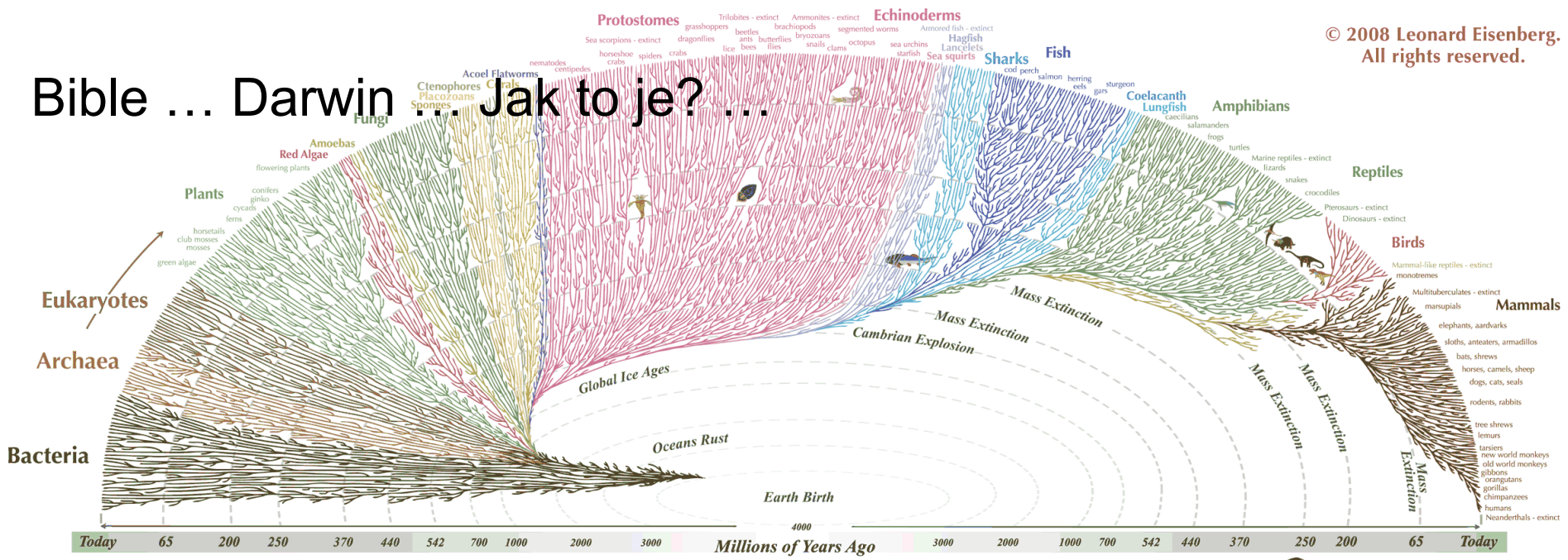
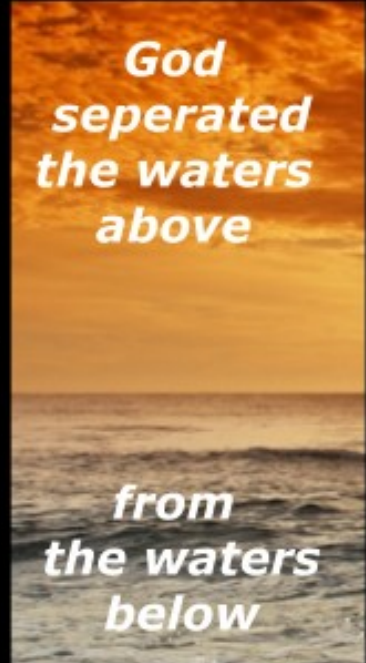


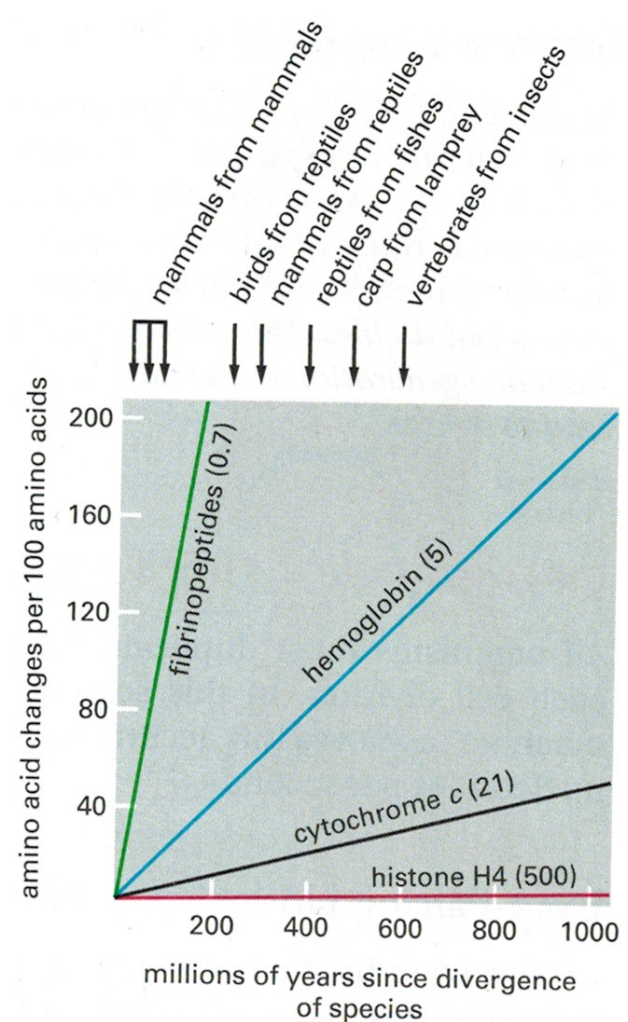
# Bible ... Darwin ... Jak to je? ...



All the major and many of the minor living branches of life are shown on this diagram, but only a few of those that have gone extinct are shown. Example: Dinosaurs - extinct

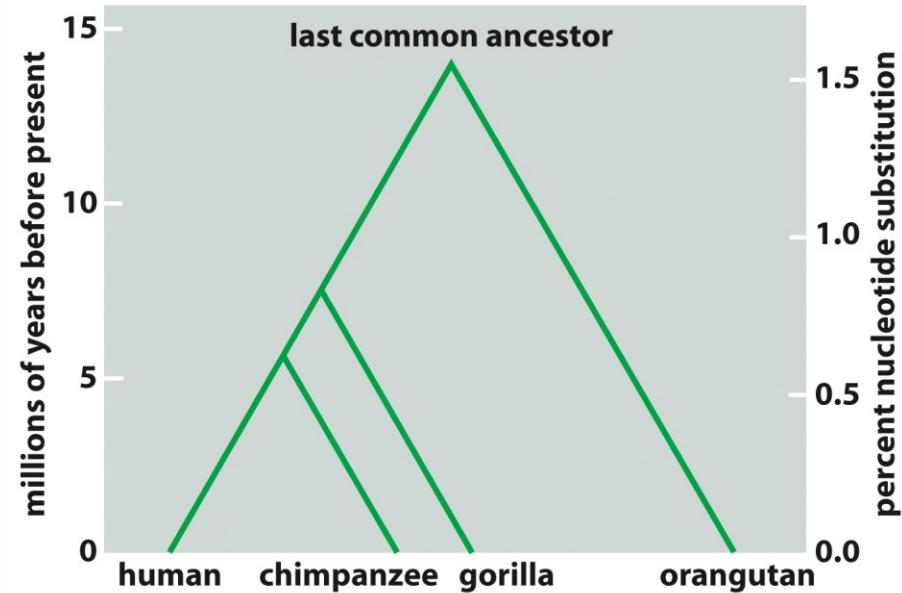
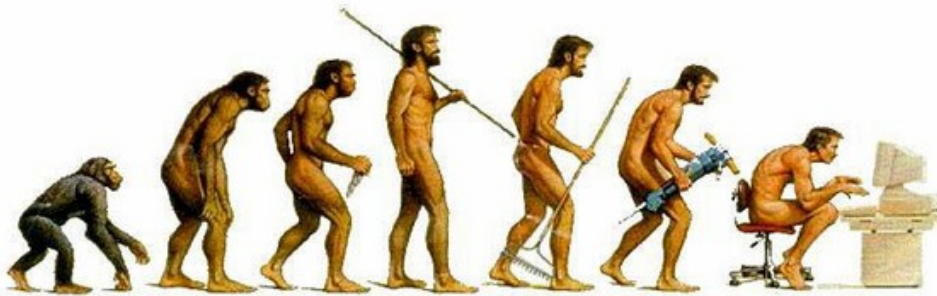


- srovnání genomů/proteomů ukazuje (podobně jako morfologie) na změnu („evoluci“) proteinů v čase
- divergence druhů koreluje do určité míry s konzervací/divergencí DNA/proteinových sekvencí
- DNA je replikována s relativně vysokou přesností (1 změna na  $10^9$  nukleotidů – cca 4000  $TNR_{12}$  písmen na A4 stranu – 8000  $TNR_{12}$  na list – 500 listů/balík – cca 250 balíčků)



- frekvence mutací DNA je +/- stejná, ale různé proteiny jsou různě změněné – (např. 6 ze 7 změn v cytochromu C jej poškodí; kvasinkový a lidský ubiquitin se liší třemi AMK)
- takto konzervované proteiny jsou lehce identifikovatelné v různých organismech a mají „homologní“ (**ortologní**) funkci

- velmi příbuzné sekvence DNA/proteinů mezi člověkem a šimpanzem (nejbližší člověku) díky „krátké“ době, po kterou mohlo k mutacím docházet – jsou zachovány i nukleotidy ve 3. pozici synonymních kodonů (kodony specifikující stejnou AMK)



- přesný fylogenetický vztah (strom) mezi člověkem a primáty bylo možné určit až dle sekvenací (ne podle morfologických znaků apod.)



- pro vzdálenější organismy (člověk a myš) je sekvence odlišnější (DNA je odlišnější než proteiny – pro 1AMK více kodonů; intron je odlišnější než exon – intr. nekóduje protein)

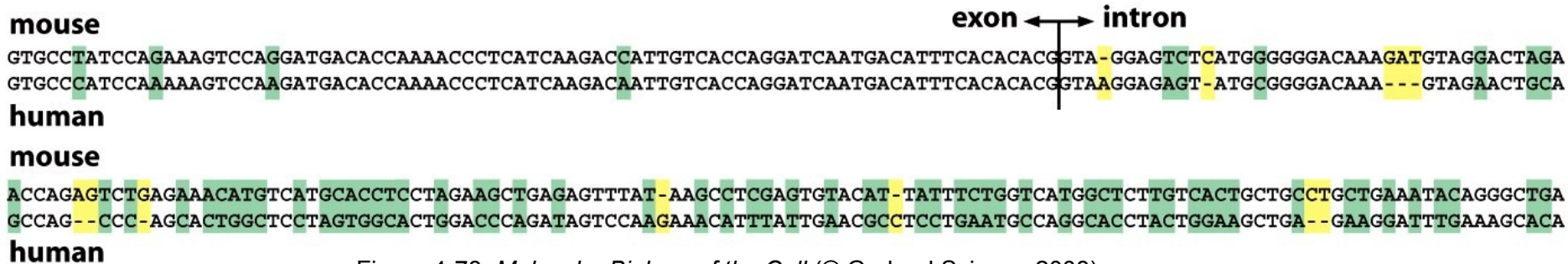


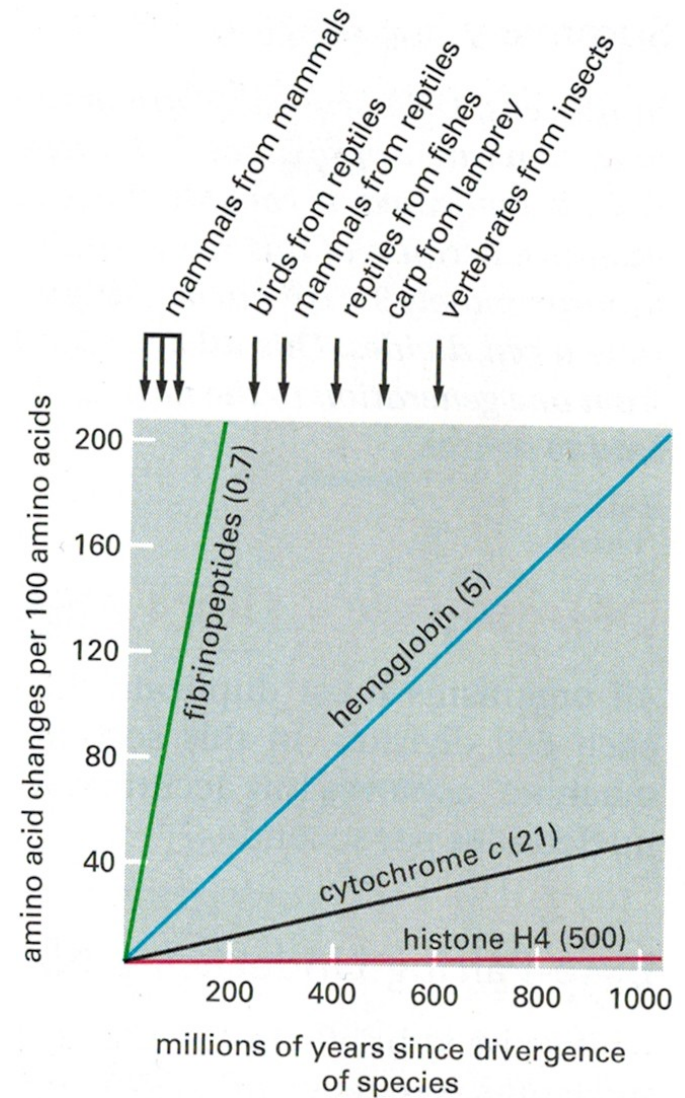
Figure 4-78 *Molecular Biology of the Cell* (© Garland Science 2008)

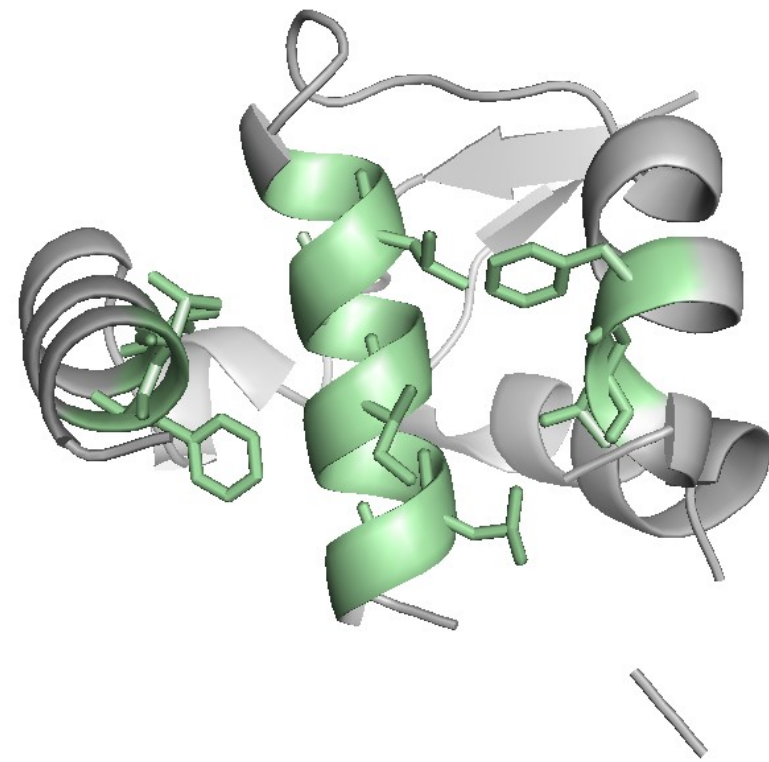
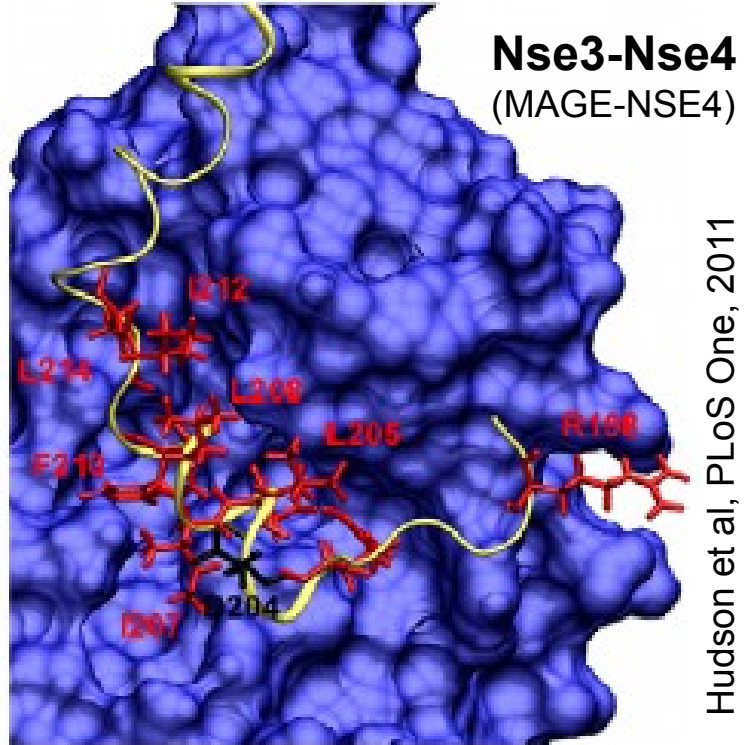
- čím důležitější funkce proteinu pro organismus (od nejzákladnějších po specializované) tím více je konzervovaná sekvence proteinu respektive jeho domén (velká podobnost i mezi odlišnými organismy)
- konzervace je důsledkem „purifikační selekce“ (eliminace jedinců s mutacemi v esenciálních/důležitých funkcích – většinou proteinech)

- 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, DNA polymerázy, histony, ribosomální proteiny, ... („drží“ základní systém)

- konzervované proteiny jsou stabilní, optimální pro svoji funkci (enzymatickou aktivitu, pro interakce s partnery ... ko-evoluují celé komplexy)

