Evolution of function, beyond similar phylogenetic profiles and only functional change after gene duplication

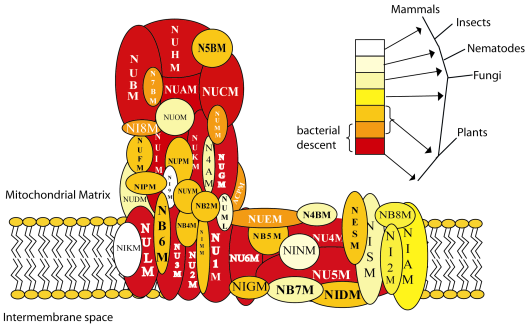
- Exceptions phylogenetic profiles
  - Retention of functionally differentiated paralogs
  - Multi functional proteins
  - Motif-protein co-evolution
  - Anti-correlating proteins
- Evolution of regulation
  - Evolution of Genetic interactions
  - Evolution of (co-)regulation
  - Evolution of phosphorylation & summary evolution of function
- Where do novelty/innovations come from some final thoughts

Explaining discordant phylogenetic profiles of proteins that interact

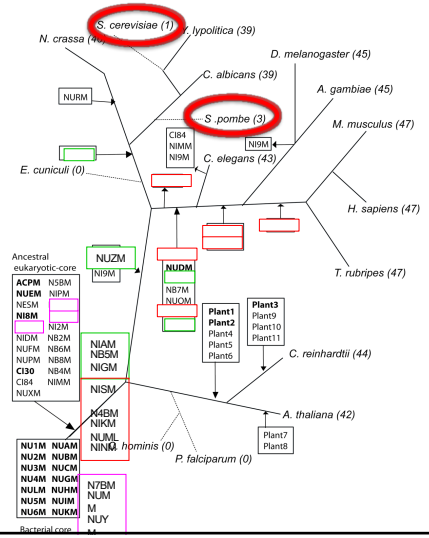
- (we could also just say that evolution is flexible and proteins change function; which I am not going to argue with but (A) conservation of interaction and (B) this is a "just so", non testable explanation)
- "*Happy families are all alike; every unhappy family is unhappy in its own way.*" (from Leo Tolstoy's book Anna Karenina, which begins with this statement)
- Case stories and large scale studies
- And what does it tell us about evolution of function?

discordant phylogenetic profiles because of lineage/group specific duplications (inparalogs) that changed their function

Evolution of Complex I



Gabaldon et al, J. Mol. Biol 2005



Reconstructing Complex I evolution by mapping the variation onto a phylogenetic tree. After an initial "surge" in complexity (from 14 to 35 subunits in early eukaryotic evolution) new subunits have been gradually added and incidentally lost., most other loss is large scale

Huynen et al, BBA 2009

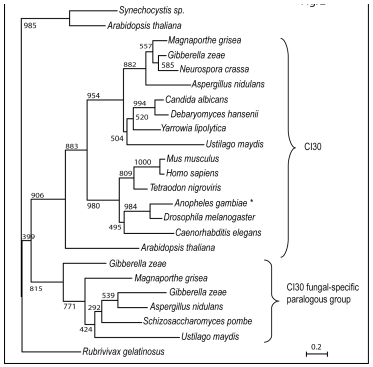
Cardol , BBA 2012

Yip et al, JBC 2011

"Exceptions" in the perfect co-evolution of Complex I

Complex I loss is not always "complete", *S.cerevisiae* and *S.pombe* have retained 1 and 3 proteins

Phylogeny of a "remaining" complex I protein in pombe



?

"The Complex I assembly protein CI30 has been duplicated in the Fungi.
This can explain the presence of a CIA30-homolog in Complex I-less *S.pombe*"

This principle is also recognized for phylogenetic profile function prediction.

Cell Reports
Resource

OPEN ACCESS
CellPress

Systematic Discovery of Human Gene Function and Principles of Modular Organization through Phylogenetic Profiling

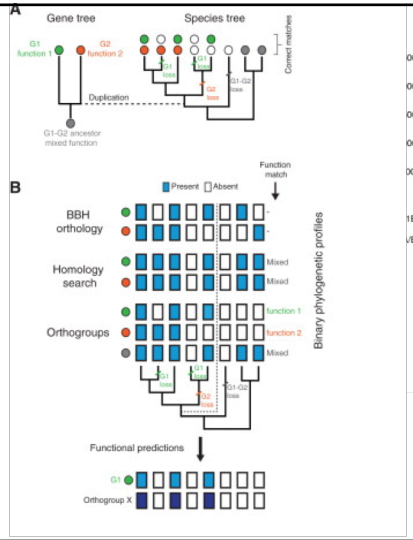
Gautam Dey,[1] Ariel Jaimovich,[1] Sean R. Collins,[1,2] Akiko Seki,[1] and Tobias Meyer[1,*]
[1]Department of Chemical and Systems Biology, Stanford University, Stanford, CA 94305, USA
[2]Present address: Department of Microbiology and Molecular Genetics, University of California, Davis, Davis, CA 95616, USA
*Correspondence: tobias1@stanford.edu
http://dx.doi.org/10.1016/j.celrep.2015.01.025

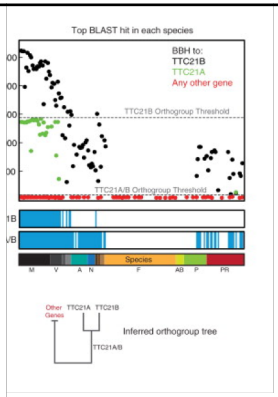
SUMMARY

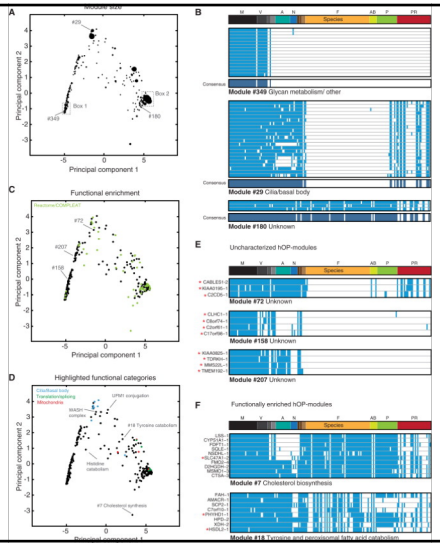
Functional links between genes can be predicted using phylogenetic profiling, by correlating the appearance and loss of homologs in subsets of species. However, effective genome-wide phylogenetic profiling has been hindered by the large fraction of human genes related to each other through historical duplication events. Here, we overcame this challenge by automatically profiling over 30,000 groups of homologous human genes (orthogroups) representing the entire protein-coding genome across 177 eukaryotic

functional screens that are often difficult to develop or cannot be performed for processes that are not well understood.

