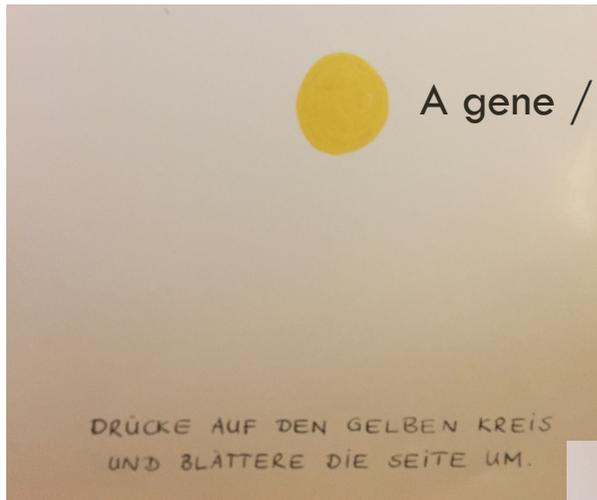


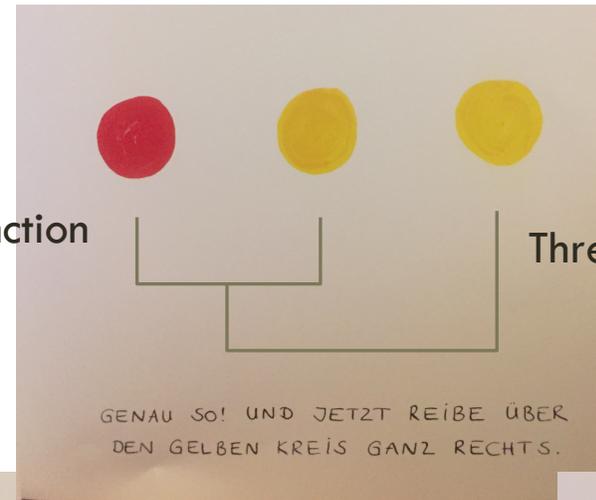
WHAT COMPARATIVE GENOMICS TELL US ABOUT THE EMERGENCE OF HUMAN PATHOGENS

Ingo Ebersberger
Applied Bioinformatics Group
Inst of Cell Biology and Neuroscience
GU Frankfurt

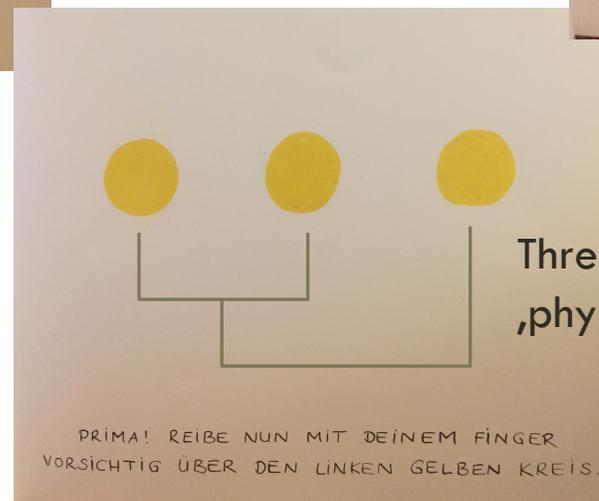
DOTS...



Change in function

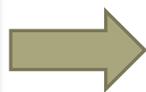


A taxon-gene matrix

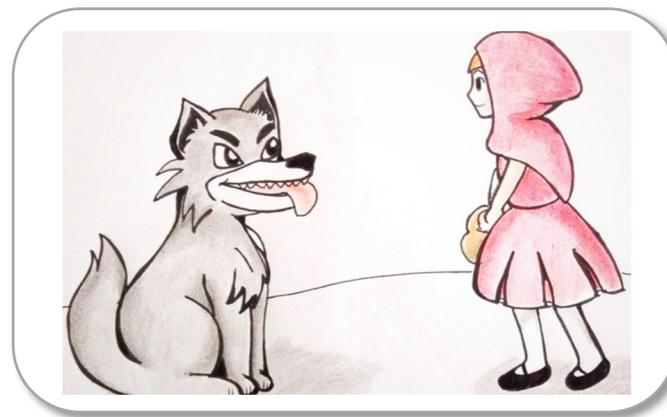


PART 1 – WORKING WITH GENOMES IS LIKE WORKING WITH TEXT

Volt egyszer egy kedves, aranyos kislány; aki csak ismerte, mindenki kedvelte, de legjobban mégis a nagymamája szerette: a világ minden kincsét neki adta volna. Egyszer vett neki egy piros bársonysapkát. A kislánynak annyira tetszett a sapka, hogy mindig csak ezt hordta; el is nevezték róla Piroskának. Piroskáék bent laktak a faluban, nagymama pedig kint az erdőben, egy takaros kis házban. Egy szép napon azt mondja Piroskának az édesanyja: – Gyere csak, kislányom! Itt van egy kalács meg egy üveg bor, vidd el a nagymamának. Beteg is, gyöngye is szegényke, jól fog esni neki. Indulj szaporán, mielőtt beáll a hőség. Aztán szépen, rendesen menj, ne szaladgálj le az útról, mert elesel, és összetörök az üveg, kifolyik a bor, és akkor mit iszik a nagymama! Ha pedig odaérsz, ne báméskodj összevissza a szobában; az legyen az első dolgod, hogy illedelmesen jó reggelt kívánj. – Bízzad csak rám, édesanyám, minden úgy lesz, ahogy mondd – felelte Piroska az intelemre, azzal karjára vette a kosárkát, és útnak indult. Átvágott a mezőn, beért az erdőbe; hát ki jön szembe vele? Nem más, mint a farkas. – Jó napot, Piroska! – köszönt rá a kislányra. Az meg mosolyogva, jó szívvel felelte: – Neked is, kedves farkas! – Nem tudta még, milyen alattomos, gonosz állattal van dolga. – Hová ilyen korán, lelkecském? – szívélyeskedett tovább a farkas. – Nagymamához. – Aztán mit viszel a kosaradban? – Bort meg kalácsot. Tegnap sütöttük; szegény jó nagymama gyöngye is, beteg is, jól fog tenni neki, legalább egy kicsit erőre kap tőle. – És hol lakik a nagymama, Piroska? – Itt az erdőben, a három tölgyfa alatt. Biztosan ismered a házat, mogyorósövény van körülötte.



There was once a lovely, cute little girl; whoever knew her was loved by all, but still loved by her grandmother best: she would have given her all the treasures of the world. He once bought her a red velvet cap. The little girl liked the hat so much that she always wore it only; it was also named Piroška. The Redheads lived inside the village, and Grandma was out in the woods in a neat little house. One fine day Piroška's mother says, "Come on, baby girl!" Here's a cake and a bottle of wine, take it to your grandmother. He is both sick and weak, he will do well. Start fast before the heat sets in. Then go nicely, properly, don't run off the road because you will fall and break the bottle, the wine will flow out, and then what will your grandmother drink! And when you get there, don't stare at the room; your first thing to do is to have a decently good morning. "Trust me, my mother, everything will be as you say," Piroška replied to the rebuke, took the basket in her arms and set off. He cut through the field, entered the woods; so who will face him? He is nothing but a wolf. - Good afternoon, Little Red Riding Hood! ...



WORKING WITH GENOMES IS LIKE WORKING WITH TEXT

```
>HiC_scaffold_1
TTTAGCTCACTTGCTTATTTTGTACATTCAGGATATACATATGGGTAAA
ACAACACTATGCGCTCTAATAACATCTCAAGTATTTTCACATAGTAAAT
TCAGTTATTTCCCGGTGCTAGACCATGAAACAATAAAACAATAAGGGTA
AATGCTAACGCTAGCTCTGGAAATACACTTGACACCCAAAACAACATTC
AAACTGAGTCTTGGAACACTCTTCTTGTCTACAAAGCTAACCGAAAGG
TTTATCTGACCTTTAGTACTATGTCAACCCATTCTTAGATGTTTTATGC
CACATTTAACGTACCTGTGGATCAGTAAGGAACGGACACAGGATTAGT
GTTATCAACTCAGGCGCTGTATTTAGTCAATACTGCCTTGAGTTGGGAG
TTTTCTACTAGCCACAGCGACACATGCTGTGCGACGAAGGTAACATA
CATTCAACCCAAAGGTTTGAAGAAAAAGAAAAGTTATTTATAAAGTAGCT
GACGATTTAGATGTTGTACAAAAAATAAAAAATAAGGGGTTGTCAAGT
TTTATATTAACCTTCAAATAGAGGCTATTTTCCCATGCTCTTTATTAATA
AAATAAAAAAATAAGGGGGTACGTGATGACACCGAGTAAGATGACTG
CTTAAACCTGCGGCTGCGGCACCACTAGCTCACATTTGTGTGAACATAT
AGGCATTTGCAACTTATTACACGTTTTTTTTTCCCCACAAGTTGCCTAT
TCTAAGCGTCTTCCACTACCAATATGCCGAATCCCCGGAAGAAACGAAA
ATGCGTCTCAACAACAGCAAAGCGCCAGGGCTAGCGCTAGCTTAAAG
CCCCCACGCCTCGCCTGATGACGCAGATTCCATCTCCACTGCTAAAAT
ACTCGAGCTTCTTCCAGAACTATCAGACGATCAAGCCAAACACTTTGC
AGACATCCGCAAGGACGTCTCAGACATAAGAAGAGCTATCGAGGCGAT
TGAGGGTAAGCTAGCCGACATGCTGACAAGGATGACTAGCGCAGAGGC
TCGCCTTGACATGCTGGAGGAGGCGGAGGGACGGCGCCGCAACGCGC
CGCCGGCCTCAGCCTCTGAAGTTGAAAACTGAACGCCAAATTAAGT
AGGTCGAGGACCGGGAAAGGCGAGTGAATCTACGCATTTATGGTTCC
```



Genome Annotation

HiC_scaffold_1	maker	mRNA	1677	5926	.	-	.	ID=augustus-gene-0.1
HiC_scaffold_1	maker	exon	1677	1769	.	-	.	ID=augustus-gene-0.1
HiC_scaffold_1	maker	exon	1960	2079	.	-	.	ID=augustus-gene-0.1
HiC_scaffold_1	maker	exon	2168	2232	.	-	.	ID=augustus-gene-0.1
HiC_scaffold_1	maker	exon	2436	2503	.	-	.	ID=augustus-gene-0.1
HiC_scaffold_1	maker	exon	3300	3436	.	-	.	ID=augustus-gene-0.1
HiC_scaffold_1	maker	CDS	5909	5926	.	-	0	ID=augustus-gene-0.1
HiC_scaffold_1	maker	CDS	3300	3436	.	-	0	ID=augustus-gene-0.1
HiC_scaffold_1	maker	CDS	2436	2503	.	-	1	ID=augustus-gene-0.1
HiC_scaffold_1	maker	CDS	2228	2232	.	-	2	ID=augustus-gene-0.1
HiC_scaffold_1	maker	3' UTR	2168	2227	.	-	.	ID=augustus-gene-0.1

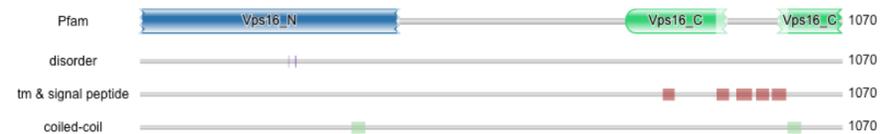
Protein Sequence

```
>maker-HiC_sc_1_gene-0.1
MCWNLKDGRLRDSLVSAAAPYGGPIA
LLREPHRRSPSSRPQLEIYSASGV
GIASFPPWKSQVVLGWTVSDLELL
CIQEDGSVLIYDLFGSFKRHFSMG
QEVVQSQVLEAKVFHSPYGTGVAI
VTGSSRFTLATNIDDLKLRRLPEV
PGLQKPCSWVVLTDQDRQTKVLLS
NGSELFILDSTCTAVCPPGLGPQ
AGSVVHMSVSFSYKYLAL.....
```

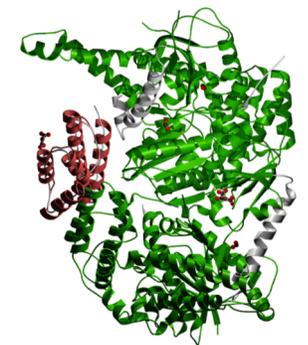
Gene Structure



Domain Architecture

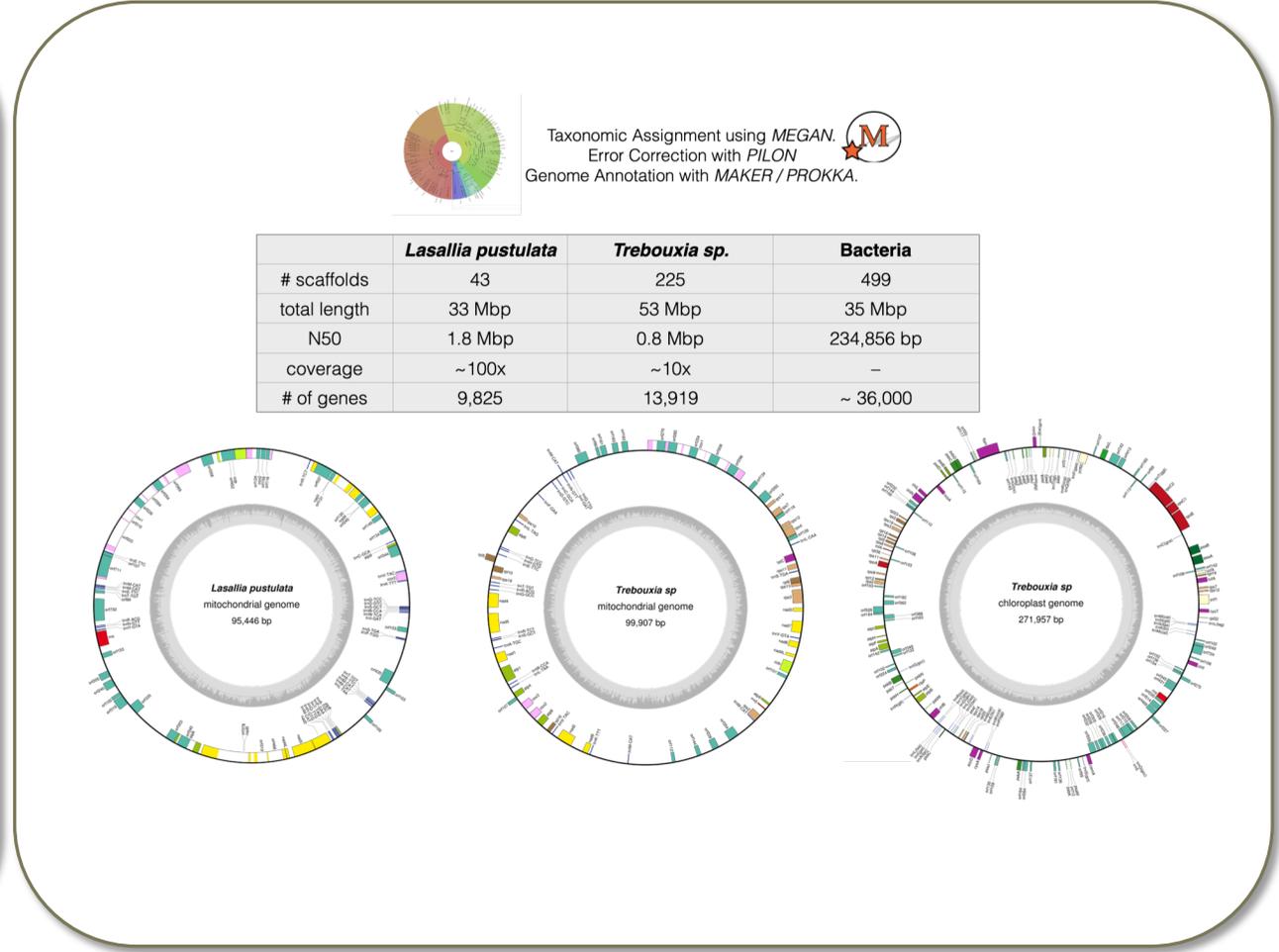
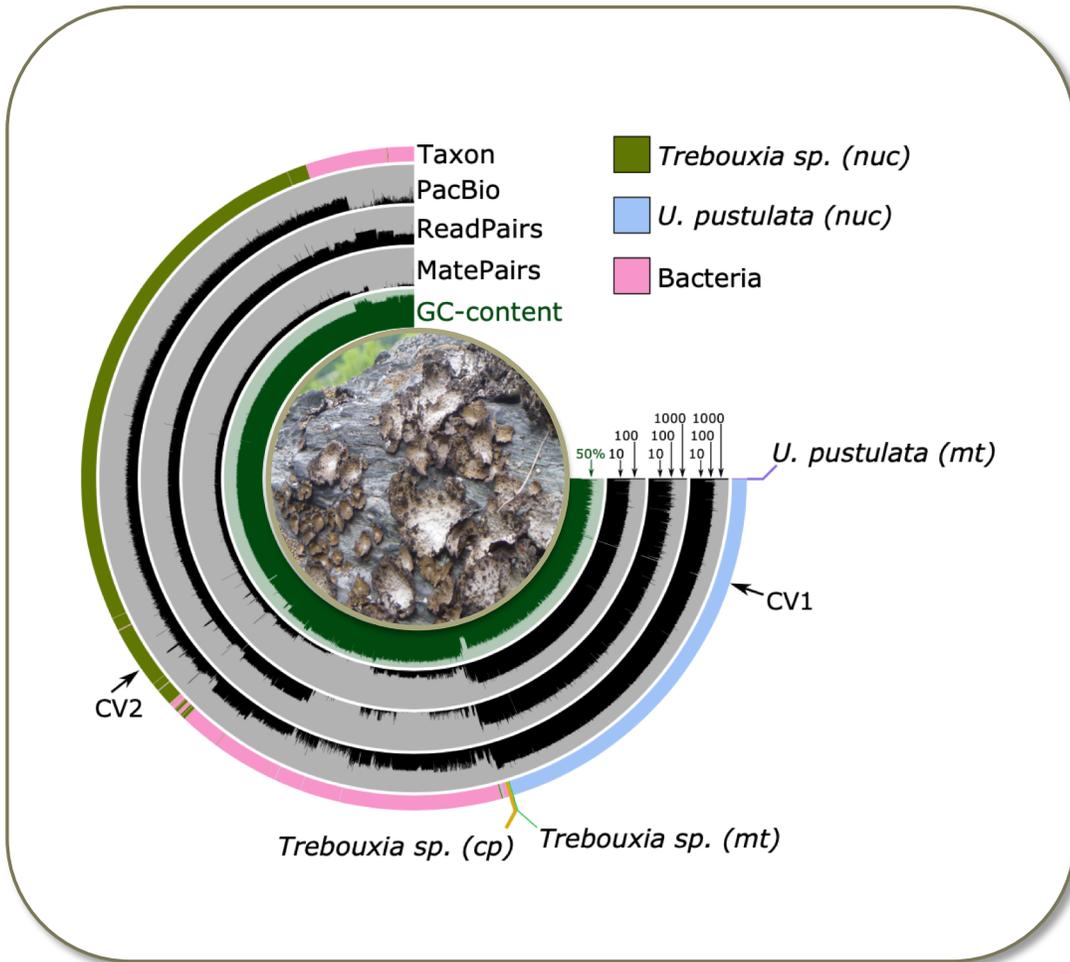


3-D Structure

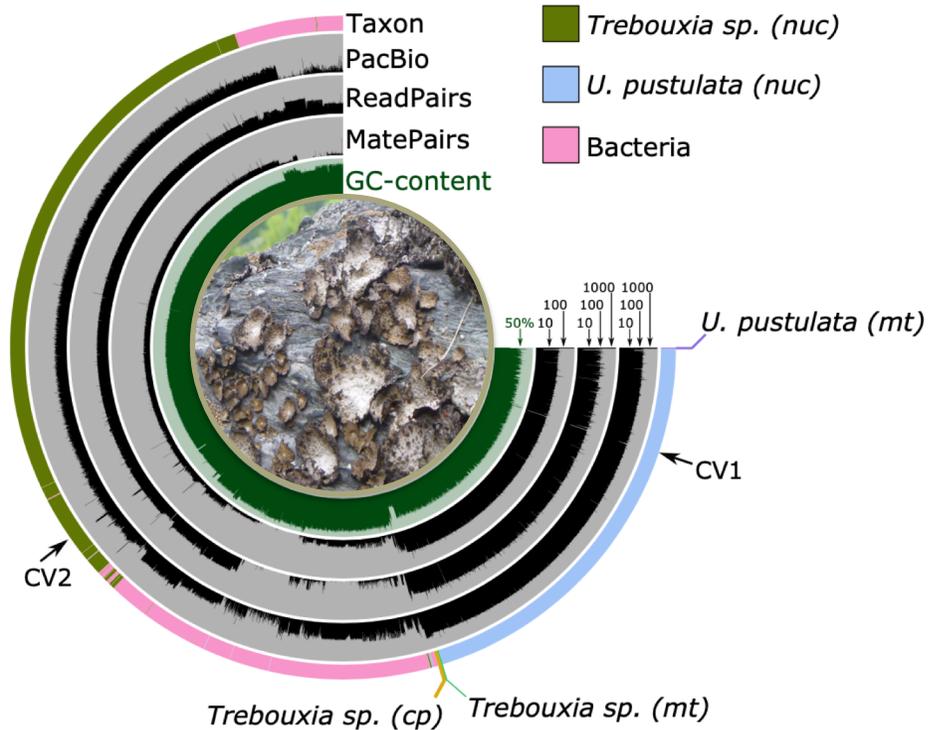


PART 0 – GET THE TEXT

RECONSTRUCTING THE HOLO-GENOME OF THE LICHEN *U. PUSTULATA*

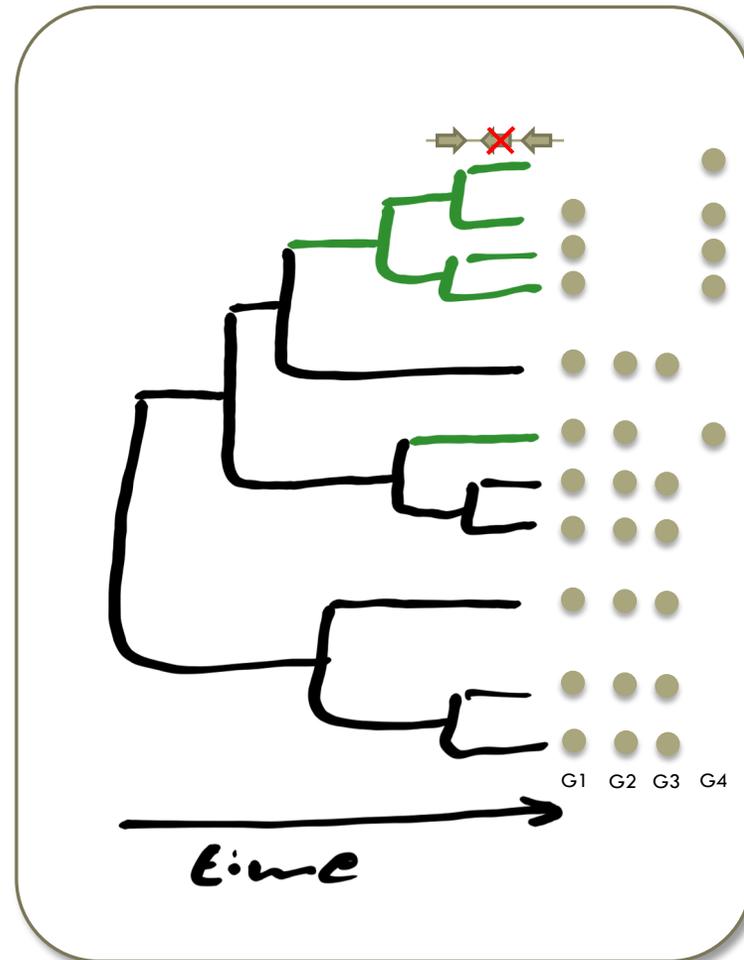


RECONSTRUCTING THE HOLO-GENOME OF THE LICHEN *U. PUSTULATA*

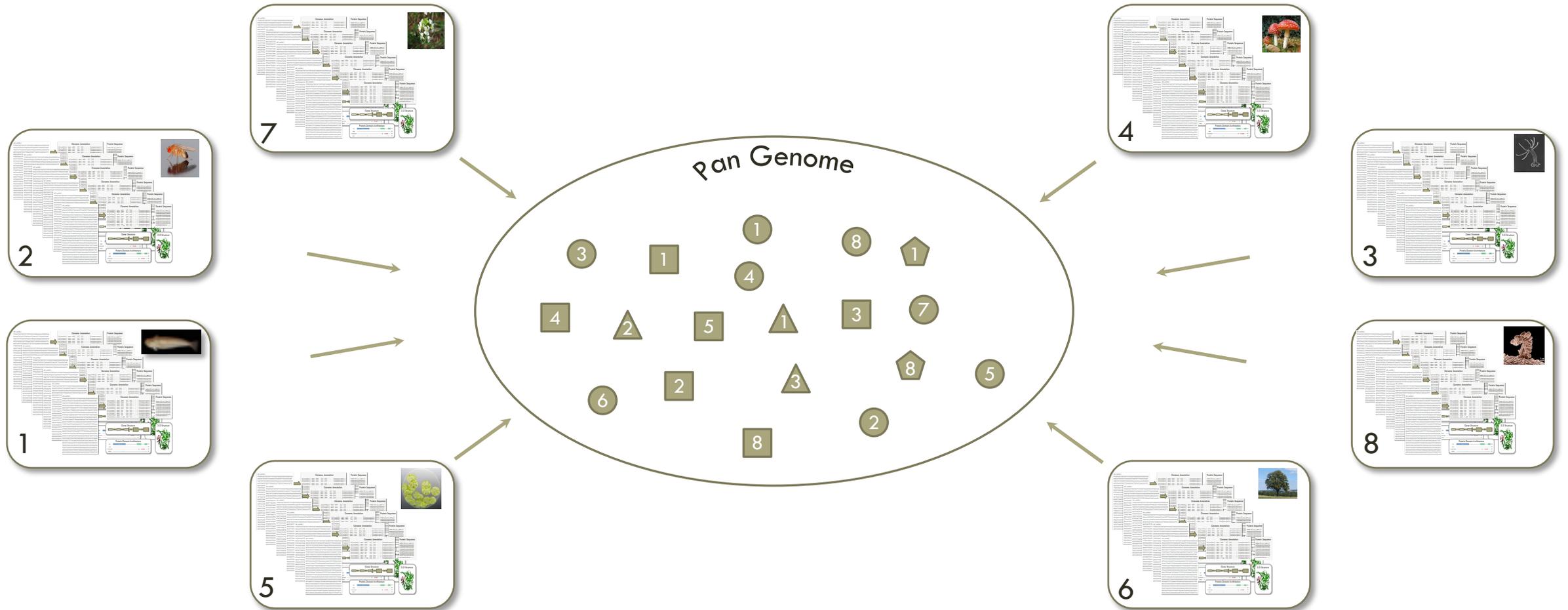


	Species ^a	Size (Mb)	Scaffolds	N50 (Mb)	Genes	Missing BUSCO (%) ^b	FGMP: HCE (%) ^c	FGMP: Proteins (%) ^c
Fungus	<i>U. muehlenbergii</i>	34.6	7	7.0	8,822	1.3	90.3	94.9
	<i>A. radiata</i>	33.5	17	2.2	na	3.0	87.1	97.8
	<i>U. pustulata</i> ^M	33.5	43	1.8	9,825	3.6	90.3	96.8
	<i>G. flavorubescens</i>	34.5	36	1.7	10,460*	1.5*	77.4	97.3
	<i>X. parietina</i>	31.9	39	1.7	11,065	1.4*	77.4	96.6
	<i>C. metacorallifera</i>	36.7	30	1.6	10,497*	3.0*	83.9	97.3
	<i>C. macilenta</i>	37.1	240	1.5	10,559*	2.7*	80.6	96.3
	<i>P. furfuracea</i>	37.8	46	1.2	8,842	1.8	93.5	97.1
	<i>R. intermedia</i>	26.2	198	0.3	na	3.3	87.1	97.3
	<i>E. prunastri</i>	40.3	277	0.3	10,992	1.3*	87.1	96.6
	<i>C. rangiferina</i>	35.7	1,069	0.3	na	2.5	80.6	98.0
	<i>C. grayi</i>	34.6	414	0.2	11,388	3.0*	87.1	96.8
	<i>E. pusillum</i>	36.8	908	0.2	9,238	3.9*	80.6	96.0
	<i>L. hispanica</i>	41.2	1,619	0.1	8,488	1.6	90.3	97.3
	<i>R. peruviana</i>	27.0	1,657	<0.1	9,338*	6.7*	80.6	95.4
	<i>L. pulmonaria</i>	56.1	1,911	<0.1	15,607	1.5*	83.9	97.0
	<i>C. uncialis</i>	32.9	2,124	<0.1	10,902*	5.3*	87.1	97.1
	<i>C. linearis</i> ^M	19.5	2,703	<0.1	na	25.0	51.6	83.8
	<i>A. sarmentosa</i> ^M	40.0	915	<0.1	na	21.9	58.1	83.3
	Alga	<i>T. gelatinosa</i> ^L	61.7	848	3.5	na	68.7	na
<i>C. subellipsoidea</i> ^F		48.8	29	2.0	9,851	2.4	na	na
<i>Chlorella</i> sp. A99 ^S		40.9	82	1.7	8,298	18.4	na	na
<i>Trebouxia</i> sp. ^{L,M}		52.9	217	0.8	13,919	13.9	na	na
<i>A. glomerata</i> ^L		55.8	151	0.8	10,025	12.4	na	na
<i>A. protothecoides</i> ^F		22.9	374	0.3	7,016	12.2	na	na
<i>Trebouxia</i> sp. TZW2008 ^L		69.3	677	0.2	na	14.8	na	na
<i>Helicosporidium</i> sp. ^S	12.4	5,666	<0.1	6,035	50.8	na	na	

