

Marine Biological Laboratories
Workshop in Molecular Evolution

Adaptive protein evolution: Detecting changes in selection

Belinda Chang

Department of Ecology & Evolutionary Biology
Department of Cell & Systems Biology
University of Toronto

@opsinlab

Adaptive protein evolution

[1]

Phylogenetic approaches to the study of protein structure and function

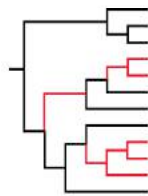
- Ancestral protein reconstruction
 - Computational analyses of selection (dN/dS)
- > Combining computational with experimental approaches allows us to test hypotheses of selection in protein evolution

Adaptive protein evolution

[2]

ADAPTIVE EVOLUTION:

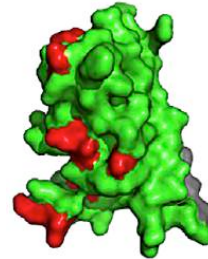
Combining computational and experimental approaches to test hypotheses of selection in protein evolution



```

GLELHPDYKTWPEQVCSFLRRGGF
GPELHPDHKTWPEQVCSFLRRGGF
GLELHPDYKTWPEQVCSFLRRGGF
GLELHPDYKTWPEQVCSFLRRGGF
GLELHPDYKTWPEQVCSFLRRGGF
GLELHDYKTWPEQVCSFLRRGGF
GLELHPDYKTWPEQVCSFLRRGGF
GLELHPDYKTWPEQVCSFLRRGGF
GLELHPDYKTWPEQVCSFLRRGGF
GLELHPDYKTWPEQVCSFLRRGGF
GLELHPDYKTWPEQVCSFLRRGGF
GLELHPDYKTWPEQVCSFLRRGGF
GLELDPDYKTWPEQVCSFLRRGGF

```



- Testing hypotheses of selection in a phylogenetic framework
- Identification of positively selected sites/regions
- Experimental studies of shifts in protein function

Adaptive protein evolution

(5)

Testing hypotheses of selection in protein evolution: Combining computational and experimental approaches

1. Formulate hypothesis of selection acting on evolution of particular group of sequences
2. Collect target group of sequences
3. Test hypothesis of selection using codon models in a phylogenetic framework
4. Identify amino acid sites/regions that have been targets of selection
5. Formulate hypothesis of selection acting on particular aspect of protein function
6. Select proteins/mutants for experimental studies
7. Test hypothesis of selection using experiments investigating shifts in protein function

Adaptive protein evolution

(6)

Testing hypotheses of selection in protein evolution: Combining computational and experimental approaches

1. Formulate hypothesis of selection acting on evolution of particular group of sequences
2. Collect target group of sequences
3. Test hypothesis of selection using codon models in a phylogenetic framework
4. Identify amino acid sites/regions that have been targets of selection
5. Formulate hypothesis of selection acting on particular aspect of protein function
6. Select proteins/mutants for experimental studies
7. Test hypothesis of selection using experiments investigating shifts in protein function

Adaptive protein evolution

[7]

Testing hypotheses of adaptive evolution in protein function: Combining computational and experimental approaches

Computational analyses

- Statistical test for positive, or divergent selection in a phylogenetic context
- Identifies amino acid residues that may have been the targets of positive selection
- Does NOT prove adaptive evolution -> Need function

Experimental studies

- Empirical test for shifts in protein function as a consequence of substitutions at sites targeted by selection
- Does NOT prove anything about selection & adaptive evolution, need evolutionary context
- Ideally combined with ancestral reconstruction and mutagenesis studies

Adaptive protein evolution

[8]

Examples of positive selection on proteins: Experimental models

Host-viral arms races

- Primate intracellular immune response TRIM5a, protein kinase R (Sawyer et al. 2005, Elde et al. 2009)
- Potato virus Y (Moury & Simon 2011)
- Neutrophil-mediated host response to infection (Loughran et al. 2012)

Sperm-egg recognition proteins

- Abalone lysin, egg coat (Aagaard et al. 2013)

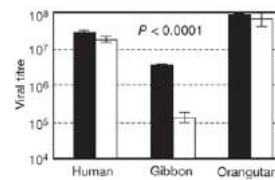
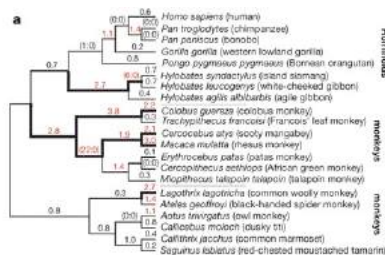
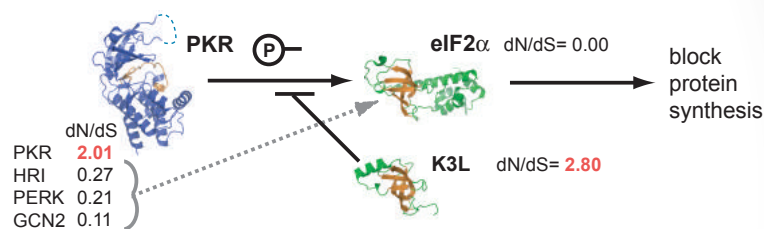
Sensory proteins

- Visual pigments in whales, fishes, birds, bats (Dungan & Chang 2017, Hauser et al. 2017)

Transcription factors in development

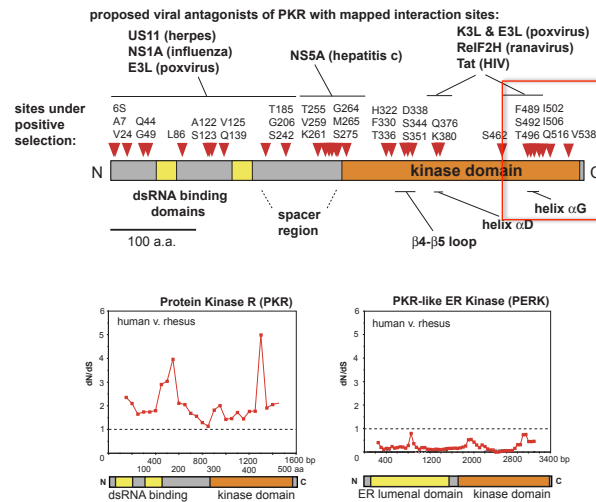
- Stem cell pluripotency (Baker et al. 2016)