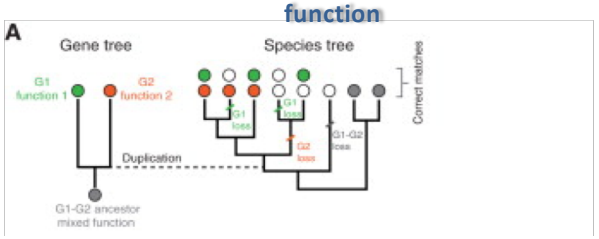
A completely independent approach to predicting gene function was first introduced in bacteria by linking genes based on the joint presence or absence of their orthologs in different species (Pellegrini et al., 1999), defined here as genes with sequence homology derived from a single common ancestor (Gabaldón and Koonin, 2013) (Supplemental Experimental Procedures). This approach, termed phylogenetic profiling, is built on the premise that genes that function together are gained and lost together in evolution. The subsequent extension of phylogenetic profiling to eukaryotic species led to the discovery of cilia genes (Avidor-Reiss et al., 2004), genes linked to Ca^{2+} influx into mito-







what do we learn about evolution of function from discordant phylogenetic profiles bc of lineage specific duplications that changed their function

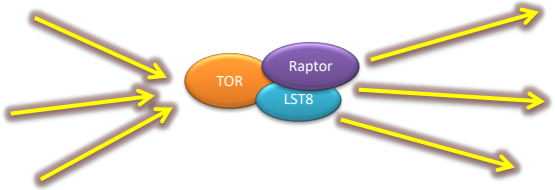


- Change of function after duplication. (=evolution).
- For the original protein. *Evolution by loss.* No change in "function"

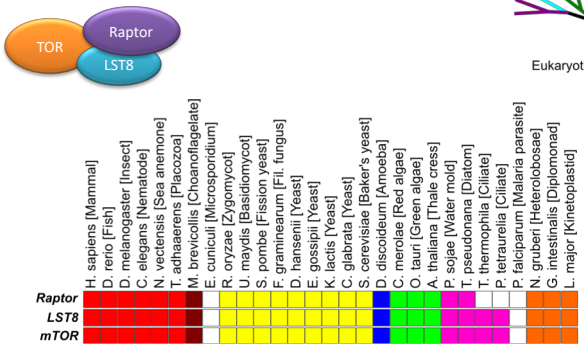
Discordant phylogenetic profiles because of multifunctional proteins

TOR1 complex

- Kinase
- Regulates growth
- Mutations of TOR1 components involved in Cancer

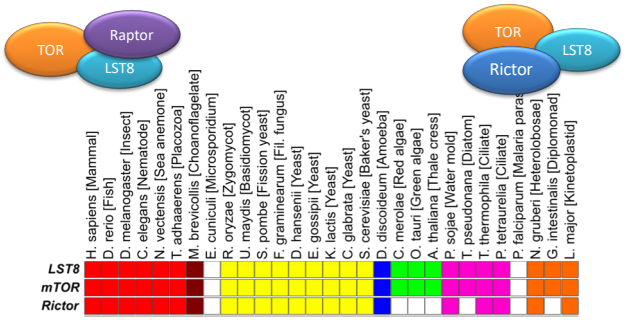


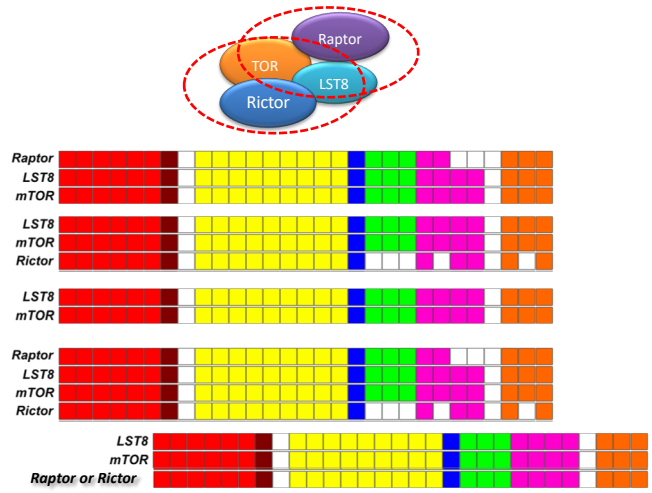
Evolution of TOR



Does evolution of TOR make more sense if we consider the whole network of interactions: TOR2 complex

- TOR2 is involved in rearrangement of cytoskeleton





OPEN ACCESS Freely available online

PLOS | COMPUTATIONAL BIOLOGY

Shared Protein Complex Subunits Contribute to Explaining Disrupted Co-occurrence

Adrian Schneider[1], **Michael F. Seidl**[1,2], **Berend Snel**[1,2]*

1 Theoretical Biology and Bioinformatics, Utrecht University, Utrecht, The Netherlands, 2 Centre for BioSystems Genomics, Wageningen, The Netherlands

Abstract

The gene composition of present-day genomes has been shaped by a complicated evolutionary history, resulting in diverse distributions of genes across genomes. The pattern of presence and absence of a gene in different genomes is called its phylogenetic profile. It has been shown that proteins whose encoding genes have highly similar profiles tend to be functionally related: As these genes were gained and lost together, their encoded proteins can probably only perform their full function if both are present. However, a large proportion of genes encoding interacting proteins do not have matching profiles. In this study, we analysed one possible reason for this, namely that phylogenetic profiles can be affected by multi-functional proteins such as shared subunits of two or more protein complexes. We found that by considering triplets of proteins, of which one protein is multi-functional, a large fraction of disturbed co-occurrence patterns can be explained.

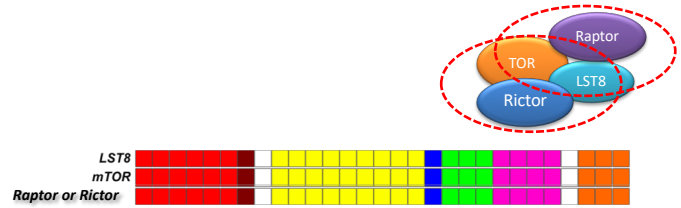
Citation: Schneider A, Seidl MF, Snel B (2013) Shared Protein Complex Subunits Contribute to Explaining Disrupted Co-occurrence. PLoS Comput Biol 9(7): e1003124. doi:10.1371/journal.pcbi.1003124

Editor: Christian von Mering, University of Zurich and Swiss Institute of Bioinformatics, Switzerland

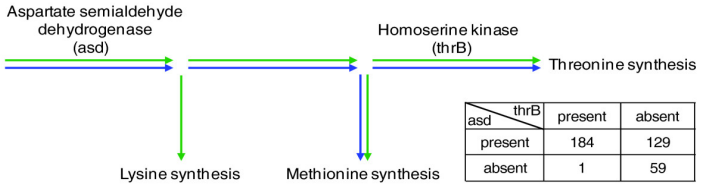
Received September 21, 2012; **Accepted** May 17, 2013; **Published** July 18, 2013