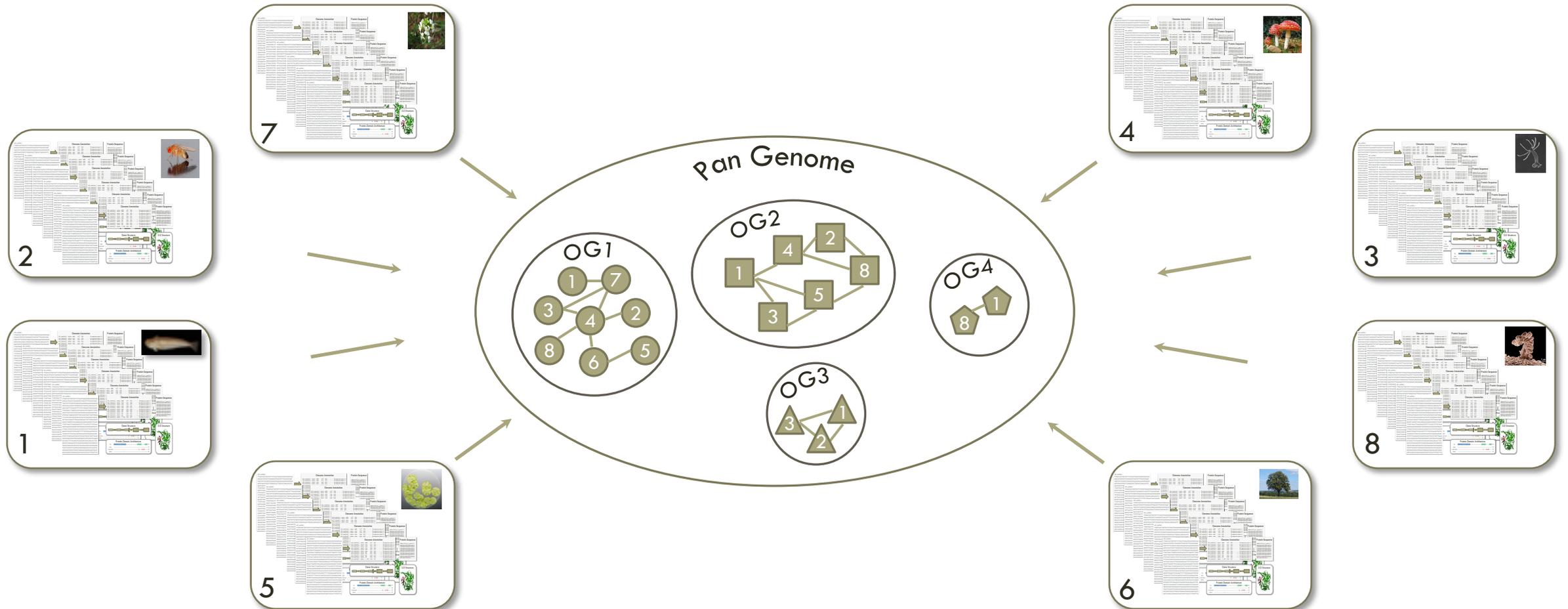
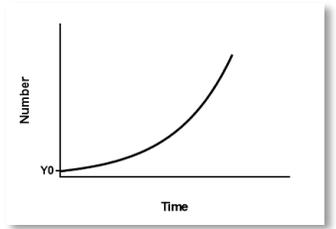
WHAT IS SPECIAL/DIFFERENT IN *U. PUSTULATA*?



FROM PAN GENOME...

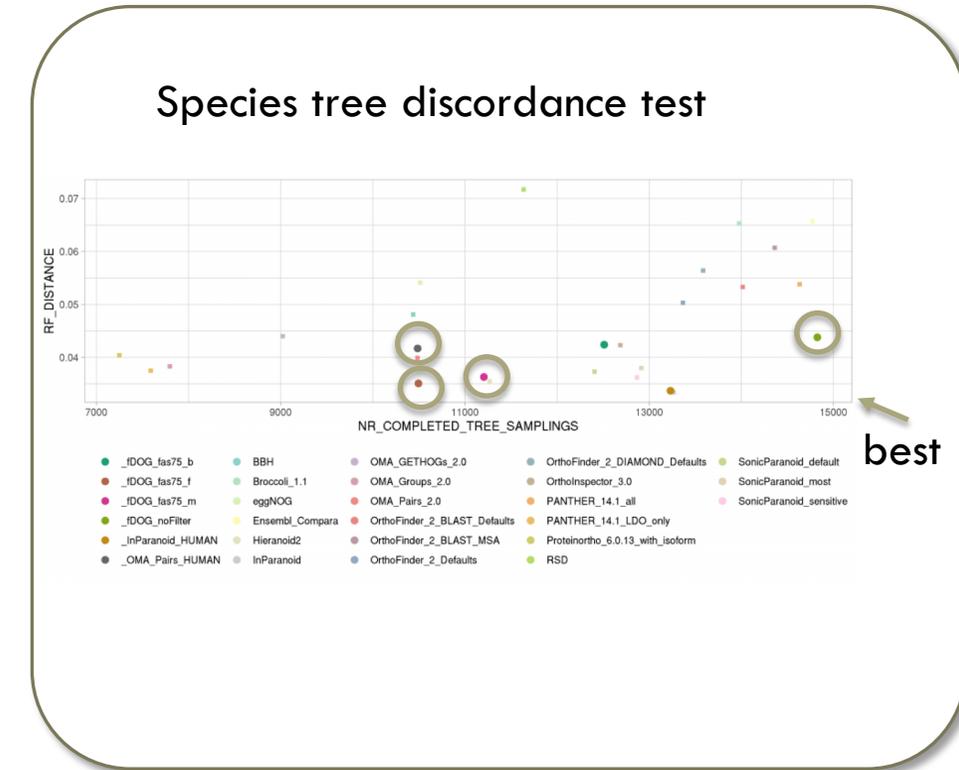
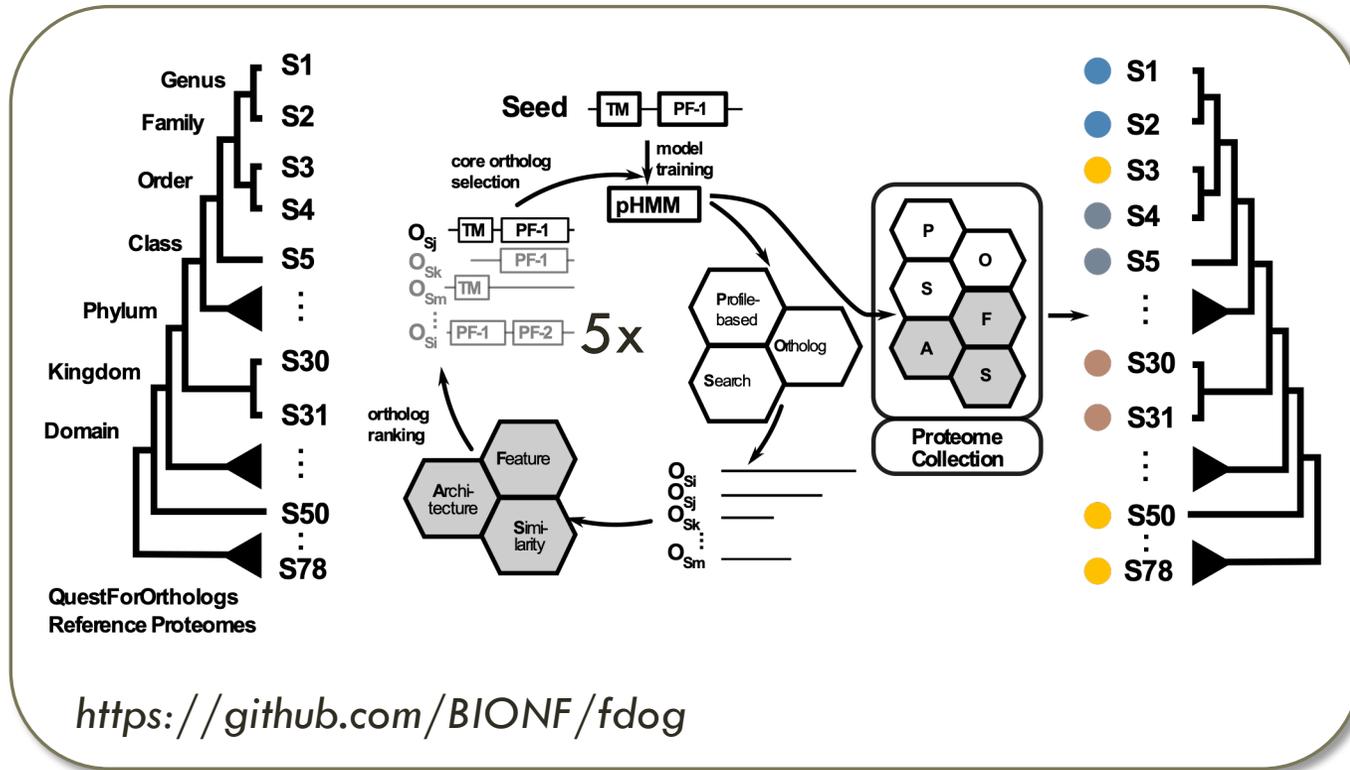


FROM PAN GENOME TO ORTHOLOGOUS GROUPS



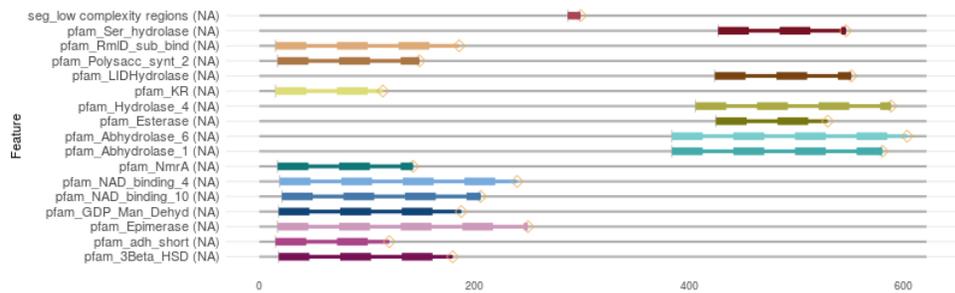
* in fact, we should replace ,genome' with ,gene set'

FDOG – FEATURE-AWARE PHYLOGENETIC PROFILING OF PROTEINS USING HMMS



IDENTIFYING GENES THAT DIVERGED IN THEIR FUNCTION SCORING ORTHOLOGS BY THEIR FEATURE ARCHITECTURE SIMILARITY

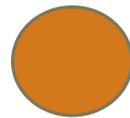
Seed : *Acinetobacter baumannii* ATCC 19606



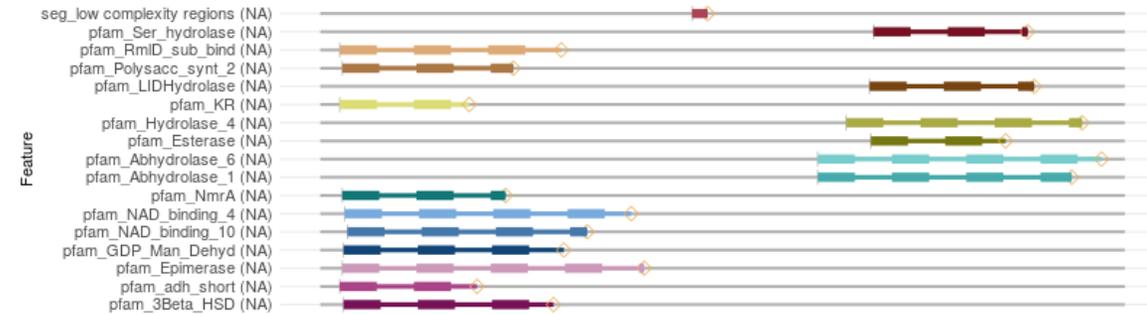
FAS=1



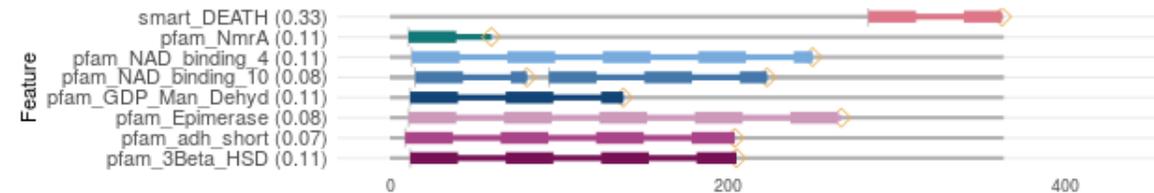
FAS=0.4



O₁: *Acinetobacter baumannii* AYE



O₂: *Acinetobacter guillouiae*



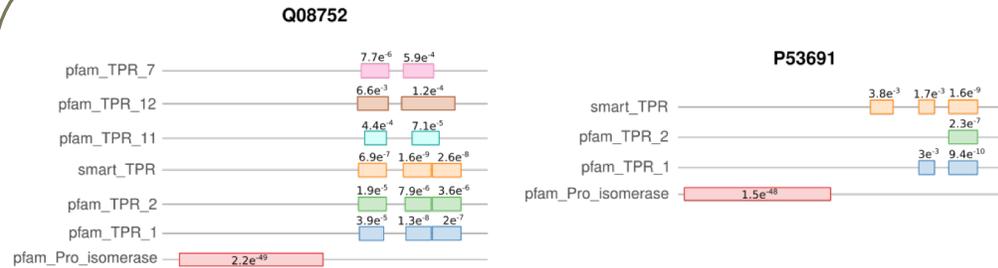
Penalize

- Absence of seed features in ortholog
- Differences in number of feature instances
- Differences in relative position of feature instances

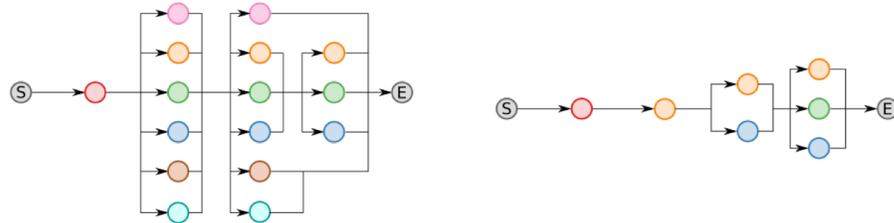
Weigh features according to their frequency in a proteome
Score in the interval of [0,1]

LINEARIZATION OF REDUNDANT DOMAIN ANNOTATIONS

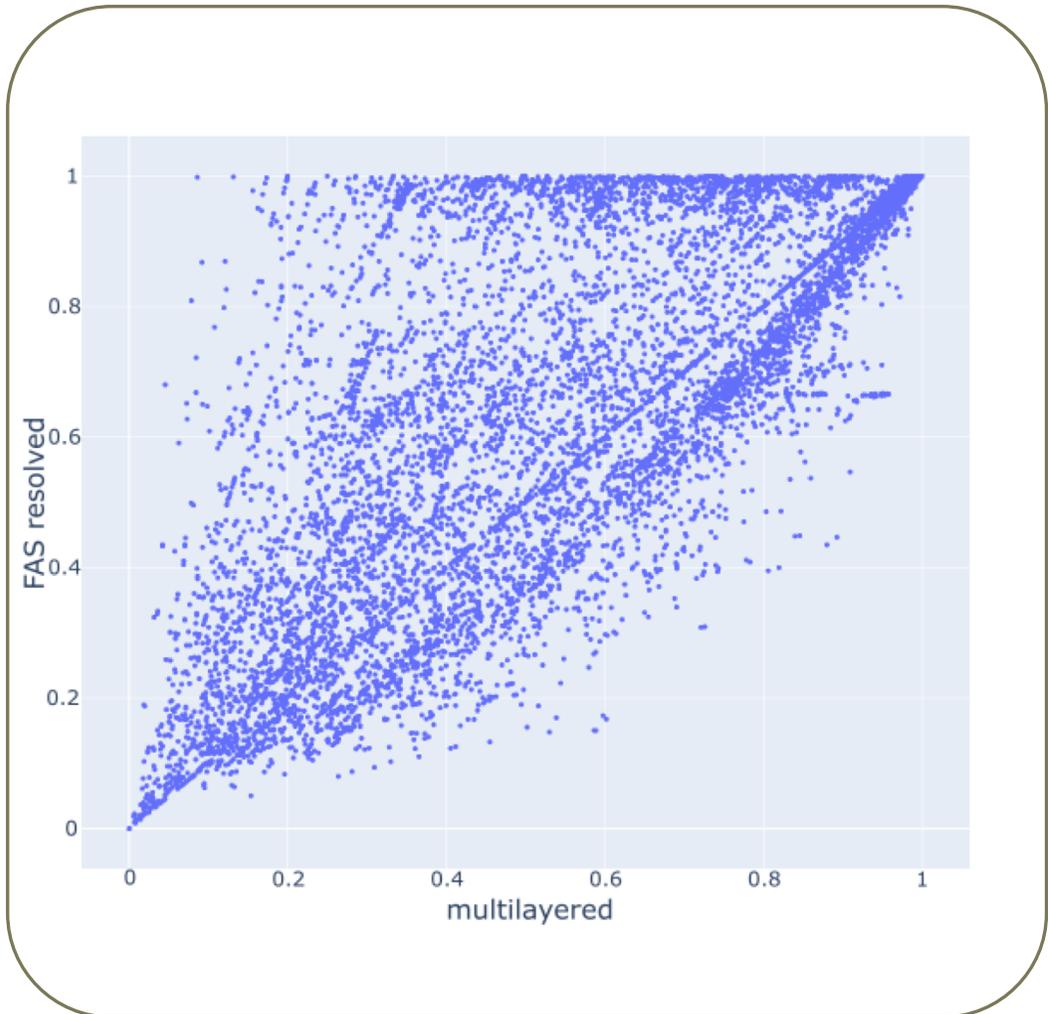
A) Unresolved Architectures



B) Build Architecture Graphs

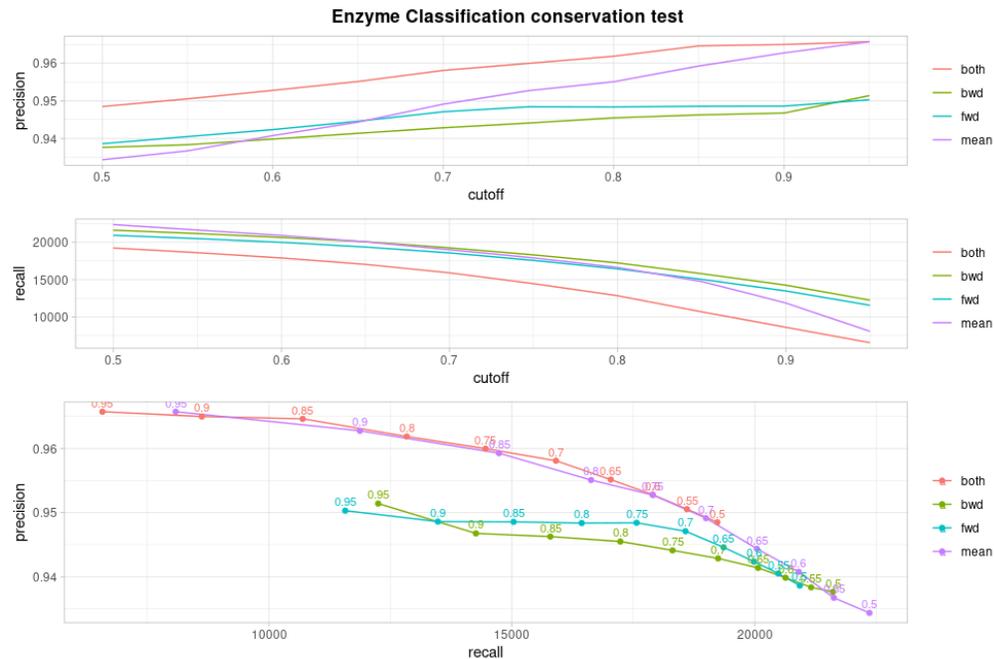


C) Linearization

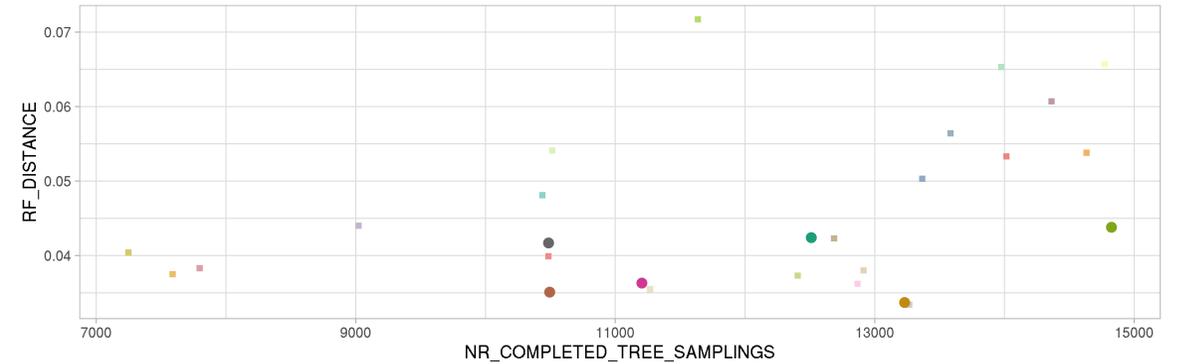


BENCHMARKING FDOG

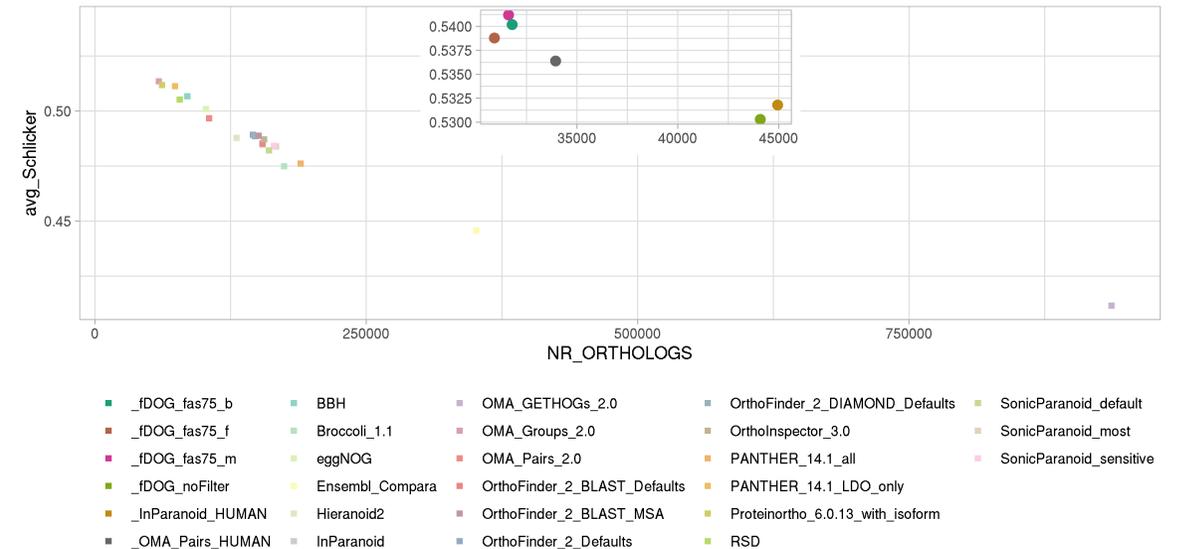
FAS cutoff



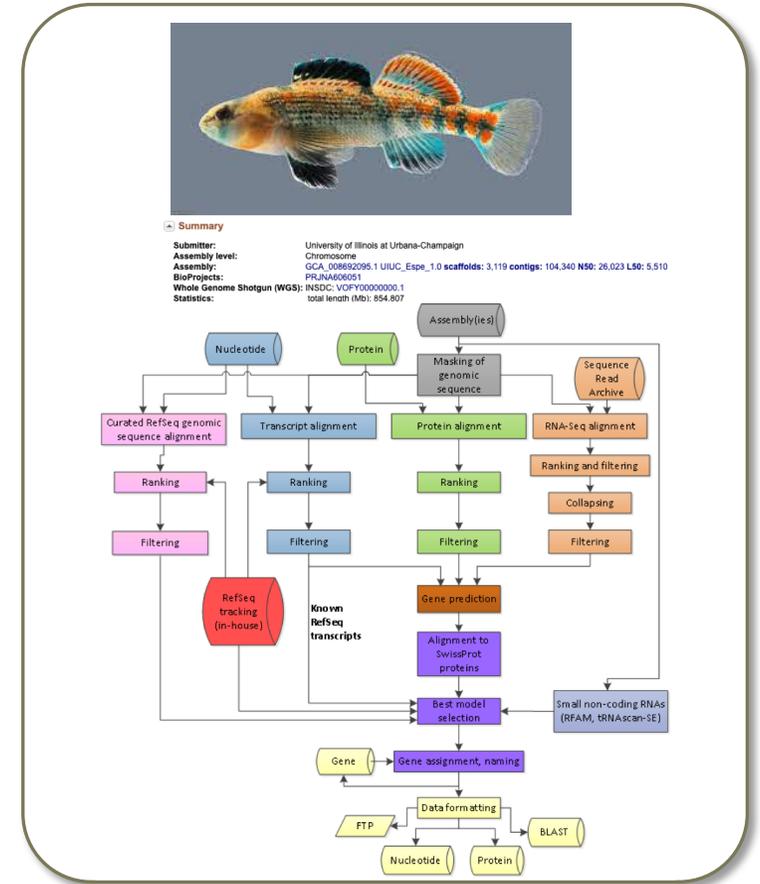
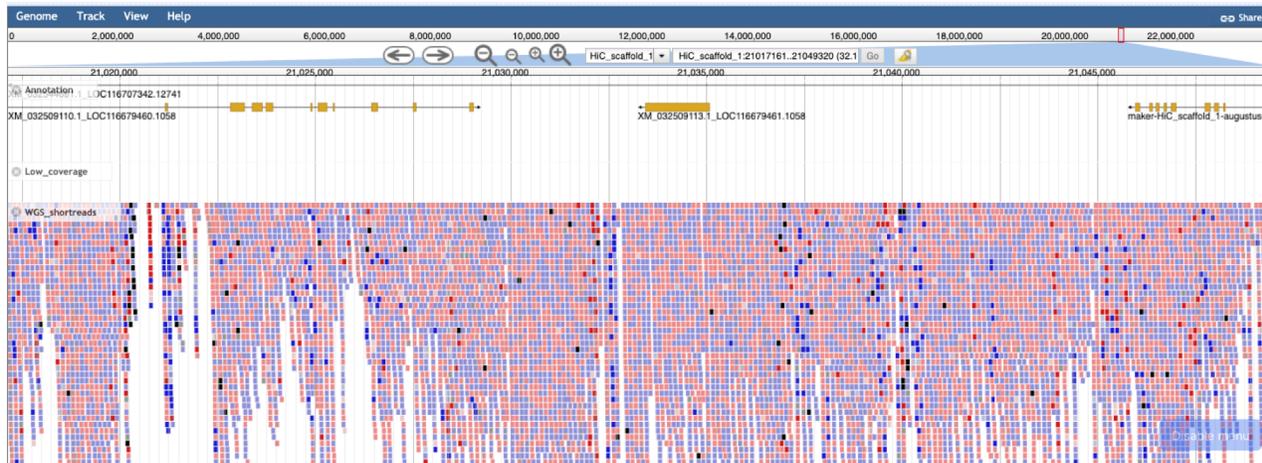
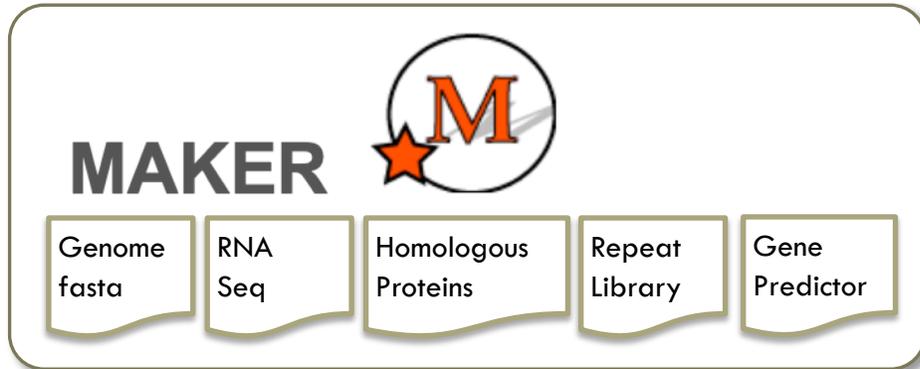
Species tree discordance test