Primate protein kinase R locked in arms race with viral protein mimic K3L



Elde et al. Nature 2009

Positive selection has shaped PKR in primate evolution

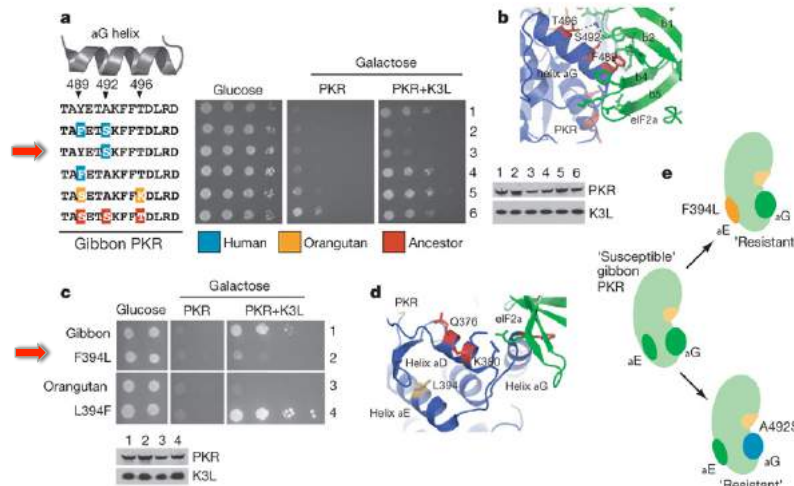


Elde et al. Nature 2009

Adaptive protein evolution

[14]

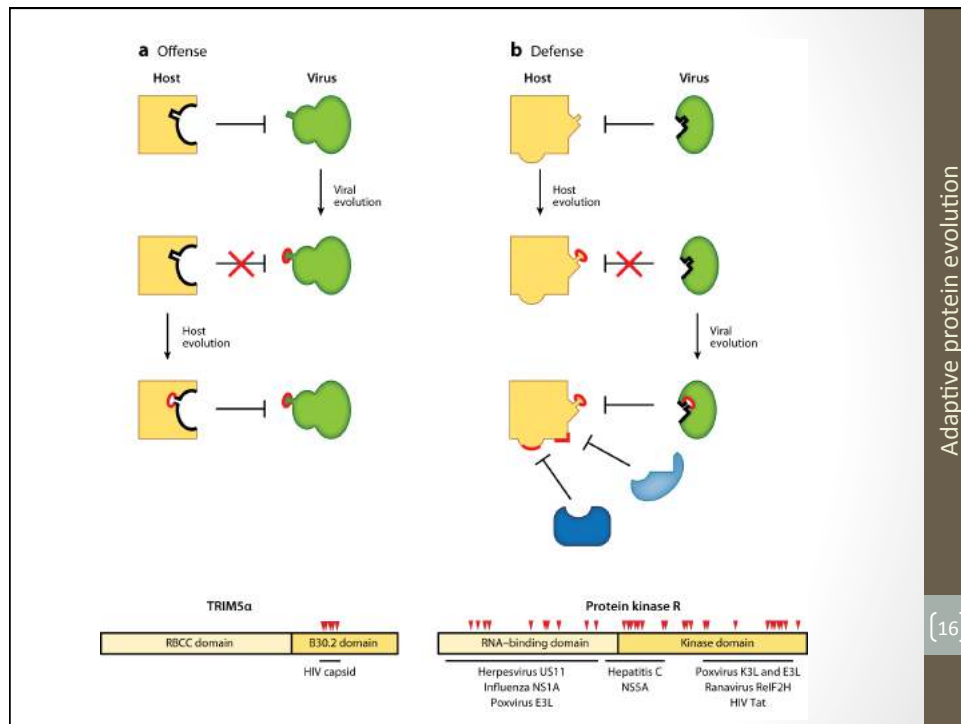
Positively selected sites on distinct surfaces of the PKR kinase domain are crucial to K3L resistance



Elde et al. Nature 2009

Adaptive protein evolution

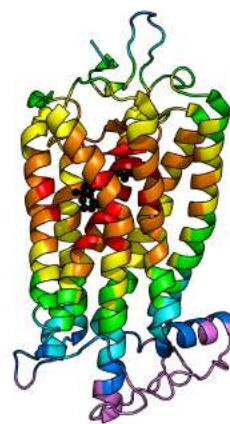
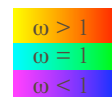
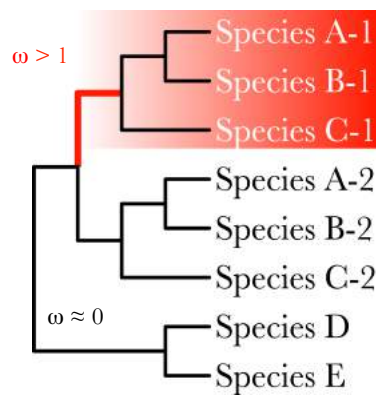
[15]



Beyond random sites and branch-sites models: Clade models

The form and strength of selection may vary across both the phylogeny and the protein.

Codon models can incorporate variation across sites and lineages, and can be used to detect positive and divergent selection.



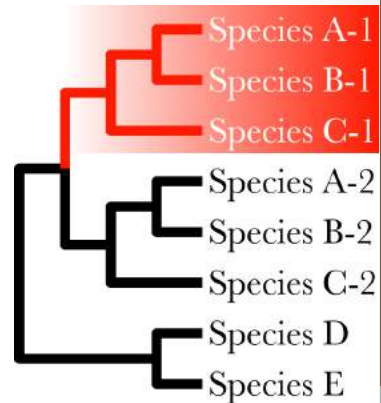
Adaptive protein evolution

[17]

Codon models of divergent selection: Clade model C (CmC)

CmC assumes sites evolve in three ways:

1. purifying selection ($0 < \omega_0 < 1$) constrains evolution at some sites, operating consistently across the tree.
2. some sites evolve free of constraint (neutrally: $\omega_1 = 1$) across the entire tree.
3. the remaining sites evolve divergently ($\omega_2 \neq \omega_3$) between particular sections of the tree.



