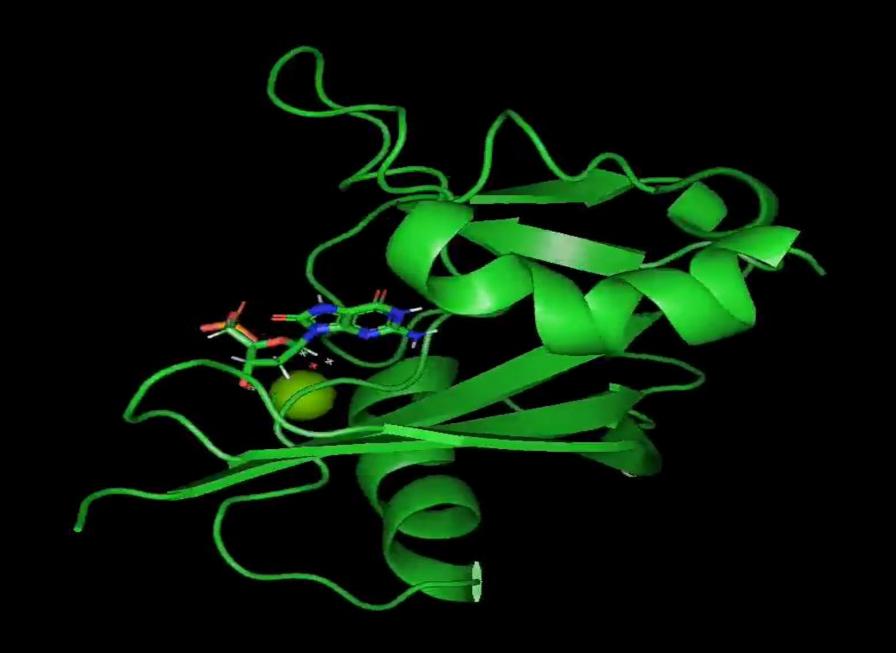
NUDIX水解酶在 大腸桿菌中的演化

連啟竣 鄭貽生 教授 魏宏仁 老師









DONKO 定義-Superfamily 蛋白質超家族 具有共同祖先的蛋白質群 重力機・ビックルドベ

演化是否存在一個規律

功能、結構、序列...



第一步:定義問題方向



EUPNKO

大腸桿菌的NUDIX Hydrolase的演化是 否存在一個規律

定義二 Motif 結構域

蛋白質中,具有共同胺基酸 序列之區域

定義一つりりに必 NUDIX hydrolase NUDIX 水解酶

具有特定MOTIF的SUPERFAMILY, 並水解核苷酸衍生物

7KX MuTT

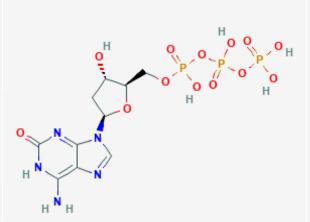
GooooEooooREUoEEoGU

將成為NUDIX BOX 所在的位置之分辨依據

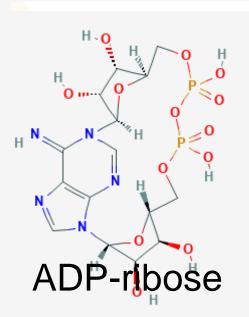
EUONKO

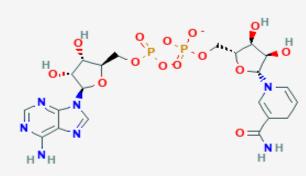
8-oH-dGTP

Ap2A

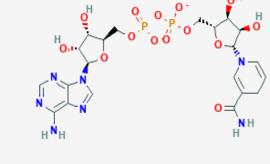


2-OH-dATP

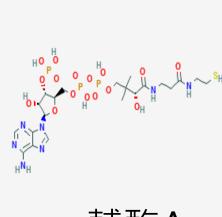




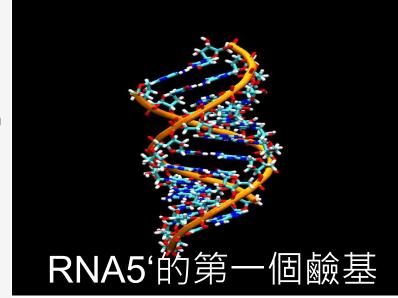
NADH



GDP-mannose



輔酶A



是否存在一個分類的方法?

是否可以圖像化?

並藉由圖像尋找演化的趨勢

對於每一個物種, 同時具有不同功能的NUDIX

每個NUDIX是為細胞功能而演化



類似的序列意味類似的功能

↑ ↑

透過不同功能的蛋白質的序列,尋找其演化的趨勢

問題:

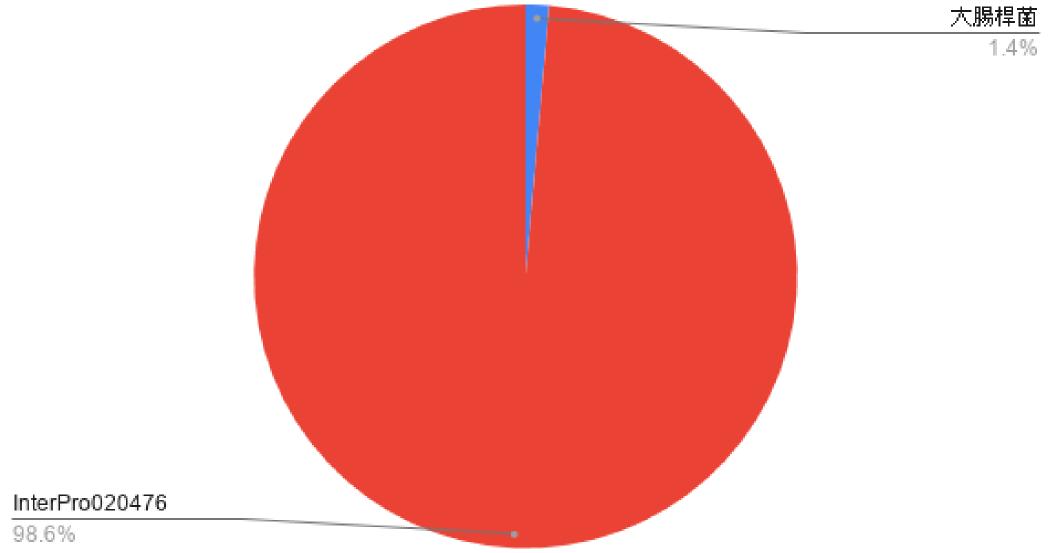
大腸桿菌?

所有物種?