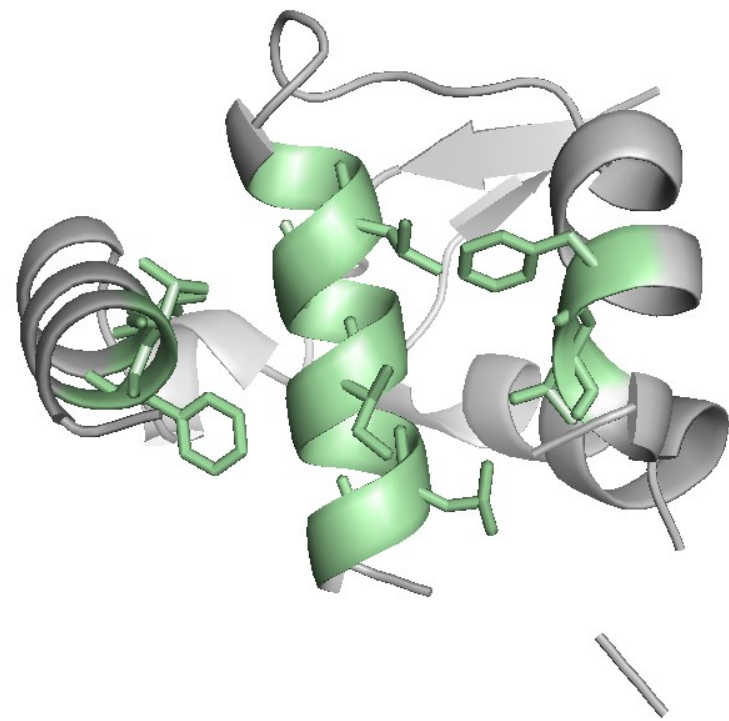
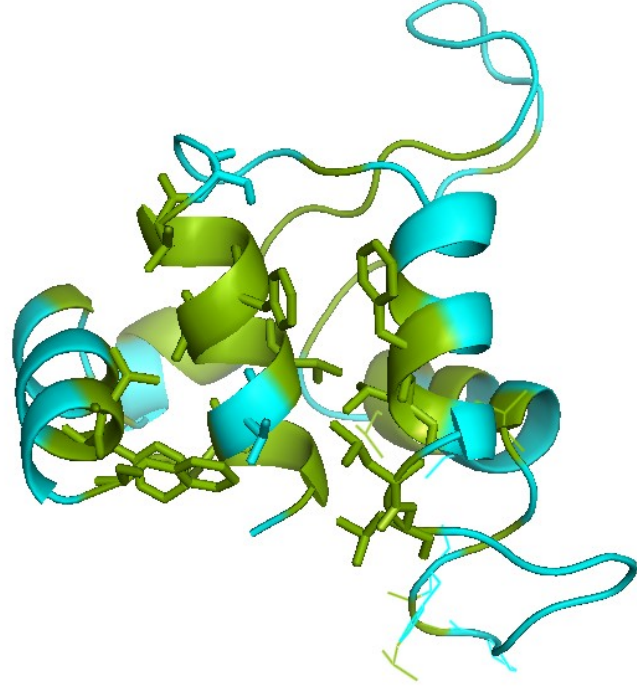
- selekční tlak na stabilitu a funkci - nutnost zachování funkce neposkytuje příliš prostoru pro evoluci/rozvoj nových vlastností (ale neznamená ani selekci na „nejstabilnější“ či „nejaktivnější“ = určitá volnost)





## ...RYBY, OBOŽIVELNÍCI, PTÁCI

N.v.	GLLMVLVLSVILM	SxxxxN-YTSLW	HFLK	KMGLEPKKEHF											
D.r.	GLLFLVLSVIFM	KGGTIK	ENLVWNIL	KKLRDP	PGEKHDEF	GDV	--	KKVVTEEFVR	QKYLEY	GKIPHTE	--	PVEYEFWR	GLRAEKE	--	VSKDKLLEFVG
T.n.	GLLFLVLSVIFM	KGGAVR	DSVVMNLL	KKLRVQP	GERHEPEF	GEV	--	KRVVMEEFVR	QRYLBC	NRIPHTE	--	ELHEEFRW	GQRADTE	--	VSKTKILEFMA
X.t.	GLLMVILSLIFM	KGNTAK	ESAVWML	RRLRIE	PAEKHSDF	GDV	--	KKLITEEFVK	QKYLEY	SKVLHTD	--	PVEYEFWR	GQRAFKE	--	TSKMQVLEFVS
G.g.	GLLIVILSFIFM	KGNSAK	DSAVWFL	RRLRVHP	GEKHEVDF	GDV	--	KKLVMEEFVR	QKYLEI	TIPIPLD	--	PPPTNFCW	GPRAAKE	--	TSKKDILSFVA
O.a.	GLLMVILSLIFM	KGSATN	ESVIWFL	RKLRVD	TRERHEV	FDV	--	KKLVTEEFVR	QKYLEY	NRIPHTE	--	PVEYEFQW	GARATKE	--	TTKMQVLFVFA
M.d.	GLLMVILSLIFM	KGNSAR	ESLVWDL	KKLRVD	PEKRHKTF	GDV	--	KKLVKDEFVR	QKYLEY	IRVPHSE	--	PEYEFLLW	GPRAAHE	--	TSKMQVLRVFA
S	GLLMIVLGLIFM	KGNTIK	ETEVDL	RRLGVY	PTKKHLIF	GDP	--	KKLITEEFVR	QRYLEY	RRIPTH	--	PVDYELQW	GPRTNLE	--	TSKMKVKFVA
A	GLLMIVLGLIFM	KGNTVK	ETEVDL	RRLGVY	PTKKHLV	GDP	--	KKLITDEFVR	QRYLEY	RRIPTH	--	PVDYELQW	GPRTKLE	--	TSKMKVKFVA
V	GLLMIVLGLIFM	KGNTIK	ETEVDL	RRLGVY	PTKKHLV	GDP	--	KKLITDEFVR	QRYLEY	RRIPTH	--	PVDYEFQW	GPRTNLE	--	TSKMKVKFVA
C	GLLMIVLGLIFM	KGNTIK	ETEVDL	RRLGVY	PTKKHLV	GDP	--	KKLITDEFVR	QRYLEY	RRIPTH	--	PVDYEFQW	GPRTNLE	--	TSKMKVKFVA
C	GLLMIVLGLIFM	KGNTIK	ETEVDL	RRLGVY	PTKKHLV	GDP	--	KKLITDEFVR	QRYLEY	RRIPTH	--	PVDYEFQW	GPRTNLE	--	TSKMKVKFVA
f	GLLMIVLGLIFM	KGNTIT	ETEVDL	RRLGVY	PTKKHLV	GDP	--	KKLITDEFVR	QRYLEY	RRIPTH	--	PVDYELQW	GPRTNLE	--	TSKMKVKFVA
M	GLLMIVLGLIFM	KGNTIT	ETEVDL	RRLGVY	PTKKHLV	GDP	--	KKLITDEFVR	QRYLEY	RRIPTH	--	PVDYELQW	GPRTNLE	--	TSKMKVKFVA
g1	GLLMIVLGLIFM	KGNTVK	ETEVDL	RRLGVY	PTKKHF	IFGDP	--	KKLITDEFVR	QRYLEY	RRIPTH	--	PVDYEFQW	GPRTNLE	--	TSKMKVKFVA
MI	GLLMIVLGLIFM	KGNTVK	ETEVDL	RRLGVY	PTKKHF	IFGDP	--	KKLITDEFVR	QRYLEY	RRIPTH	--	PVDYEFQW	GPRTNLE	--	TSKMKVKFVA
P	GLLMIVLGLIFM	KGNTIK	ETEVDL	RRLGVY	PTKKHLV	GDP	--	KKLITDEFVR	QRYLEY	RRIPTH	--	PVDYEFQW	GPRTNLE	--	TSKMKVKFVA
a	GLLMIVLGLIFM	KGNTIK	ETEVDL	RRLGVY	PTKKHLV	GDP	--	KKLITDEFVR	QRYLEY	RRIPTH	--	PVDYEFQW	GPRTNLE	--	TSKMKVKFVA
P	GLLMIVLGLIFM	KGNTIK	ETEVDL	RRLGVY	PTKKHLV	GDP	--	KKLITDEFVR	QRYLEY	RRIPTH	--	PVDYEFQW	GPRTNLE	--	TSKMKVKFVA
t.	GLLMIVLGLIFM	KGNTIK	ETEVDL	RRLGVY	PTKKHLV	GDP	--	KKLITDEFVR	QRYLEY	RRIPTH	--	PVDYEFQW	GPRTNLE	--	TSKMKVKFVA
H	GLLMIVLGLIFM	KGNTIK	ETEVDL	RRLGVY	PTKKHLV	GDP	--	KKLITDEFVR	QRYLEY	RRIPTH	--	PVDYEFQW	GPRTNLE	--	TSKMKVKFVA
s.	GLLMIVLGLIFM	KGNTIK	ETEVDL	RRLGVY	PTKKHLV	GDP	--	KKLITDEFVR	QRYLEY	RRIPTH	--	PVDYEFQW	GPRTNLE	--	TSKMKVKFVA



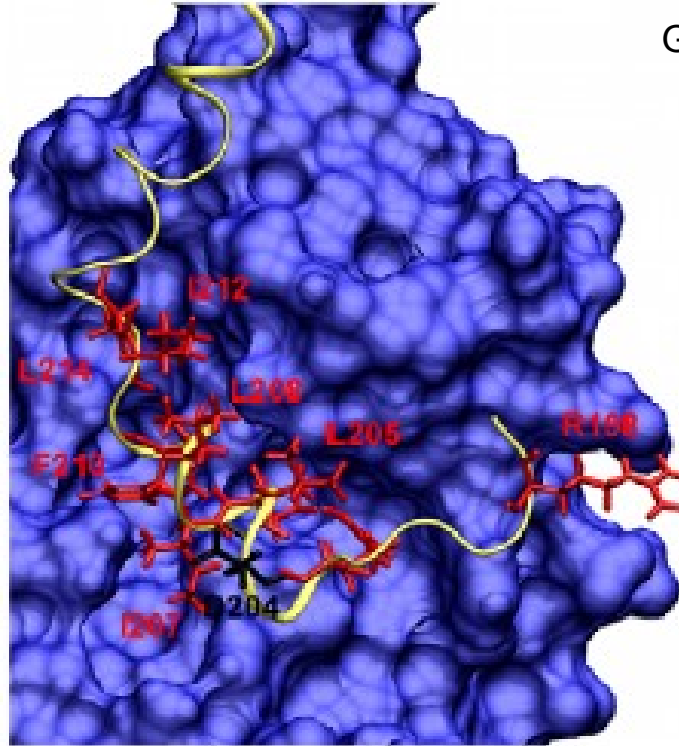
## KVASINKY ... určitá volnost v obsazení AMK

S. p.	GFLMTVIAFIAY	SHC	SVG	HSEL	QSF	EQE	LLT	---	E	ETTP	LHLD	-ITR	SL	SL	LV	RO	GYL	---	DR	KDDT	---	HN	QF	V	Y	Y	I	---	GS	RA	VT	E	IS	-	IE	GL	KS	FV	TE	FF																																												
Ani.	GLYTFIIAVILL	NGG	TIQ	EOK	LDR	YLS	R	MNA	---	E	QFT	P	VER	---	T	HL	L	Q	R	L	C	K	E	G	Y	---	V	KN	R	E	M	D	G	-	G	DE	I	E	Y	M	V	---	G	PR	G	K	VE	V	G	-	AR	GW	AG	LV	RE	V																												
Pyr.	ALYTTVIAFI II	SGG	IIP	EGK	LR	AL	R	R	MNA	---	Q	Q	T	P	L	G	T	---	K	K	T	L	A	A	M	V	K	D	G	Y	I	---	V	K	V	K	D	V	S	G	T	E	T	I	D	Y	I	V	---	G	PR	G	K	VE	V	G	-	GE	GW	AQ	F	IR	AV																					
N. f.	GLYSFIIAVILL	NGG	SLP	EOK	L	E	R	Y	L	K	R	T	N	A	---	D	T	Y	T	P	V	D	R	---	T	R	F	L	Q	R	L	C	K	E	G	Y	---	I	R	N	R	E	M	D	G	-	G	E	E	I	E	Y	M	V	---	G	PR	G	K	VE	V	G	-	VO	GW	AG	LV	RE	V															
Ate.	GLYTFIIALILL	NGG	SLP	EOK	L	E	R	Y	L	Q	R	T	N	A	---	D	T	Y	T	P	I	D	R	---	T	R	F	L	Q	R	L	C	K	D	G	Y	---	V	R	N	R	E	M	D	S	-	G	E	E	V	E	Y	M	V	---	G	PR	G	K	I	E	V	G	-	TQ	GW	AG	LV	RE	V														
Acl.	GLYSFIIAVITL	NGG	SLP	EOK	L	E	R	Y	L	K	R	T	N	A	---	D	T	Y	T	P	I	D	R	---	T	R	F	L	Q	R	L	C	K	E	G	Y	---	L	R	T	R	E	M	D	G	-	G	E	E	V	E	Y	L	V	---	G	PR	G	K	I	E	V	G	-	VO	GW	AG	LV	RE	V														
C. i.	ALYTFIISLITL	SGG	SLA	EOK	L	D	R	Y	L	R	R	V	N	A	---	D	T	Y	T	P	L	D	R	---	T	E	K	L	A	R	L	C	K	D	G	Y	---	V	R	N	R	D	V	D	G	-	G	E	E	V	E	Y	L	V	---	G	PR	G	K	I	E	V	G	-	TE	GW	AG	L	T	R	E	V												
P. a.	ALYTTIISLITI	SGG	E	L	S	D	T	R	L	R	R	H	L	A	R	L	N	A	A	E	Y	M	P	S	M	N	P	N	D	P	x	-	T	D	V	V	L	Q	R	M	I	K	H	G	Y	---	V	R	Q	V	D	N	R	G	x	D	D	D	S	T	T	T	H	V	---	G	PR	G	K	A	E	V	P	-	KE	S	I	AG	F	V	R	T	I	Y
N. c.	GLYTMLIAIITL	SGG	E	L	S	D	P	R	L	R	R	Y	L	T	R	L	N	A	A	x	P	M	N	E	N	A	P	S	E	K	-	T	E	L	V	L	Q	R	M	T	K	Q	G	Y	---	V	R	V	A	D	N	x	A	G	D	D	D	A	I	T	T	H	V	---	G	PR	G	K	VE	V	D	-	NE	A	I	AA	V	V	RE	V				
M. g.	GLYSMIVTIIQL	N	R	G	E	L	S	D	P	K	L	K	R	Y	L	Q	R	L	N	A	---	E	T	N	T	P	V	E	K	-	T	E	L	L	L	Q	R	L	I	R	Q	N	Y	I	---	V	K	T	V	E	R	N	x	-	D	D	A	I	T	T	R	V	---	G	PR	A	K	Q	E	L	T	-	DE	A	M	A	S	I	V	R	D	V		

O  
B  
R  
A  
T.

GLLMLVLSVILM	S	x	x	x	N	-	Y	T	S	L	W	H	F	L	K	K	M	G	L	E	P	K	K	E	H	E	V	F	G	D	P	-	E	K	L	I	A	Q	E	F	T	R	O	G	Y	D	E	R	R	K	V	T	G	G	E	-	E	A	T	F	E	Y	S	W	---	G	S	R	S	N	K	E	---	L	T	K	R	K	V	L	E	F	V	S
GLLEFVILSVIFM	K	G	G	T	I	K	-	E	N	L	V	W	N	I	L	K	K	L	R	L	D	P	G	E	K	H	D	E	F	G	D	V	-	K	K	V	V	T	E	E	F	V	R	Q	K	Y	L	E	Y	G	K	I	P	H	T	E	-	F	V	E	Y	E	F	R	W	---	G	L	R	A	E	K	E	---	V	S	K	L	K	L	E	F	V	G
GLLEFVILSVIFM	K	G	G	A	V	R	-	D	S	V	W	N	I	L	K	K	L	R	V	Q	P	G	E	R	H	E	P	E	F	G	E	V	-	K	R	V	M	E	E	F	V	R	Q	R	Y	L	E	C	N	R	I	P	H	T	E	-	E	L	E	H	E	F	R	W	---	G	Q	R	A	D	T	E	---	V	S	K	T	K	I	L	E	F	M	
GLLMVILSLIFM	K	G	N	T	A	K	-	E	S	A	V	W	E	M	L	R	R	L	R	I	E	P	A	E	K	H	S	D	F	G	D	V	-	K	K	L	I	T	E	E	F	V	K	Q	K	Y	L	E	Y	S	K	V	L	H	T	D	-	F	V	E	Y	E	F	R	W	---	G	Q	R	A	F	K	E	---	T	S	K	M	Q	V	L	E	F	V
GLLIVILSFIFM	K	G	N	S	A	K	-	D	S	A	V	W	E	F	L	R	R	L	R	V	H	P	G	E	K	H	E	V	F	G	D	V	-	K	K	L	V	M	E	E	F	V	R	Q	K	Y	L	E	I	T	P	I	P	L	T	D	-	P	P	E	F	N	F	Q	W	---	G	P	R	A	A	K	E	---	T	S	K	K	D	I	L	S	F	V
GLLMVILSLIFM	K	G	S	A	T	N	-	E	S	V	I	W	E	T	L	R	K	L	R	V	D	T	R	E	R	H	E	V	F	G	D	V	-	K	K	L	V	T	E	E	F	V	R	Q	K	Y	L	E	Y	N	R	I	P	H	T	E	-	F	V	E	F	E	F	Q	W	---	G	A	R	A	T	K	E	---	T	T	K	M	Q	V	L	N	F	V
GLLMVILSLIFM	K	G	N	S	A	R	-	E	S	L	W	D	V	L	K	K	L	R	V	D	P	E	K	R	H	K	T	F	G	D	V	-	K	K	L	V	K	D	E	F	V	R	Q	K	Y	L	E	Y	I	R	V	P	H	S	E	-	P	P	E	Y	E	F	L	W	---	G	P	R	A	A	H	E	---	T	S	K	M	Q	V	L	R	F	V	
GLLMIVLGLIFM	K	G	N	T	I	K	-	E	T	E	W	D	F	L	R	R	L	G	V	P	T	K	K	H	L	I	F	G	D	P	-	K	K	L	I	T	E	E	F	V	R	Q	R	Y	L	E	Y	R	R	I	P	H	T	D	-	F	V	D	Y	E	L	Q	W	---	G	P	R	T	N	L	E	---	T	S	K	M	K	V	L	K	F	V		