Adaptive protein evolution

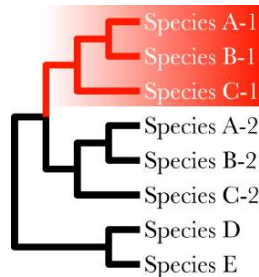
[18]

Bielawski & Yang (2004)

Codon model tests of divergent selection

CmC assumes sites evolve in three ways:

1. always purifying selection
($0 < \omega_0 < 1$)
2. always neutral
($\omega_1 = 1$)
3. divergent
($\omega_2 \neq \omega_3$)
- 3 (alt). no divergence
($\omega_2 > 0$)

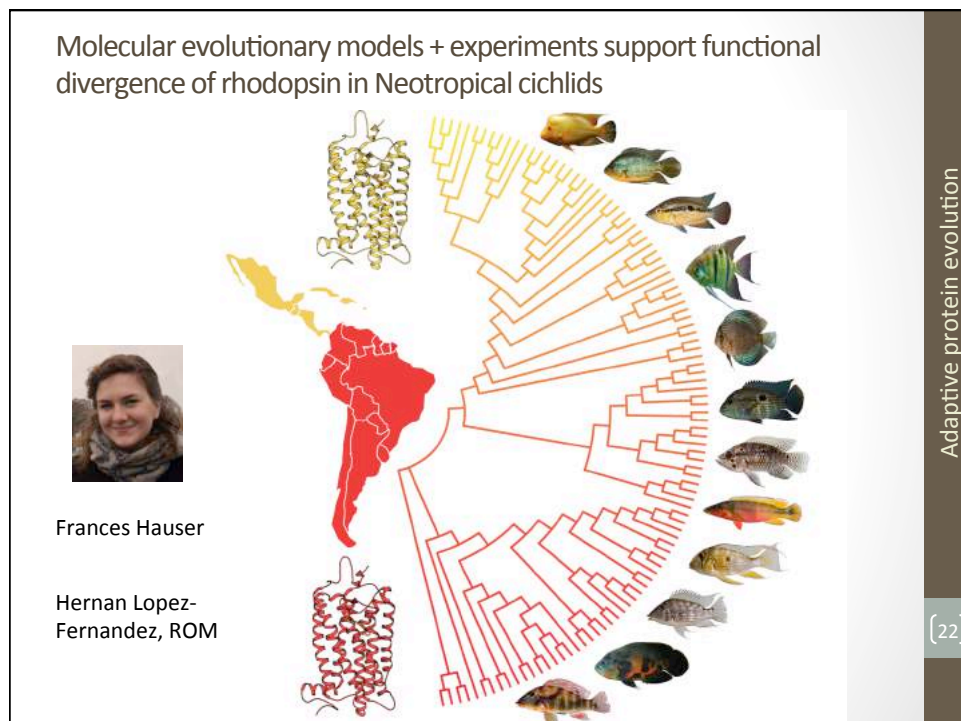
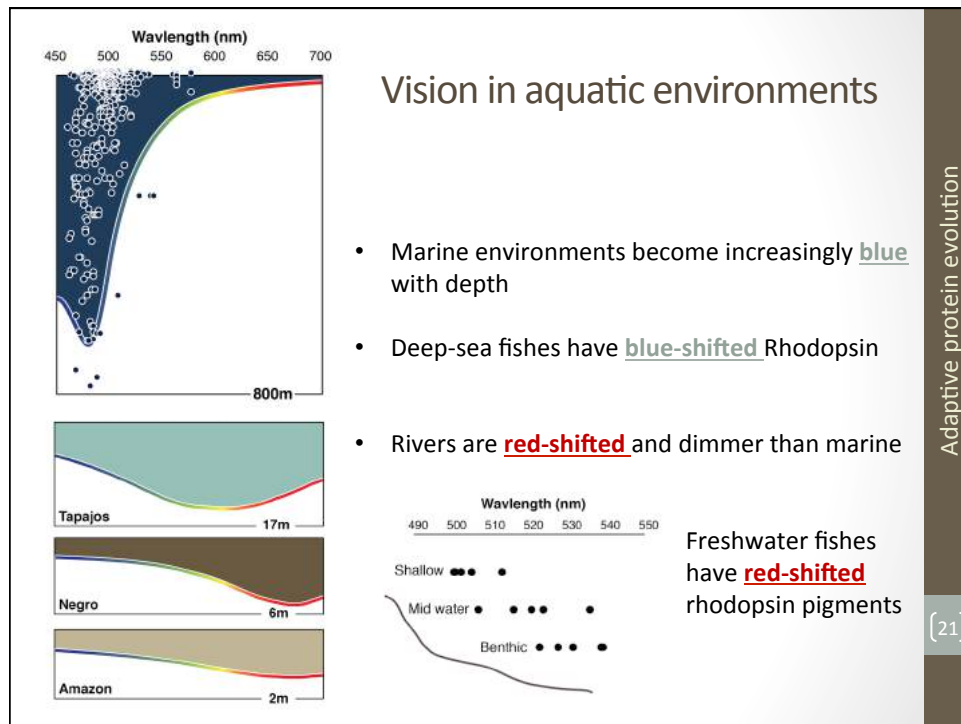


