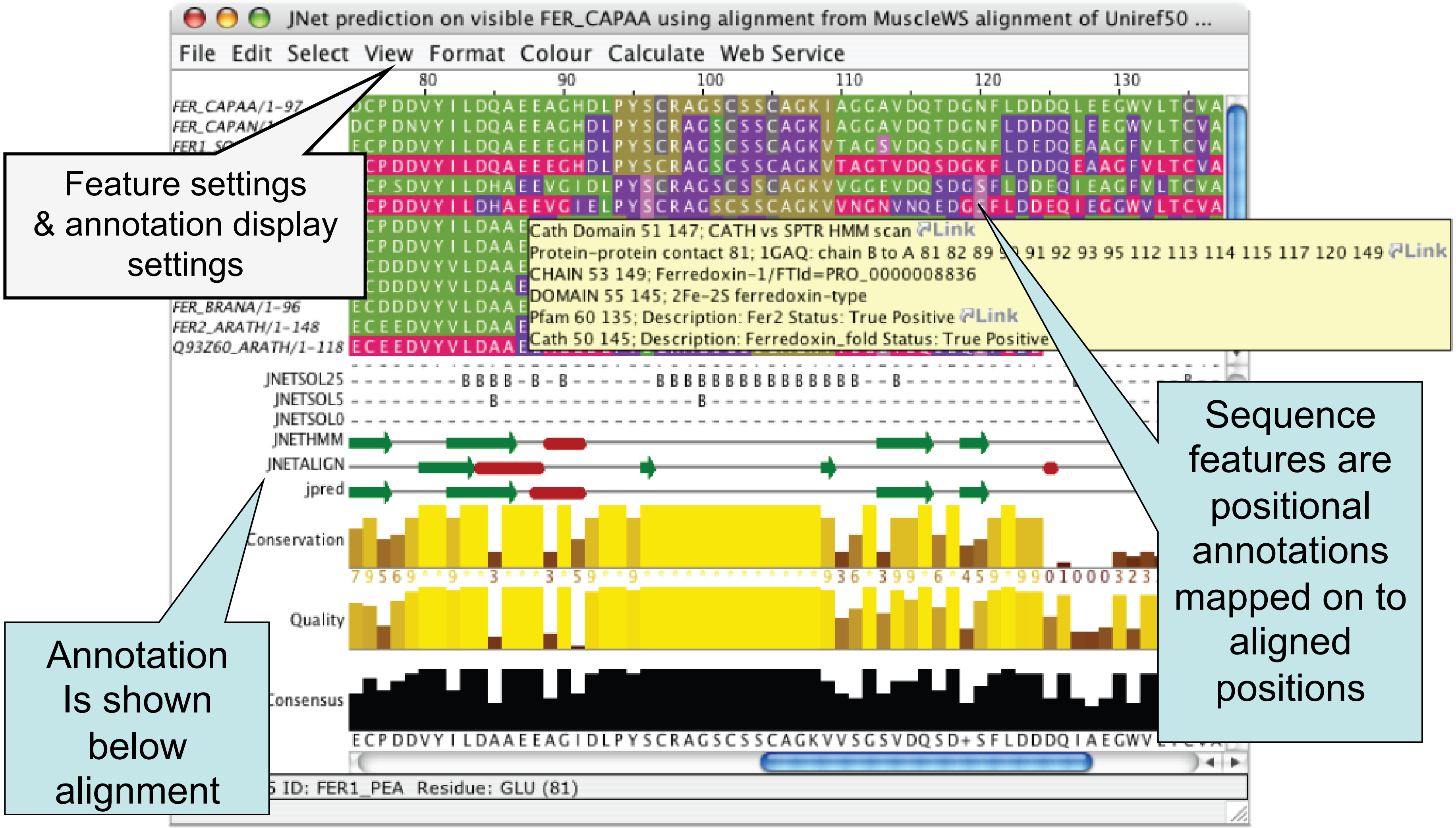
Exercise 26

- Tasks
  - Perform a variety of Jnet predictions
    - Note the effect of hidden regions
    - Learn about sequence associated annotation
  - Save your results for the next exercise
- Questions
  - …

Exercise 26

- Tasks
  - – Perform a variety of Jnet predictions
    - Note the effect of hidden regions
    - Learn about sequence associated annotation
  - – Save your results for the next exercise
- Questions
  - – What other data does Jnet provide ?
  - – Which is better – a PSI blast prediction or an MSA based prediction ?
  - – What happens when you have hidden regions ?

Alignment Annotation and sequence features.



JNet prediction on visible FER_CAPAA using alignment from MuscleWS alignment of Uniref50 ...

Cath Domain 51 147; CATH vs SPTR HMM scan Link
Protein-protein contact 81; 1GAQ: chain B to A 81 82 89 9 91 92 93 95 112 113 114 115 117 120 149 Link
CHAIN 53 149; Ferredoxin-1/FTId=PRO_0000008836
DOMAIN 55 145; 2Fe-2S ferredoxin-type
Pfam 60 135; Description: Fer2 Status: True Positive Link
Cath 50 145; Description: Ferredoxin_fold Status: True Positive

ID: FER1_PEA Residue: GLU (81)

Feature settings & annotation display settings

Sequence features are positional annotations mapped on to aligned positions

Annotation Is shown below alignment

Creating, editing and using annotation.
Exercise 28 (Sec. 2.8.4)

- Tasks
  - Manually annotate some columns using the interactive editing functions
  - Learn about jalview annotation files
    - How to change the appearance of quantitative data.
    - Understand how to create sequence associated annotation
- Questions
  - What other things can be defined in jalview annotation files ?

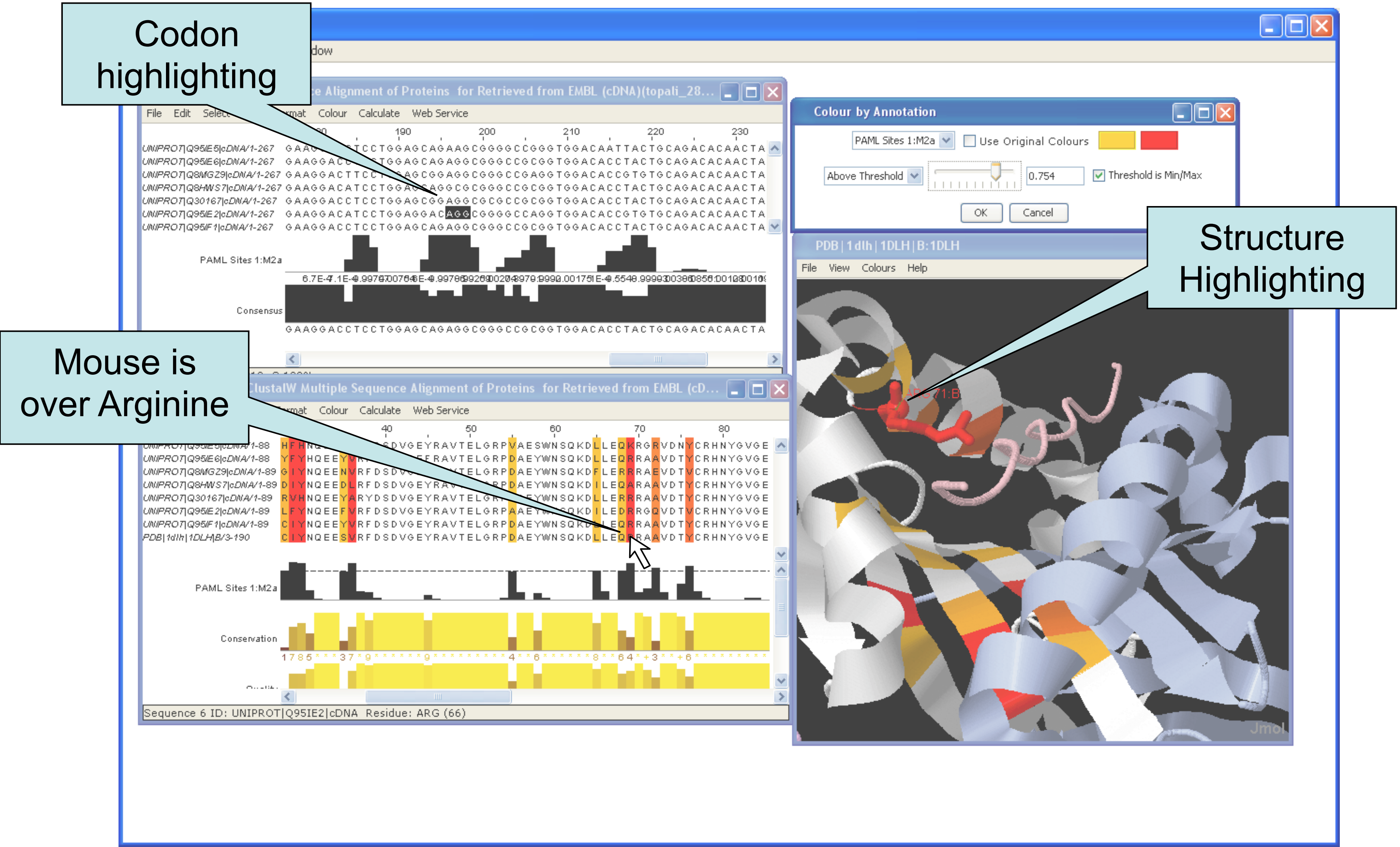
- Session 3
  - Sequence DB refs and Sequence Features
  - Protein secondary structure prediction
  - Alignment annotation

- Session 4
  - Working with structures
    - Viewing 3D structures
    - Viewing RNA