- selekční tlak na strukturu, na povrch = interakce
- 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

Nse4

aa75 aa103

S.p.	L	H	I	G	R	P	K	F	I	E	L	F	T	K	N	I	K	Q	F	L	N	Y	P	T	S	H	S	
A.n.	L	x	-	S	A	G	I	D	V	D	E	F	V	S	K	C	I	S	F	M	R	R	A	P	A	D	S	Q
N.f.	L	x	-	S	A	G	I	D	V	D	E	F	V	S	K	C	I	S	F	M	R	R	G	P	S	D	P	D
A.t.	L	x	-	T	A	R	I	D	V	D	E	F	V	S	K	C	I	S	F	M	R	R	A	P	G	D	S	Q
A.o.	L	x	-	S	A	G	I	D	V	D	E	F	V	S	K	C	I	S	F	M	C	R	A	P	E	D	S	Q
A.c.	L	x	-	S	A	G	I	D	V	D	E	F	V	S	K	C	I	S	F	M	R	R	G	P	S	D	P	E
D.r.	L	H	A	E	G	S	S	F	D	P	S	A	F	A	E	H	L	L	S	F	M	D	L	N	R	L	E	D
X.l.	L	H	A	D	M	T	V	F	D	P	T	S	F	T	A	D	L	L	S	F	M	G	L	N	R	M	E	S
G.g.	L	R	S	E	I	T	T	F	D	S	L	T	F	A	E	D	L	L	T	F	M	G	I	N	R	T	E	T
M.d.	L	H	S	D	M	T	L	F	D	P	S	S	F	A	E	D	L	L	K	F	M	G	L	N	R	L	E	V
C.f.	L	R	S	D	L	S	S	F	D	L	R	Y	V	E	T	L	L	T	H	M	G	V	N	P	L	E	A	E
M.m.	L	R	S	D	L	S	S	F	D	L	R	Y	V	E	T	L	L	T	H	M	G	V	N	P	L	E	A	E
H.s.	L	R	S	D	L	S	S	F	D	L	R	Y	V	E	T	L	L	T	H	M	G	V	N	P	L	E	A	E

NSE4b/EID3

D.n.	L	N	S	D	M	S	S	F	D	P	S	A	F	A	E	D	L	L	F	F	F	V	G	L	N	R	M	E	D
C.f.	L	N	S	D	M	S	S	F	D	P	S	A	F	A	E	D	L	L	F	F	F	V	G	L	N	R	M	E	D
M.m.	L	N	T	D	M	N	F	F	D	P	S	A	F	C	D	L	L	L	L	F	V	G	F	N	V	E	E	E	E
H.s.	L	N	S	D	M	N	F	F	D	P	S	A	F	C	D	L	L	F	F	V	G	L	N	R	M	E	G	D	
aa106++	+++	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	++	+	+	aa135					

core region

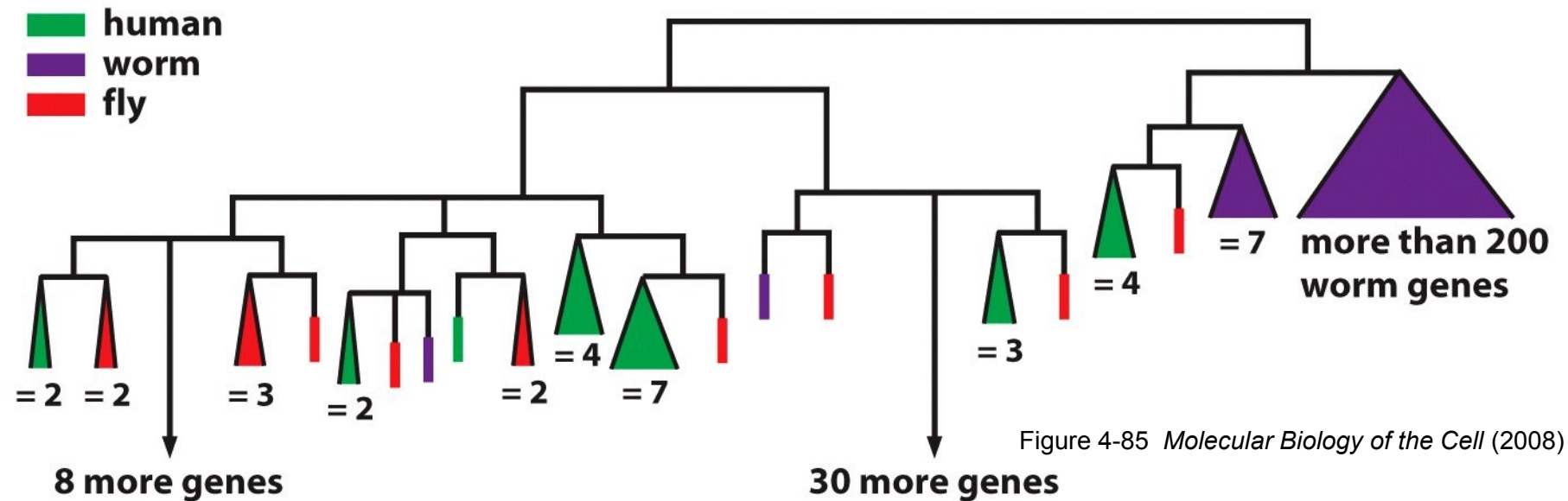
KVASINKY ... určitá volnost v obsazení AMK

S.p.	G	F	L	M	T	V	I	A	F	I	A	V	-	S	H	C	S	V	G	-	H	S	E	L	Q	S	F	L	Q	E	L	L	T	-	-	E	T	T	P	L	H	L	D	-	I	T	R	S	L	S	L	V	R	O	G	Y	L	-	D	R	V	K	D	D	T	-	-	H	N	Q	F	V	Y	Y	I	-	G	S	R	A	V	T	E	I	S	-	I	E	G	L	K	S	F	V	T	E	F	F		
Ani.	G	L	Y	T	F	I	I	A	V	I	L	L	-	N	G	G	T	L	Q	-	E	Q	K	I	D	R	Y	L	S	R	M	N	A	-	-	E	Q	F	T	P	V	E	R	-	T	H	L	L	Q	R	L	C	K	E	G	Y	L	-	V	K	N	R	E	M	D	G	-	G	D	E	I	I	E	Y	M	V	-	G	P	R	G	K	V	E	V	G	-	A	R	G	W	A	G	L	V	R	E	V	Y	
Pyr.	A	L	Y	T	T	V	I	A	F	I	I	I	-	S	G	G	I	I	P	-	E	G	K	I	D	R	A	L	R	R	M	N	A	-	-	D	Q	T	T	P	L	G	T	-	K	K	T	L	A	A	M	V	K	D	G	Y	I	-	V	K	V	K	D	V	S	G	G	T	E	T	I	D	Y	I	V	-	G	P	R	G	K	V	E	V	G	-	G	E	G	W	A	Q	F	I	R	A	V	Y		
N.f.	G	L	Y	S	F	I	I	A	V	I	L	L	-	N	G	G	S	L	P	-	E	Q	K	L	E	R	Y	L	K	R	T	N	A	-	-	D	T	Y	T	P	W	D	R	-	T	D	R	F	L	Q	R	L	C	K	E	G	Y	L	-	I	R	N	R	E	M	D	G	-	G	E	E	I	I	E	Y	M	V	-	G	P	R	G	K	V	E	V	G	-	V	Q	G	W	A	G	L	V	R	E	V	Y
Ate.	G	L	Y	T	F	I	I	A	L	I	L	L	-	N	G	G	S	L	P	-	E	Q	K	L	E	R	Y	L	Q	R	T	N	T	-	-	D	T	Y	T	P	I	D	R	-	T	D	R	F	L	Q	R	L	C	K	D	G	Y	L	-	V	R	N	R	E	M	D	S	-	G	E	E	V	I	E	Y	M	V	-	G	P	R	G	K	I	E	V	G	-	T	Q	G	W	A	G	L	V	R	E	V	Y
Ac1.	G	L	Y	S	F	I	I	A	V	I	T	L	-	N	G	G	S	L	P	-	E	Q	K	L	E	R	Y	L	K	R	T	N	A	-	-	D	T	F	T	P	I	D	R	-	T	D	R	F	L	Q	R	L	C	K	E	G	Y	L	-	L	R	T	R	E	M	D	G	-	G	E	E	V	W	E	Y	L	V	-	G	P	R	G	K	I	E	V	G	-	V	Q	G	W	A	G	L	V	R	E	V	Y
C.i.	A	L	Y	T	F	I	I	S	L	I	T	L	-	S	G	G	S	L	A	-	E	Q	K	I	D	R	Y	L	R	R	V	N	A	-	-	D	T	Y	T	P	L	D	R	-	T	E	K	L	L	A	R	L	C	K	D	G	Y	L	-	V	R	N	R	D	V	D	G	-	G	E	E	V	W	E	Y	L	V	-	G	P	R	G	K	I	E	V	G	-	T	E	G	W	A	G	L	T	R	E	V	Y
P.a.	A	L	Y	T	T	I	I	S	L	I	T	I	-	S	G	G	E	L	S	-	D	T	R	L	R	R	H	L	A	R	L	N	A	A	E	Y	M	P	S	M	N	P	N	D	F	x	-	T	D	V	L	Q	R	M	I	K	H	G	Y	L	-	V	R	W	D	N	R	G	x	D	D	D	S	T	T	H	V	-	G	P	R	G	K	A	E	V	P	-	K	E	S	I	A	G	F	V	R	T	I	Y
N.c.	G	L	Y	T	M	L	I	A	I	I	T	L	-	S	G	G	E	L	S	-	D	P	R	L	R	R	Y	L	T	R	L	N	A	A	x	F	M	N	E	N	A	P	S	E	K	-	T	E	L	V	L	Q	R	M	T	K	O	G	Y	L	-	V	R	V	A	D	N	x	A	G	D	D	A	I	T	H	V	-	G	P	R	G	K	V	E	V	D	-	N	E	A	I	A	A	V	R	E	V	Y	
M.g.	G	L	Y	S	M	I	V	T	I	I	Q	L	-	N	R	G	E	L	S	-	D	P	K	L	K	R	Y	L	Q	R	L	N	A	-	-	E	T	N	T	P	V	E	K	-	T	D	L	L	L	Q	R	L	I	R	Q	N	Y	I	-	V	K	T	V	E	R	N	x	-	D	D	A	I	T	H	R	V	-	G	P	R	A	K	O	E	L	T	-	D	E	A	M	A	S	I	V	R	D	V	Y	

O  
B  
R  
A  
T.

G	L	L	M	L	V	L	S	V	I	L	M	-	S	x	x	x	x	N	-	Y	T	S	L	W	H	F	L	K	K	G	L	E	P	K	K	E	H	E	V	F	G	D	P	-	E	K	L	I	A	Q	E	F	T	R	O	G	Y	L	D	R	R	K	V	T	G	G	E	-	-	E	A	T	F	E	Y	S	W	-	G	S	R	S	N	K	E	-	-	-	L	T	K	R	K	V	L	E	F	V	S	-	-	-			
G	L	L	F	V	I	L	S	V	I	F	M	-	K	G	G	T	I	K	-	E	N	L	V	W	N	I	L	K	K	L	R	L	D	P	G	E	K	H	D	E	F	G	D	V	-	-	K	K	V	V	T	E	E	F	V	R	Q	K	Y	L	E	Y	G	K	I	P	H	T	E	-	-	-	P	V	E	Y	E	F	R	W	-	G	L	R	A	E	K	E	-	-	-	V	S	K	L	K	L	E	F	V	G	-	-	-	
G	L	L	F	V	I	L	S	V	I	F	M	-	K	G	G	A	V	R	-	D	S	V	W	N	I	L	K	K	L	R	V	Q	P	G	E	R	H	E	P	F	G	E	V	-	-	K	R	V	M	E	E	F	V	R	Q	R	Y	L	E	C	N	R	I	P	H	T	E	-	-	-	E	L	E	H	E	F	R	W	-	G	Q	R	A	D	T	E	-	-	-	V	S	K	T	K	I	L	E	F	M	A	-	-	-		
G	L	L	M	V	I	L	S	L	I	F	M	-	K	G	N	T	A	K	-	E	S	A	V	W	E	M	L	R	R	L	R	I	E	P	A	E	K	H	S	D	F	G	D	V	-	-	K	K	L	I	T	E	E	F	V	K	Q	K	Y	L	E	Y	S	K	V	L	H	T	D	-	-	-	E	Y	E	Y	E	F	R	W	-	G	Q	R	A	F	K	E	-	-	-	T	S	K	M	Q	V	L	E	F	V	S	-	-	-
G	L	L	I	V	I	L	S	F	I	F	M	-	K	G	N	S	A	K	-	D	S	A	V	W	E	F	L	R	R	L	R	V	H	P	G	E	K	H	E	V	F	G	D	V	-	-	K	K	L	V	M	E	E	F	V	R	Q	K	Y	L	E	I	T	P	I	P	L	T	D	-	-	-	P	P	E	F	N	F	Q	W	-	G	P	R	A	A	K	E	-	-	-	T	S	K	K	D	I	L	S	F	V	A	-	-	-
G	L	L	M	V	I	L	S	L	I	F	M	-	K	G	S	A	T	N	-	E	S	V	I	W	E	T	L	R	K	L	R	V	D	T	R	E	R	H	E	V	F	G	D	V	-	-	K	K	L	V	T	E	E	F	V	R	Q	K	Y	L	E	Y	N	R	I	P	H	T	E	-	-	-	P	V	E	F	E	F	Q	W	-	G	A	R	A	T	K	E	-	-	-	T	T	K	M	Q	V	L	N	F	V	A	-	-	-
G	L	L	M	V	I	L	S	L	I	F	M	-	K	G	N	S	A	R	-	E	S	L	W	D	V	L	K	K	L	R	V	D	P	E	K	R	H	K	T	F	G	D	V	-	-	K	K	L	V	K	D	E	F	V	R	Q	K	Y	L	E	Y	I	R	V	P	H	S	E	-	-	-	P	P	E	Y	E	F	L	W	-	G	P	R	A	A	H	E	-	-	-	T	S	K	M	Q	V	L	R	F	V	A	-	-	-	
G	L	L	M	I	V	L	G	L	I	F</																																																																																															

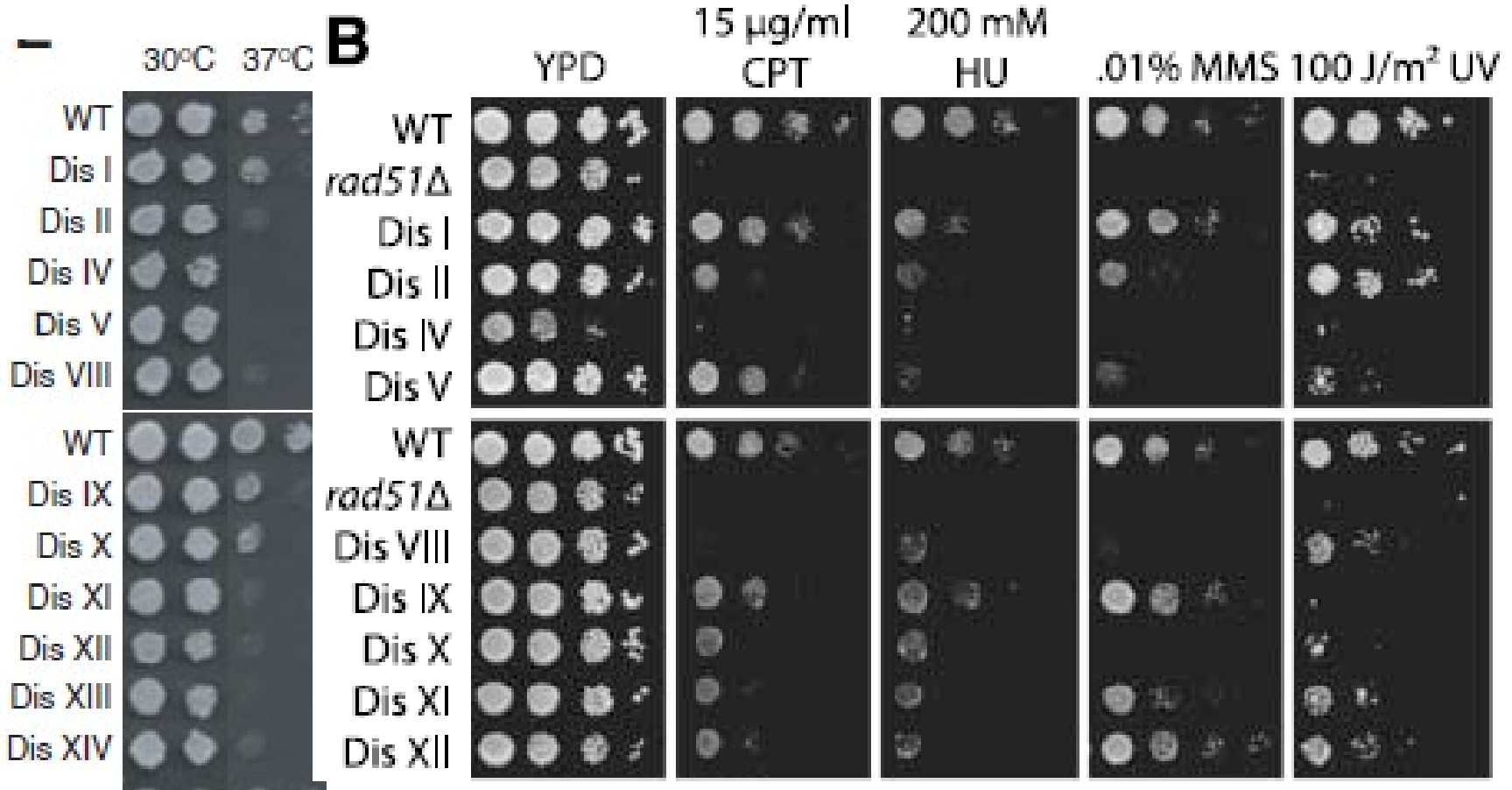
- selekční tlak na stabilitu a funkci - nutnost zachování funkce neposkytuje příliš prostoru pro evoluci/rozvoj nových vlastností  
 - pro rozvoj nových vlastností, nových druhů - spíše než druhově specifické mutace proteinů lze vidět **expansi různých genových/proteinových rodin** v různých živočišných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