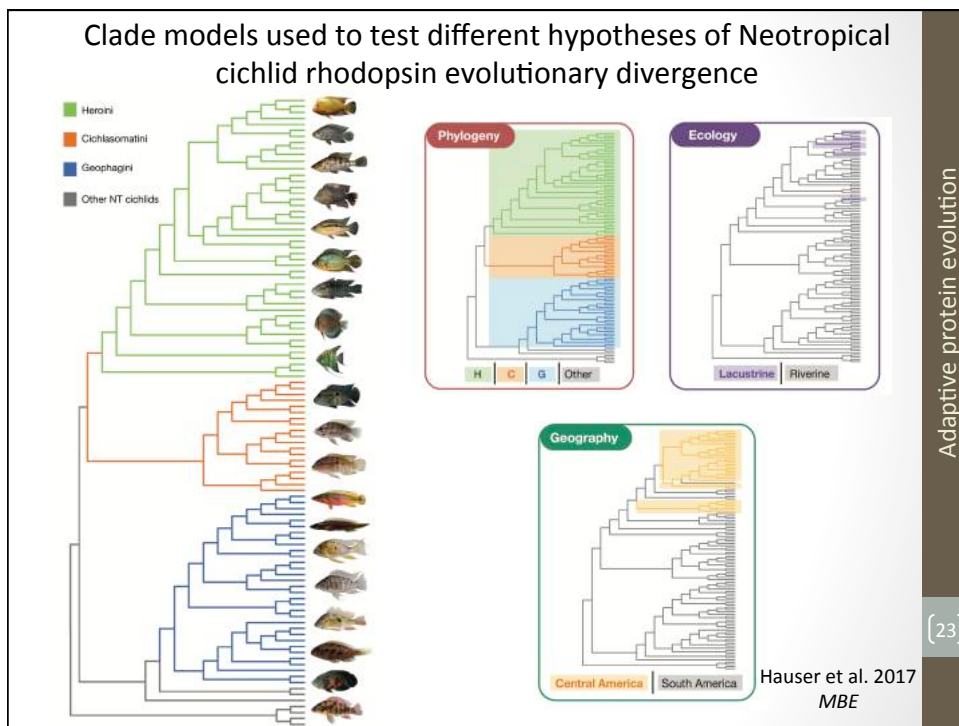
An improved null model for testing for divergent selection, M2a_rel model (Weadick & Chang 2012).

M2a_rel retains the third site class, but simply estimates a single ω ratio that applies to all branches of the phylogeny.

Adaptive protein evolution

[20]





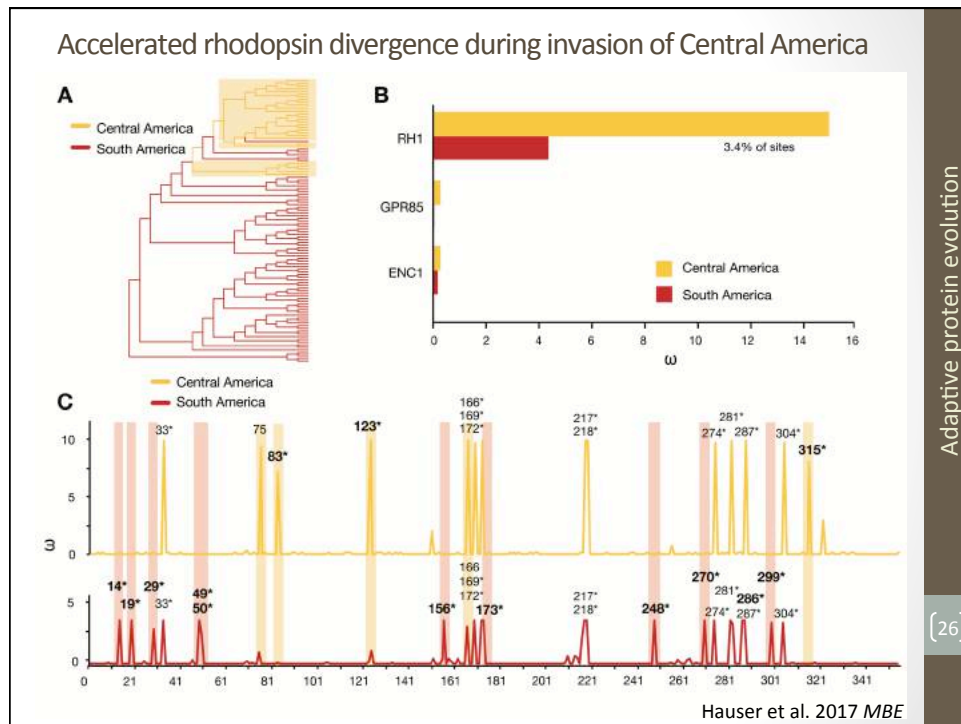
Adaptive protein evolution

(23)

Adaptive protein evolution												
Model (foreground clade)	np	lnL	k	AIC	ΔAIC	Parameters			Null	LRT	df	P
M2a_rel	210	-5909.85	3.12	12239.70	17.46	p: 0.868	0.091	0.041	M1a	250.160	2	0.0000
C+H+G	213	-5904.67	3.06	12235.34	13.10	w: 0.009	1	5.470	M2a_rel	10.360	3	0.0157
						site 0	1	2				
						proportion 0.868	0.097	0.034				
						background 0.009	1	3.979				
						Cichlasomatini 0.009	1	6.357				
Heroini 0.009	1	8.710										
Geophagini 0.009	1	4.100										
Central America (clade)	211	-5900.12	2.41	12222.24	7.34	site 0	1	2	M2a_rel	19.464	1	0.0000
						proportion 0.868	0.097	0.034				
						background 0.009	1	4.476				
						foreground 0.009	1	11.660				
Cichlasomatini	211	-5909.84	3.12	12241.68	19.44	site 0	1	2	M2a_rel	0.020	1	0.8875
						proportion 0.868	0.091	0.040				
						background 0.009	1	5.430				
						foreground 0.009	1	5.710				
Heroini	211	-5906.21	3.09	12234.42	12.18	site 0	1	2	M2a_rel	7.280	1	0.0070
						proportion 0.868	0.095	0.037				
						background 0.009	1	4.630				
						foreground 0.009	1	8.020				
Geophagini	211	-5907.70	3.10	12237.40	15.16	site 0	1	2	M2a_rel	4.300	1	0.0381
						proportion 0.868	0.094	0.038				
						background 0.009	1	6.520				
						branch 0.009	1	4.200				
Central America	211	-5896.45	3.07	12214.90	0.00	site 0	1	2	M2a_rel	26.800	1	0.0000
						proportion 0.868	0.097	3.440				
						background 0.009	1	4.500				
						Central America 0.009	1	14.800				
Lake-dwelling	211	-5909.02	3.11	12240.03	25.13	site 0	1	2	M2a_rel	1.670	1	0.1963
						proportion 0.868	0.092	0.040				
						background 0.009	1	5.410				
						foreground 0.009	1	10.250				