Protein Structure and Jalview
Section 2.1

- Jalview includes the Jmol molecular graphics viewer
  - Structures can be coloured by their aligned sequences
  - Position of mouse highlighted in sequence or structure

Structure shaded by sequence



Codon highlighting

Mouse is over Arginine

Structure Highlighting

Colour by Annotation

PAML Sites 1:M2a ☐ Use Original Colours 🟨 🟥

Above Threshold 0.754 ☑ Threshold is Min/Max

OK Cancel

Sequence 6 ID: UNIPROT|Q95IE2|cDNA Residue: ARG (66)

Associating structures with sequences

- Local PDB file
  - – Attach PDB file to sequence manually
  - – drag and drop to match files to sequences by ID
- Structures in the PDB database
  - – Provide PDB id (and chain) for sequence
  - – Discover references *via* sequence database

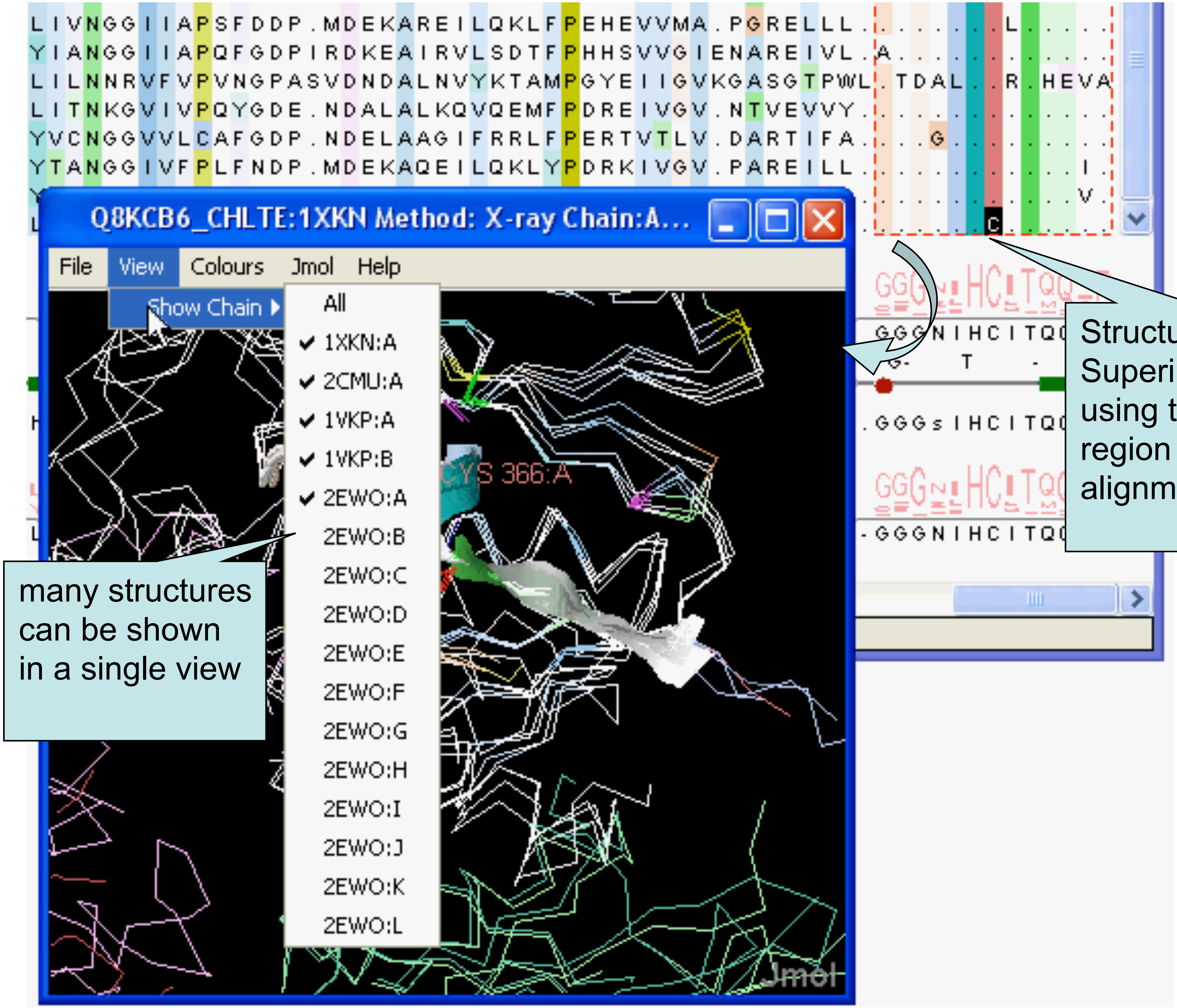
Protein Structures in Jalview
Sec 2.1. Exercise 14

- Task
  - Discover PDB structures for ferredoxin sequence(s)
    - **Note use**

      Fetch Database Refs->UNIPROT->Uniprot
  - Save and load structures and manipulate colouring

Protein Structures in Jalview
Sec 2.1. Exercise 14

- Task
  - Discover PDB structures for ferredoxin sequence(s)
  - Save and load structures and manipulate colouring
- Questions
  - How does Jalview match up sequence data to structural data



Structures can be Superimposed using the visible region of alignment

many structures can be shown in a single view

Superposing Structures using Alignments
Sec 2.1.4 – Exercise 15

- Task
  - Align structures using the ferredoxin alignment

If ‘View all N structures’ doesn’t align structures:

- **Use Jmol->Align menu**

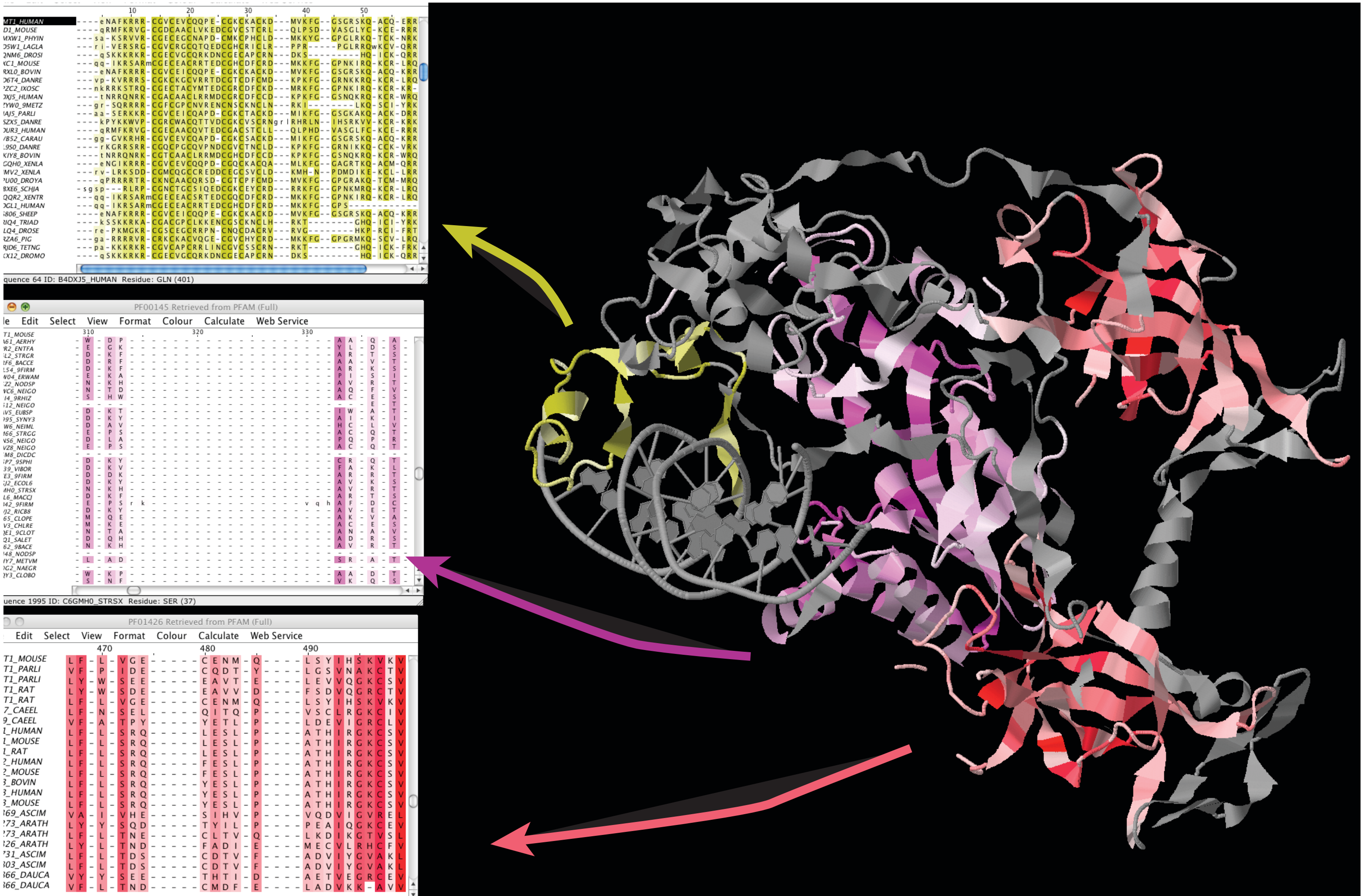
  - Experiment with views to control what part of the alignment is used to superimpose the structures

Superposing Structures using Alignments
Sec 2.1.4 – Exercise 15

- Task
  - Align structures using the ferredoxin alignment
  - Experiment with views to control what part of the alignment is used to superimpose the structures
- Questions
  - What colourscheme would highlight the conserved parts of the structures ?
  - Which view gave the ‘best’ structure superposition ?
    - How did you decide this ?

