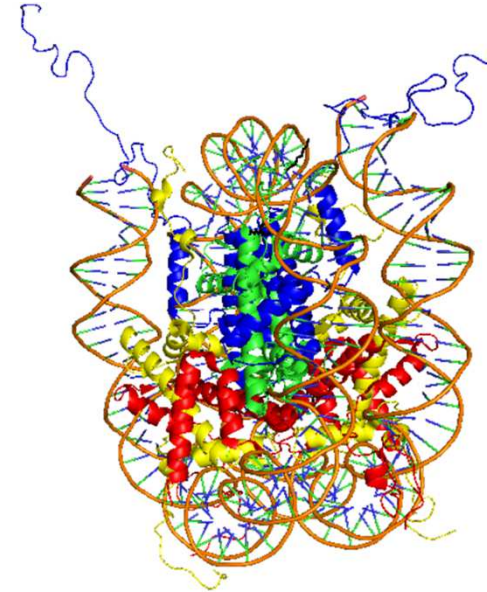


Zkouška: - test + přednáška

- Úvod - Analýza proteinu
 - Domény
 - fold-struktura (sekundární, PDB)
 - v PyMolu připravit 3D strukturu
 - Interakce (IntAct...)
 - Komplexy
 - Funkce
 - Lokalizace
 - **Evoluce** (alignment sekvencí)
- Konkrétní nová data – článek (< 5 let) o komplexu (nebo proteinu) - *konzultace*

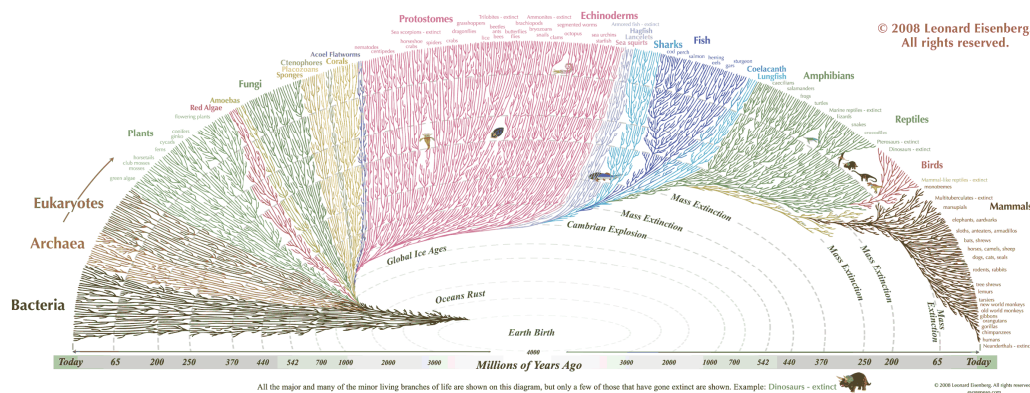


Ujasnit si souvislosti, rozšířit si znalosti, aplikovat poznatky z přednášek ...

4.05.2023 přednáška „Přehled nástrojů pro bioinformatickou analýzu komplexů“
11.05.2023 9.00 hod C2-2.11 zkouška (test + prezentace)

Evolve

- podobně jako srovnání morfologie organismů (fenomů), také srovnání genomů/proteomů ukazuje na vývoj „evoluci“ v čase (genů/proteinů)
- divergence druhů koreluje do značné míry s konzervací/divergencí DNA/proteinových sekvencí
- způsobené „nestabilitou“ DNA (... oprava poškození)
- DNA je replikována s relativně vysokou přesností (1 změna na 10^9 nukleotidů – cca4000 _{TNR12} písmen na A4 stranu – 8000 _{TNR12} na list – 500listů/balík – cca 250 balíků)
- umožňují „vylepšující“ změny (adaptace)



mutace virů ukazují jejich úlohu pro „adaptabilitu“ na prostředí ... hlavní roli však hrají jiné změny ...

Evolve – mutace

- velmi příbuzné sekvence DNA/proteinů mezi člověkem a primáty díky „krátké“ době, po kterou mohlo k mutacím docházet (jsou zachovány i nukleotidy ve 3. pozici synonymních kodonů - odhad 1/400AMK protein za 200,000 let)

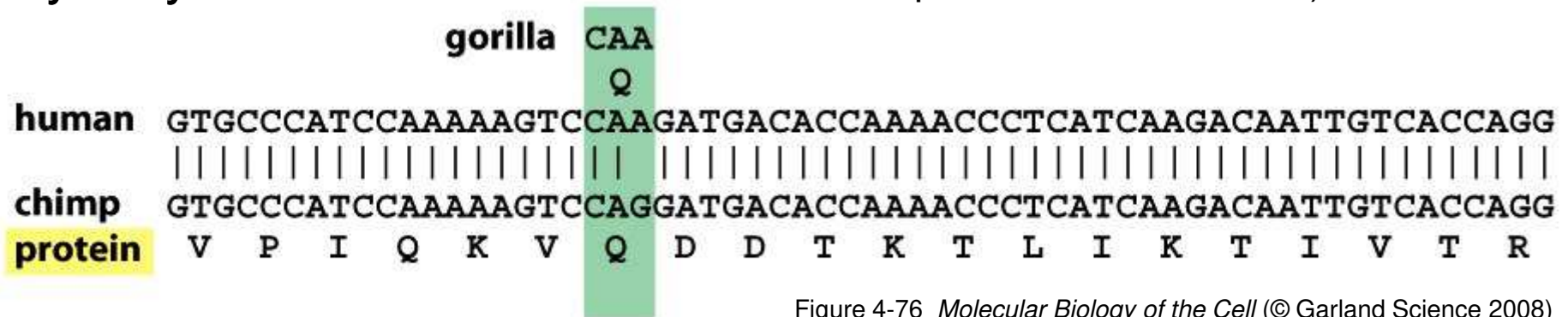
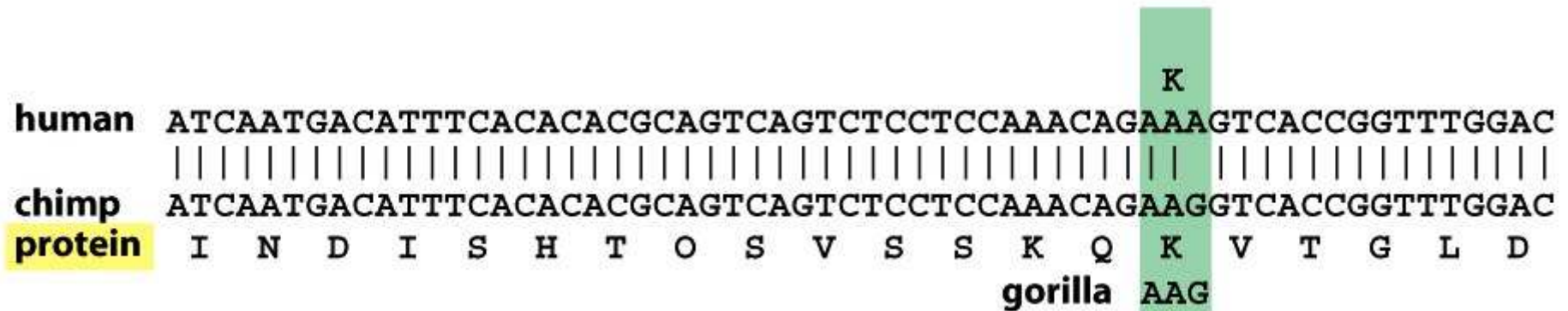


Figure 4-76 *Molecular Biology of the Cell* (© Garland Science 2008)



- druhovou rozdílnost nezpůsobuje mnoho mutací v sekvencích proteinů (nemohlo jich tolik vzniknout), ale ...

- odlišnost druhů je dána spíše rozdílnou regulací (nekódujícími sekvencemi), tj. rozdílem v **expresi proteinů** než rozdílem v sekvenci proteinů (tj. rozdílnou funkcí proteinů)
- rozdílná exprese tj. rozdílné proteomy v buňkách podmiňují odlišnost buněk v organismu (svaly, játra ...) i odlišnost buněk v čase a prostoru (morfogeneze – odlišný vzhled, vývoj ...
mozek => menší problém když se modifikuje/zmutuje program morfogeneze než když zmutuje protein => vliv na funkci)

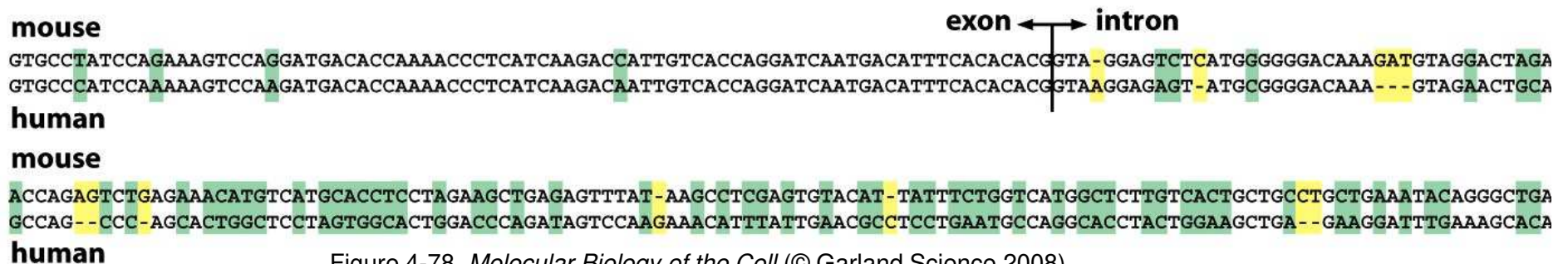
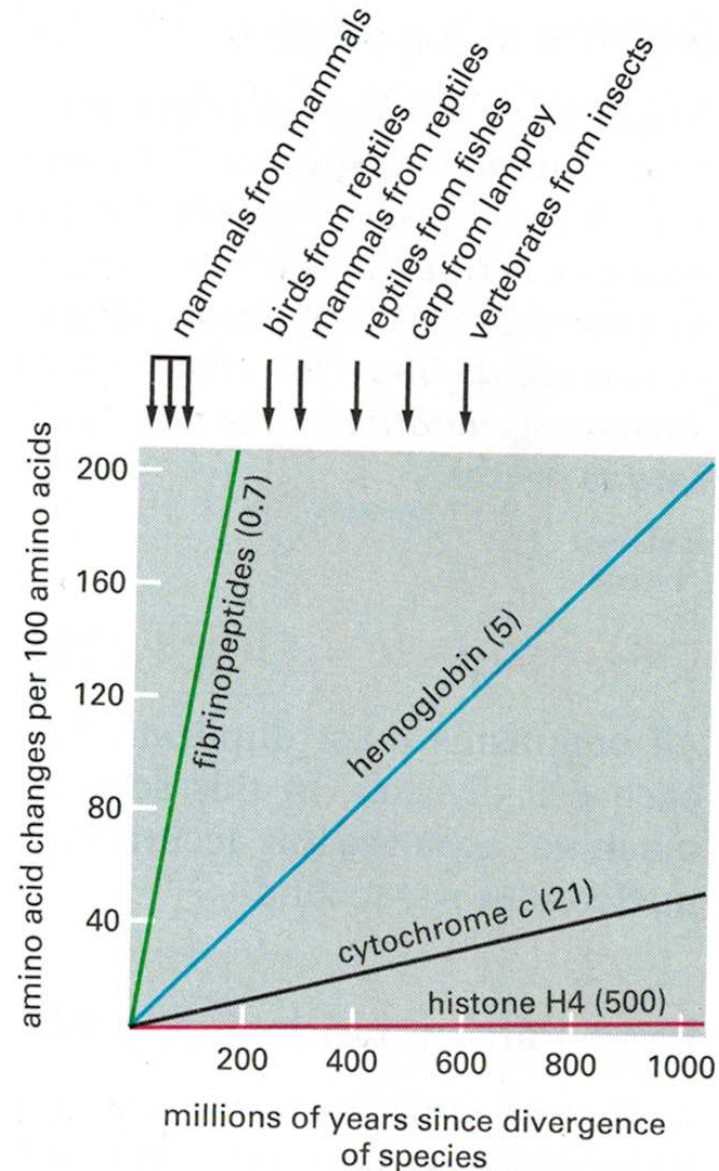


Figure 4-78 *Molecular Biology of the Cell* (© Garland Science 2008)

- srovnání sekvencí (člověk a myš) ukazuje odlišnější **nekódující sekvence** než protein-kódující (regulační sekvence jako promotor ..., intron je odlišnější než kódující exon) - silnější selekční tlak na proteiny