含量





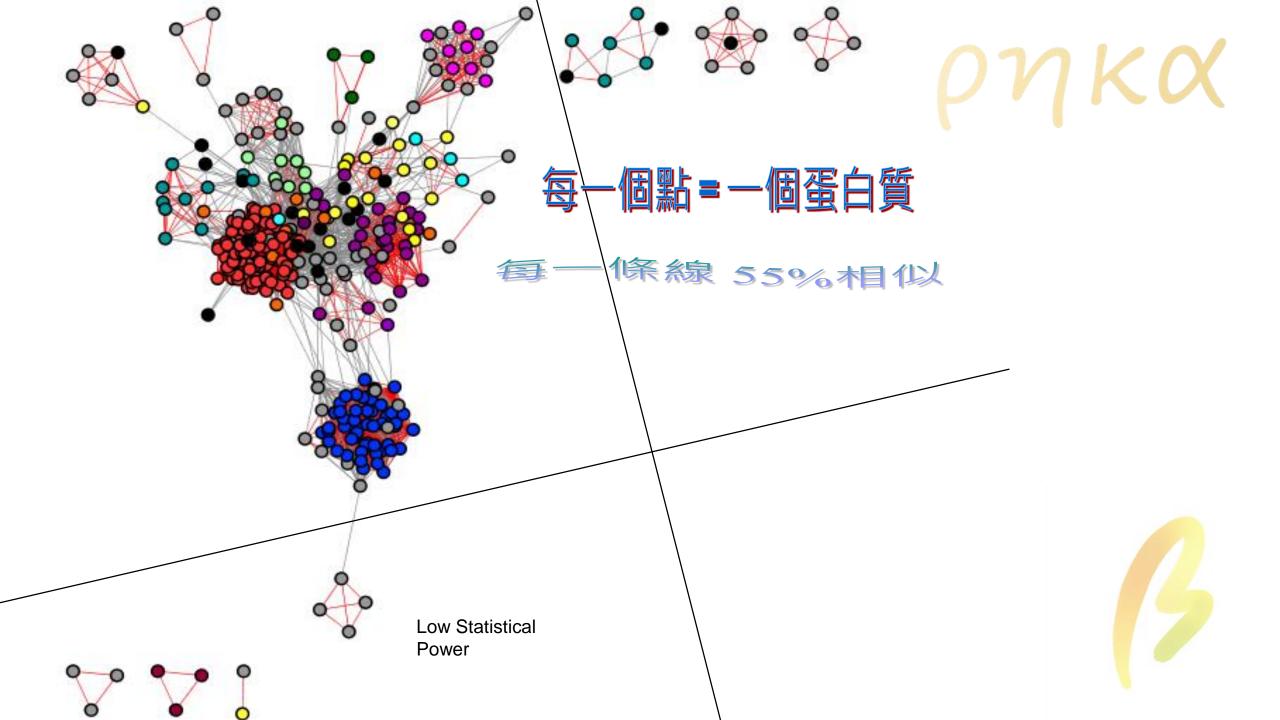
EUPNKO

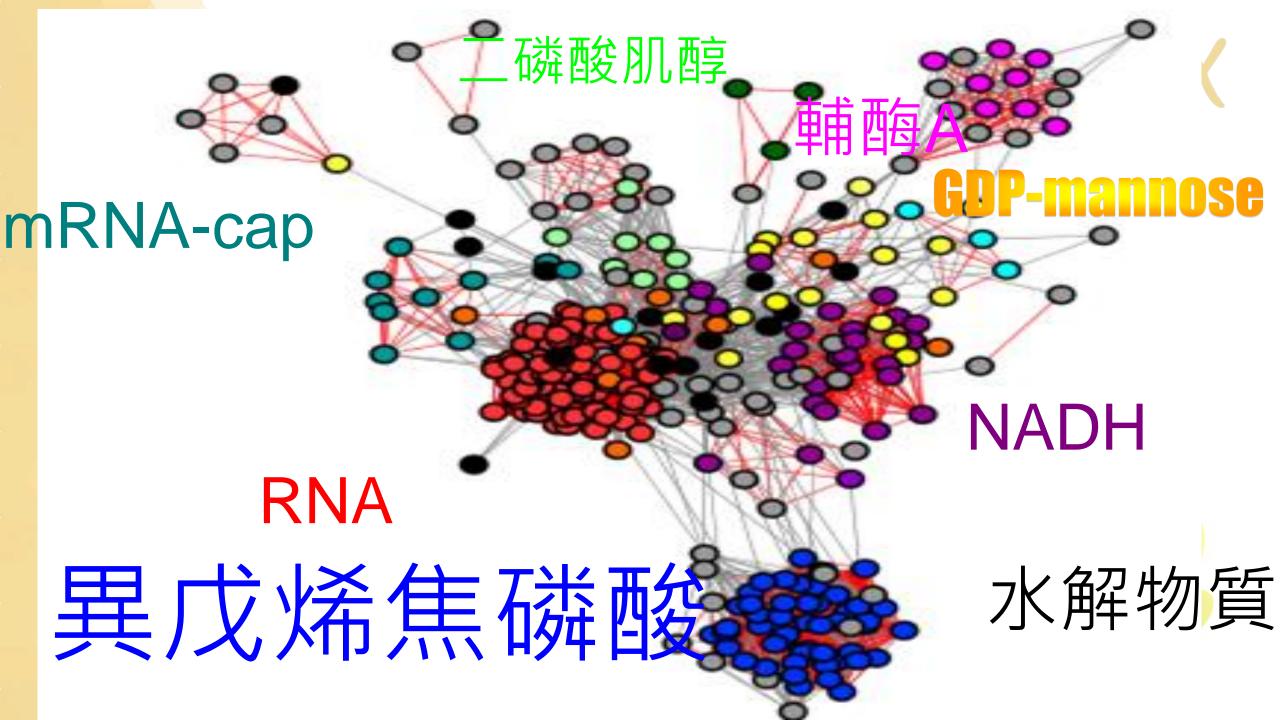
避免取樣偏誤

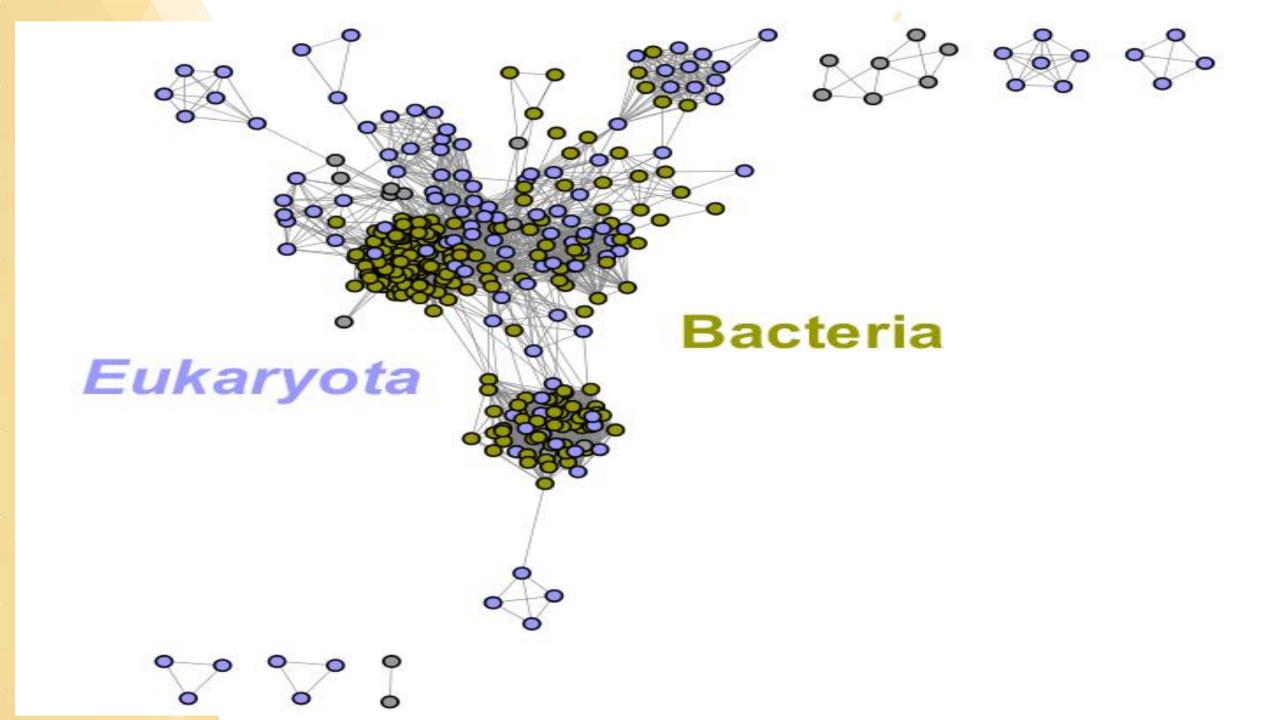
所有已知 推論 大陽桿菌

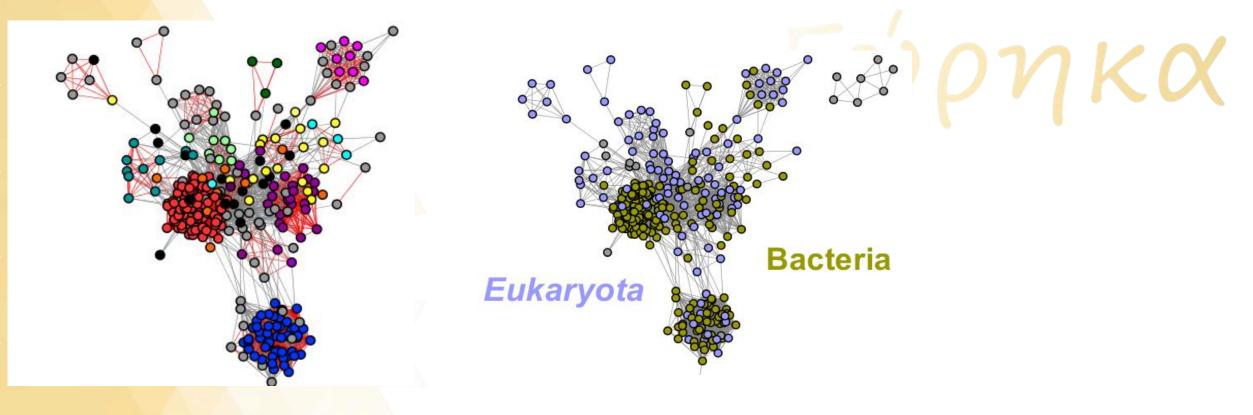
蒐集(手動挑選) 07/K(X) (InterPro) ① ① 資料處理(手動參數) (EFI-EST) ① ①

可視化(手動染色) (Cytoscape)









水解RNA為共祖蛋白最有可能的功能

- 一、演化趨向多樣化
- 1.紅色連接最多 (藍色<紅色)
- 二、物種演化順序確定紅色非真核

Ευρηκα

水解RNA為共祖蛋白最有可能的功能

為何有些NUDIX水解酶可以水解超過一種核苷酸 分化過程中功能遺失

大陽桿菌りかんの

NudA (Mutt)	8-OH-dGTP
NudB	2-OH-dATP \ dNTP
NudC	NADH \ Ap ₂ A
NudD	GDP-mannose (sugar)
NudE	Ap₃A ` ADP-ribose
NudF	ADP-sugar

大陽桿菌シの水の

NudG	5-OH-dCTP \ 2-OH-dATP \ 8-OH-dGTP
NudH	RNA
YeaB	未知
YmfB	TPP
YffH	GDP-mannose
YfaO	dUTP \ dCTP \ dTTP
YfcD	未知

定義四 Alignment 序列比對(排齊)

在最少刪減、延長下在不同序列 找到最多相似處的方法

```
1.0
                   2.0
                              30
                                        4.0
                                                    50
MTASAOPRGR RPGVGVGVVV TSCKHPRCVL LGKRKGSVGA GSFOLPGGHL
        60
                  70
                             80
                                         90
                                                   100
EFGETWEECA ORETWEEAAL HLKNVHFASV VNSFIEKENY HYVTILMKGE
       110
                  120
                            130
                                       140
VDVTHDSEPK NVEPEKNESW EWVPWEELPP LDQLFWGLRC LKEQGYDPFK
      160
EDLNHLVGYK GNHL
```