Colouring structures using many multiple alignments

Sect 2.1.5. Exercise 16



DNA and Protein in Jalview

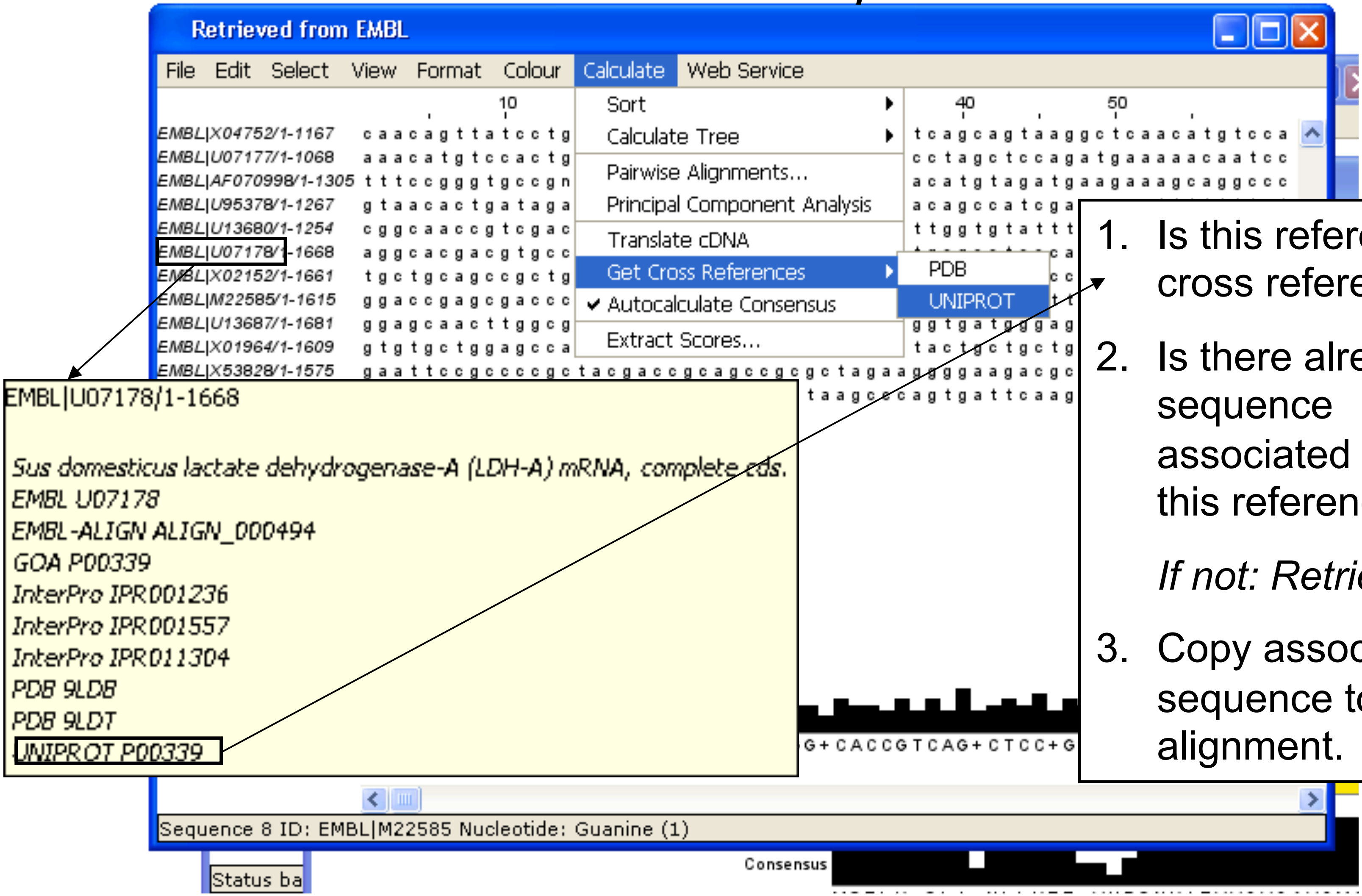
- Discussed in Section 2.10 of manual
- From DNA to Protein
  - Calculations => Translate cDNA
  - View protein annotation on exons using EMBL records
- From protein to DNA
  - Recover DNA for proteins using EMBL cross references

Protein Feature visualization on DNA
Section 2.10, exercise 33

- Task
  - Retrieve a DNA contig and visualize features from UNIPROT at their coding positions.
- Question
  - What fields in an EMBL sequence record can Jalview use ?

Semantic Processing: Database Reference Tracing

'get me the sequences from database ***blah*** *for the selected sequences'*



EMBL|U07178/1-1668

Sus domesticus lactate dehydrogenase-A (LDH-A) mRNA, complete cds.
EMBL U07178
EMBL-ALIGN ALIGN_000494
GOA P00339
InterPro IPR001236
InterPro IPR001557
InterPro IPR011304
PDB 9LDB
PDB 9LDT
UNIPROT P00339

Sequence 8 ID: EMBL|M22585 Nucleotide: Guanine (1)

1. Is this reference a cross reference ?
2. Is there already a sequence associated with this reference ?

   If not: Retrieve it.
3. Copy associated sequence to new alignment.

Jalview 2.8 and RNA 2^{nd}-ary Structure

Sequence position 156 CG 45%; GC 18%; UA 15%; AU 9%; UG 7%; UU 0%; GU 0%; A(

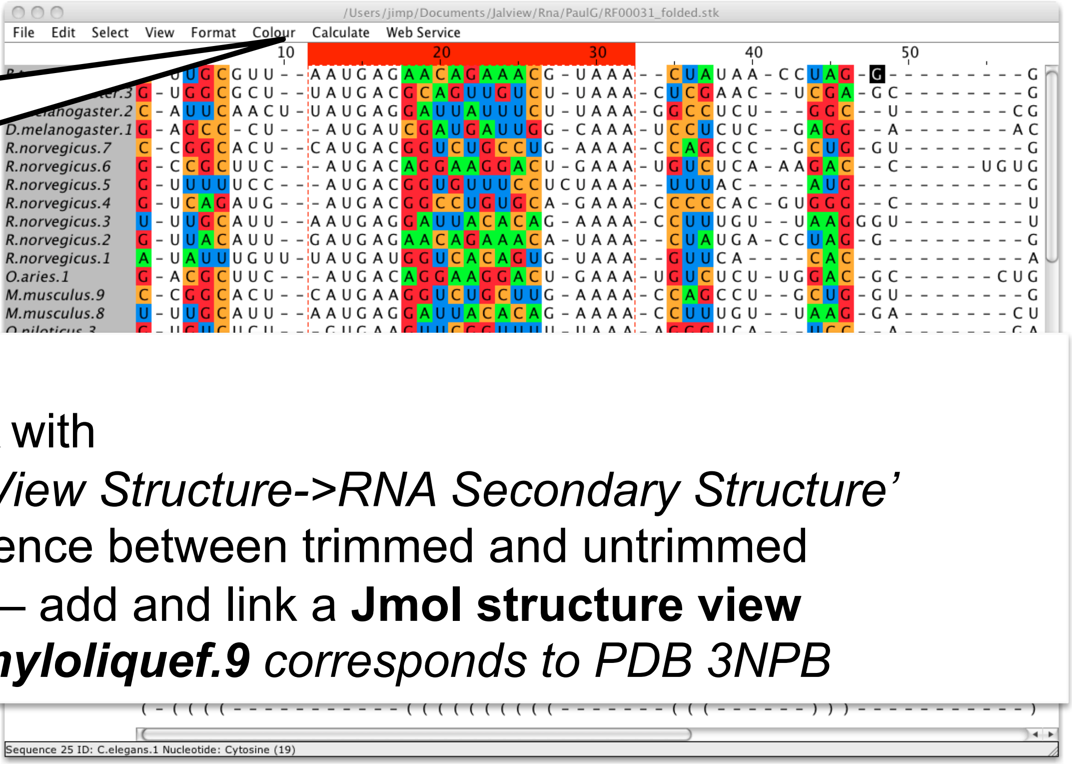
Structure Consensus Logo: Shows base pair distribution at each paired position in a given RNA secondary structure.

Try it now:

1. **Import RF00162 from *Rfam (Full)***
2. Select 'Colour by RNA Helices'