What do discordant functional profiles caused by multifunctional proteins tell us about the evolution of function

- One of the functions was not necessary anymore. That function is part of one protein and another protein, those are lost.
- *Evolution by loss.* No change in "function"

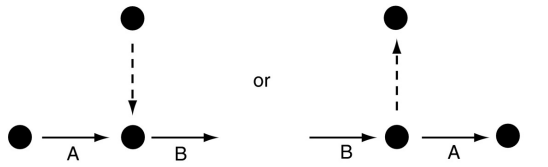


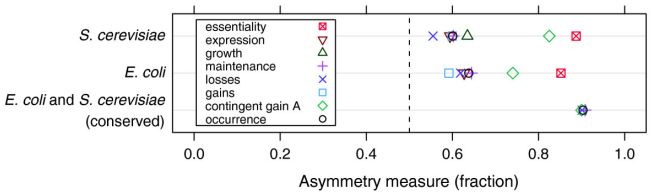
Asymmetric functional/metabolic relations explain discordant phylogenetic profiles



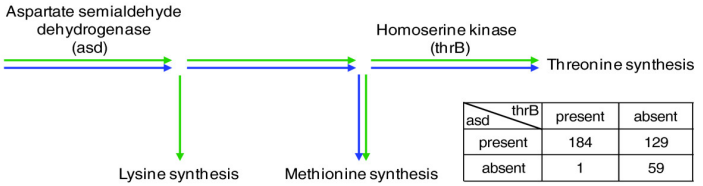
asd \ thrB	present	absent
present	184	129
absent	1	59

Asymmetric relationships between proteins shape genome evolution.
Notebaart RA, Kensche PR, **Huynen MA**, Dutilh BE.
Genome Biol. 2009 Feb 12;10(2):R19.





What do discordant functional profiles caused by assymmetric functional/metabolic proteins tell us about the evolution of function

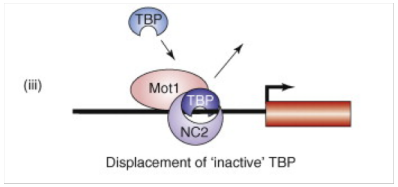


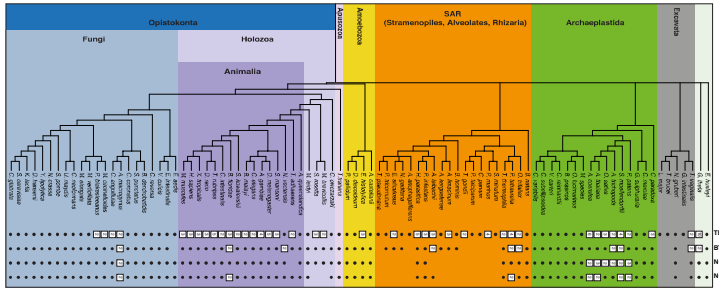
asd \ thrB	present	absent
present	184	129
absent	1	59

- Either functions is not always necessary so loss or (re-)gain through HGT

MOTIF-PROTEIN PHYLOGENETIC PROFILES

Lack of co-evolution (phylogenetic profile similarity) between TBP and MOT1/NC2





Koster, Snel and Timmers *Cell* 2015

Organisms without MOT1/NC2 tend to lose one of the critical phenylalanines, this explains how they cope AND reveals co-evolution between presence / absence of a gene and residues in another gene

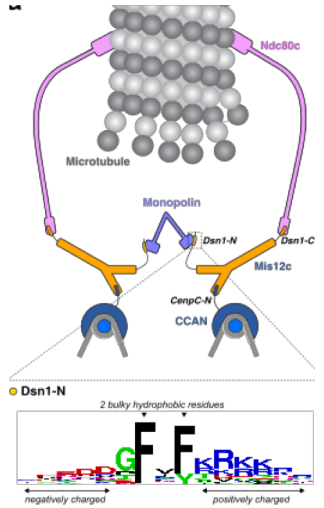
	(184-199)	(207-221)	(278-290)	(294-309)
A. thaliana	RNAEYNPK-R-FAAVIMR	ALVFASGKMVCTGAK	SYEPELFPGL--IYR	PKIVLLIFVSGKIVIT
B. dendrobatidis	RNAEYNPK-R-FAAVIMR	ALIFASGKMVVTGAK	SYEPELFPGL--IYR	PKIVLLIFVSGKIVLT
C. elegans	RNAEYNPK-R-FAAVIMR	ALIFSSGKMVCTGAK	TYEPELFPGL--IYR	PRVVLLIFVSGKVVIT
C. reinhardtii	RNAEYNPK-R-FAAVIMR	ALIFASGKMVCTGAK	SYEPELFPGL--IYR	PKIVLLIFVSGKVVLT
D. melanogaster	RNAEYNPK-R-FAAVIMR	ALIFSSGKMVCTGAK	SYEPELFPGL--IYR	PRIVLLIFVSGKVVLT
D. rerio	RNAEYNPK-R-FAAVIMR	ALIFSSGKMVCTGAK	SYEPELFPGL--IYR	PRIVLLIFVSGKVVLT
H. sapiens	KNAEYNPK-R-FAAVIMR	ALIFSSGKMVCTGAK	SYEPELFPGL--IYR	PRIVLLIFVSGKVVLT
M. leidyi	RNAEYNPK-R-FAAVIMR	ALLFNSGKMVCTGAK	HYEPELFPGL--IYR	PKIVLLIFVSGKVVLT
P. infestans	RNAEYNPK-R-FSAVIMR	ALIFGSGKIVIIGGT	SYEPELFPGL--IYK	PKLILLIFVSGKIVLC
P. patens	RNAEYNPK-R-FAAVIMR	ALIFASGKMVCTGAK	SYEPELFPGL--IYR	PKIVLLIFVSGKIVLT
S. cerevisiae	RNAEYNPK-R-FAAVIMR	ALIFASGKMVVTGAK	SYEPELFPGL--IYR	PKIVLLIFVSGKIVLT

	(184-199)	(207-221)	(278-290)	(294-309)
A. anophagefferens	RNAEYNPK-K-FAAVIMR	ALVFSSGKMVITGAK	SYEPELFPGL--IYR	PKTVLLLFVSGRMVLT
C. parvum	RNAEYNPK-K-YAAVVMR	GLLFRSGRLMITGAR	NYEPELFPGL--VYR	TKAVLLLFVSGKVIVT
G. intestinalis	LTADYNP--R-FPTVRVR	ISVGHGHCTIFGCE	MYQPEIMPSLQVVFK	RNICCSVFADGQVTIV
L. major	RNVEYMPNNR-LFSATMR	VMDGNSGALSIIGAA	SYEPERFNGC--VLR	WSVSCSVFVIGKVQLM
P. falciparum	RNAEYNPS-K-LNTLIIR	ALIFKNGRIMLTGTR	NYEPELFAGL--VYR	LKSVILIFVSGKIIIT
P. marinus	RNAEYDPS-K-FPCVRIR	IAVFRSGKIQATGAA	AYEPSRIPAV--VLR	RGVIVDVFSTGRVSMK
S. minutum	RNAEYNPQ-KGVVRMLLR	ATLFASGNVHCC-VK	LIVPDVGCAA--SLF	PPCSMQIDASGKLTAV
T. brucei	RNVEFTPRVR-GFAATLR	VGNFDSGLLGIIGST	SYEPDRFSGC--IVR	WQVCCIVFVIGKVIVL
T. gondii	GNSVYNPE-E-FHSVRVD	IVLFSNGKMMGTGAN	DYEPERFPGA--RVK	P-VTLQLFSTGNVTLT
T. pseudonana	RNTIEFNPR-R-FAAVIMR	ALLFSSGRMVITGTK	SYEPELFPGL--IYR	PRVVLLIFVSGKVVIT