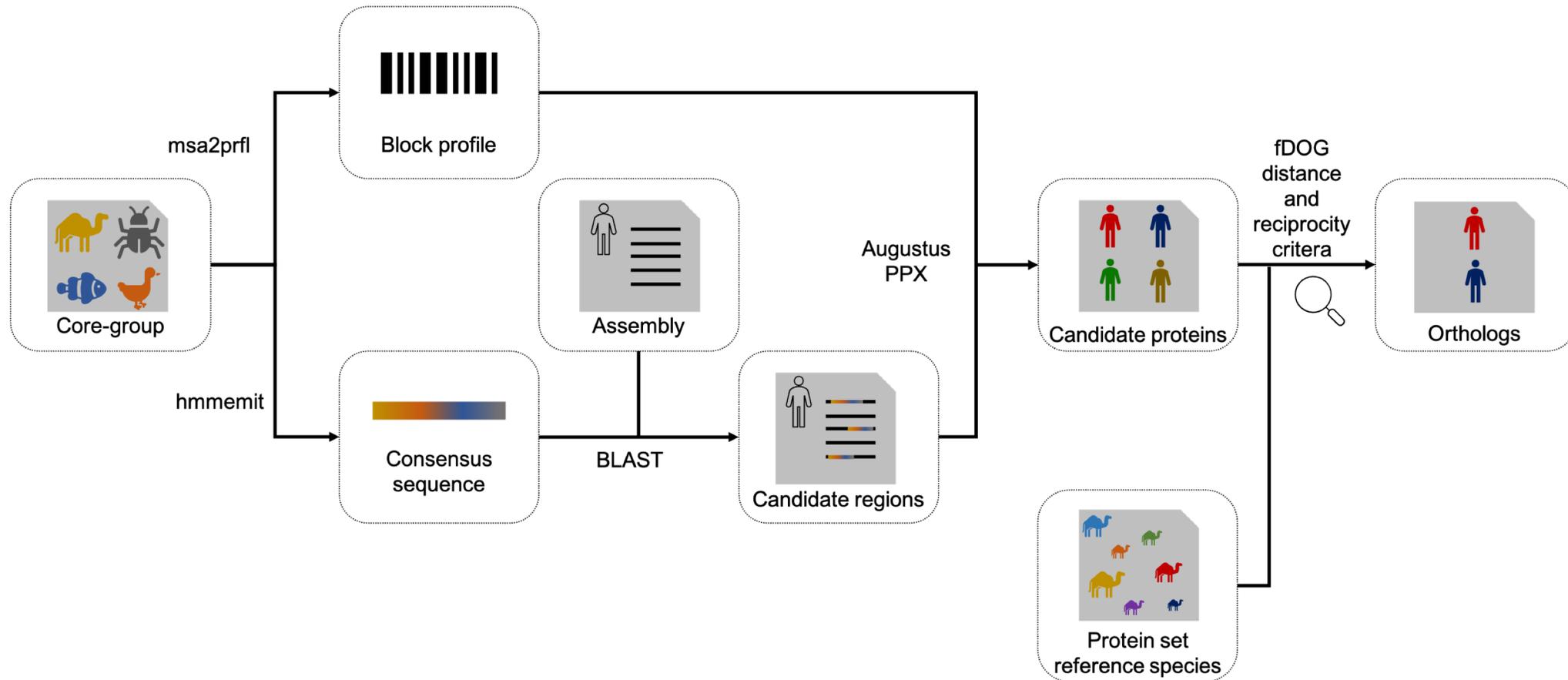
Functional similarity (GO)



ANNOTATING GENOMES – HOW COMPREHENSIVE CAN YOU BE?



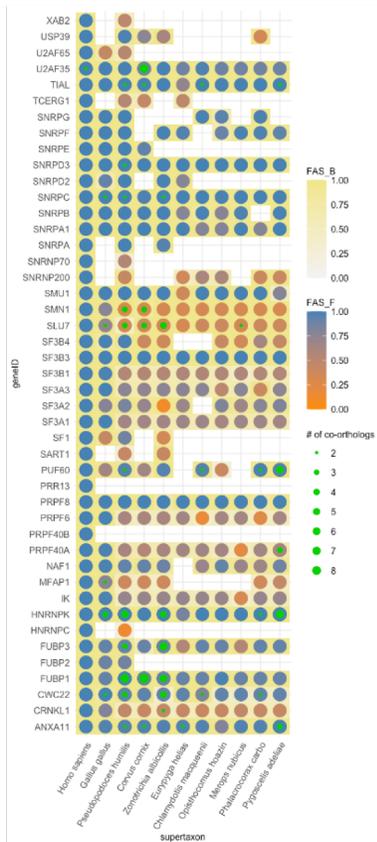
FDOG ASSEMBLY — TARGETED SEARCH FOR ORTHOLOGS IN UNANNOTATED GENOMES



fDOG ASSEMBLY — A FIRST BENCHMARK

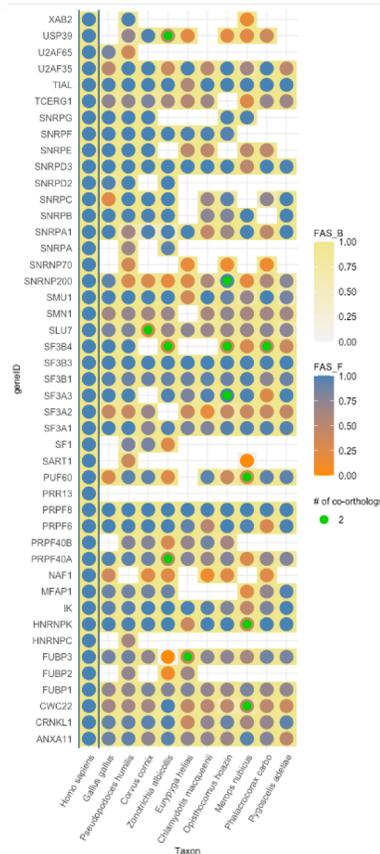
Using annotated gene sets from
NCBI Genome (RefSeq)

A: fDOG

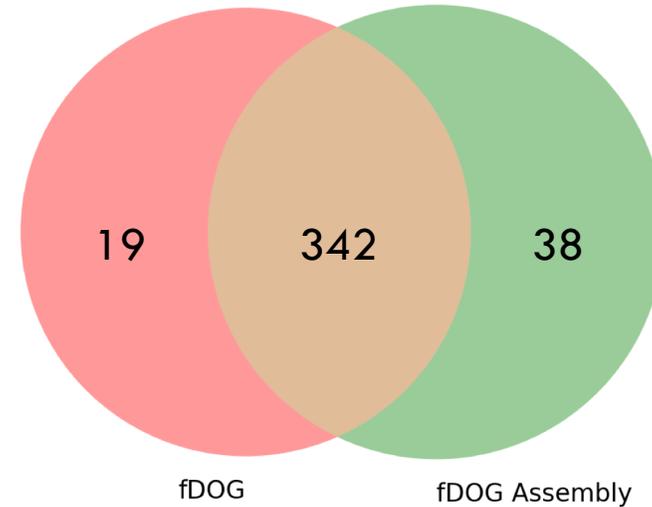


Using the genome fasta file only

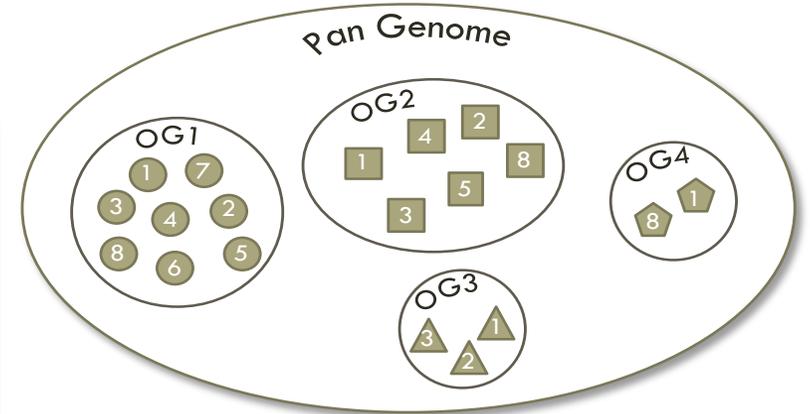
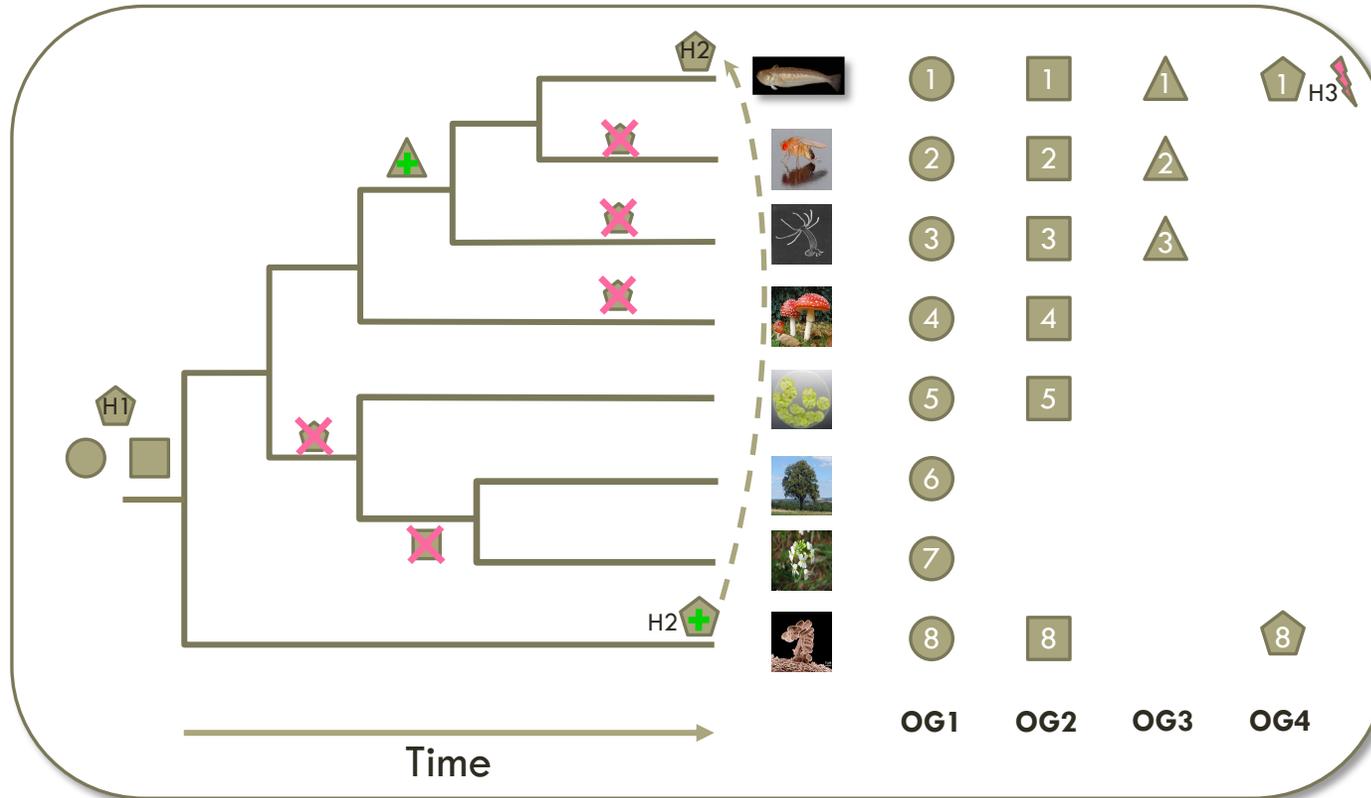
B: fDOG Assembly



Statistics:
genes: 45
B found all orthologs present in A: 33 genes
B found not all orthologs present in A: 12 genes



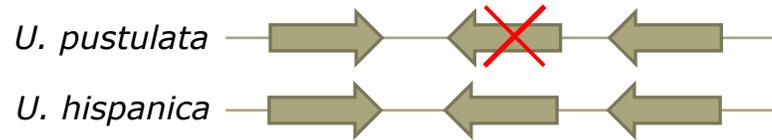
FROM ORTHOLOGOUS GROUPS TO PHYLOGENETIC PROFILES – GENE LOSS?!



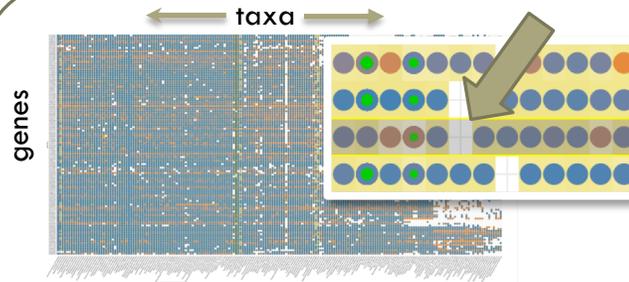
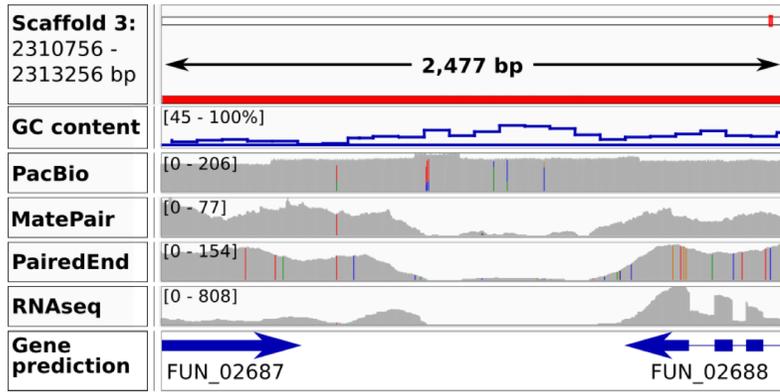
OG1 – core gene
OG2 – gene loss
OG3 – gene gain
OG4 – multiple independent gene loss (H1)
 / horizontal gene transfer (H2)/
 contamination (H3)

ABSENCE OF GENES ANCIENT GENES – EVOLUTIONARY SIGNAL OR ARTEFACT?

1. Micro-synteny analysis



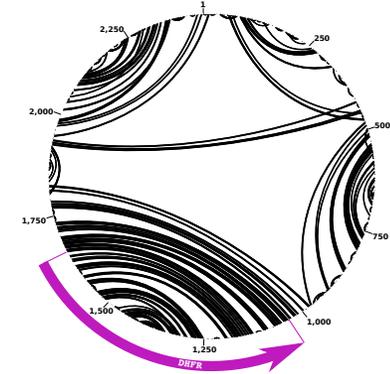
2. Assembly statistics



Dihydrofolate Reductase

Key enzyme in folate metabolism. Essential for de novo glycine and purine synthesis. Mutation is complemented by human DHFR.

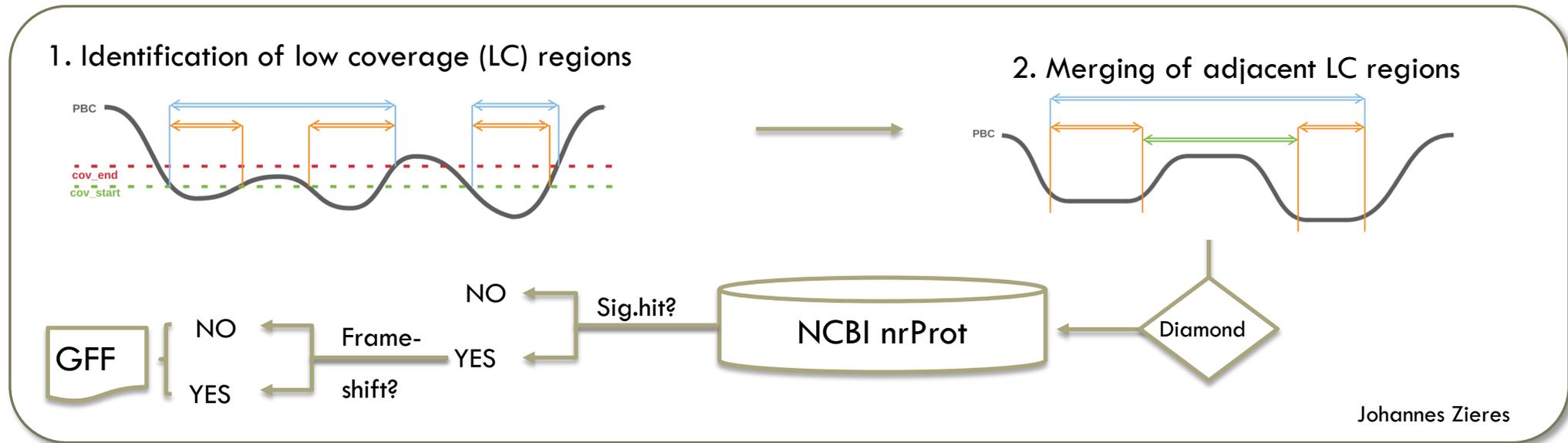
3. Secondary structure prediction



GC rich inverted repeats:

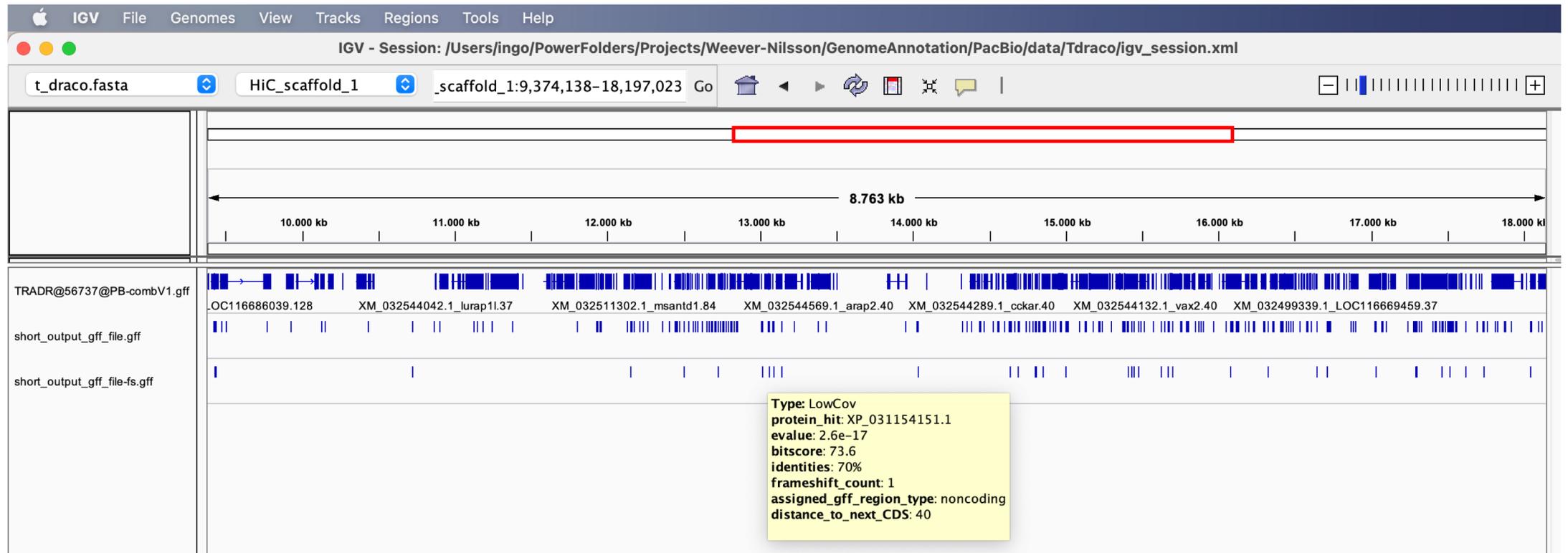
- Illumina fails (DNA & RNA seq)
- Non-random sequencing error in PacBio
- Assembly polishing fails
- Uncorrected indels mimic frame shift mutations
- Wrong inference of gene loss

UNCORRECTED FRAMESHIFT ERRORS – RARE EXCEPTIONS?

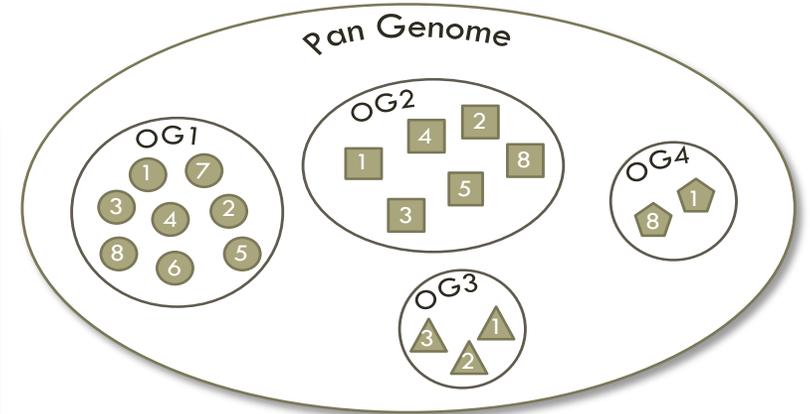
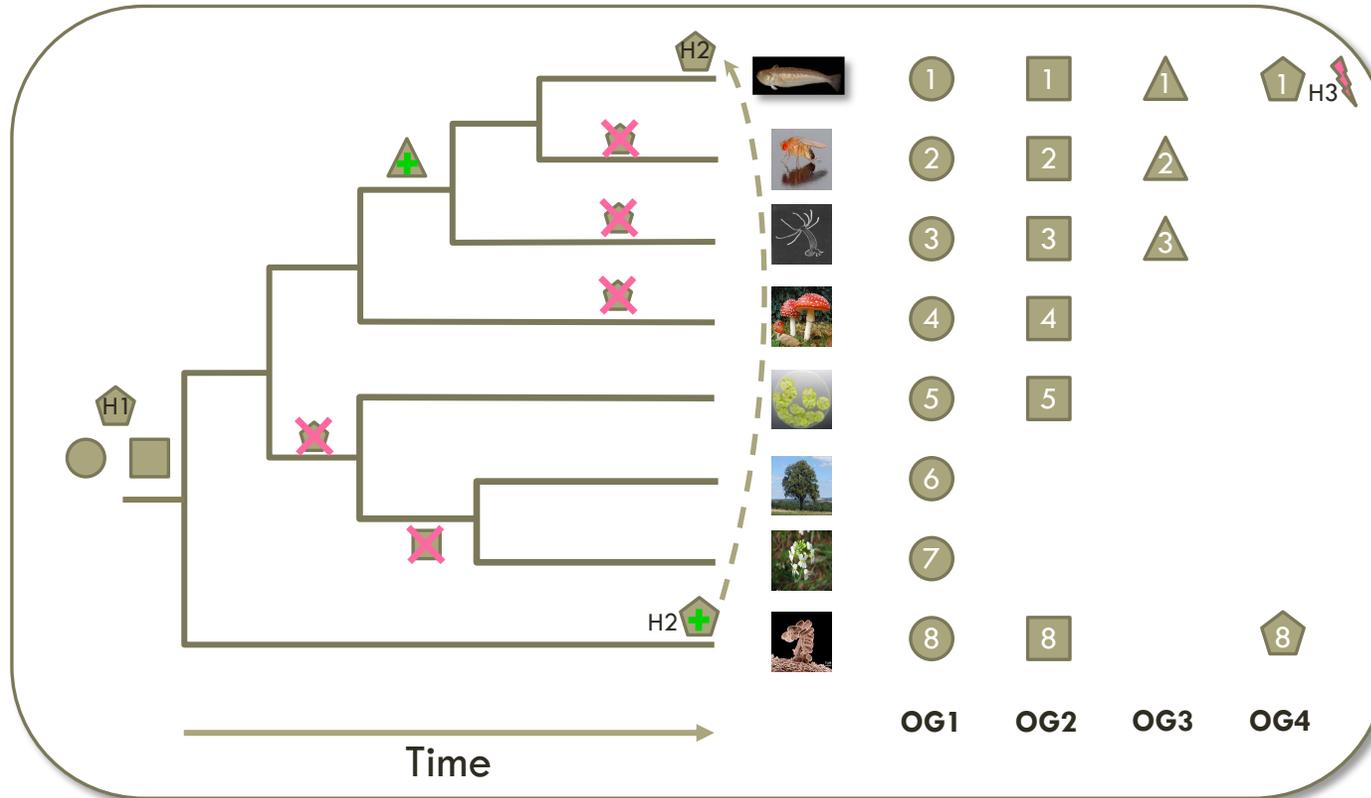


	detect_regions			merge_close_reg		database_comparison			
	cov_start	cov_end	detected regions	merge_distance	regions after merging	min_lenght	Diamond hits	detected frameshifts	regions with frameshifts
short read	15	18	435.310	150	298.957	300	26.611	8810	4846
long read	25	30	48.116		15.561				

NON-RANDOM SEQUENCING ERRORS — RARE EXCEPTIONS?

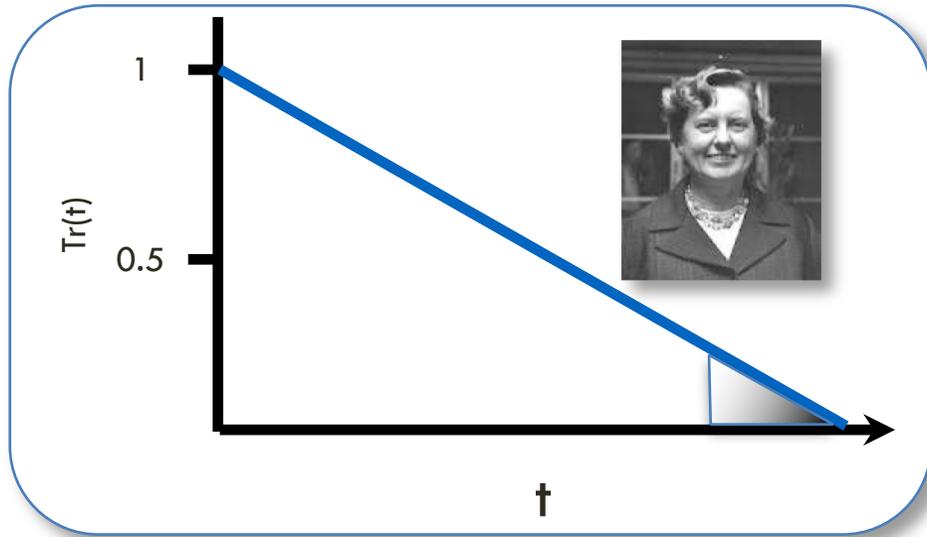


FROM ORTHOLOGOUS GROUPS TO PHYLOGENETIC PROFILES – GENE LOSS?!