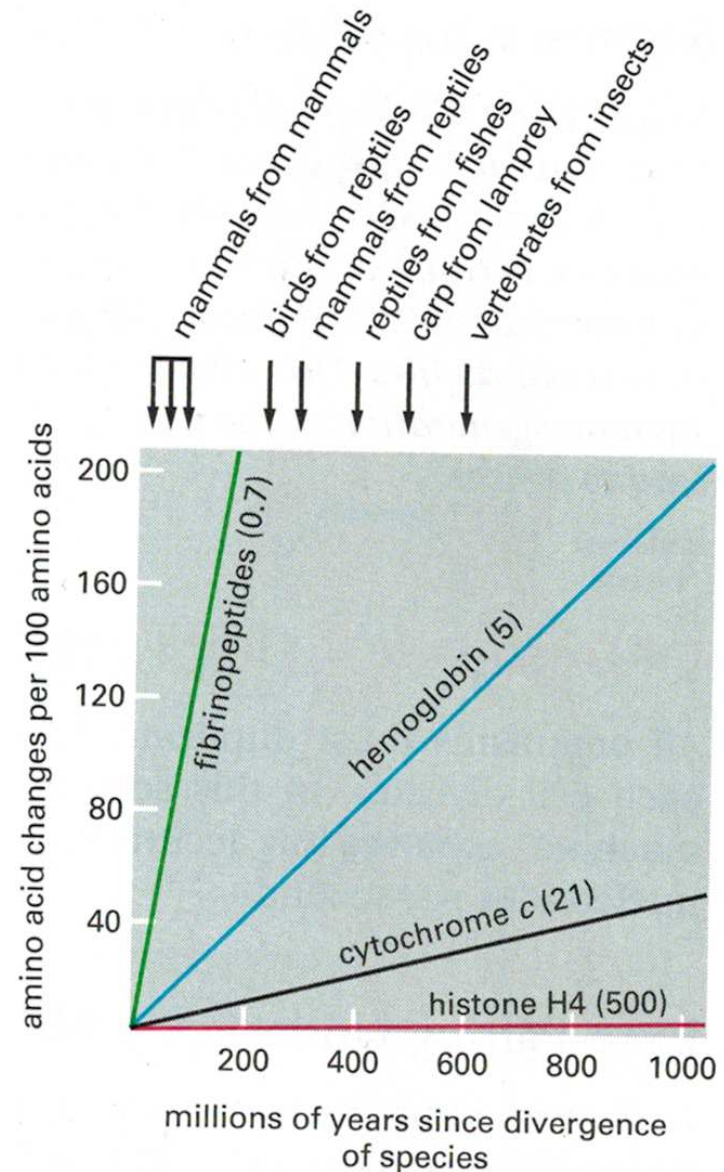
Evolve – selekce

- frekvence mutací by byla +/- stejná podél molekuly DNA ... ale různé proteiny jsou různě změněné díky **selekčním tlakům**
 - (histony ... kvasinkový a lidský ubikvitin se liší třemi AMK, 6 ze 7 změn v cytochromu C jej poškodí)
 - čím důležitější a komplexnější funkce proteinu pro organismus (od nejzákladnějších po specializované) tím více je konzervovaná sekvence proteinu (respektive jeho domén)



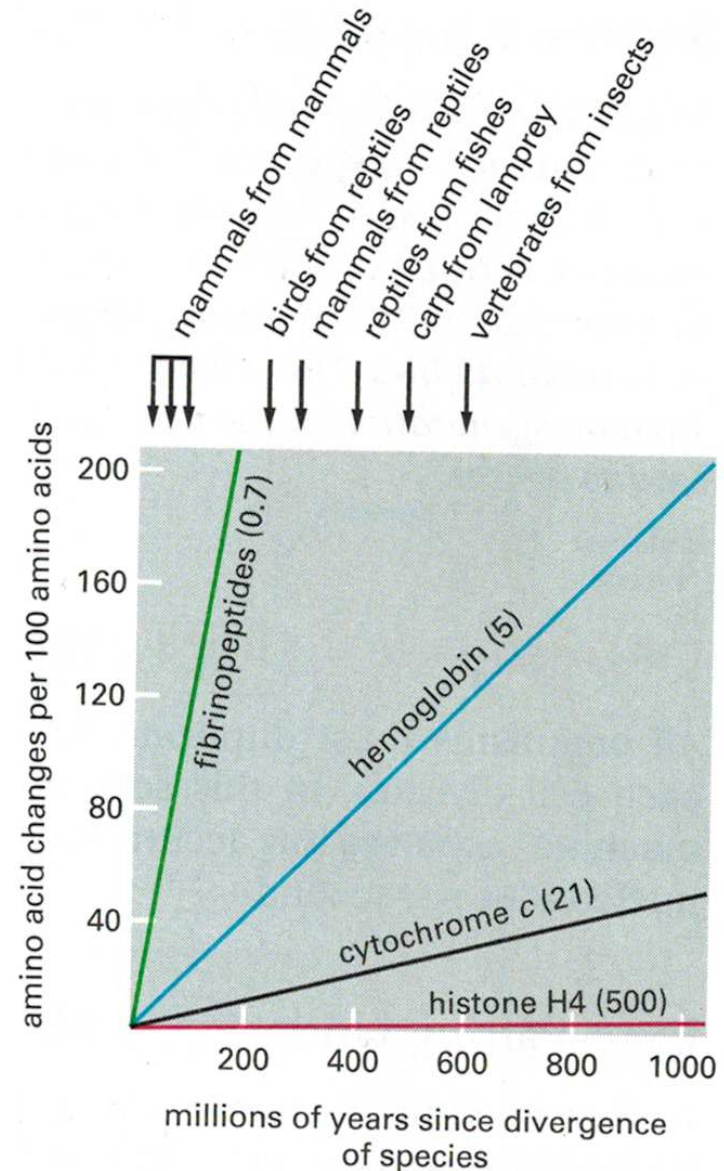
Evolve – selekce

- takto konzervované proteiny jsou lehce identifikovatelné v různých organismech a mají „homologní“ (**ortologní**) funkci
- např. některé lidské proteiny funkčně zastoupí zmutované/deletované kvasinkové proteiny (Y2H cytotrap)
- konzervace je důsledkem „purifikační selekce“ (eliminace buněk/jedinců s mutacemi v esenciálních/důležitých funkcích)
- mutace u pacientů s různými syndromy ...



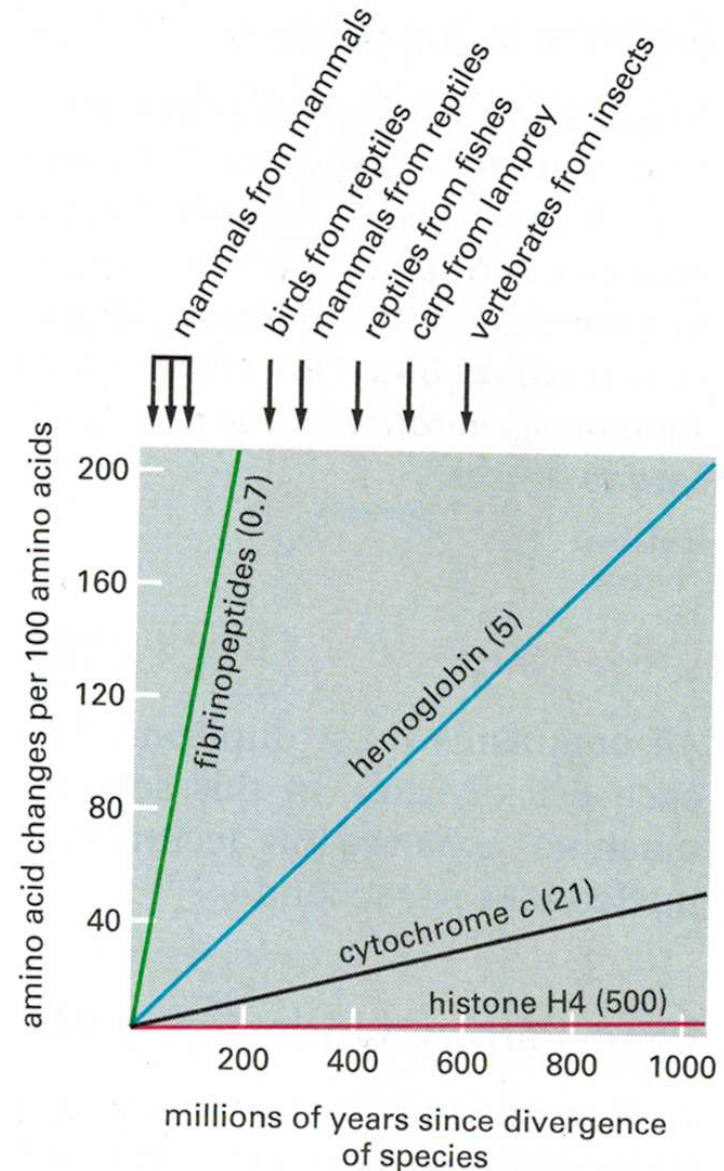
Evolve – selekční tlak

- purifikační tlak na proteinové sekvence pro udržení jejich funkce:
- „tlak“ na **strukturu** proteinu (více mutací ve smyčkách),
- selekční tlak na povrch proteinu = tj. na **interakce** a **PTM**
- mutace, které „neruší“ jsou **neutrální** (protein je částečně modifikován)
- modifikace je kompenzována (i později) mutací partnera
- změní se v čase a v budoucnu může přinést novou vlastnost



Evolve – selekční tlak

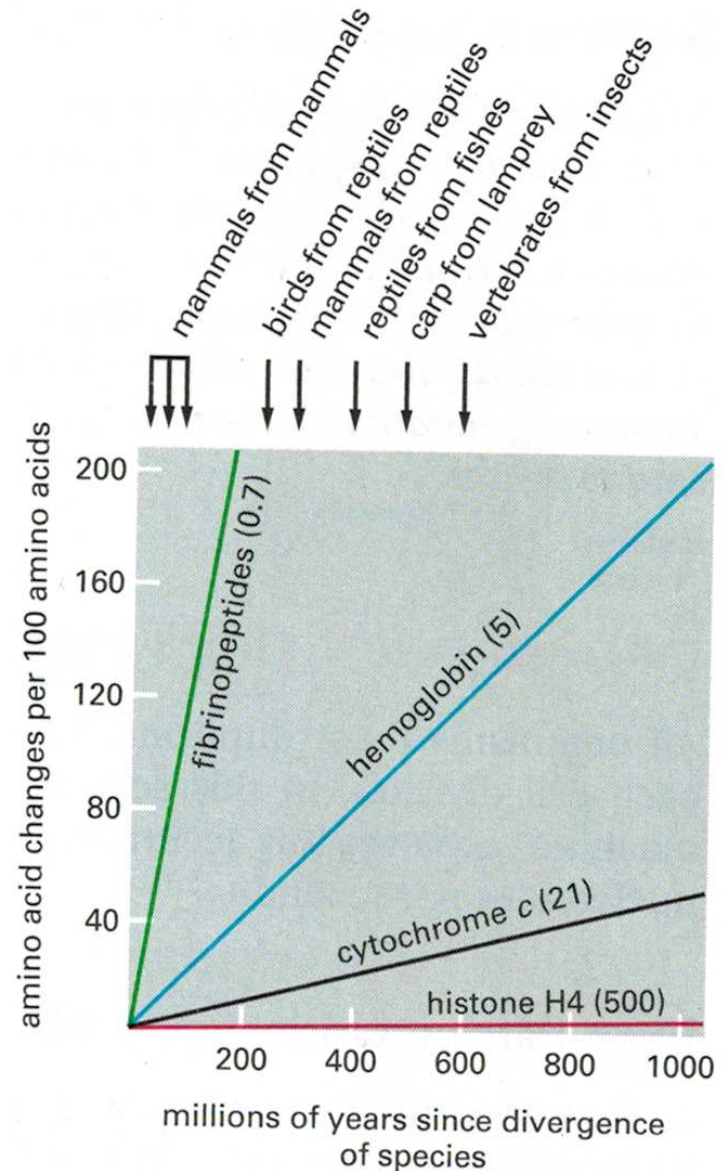
- purifikační tlak na proteinové sekvence pro udržení jejich funkce:
- „tlak“ na **strukturu** proteinu (více mutací ve smyčkách),
- selekční tlak na povrch proteinu = tj. na **interakce** a **PTM**
- nejpomaleji se mění proteiny, které jsou zapojeny do nejvíce interakcí s dalšími proteiny (limitována, jak struktura, tak povrch) - není příliš prostoru pro změny - např. ubikvitin, histony ... („drží“ základní systém)



Evolve – selekce

- nutnost zachování funkce neposkytuje příliš prostoru pro evoluci/rozvoj nových vlastností (ale neznamena ani selekci na „nejstabilnější“ či „nejaktivnější“ = určitá volnost)
- neutrální mutace neruší strukturu či interakce proteinu
- taková změna ale může v budoucnu přinést novou vlastnost

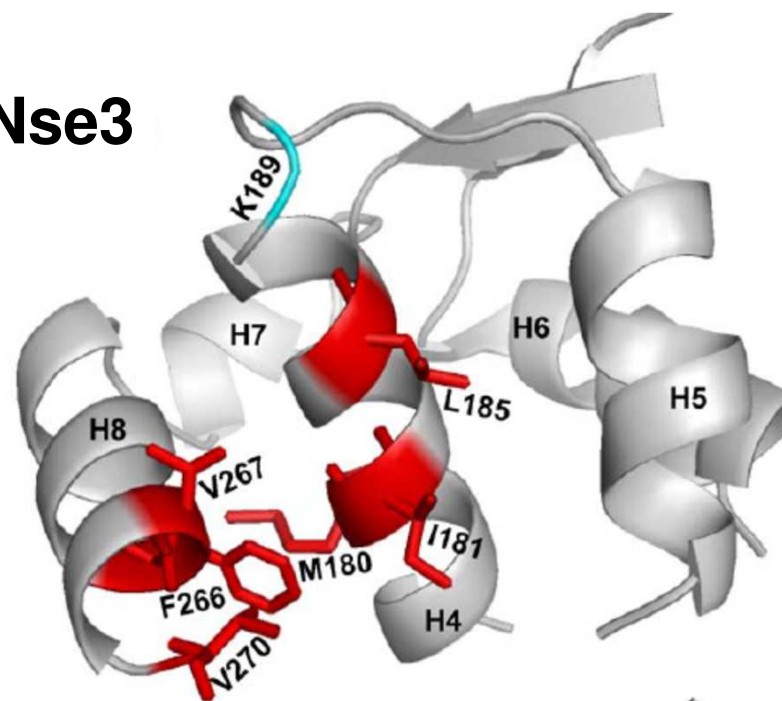
v alignmentu se podívejte na pozice konzervovaných AMK v 3D modelu/struktuře



blízke organismy

- mutace specifické AMK na Ala zruší Nse3-Nse4 interakci
- „mutace“/změny těchto AMK na podobné (hydrofobní) AMK tyto interakce neruší

Nse3



Hudson et al, PLoS One, 2011

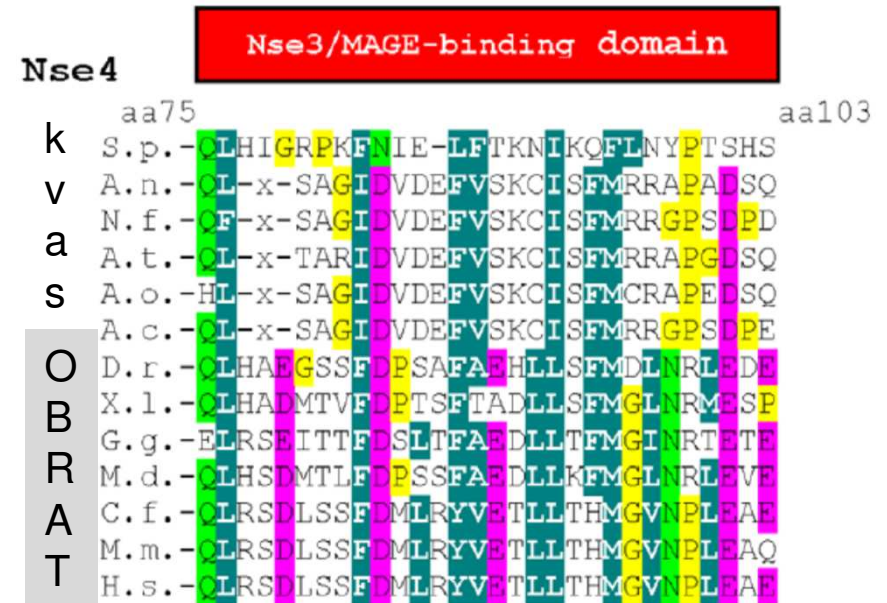
...RYBY, OBOŽIVELNÍCI, PTÁCI ...