RALEE style colouring highlights distinct stems and helices

Selections and mouse positions shared between alignment view and VARNA


/Users/jimp/Documents/Jalview/Rna/PaulG/RF00031_folded.stk

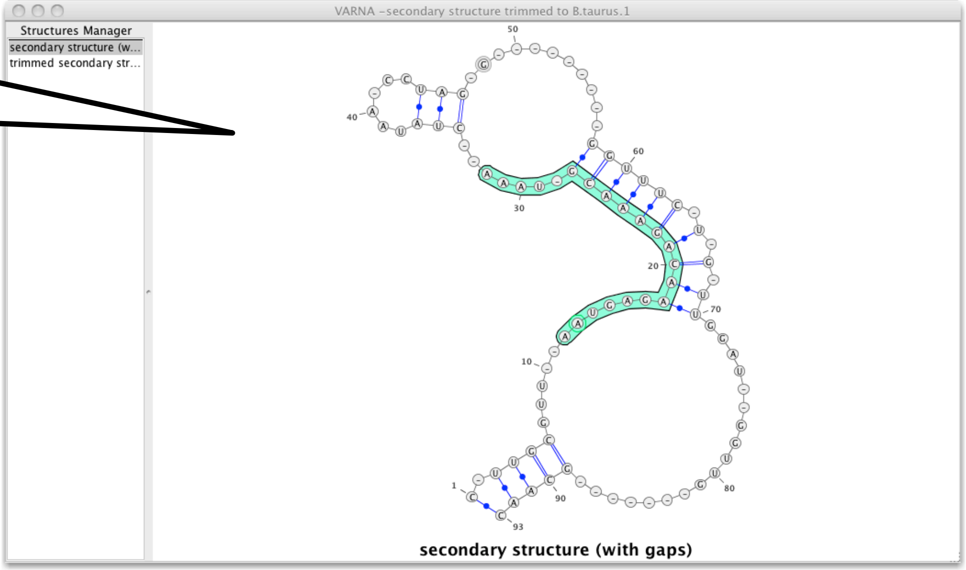
```
                   UUGCGUU--AAUGAGAACAGAAACG-UAAA--CUAUAA-CCUAG-G---------G
ter.3              G-UGGCGCU--UAUGACGCAGUUGUCU-UAAA-CUCGAAC--UCGA-GC--------G
lanogaster.2       C-AUUCAACU-UAUGAGGAUUAUUUCU-UAAA-GGCCUCU---GGC--U-------CG
D.melanogaster.1   G-AGCC-CU---AUGAUCGAUGAUUGG-CAAA-UCCUCUC--GAGG--A-------AC
R.norvegicus.7     C-CGGCACU--CAUGACGGUCUGCCUG-AAAA-CCAGCCC--GCUG-GU--------G
R.norvegicus.6     G-CCGCUUC---AUGACAGGAAGGACU-GAAA-UGUCUCA-AAGAC--C-----UGUG
R.norvegicus.5     G-UUUUUCC---AUGACGGUGUUUCCUCUAAA--UUUAC----AUG-----------G
R.norvegicus.4     G-UCAGAUG---AUGACGGCCUGUGCA-GAAA-CCCCCAC-GUGGG--C--------U
R.norvegicus.3     U-UUGCAUU--AAUGAGGAUUACACAG-AAAA-CCUUUGU--UAAGGGU--------U
R.norvegicus.2     G-UUACAUU--GAUGAGAACAGAAACA-UAAA--CUAUGA-CCUAG-G---------G
R.norvegicus.1     A-UAUUUGUU-UAUGAUGGUCACAGUG-UAAA--GUUCA----CAC-----------A
O.aries.1          G-ACGCUUC---AUGACAGGAAGGACU-GAAA-UGUCUCU-UGGAC-GC------CUG
M.musculus.9       C-CGGCACU--CAUGAAGGUCUGCUUG-AAAA-CCAGCCU--GCUG-GU--------G
M.musculus.8       U-UUGCAUU--AAUGAGGAUUACACAG-AAAA-CCUUUGU--UAAG-GA-------CU

                   (-((((-----------((((((((((-------(((------)))-----------)
```

Sequence 25 ID: C.elegans.1 Nucleotide: Cytosine (19)

Try it now:

1. Open VARNA with
   'Structure->View Structure->RNA Secondary Structure'
2. Explore difference between trimmed and untrimmed
3. Bonus points – add and link a **Jmol structure view**
   Bacillus_amyloliquef.9 *corresponds to PDB 3NPB*

VARNA has a wide range of 2D RNA plots and supports interactive annotation

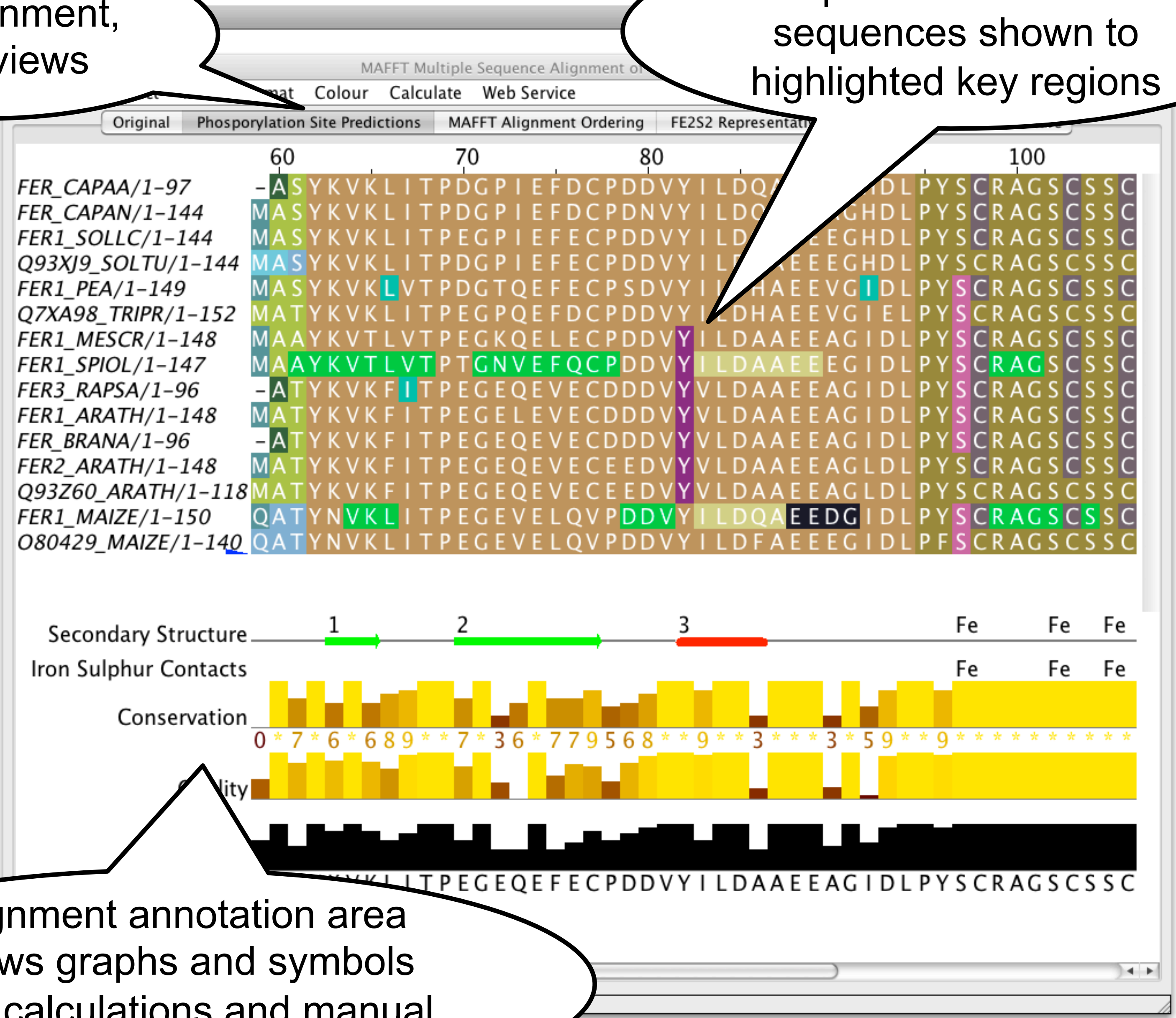

VARNA –secondary structure trimmed to B.taurus.1

Structures Manager
secondary structure (w...
trimmed secondary str...

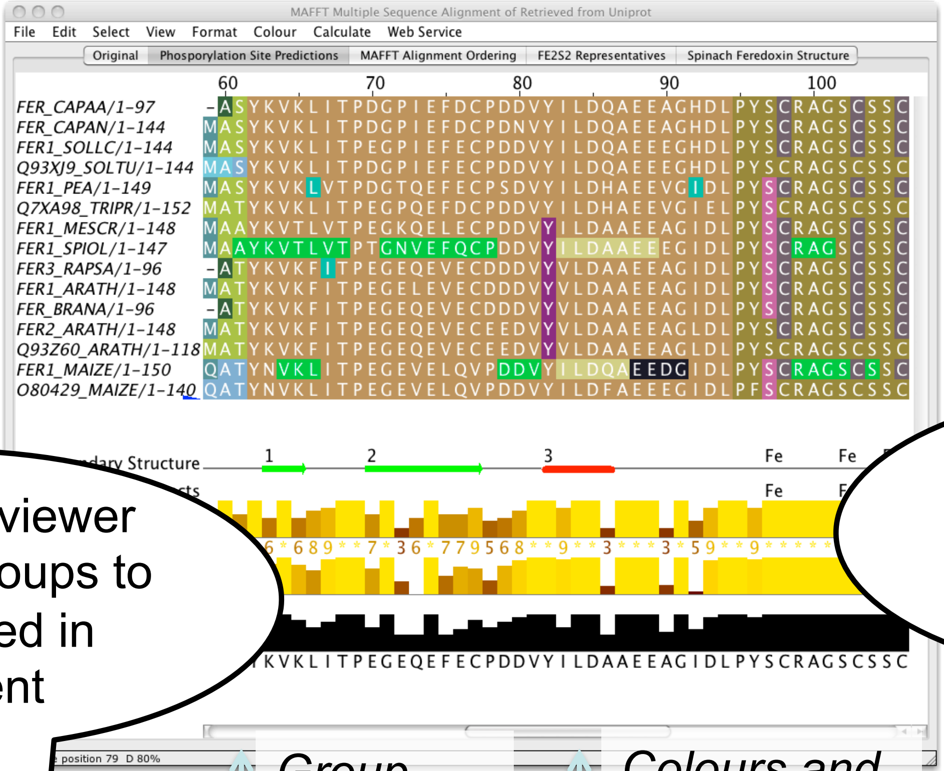
secondary structure (with gaps)

VARNA: Interactive drawing and editing of the RNA secondary structure
Kévin Darty, Alain Denise and Yann Ponty
Bioinformatics (2009) **25** 1974-1975