(příklad superrodiny jaderných hormonálních receptorů)

- **duplikace a divergence** jsou hlavními tahouny evolučních procesů (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 cca 2000 foldů)

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

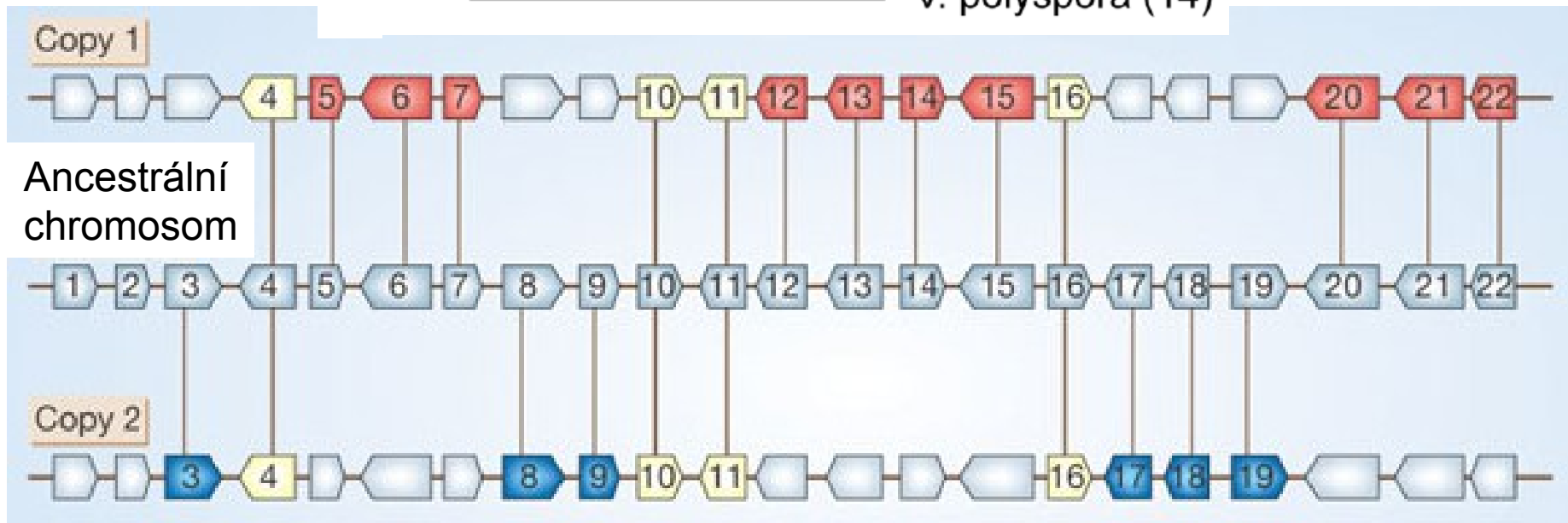
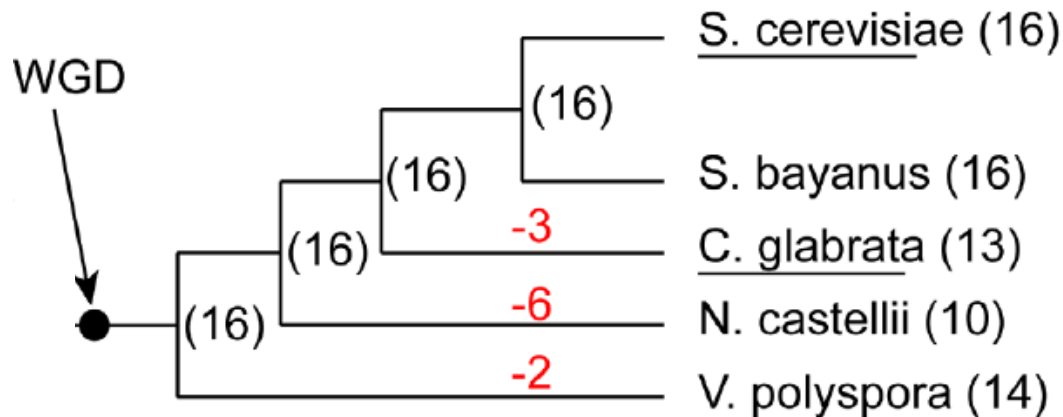


Shelzer et al, Science (2011)

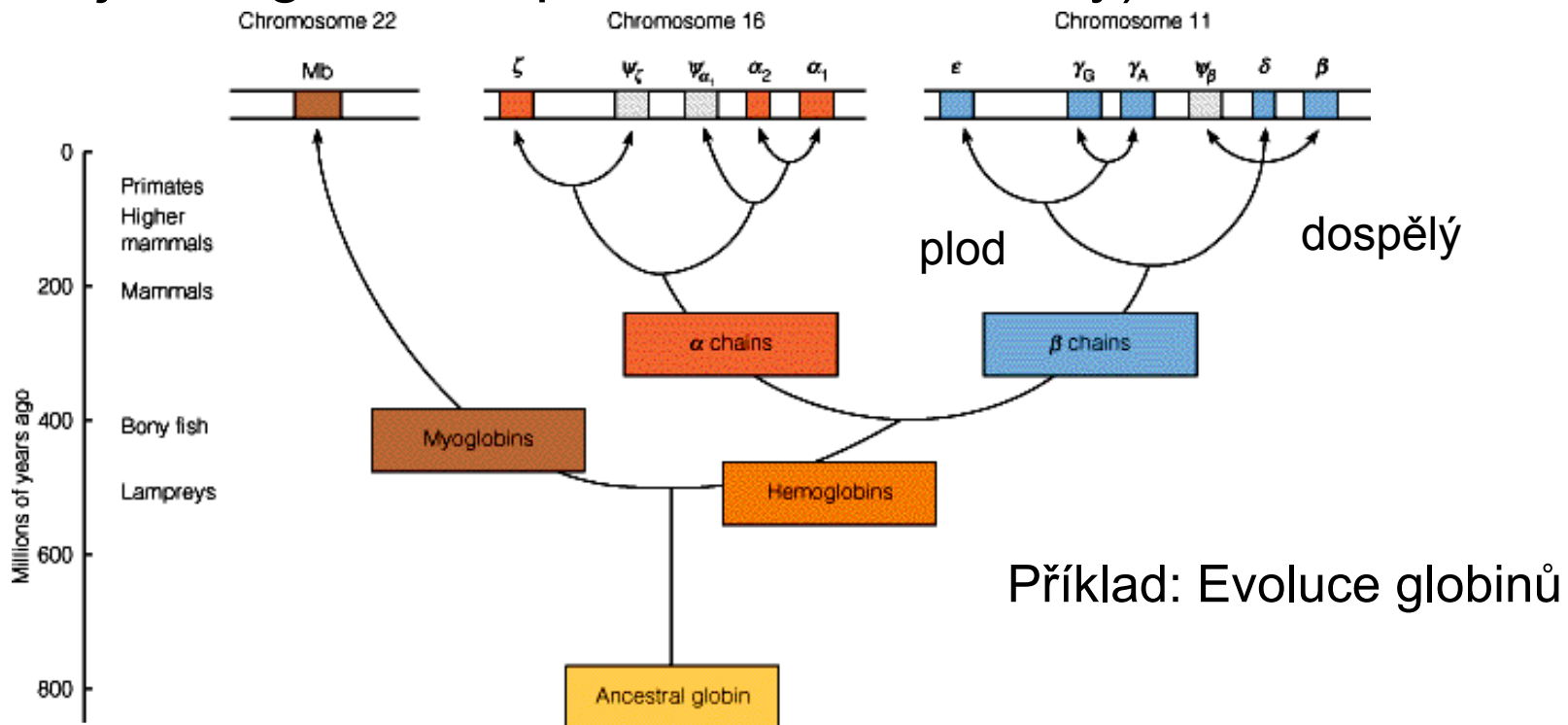
- nadbytek proteinů (např. transkripčních faktorů) nebo nerovnováha podjednotek komplexů znamená disregulaci některých procesů

# Celogenomová duplikace u kvasinek

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



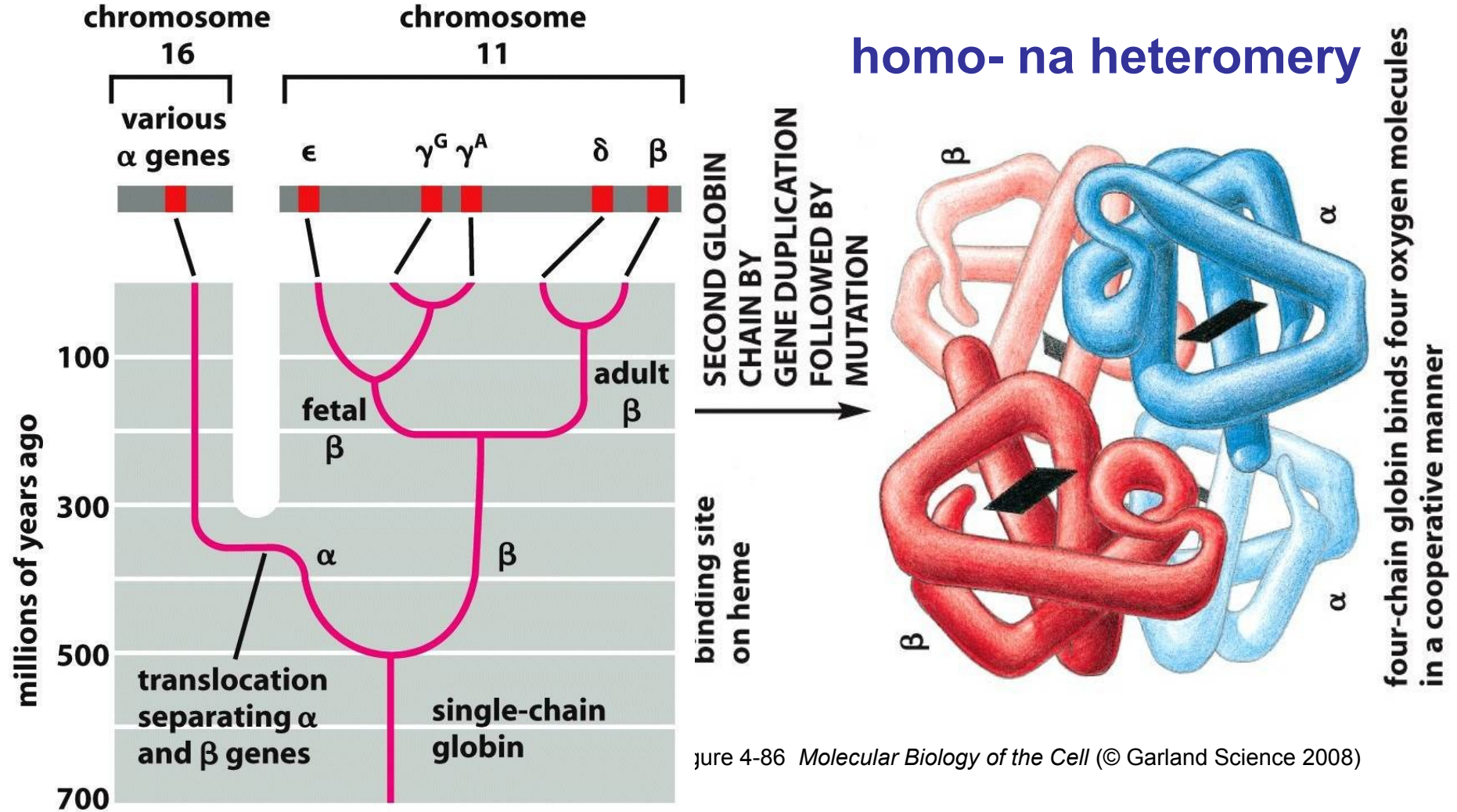
- duplikace – **reverzní transkripcí** a integrací DNA (pouze individuální geny nebo pouze domény = exony)
- na počátku stejné sekvence = stejné funkce (pod jinými **promotory** – jiný lokus tzn. jiné „okolí“ zaintegrované „mRNA“ tzn. jiná regulace exprese – „nové“ buňky)



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

# Evoluce rodiny globinů (duplikace)

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

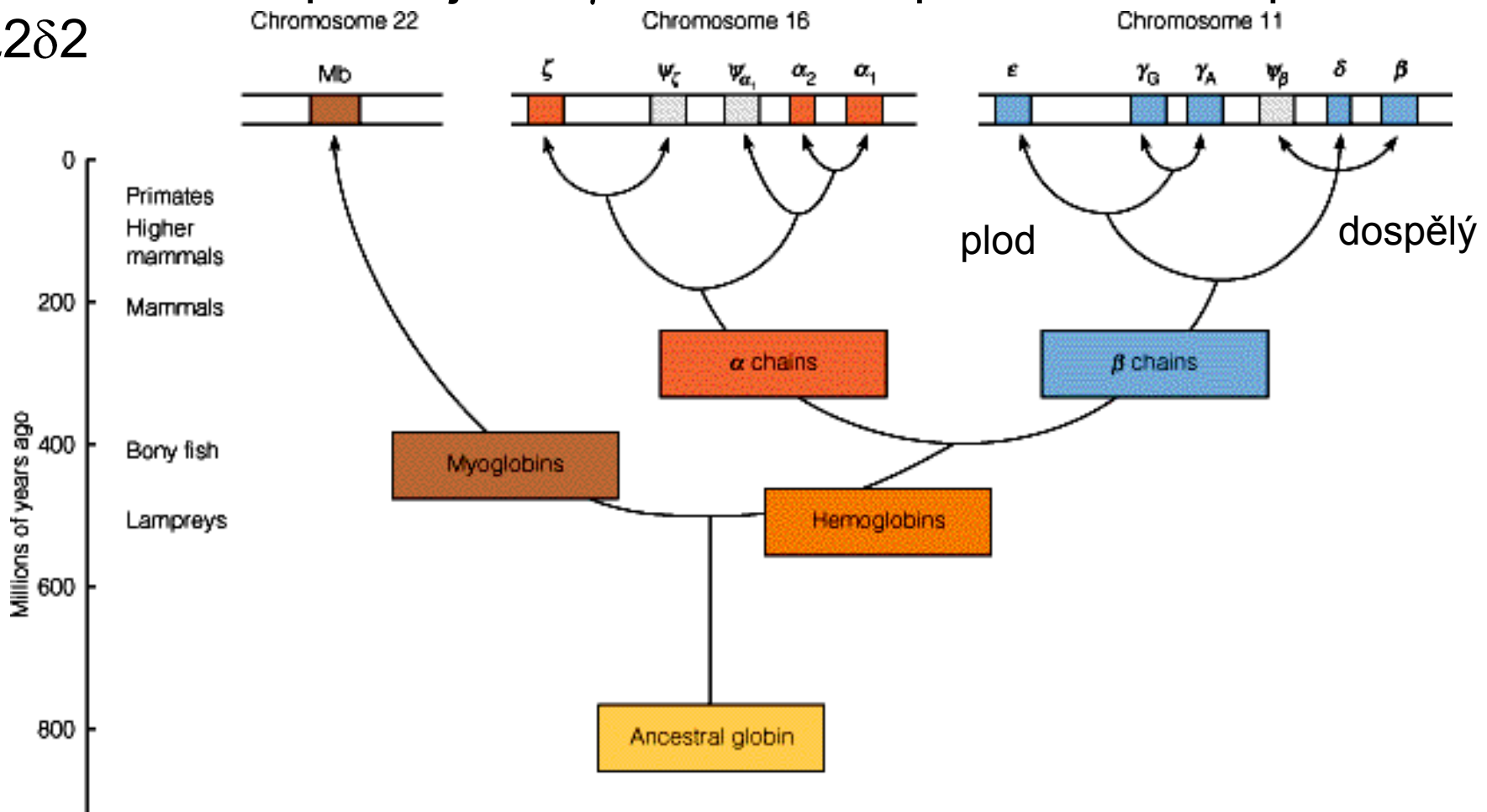


- vyšší obratlovci mají  $\alpha$ - a  $\beta$ -globin (tvoří  $\alpha_2\beta_2$  komplex – účinnější přenos kyslíku) – vyvinul se u vyšších ryb (500Mya)

# Evoluce rodiny globinů (duplikace)

- u savců se dále duplikoval  $\beta$ -globin (**paralog**) - exprimován specificky v embryu – má vyšší afinitu ke kyslíku a napomáhá přenosu kyslíku z krve matky do krve plodu

- dále se duplikoval a specializoval se na časná vývojová stádia  $\alpha_2\varepsilon_2$  a pozdější  $\alpha_2\gamma_2$  - k další duplikaci došlo u primátů  $\alpha_2\delta_2$