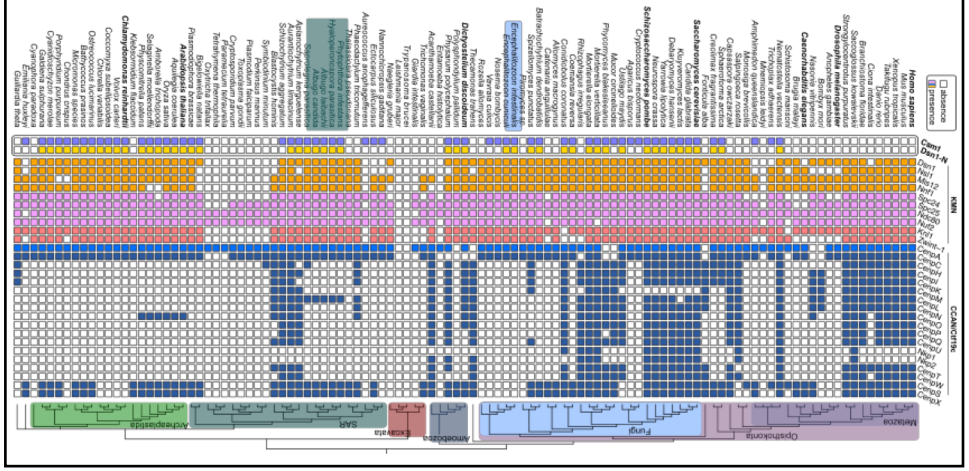
ref

Koster, Snel and Timmers *Cell* 2015

Csm1 is a LECA kinetochore subunit of the Monopolin complex lost in higher animals that interacts with Dsn1



The phylogenetic profiles of the motif Dsn1-N and Csm1 are highly similar



Disruption of phylogenetic profile similarity; what have we learned about function?

- The interaction/function is ancestral
- Orthologs differentiate in function by loss of interaction and the function associated with this interaction (*cf.* multifunctional proteins)
- Potentially useful tool to predict interaction motifs

Non-orthologous gene displacement/analogous proteins explain discordant phylogenetic profiles

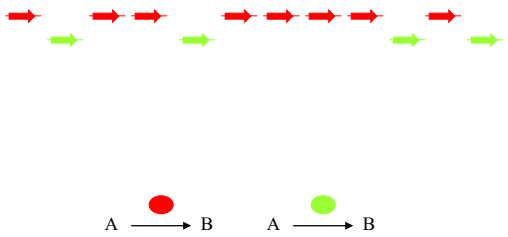
- First systematic analysis on *M.genitalium* (Koonin et al., Trends Genet. 1997)

TABLE 1. Non-orthologous genes coding for the same function in *Mycoplasma genitalium* and *Haemophilus influenzae*

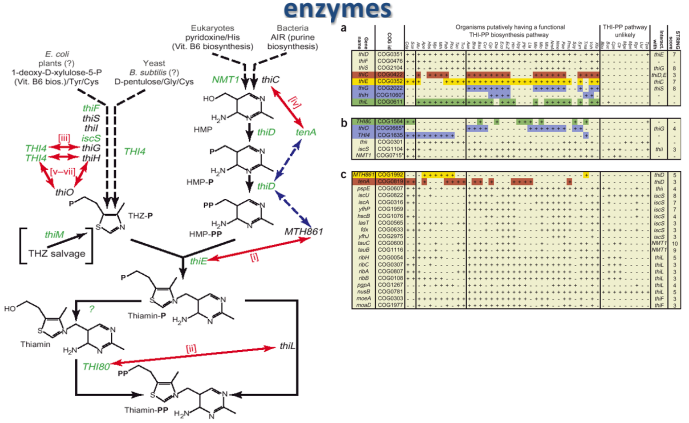
	M. genitalium		*H. influenzae*		
Enzyme	**Gene[a]**	**Orthologs[b]**	**Gene[a]**	**Orthologs[b]**	**Comment[c]**
No sequence similarity between *M. genitalium* and *H. influenzae* proteins					
Phosphoglycerate mutase	MG430 (*ybO*)	PMGI_BACSU PMGI_ECOLI PMGI_MAIZE	HI0757 (*gpmA*)	PMG1_ECOLI PMGM_HUMAN not in G(+)	*Escherichia coli* encodes both types of enzymes
L-lactate dehydrogenase	MG460	LDH_BACSU LDHM_HUMAN	HI1739B (*lctD* or *lldD*)	LLDD_ECOLI G(+)	The HI enzyme is distantly related to eukaryotic cytochrome B2
Lipoate-protein ligase	MG270	LPLA_ECOLI SCYJL046W_1	HI0027 (*lipB*)	LIPB_ECOLI S51458 (yeast)	*E. coli* and yeast encode both types of enzymes
Nucleoside diphosphate kinase	MG264[d] MG268[d]	None	HI0876 (*ndk*)	NDK_ECOLI NDKB_HUMAN	The two predicted kinases in MG are candidates for this indispensable activity
DNA polymerase, repair	MG261 (*dnaE*)	DP3A_HAEIN DP3A_ECOLI	HI0856 (*polA*)	DPO1_ECOLI DPO1_MYCTU	MG encodes two homologs of DNA polymerase III. MG261 is the likely repair polymerase as it belongs to a putative repair operon[6]
RNase H	MG262[d]	DPO1_BACCA DPO1_HAEIN	HI0138 (*rnhA*); HI1059 (*rnhB*)	RNH_ECOLI RNH1_YEAST RNH2_ECOLI MG326_1 (*M. capricol.*) SC23CDS_13 (yeast)	MG262 is homologous to the 5′–3′ exonuclease domain of DNA polymerase I. It is predicted to replace the two unrelated RNases H of HI in primer removal during DNA replication
Glycyl-tRNA synthetase	MG251	SYG_HUMAN	HI0927 (*glyQ*) HI0924 (*glyS*)	SYGA_ECOLI SYGB_ECOLI CTU20547_1 (Chlamydia) G(–)	The MG enzyme contains one subunit, the HI counterpart two
Paralogs in *M. genitalium* and *H. influenzae*					
Prolyl-tRNA synthetase	MG283	YHI0_YEAST	HI0729 (*proS*)	SYP_ECOLI YER7_YEAST	Yeast encodes both types of enzymes
Cytidine deaminase	MG052	CDD_BACSU CDD_HUMAN	HI1350 (*cdd*)	CDD_ECOLI	The MG cytidine deaminase is more closely related to eukaryotic enzymes than to those from G(+) bacteria