Adaptive protein evolution

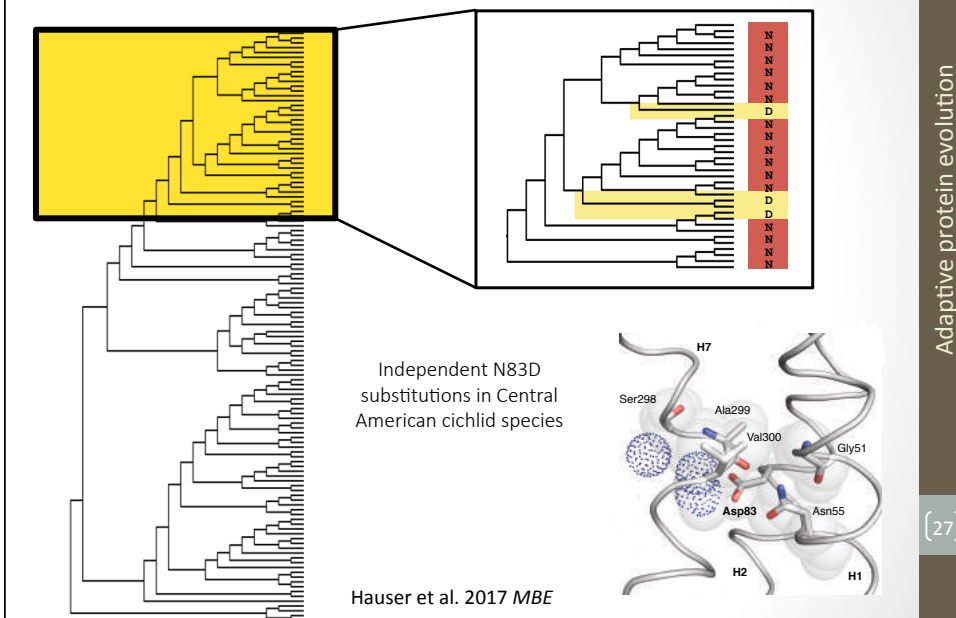
(25)



Adaptive protein evolution

[26]

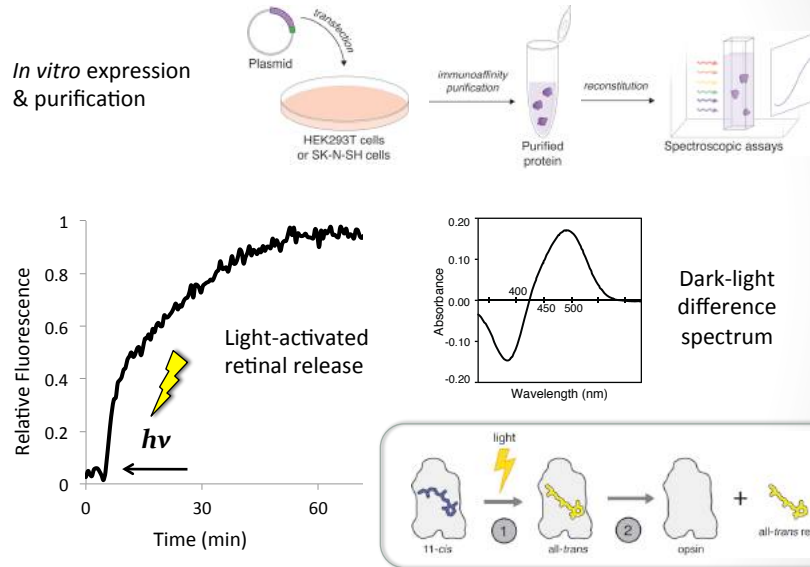
Rhodopsin site 83 identified as under positive selection and undergoes independent transitions in Central American cichlids



Adaptive protein evolution

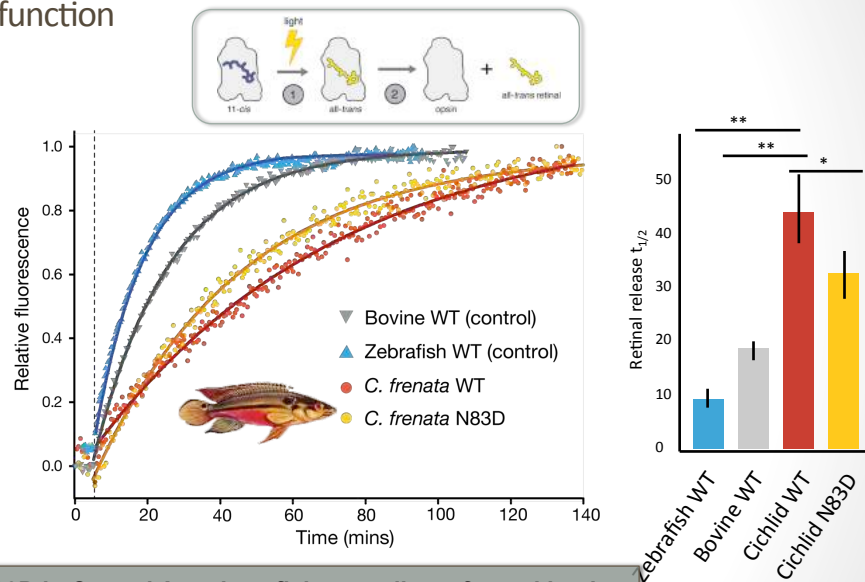
[27]