N.v.	GLLMLVLSVILM-SxxxxN-YTSLWHFLKMGLEPKKEHEVFGDP--EKLIACEFTROGYLERRKVTGGE--EATEEYSW--GSRSNKE---LTKRKVLEFVFS:	
D.r.	GLLFVILSVIFM-KGGTIK-ENLVWNTLKKLRDPGEKHDEFQDV--KKVVTEEFVROKYLBYGKIPHTE--PVEYEFRW--GLRAEKE---VSKLKILLEFVG:	
T.n.	GLLFVILSVIFM-KGGAVR-DSVVMNLLKKLRVQPGERHEPEFGEV--KRVVMEEFVRORYLBCNRIPHTE--PLEHEFRW--GQRADTE---VSKTKILEFMA:	
X.t.	GLLMVILSLIFM-KGNTAK-ESAVWEMLRRLRIEPAEKHSDFQDV--KKLITEEFVKOKYLYSKVLHTD--PVEYEFRW--GQRAFKE---TSKMQVLEFVS:	
G.g.	GLLIVILSFIFM-KGNSAK-DSAVWDFLRRLRVHPGEKHEVFGDV--KKLVMEEFVROKYLHITPIPLTD--PPRFNFQW--GPRAAKE---TSKKDILSFVA:	
O.a.	GLLMVILSLIFM-KGSATN-ESVIWDFLRRLRVDTRERHEVFGDV--KKLVTEEFVROKYLBYNRRIPHTE--PVEYEFQW--GARATKE---TTKMQVLEFVA:	
M.d.	GLLMVILSLIFM-KGNSAR-ESLVWDFLRRLRVDPEKRHKTFQDV--KKLVKDEFVROKYLBYIRVPHSE--PPEYEFQW--GPRAAHE---TSKMQVLEFVA:	
E.e.	GLLMIVLGLIFM-KGNTIK-ETEVMDFLRRLGVYPTKKHLIFGDP--KKLITEEFVRORYLBYRRIPHTD--PVDYELQW--GPRTNLE---TSKMKVLEFVA:	
S	Q.p.	GLLMIVLGLIFM-KGNTVK-ETEVMDFLRRLGVYPTKKHLVFGDP--KKLITDFVRORYLBYRRIPHTD--PVDYELQW--GPRTKLE---TSKMKVLEFVA:
A	M.cr	GLLMIVLGLIFM-KGNTIK-ETEVMDFLRRLGVYPTKKHLVFGDP--KKLITDFVRORYLBYRRIPHTD--PVDYEFQW--GPRTNLE---TSKMKVLEFVA:
V	B.t.	GLLMIVLGLIFM-KGNSIK-ETEVMDFLRRLGVHPTKKHLIFGDP--KKLITDFVRORYLBYRRIPHTD--PVDYELQW--GPRTNLE---TSKMKVLEFVA:
C	Ev.c.	GLLMIVLGLIFM-KGNTIK-ETEVMDFLRRLGVSPPTKKHLIFGDP--KKLITDFVRORYLBYRRIPHTD--PVDYEFQW--GPRTNLE---TSKMKVLEFVA:
	C.f.	GLLMIVLGLIFM-KGNTIK-ETEVMDFLRRLGVYPTKKHLVFGDP--KKLITEEFVRORYLBYRRIPHTD--PVDYEFQW--GPRTNLE---TSKMKVLEFVA:
	H.c.	GLLMIVLGLIFM-KGNTIT-ETEVMDFLRRLGVYPTKKHLIFGDP--KKLITDFVRORYLBYRRIPHTD--PVDYELQW--GPRTNLE---TSKMKVLEFVA:
	Mmag1	GLLMIVLGLIFM-KGNTIT-ETEVMDFLRRLGVYPTKKHLIFGDP--KKLITDFVRORYLBYRRIPHTD--PVDYELQW--GPRTNLE---TSKMKVLEFVA:
	Ml--	GLLMIVLGLIFM-KGNTVK-ETEVMDFLRRLGVYPTKKHIFGDP--KKLITDFVRORYLBYRRIPHTD--PVDYEFQW--GPRTNLE---TSKMKVLEFVA:
	P.a.	GLLMIVLGLIFM-KGNTIK-ETEVMDFLRRLGVYPTKKHLIFGDP--KKLITDFVRORYLBYRRIPHTD--PVDYEFQW--GPRTNLE---TSKMKVLEFVA:
	P.t.	GLLMIVLGLIFM-KGNTIK-ETEVMDFLRRLGVYPTKKHLIFGDP--KKLITDFVRORYLBYRRIPHTD--PVDYEFQW--GPRTNLE---TSKMKVLEFVA:
	H.s.	GLLMIVLGLIFM-KGNTIK-ETEAWDFLRRLGVYPTKKHLIFGDP--KKLITDFVRORYLBYRRIPHTD--PVDYEFQW--GPRTNLE---TSKMKVLEFVA:

„kompenzační“ mutace?

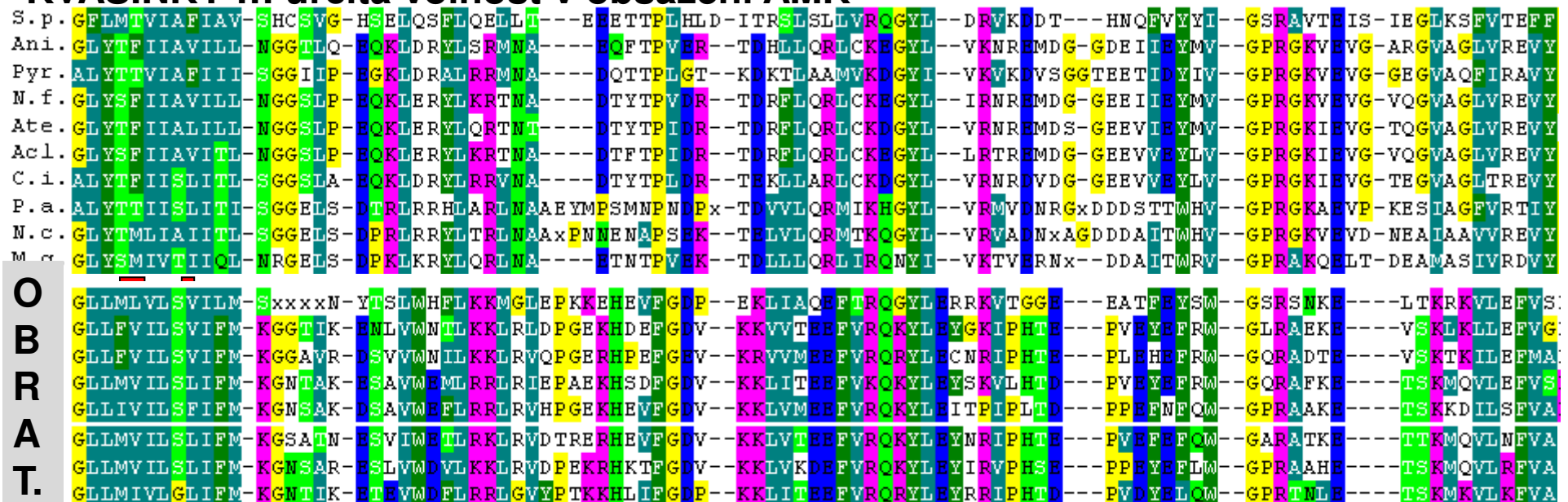
- „mutace“/změny těchto AMK na odlišné (polární Thr nebo posun motivu o 1AMK na konci)
- pořad vytváří stejný komplex (interakční partner se také „mění“)
- „mutace“/změny zřejmě koevolvuji
- **teorie kompenzačních mutací ...**

NSE4 subfamilies

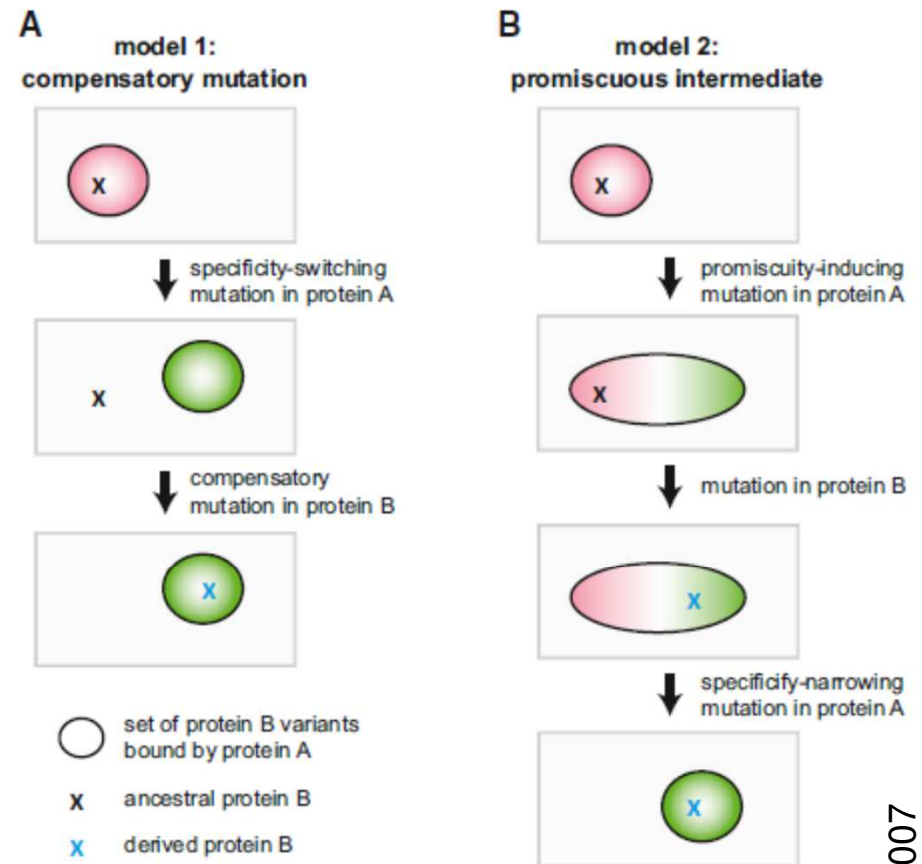
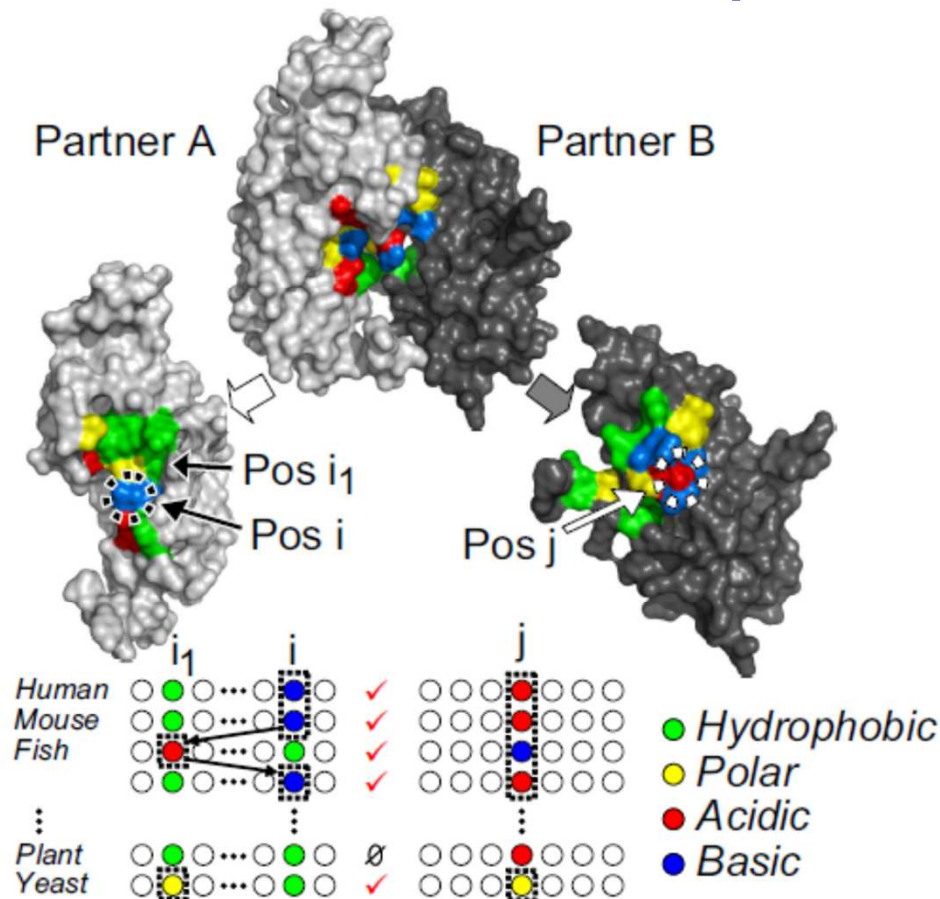