The opposite of co-occurrence: anti-correlation / complementary patterns: predicting analogous enzymes

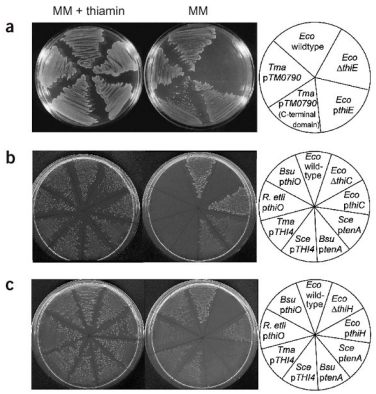
Genes with complementary phylogenetic profiles could have a similar biochemical function.



Complementary patterns in thiamin biosynthesis predict analogous enzymes

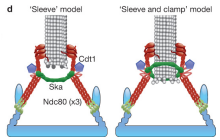


Prediction of analogous enzymes is confirmed

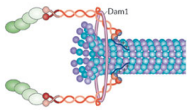


Ska & Dam1: functional counterparts

- KT-MT attachments
  - Dependent on Ndc80
    - Interaction with loop?
  - Tracking of depolymerizing microtubules

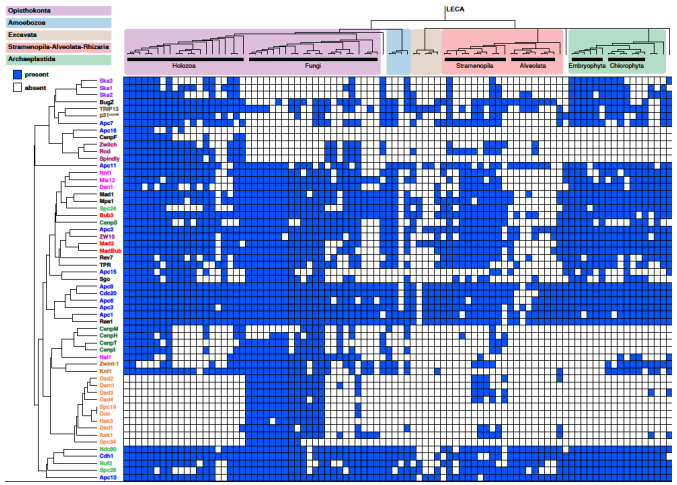


Matson & Stukenberg (2012)

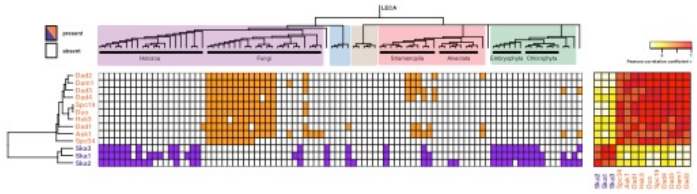


Lampert & Westermann (2011)

→ Orthologs of Ska (3 subunits) and Dam1 (10 subunits) across 94 genomes

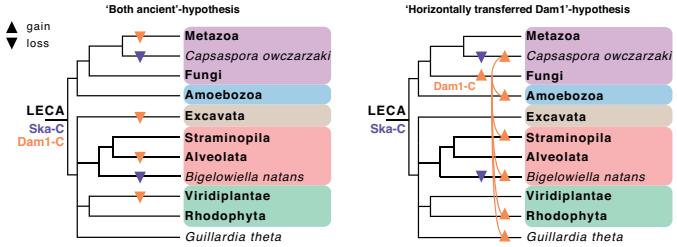


Ska & Dam1 across eukaryotes: intracomplex correlation and intercomplex anticorrelation

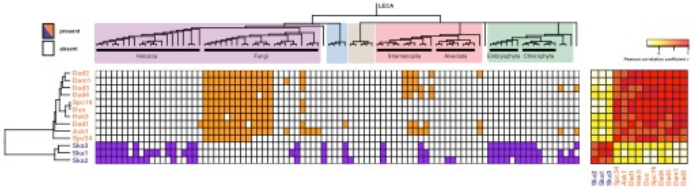


- Ska complex subunits in i.e. Metazoa, Chytridiomycota, Apusozoa, Archaeplastids and some SAR.
- Dam1 complex subunits in most fungal lineages, Filasteria, Amoebozoa, various Stramenopila, Rhizaria, red algae, Cryptophyta.

Alternative evolutionary scenarios



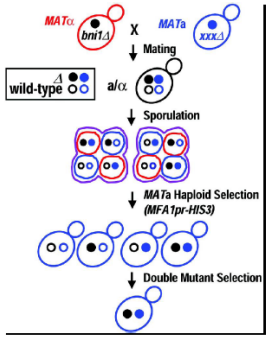
what do we learn about evolution of function from analogous enzymes

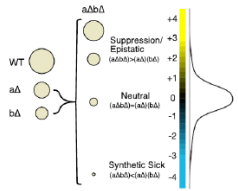


- The function is "conserved", there is no evolution of function (for the network / organisms) (???)
- But there is evolution of protein/gene with similar functionality (and where does the analogous protein come from?) (but also perhaps a lot of *evolution by loss*)
- And why?

EVOLUTION OF GENETIC INTERACTIONS

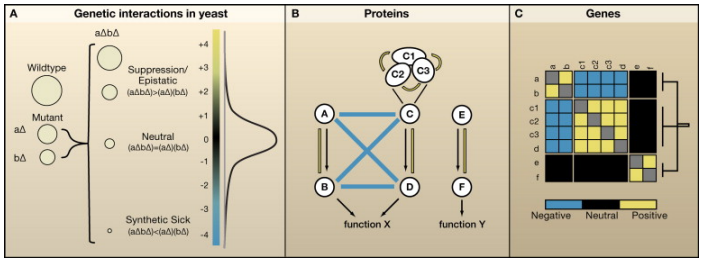
Genetic interactions





Tong et al. 2001 Science
Schuldiner et al. 2005 Cell

45



Negative / syntetic lethal / aggravating
Positive / buffering / alleviating

Cell

Functional Repurposing Revealed by Comparing *S. pombe* and *S. cerevisiae* Genetic Interactions