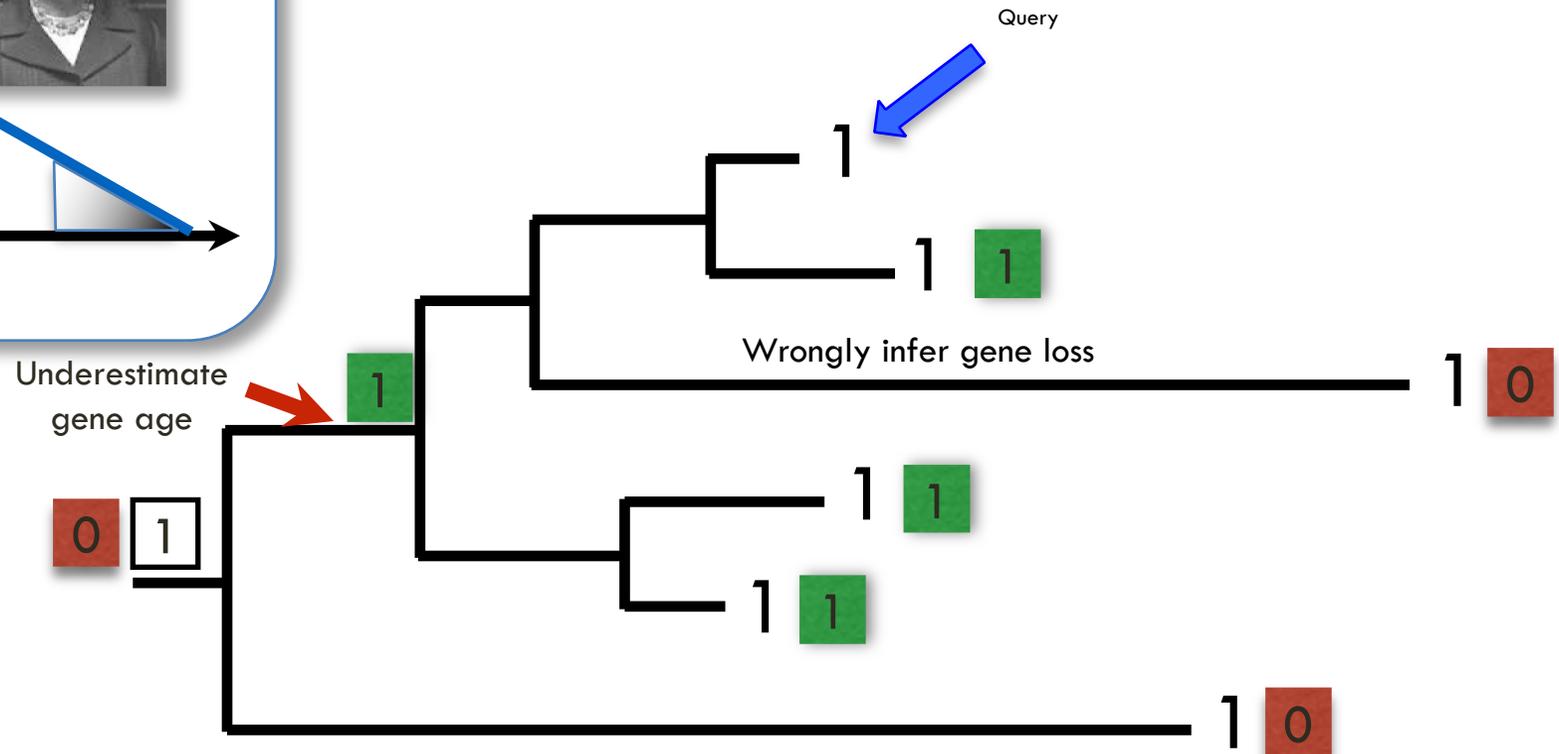


OG1 – core gene
OG2 – gene loss
OG3 – gene gain
OG4 – multiple independent gene loss (H1) / horizontal gene transfer (H2) / contamination (H3)

WHAT HAPPENS WHEN SEQUENCE SIMILARITY DECAYS AS A FUNCTION OF TIME?



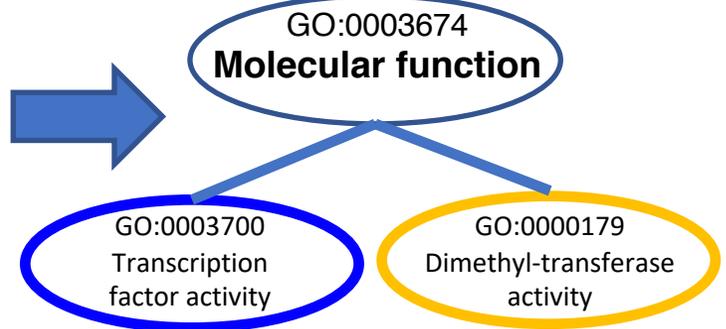
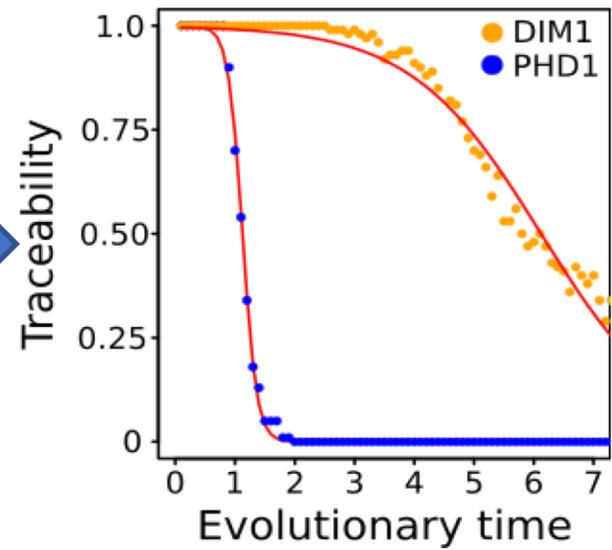
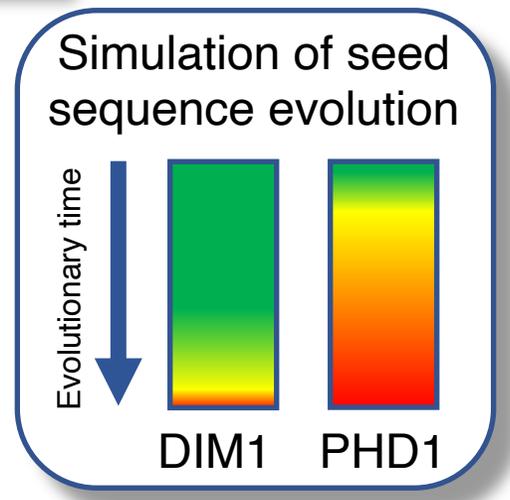
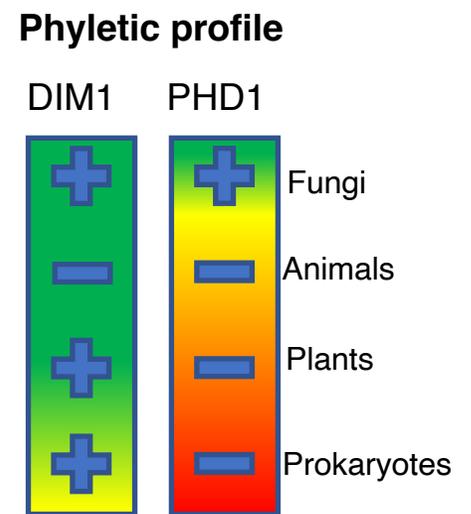
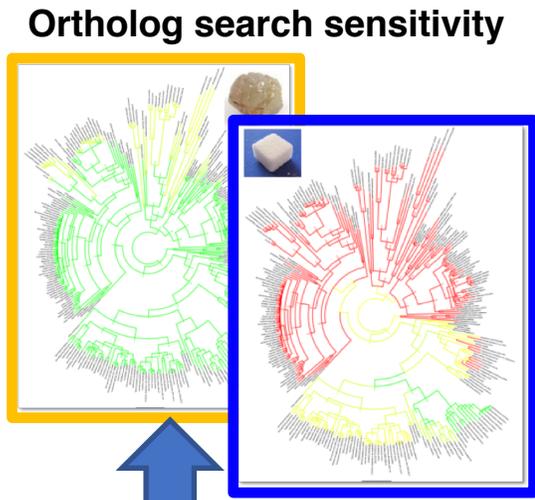
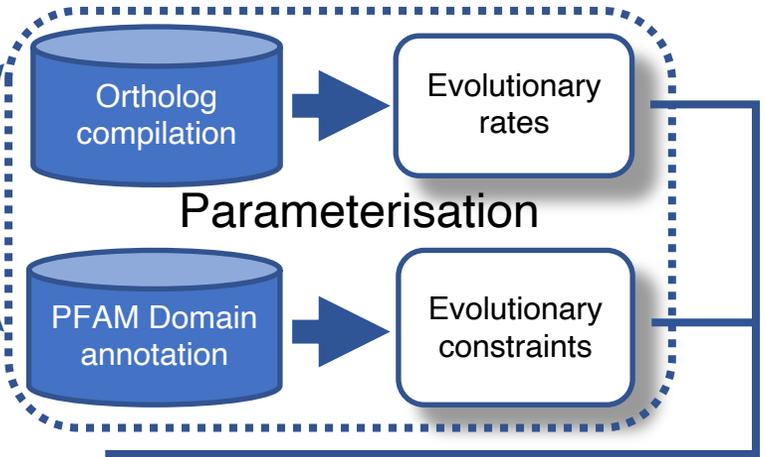
Twilight zone: Extent of sequence similarity is too low to differentiate between chance and common ancestry.



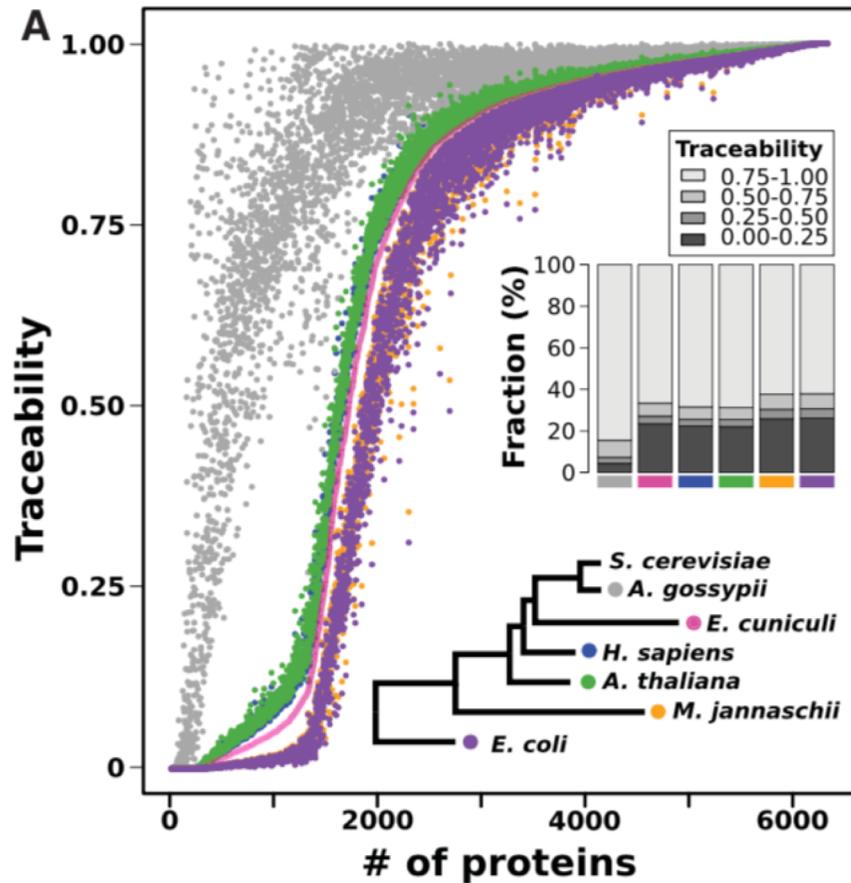
PROTRACE - A SIMULATION BASED STRATEGY TO ASSESS PROTEIN TRACEABILITY

Seed1 - DIM1
 >DIM1 YPL266W
 MGKAAKKKYSGATSSKQVSAEKHLSSVFKFNTDL
 GQHILKNPLVAQGIVDKAQIRPSDVVLEVGGPTGN
 LT
 EK
 LIN
 SA
 LE
 ST
 KE
 VG

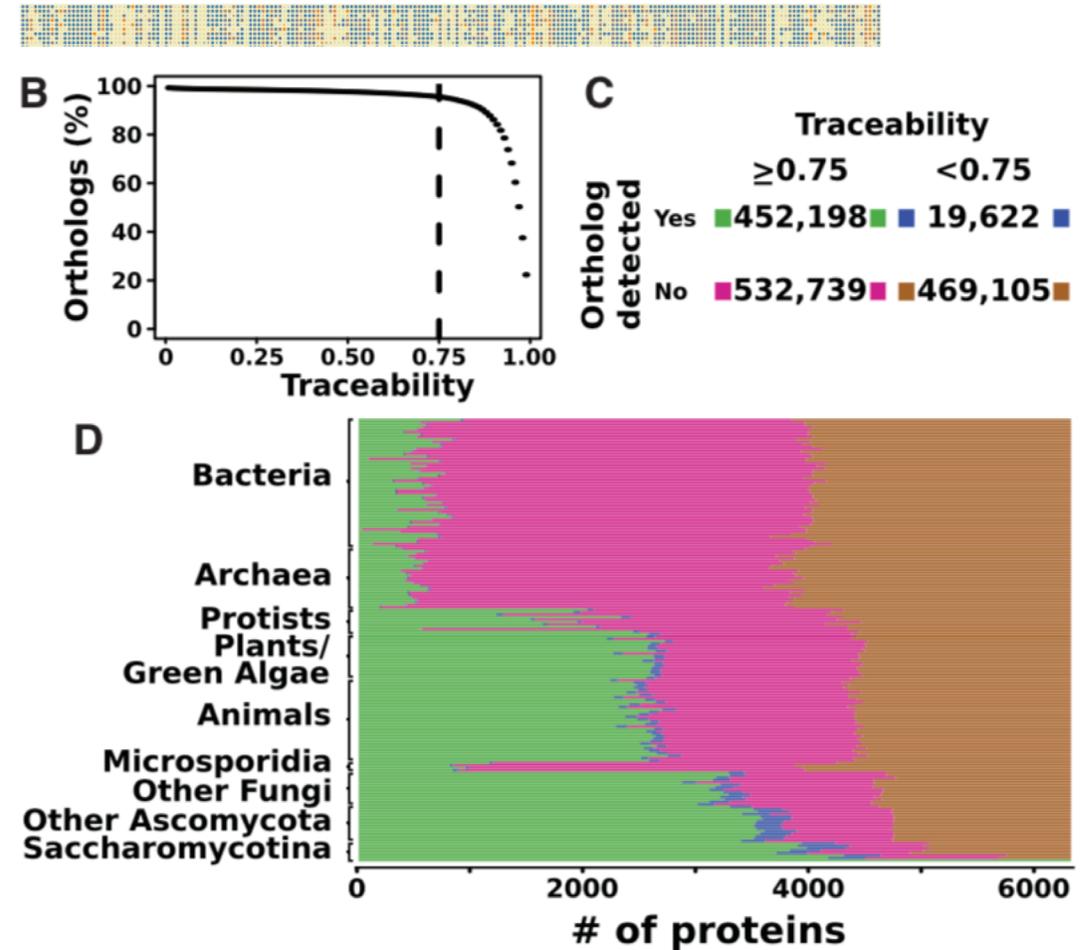
Seed2 - PHD1
 >PHD1 YKL043W
 MYHVP EMRLHYPLVNTQSNAAITPTRSYDNTLPS
 FNELSHQSTINLPFVQRETPNAYANVAQLATSPTQ
 AKSGYYCRYAVPFPTYQQPQSPYQQAVLPYAT
 IPNSNFQPSFFVMAVMPPEVQFDGSFLNTHPH
 TELPPIQNTNDTSVARPNLKSIAAASPTVATTR
 TPGVSTSVLKPRVITTMWEDENTICYQVEANGIS
 VVRRADNNMINGTKLLNVTKMTRGRRDGILRSEK
 VREVVKIGSMHLKGVWIPFERAYILAQREQILDHLY
 PLFVKDIESIVDARKPSNKASLTPKSSPAPIKQEPS
 DNKHEIATEIKPKSIDALSNGASTQGAGELPHLKIN
 HIDTEAQTSRACKNELS



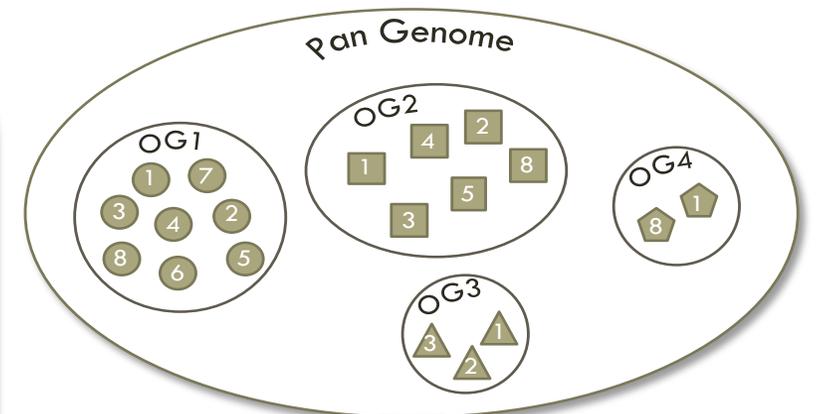
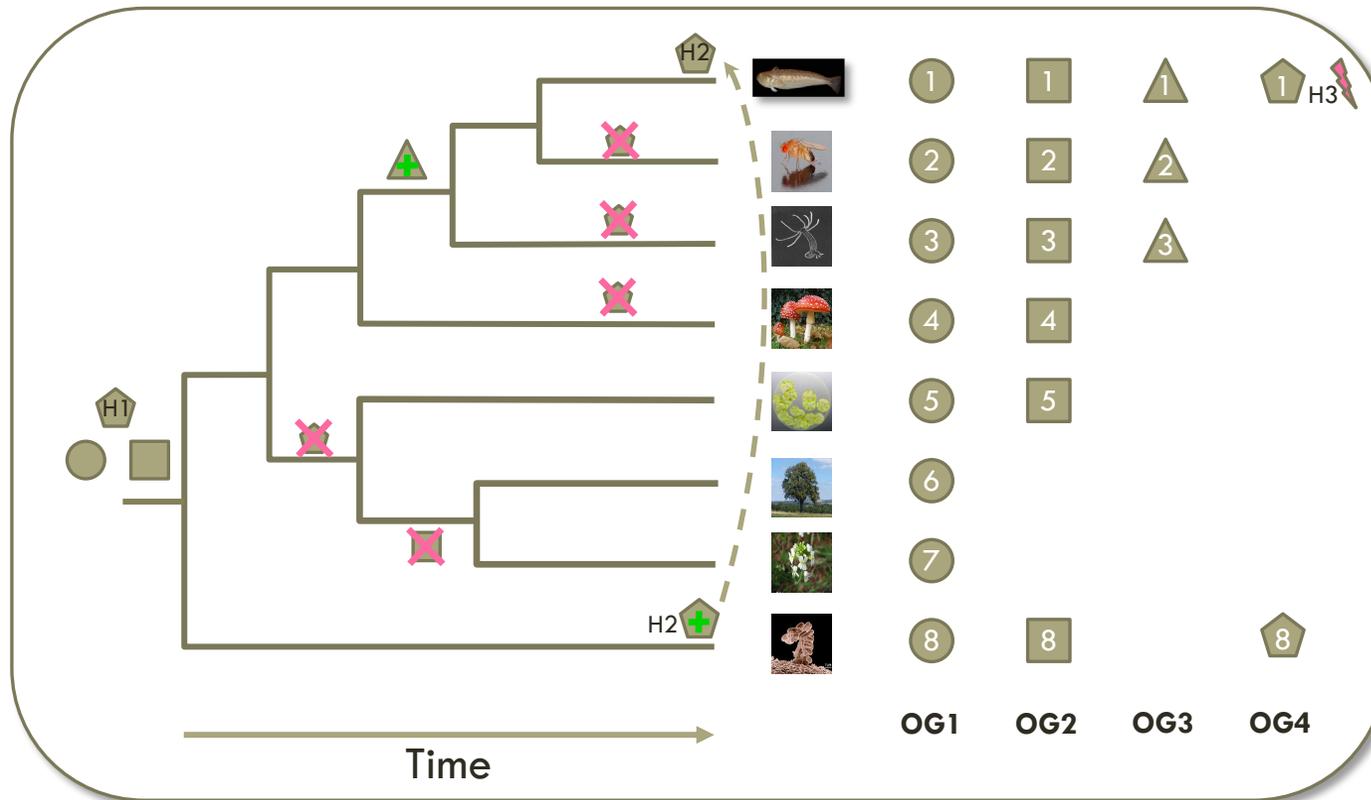
THE EVOLUTIONARY TRACEABILITY OF YEAST PROTEINS



↑ 232 Reference Species

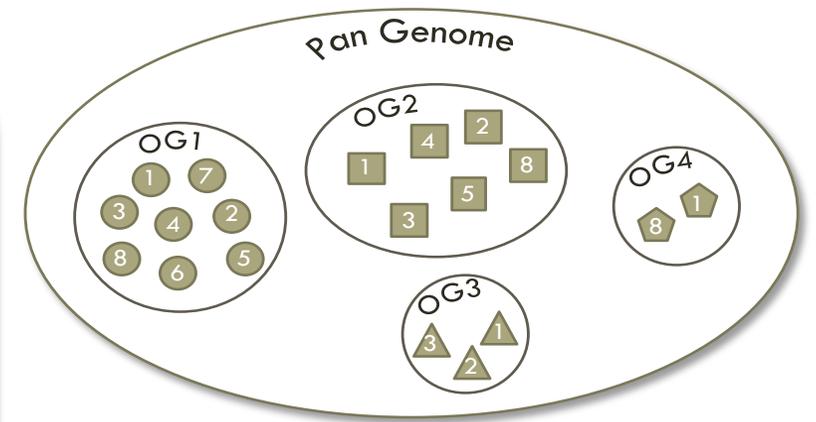
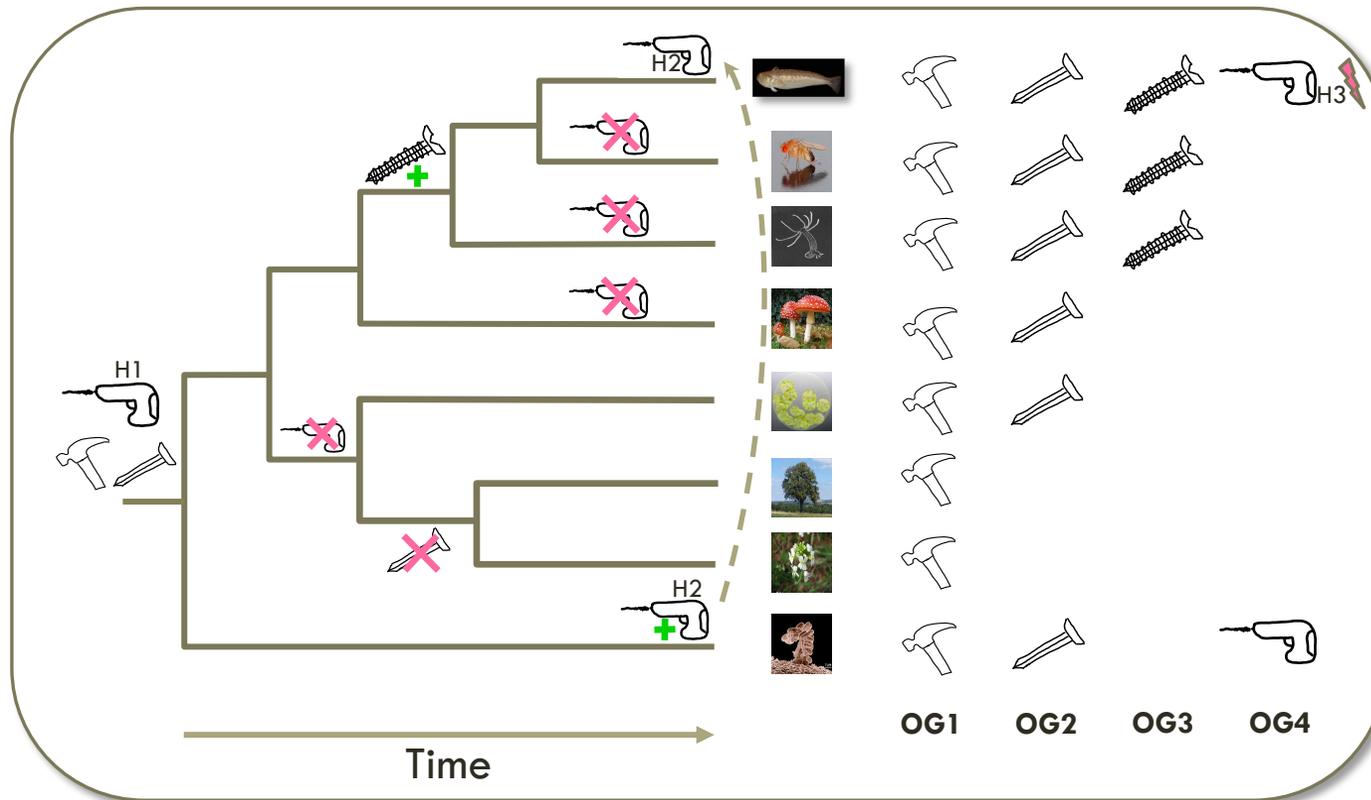