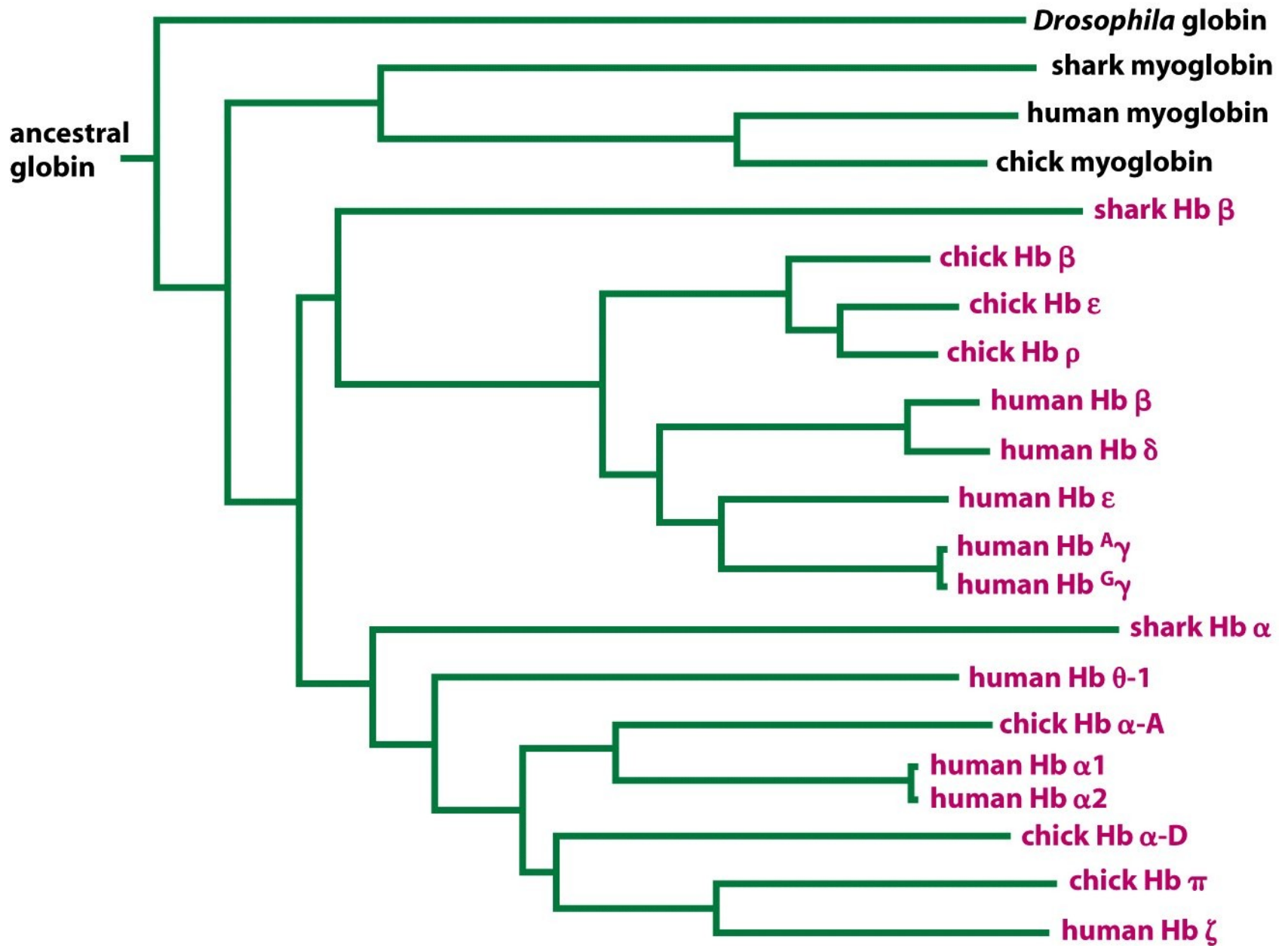
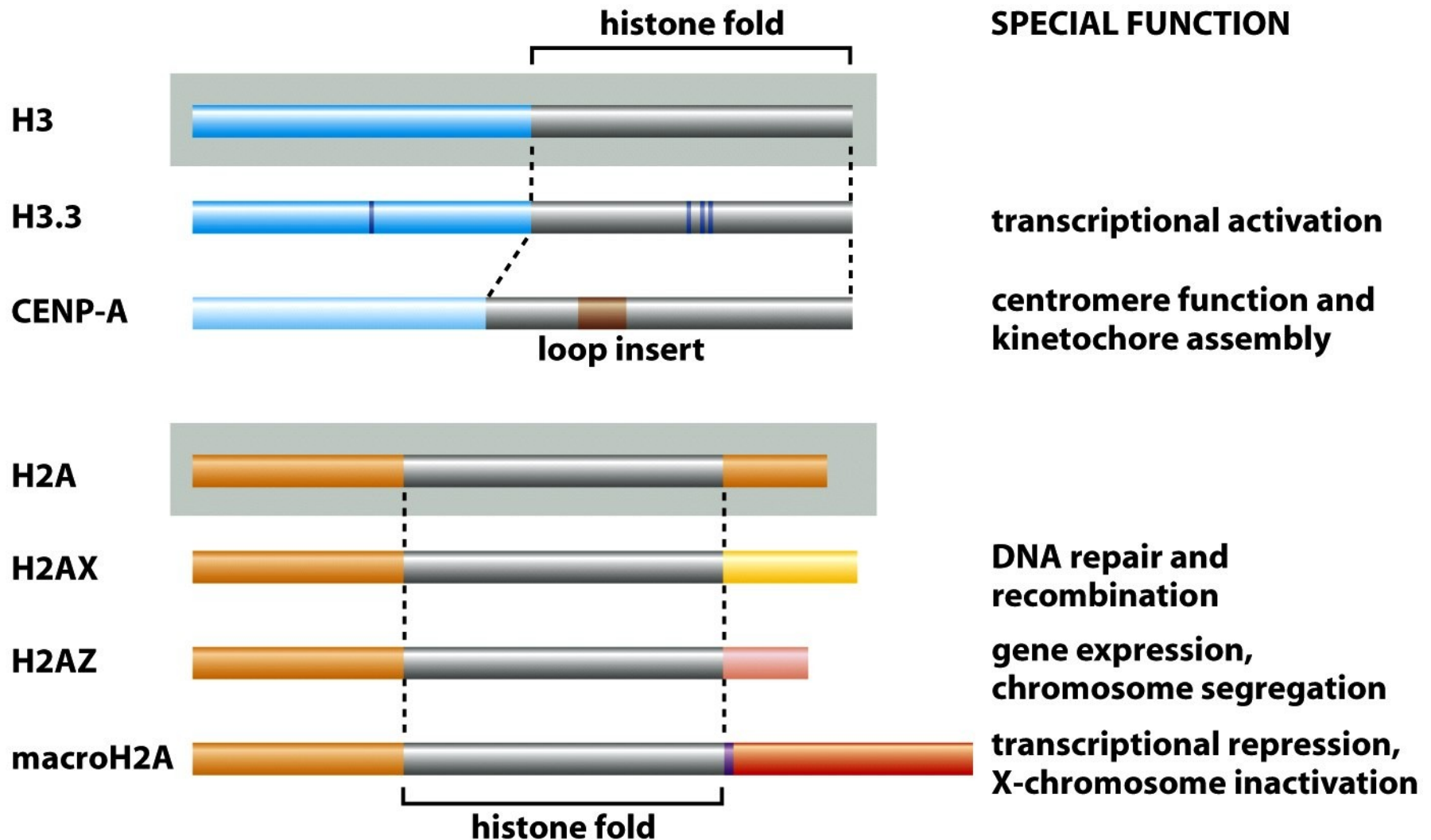


Figure 1-26 *Molecular Biology of the Cell*, Fifth Edition (© Garland Science 2008)

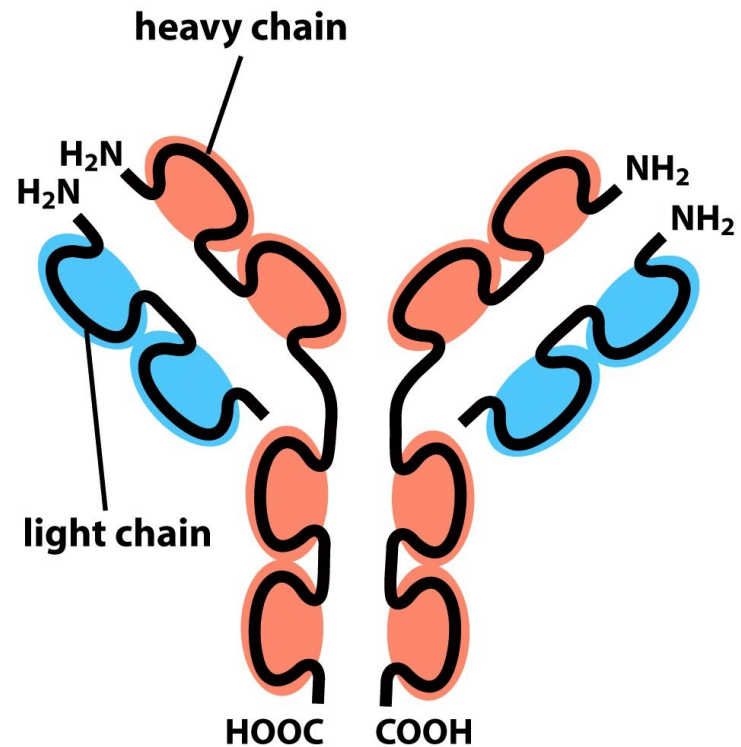
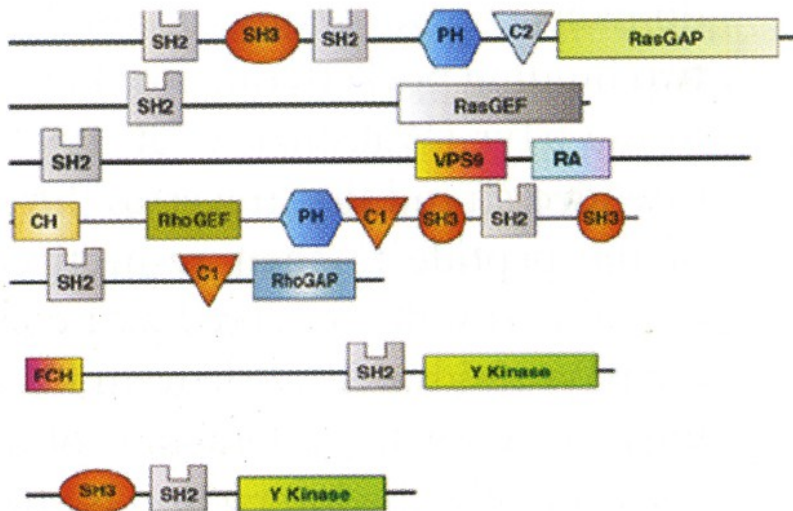


- následují **mutace** – modifikující původní protein na novou variantu (podobná funkce)

Příklad: histony - histon fold + odlišné sekvence (odlišně PTM) odlišné funkce (stále v rámci nukleosomů/chromatinu)



- duplikace celého genu nebo pouze **domény** – většina domén v proteinech obratlovců pochází/existuje v bezobratlých – 7% lidských proteinů/domén je specifických pro obratlovce)
- opakování domén (imunoglobuliny, kolageny apod.) nebo spojování různých domén
- fúze domén naznačuje funkční příbuznost nebo dokonce přítomnost v komplexu (v organismu s nefúzovanými doménami - ChimeraDB)



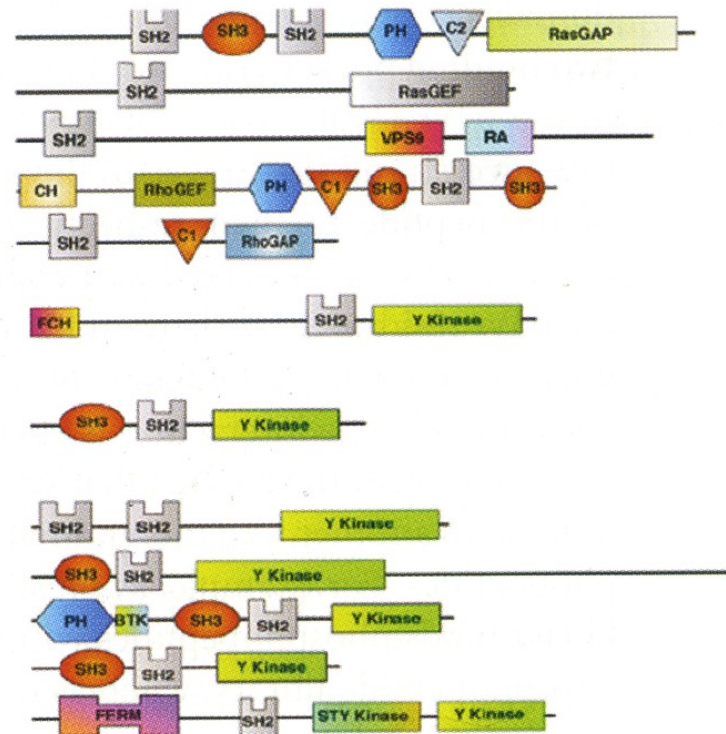
- duplikace **domény**
- domény jsou většinou kódovány **introny** – bez intronů by bylo obtížnější přesně vybrat pouze „doménovou“ část sekvence
- duplikací a „**shuffling**“ **domén** – poskládají se nové geny/proteiny – vytváří nová funkční/fyzická provázání (interakce => „networks“)

Small GTPase  
Signaling

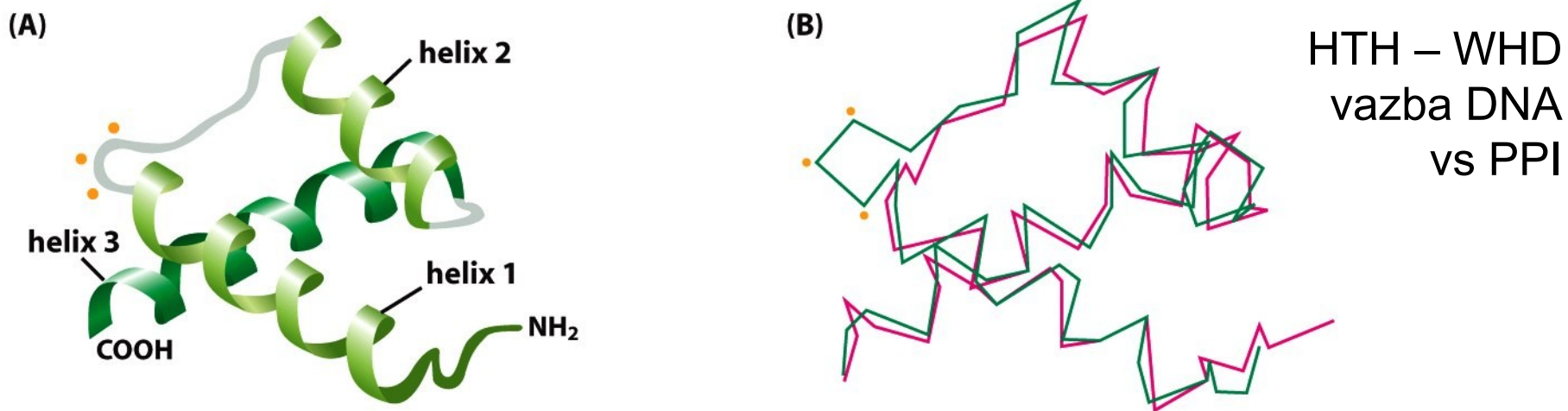
Ras-GAP  
Nsp1,2,3  
Rin1  
Vav1,2,3  
Chimerin

Kinases

Fps, Fer  
Src, Csk, Ctk/Hyl,  
Fgr, Fyn, Yes, Hck,  
Lck, Lyn, Bik, Frk,  
Brk, DJ697K14.1  
Zap70, Syk  
c-Abl, Arg/Abl2  
Btk, Tec, Itk, Bmx  
Txk  
Jak1,2,3,Tyk2



- duplikací a „shuffling“ domén – poskládají se nové geny/proteiny
- nová doména (protein) není pod tlakem „funkčním“, ale zůstává pod tlakem „strukturním“ („misfold“ = degradace)



(C) **Hydrofobní – struktura, bazické – vazba DNA**

yeast H<sub>2</sub>N

GHRFTKENVRILESWFAKNIENPYLDTKGLENLMKNTSLSRIQIKNWVSNRRRKEKTI

RTAFSSÉOLARLKREFNEN - - - RYLTERRRQQLSSELGLNEAQIKIWFQNKRAKIKKS COOH

*Drosophila*

Figure 3-13 *Molecular Biology of the Cell* (© Garland Science 2008)

- zůstává hydrofobní profil (hydrofobní AMK jsou uvnitř a drží strukturu domény) – mění se povrch (můžete hledat „paralogy“ i podle 3D struktury, sekvence <25% ident. – funkce ? odlišná)