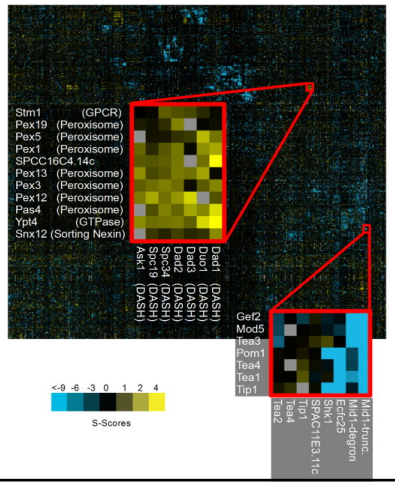
Adam Frost,[1,*] Marc G. Elgort,[1] Onn Brandman,[2,3,4] Clinton Ives,[2,3,4] Sean R. Collins,[7] Lakshmi Miller-Vedam,[2,3,4] Jimena Weibezahn,[2,3,4] Marco Y. Hein,[5] Ina Poser,[6] Matthias Mann,[5] Anthony A. Hyman,[6] and Jonathan S. Weissman[2,3,4]

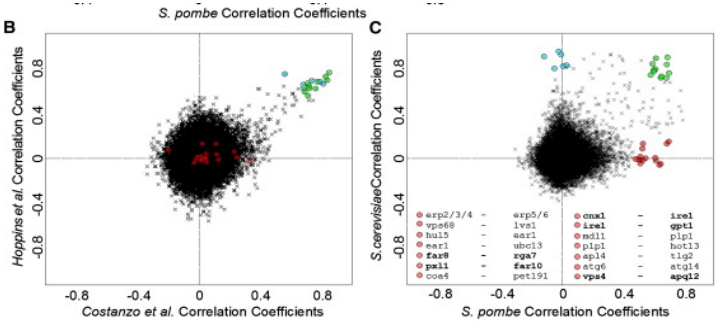
"generate 774,309 double mutants "

But …

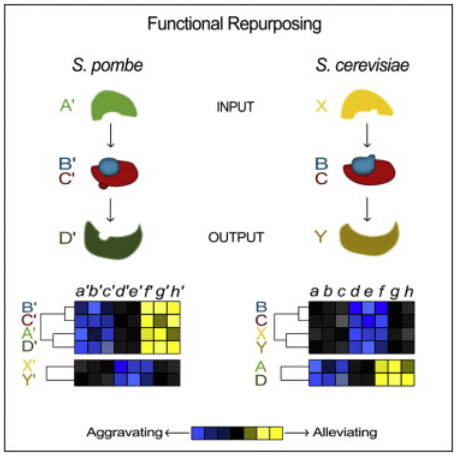
"Our *Sp* map identified > 700 high-confidence gene-to-gene correlations indicative of genes with related functions"

Genetic interaction correlations





We present a genetic interaction map of pairwise measures including ~40% of nonessential *S. pombe* genes. By comparing interaction maps for fission and budding yeast, we confirmed widespread conservation of genetic relationships **within** and **between** complexes and pathways. i.e. *the data is of high enough quality to reliably (consistently) presence or absence of "function"*

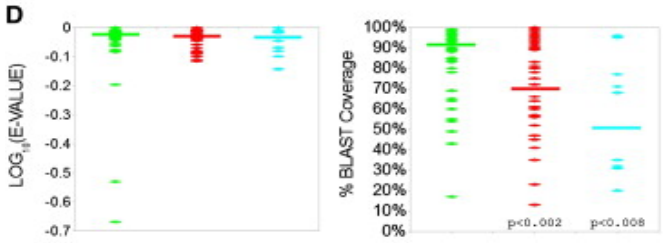


an important subset of orthologous complexes that have undergone functional "repurposing": the evolution of divergent functions and partnerships

Example ESCRT

- the *e*ndosomal *s*orting *c*omplex *r*equired for *t*ransport (ESCRT) genes in endosomal maturation
- Also a role in cytokinesis in pombe (and animals) but not in cerevisiae
- Extensive experimental validation
- ? Loss of function in yeast
- **Different behavior for intra complex vs inter-complex interactions in evolution:** *within module/complex interactions are conserved but regulation and role of module for the cell evolves*

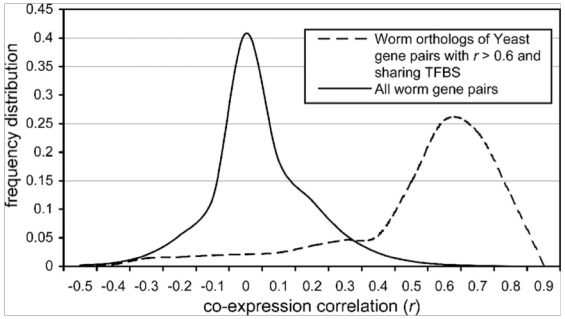
Change in function between orthologs does not seem to depend on sequence identity but does seem to depend on sequence domain/motif composition



Lower amino acid similarity did not correlate with repurposing (Figure 2D, left), but lower percentage coverage (i.e., additional motifs or domains present in only one of the orthologs) did correlate with apparent repurposing

EVOLUTION OF (CO-)REGULATION

"Co-regulation" is quite well conserved (if the genes are conserved) -> co-regulation indicates "same complex" "close together in a pathway"

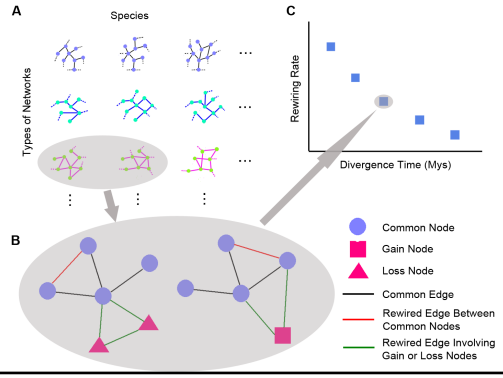


https://academic.oup.com/nar/article/32/16/4725/1023281

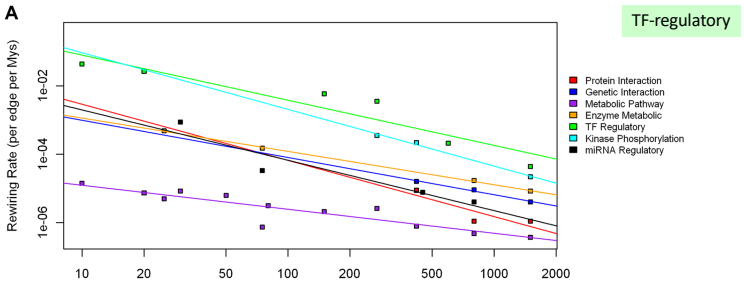
OPEN ACCESS Freely available online PLoS COMPUTATIONAL BIOLOGY

Measuring the Evolutionary Rewiring of Biological Networks

Chong Shou[1], Nitin Bhardwaj[2], Hugo Y. K. Lam[1], Koon-Kiu Yan[2], Philip M. Kim[3], Michael Snyder[4], Mark B. Gerstein[1,2,5]*



Conserrvation of TF-target relations?