10 2.0 30 40 5.0 MYWSNQITRR LGERVQGFMS GISPQQMGEP EGSWSGKNPG TMGASRLYTL 70 80 60 90 100 VLVLQPQRVL LGMKKRGFGA GRWNGFGGKV QEGETIEDGA RRELQEESGL 110 120 130 140 150 TVDALHKVGQ IVFEFVGEPE LMDVHVFCTD SIQGTPVESD EMRPCWFQLD 160 170 180 190 QIPFKDMWPD DSYWFPLLLQ KKKFHGYFKF QGQDTILDYT LREVDTV

```
sp|Q9NV35|NUD15 HUMAN
                          ------<mark>VVV</mark>TSCKHPRCVL 30
sp|P36639|80DP HUMAN
                    MYWSNQITRRLGERVQGFMSGISPQQMGEPEGSWSGKNPGTMGASRLYTLVLVLQPQRVL 60
                                         sp|Q9NV35|NUD15 HUMAN
                    LGKRKGSVGAGSFQLPGGHLEFGETWEECAQRETWEEAALHLKNVHFAS-VVNSFIEKEN 89
sp|P36639|80DP HUMAN
                    LGMKKRGFGAGRWNGFGGKVQEGETIEDGARRELQEESGLTVDALHKVGQIVFEFVGEPE 120
                    sp|Q9NV35|NUD15 HUMAN
                    YHYVTILMKGEVDVTHDSEPKNVEPE-KNESW---EWVPWEELPPLDQLFWGLRCLK--E 143
sp|P36639|80DP HUMAN
                    LMDVHVFCTDSIQ-----GTPVESDEMRPCWFQLDQIPFKDMWPDDSYWFPLLLQKKKF 174
                      sp|Q9NV35|NUD15 HUMAN
                    OGYDPFKEDLNHLVGYKGNHL---
                                          164
sp|P36639|80DP HUMAN
                    HGYFKFQGQ-DTILDYTLREVDTV
                                          197
```





相異性高作比對會有 甚麼問題?



對於每一個物種,可時具有不同功能的NUDIX

每個NUDIX是為 細胞功能而演化 而非跨物種的演化



ΩÛ

透過不同功能的蛋白質的序列(非物種) 尋找其演化的趨勢

結構較序列演化上緩慢

結構更多相似處,比對更加準確

ÛÛ

使用mTM-align



```
cov
                             pid
                                      M-GASRLYTLVLVLQ---PQRVLLGMKKRGF-GAGRWNGFGGKVQEGETIEDGARRELQEESGLTVDALHKVGQIVFEF
1 mth1.pdb
                   100.0% 100.0%
                    86.5% 17.3%
2 mth2.pdb
  consensus/100%
                                      ..ttstlhshvlvhp...PppVLLGh+.+G..GAGpaph.GG+lp.GEThE-sApREh.EDuuLplcslthhsu.ls..h.
                                      ..ttstlhshVlVhp...PppVLLGh+.+G..GAGpaph.GG+lp.GEThE-sApRSh.
                                                                                                       . Duu Lplesl hsu.ls..h.
  consensus/90%
                                      ..ttstlhshVlVhp...PppVLLGh+.+G..GAGpaph.GG+lp.GEThB-sApRBh.BBuuLplcslHhsu.ls..h.
..ttstlhshVlVhp...PppVLLGh+.+G..GAGpaph.GG+lp.GEThB-sApRBh.BBuuLplcslHhsu.ls..h.
  consensus/80%
  consensus/70%
                             pid 81
                                                                                                                             . 160
                      COV
                                      GEP---ELMDVHVFCTDSI--QG-TPVES-DE-M-RPCWFQLDQIP-FKDMWPDDSYWF--PLLLQK-KK-FH--GYFKF
1 mth1.pdb
                  100.0% 100.0%
                                      ---KENYHYVTILMKGEVDVTHDSEPKNVEPEKNESWEWVPWEELPPLDQLFWGLRCLKEQGYDPF<mark>K</mark>EDLNHLVGYKGN-
2 mth2.pdb
                    86.5% 17.3%
                                      .....hhsshlhps-s...ps.pphps.so...p.pwh.h-plp.hcpha.s.phhh..sh...k.ch.ah..hhhp.
  consensus/100%
                                      .....hhsshlhps-s...ps.pphps.so...p.pwh.h-plp.hcpha.s.phhh..sh...k.ch.ah..hhhp.
  consensus/90%
                                      .....hhsshlhps-s...ps.p\mathbf{p}hps.s\mathbf{p}...p.p\mathbf{w}h.h-pl\mathbf{p}.hcpha.s.phhh..sh...\mathbf{x}.ch.ah..hhhp.
  consensus/80%
  consensus/70%
                                      .....hhsshlhps-s...ps.pphps.sp...p.pmh.h-plp.hcpha.s.phhh..sh...k.ch.ah..hhhp.
                      COV
                             pid 161
                                                       ] 177
1 mth1.pdb
                   100.0% 100.0%
                                      QGQDTILDYTLREVDTV
2 mth2.pdb
                    86.5% 17.3%
                                      -----HL-----
  consensus/100%
                                      .....ch...ch
  consensus/90%
                                      .....ch....ch
  consensus/80%
                                      .....ch...ch
  consensus/70%
                                      .....ch....ch
```

```
pid
                   COV
                100.0% 100.0%
                                                   ORVLLGMKKRGF-GAGRWNGFGGKVOEGETIEDGARRELOEESGLTVDALHKVGOIVFEFV
1 mth1.pdb
2 mth2.pdb
                 86.5% 17.3%
  consensus/100%
                                                               .GAGpaph.GG+lp.GEThE-sApReh.EEuuLplcslChsu.ls..h.
  consensus/90%
  consensus/80%
  consensus/70%
                          pid 81
                   COV
                                 GEP ---ELMDVHVFCTDSI--QG-TPVES-DE-M-RPCWFQLDQIP-FKDMWPDDSYWF--PLLLQK-KK-FH--GYFKF
1 mth1.pdb
                100.0% 100.0%
2 mth2.pdb
                 86.5% 17.3%
                                 ---KENYHYVTILMKGEVDVTHDSEPKNVEPEKNESWEMVPWEELPPLDQLFWGLRCLKEQGYDPFKEDLNHLVGYKGN-
                                 .....hhsshlhps-s...ps.pPhps.sD...p.pWh.h-plP.hcpha.s.phhh..sh...K.ch.ah..hhhp.
  consensus/100%
                                 .....hhsshlhps-s...ps.pPhps.sD...p.pWh.h-plP.hcpha.s.phhh..sh...K.ch.ah..hhhp.
  consensus/90%
                                 .....hhsshlhps-s...ps.pPhps.sD...p.pWh.h-plP.hcpha.s.phhh..sh...K.ch.ah..hhhp.
  consensus/80%
                                 .....hhsshlhps-s...ps.pphps.sp...p.pmh.h-plp.hcpha.s.phhh..sh...k.ch.ah..hhhp.
  consensus/70%
                                                ] 177
                         pid 161
                   COV
                                 OGODTILDYTLREVDTV
1 mth1.pdb
                100.0% 100.0%
2 mth2.pdb
                                 -----HL-----
                 86.5% 17.3%
  consensus/100%
                                 ......ch......
                                                           GAP(-) 比較少
  consensus/90%
  consensus/80%
                                 .....ch...ch
                                 ....ch...ch
  consensus/70%
```

```
sp|Q9NV35|NUD15 HUMAN
                                          --MTASAQPRGRRPGVGVG-----VVVTSCKHPRCVL 30
sp|P36639|80DP HUMAN
                       MYWSNQITRRLGERVQGFMSGISPQQMGEPEGSWSGKNPGTMGASRLYTLVLVLQPQRVL 60
                                                                       ** * * *
sp|Q9NV35|NUD15 HUMAN
                       LGKRKGSVGAGSFQLPGGHLEFGETWEECAQRETWEEAALHLKNVHFAS-VVNSFIEKEN 89
sp|P36639|80DP HUMAN
                       LGMKKRGFGAGRWNGFGGKVQEGETIEDGARRELQEESGLTVDALHKVGQIVFEFVGEPE 120
                       sp|Q9NV35|NUD15 HUMAN
                       YHYVTILMKGEVDVTHDSEPKNVEPE-KNESW---EWVPWEELPPLDQLFWGLRCLK--E 143
sp|P36639|80DP HUMAN
                       LMDVHVFCTDSIQ-----GTPVESDEMRPCWFQLDQIPFKDMWPDDSYWFPLLLQKKKF 174
                                        _ ** : _ _ * : : * * * *
                          * :: ...:
sp|Q9NV35|NUD15 HUMAN
                       OGYDPFKEDLNHLVGYKGNHL---
                                                 164
sp|P36639|80DP HUMAN
                                                 197
                       HGYFKFQGQ-DTILDYTLREVDTV
```