WRAP UP



```
FER_CAPAA/1-97       -ASYKVKLITPDGPIEFDCPDDVYILDQA     DLPYSCRAGSCSSC
FER_CAPAN/1-144      MASYKVKLITPDGPIEFDCPDNVYILDQ    GHDLPYSCRAGSCSSC
FER1_SOLLC/1-144     MASYKVKLITPEGPIEFECPDDVYILD   EEGHDLPYSCRAGSCSSC
Q93XJ9_SOLTU/1-144   MASYKVKLITPDGPIEFECPDDVYIL  AEEEGHDLPYSCRAGSCSSC
FER1_PEA/1-149       MASYKVKLVTPDGTQEFECPSDVYI  HAEEVGIDLPYSCRAGSCSSC
Q7XA98_TRIPR/1-152   MATYKVKLITPEGPQEFDCPDDVYILDHAEEVGIELPYSCRAGSCSSC
FER1_MESCR/1-148     MAAYKVTLVTPEGKQELECPDDVYILDAAEEAGIDLPYSCRAGSCSSC
FER1_SPIOL/1-147     MAAYKVTLVTPTGNVEFQCPDDVYILDAAEEEGIDLPYSCRAGSCSSC
FER3_RAPSA/1-96      -ATYKVKFITPEGEQEVECDDDVYVLDAAEEAGIDLPYSCRAGSCSSC
FER1_ARATH/1-148     MATYKVKFITPEGELEVECDDDVYVLDAAEEAGIDLPYSCRAGSCSSC
FER_BRANA/1-96       -ATYKVKFITPEGEQEVECDDDVYVLDAAEEAGIDLPYSCRAGSCSSC
FER2_ARATH/1-148     MATYKVKFITPEGEQEVECEEDVYVLDAAEEAGLDLPYSCRAGSCSSC
Q93Z60_ARATH/1-118   MATYKVKFITPEGEQEVECEEDVYVLDAAEEAGLDLPYSCRAGSCSSC
FER1_MAIZE/1-150     QATYNVKLITPEGEVELQVPDDVYILDQAEEDGIDLPYSCRAGSCSSC
O80429_MAIZE/1-140   QATYNVKLITPEGEVELQVPDDVYILDFAEEEGIDLPFSCRAGSCSSC

Conservation         0*7*6*689**7*36*779568**9**3***3*59**9**********
                           LITPEGEQEFECPDDVYILDAAEEAGIDLPYSCRAGSCSSC
```



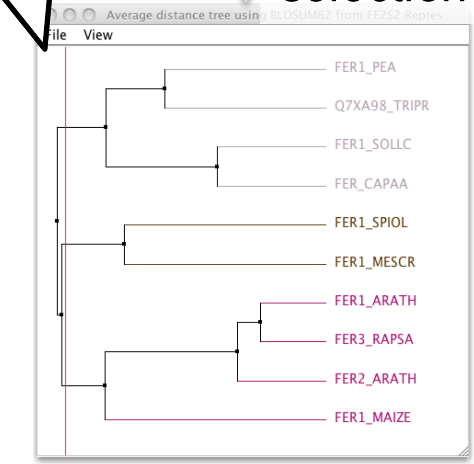
```
FER_CAPAA/1-97       -ASYKVKLITPDGPIEFDCPDDVYILDQAEEAGHDLPYSCRAGSCSSC
FER_CAPAN/1-144      MASYKVKLITPDGPIEFDCPDNVYILDQAEEAGHDLPYSCRAGSCSSC
FER1_SOLLC/1-144     MASYKVKLITPEGPIEFECPDDVYILDQAEEEGHDLPYSCRAGSCSSC
Q93XJ9_SOLTU/1-144   MASYKVKLITPDGPIEFECPDDVYILDQAEEEGHDLPYSCRAGSCSSC
FER1_PEA/1-149       MASYKVKLVTPDGTQEFECPSDVYILDHAEEVGIDLPYSCRAGSCSSC
Q7XA98_TRIPR/1-152   MATYKVKLITPEGPQEFDCPDDVYILDHAEEVGIELPYSCRAGSCSSC
FER1_MESCR/1-148     MAAYKVTLVTPEGKQELECPDDVYILDAAEEAGIDLPYSCRAGSCSSC
FER1_SPIOL/1-147     MAAYKVTLVTPTGNVEFQCPDDVYILDAAEEEGIDLPYSCRAGSCSSC
FER3_RAPSA/1-96      -ATYKVKFITPEGEQEVECDDDVYVLDAAEEAGIDLPYSCRAGSCSSC
FER1_ARATH/1-148     MATYKVKFITPEGELEVECDDDVYVLDAAEEAGIDLPYSCRAGSCSSC
FER_BRANA/1-96       -ATYKVKFITPEGEQEVECDDDVYVLDAAEEAGIDLPYSCRAGSCSSC
FER2_ARATH/1-148     MATYKVKFITPEGEQEVECEEDVYVLDAAEEAGLDLPYSCRAGSCSSC
Q93Z60_ARATH/1-118   MATYKVKFITPEGEQEVECEEDVYVLDAAEEAGLDLPYSCRAGSCSSC
FER1_MAIZE/1-150     QATYNVKLITPEGEVELQVPDDVYILDQAEEDGIDLPYSCRAGSCSSC
O80429_MAIZE/1-140   QATYNVKLITPEGEVELQVPDDVYILDFAEEEGIDLPFSCRAGSCSSC
```

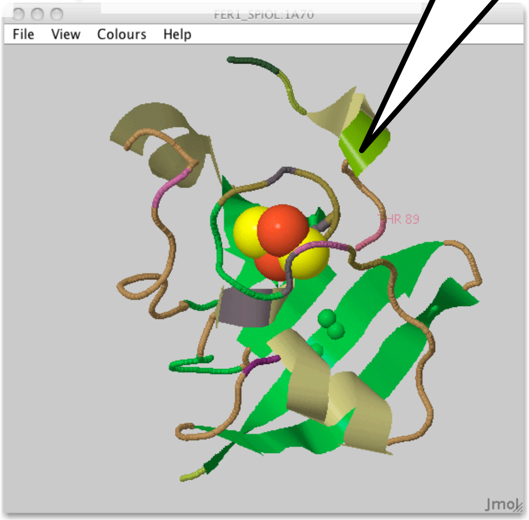
Linked tree viewer allows subgroups to be identified in alignment

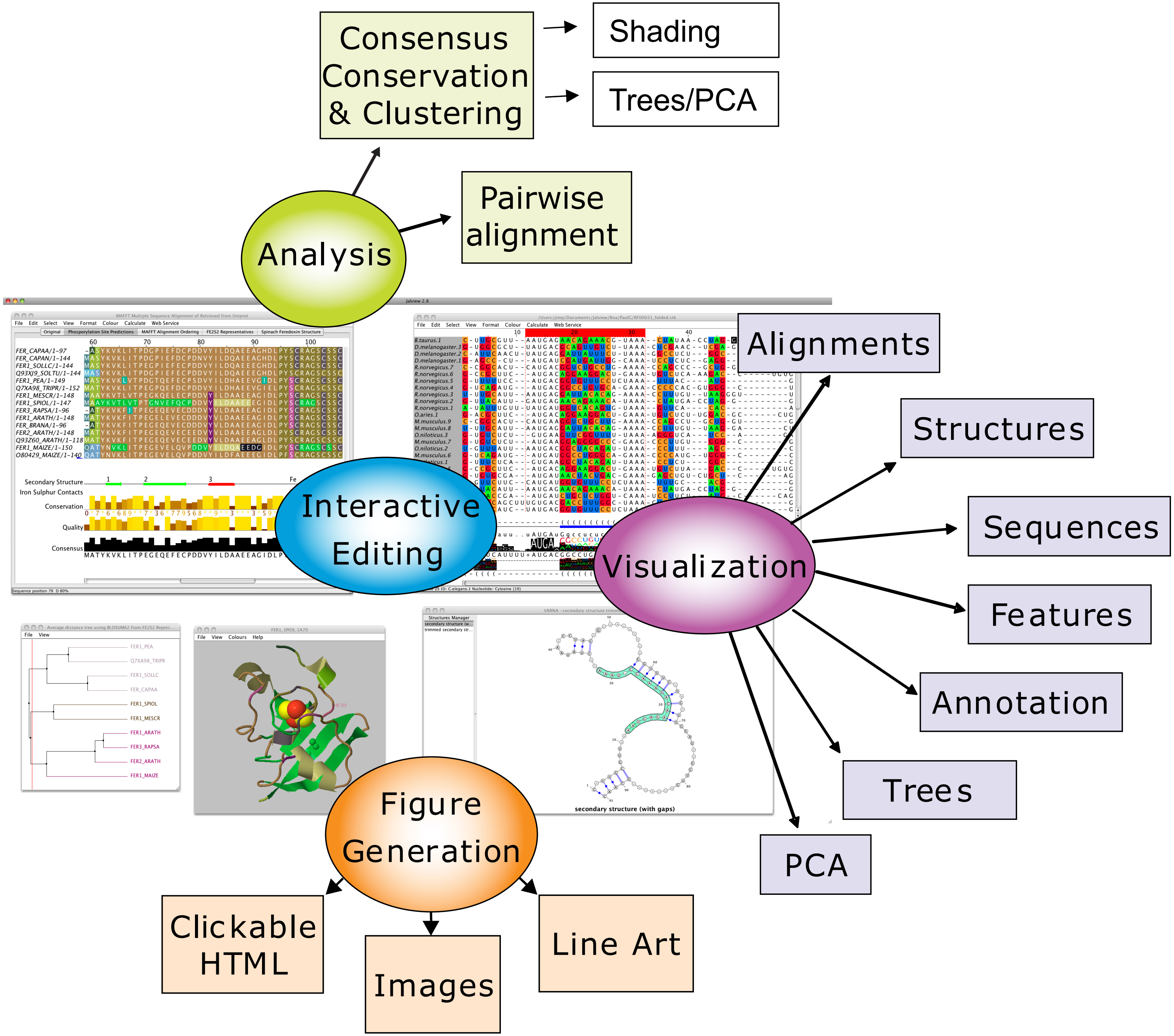
Linked Jmol viewer shows one or more structures coloured by alignment