Guerineau et al, PLoS One, 2012

KVASINKY ... určitá volnost v obsazení AMK



Vazební partneři ko-evoluují

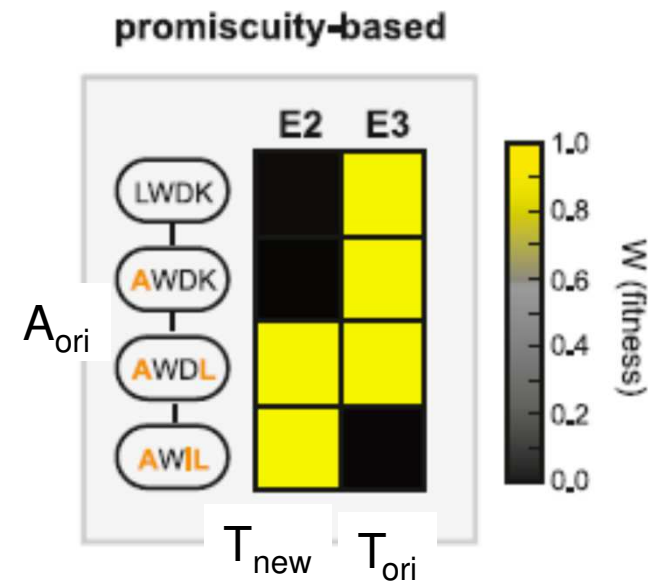
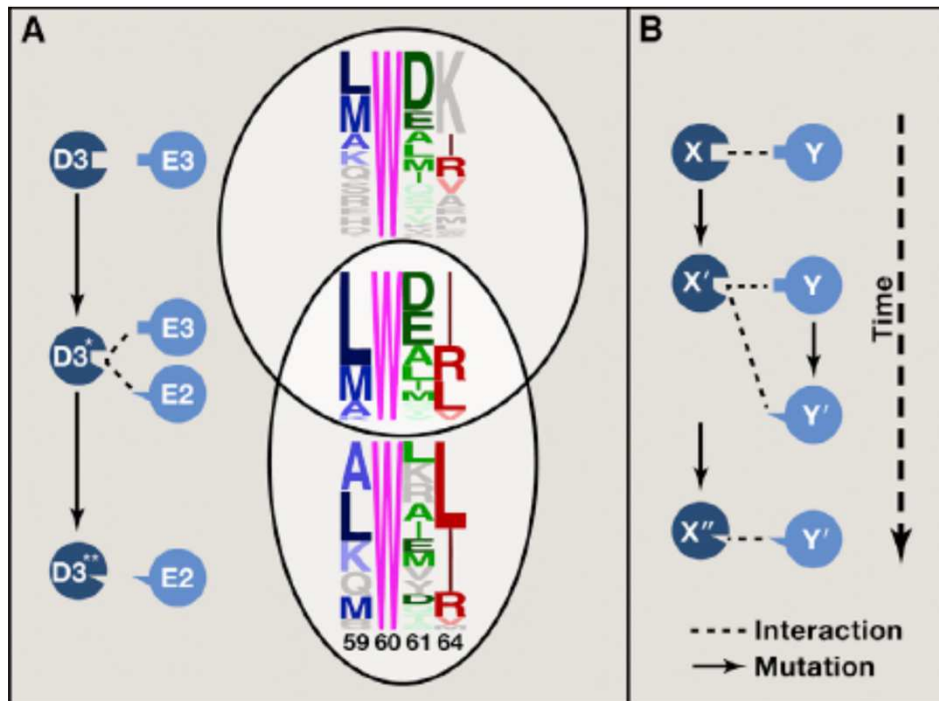


Aakre et al, Cell, 2016

- nutnost zachování funkce nesevďčí o „kompenzačních mutacích“ (mutace v jednom z proteinů přímo kompenzována mutací v partnerském proteinu) – „kompenzace“ přichází postupně přes „**promiscuous intermediate**“ mutace

„promiskuitní“ mutace

- „promiscuous intermediate“ může rozšířit interakční spektrum = může interagovat i s **duplikovaným** proteinem (např. tkáňově specifickým – specifický komplex)
- později se může „oddělit“ a vytvořit nový komplex (paralelní ko-evoluce = drift)



Vznik proteinových rodin

- selekční tlak na stabilitu a funkci - nutnost zachování funkce neposkytuje příliš prostoru pro evoluci/rozvoj nových vlastností
- pro rozvoj nových vlastností, nových druhů - spíše než druhově specifické mutace proteinů lze vidět **expansi různých genových/proteinových rodin** v různých živočišných druzích (různé kopie mají různé funkce - **paralogy**)

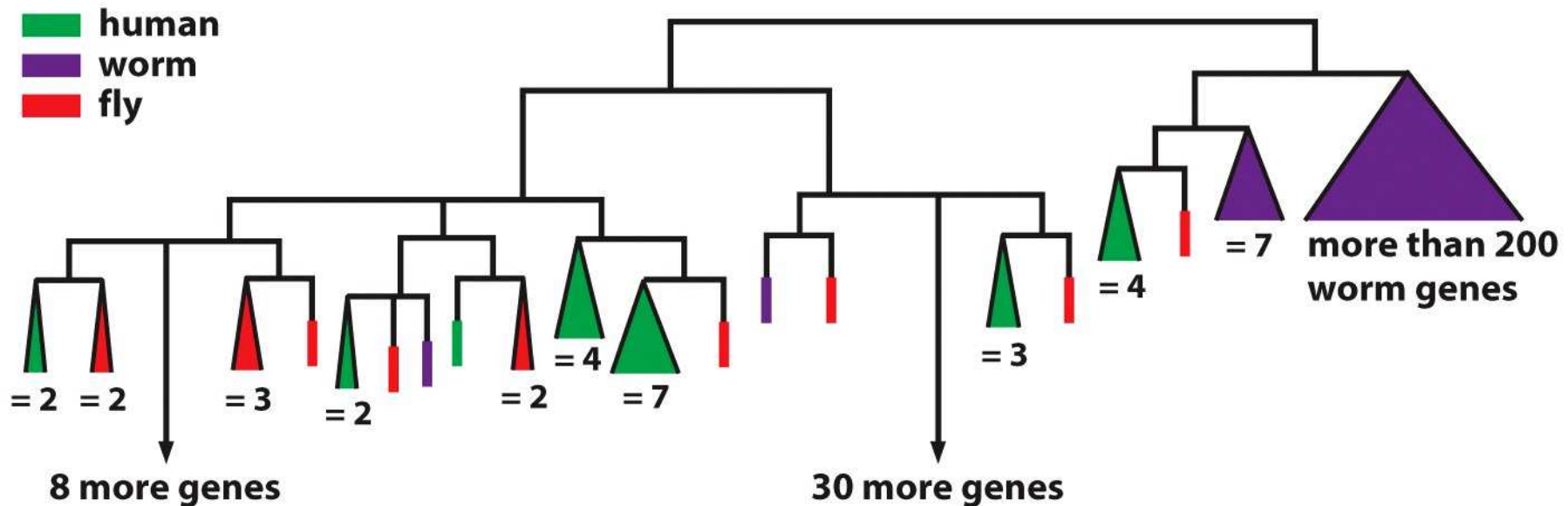
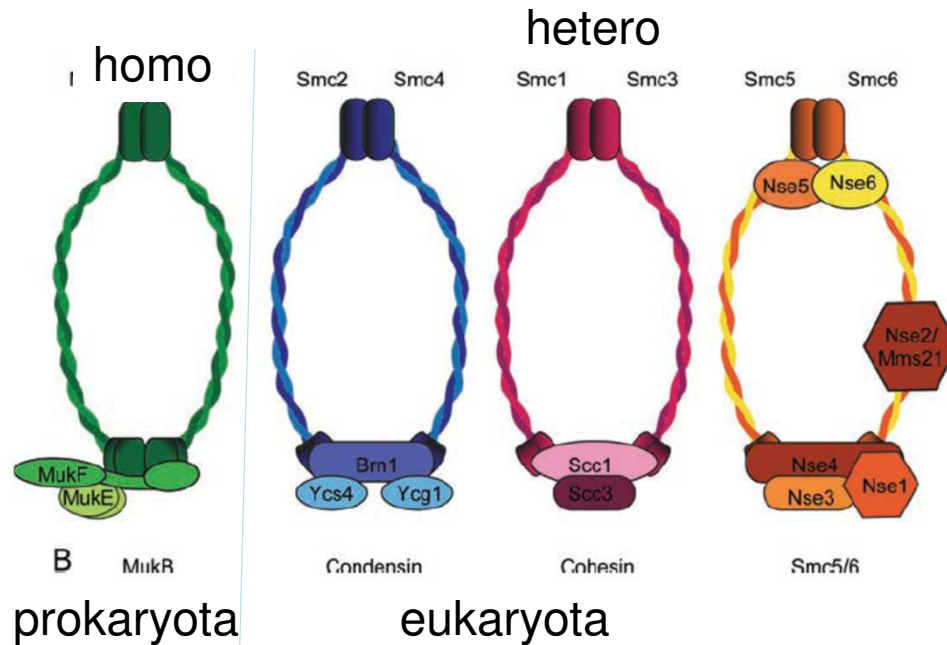


Figure 4-85 *Molecular Biology of the Cell* (2008)

(příklad superrodiny jaderných hormonálních receptorů)

Duplikace a divergence (neo- a sub-funkcionalizace)

- hlavními tahouny evolučních procesů jsou **duplikace** (všechny geny/proteiny jsou „potomky“ několika ancestrálních genů/proteinů (foldů), které existovaly v nejčasnějších živých formách
- (nyní cca 1000 foldů na >100000 struktur v PDB, odhad je cca 2000 foldů))

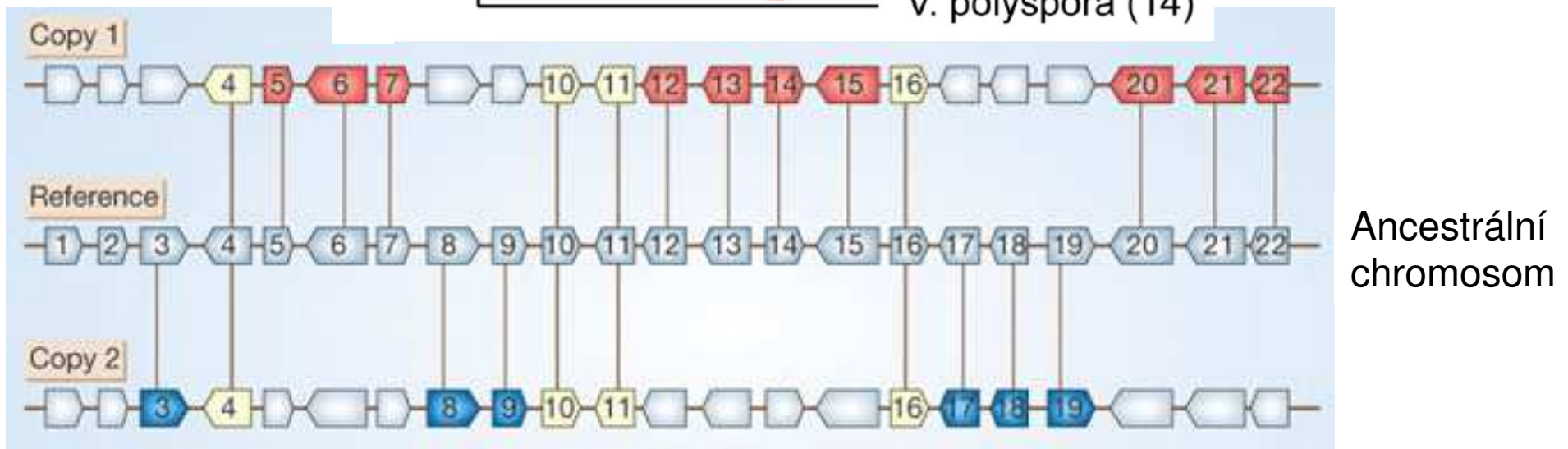
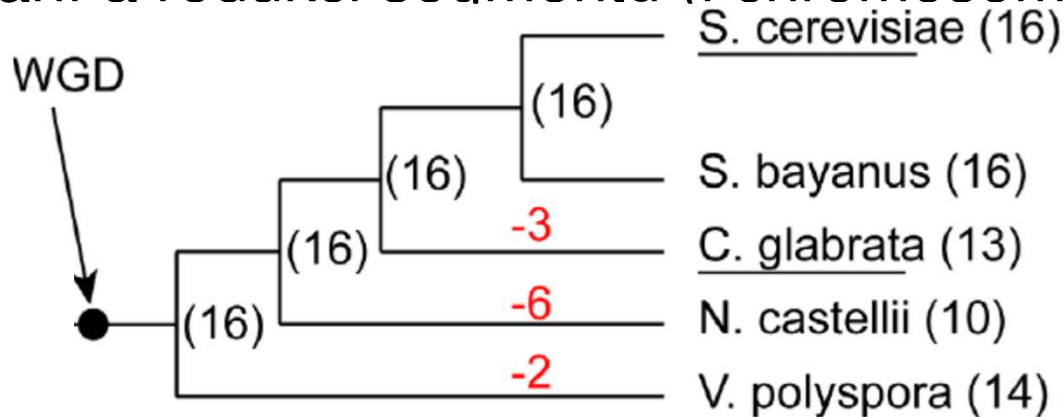