FROM ORTHOLOGOUS GROUPS TO PHYLOGENETIC PROFILES



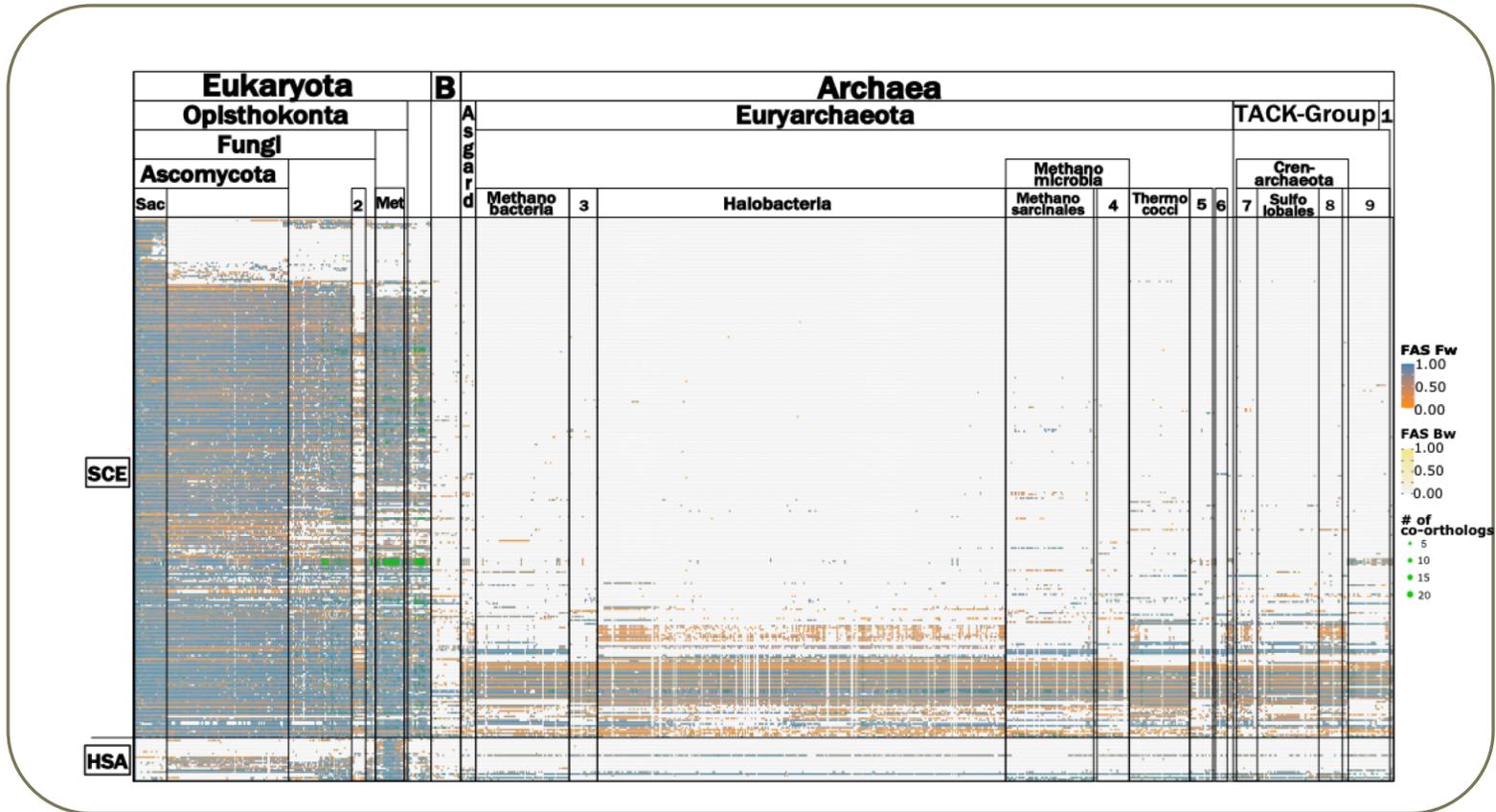
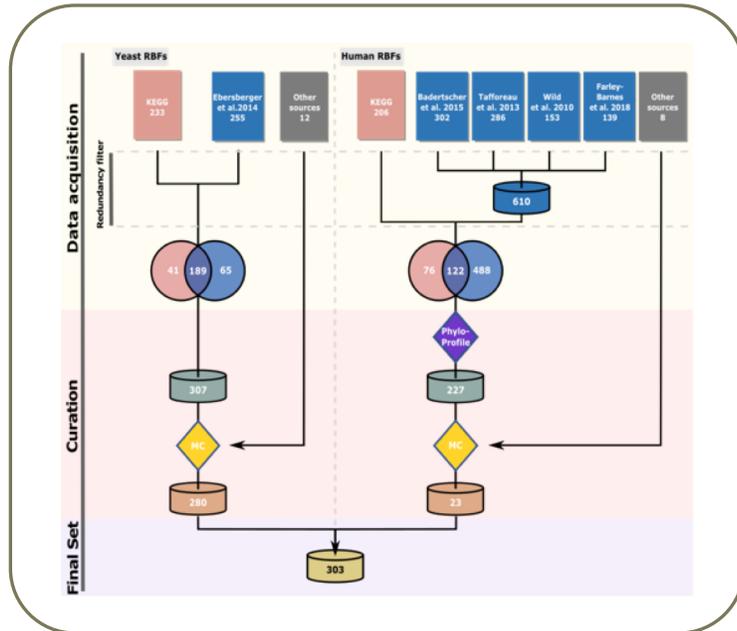
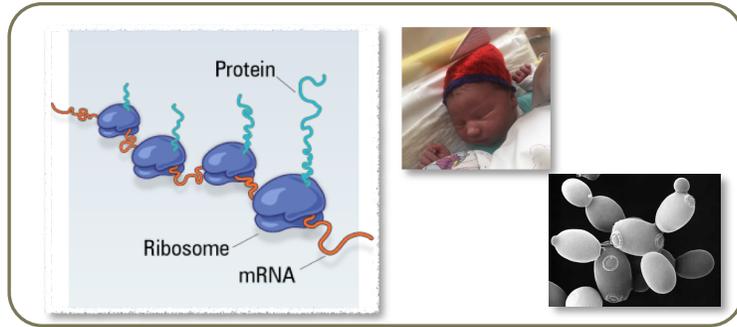
OG1 – core gene
OG2 – gene loss
OG3 – gene gain
OG4 – multiple independent gene loss (H1) / horizontal gene transfer (H2) / contamination (H3)

FROM PHYLOGENETIC PROFILES TO EVOLUTION OF FUNCTION

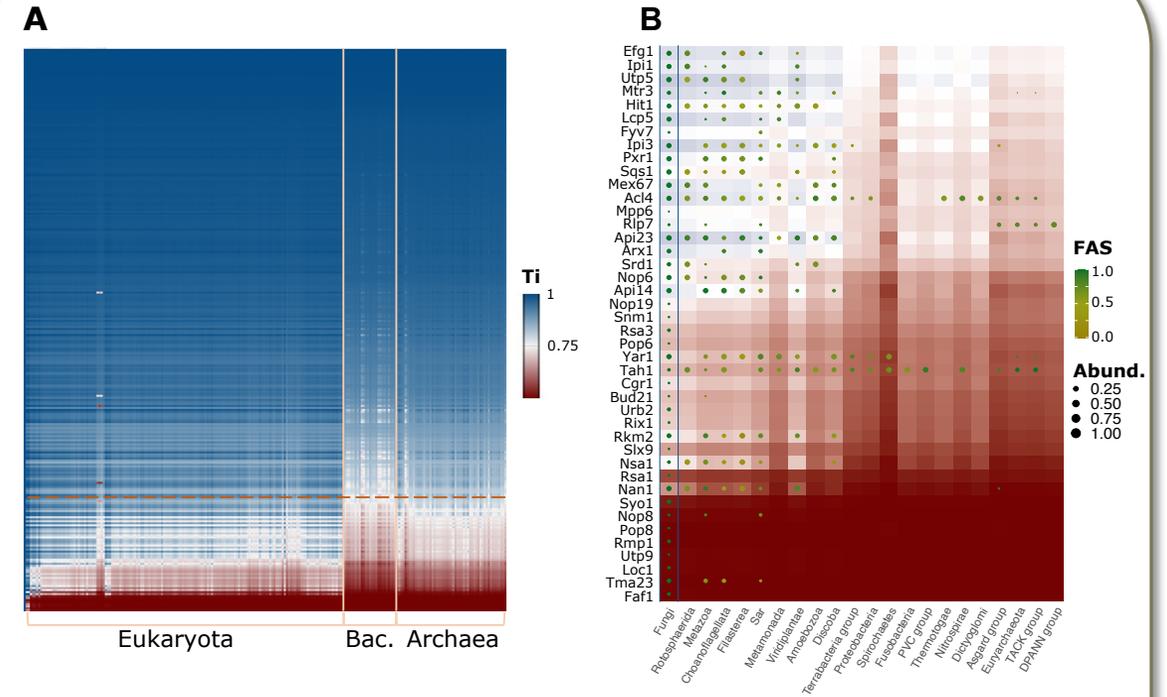
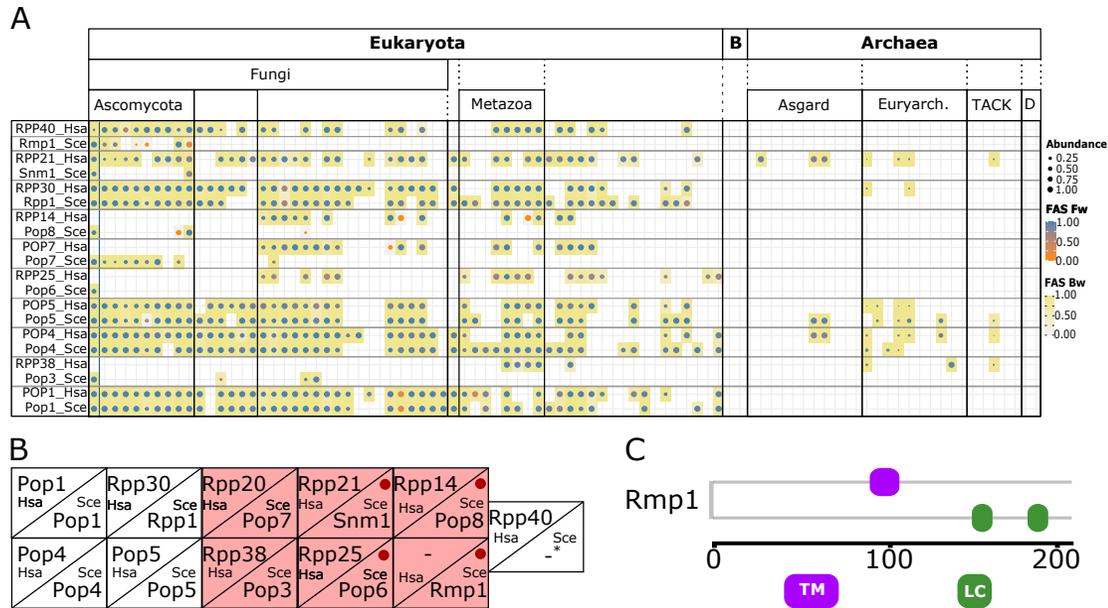


OG1 – core gene
OG2 – gene loss
OG3 – gene gain
OG4 – multiple independent gene loss (H1)
 / horizontal gene transfer (H2)/
 contamination (H3)

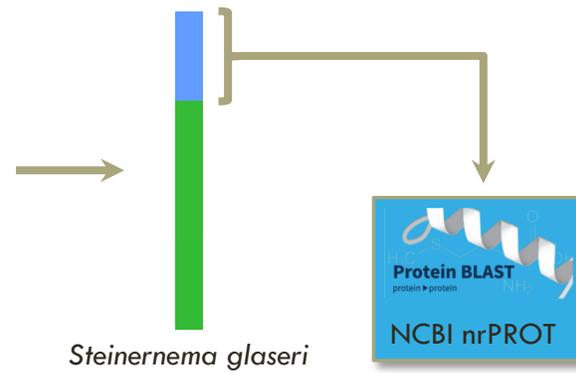
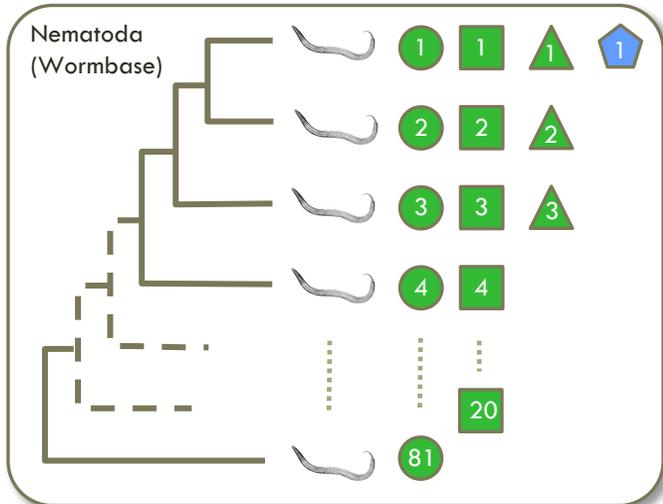
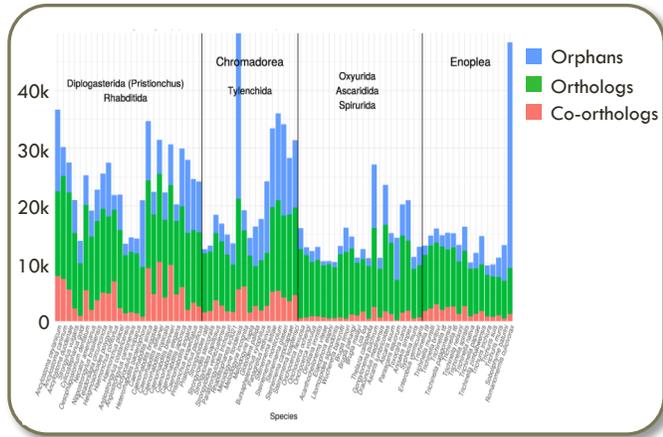
PHYLOGENETIC PROFILES OF 303 EUKARYOTIC RIBOSOME BIOGENESIS FACTORS ACROSS 900 SPECIES



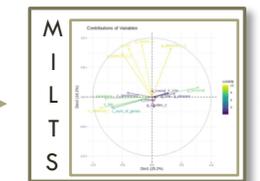
THE RNASE-MRP COMPLEX IN YEAST AND HUMANS – WHEN LIMITED TRACEABILITY CONFOUNDS THE ANALYSIS



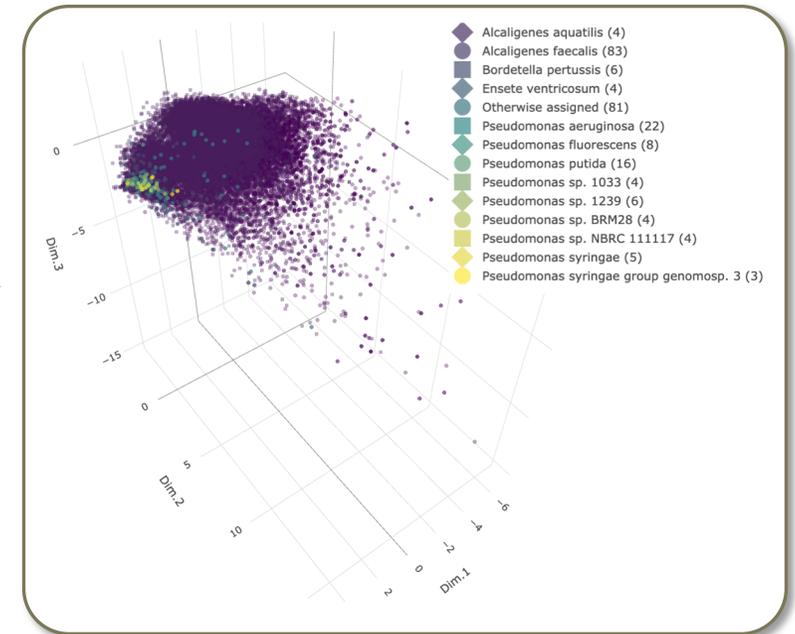
H3 — ARE CONTAMINATIONS IN GENE SETS AN ISSUE?



Taxonomic assignment



Gene set statistics



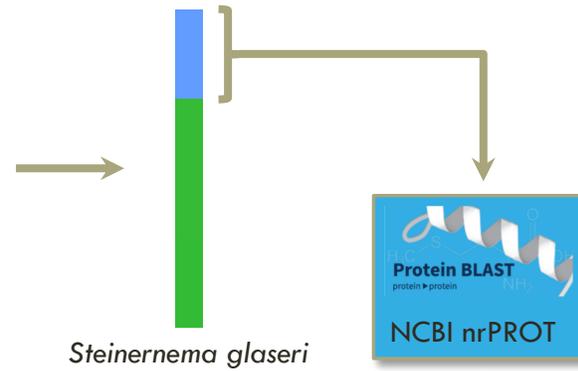
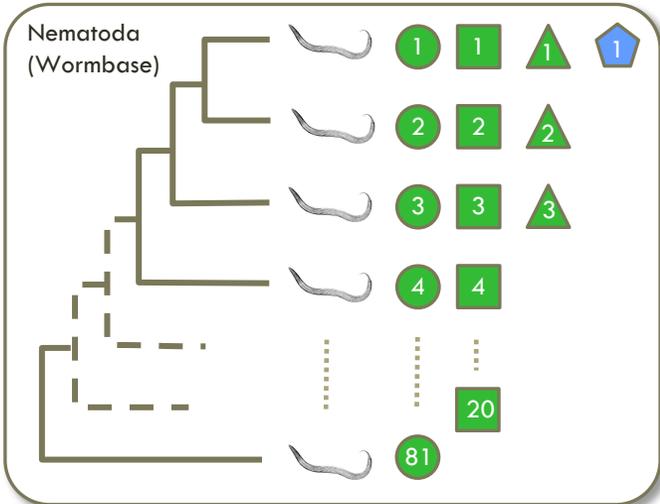
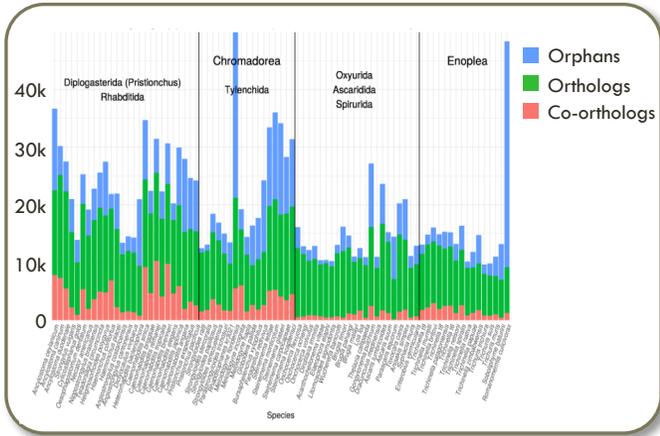
G-NOME - AN ONLINE RESOURCE FOR GENOMES AND GENE SETS -

The screenshot displays the G-NOME online resource dashboard. At the top, there is a navigation bar with the logo 'tbg translational biodiversity genomics' and menu items: 'Dashboard', 'Assemblies', 'Tools', and 'Downloads'. On the right side of the navigation bar, there are icons for a notification bell and a user profile. Below the navigation bar, the main content area is titled 'Dashboard' and contains five organism entries, each with a thumbnail image, a title, a status indicator, and a description of the assembly.

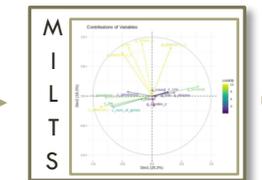
Organism	Assembly ID	Status	Description
<i>Trichinella nelsoni</i>	Tnel1	B, F, R, M	New BUSCO results by user1 on 04.04.2021
<i>Trichinella nelsoni</i>	Tnel1	B, F, R, M	New assembly by user1 on 04.04.2021
<i>Trichinella papuae</i>	Tpap1	B, F, R, M	New assembly by user1 on 04.04.2021
<i>Trachinus draco</i>	Tdra1	B, F, R, M	New assembly by user2 on 13.04.2021
<i>Trichinella patagoniensis</i>	Tpat1	B, F, R, M	New assembly by user1 on 04.04.2021

At the bottom of the dashboard, there is a footer bar with the following links: 'Contact', 'Impressum', and 'Help'.

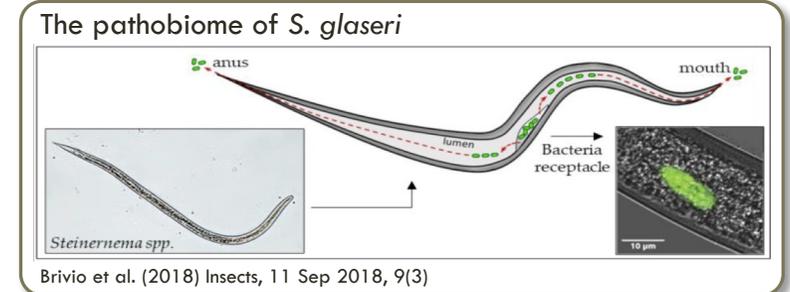
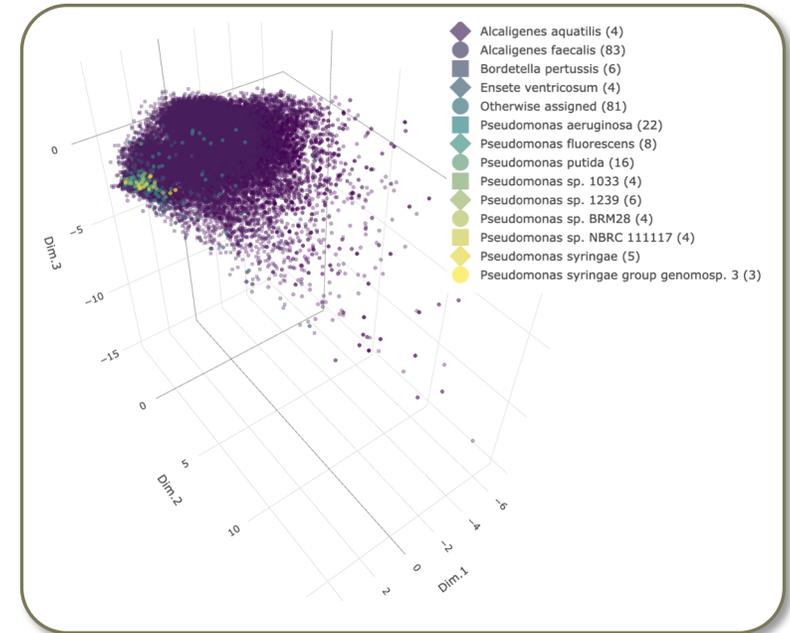
'SPECIES SPECIFIC' GENE SETS ARE OFTEN 'COMMUNITY SPECIFIC' GENE SETS



Taxonomic assignment



Gene set statistics



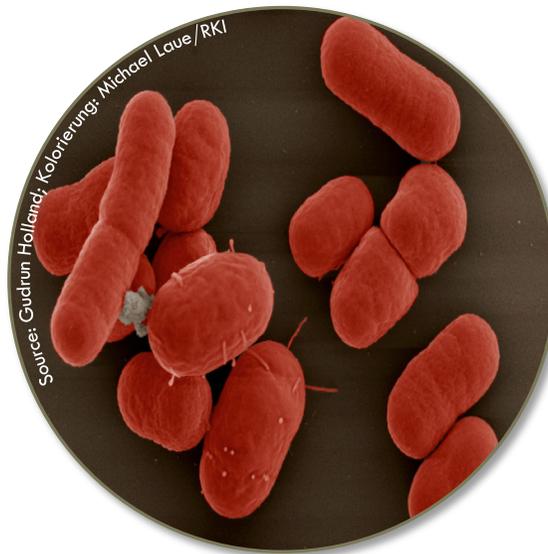
PART 2 – WHAT MAKES ACINETOBACTER BAUMANNII PATHOGENIC/VIRULENT?



WHAT CHARACTERIZES PATHOGENIC ACINETOBACTER?



Adhesion & biofilm formation



Source: Gudrun Holland; Kolorierung: Michael Loeu/RKI



Protection against...

- ROS/NOS
- Antimicrobials
- Desiccation
- ...

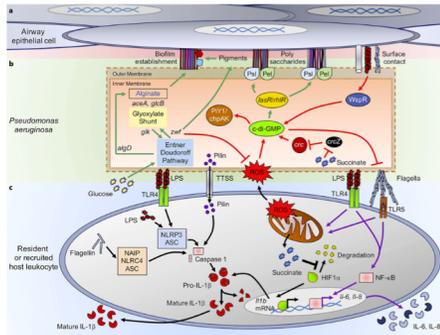


Nutrient acquisition



HOW TO CHARACTERIZE PATHOGENIC ACINETOBACTER?

Candidate



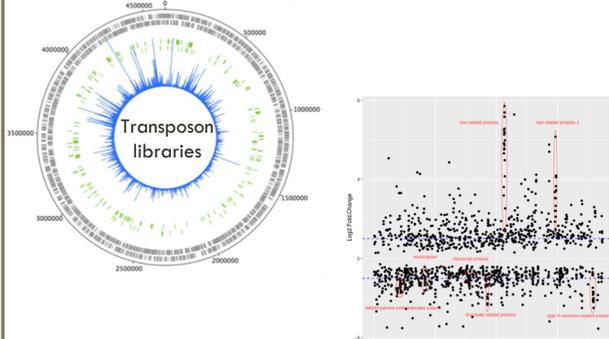
Comparison to other pathogens



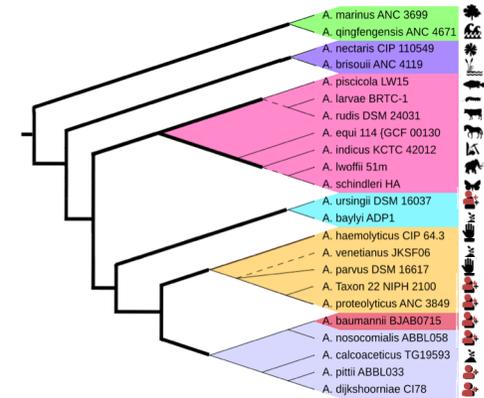
Mapping of virulence factors

Genome-wide screens

Random mutagenesis

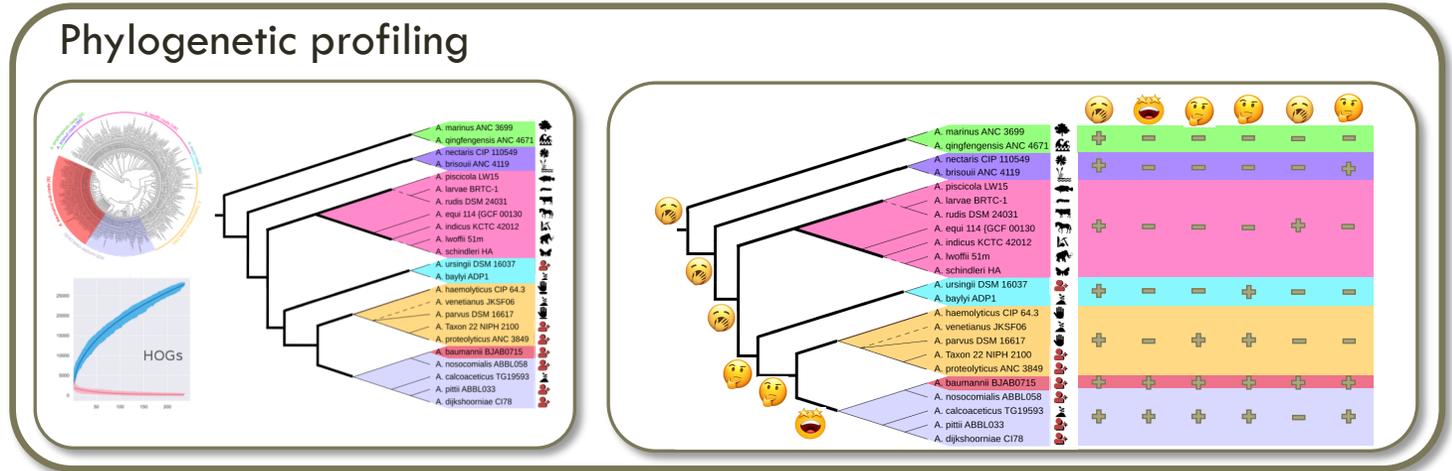
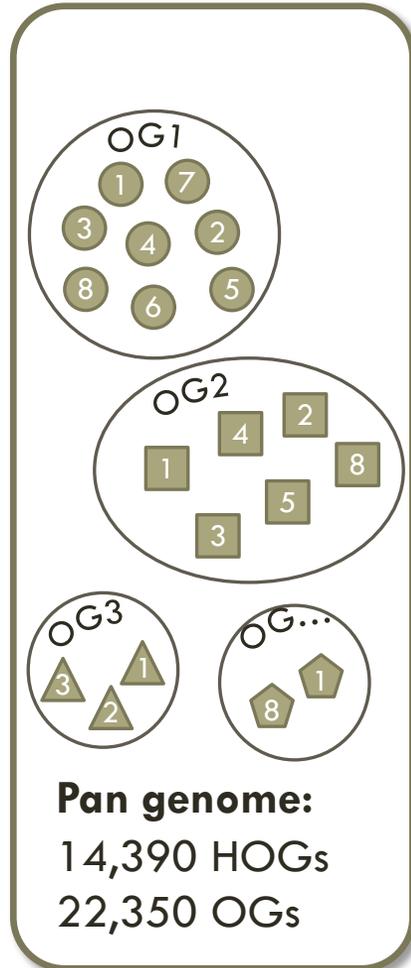
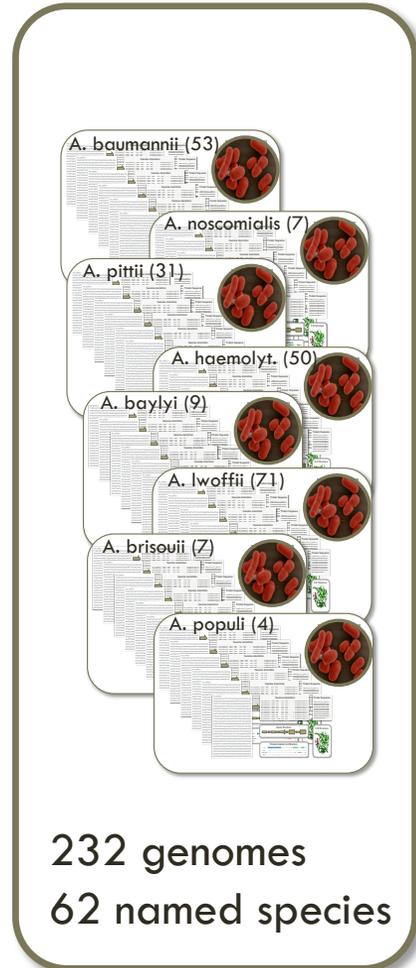