Group selections

Colours and mouseovers







Analysis

Consensus Conservation & Clustering

Shading

Trees/PCA

Pairwise alignment

Interactive Editing

Visualization

Alignments

Structures

Sequences

Features

Annotation

Trees

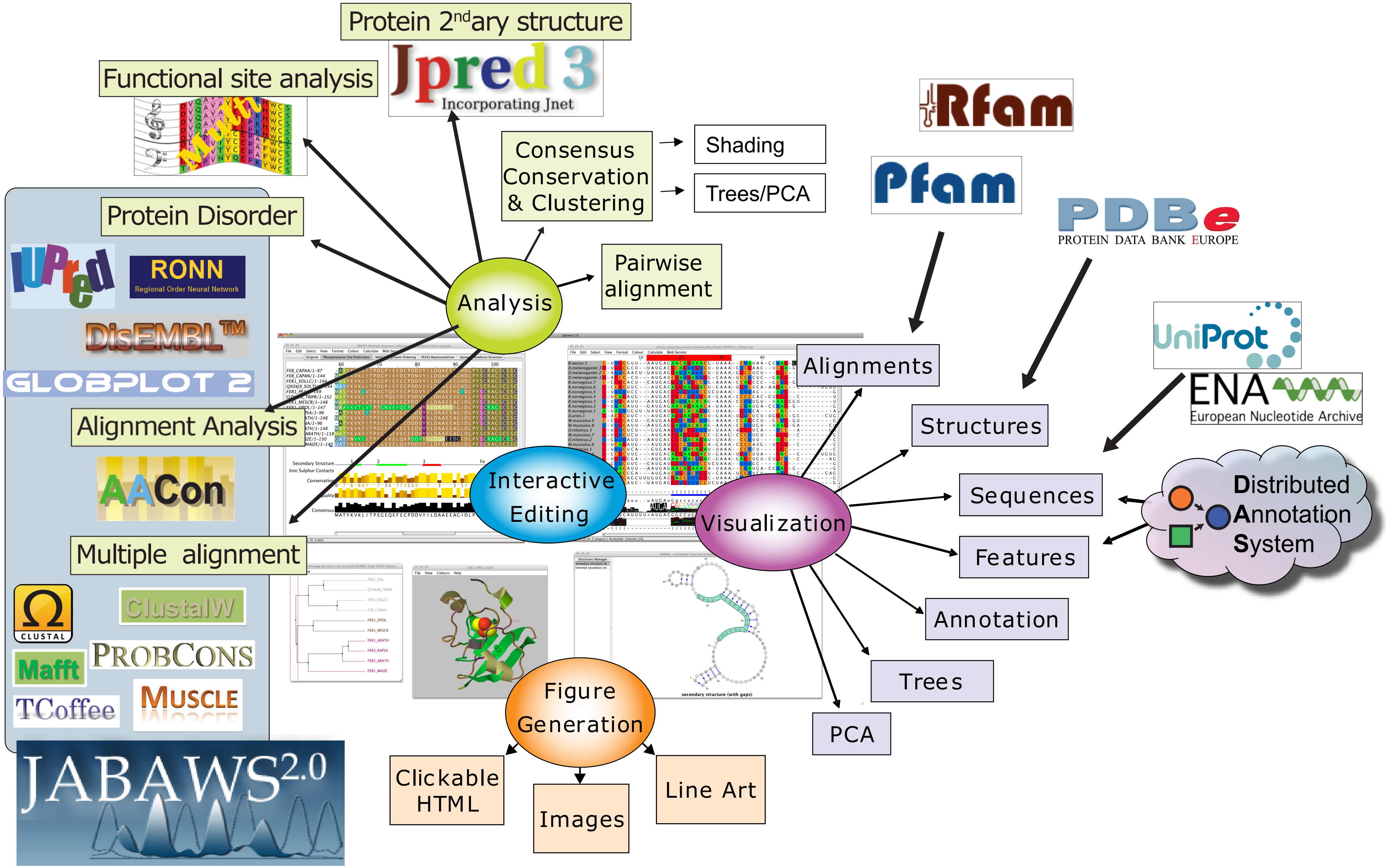
PCA

Figure Generation

Clickable HTML

Images

Line Art



Protein 2nd ary structure

Jpred 3
Incorporating Jnet

Functional site analysis

Multi

Protein Disorder

IUPred

RONN
Regional Order Neural Network

DisEMBL™

GLOBPLOT 2

Alignment Analysis

AACon

Multiple alignment

CLUSTAL

ClustalW

Mafft

ProbCons

TCoffee

Muscle

JABAWS 2.0

Consensus
Conservation
& Clustering

Shading

Trees/PCA

Pairwise
alignment

Analysis

Interactive
Editing

Visualization

Figure
Generation

Clickable
HTML

Images

Line Art

Rfam

Pfam

PDBe
PROTEIN DATA BANK EUROPE

UniProt

ENA
European Nucleotide Archive

Alignments

Structures

Sequences

Features

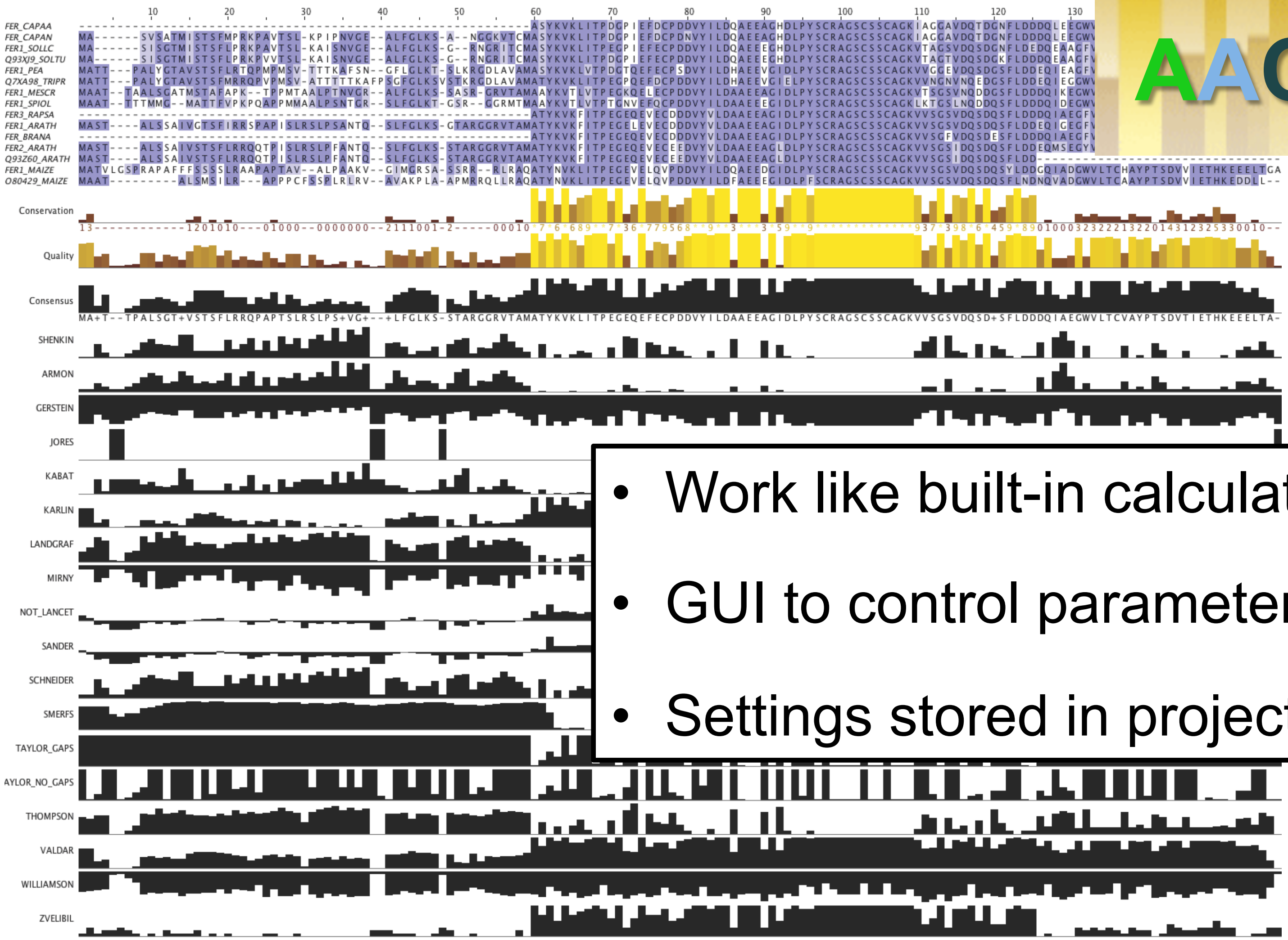
Annotation

Trees

PCA

Distributed
Annotation
System

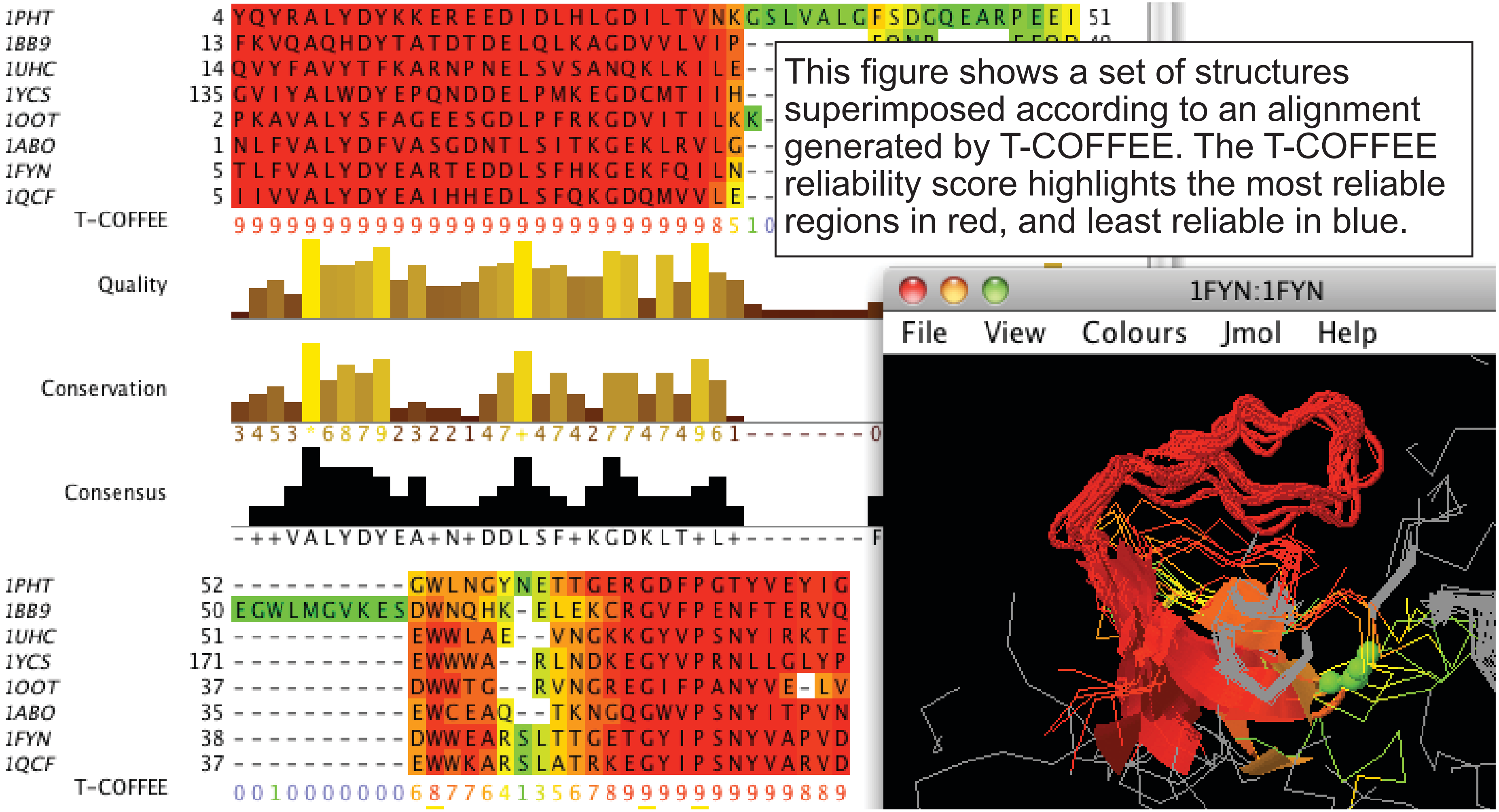
18 new alignment conservation calculations provided as web services



AACon

- Work like built-in calculations
- GUI to control parameters
- Settings stored in project file

T-COFFEE alignment reliability scores