**B**

# Hydrofobní – struktura, hydrofobní povrch – PPI

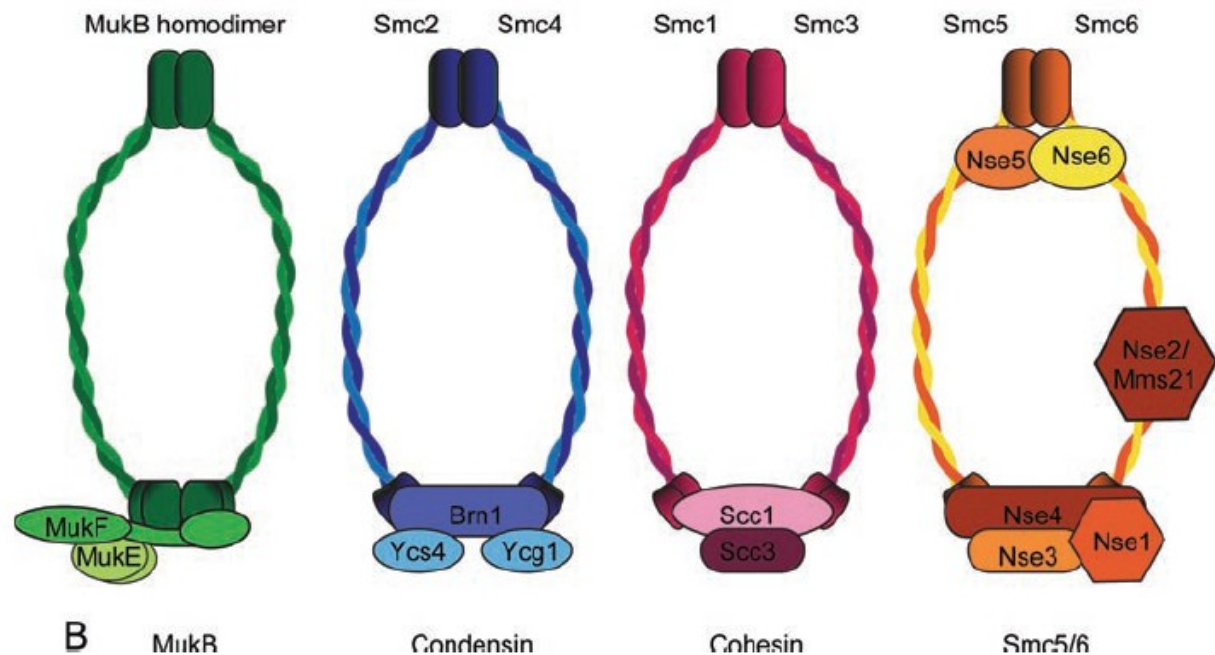
Species	Sequence	Position	Species	Sequence	Position
<i>Bacillus s.</i>	EEVQVKIDTFEGP	1-13	<i>Bacillus s.</i>	LDLLLHLINRL	14-26
<i>Mycoplasma p.</i>	LEENVFLEDYNGT	1-13	<i>Mycoplasma p.</i>	LPNLELLIDK	14-26
<i>Mycobacterium t.</i>	AGFRVRLTNFEGP	1-13	<i>Mycobacterium t.</i>	FDLLLQLIFAH	14-26
<i>Fusobacterium n.</i>	EEVVKLNNEFEGP	1-13	<i>Fusobacterium n.</i>	FDLLLNLIENK	14-26
<i>Pseudomonas a.</i>	MEVFLEAFEGP	1-13	<i>Pseudomonas a.</i>	LDLLLVLIRKQ	14-26
<i>Rhodospirillum r.</i>			<i>Rhodospirillum r.</i>	MLLTLARSQ	14-26
<i>Treponema p.</i>	PVQEFKLSQFEGP	1-13	<i>Treponema p.</i>	LDLLLFLLIKN	14-26
<i>Thermotoga m.</i>	MDLVFKLVPEFEGP	1-13	<i>Thermotoga m.</i>	LDLLLVLVRRK	14-26
<i>Prochlorococcus m.</i>	LNRGADSGARLA	1-13	<i>Prochlorococcus m.</i>	IRLLQDAAEGR	14-26
<i>Thermosynechococcus e.</i>	PFVMSQSFADTA	1-13	<i>Thermosynechococcus e.</i>	IDILIELAERG	14-26
<i>Chloroflexus a.</i>	MPYAVTLPEFTGP	1-13	<i>Chloroflexus a.</i>	LDLLLRLEIRA	14-26
<i>Cytophaga h.</i>	ASFELQLPLFAGP	1-13	<i>Cytophaga h.</i>	FDLLFLIERD	14-26
<i>Aquifex a.</i>	MKYFPHTEHP	1-13	<i>Aquifex a.</i>	FSLVLPLIEEG	14-26
<i>Archaeoglobus f.</i>			<i>Archaeoglobus f.</i>	MAKRG	14-26
<i>Pyrococcus a.</i>	MERFEPEVTP	1-13	<i>Pyrococcus a.</i>	VDILLQLVKG	14-26
<i>Methanococcus j.</i>	MIDSNFDIVLN	1-13	<i>Methanococcus j.</i>	VRMIKESIEKK	14-26
<i>Thermoplasma v.</i>	MRIDRSVDFVKSI	1-13	<i>Thermoplasma v.</i>	ILEILSLCSDG	14-26
<i>Homo s.</i>	MFYAHFVLKRRGP	1-13	<i>Homo s.</i>	LAKIWLAAHWDK	14-26
<i>Drosophila m.</i>	MFYEHIIIAKKGP	1-13	<i>Drosophila m.</i>	LARIWLAHWDK	14-26
<i>Caenorhabditis e.</i>	MFYADFVLKRRGP	1-13	<i>Caenorhabditis e.</i>	LSKVWLAHWEK	14-26
<i>Schizosaccharomyces p.</i>	MFYSEALIKRRGP	1-13	<i>Schizosaccharomyces p.</i>	LAKVWLAHWEK	14-26
<i>Saccharomyces c.</i>	RLIVLRLATNKG	1-13	<i>Saccharomyces c.</i>	LAQIWLASNMS	14-26
<i>Arabidopsis t.</i>	MFYSHTLLARKGP	1-13	<i>Arabidopsis t.</i>	LGTVMCAAHVHG	14-26
<i>Leishmania m.</i>	MPFSTIYVLTKKGP	1-13	<i>Leishmania m.</i>	LAKVWLAHWDK	14-26
<i>Encephalitozoon c.</i>	MSKAMGALCLANP	1-13	<i>Encephalitozoon c.</i>	QDVLKAFVVER	14-26
<i>Homo s.</i>	MFYYPNVLQRHTG	1-13	<i>Homo s.</i>	FATIWLAATRGS	14-26
<i>Drosophila m.</i>	SAPGQQRNREIAD	1-13	<i>Drosophila m.</i>	VDVAVCKQITE	14-26
<i>Caenorhabditis e.</i>	MVYSAEVIRKDAV	1-13	<i>Caenorhabditis e.</i>	FHVAVLGTGDSK	14-26
<i>Schizosaccharomyces p.</i>	FYNQDVLTKKGG	1-13	<i>Schizosaccharomyces p.</i>	MGIWLAATLSKSH	14-26
<i>Saccharomyces c.</i>	LSLNFKDDKRYK	1-13	<i>Saccharomyces c.</i>	LTTVMWLLSALGNS	14-26
<i>Arabidopsis t.</i>	MFYSHQQLARKAP	1-13	<i>Arabidopsis t.</i>	LGQIWMATLHA	14-26
<i>Homo s.</i>			<i>Homo s.</i>	MEDVEARFAHLLQ	14-26
<i>Drosophila m.</i>	VAKCIRAYDEQQ	1-13	<i>Drosophila m.</i>	LDLWLVLEELS	14-26
<i>Caenorhabditis e.</i>	AIMTRNAPPGQES	1-13	<i>Caenorhabditis e.</i>	IDLAVLTPAK	14-26
<i>Arabidopsis t.</i>	MSHGGGEVVRGER	1-13	<i>Arabidopsis t.</i>	IHTVQ	14-26
<i>Plasmodium f.</i>	MSTZDELNLLIQNLQ	1-13	<i>Plasmodium f.</i>	KCN	14-26
<i>Cryptosporidium j.</i>			<i>Cryptosporidium j.</i>		14-26
<i>Homo s.</i>			<i>Homo s.</i>		14-26
<i>Drosophila m.</i>	Scc1	1-13	<i>Drosophila m.</i>	GPLAQIWLASNMS	14-26
<i>Schizosaccharomyces p.</i>	Scc1	1-13	<i>Schizosaccharomyces p.</i>	NIPRGSV	14-26
<i>Saccharomyces c.</i>	B.s.	1-13	<i>Saccharomyces c.</i>	IQTHIAESAKEIAKASG	14-26
<i>Arabidopsis t.</i>	B.s.	1-13	<i>Arabidopsis t.</i>	[ 8 ] -EYITARTSGELLQGI	14-26
<i>Plasmodium f.</i>	B.s.	1-13	<i>Plasmodium f.</i>	IVRVYSKQATFL	14-26
<i>Encephalitozoon c.</i>	B.s.	1-13	<i>Encephalitozoon c.</i>		14-26
<i>S.p.</i>	SIDKRDLRKRYRN	1-13	<i>S.p.</i>	ESRLELVDEENN	14-26
<i>S.c.</i>	PQLEFKVLQGYRD	1-13	<i>S.c.</i>	ESEMHLKGR	14-26
<i>E.c.</i>	VMDEDEVHDRYL	1-13	<i>E.c.</i>	NLLQEMRLLKED	14-26
<i>E.h.</i>	PNDRSAVRKE	1-13	<i>E.h.</i>	LRIMQEIQQAKHELLSN	14-26
<i>D.d.</i>	IGERRELKRYRT	1-13	<i>D.d.</i>	LIQDVS	14-26
<i>A.t.</i>	IADRRAIRSQW	1-13	<i>A.t.</i>	LAATHKIK	14-26
<i>A.t.</i>	IADRRAIRSQW	1-13	<i>A.t.</i>	LAATHKIK	14-26
<i>A.t.</i>	ISDRRILRSKYL	1-13	<i>A.t.</i>	SONEIN	14-26
<i>M.m.</i>	REKCRKIRROY	1-13	<i>M.m.</i>	RQLIYTVQ	14-26
<i>M.m.</i>	OGLCRQIRHOY	1-13	<i>M.m.</i>	RALINSVQ	14-26
<i>H.s.</i>	EKCRSIRROY	1-13	<i>H.s.</i>	RQMYCVR	14-26
<i>H.s.</i>	OGLCRQIRHOY	1-13	<i>H.s.</i>	RALINSVQ	14-26

h h h h \* \*hh h hh h h h h \*h h\*h hhh h h h



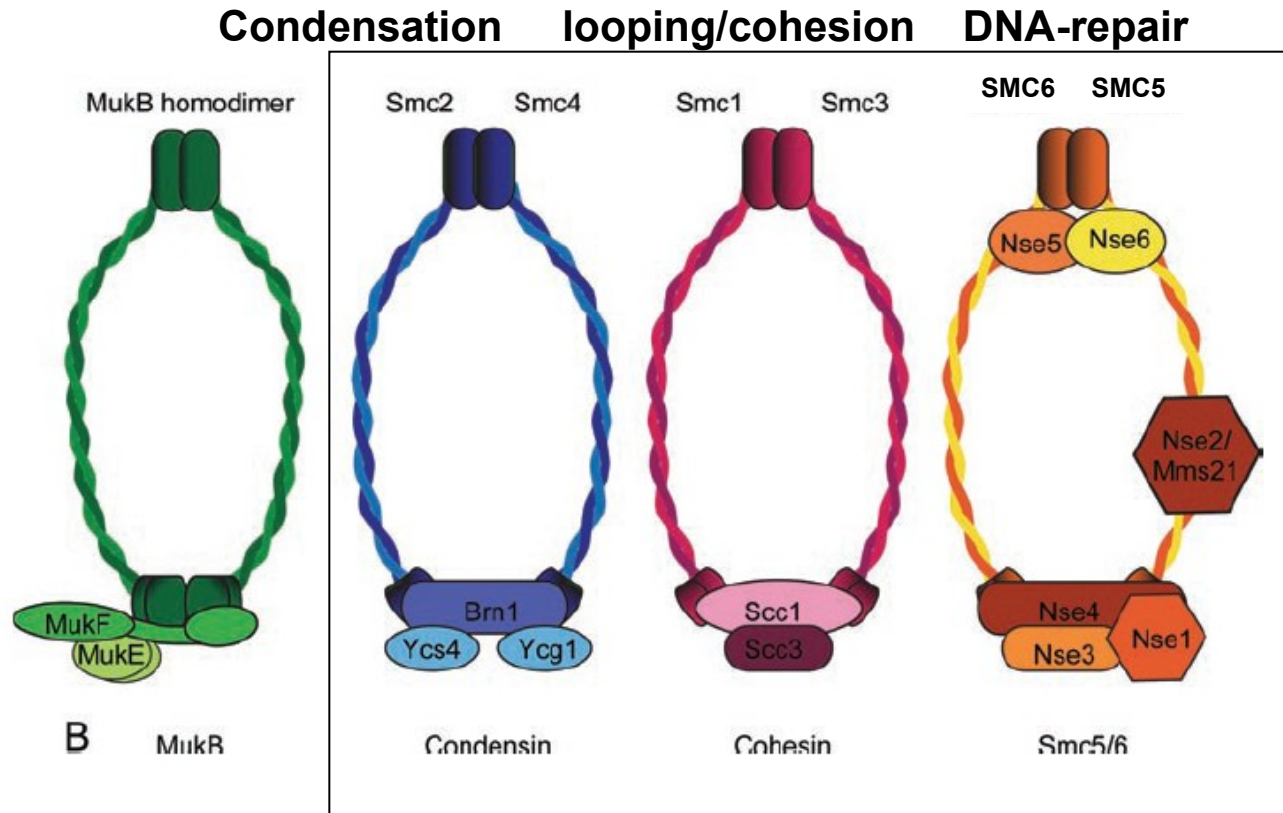
Schleifer et al,  
Palecek et al,

- „staré“ proteiny jsou konzervovány „strukturně“ a „funkčně“ – funkčně znamená většinou „konzervaci“ interakcí (většinou PPI)
- nová doména (protein) je pod tlakem „strukturním“ („misfold“ = degradace) – je ale i pod tlakem „povrchovým“ tj. tendencí povrchů interagovat (zvláště hydrofobní povrchy – hydrofobní povrch je rozeznáván chaperony a bez interakčního partnera degradován)
- z toho plyne tendence duplikovaných proteinů vytvářet podobné komplexy (podobné interakce)



# Evoluce SMC komplexů

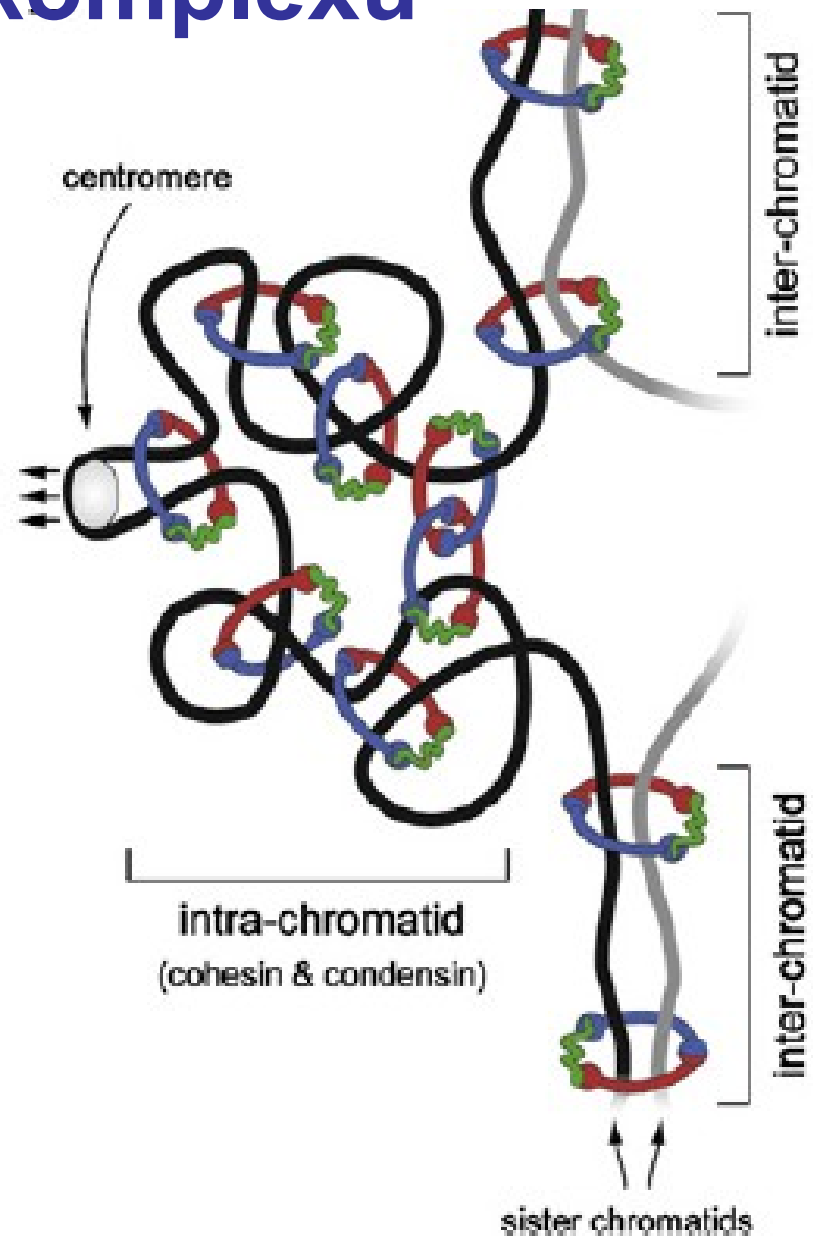
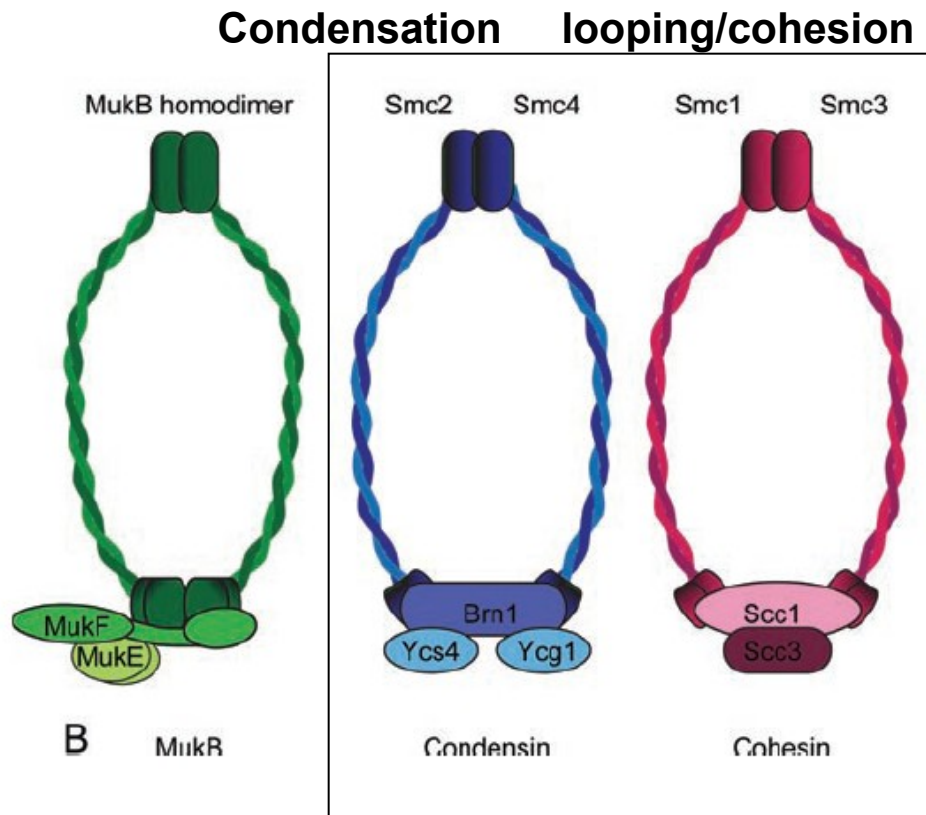
- bakterie mají po jednom komplexu složeném z homodimeru SMC/MukB a 2 Nse



- eukaryota mají 3 SMC komplexy – SMC heterodimery + kleisin (další podjednotky nepříliš konzervované) - (příklad využití konzervovaného motivu a alterace částí systému/komplexu)