對於每一個物種, 同時具有不同功能的NUDIX

ÛÛ

每個NUDIX是為 細胞功能而演化 而非跨物種的演化

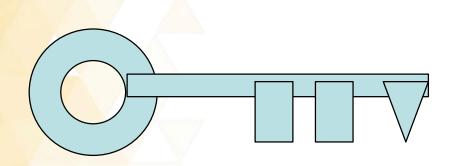


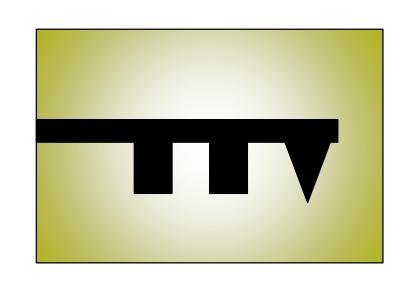
類似的結構 更加意味類 似的功能

① ①

透過不同功能的蛋白質的序列(非物種),尋找其演化的趨勢

Eupnka





類似的鑰匙孔更加意味類似的鑰匙

集(手動挑選12個結構序列) (PDBeBank)

ÛÛ

資料處理

做序列分析 (mTM-align) 做演化樹分析 (PhyML)

可視化 (IToL)

Maximum-Likelihood的演化假設

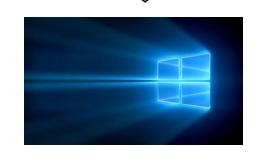
哪一個演化樹容易發生?