```
1PHT       4 YQYRALYDYKKEREEDIDLHLGDILTVNKGSLVALGFSDGQEARPEEI 51
1BB9      13 FKVQAQHDYTATDTDELQLKAGDVVLVIP--
1UHC      14 QVYFAVYTFKARNPNELSVSANQKLKILE--
1YCS     135 GVIYALWDYEPQNDDELPMKEGDCMTIIH--
1OOT       2 PKAVALYSFAGEESGDLPFRKGDVITILKK-
1ABO       1 NLFVALYDFVASGDNTLSITKGEKLRVLG--
1FYN       5 TLFVALYDYEARTEDDLSFHKGEKFQILN--
1QCF       5 IIVVALYDYEAIHHEDLSFQKGDQMVVLE--
    T-COFFEE 9999999999999999999999999998510

     Quality

Conservation
             3453*6879232214 7+474277474961-------0

   Consensus

             -++VALYDYEA+N+DDLSF+KGDKLT+L+--------F

1PHT      52 ----------GWLNGYNETTGERGDFPGTYVEYIG
1BB9      50 EGWLMGVKESDWNQHK-ELEKCRGVFPENFTERVQ
1UHC      51 ----------EWWLAE--VNGKKGYVPSNYIRKTE
1YCS     171 ----------EWWWA--RLNDKEGYVPRNLLGLYP
1OOT      37 ----------DWWTG--RVNGREGIFPANYVE-LV
1ABO      35 ----------EWCEAQ--TKNGQGWVPSNYITPVN
1FYN      38 ----------DWWEARSLTTGETGYIPSNYVAPVD
1QCF      37 ----------EWWKARSLATRKEGYIPSNYVARVD
    T-COFFEE 00100000006877641356789999999999889
```

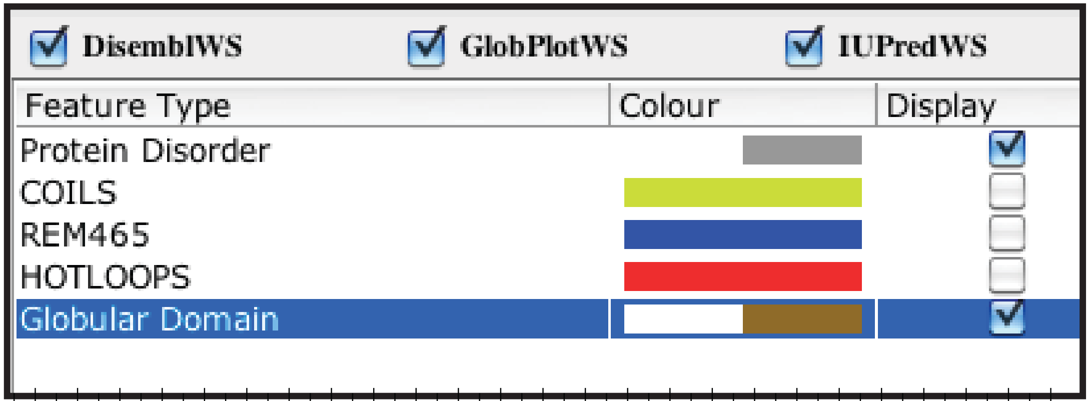
This figure shows a set of structures superimposed according to an alignment generated by T-COFFEE. The T-COFFEE reliability score highlights the most reliable regions in red, and least reliable in blue.