Marsh et al, ARB, 2015

- po duplikaci jsou oba proteiny stejné - vytváří stejný homo(di)merní komplex – později jeden protein diverguje (mutace) a vzniká hetero(di)mer
- duplikace více genů/proteinů (podjednotek) ... komplexů

Celogenomová duplikace (u kvasinek)

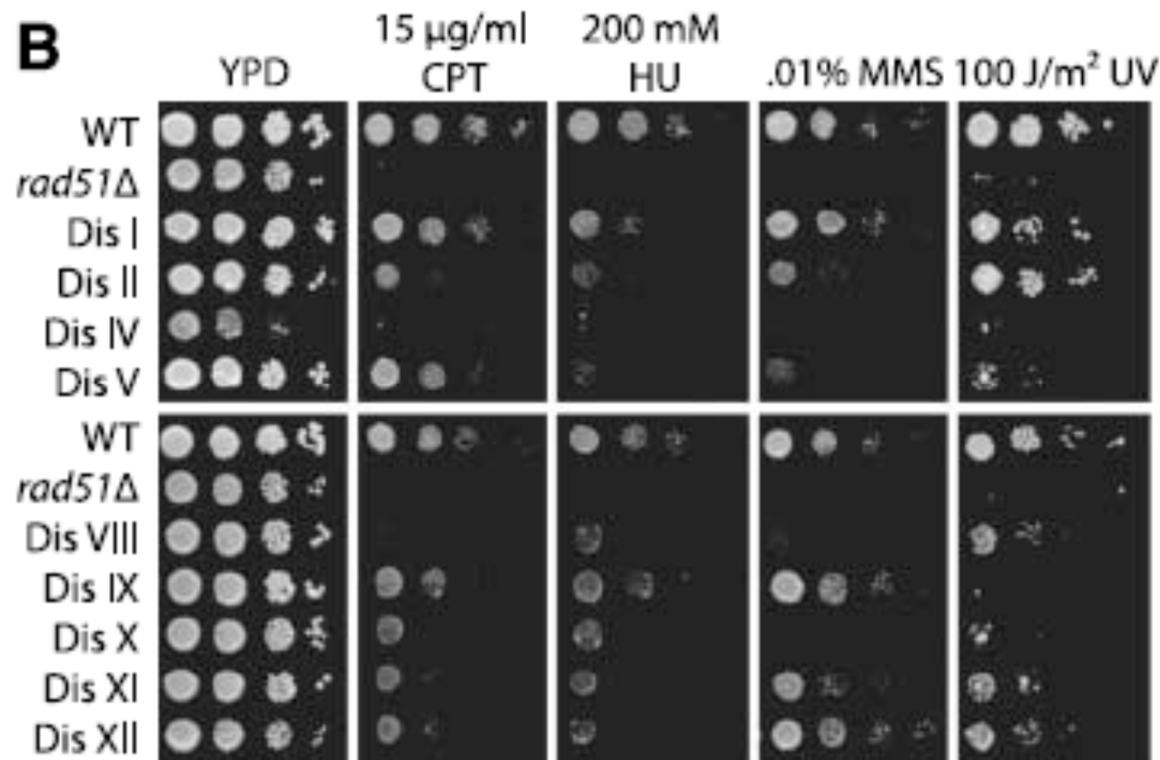
cca 30% genomu *S.c.* je „duplikováno“ => došlo k **celogenomové duplikaci** (WGD) => a poté došlo k přeskupování a redukci segmentů (i chromosomů) – polyploidie u kyttek ...



- následují **mutace** – inaktivující tj. **pseudogeny** (ustaví hladinu proteinu zpět na původní) nebo nefunkční (zatěžují expresní-chaperonový aparát – snaha odstranit)

„duplikace“ chromosomu - aneuploidie

- duplikace ALE ... – na počátku stejné sekvence = stejné funkce - vyšší hladina proteinu/ů může být toxická
- (**aneuploidie** – kvasinky s 1 chromosomem navíc, nádorové buňky)

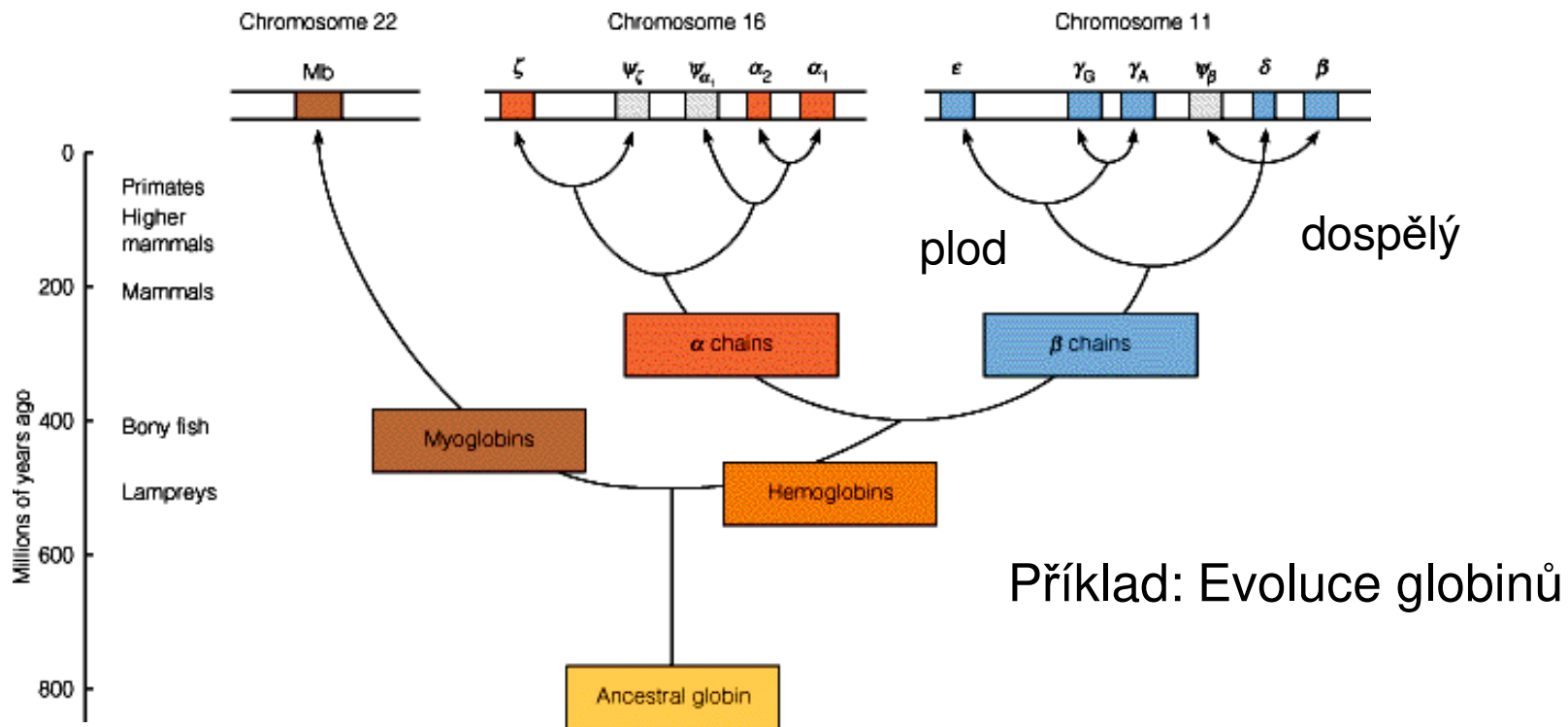


- nadbytek proteinů (např. transkripčních faktorů) nebo nerovnováha podjednotek komplexů znamená disregulaci některých procesů

Duplikace segmentů a genů

- duplikace segmentů: např. rozdíl v genomu člověka a šimpanze je především v duplikaci 30Mbp – cca 50kbp segmenty - nikoli v jednotlivých mutacích tj. SNPs
- duplikace individuálních **genů** - reverzní transkripcí a integrací DNA
- duplikace domén = exonů – přidají se do sekvence genu = protein s novým uspořádáním domén
-

Duplikace reverzní transkripcí a integrací DNA



- duplikované geny mají na počátku stejné sekvence = stejné funkce (pod jinými **promotory** – jiný lokus tzn. jiné „okolí“ zaintegrované „mRNA“ tzn. jiná regulace exprese – „nové“ buňky) – postupné mutace a další odlišení

Evoluce rodiny globinů (duplikace)

- červy, hmyz a primitivní ryby mají jeden globin (150AMK)

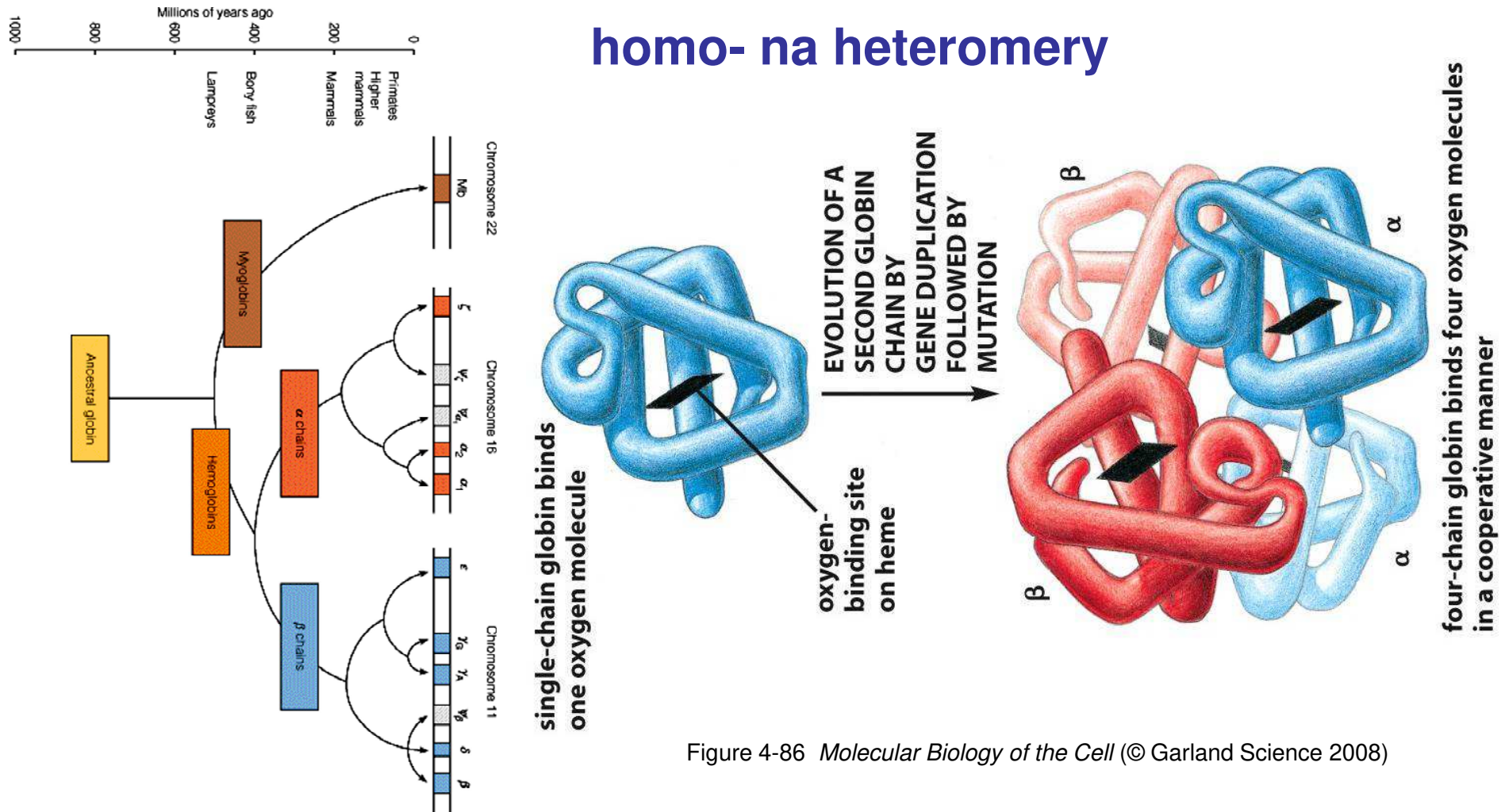
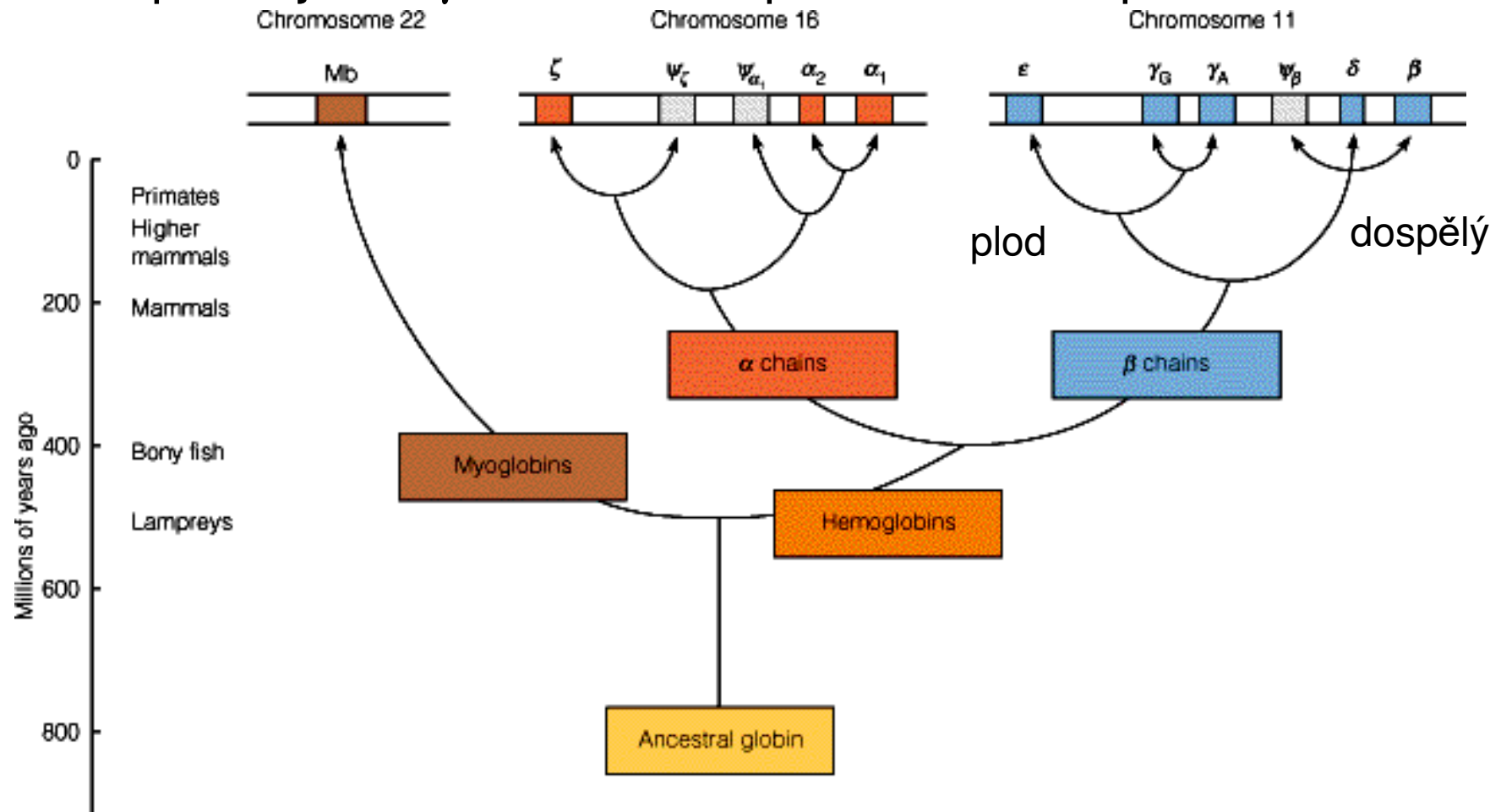