Diff. Gene expression

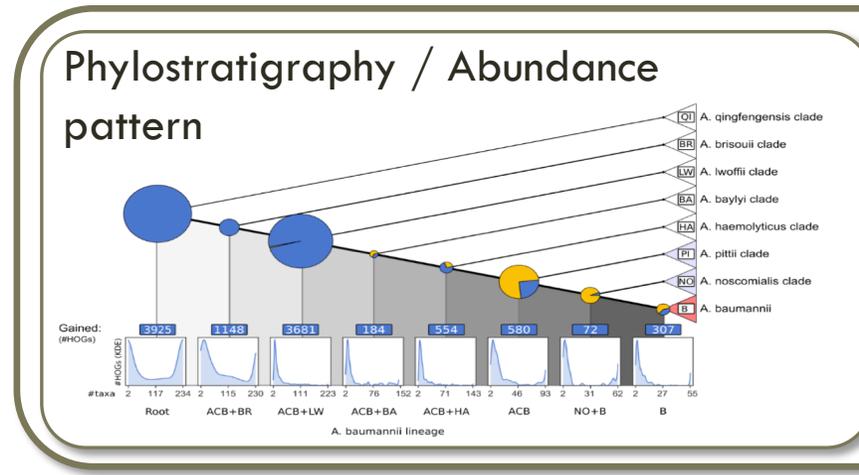
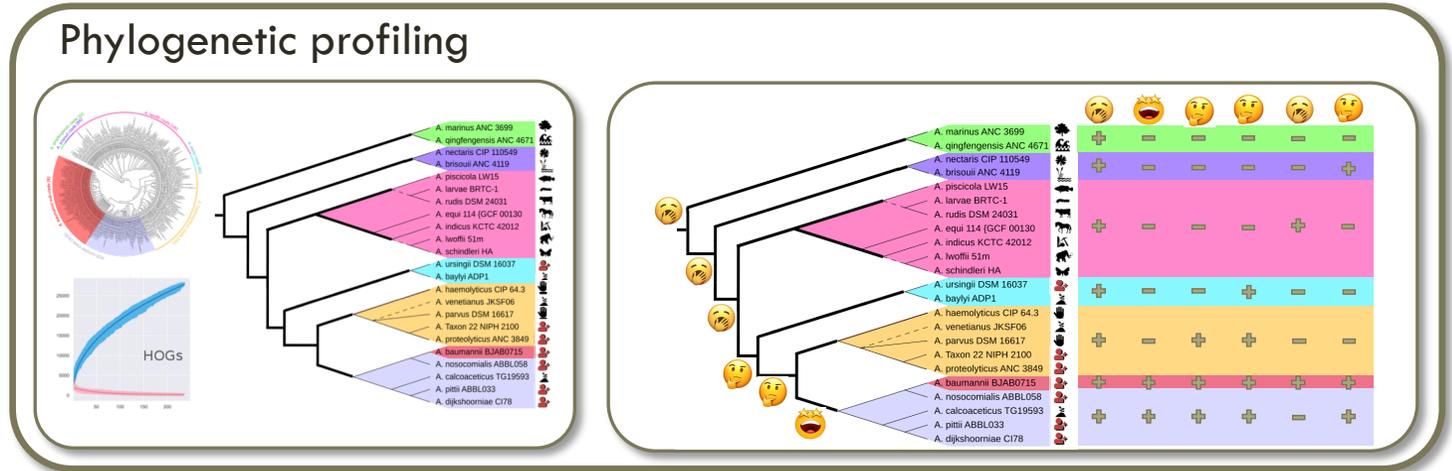
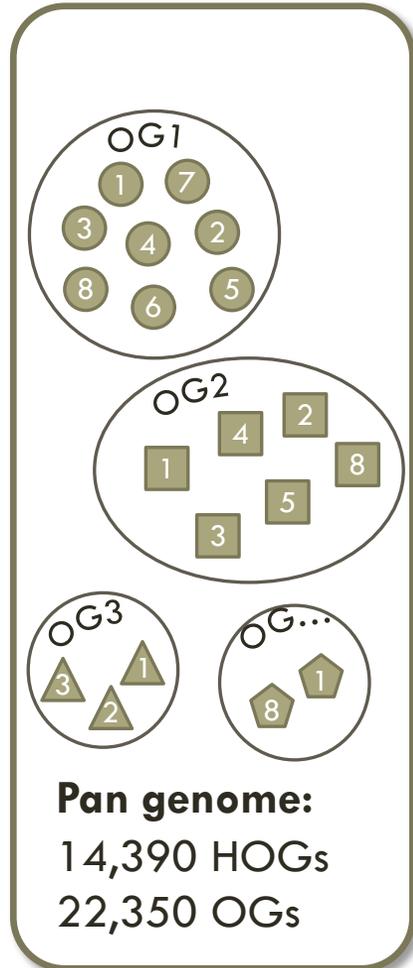
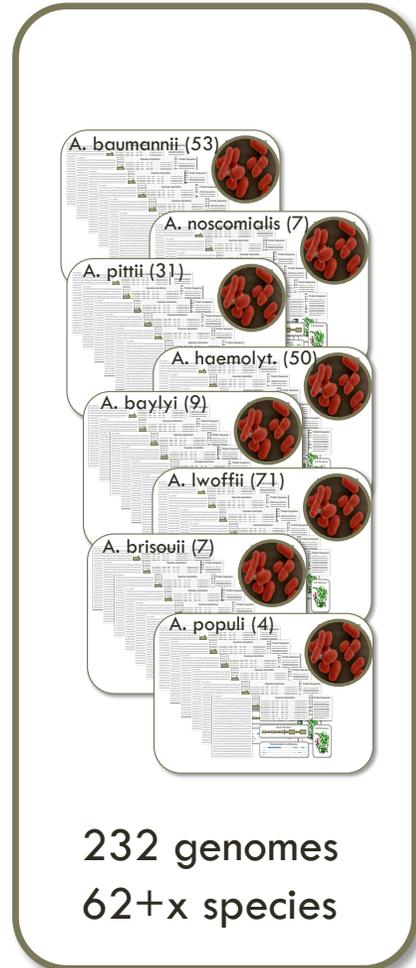


Phylogenetic Approaches

PHYLOGENETIC PROFILING THE ACINETOBACTER PAN GENOME USING ORTHOLOGS – PRIMING STAGE



PHYLOGENETIC PROFILING THE ACINETOBACTER PAN GENOME USING ORTHOLOGS – PRIMING STAGE



- ### GO overrepresentation analysis for each node
- C-PO(OH)₂ catabolism
 - Hexose metabolism
 - PQQ biosynthesis
 - L-glutamate transport
 - Amino acid transport
 - Fatty acid catabolism
 - Purin-P biosynthesis
 - oligopeptide transport
 - D-glucurate catabolism
 - Acetoin catabolism
 - Siderophore biosynthesis
 - Kdo2-lipid A biosynthesis
 - Organic acid phosphorylation
 - Viral release from host cells
 - **Quorum sensing (ACB)**

EXPLORE THE DATA VIA THE ACI-DASH (HTTPS://ACI-DASH.INGRESS.RANCHER.COMPUTATIONAL.BIO)

Companion Dashboard - Acinetobacter

Search/Select Genome:

- A. baumannii ATCC 19606 CIP 70.34 - GCF_000737145.1
- A. baumannii 25493_4 - GCF_000600815.1
- A. baumannii 25569_7 - GCF_000611385.1
- A. baumannii ATCC 19606 CIP 70.34 - GCF_000737145.1**
- A. baumannii AB030 - GCF_000746645.1
- A. baumannii A33405 - GCF_000757665.1
- A. baumannii AbH120-A2 - GCF_000761175.1

Companion Dashboard - Acinetobacter Comparative Genomics - v

Search/Select Genome:

A. baumannii ATCC 19606 CIP 70.34 - GCF_000737145.1

NCBI Taxonomy ID: 575584

Corrected Species:

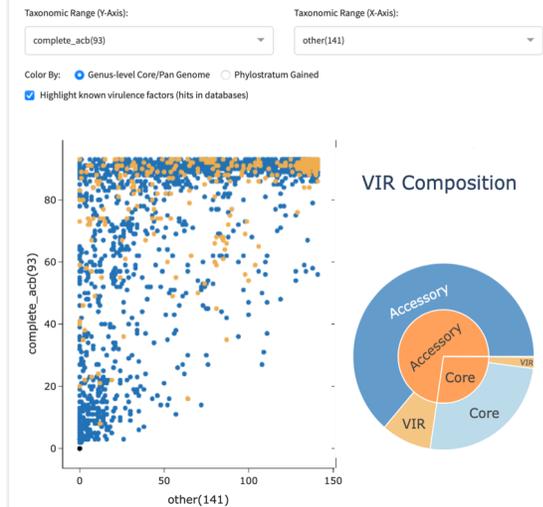
Sample Year: 1911

Isolated from: human

Publication: [Link](#)



Sampled at: unspecified place, Netherlands



EXPLORE THE DATA VIA THE ACI-DASH ([HTTPS://ACI-DASH.INGRESS.RANCHER.COMPUTATIONAL.BIO](https://aci-dash.ingress.rancher.computational.bio))

Companion Dashboard - Acinetobacter

Search/Select Genome:

- A. baumannii ATCC 19606 CIP 70.34 - GCF_000737145.1
- A. baumannii 25493_4 - GCF_000600815.1
- A. baumannii 25569_7 - GCF_000611385.1
- A. baumannii ATCC 19606 CIP 70.34 - GCF_000737145.1**
- A. baumannii AB030 - GCF_000746645.1
- A. baumannii A33405 - GCF_000757665.1
- A. baumannii AbH120-A2 - GCF_000761175.1

Companion Dashboard - Acinetobacter Comparative Genomics - v

Search/Select Genome:

A. baumannii ATCC 19606 CIP 70.34 - GCF_000737145.1

NCBI Taxonomy ID: 575584

Corrected Species:

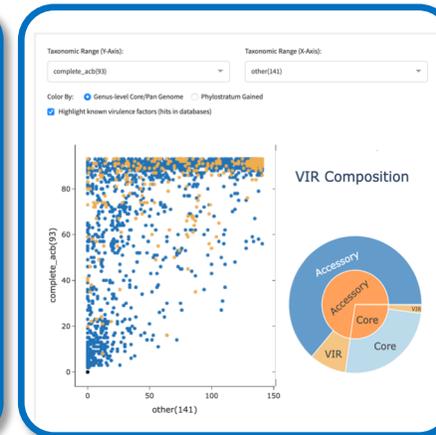
Sample Year: 1911

Isolated from: human

Publication: [Link](#)



Sampled at: unspecified place, Netherlands



NCBI RefSeq: [WP_001223275.1](#)

Predicted SCL class [GN]: IMB

This class % across HOG: 100

KEGG KO(s): [K02015](#)

KEGG Description(s): ABC.FEV.P; iron complex transport system permease protein

COG ID: [COG4605](#)

COG Letter: P

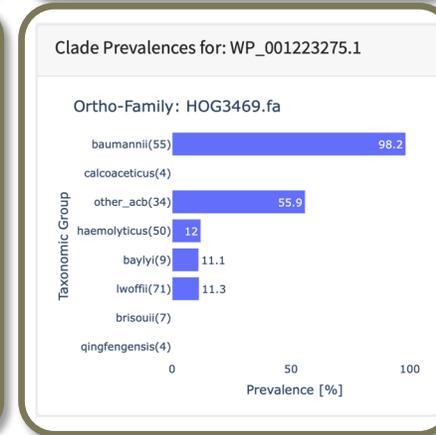
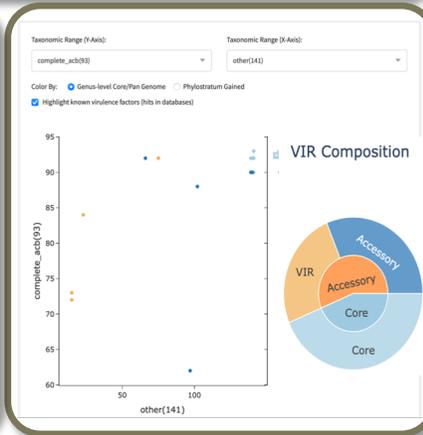
COG Description: ABC-type enterochelin transport system, permease component

VIR Factor Hit ID: [470.9257.peg.1700](#)

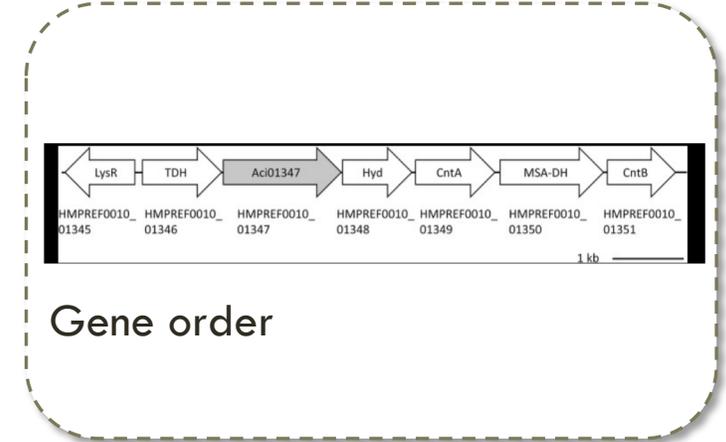
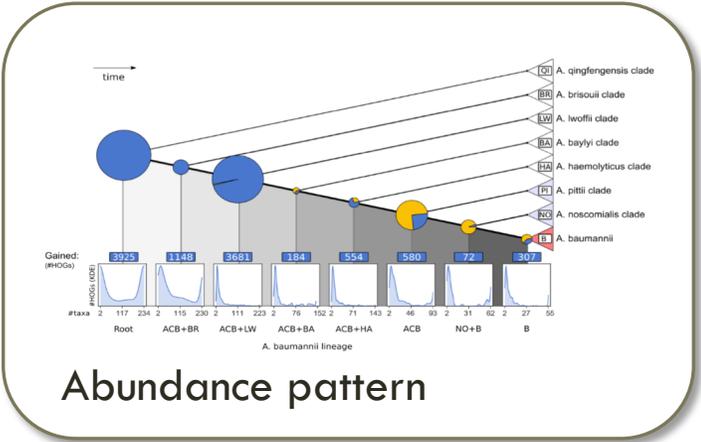
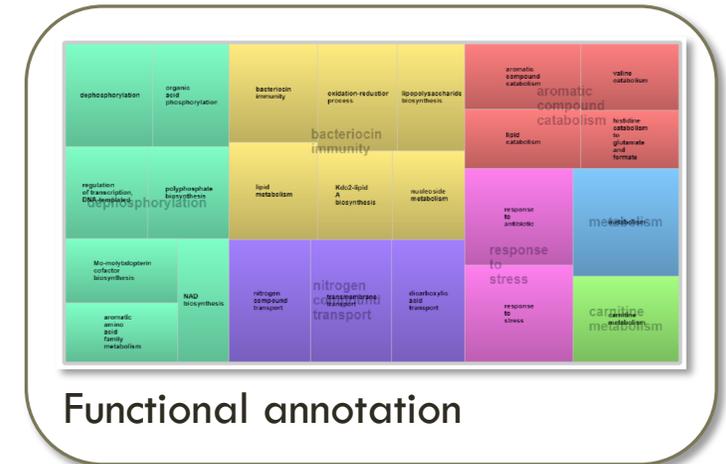
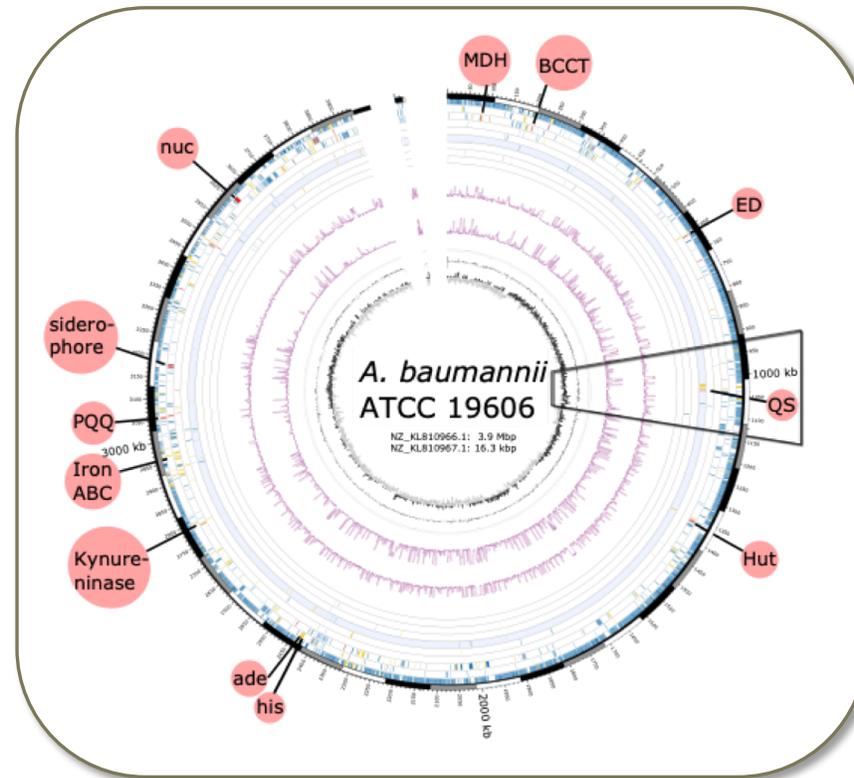
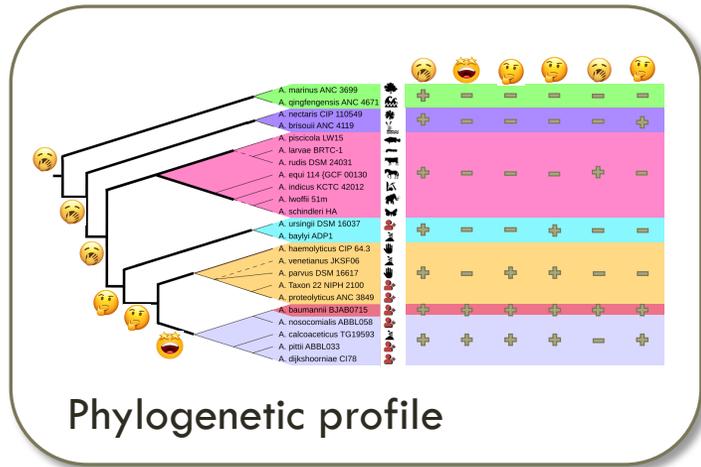
VIR evaluate: 2.8e-07

VIR Description: Ferric enterobactin transport system permease protein FepG (TC 3.A.1.14.2) [Acinetobacter baumannii strain aba_imp |

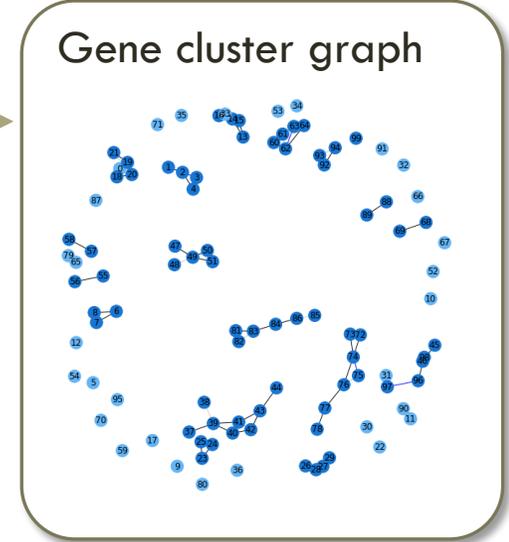
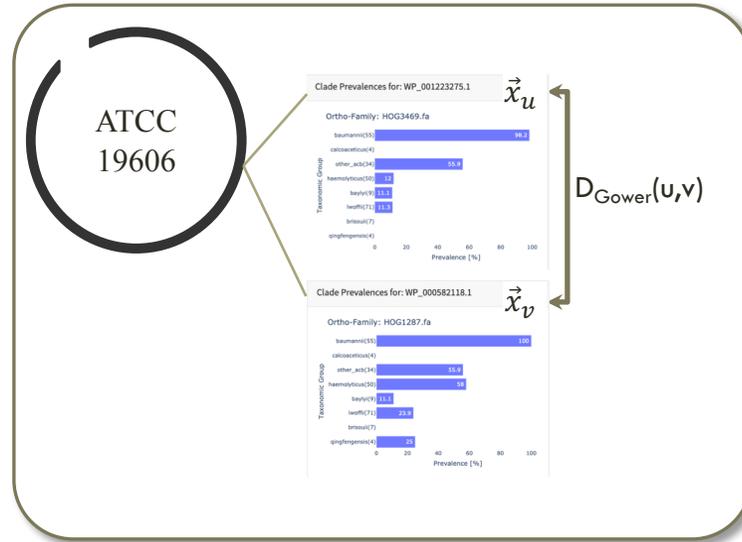
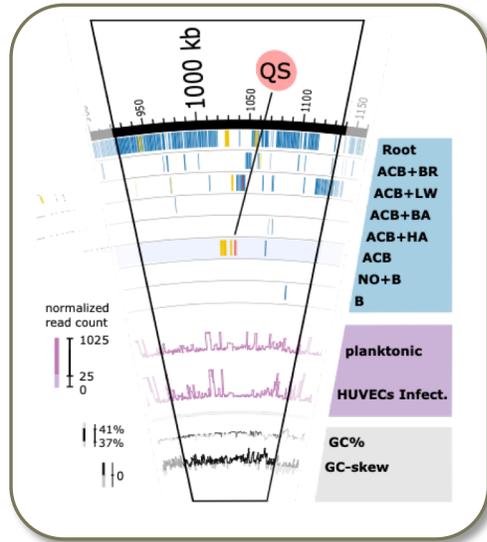
Genomic Acc	Start	End	Str	RefSeq Acc	Protein Annotation	Sym	Locus_tag	Len
filter data...					iron			
<input type="checkbox"/> NZ_KL810966.1	124349	124786	+	WP_001122847.1	ferric iron uptake transcriptional regulator	DJ41_RS04960		438
<input type="checkbox"/> NZ_KL810966.1	220700	221005	+	WP_000118514.1	non-heme iron oxygenase ferredoxin subunit	DJ41_RS05365		306
<input type="checkbox"/> NZ_KL810966.1	307134	307925	-	WP_001069886.1	iron-sulfur protein	DJ41_RS05755		792
<input type="checkbox"/> NZ_KL810966.1	417603	418832	-	WP_000118710.1	iron-sulfur cluster carrier protein ApbC	DJ41_RS06365		1230
<input type="checkbox"/> NZ_KL810966.1	916288	918141	-	WP_001278227.1	ferrous iron transporter B	DJ41_RS08690		1854
<input type="checkbox"/> NZ_KL810966.1	918138	918437	-	WP_000288742.1	ferrous iron transport protein A	DJ41_RS08695		300
<input type="checkbox"/> NZ_KL810966.1	1226177	1226512	+	WP_000993572.1	iron-sulfur cluster insertion protein ErpA	DJ41_RS10170		336



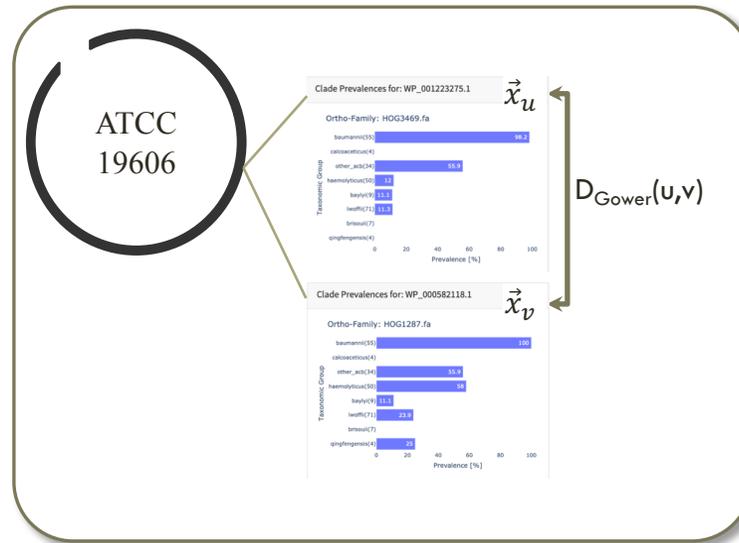
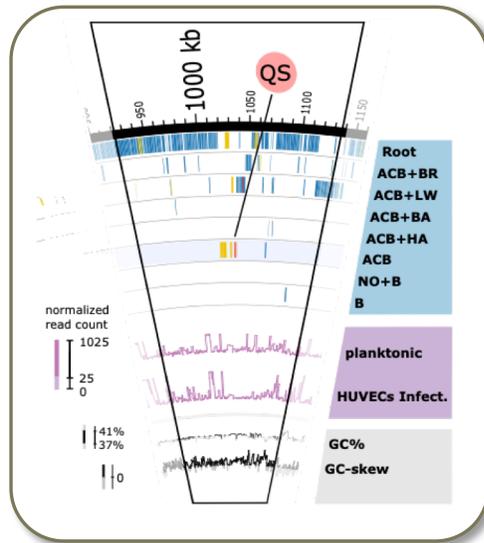
INNOVATION IN ONE CONTEXT: PROJECTION TO THE TYPE STRAIN – A. BAUMANNII ATCC 19606



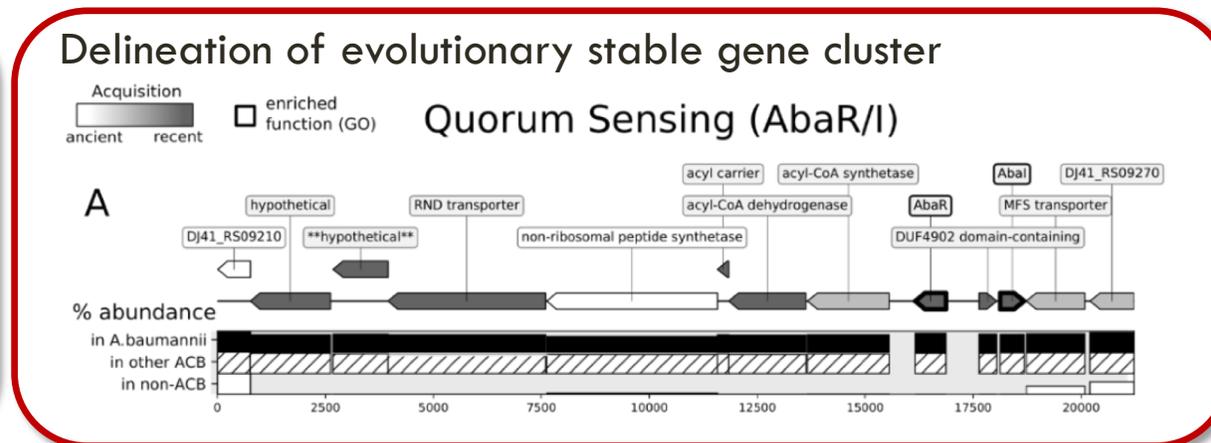
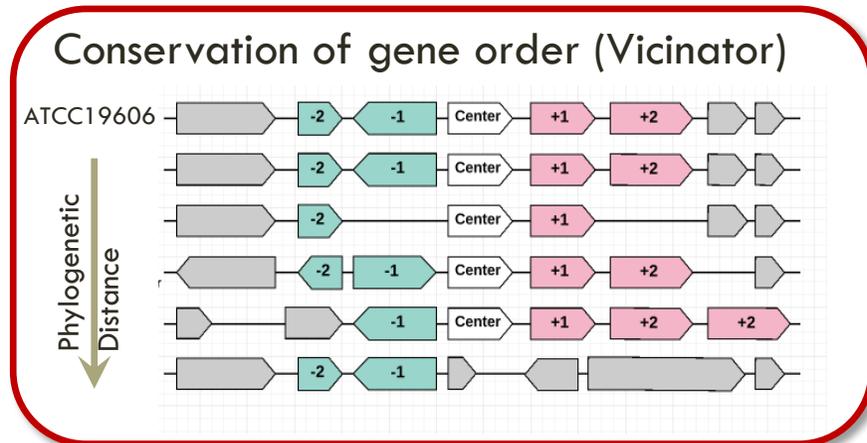
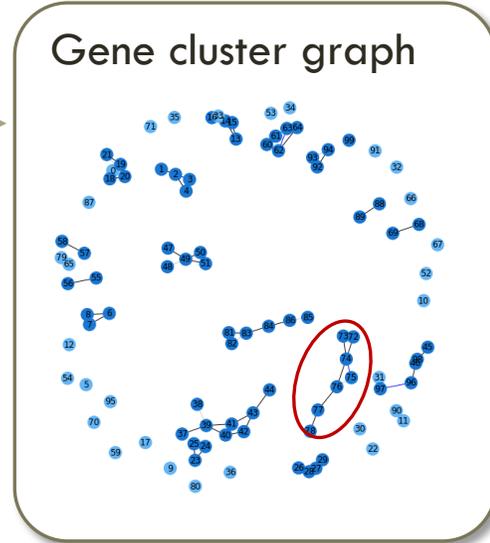
INNOVATION IN ONE CONTEXT: INFERENCE OF EVOLUTIONARY AND FUNCTIONAL(?!) UNITS