# Evoluce SMC komplexů

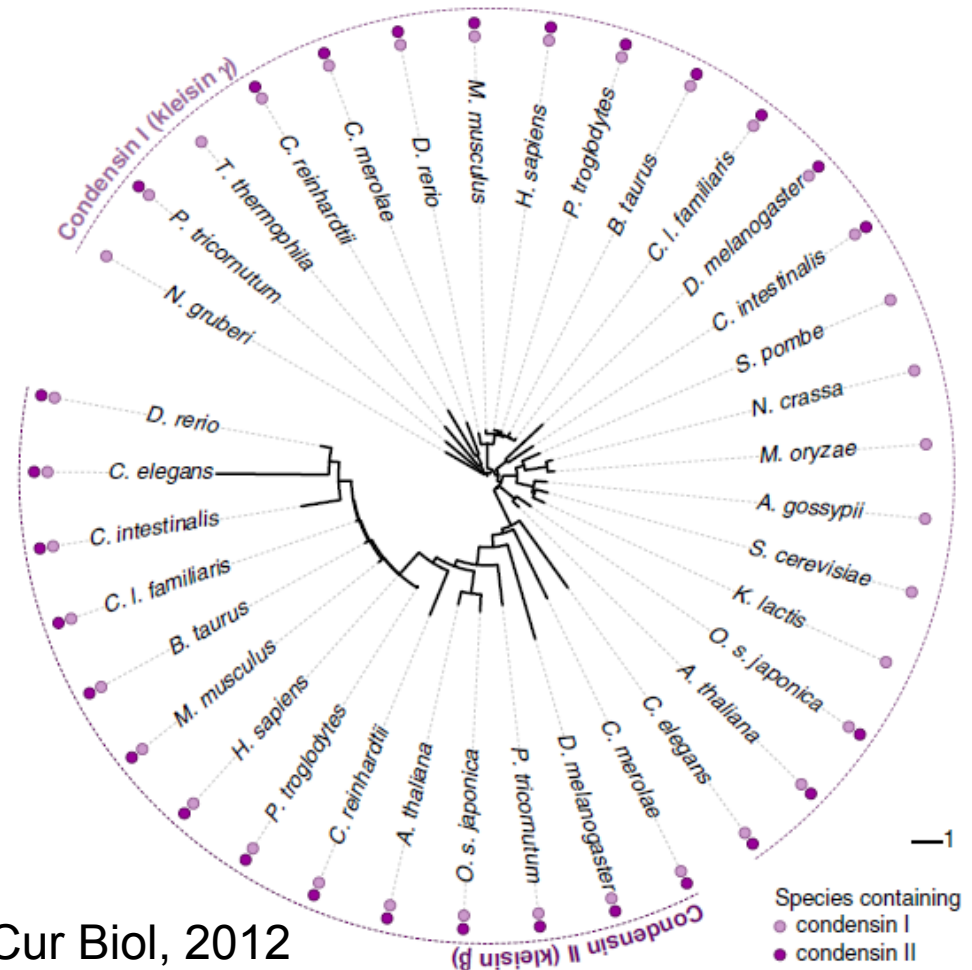
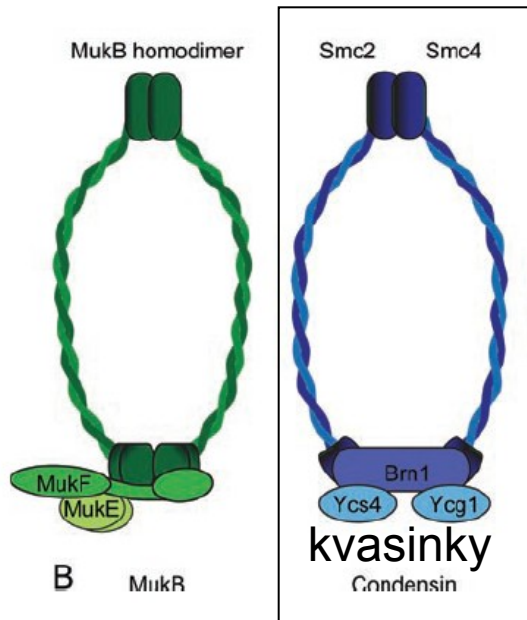
- konzervovaný motiv (kroužek)
- SMC heterodimery + kleisin



- alterace – kohese x kondensace
- (Nse podjednotky, faktory ovlivňující loading = místo, čas ...)

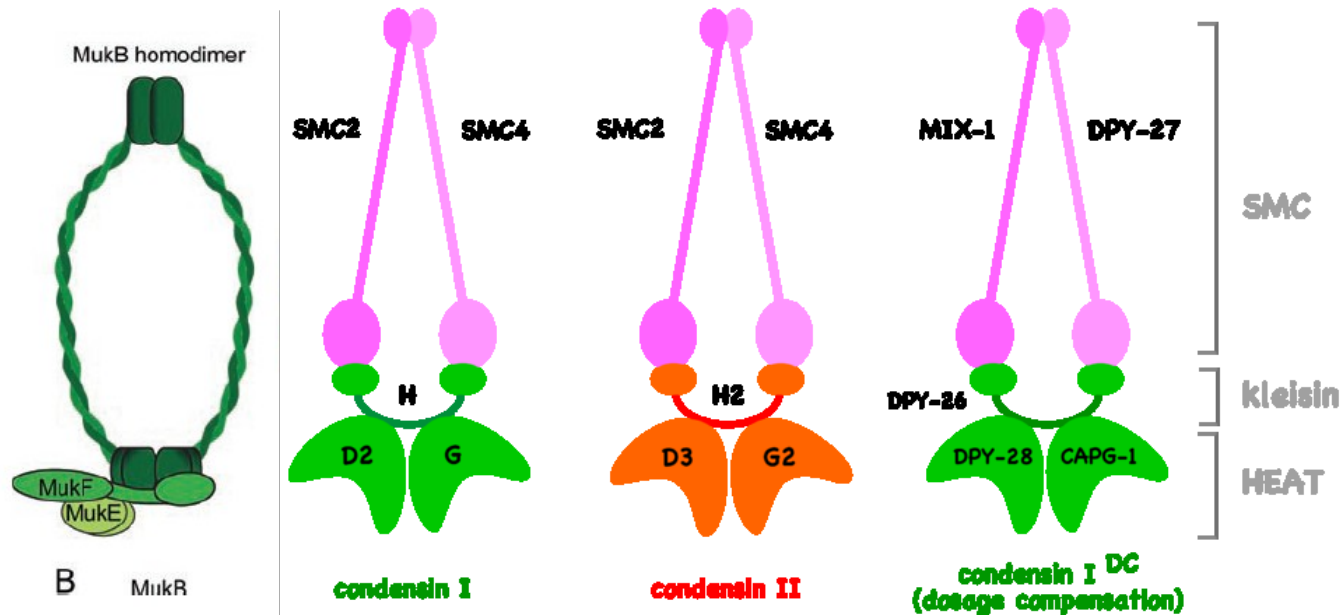


Name	<i>Saccharomyces cerevisiae</i>	<i>Schizosaccharomyces pombe</i>	<i>Drosophila</i>	Vertebrates
Smc1	Smc1	Psm1	DmSmc1	Smc1
Smc3	Smc3	Psm3	DmSmc3	Smc3
Scs1	Mcd1/Pds3	Rad21	DmRad21	Rad21
Scs3	Scs3	Psc3	DmSA	SA1 and SA2



- eukaryota mají 3 SMC kom
- kohesin = mitotický a meiot
- Scs1xRec8, navíc obratlovci
- kondensin (komplex I a II, I

Name	<i>Saccharomyces cerevisiae</i>	<i>Schizosaccharomyces pombe</i>	<i>Drosophila</i>	Vertebrates
Smc1	Smc1	Psm1	DmSmc1	Smc1
Smc3	Smc3	Psm3	DmSmc3	Smc3
Scs1	Mcd1/Pds3	Rad21	DmRad21	Rad21
Scs3	Scs3	Psc3	DmSA	SA1 and SA2



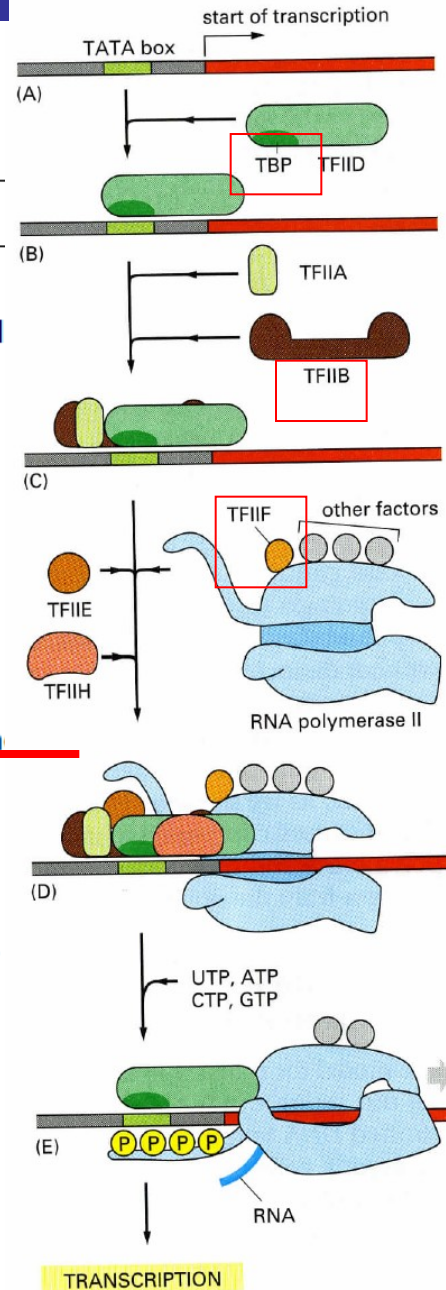
- eukaryota mají 3 SMC komplexy +
- kohesin = mitotický a meiotický (liší se kleisinem = kvasinky Scc1xRec8, navíc obratlovci Scc3 2x = SA1 a SA2)
- kondensin (I = H+D2+G; II = H2+D3+G2)
- SMC5/6 komplex (lidský = NSE4a x NSE4b-testis/meiosa specific)

# RNA pol II a podobnosti s pol I a III

TABLE 1 Components of the human general transcription machinery

Factor	Protein composition	Function
TFIIA	p35 ( $\alpha$ ), p19 ( $\beta$ ), and p12 ( $\gamma$ )	Antirepressor; stabilizes TBP-TATA complex; coactivator
TFIIB	p33	Start site selection; stabilize TBP-TATA complex; pol II/TFI
TFIID	TBP + TAFs (TAF1-TAF14)	Core promoter-binding factor Coactivator Protein kinase Ubiquitin-activating/conjugating activity Histone acetyltransferase
TFIIE	p56 ( $\alpha$ ) and p34 ( $\beta$ )	Recruits TFIIH Facilitates formation of an initiation-competent pol II Involved in promoter clearance
TFIIF	RAP30 and RAP74	<u>Binds pol II and facilitates pol II recruitment to the prom</u> Recruits TFIIE and TFIIH Functions with TFIIB and pol II in start site selection Facilitates pol II promoter escape Enhances the efficiency of pol II elongation
TFIIH	P89/XPB, p80/XPD, p62, p52, p44, p40/CDK7, p38/Cyclin H, p34, p32/MAT1, and p8/TFB5	ATPase activity for transcription initiation and promoter Helicase activity for promoter opening Transcription-coupled nucleotide excision repair Kinase activity for phosphorylating pol II CTD E3 ubiquitin ligase activity
pol II	RPB1-RPB12	<u>Transcription initiation, elongation, termination</u> Recruitment of mRNA capping enzymes Transcription-coupled recruitment of splicing and 3' end CTD phosphorylation, glycosylation, and ubiquitination

## TBP-všetchny 3 pol



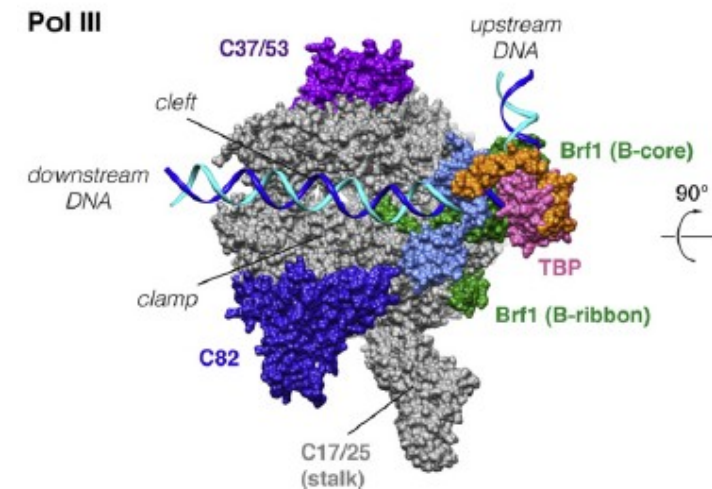
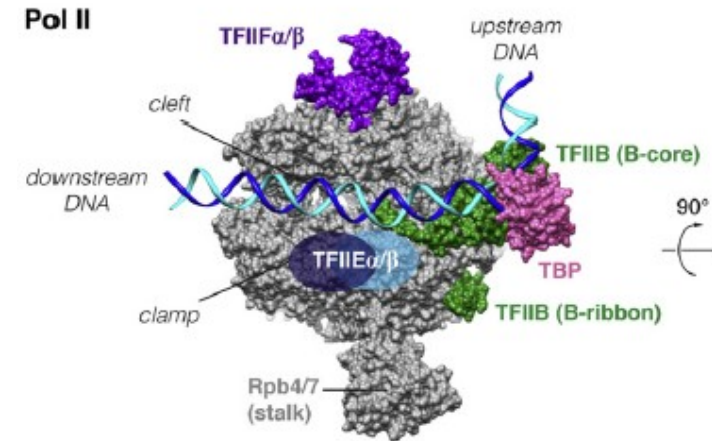
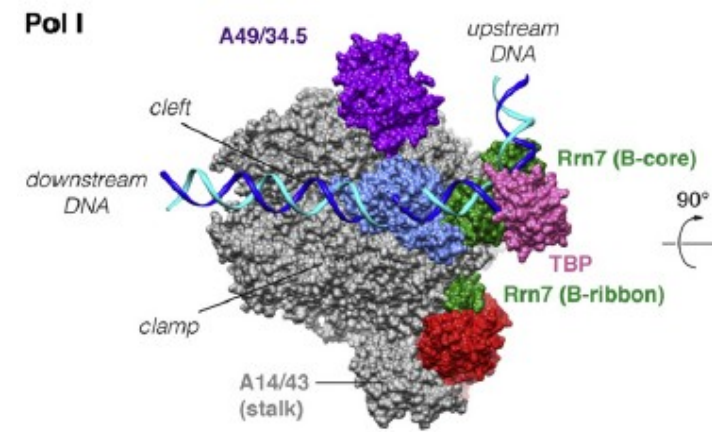
# RNA polymerasy

Pol I = 12 + A49/34.5

Pol II = 12 podjednotek (Tabulka)

Pol III = 12 + C37/53 + C82/34/31

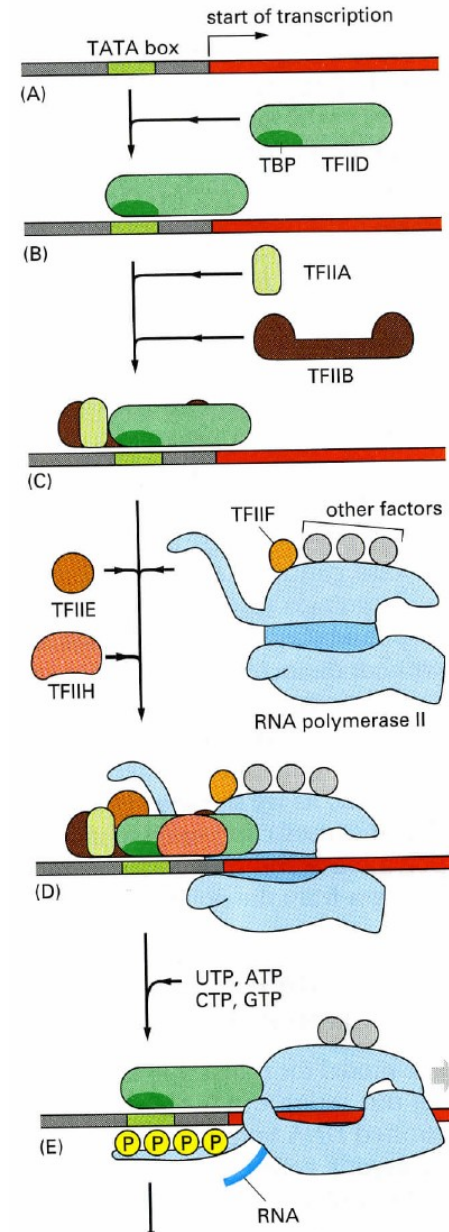
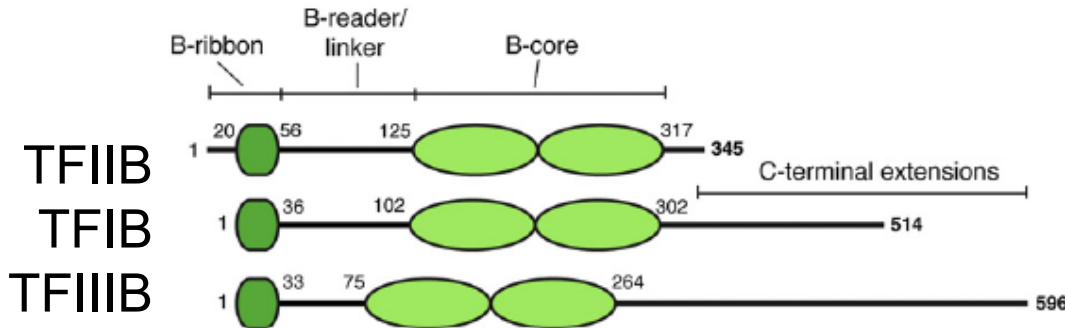
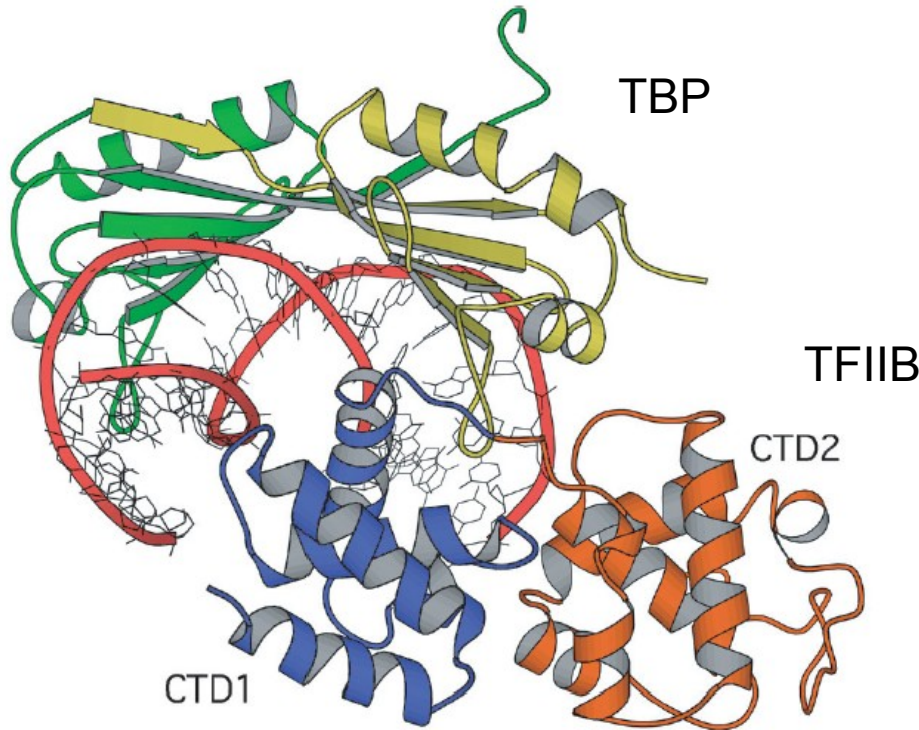
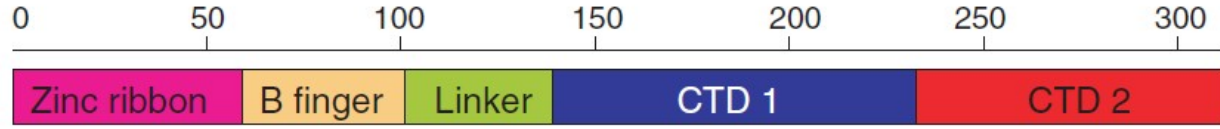
**TBP**  
**TFIIB**