Figure 4-86 *Molecular Biology of the Cell* (© Garland Science 2008)

- vyšší obratlovci mají α - a β -globin (tvoří $\alpha_2\beta_2$ komplex – účinnější přenos kyslíku) – vyvinul se u vyšších ryb (500Mya)

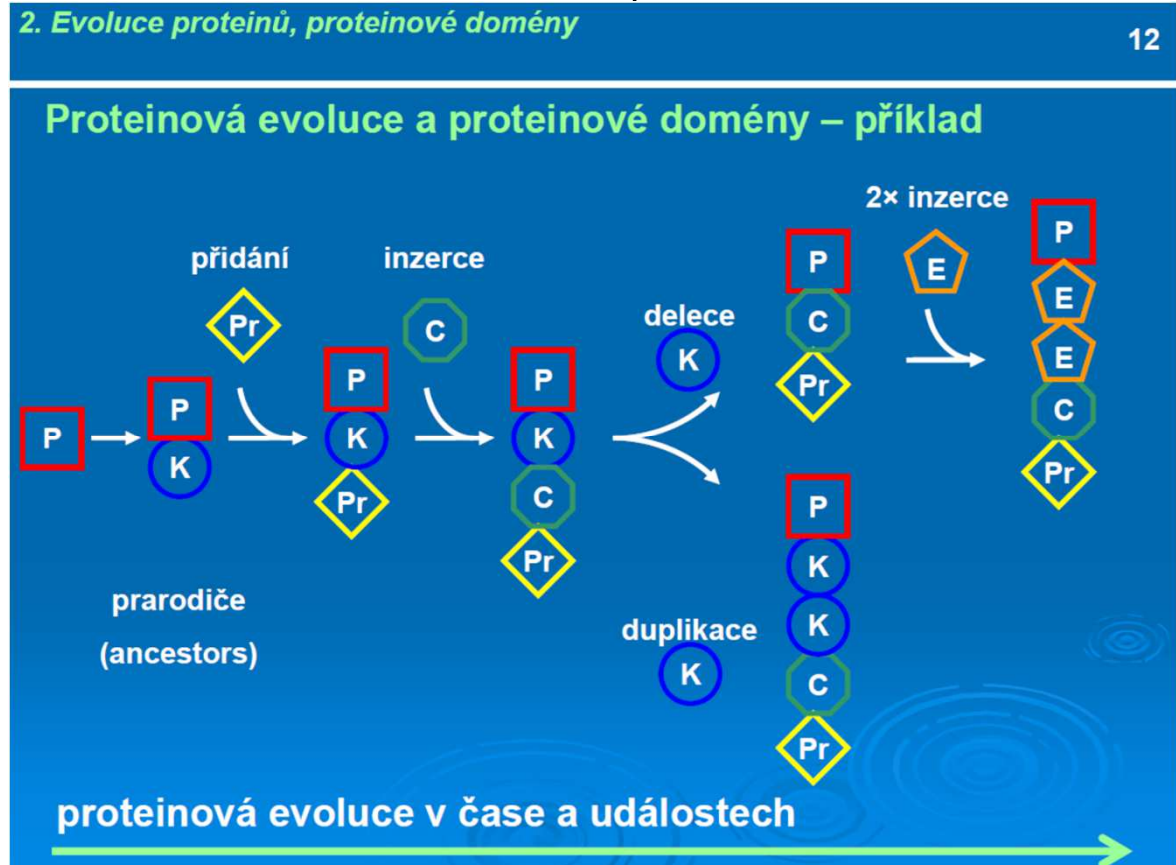
Evoluce rodiny globinů (duplikace)

- ... u savců se dále duplikoval β -globin, který je exprimován specificky v embryu – má vyšší afinitu ke kyslíku a napomáhá přenosu kyslíku z krve matky do krve plodu
- dále se duplikoval a specializoval na časná vývojová stádia $\alpha_2\varepsilon_2$ a pozdější $\alpha_2\gamma_2$ - k další duplikaci došlo u primátů $\alpha_2\delta_2$



Shuffling domén

- hranice domén jsou většinou kódovány **introny** – bez intronů by bylo obtížnější přesně vybrat pouze „doménovou“ část
- „**shuffling**“ **domén** (inzerce/delece) – poskládají se nové geny/proteiny – vytváří nová funkční/fyzická provázání (interakce => nové „prosítování“ interaktomů)



Duplikace domén

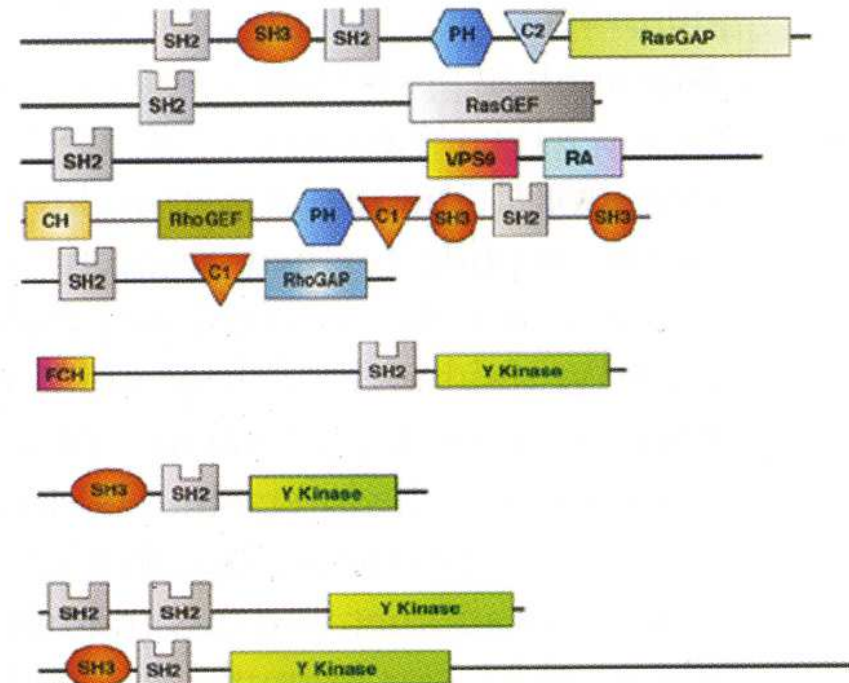
- většina domén v proteinech obratlovců pochází/existuje v bezobratlých – pouze 7% lidských proteinů/domén je specifických pro obratlovce)
- fúze domén naznačuje funkční příbuznost nebo dokonce přítomnost v komplexu (v organismu s nefúzovanými doménami - ChimeraDB)

Small GTPase
Signaling

Ras-GAP
Nsp1,2,3
Rin1
Vav1,2,3
Chimerin

Kinases

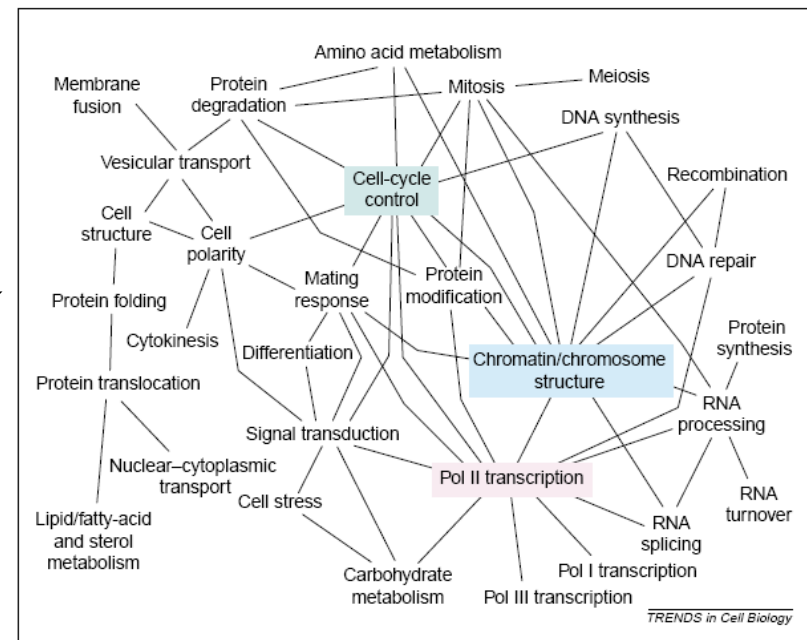
Fps, Fer
Src, Csk, Ctk/Hyl,
Fgr, Fyn, Yes, Hck,
Lck, Lyn, Blk, Frk,
Brk, DJ697K14.1
Zap70, Syk
c-Abl, Arg/Abl2



Evoluce interaktomu

- WGD vede více k redundanci a sub-funkcionalizaci
- WGD zachovává duplikáty komplexů (kvůli zachování stechiometrie podjednotek komplexu – při genové erozi se ztrácí více geny mimo komplexy)...
- více-podjednotkové komplexy jsou pod větším tlakem (musí zároveň „zachovat“ více vazeb „najednou“) než méně-podjednotkové komplexy

- duplikace jednotlivých genů vedou více k diverzifikaci funkce
- duplikace domén (shuffling) vede k neo-funkcionalizaci a k přemodulování (**rewiring**) interaktomu

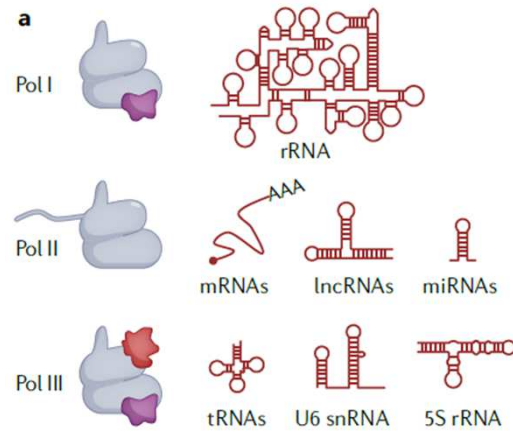


RNA polymerasy

Pol I = 11 + A49/34.5

Pol II = 11 podjednotek (Tabulka)