TF-regulatory

3.
Measuring the evolutionary rewiring of biological networks.
Shou C, Bhardwaj N, Lam HY, Yan KK, Kim PM, Snyder M, **Gerstein** MB.
PLoS Comput Biol. 2011 Jan 6;7(1):e1001050.

Table S3.

Network Type TF	*S. cerevisiae, S. bayanus*	*D. melanogaster, S. cerevisiae*
Edge change from Edge Gain	26	80
Edge change from Edge Loss	53	80
Edge change from Node Gain	60	12733
Edge change from Node Loss	306	76543

!

Regulatory evolution. Dynamic conservation?

Research **Open Access**

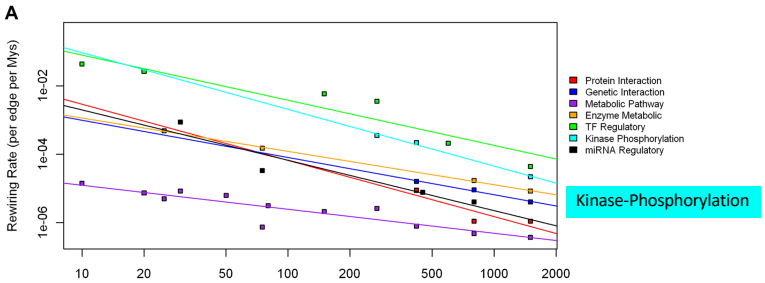
Genome adaptation to chemical stress: clues from comparative transcriptomics in *Saccharomyces cerevisiae* and *Candida glabrata*

Gaëlle Lelandais*†, Véronique Tanty‡, Colette Geneix*§, Catherine Etchebest*, Claude Jacq†‡ and Frédéric Devaux†

- "We found that although the gene expression patterns characterizing the response to drugs were remarkably conserved between the two species, part of the underlying regulatory networks differed."

EVOLUTION OF PHOSPHORYLATION

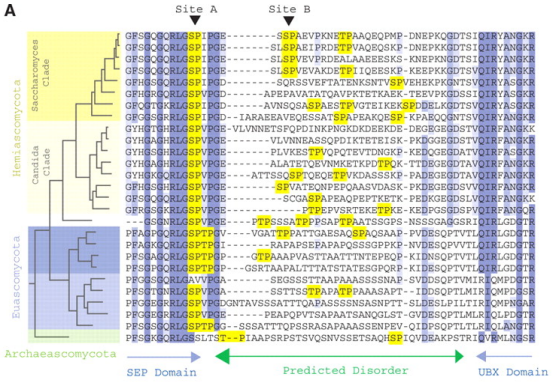
Evolution of kinase-phosphorylation targets



Kinase-Phosphorylation

3.
Measuring the evolutionary rewiring of biological networks.
Shou C, Bhardwaj N, Lam HY, Yan KK, Kim PM, Snyder M, **Gerstein** MB.
PLoS Comput Biol. 2011 Jan 6;7(1):e1001050.

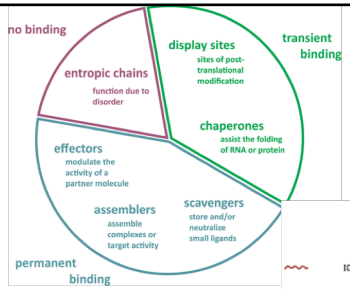
Evolution of phosphorylation



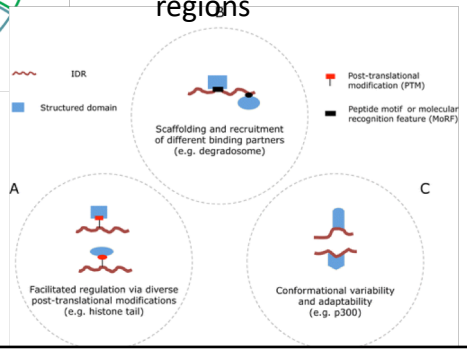
Science
AAAS

Global Analysis of Cdk1 Substrate Phosphorylation Sites Provides Insights into Evolution
Liam J. Holt *et al.*
Science **325**, 1682 (2009);
DOI: 10.1126/science.1172867

- position of most phosphorylation sites is not conserved in evolution; instead, clusters of sites shift position in rapidly evolving disordered regions.
- the regulation of protein function by phosphorylation often depends on simple nonspecific mechanisms that disrupt or enhance protein-protein interactions.
- Is similar to?

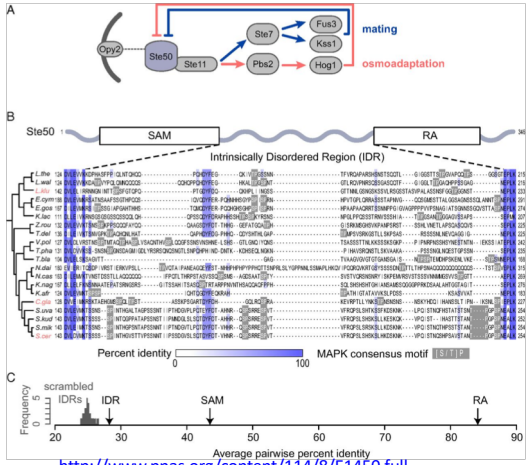


Functions of non-globular / disordered / unstructured regions



So how do they evolve? How should we think about that?

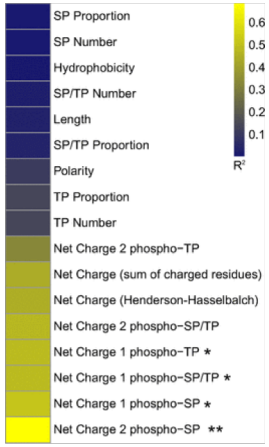
example



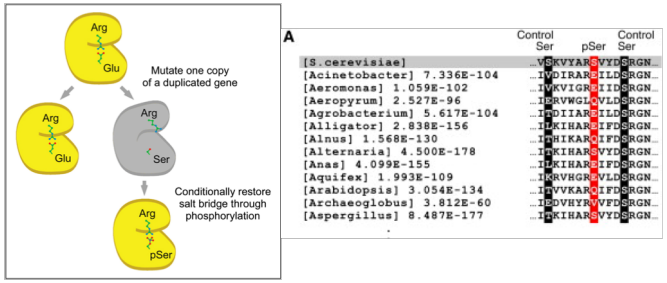
http://www.pnas.org/content/114/8/E1450.full

Diverged orthologous IDRs recapitulate *S. cerevisiae* IDR functions compared with the 5A mutant.