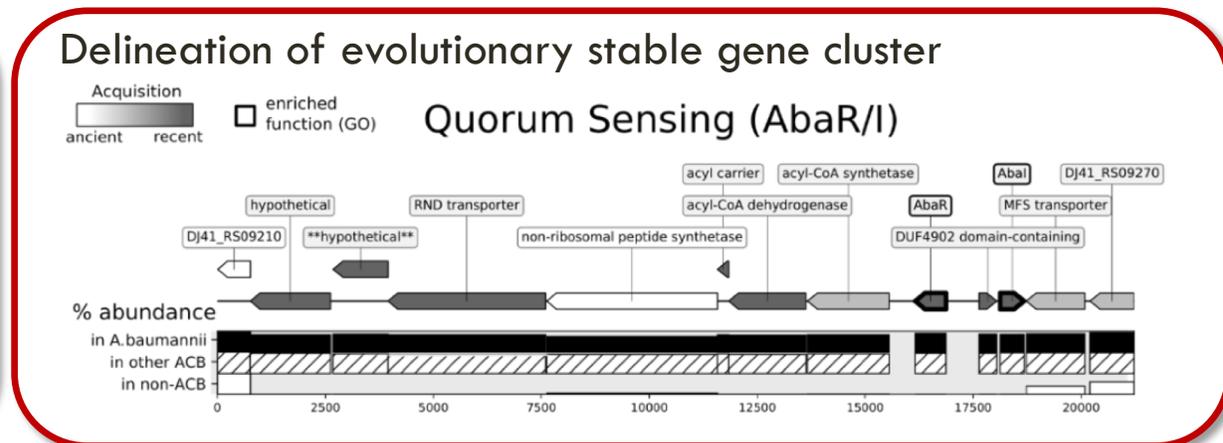
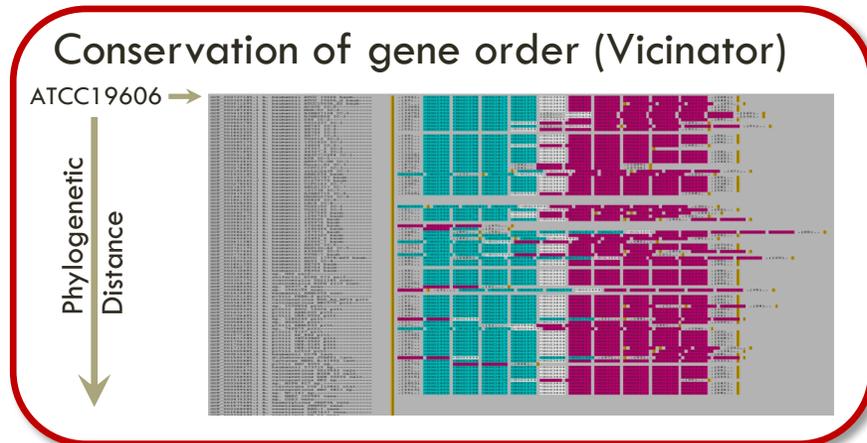
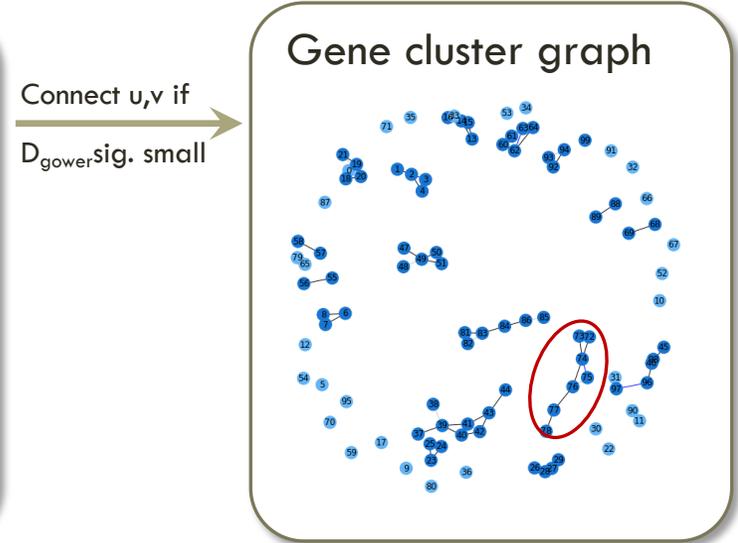
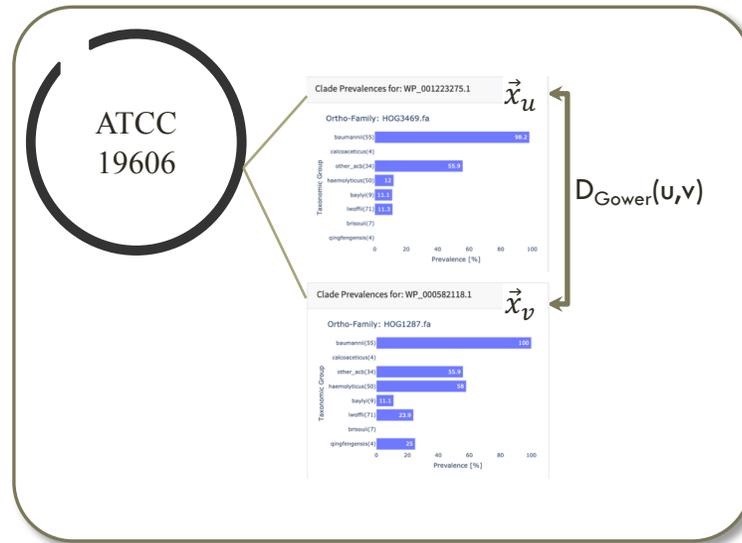
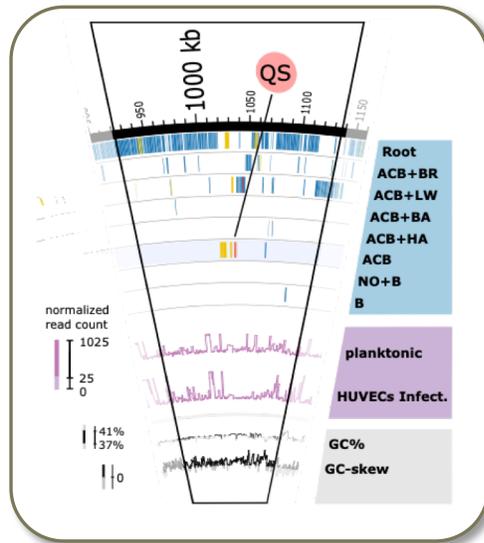
INNOVATION IN ONE CONTEXT: INFERENCE OF EVOLUTIONARY AND FUNCTIONAL(?!) UNITS



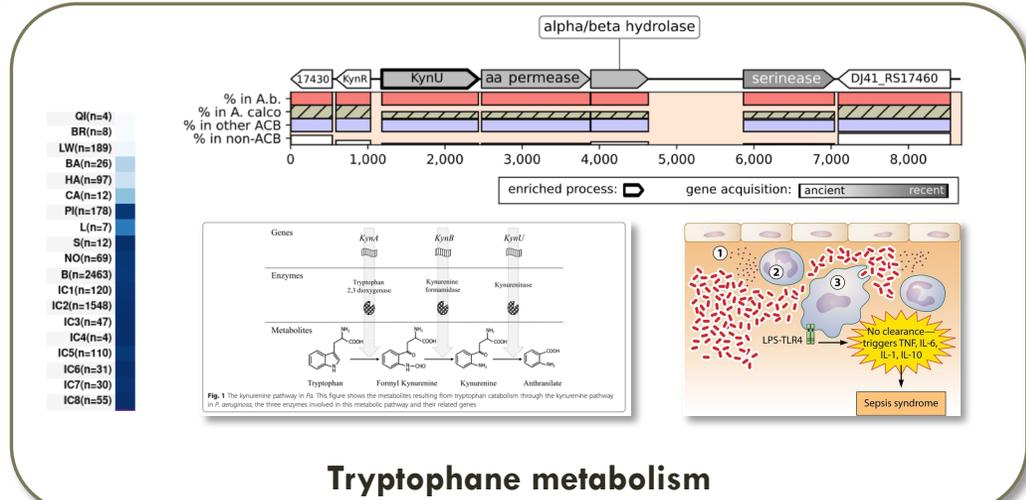
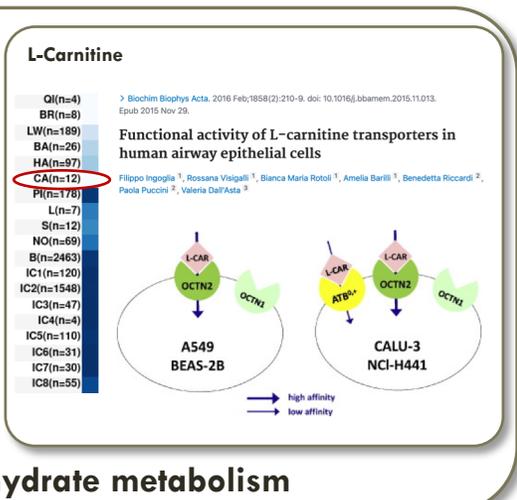
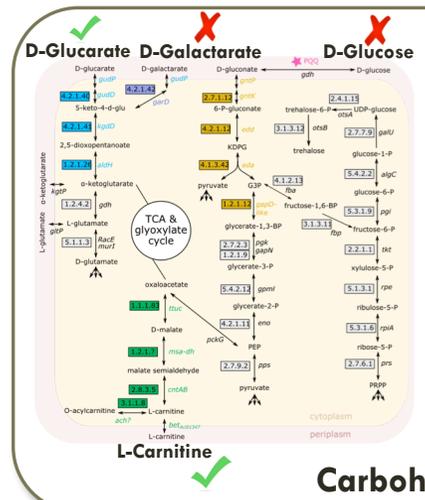
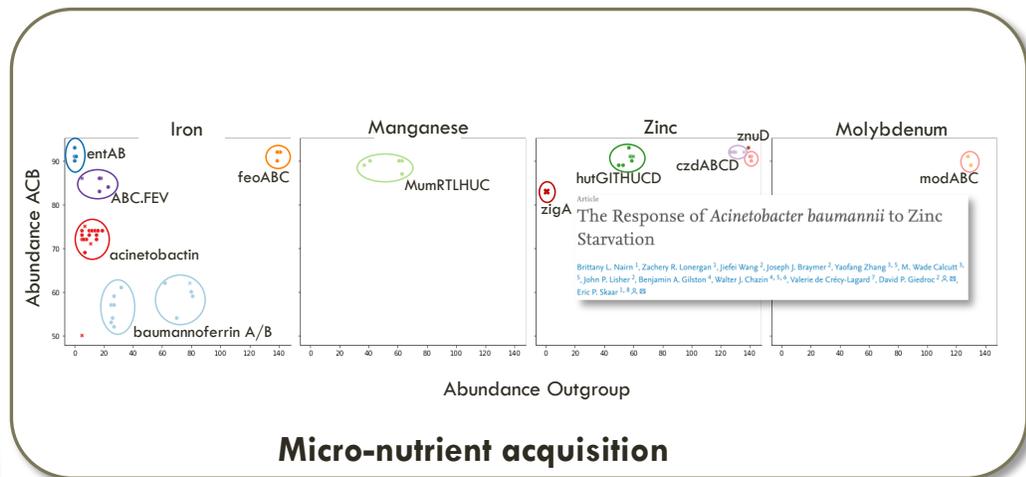
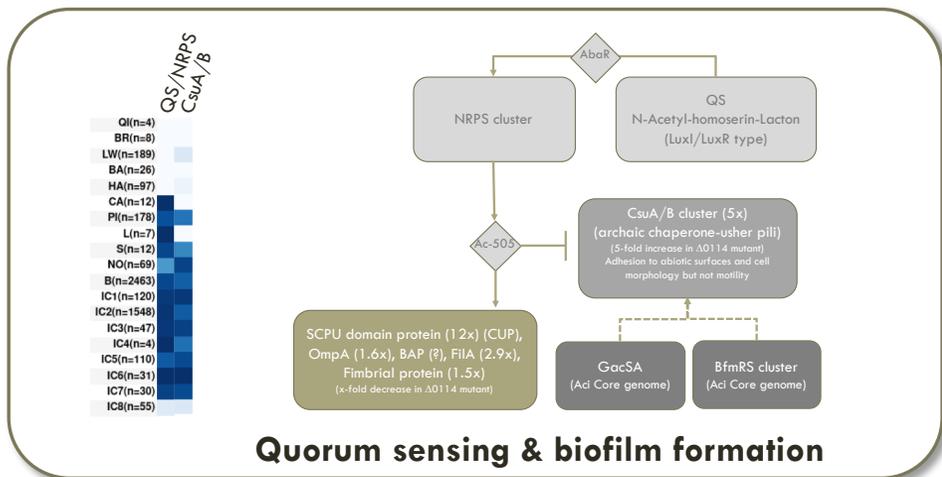
Connect u, v if $D_{Gower} \text{ sig. small}$



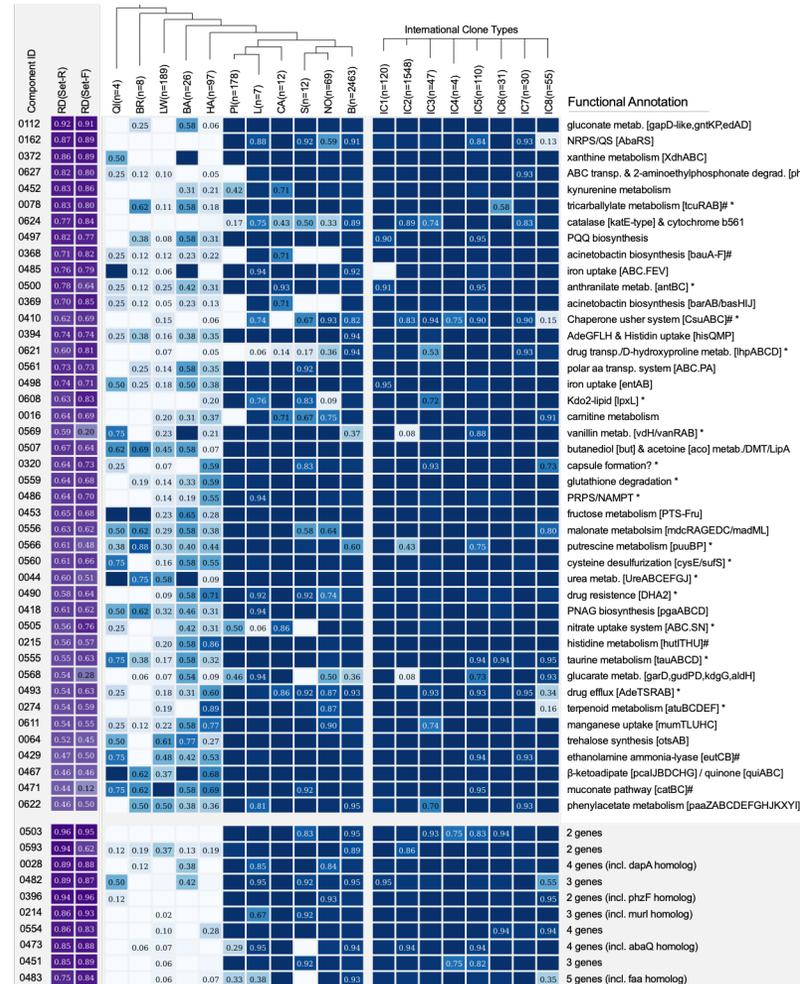
INNOVATION IN ONE CONTEXT: INFERENCE OF EVOLUTIONARY AND FUNCTIONAL(?!) UNITS



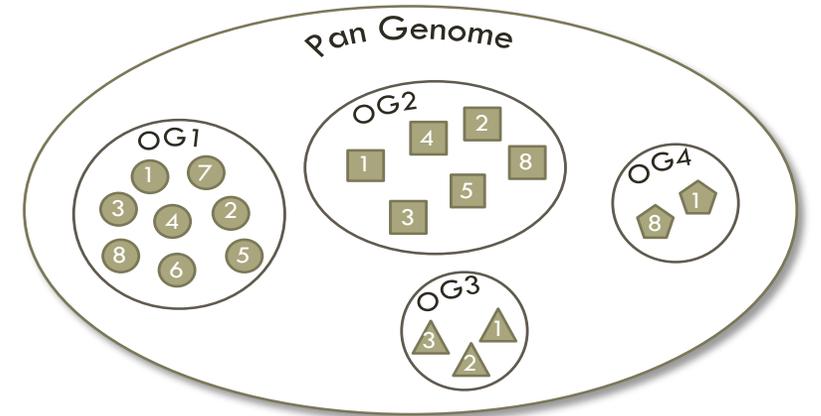
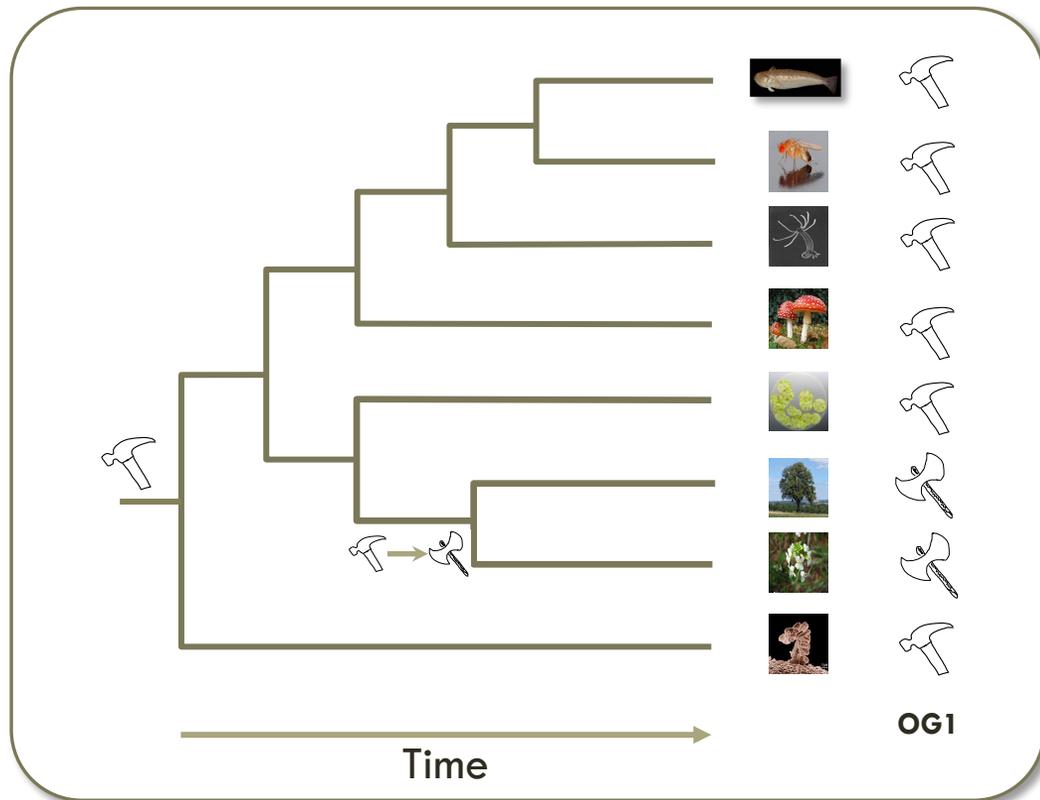
WHAT CHARACTERIZES PATHOGENIC ACINETOBACTER — AN EXCERPT



WHAT CHARACTERIZES PATHOGENIC ACINETOBACTER — MANY CLUSTERS ARE JUST A NUMBER...

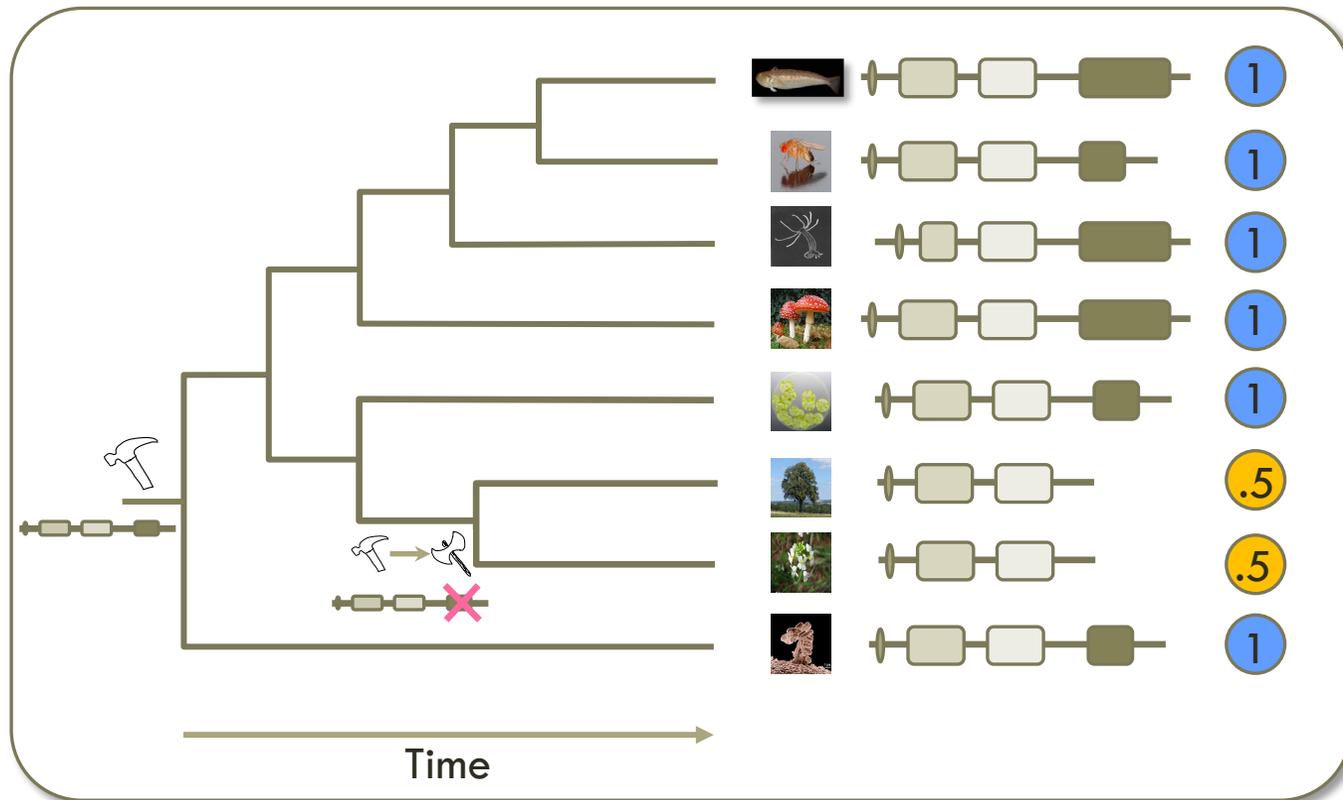


FUNCTIONAL DIVERGENCE OF ORTHOLOGOUS PROTEINS



H0 – Orthologous proteins are functionally equivalent

FEATURE ARCHITECTURES INFORM ABOUT FUNCTIONAL DIVERGENCE



Abundance Score

$$MS(S, O) = \sum_{i=1}^{N^S} (\omega_i * \min \left(\frac{N_i^S * N_i^O}{(N_i^S)^2}, 1 \right))$$

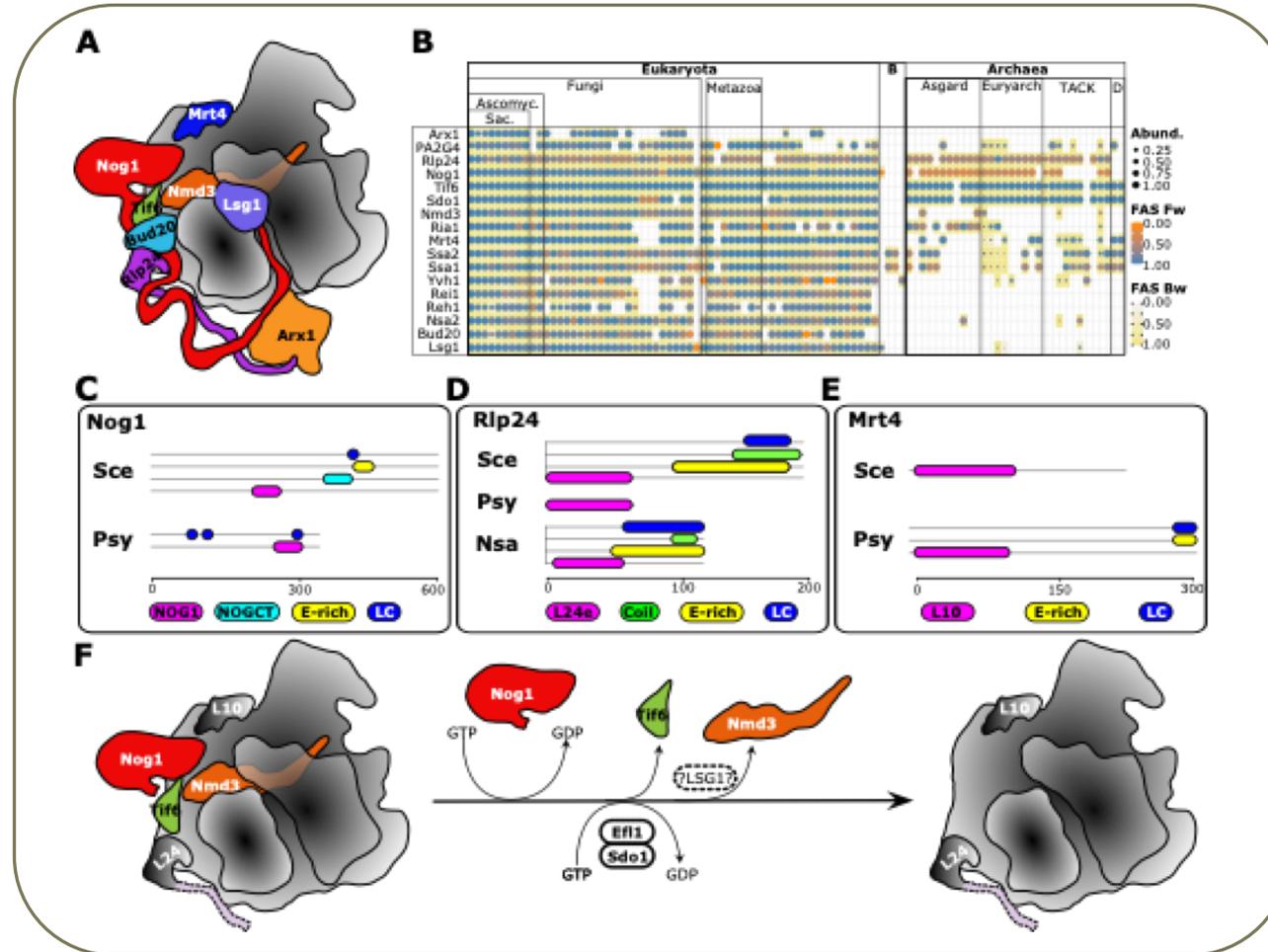
Positional Score

$$PS(S, O) = \sum_{i=1}^{N^S} \frac{\omega_i}{N_i^S} * \sum_{j=1}^{N_i^S} (1 - \min_{1 \leq l \leq N_i^O} |P_i^S, j - P_i^O, l|)$$