Spectroscopic assays of rhodopsin function



Adaptive protein evolution

The N83D mutation produces a significant shift in rhodopsin function



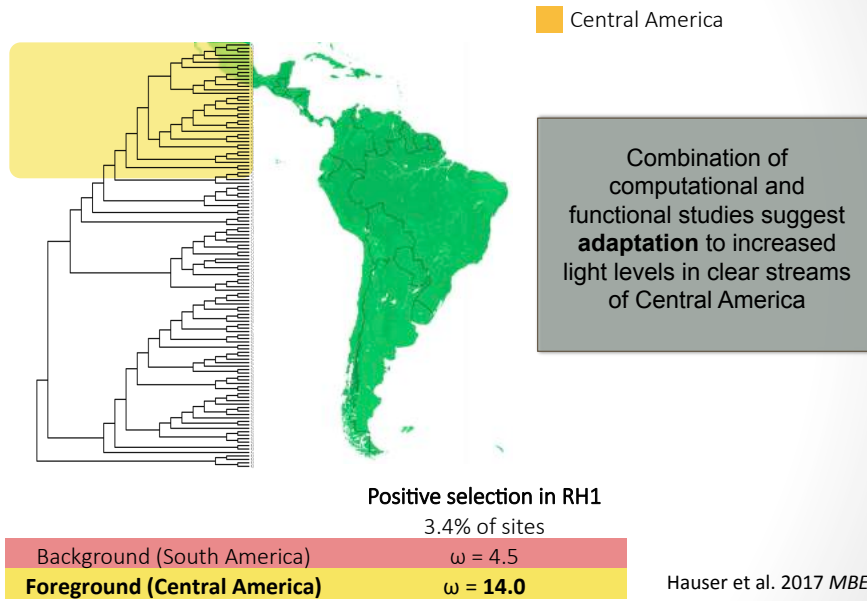
N83D in Central American fishes mediates faster kinetics

Hauser et al. 2017 MBE

Adaptive protein evolution

[29]

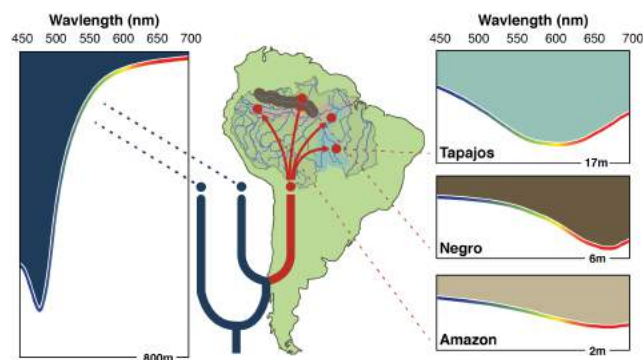
Accelerated rhodopsin divergence during invasion of Central America



Adaptive protein evolution

Rhodopsin evolution in marine-derived Amazonian anchovies

Hypothesis: Adaptive evolution in rhodopsin gene of freshwater anchovies



Alex Van Nynatten

Nathan Lovejoy
U of Toronto

- Closely related marine and freshwater clades
- **Ancestrally marine**

Van Nynatten et al. 2015 Biol Lett

Adaptive protein evolution

[31]

Vision at high altitudes: Andean catfish rhodopsin evolution



Gianni Castiglione

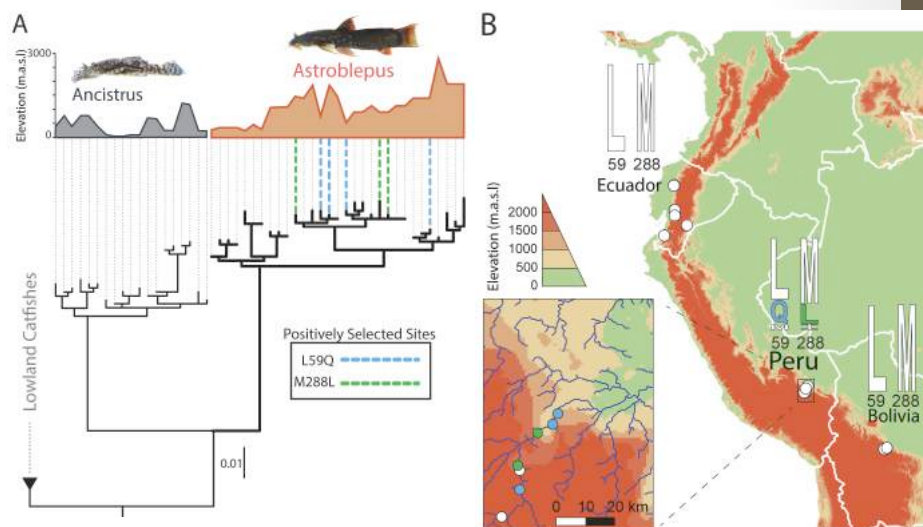
Nathan Lujan
ROM, UTSC

Castiglione et al. 2017 PNAS

Adaptive protein evolution

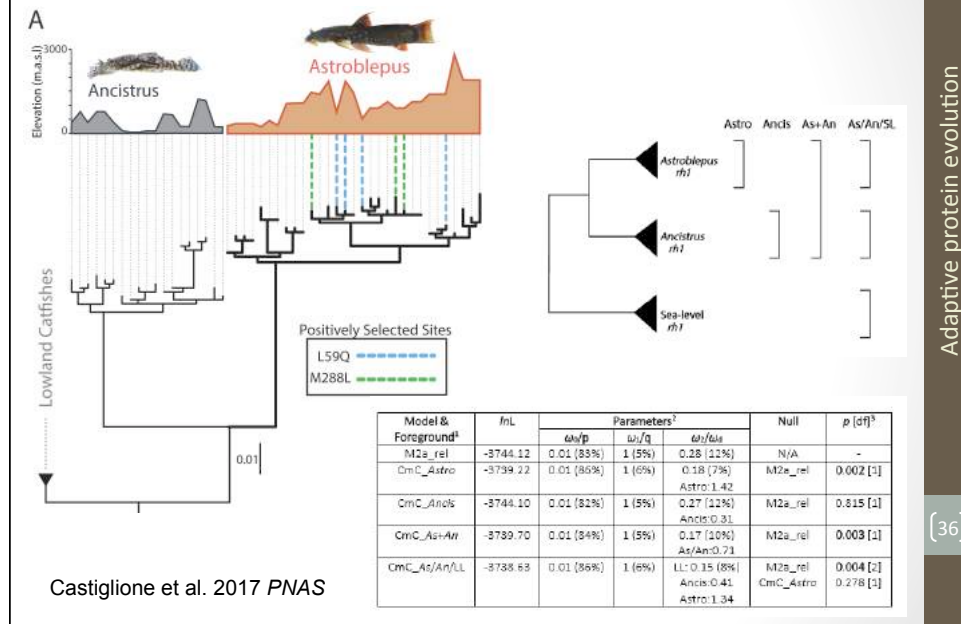
(34)