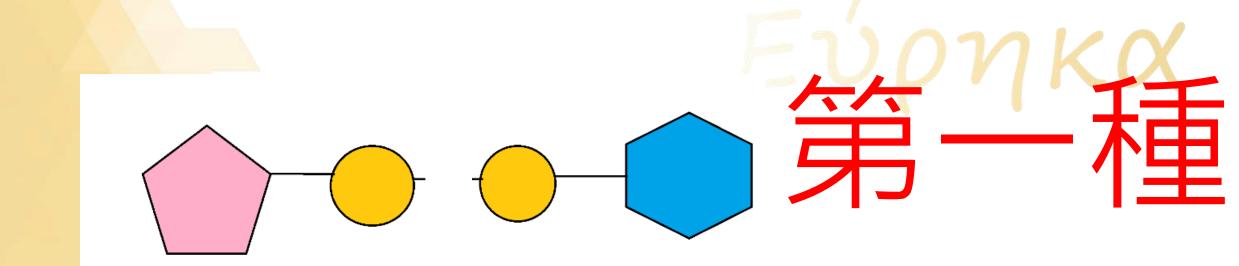


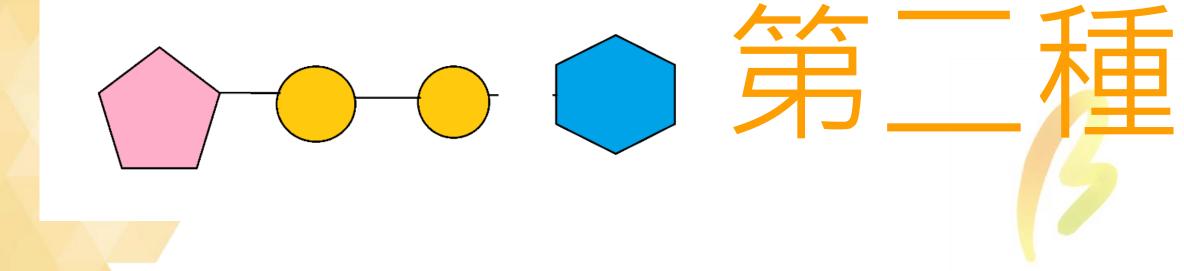
永遠選後設機率較高者

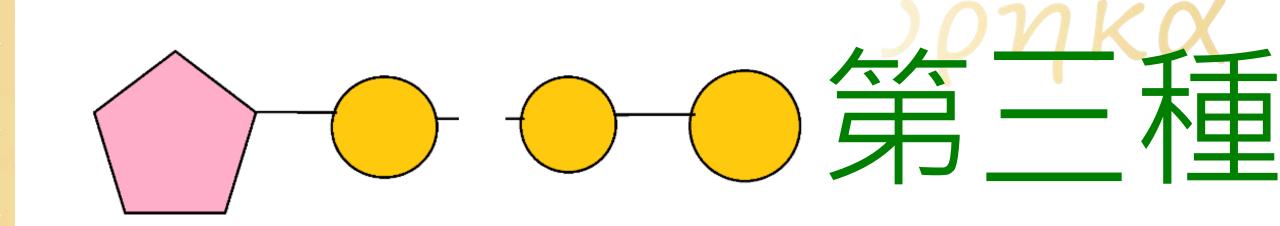


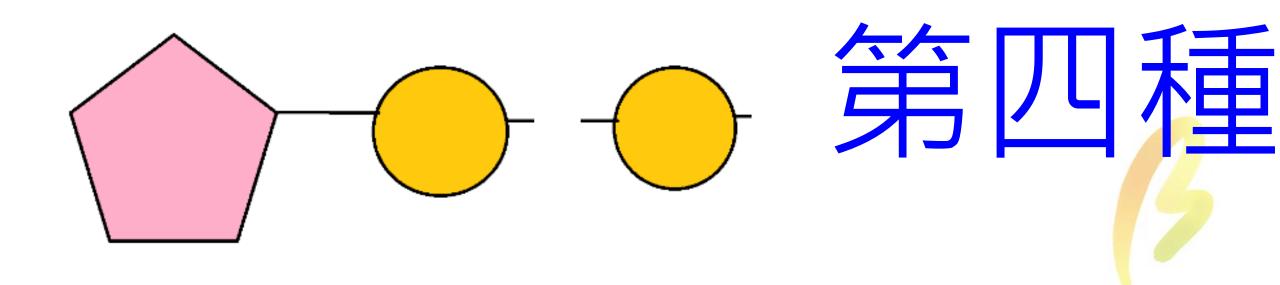


MTH1.pdb NudB.pdb NudC.pdb NudD.pdb	KIHHHH HI-DRIIEKLDHGAAYYYSH EQKLWLPKGELPYGEAANFDLYG SKS LQ-KPTIL NYETYAR SRLFTY MLKPDNLPYTFGKNDYEII ARETLYR GFFSL G-P LGS PH MTQ-QITLI KDKILSD N-YFTL
NUGE, DOD	SKSLQKPTILNYETYARSRLFTY
NudF.pdb NudG.pdb NudH.pdb	G-FFSL
NudH. pdb YeaB. pdb	РН
YeaB.pdb YfaO.pdb YffH.pdb	N-YFTLN-YFTL
ттв.рав	
cons	
MTH1.pdb NudB.pdb	
NUGL. DOD	QRALQIGEYQGEPYWLYQQQRRHDGSYRQYIDLDYGLFQLAGRGYQLAEFYRSHKYCGYCGHEYPSKTEYALCSH
NudD. pdb NudE. pdb NudF. pdb	
Nudu odb	
YeaB.pdb YfaO.pdb YffH.pdb	
YTTH.pdb YmfB.pdb	QR ALQIGEYQ GEPY WLYQQQRRHDGSYRQYIDLDYGLFQLAGRGYQLAEFYRSHKYCGYCGHEYPSKTEWALCSH YYR S ESYDLE FS NGYRRYY ER DLYRFR HRLF NGQMSHEYRR EIF A HNITYD LTR KDGEYIRHKR EYY
cons	
MTH1. odb	K-K-LOIAYGIIRN-E-NNEIFITRRAADAHMANKLEFPGGKIEM
MTH1.pdb NudB.pdb NudC.pdb NudD.pdb NudE.pdb	K
NudD. pdb NudE. pdb	T-PLYSLDFIYEN-S-RGEFLLGKRTNRPAQGYWFYPGG-RYQSYELGFSKG-LID
Nudg: Bab	
NudH. pdb YeaB. pdb	DPLDDIQADPYALVLSGYRRAAYLYALTREADPRYLLTYRSK-GQIAFPGG-SLD
YeaB.pdb YfaO.pdb YffH.pdb YmfB.pdb	
cons	
MTH1.pdb NudB.pdb NudC.pdb NudD.pdb	G-ETPEGAYYRELQEEYGI TPQHF-SLF-EKLEYEF-P
Nude: pdb	KDETLEAAFERLTMAELGLRL - PITAG-OFY-GYWOH-FY-DDNFSGTDFTTHYYYLGFRFRYSEEELL
NudE.pdb NudF.pdb NudG.pdb	EGESYEDYARREAI EEAGLIY
NudH.pdb YeaB.pdb	PGESAEOAMYRELFEEYGLSR-KDYRIL-ASTRWYLRY-KLPKRLYRWDTKPYCIGOKOKYFLLOLY-SG-DAEI AGETPTQAALREAQEEYALDP-AA-YTLLGEL-DDYFT-PYGFHYTPYLGRIA-PEALDTL
Yfa0.pdb YffH.pdb	PGERIEEALRREIREELGEOL ILSDI-TPW-TFRDD-IR-IKT-YGRQEEIYMIYLIFDCYSA-NRDI N-DEPEYCIRKAAIEETGYEYGEY-RKL-FELYM-SPGGYTELIHFFIAEYS-DN-QR-AN
ттв.рав	ADETLYEAAARELWEETGISAQPQHFI-RMHQW-IAPDKTPFLRFLFAIELE-QI-CPT
cons	
MTH1.pdb NudB.pdb	-WGKBGOPGEYMSLYGLNADD-F-PPANEPYI-AKLKRL -Y-F-T
NudC. pdb NudD. pdb NudE. pdb NudE. pdb NudE. pdb	-Y-I-DPKELLEAM/YRY-DDL
Nude: pdb	PEPL-POYRMPLAHD
NudH. pdb	- Q-A-H
YeaB.pdb YfaO.pdb YffH.pdb	-R-T-TPETAULITPTLAELRATPLTRERRILPDGTETPLTR-YPYRGLD-I-WGMTARTLHDLLE
YmfB.pdb	-WG - K - EGOPGE/MSLYG - L - NA - DD F - PP - ANEP - YI - AKLKEL - Y - F - T - E - HLAYKYLDA - PA - AA - A - L - T - K - SW - SNRQAI - E Q - Y - I - DP - KELLEANY/RY - DD - L - PP - L - L - PP - PO - YARRIE - DTYA - L - L - AS - DN Y - HA - N - SRAYFLAE - KRTGYPGL - P - EPL - POYRAYLAH - D - L - E - DP - DF - NE - A - RNYSALFLY - R - E - EDI - RYHYYSR - EQ - AYQ - WY - EE - KK I - DN - A - ASYIALOWL - Q - L - Q - A - B - EHQAL WY - SP - EE - AL - Q - Y - P - L - P - A - DI PLLEAF - M - A - N - P - P - P - P - P - P - P - P - P
cons	
MTH1.pdb	
NudB. pdb NudC. pdb	
NudD. pdb NudE. pdb NudF. pdb	\(\text{VL} - \text{K} - \text{G} - \dots - \dots - \dots \\ \text{QGRY} - \dots \\ \dots \text{NPWB} \\ \dots \\ \dots \text{NPWB} \\ \dots \\dots \\ \dots \\dots \\ \dots \\ \dots \\dots \\ \dots \\dots \\ \dots \\dots \\ \dots \\ \dots \\dots \\ \dots \\ \dots \\dots \\ \dots \\dots \\dots \\ \dots \\dots \
NudG: pdb NudH: pdb	- L - RAARPAD
YeaR odb	FY I N CRA
YfaO.pdb YffH.pdb YmfB.pdb	SH
cons	