Disorder Predictions from JABAWS

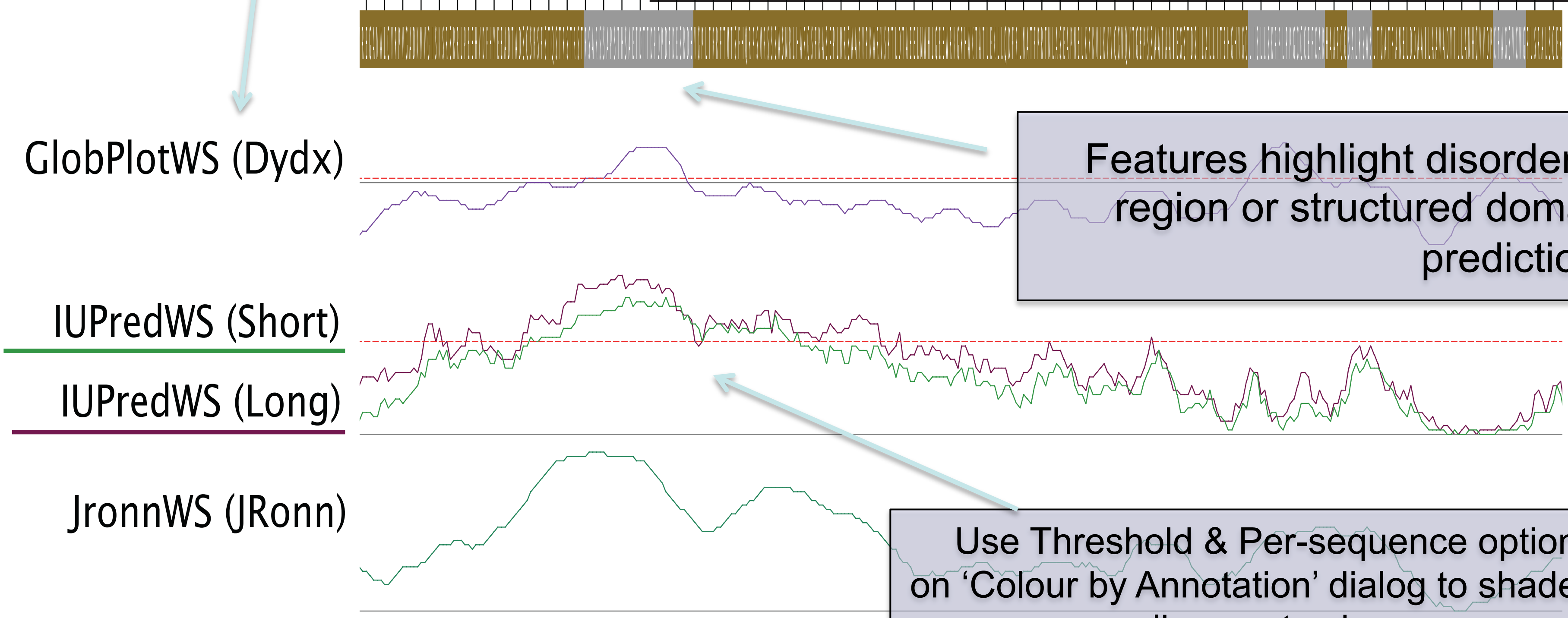
JABAWS Analysis Service

Jalview JABAWS 2.0 Client

Process results into both **annotation** and **features**



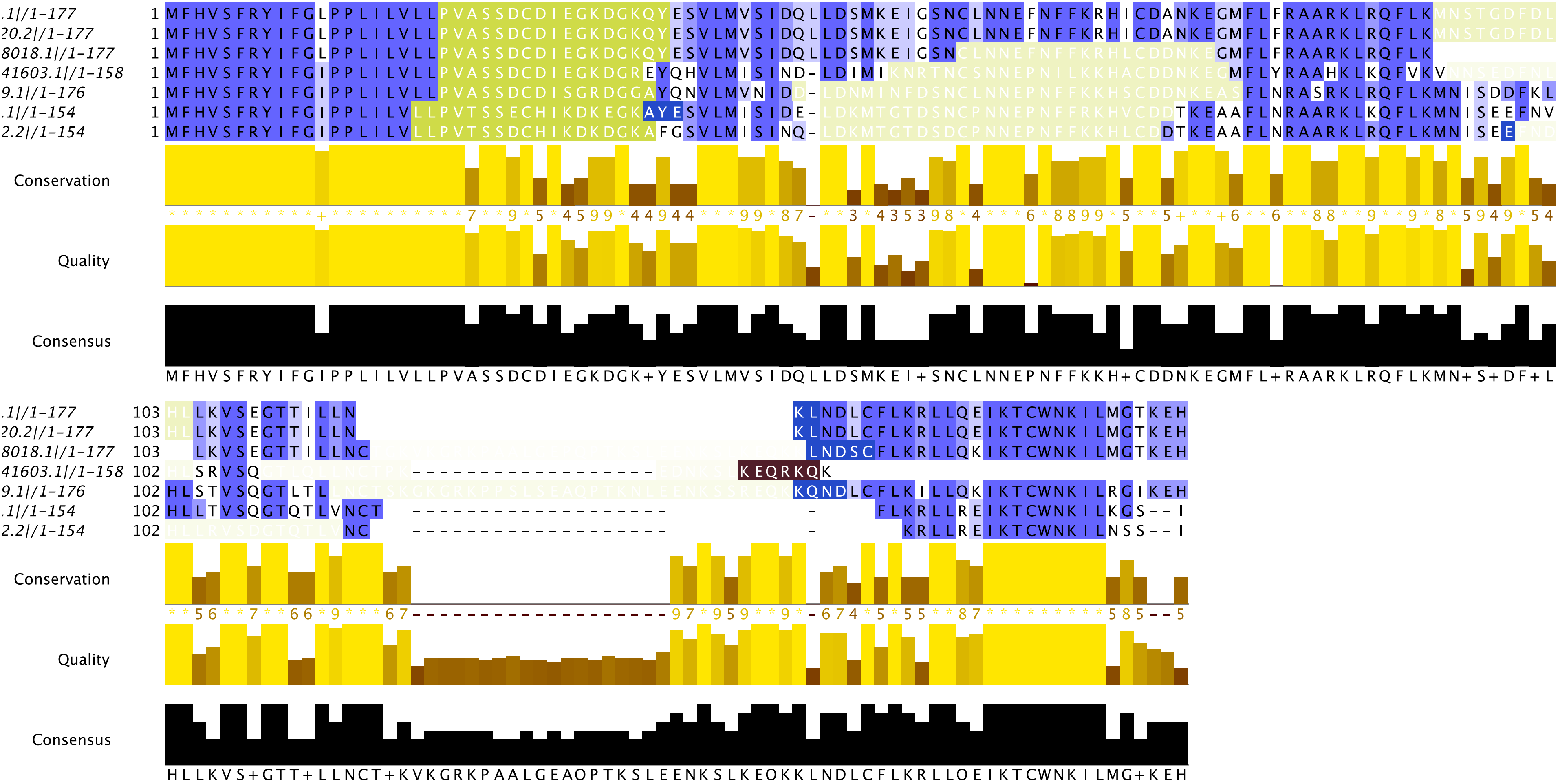
Feature Type	Colour	Display
Protein Disorder	▨	☑
COILS	🟨	☐
REM465	🟦	☐
HOTLOOPS	🟥	☐
Globular Domain	⬜🟫	☑



Features highlight disordered region or structured domain predictions

Use Threshold & Per-sequence option on ‘Colour by Annotation’ dialog to shade alignment using raw scores

Disorder in Interleukin 7



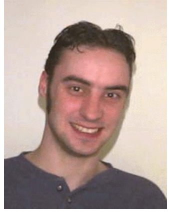
```
.1|/1-177          1 MFHVSFRYIFGLPPLILVLLPVASSDCDIEGKDGKQYESVLMVSIDQLLDSMKEIGSNCLNNEFNFFKRHICDANKEGMFLFRAARKLRQFLKMNSTGDFDL
?0.2|/1-177        1 MFHVSFRYIFGLPPLILVLLPVASSDCDIEGKDGKQYESVLMVSIDQLLDSMKEIGSNCLNNEFNFFKRHICDANKEGMFLFRAARKLRQFLKMNSTGDFDL
8018.1|/1-177      1 MFHVSFRYIFGLPPLILVLLPVASSDCDIEGKDGKQYESVLMVSIDQLLDSMKEIGSNCLNNEFNFFKRHLCDDNKEGMFLFRAARKLRQFLK
41603.1|/1-158     1 MFHVSFRYIFGIPPLILVLLPVASSDCDIEGKDGREYQHVLMISIND-LDIMIKNRTNCSNNEPNILKKHACDDNKEGMFLYRAAHKLKQFVKVNNSEDFNL
9.1|/1-176         1 MFHVSFRYIFGIPPLILVLLPVASSDCDISGRDGGAYQNVLMVNIDD-LDNMINFDSNCLNNEPNFFKKHSCDDNKEASFLNRASRKLRQFLKMNISDDFKL
.1|/1-154          1 MFHVSFRYIFGIPPLILVLLPVTSSECHIKDKEGKAYESVLMISIDE-LDKMTGTDSNCPNNEPNFFRKHVCDDTKEAAFLNRAARKLKQFLKMNISEEFNV
2.2|/1-154         1 MFHVSFRYIFGIPPLILVLLPVTSSDCHIKDKDGKAFGSVLMISINQ-LDKMTGTDSDCPNNEPNFFKKHLCDDTKEAAFLNRAARKLRQFLKMNISEEFND
Conservation
                     **********+**********7**9*5*4599*44944***99*87-**3*435398*4***6*8899*5**5+**+6**6**88**9**9*8*5949*54
Quality

Consensus
                     MFHVSFRYIFGIPPLILVLLPVASSDCDIEGKDGK+YESVLMVSIDQLLDSMKEI+SNCLNNEPNFFKKH+CDDNKEGMFL+RAARKLRQFLKMN+S+DF+L

.1|/1-177        103 HLLKVSEGTTILLN                                KLNDLCFLKRLLQEIKTCWNKILMGTKEH
?0.2|/1-177      103 HLLKVSEGTTILLN                                KLNDLCFLKRLLQEIKTCWNKILMGTKEH
8018.1|/1-177    103   LKVSEGTTILLNCTGKVKGRKPAALGEPQPTKSLEENKSLKEQKK LNDSCFLKRLLQKIKTCWNKILMGTKEH
41603.1|/1-158   102 HLSRVSQGTLQLLNCTPK------------------EDNKSLKEQRKQK
9.1|/1-176       102 HLSTVSQGTLTLLNCTSKGKGRKPPSLSEAQPTKNLEENKSSREQKKQNDLCFLKILLQKIKTCWNKILRGIKEH
.1|/1-154        102 HLLTVSQGTQTLVNCT  -------------------   -     FLKRLLREIKTCWNKILKGS--I
2.2|/1-154       102 HLLRVSDGTQTLVNC   -------------------   -       KRLLREIKTCWNKILNSS--I
Conservation
                     **56**7**66*9***67-------------------97*959**9*-674*5*55**87*********585--5
Quality

Consensus
                     HLLKVS+GTT+LLNCT+KVKGRKPAALGEAQPTKSLEENKSLKEQKKLNDLCFLKRLLQEIKTCWNKILMG+KEH
```

Jalview

www.jalview.org

<u>The Jalview developers</u>

Michele Clamp
Harvard, USA.



James Cuff
Harvard, USA.

Steve Searle
Sanger, UK

Andrew Waterhouse
Basel, Switzerland.

<u>RNA Features</u>

Lauren Lui
UC Santa Cruz, USA.

Jan Engelhardt
Univ. Leipzig, Germany.

Yann Ponty (**VARNA**)

SUMMER OF CODE // 2010
code.google.com/soc
Google

Google
SUMMER OF CODE
2011

<u>T-COFFEE Scores</u>

Paolo di Tomasso
Notredame Group, CRG, Spain.

Geoff Barton (**Money**)
David Martin (**Teaching**)
Peter Troshin (**JABAWS**)
Barry Strachan (**logo**)
Tom Walsh (**Apache**)
Ryan Maclaughlan (**CSS**)
Andrew Millar (**Drupal**)

All the Jalview users, and …

BBSRC
bioscience for the future