Positive divergent selection in rhodopsin of high altitude catfishes



Castiglione et al. 2017 PNAS

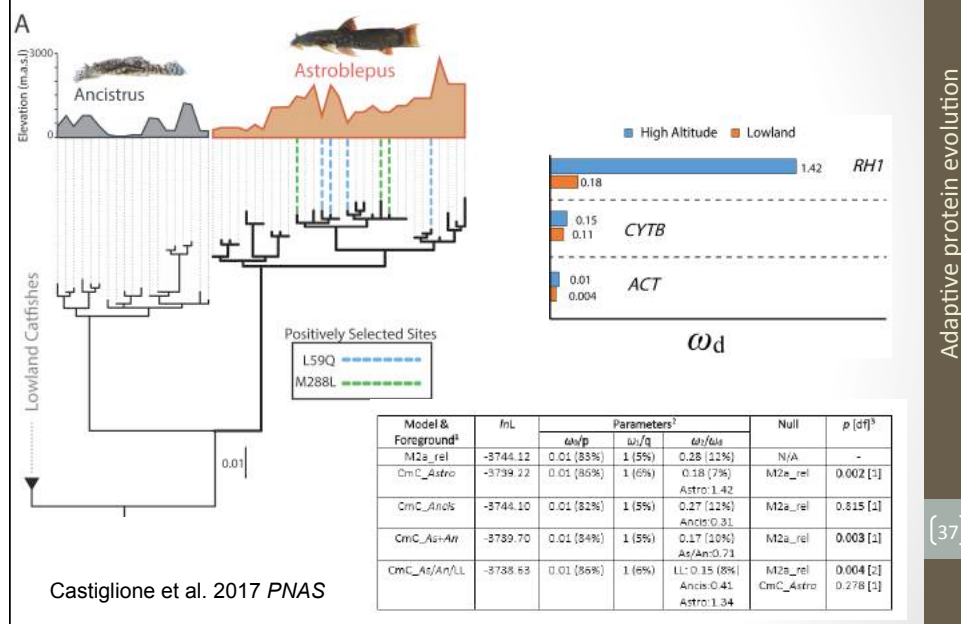
Positive divergent selection in rhodopsin of high altitude catfishes



Adaptive protein evolution

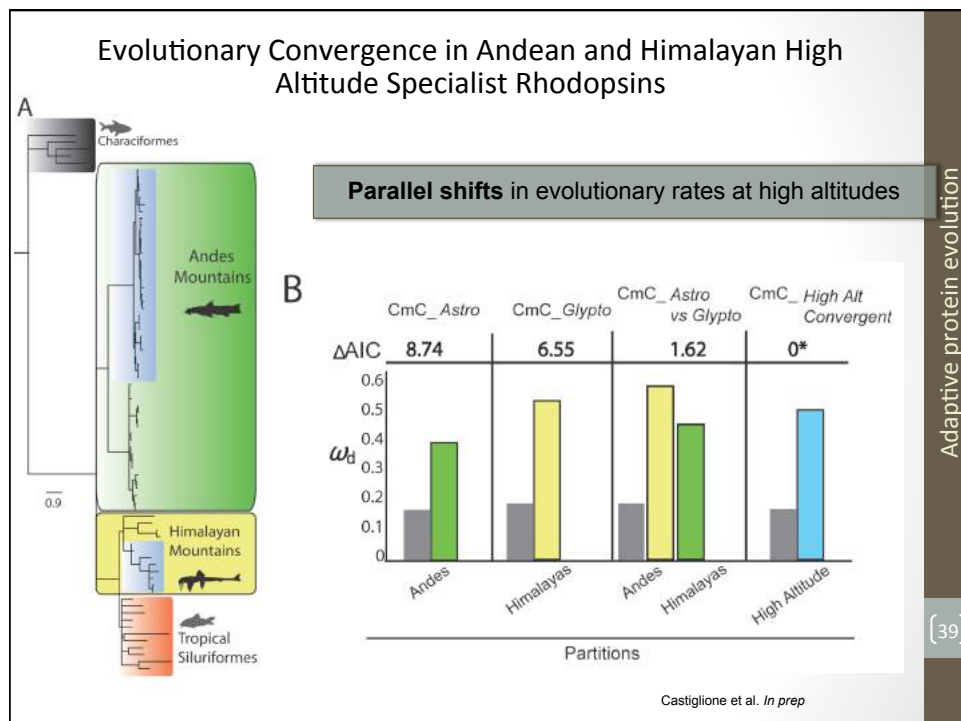
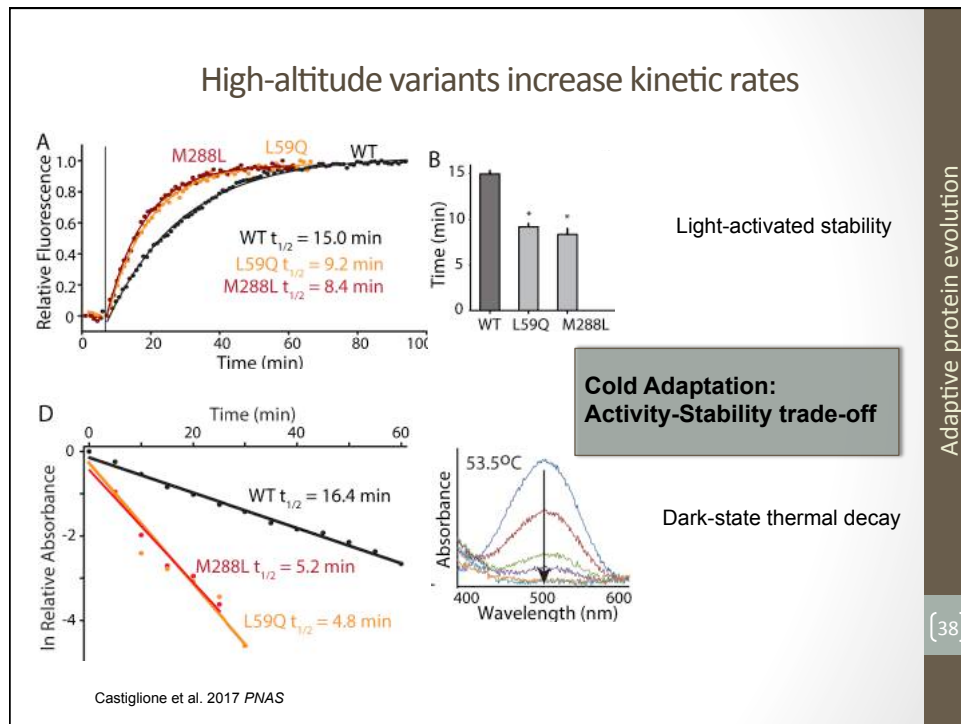
(36)