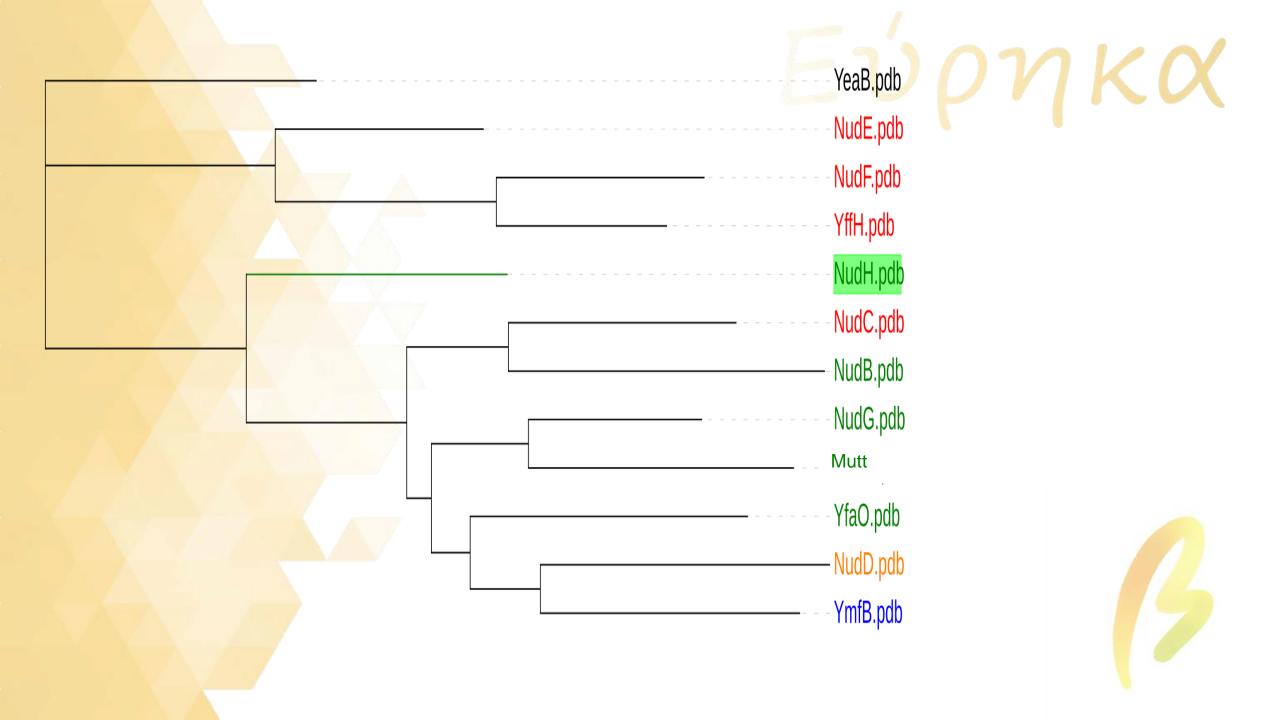




EUPNKO

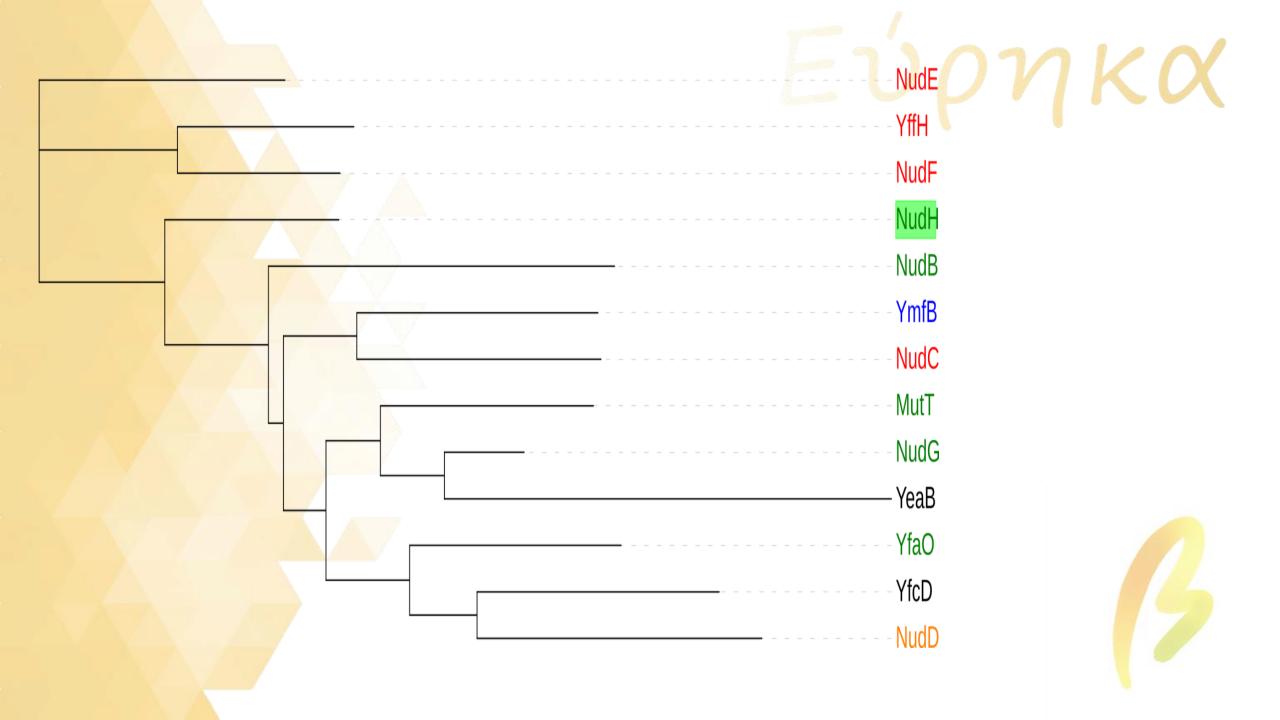
Structure Only





Sequence Only





NudA NudG



Evpnka

NudA (Mutt)	8-OH-dGTP
NudB	2-OH-dATP \ dNTP
NudC	NADH \ Ap2A
NudD	GDP-mannose (sugar)
NudE	Ap₃A ` ADP-ribose
NudF	ADP-sugar

		ONKO
NudG	5-OH-dCTP \ 2-OH-	
	dATP \ 8-OH-dGTP	
NudH	RNA	
YeaB	未知	
YmfB	TPP	
YffH	GDP-mannose	
YfaO	dUTP \ dCTP \ dTTP	
YfcD	未知	16