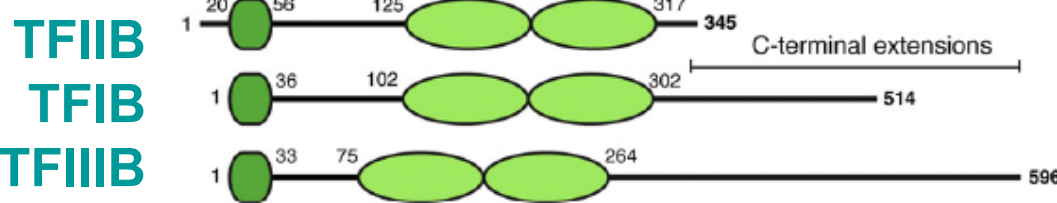
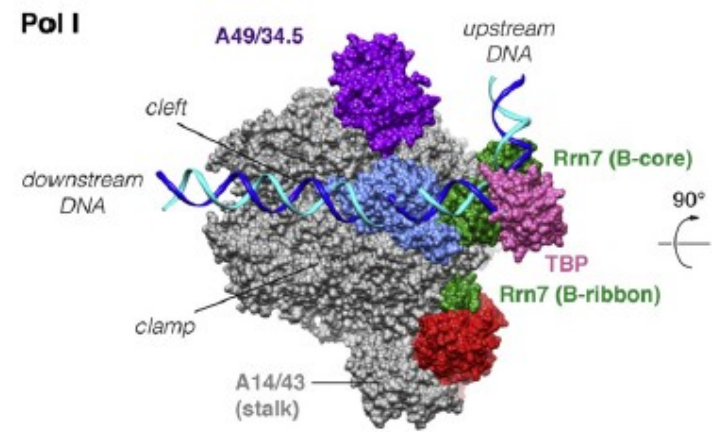


**Table 1. Yeast RNA Polymerase Subunits and Initiation Factor Homologies**

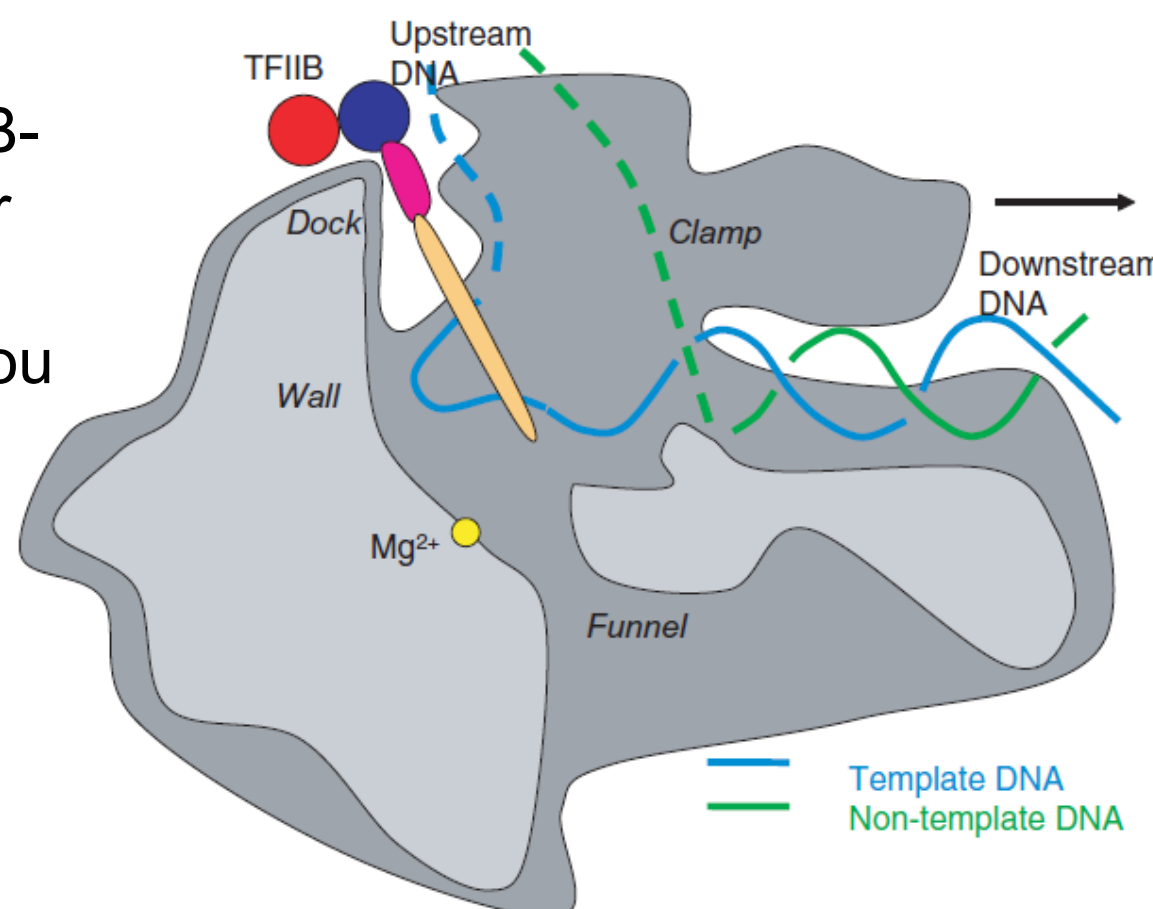
Pol II	Pol I	Pol III	Function
<b>Polymerase Core</b>			
Rpb1	A190	C160	Active center
Rpb2	A135	C128	Active center
Rpb3	AC40	AC40	
Rpb11	AC19	AC19	
Rpb9	A12.2 N ribbon	C11 N ribbon	RNA cleavage
TFIIS C-ribbon <sup>a</sup>	A12.2 C ribbon	C11 C ribbon	RNA cleavage
Rpb5	Rpb5	Rpb5	
Rpb6	Rpb6	Rpb6	
Rpb8	Rpb8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

- TFIIB (C-konec) váže TBP a 6-7bp up- a downstream od ohybu DNA přes cukrfosfátovou kostru (kolem TATA boxu)

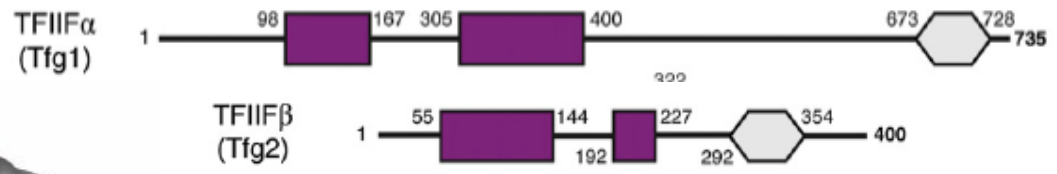
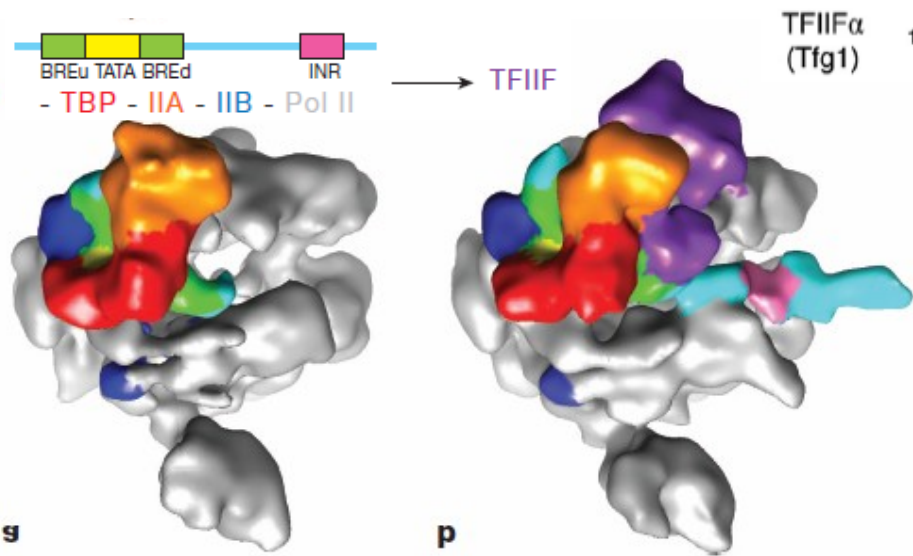




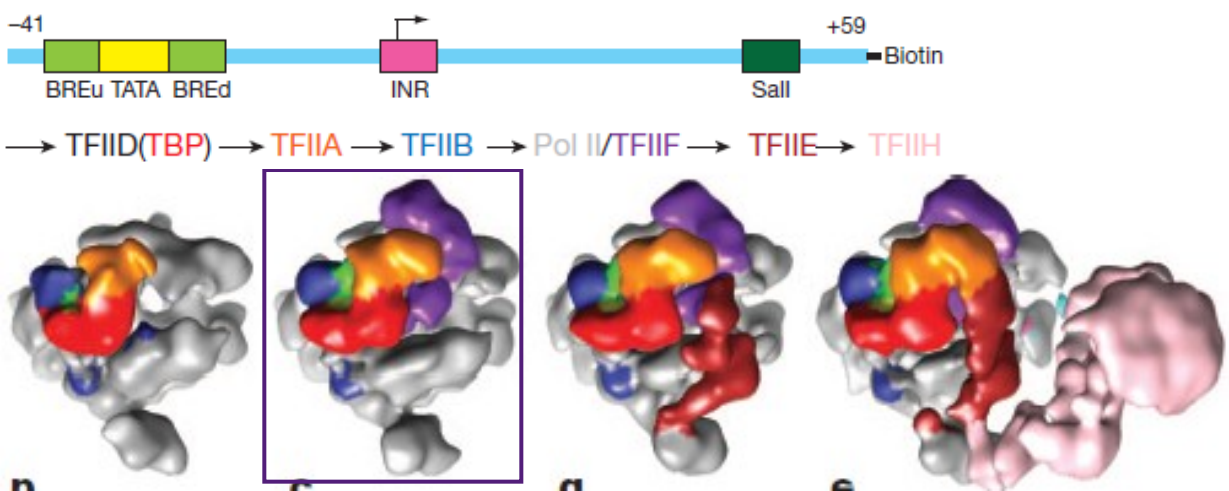
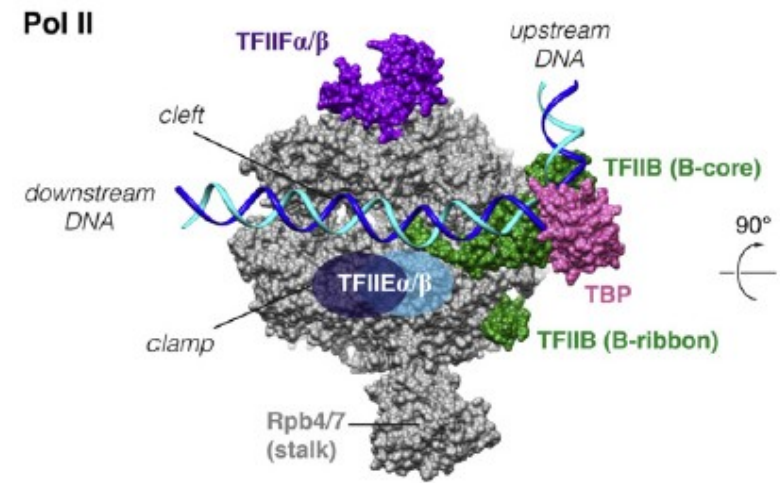
- Zn-ribbon s dock, B-finger/reader RNA tunel, B-core/CTD1 s wall, B-linker s CC clamp
- podobné natolik, že vážou TBP + ribbon a linker mohou být zaměněny pro kteroukoli polymerásu – zatímco reader je specifický pro určitou polymerásu



- **TFIIF** se váže na pol II, stabilizuje DNA v prohlubni/cleft RNA pol II a pomáhá TFIIIB s nastavením startu
- dimerizace a WHD (váže přímo DNA: BRE<sub>downstream</sub>)



Dimerizace WHD

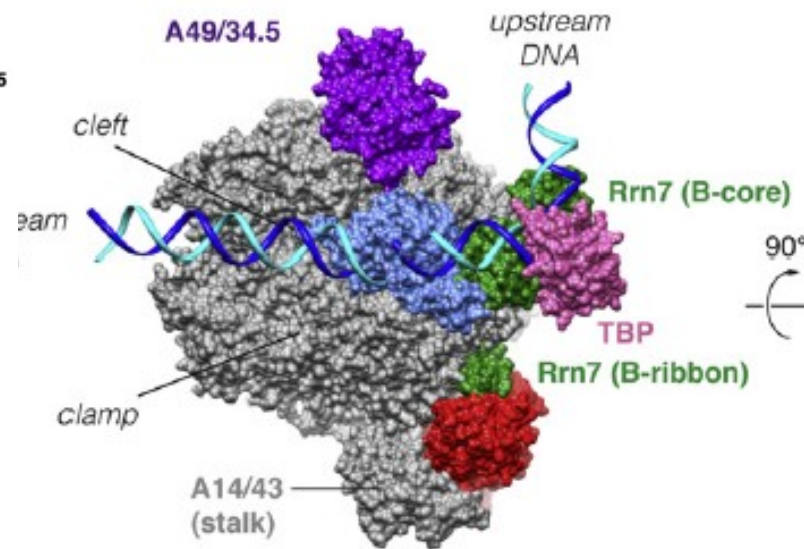
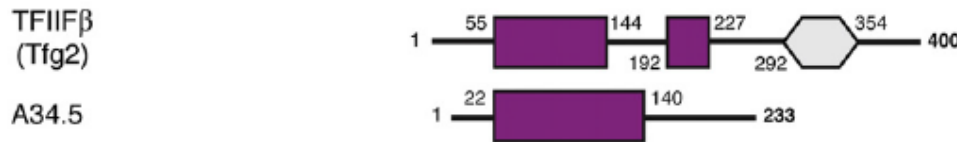


- **TFIIF** se váže na pol II, stabilizuje DNA v prohlubni/cleft RNA pol II a pomáhá TFIIB s nastavením startu

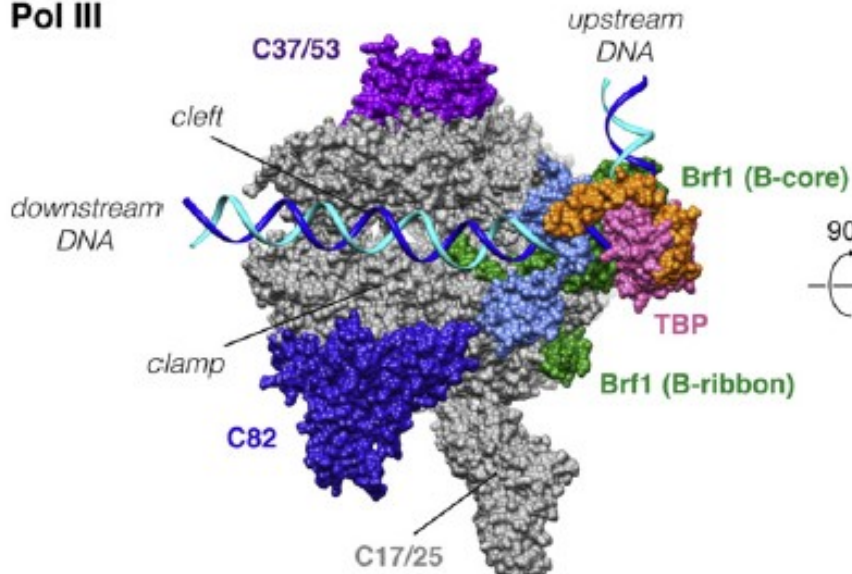
- dimerizace a WHD (váže přímo DNA: BRE<sub>downstream</sub>)



Dimerizace WHD



Pol III



Pol II