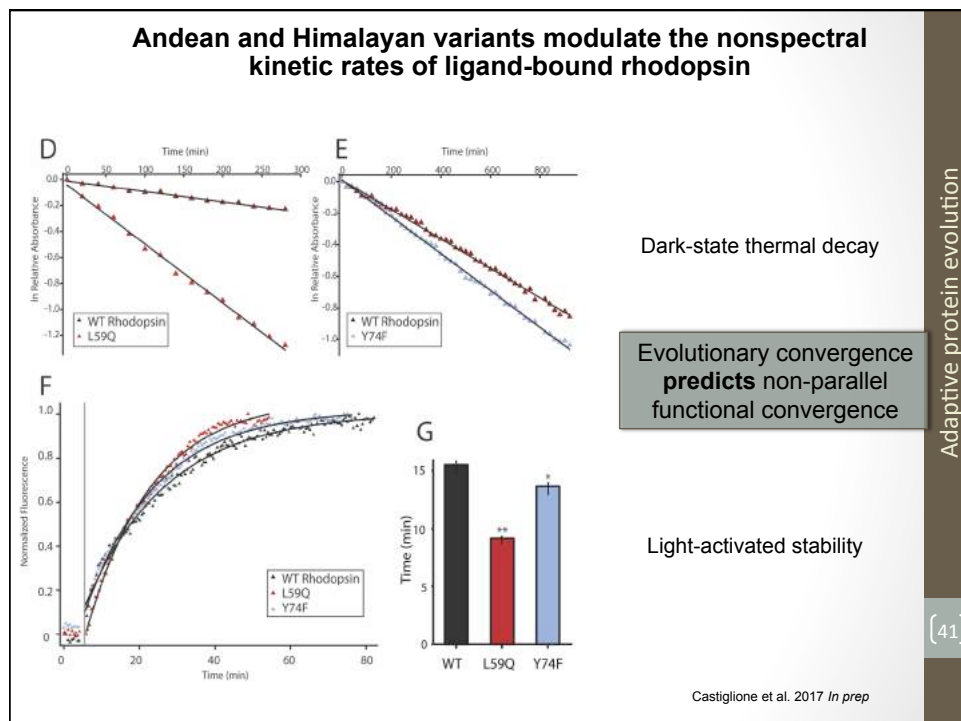
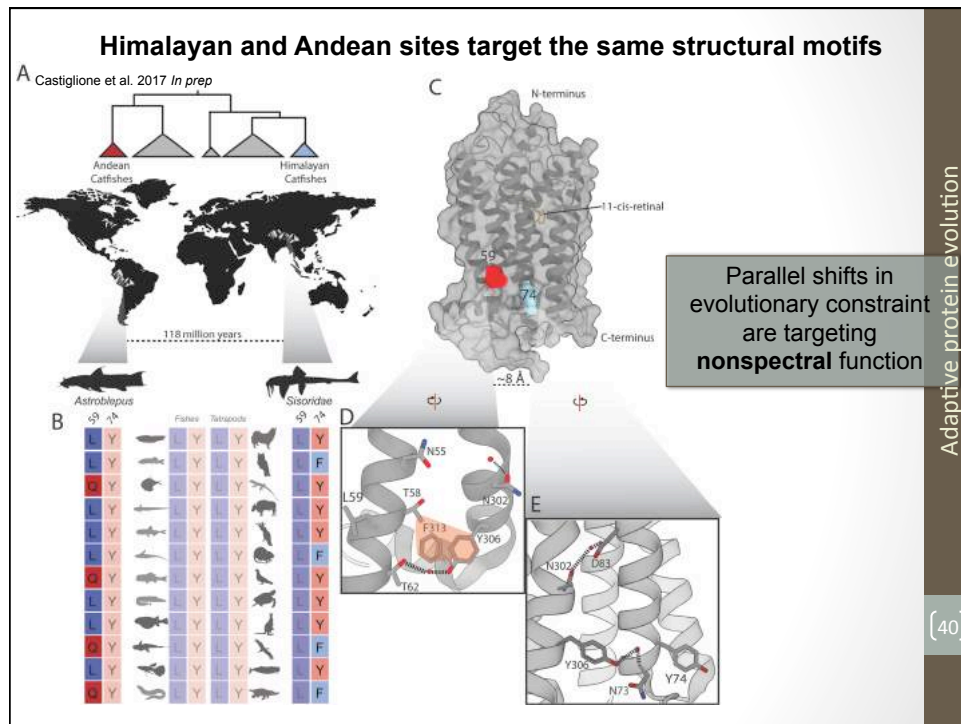
Positive divergent selection in rhodopsin of high altitude catfishes



Adaptive protein evolution

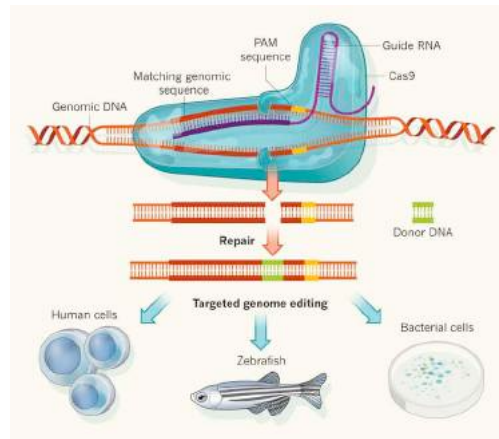
(37)





The future of studies in adaptive molecular evolution?

Combining computational and experimental approaches to study adaptive evolution within a hypothesis testing framework enables us to draw conclusions about adaptive molecular evolution that would not otherwise be possible with either method alone



CRISPR-Cas
system