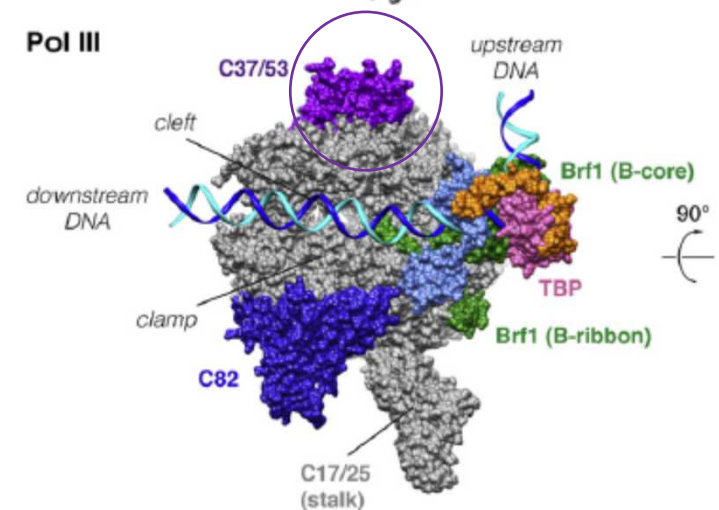
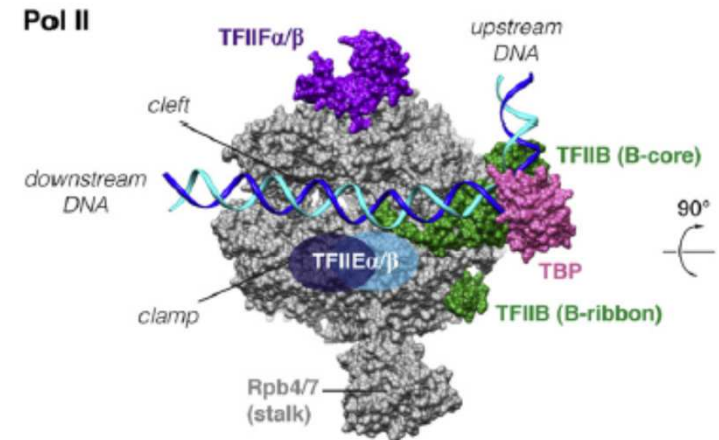
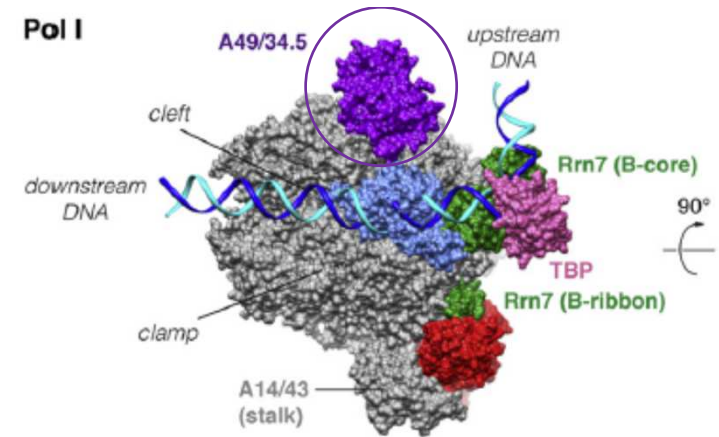
Pol III = 11 + C37/53 + C82/34/31



TFIIF
TBP
TFIIB
TFIIE

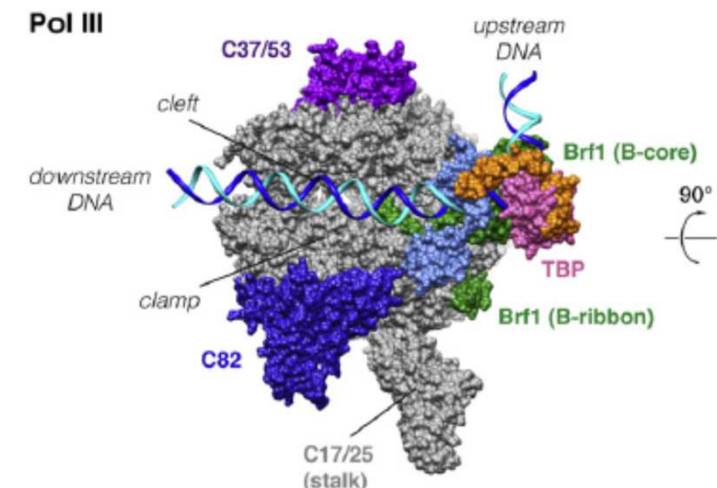
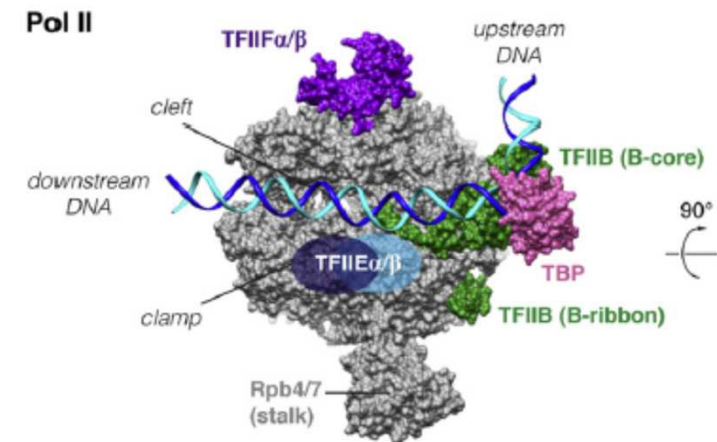
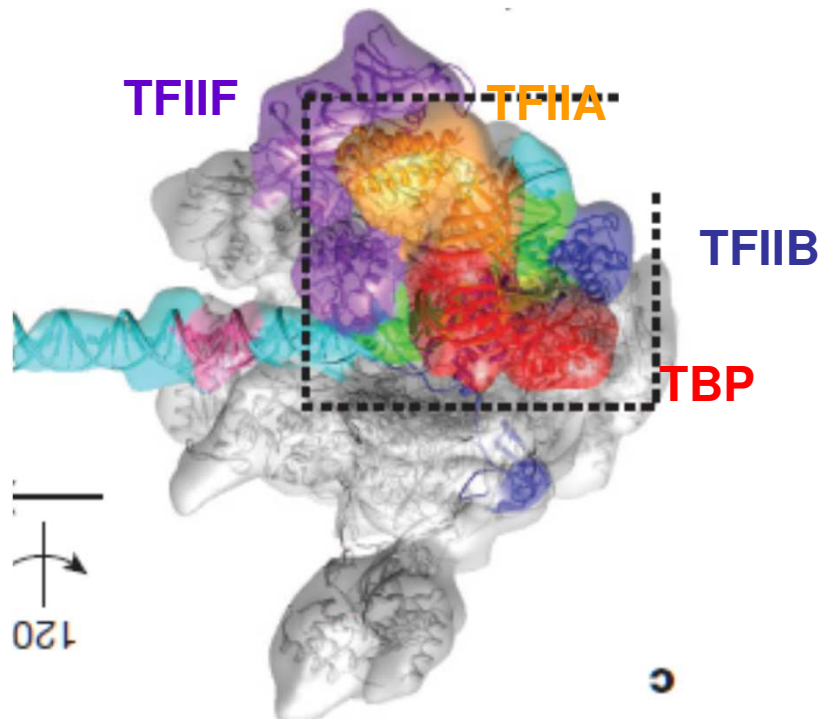
Table 1. Yeast RNA Polymerase Subunits and Initiation Factor Homologies

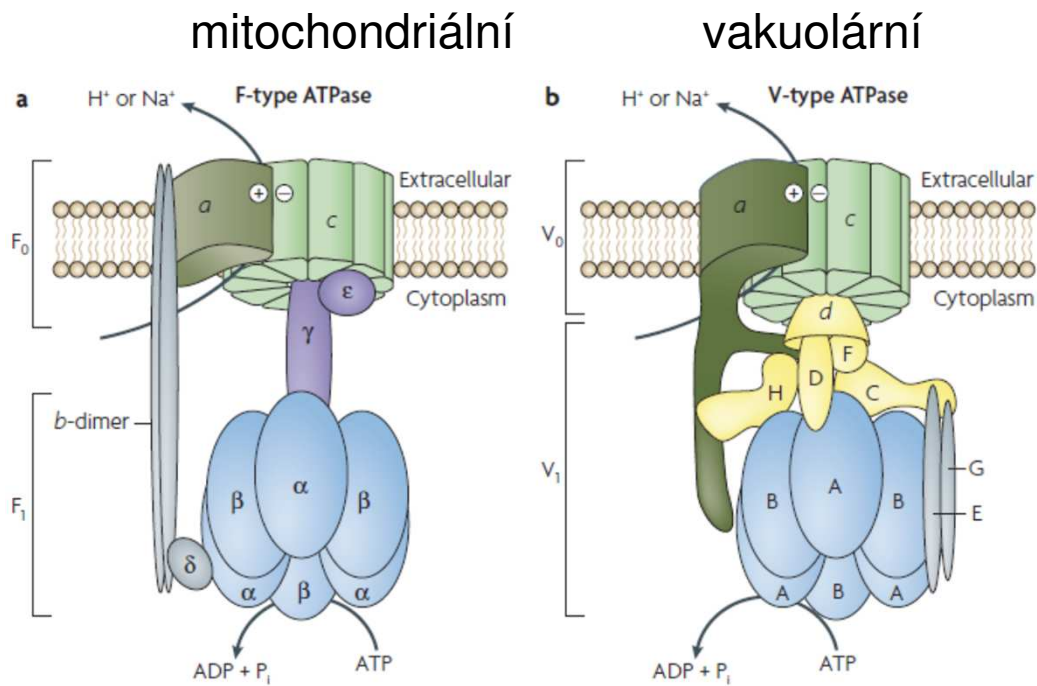
Pol II	Pol I	Pol III	Function
Polymerase Core			
Rpb1	A190	C160	Active center
Rpb2	A135	C128	Active center
Rpb3	AC40	AC40	
Rpb11	AC19	AC19	
Rpb9	A12.2 N ribbon	C11 N ribbon	RNA cleavage
TFIIS C-ribbon ^a	A12.2 C ribbon	C11 C ribbon	RNA cleavage
Rpb5	Rpb5	Rpb5	
Rpb6	Rpb6	Rpb6	
Rpb8	Rpb8	Rpb8	
Rpb10	Rpb10	Rpb10	
Rpb12	Rpb12	Rpb12	
Polymerase Stalk			
Rpb4	A14	C17	Initiation complex formation
Rpb7	A43	C25	Initiation complex formation



Vannini & Cramer, Mol Cell, 2012

- **TFIIF** se váže na pol II, stabilizuje DNA v prohlubni/cleft pol II a pomáhá TFIIB s nastavením startu
- v pol I a III jsou paralogy TFIIF součástí komplexu polymerázy





vytváří ATP

vytváří pH gradient

podobné proteinové komplexy – tzv. AAA ATPasy jsou součástí jiných komplexů (v jiných procesech):

- sekreční systém (T3SS injectisom)
- pohon bičičků
- „denaturace“ DNA: helikázy (MCM ...)
- v opravě poškozené DNA: Rad51, RecA