&

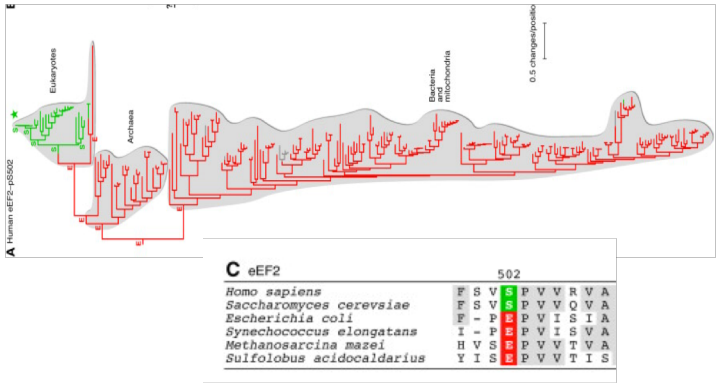


Other class of phosphosites



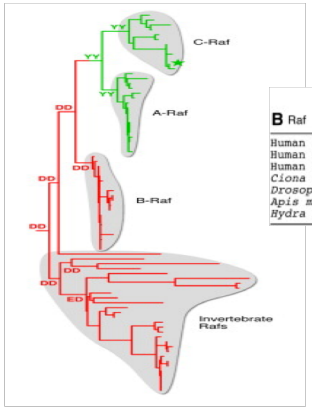
https://www.sciencedirect.com/science/article/pii/S0092867411012141

Other class of phosphosites: eEF2



https://www.sciencedirect.com/science/article/pii/S0092867411012141

Other class of phosphosites: Raf



B Raf

	340 341											
Human A-Raf	D	S	G	Y	Y	W	E	V	P	P	S	E	V
Human C-Raf	D	S	S	Y	Y	W	E	I	E	A	S	E	V
Human B-Raf	D	S	S	D	D	W	E	I	P	D	G	Q	I
Ciona intestinalis Raf	D	S	N	E	D	W	E	V	P	E	E	E	I
Drosophila melanogaster Raf	S	S	E	E	N	W	N	I	L	A	E	E	I
Apis mellifera Raf	E	S	I	E	D	W	E	I	P	A	D	E	I
Hydra magnipapillata Raf	D	S	N	D	D	W	E	I	K	E	G	E	I

https://www.sciencedirect.com/science/article/pii/S0092867411012141

"dynamic conservation" "neutral-rewiring & conserved output/function"

- Function / output is conserved but exact wiring / positions is not
- Also implied to play a large role in evolution of transcription factor binding sites.
- i.e. in normal (globular) protein sequence evolution conservation of function implies conservation of sequence/structure, neutrality means similar amino acids (or synonymous substitutions) but for other units of function it could be higher level (conservation of charge and length, conservation of co-expression*) and dynamics at lower level

* When and why (role) a protein/module/complex does its thing will evolve a lot more than module-membership and module molecular activity

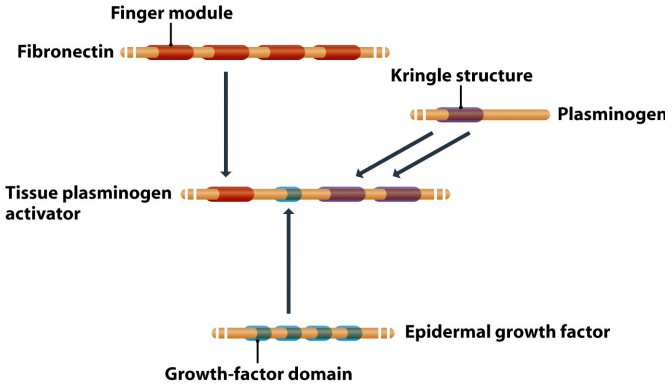
Evolution of function: grand summary

- Strong interplay between network and genome evolution
  - Within pathways/complexes (**modules**) evolve by loss and gain of genes (from the genome!) but little rewiring (as in loss or gain of co-expression/interaction)
    - Most differences in networks are due to gain and loss of genes from the genome!
  - Also gain (and "loss") of module membership after duplication followed by rapid functional substitutions
- Regulatory relations "dynamic conservation"
  - At "shorter" evolutionary distances, change in wiring, but same output ("function")
  - At longer distances repurposing of when / how modules are needed
    - Between module relations are less conserved than within module
  - (also "applies" to intrinsically disorderd proteins, and a subset of phosphosites)

So where does "new stuff" come from (besides duplication)

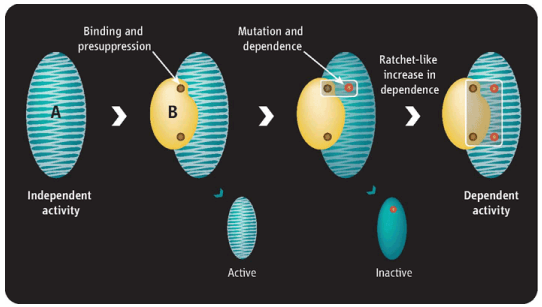
- Duplication / invention of new genes, & domain-recombination
- Inflation-contraction / biphasic model of genome evolution: e.g. eukaryogenesis, origin of animals, origin of vertebrates (mix of duplication, innovation, vertical inheritance)
- Constructive neutral evolution
- Function evolution is often episodic: rapid emergence of new functions, long periods of conservative evolution
- Exception: Arms-race processes (genetic conflict, host-pathogen) adaptive evolution is much more frequent

"new proteins" from duplication & domain recombination



92

Accumulation of complexity: a neutral explanation



fixation of neutral or slightly deleterious features as a general and unavoidable source of complexity in taxa with small populations

Science. 2010 Nov 12;330(6006):920-1.
Cell biology. Irremediable complexity?
Gray MW, Lukes J, Archibald JM, Keeling PJ, Doolittle WF.

How to falsify?

e.g. Neurospora mito-TyrRS

- *Neurospora* mitochondrial genome encodes several introns which require a tyrosyl tRNA synthetase (TyrRS) to splice.
- "to compensate for structural defects acquired by the intron sequences "
- BUT Introns with defects arising -> negative selection
- ? Reverse: first binding (fortuitously or for reason unrelated to splicing)—> accumulation of mutations in the intron that inactivate splicing, if TyrRS not bound.
- Because the compensatory / suppressive activity exists before mutation "presuppression,"
- the protein dependence by the intron could be selectively neutral (or slightly disadvantageous)

"Constructive neutral evolution"

- Suggested that many taxon specific subunits (taxon specifc proteins that are a subunit in a complex) are regulatory subunits
- Hypothesis: neutrally added but necessary subunits could have been appropriated as regulatory subunits or "assembly" factors?
- "*Finally, and to me most interestingly, how can we combine multi-level selection theory with reasoning about introns as adaptations (Doolittle, 1987, Cold Spr Hbr Symp Quant Biol 52: 907–913)? It may well be that multicellular eukaryotes of a certain type (us, for instance) have gained considerable evolvability (and consequent diversity) from having alternatively spliceable introns. But clearly, introns were not added to the genome of LECA so that more than a billion years later this advantage could be realized. Authors are (although too circumspectly in my opinion) down on such teleological rationalizing, but might we imagine such evolvability to be an adaptation at some much higher level (clades above species, Doolittle 2017; Phil Sci 84: 275–295)?*"