Ευρηκα

NudE NudF YffH



		0201/0/
NudA (Mutt)	8-OH-dGTP	PYKK
NudB	2-OH-dATP \ dNTP	
NudC	NADH \ Ap ₂ A	
NudD	GDP-mannose (sugar)	
NudE	Ap₃A ` ADP-ribose	
NudF	ADP-sugar	



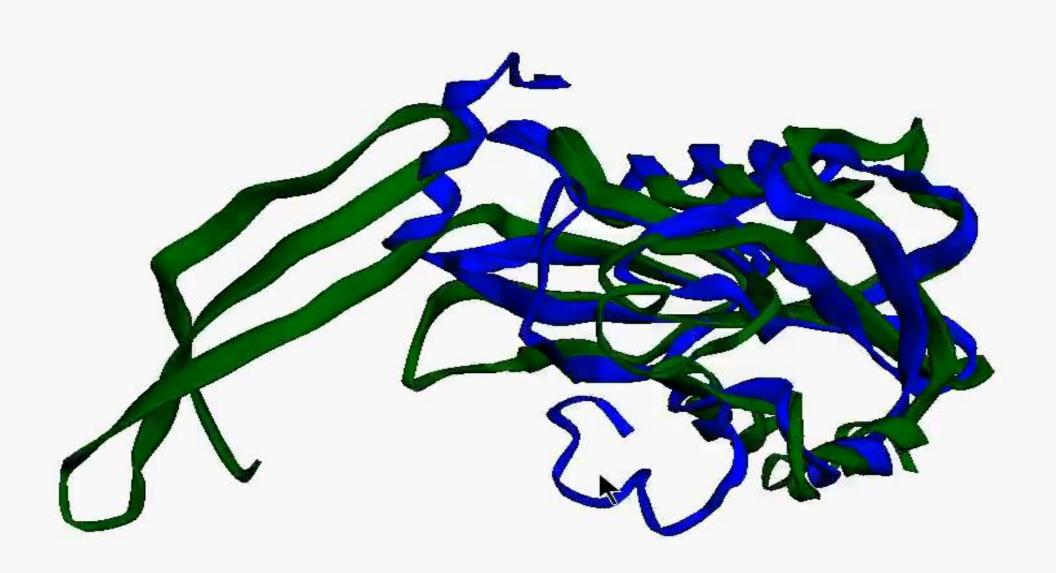
NudG	5-OH-dCTP \ 2-OH- dATP \ 8-OH-dGTP
NudH	RNA
YeaB	未知
YmfB	TPP
YffH	GDP-mannose
YfaO	dUTP \ dCTP \ dTTP
YfcD	未知

		0201/0/
NudA (Mutt)	8-OH-dGTP	pyka
NudB	2-OH-dATP \ dNTP	
NudC	NADH \ Ap2A	
NudD	GDP-mannose (sugar)	
NudE	Ap₃A ` ADP-ribose	
NudF	ADP-sugar	

是否存在錯誤? YffH、NudD看似功能相似 卻不被演化樹分在一起。

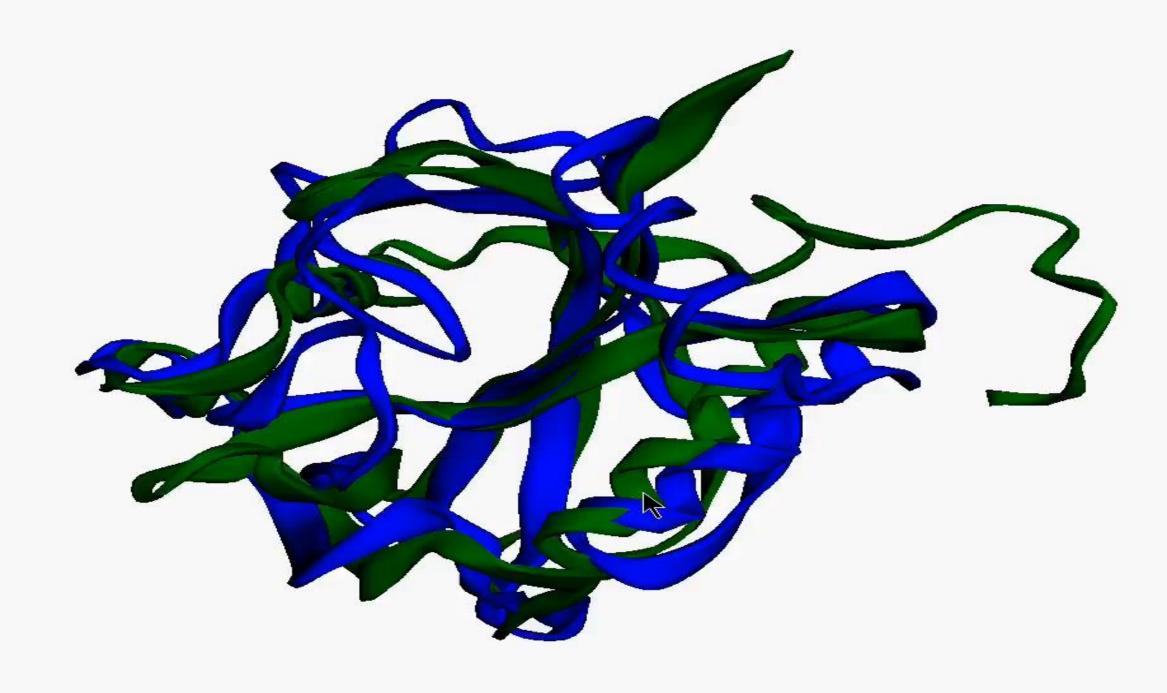
MudD MffH





MudD MnfB





假設

- LUPNKO
- 一、演化趨向多樣化
- 一、每個NUDIX是為細胞功能而演化
- 一、結構較序列演化上緩慢
- 四、類似的結構更加意味類似的功能
- 五、演化樹永遠選後設機率較高者

結論



- 一、水解RNA為共祖蛋白最有可能的功能
- 二、比對發現功能十分相似的NUDIX水解酶



未來展望



- 一、催化的點在序列的何處? (功能性)
- 二、等到有更多資料再畫一次SSN



感謝鄭貽生 教授 - 2000 KX 提供研究方向與器材

實驗室: 台大生科館1121室

特色: MTH1----癌症治療曙光

纖維素合成酶----生質能源之希冀

植物去乙醯基酶---植物對基因的調控

優勢: 可以利用結構來了解功能!

謝謝各位的時間與鼓勵!

感激不盡!