Mulkiđjanian et al, NRM, 2007

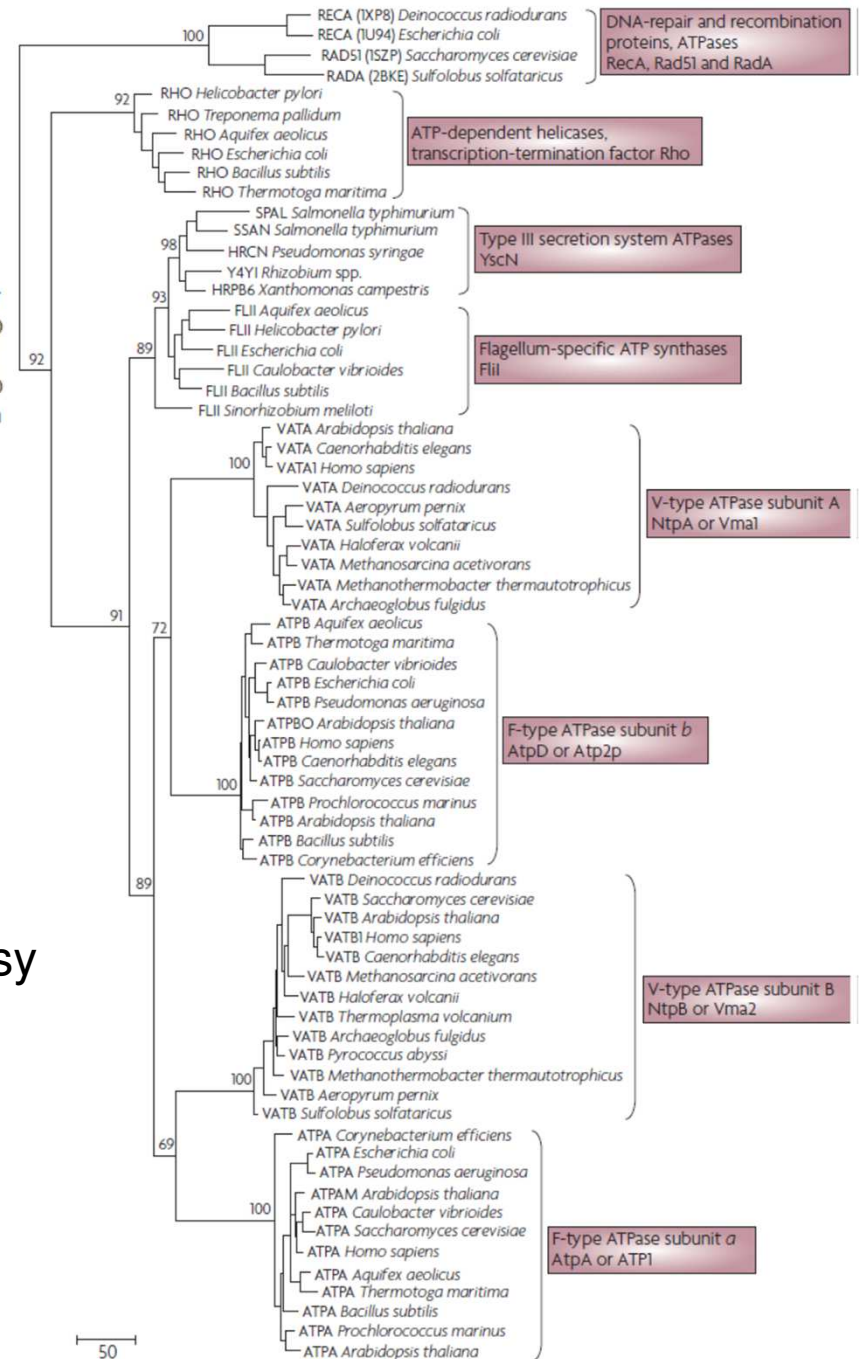
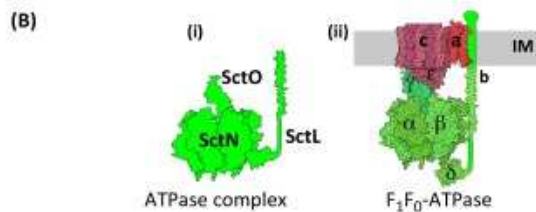
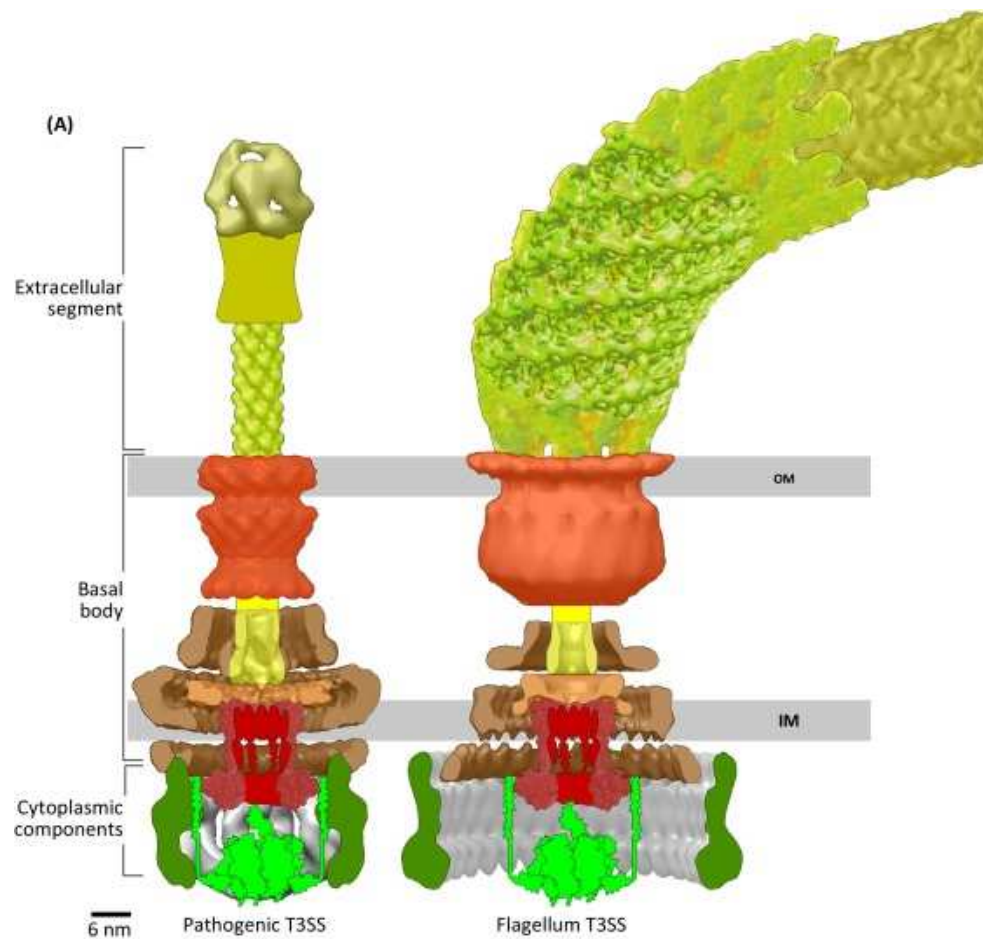


Figure 3 | Phylogenetic tree of the catalytic subunits of the F- and V-type ATPases and related P-loop ATPases. The protein sequences were retrieved from [GenBank](http://www.ncbi.nlm.nih.gov/genbank/) (for the RecA family, four



Trends in Biochemical Sciences

- ATPasové pumpy ...
- sekreční systém (T3SS injectisom)
- pohon bičičků ...

Mulkiđjanian et al, NRM, 2007

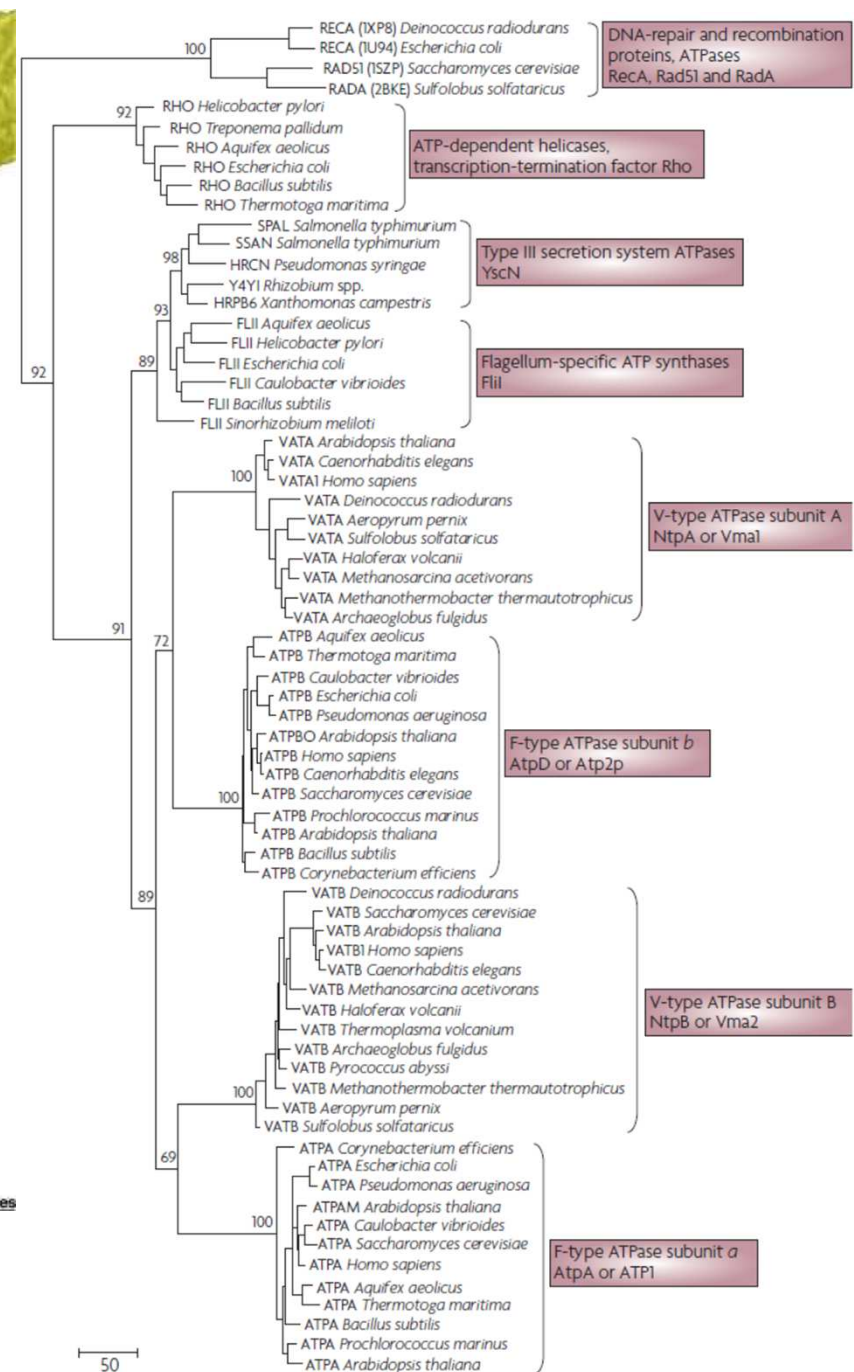
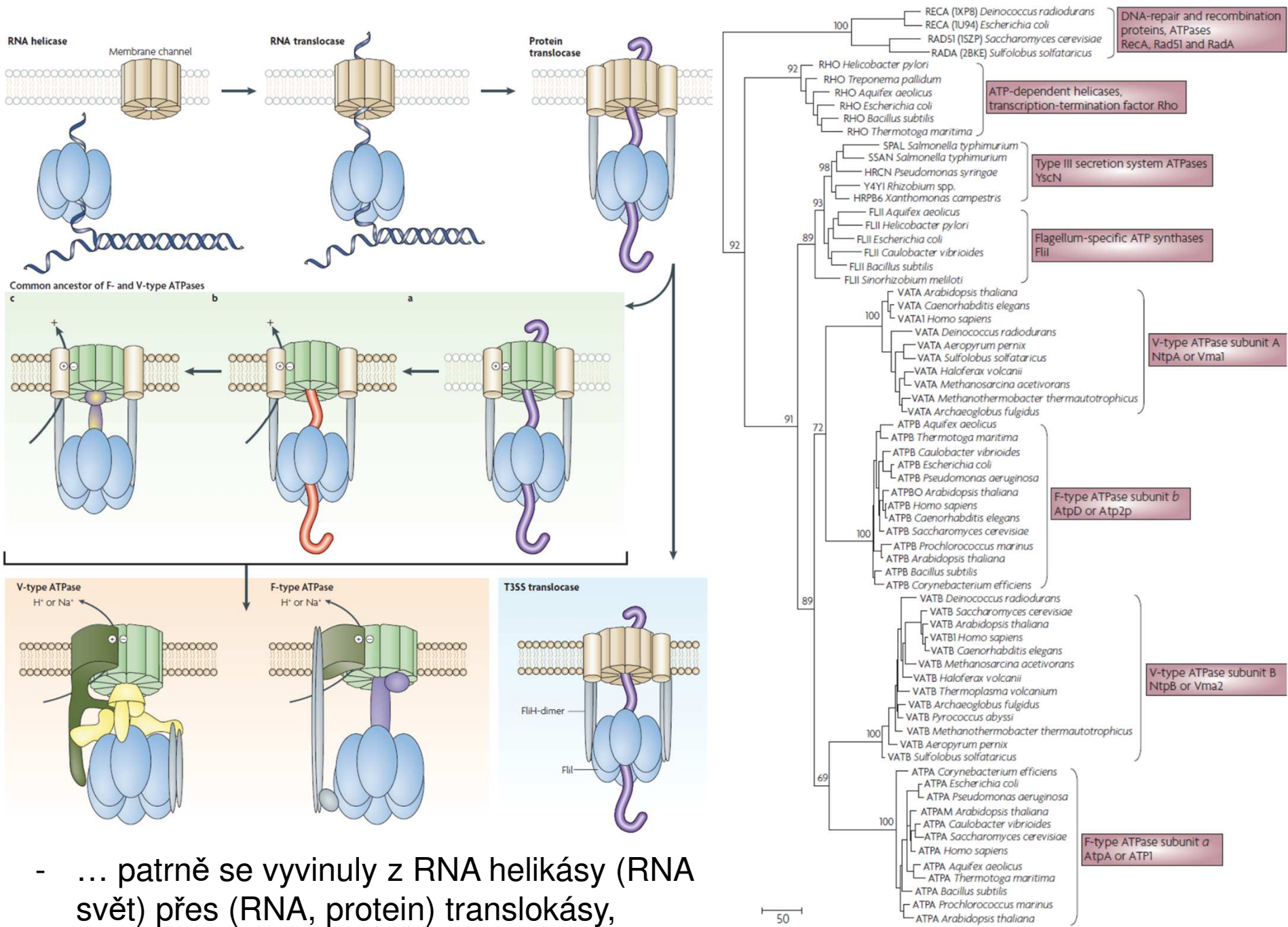


Figure 3 | Phylogenetic tree of the catalytic subunits of the F- and V-type ATPases and related P-loop ATPases. The protein sequences were retrieved from GenBank (for the RecA family, four



- ... patrně se vyvinuly z RNA helikázy (RNA svět) přes (RNA, protein) translokázy, ATPasové pumpy ...

Figure 3 | Phylogenetic tree of the catalytic subunits of the F- and V-type ATPases and related P-loop ATPases. The protein sequences were retrieved from GenBank (for the RecA family, four

Souhrn

Průběh evoluce je ovlivněn mnoha faktory

- Selekcční tlaky udržující funkční a stabilní (dříve vytvořené) proteiny a komplexy
- Mutace modifikující proteiny/komplexy (drift) jsou eliminovány u esenciálních proteinů
- Zatímco duplikované mohou podléhat mutacím „volněji“
- Duplikace a neofunkcionalizace (mutace) je hnací silou vzniku nových proteinů (nové kombinace domén) - komplexů – funkcí – typů buněk - organismů

Praktické implikace (pro zkoušku)

- analýza sekvenční podobnosti (stupeň konzervace)
 - napoví o přítomnosti domén (alignment – podobné AMK, doplnit analýzou sek. a terc. struktury)
 - projekce podobnosti do 3D modelu (ConSurf, PatchFinder) - konzervovaná struktura tj. vnitřní AMK držící fold (u ortologů i paralogů)
 - konzervované PPI kontaktní zóny na povrchu proteinu (povrchové AMK jsou konzervované pouze u ortologů – ne u paralogů - pokud jsou v alignmentu i paralogy, pak podobnost nevidíte)