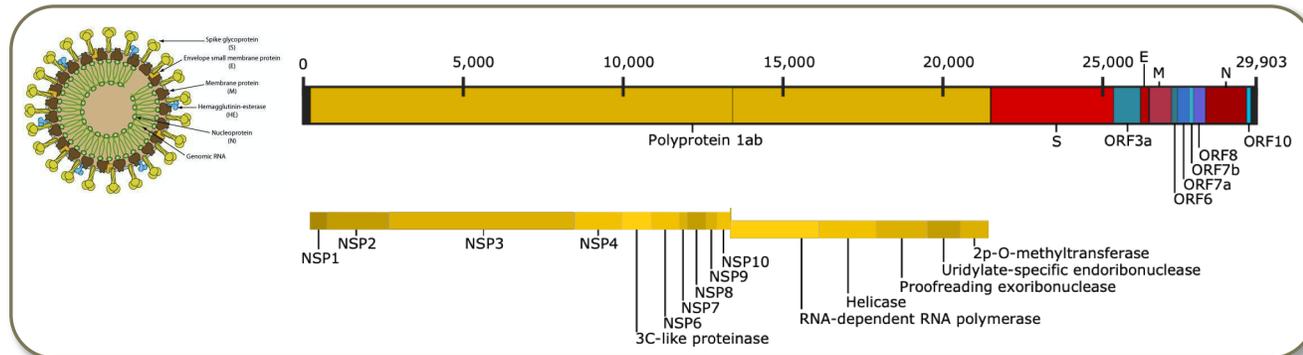
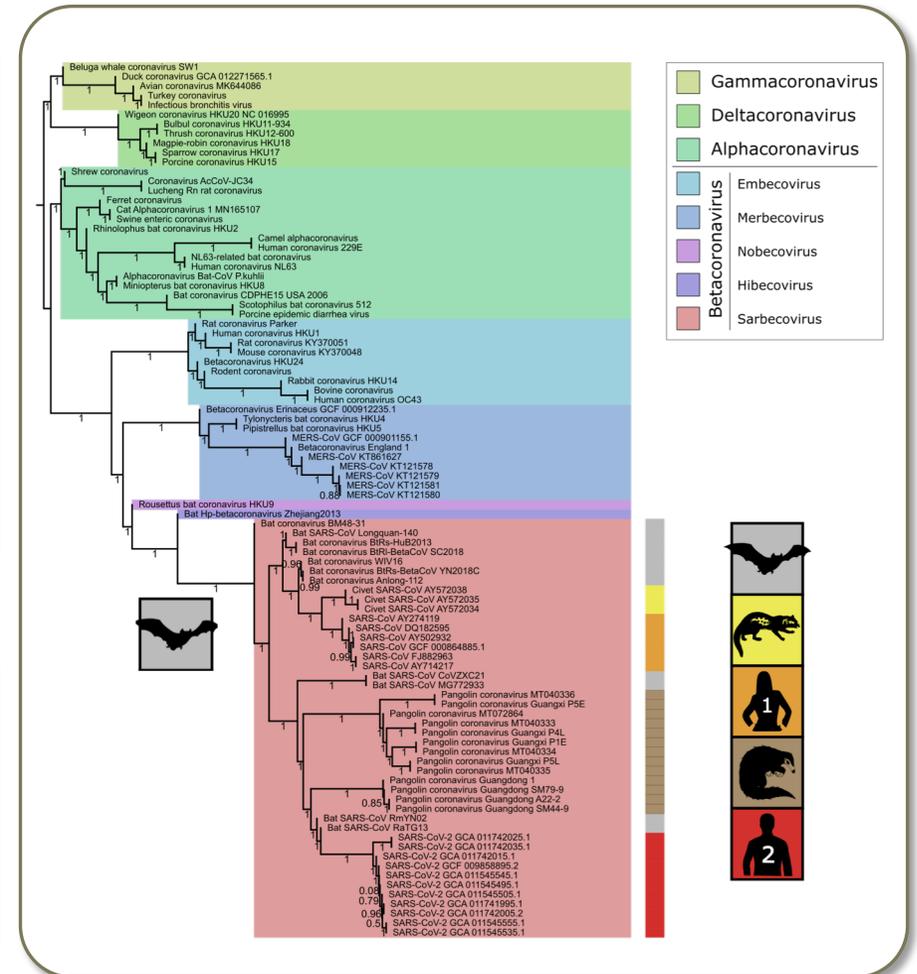
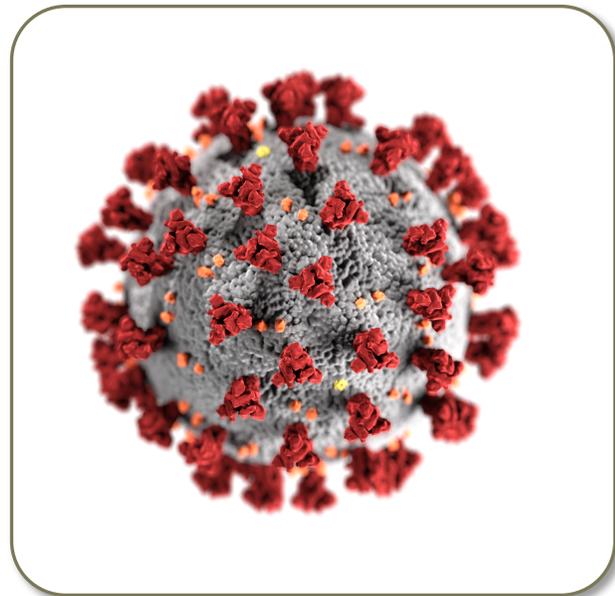
Feature Architecture Similarity Score [0,1]

$$FAS(S, O) = \alpha * MS + \beta * PS$$

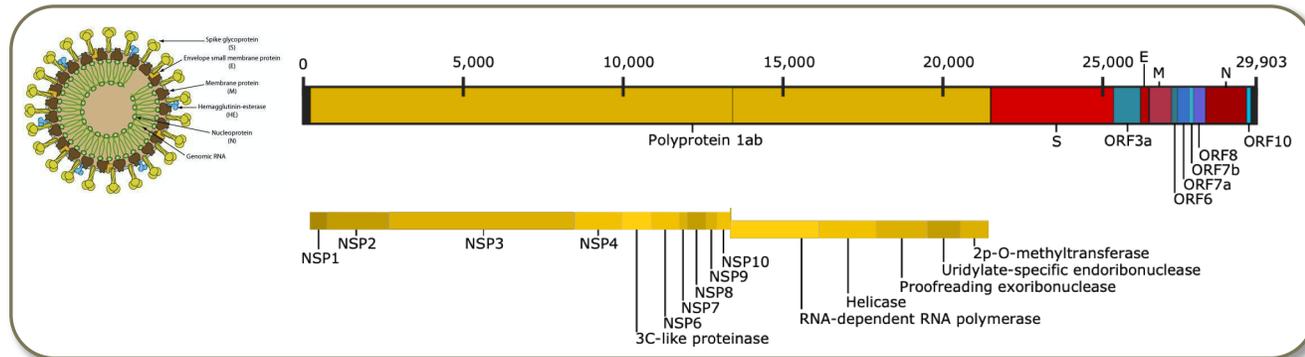
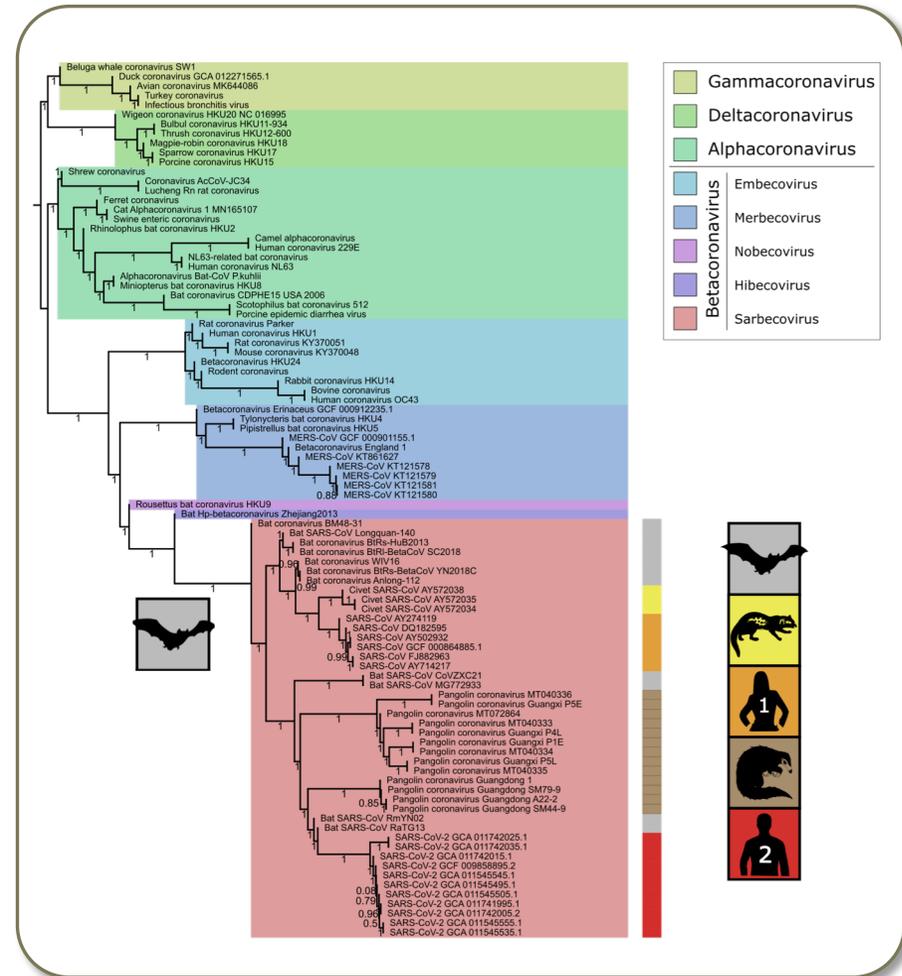
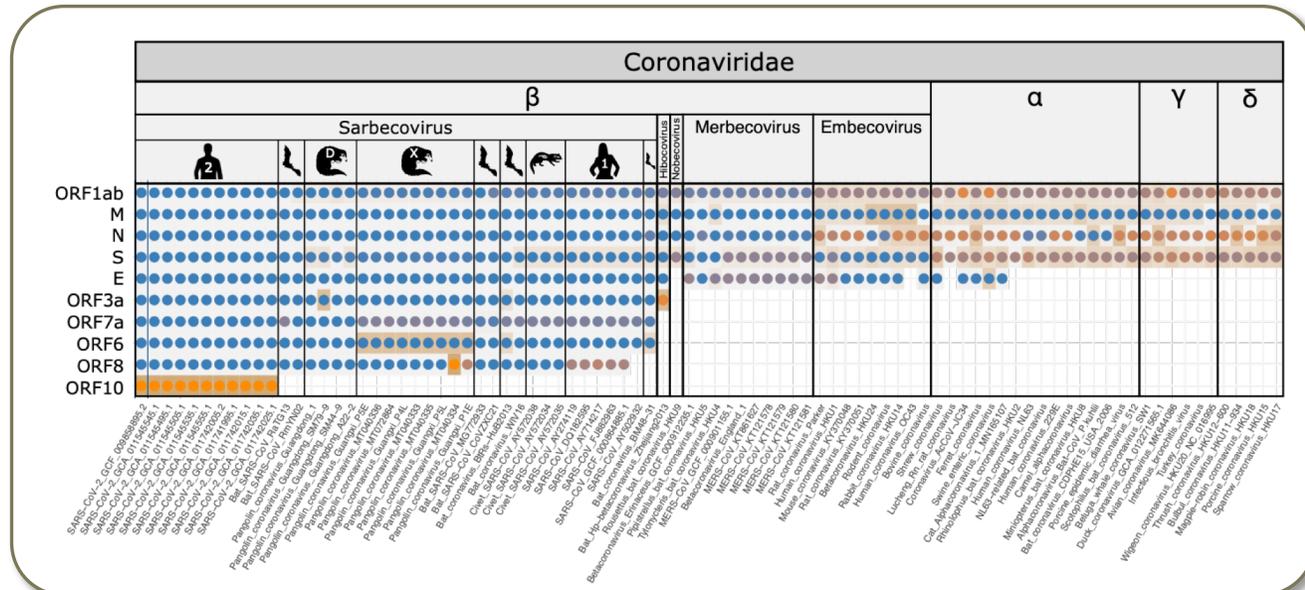
LARGE SUBUNIT MATURATION IN ARCHAEA AND EUKARYOTES – EVOLUTION OF COMPLEXITY



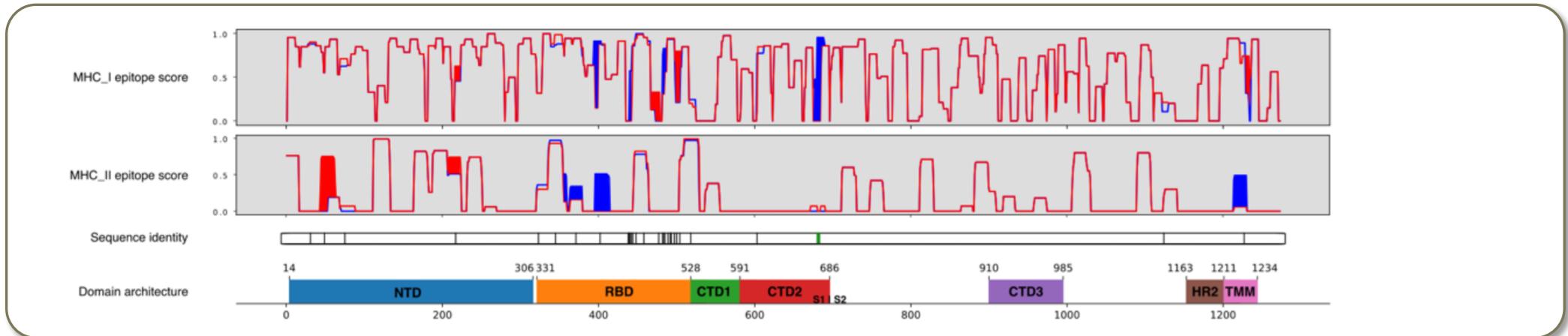
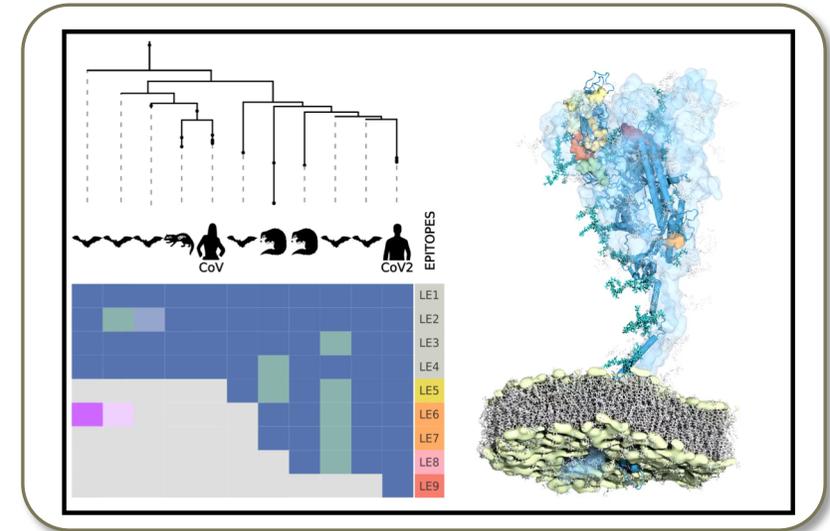
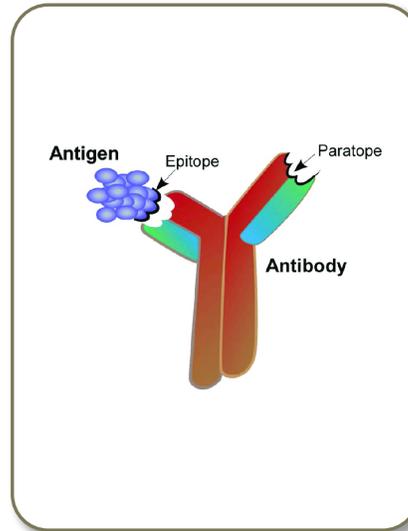
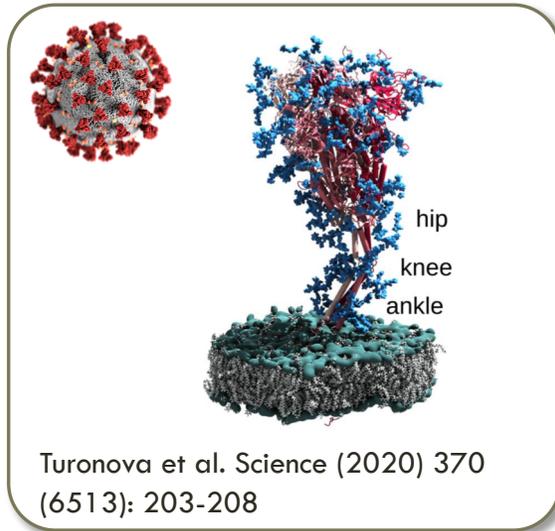
PART 3 – THE EVOLUTIONARY MAKING OF SARS-COV-2



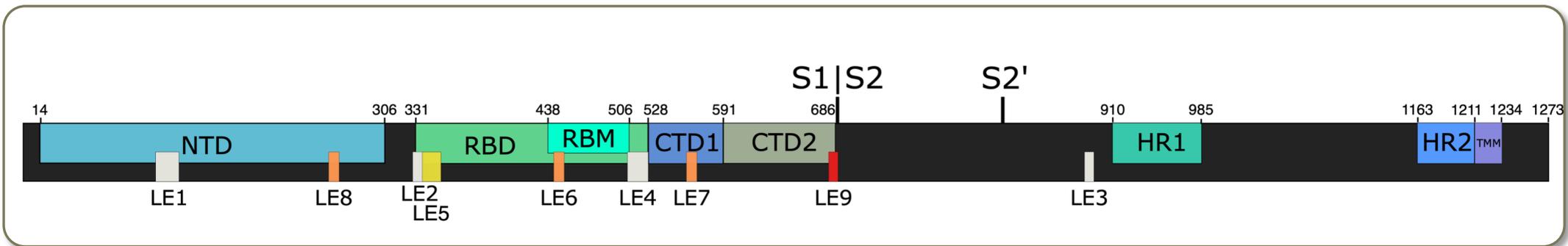
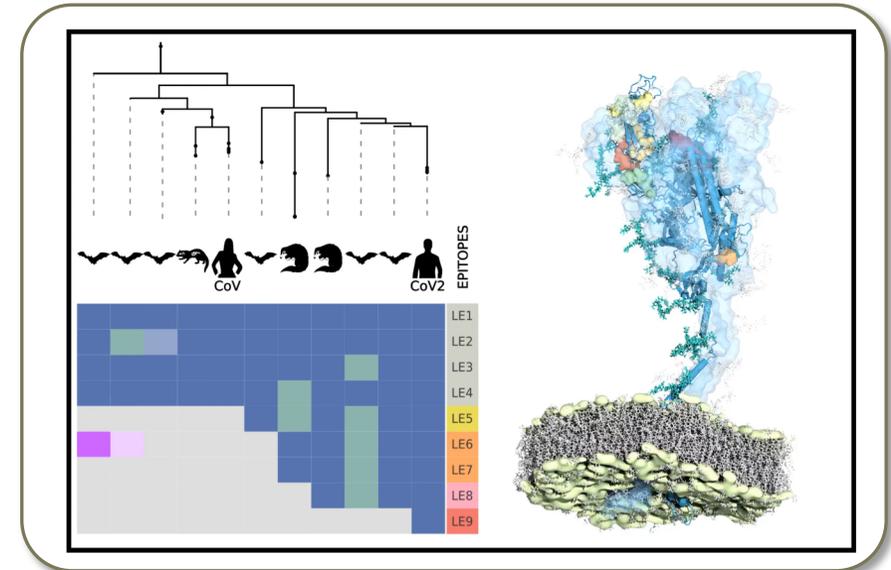
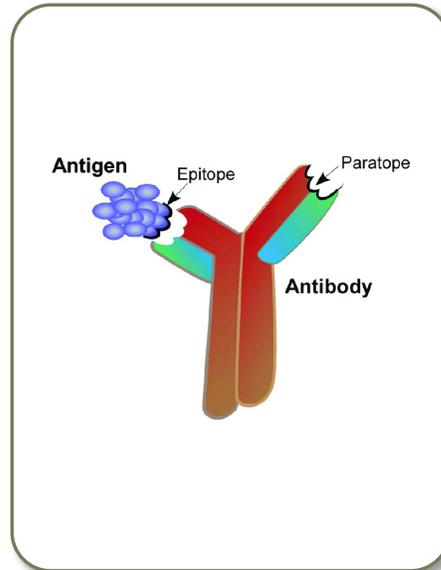
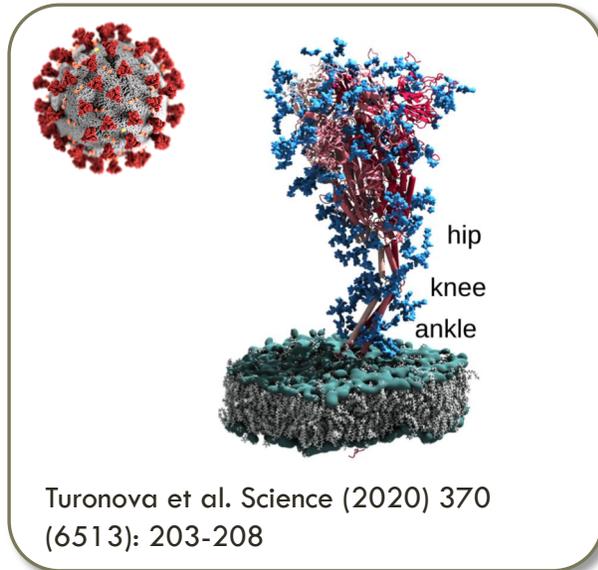
THE EVOLUTIONARY HISTORY OF SARS-COV-2 AND OF ITS PROTEOME



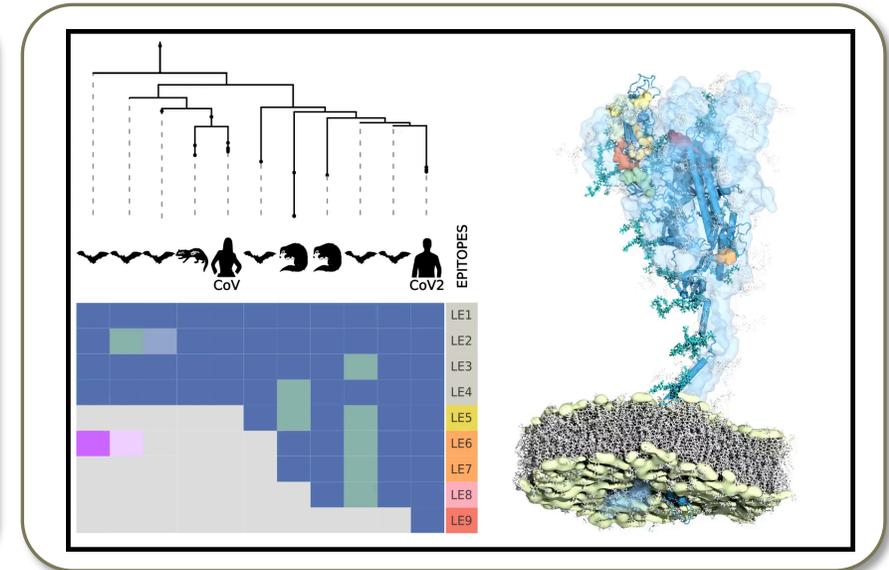
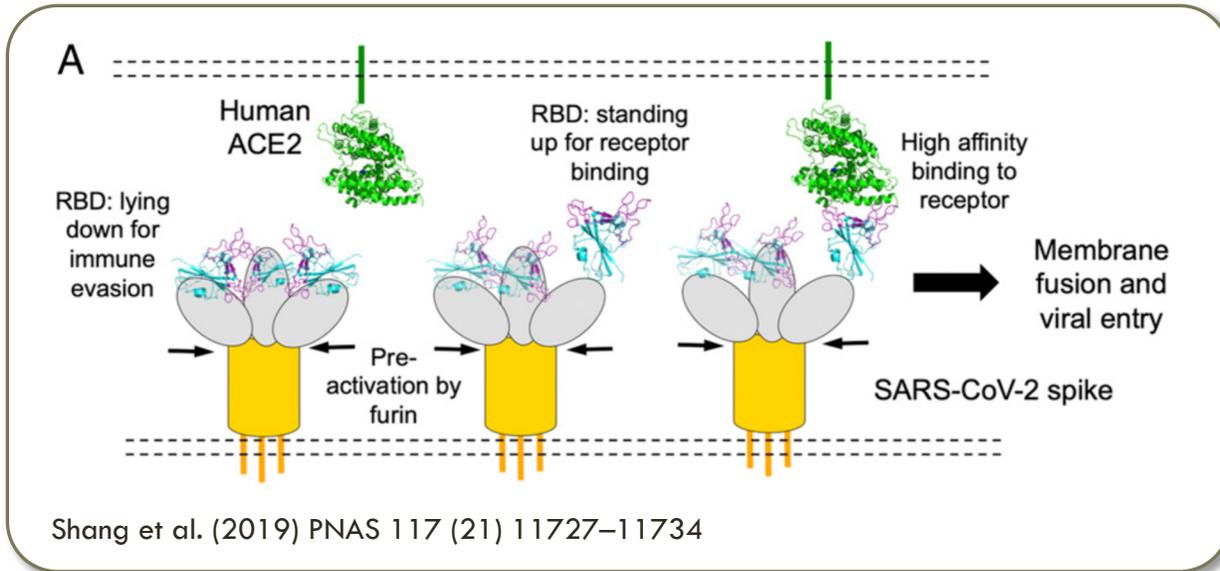
INCREASING RESOLUTION TO THE SUB-DOMAIN LEVEL — THE EPITOPE LANDSCAPE OF THE SPIKE



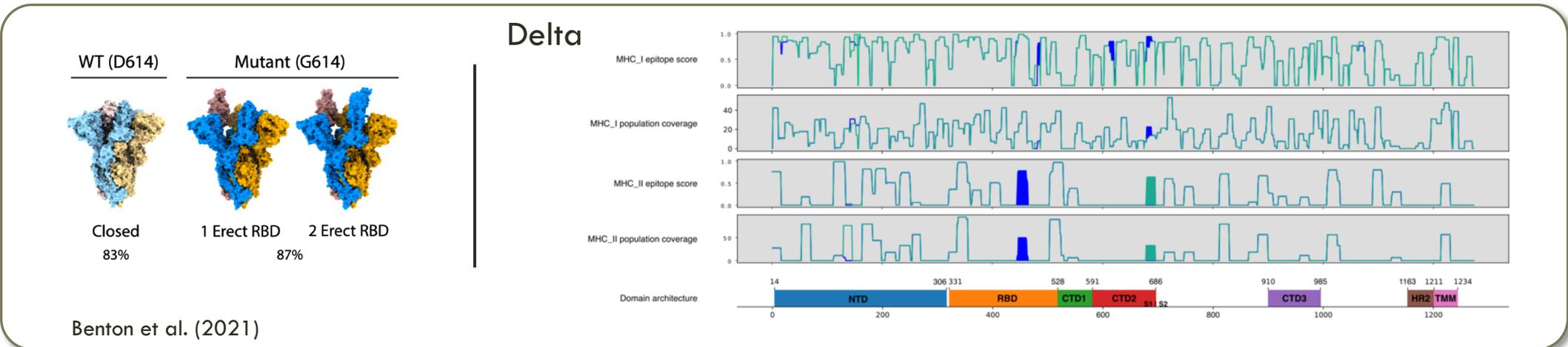
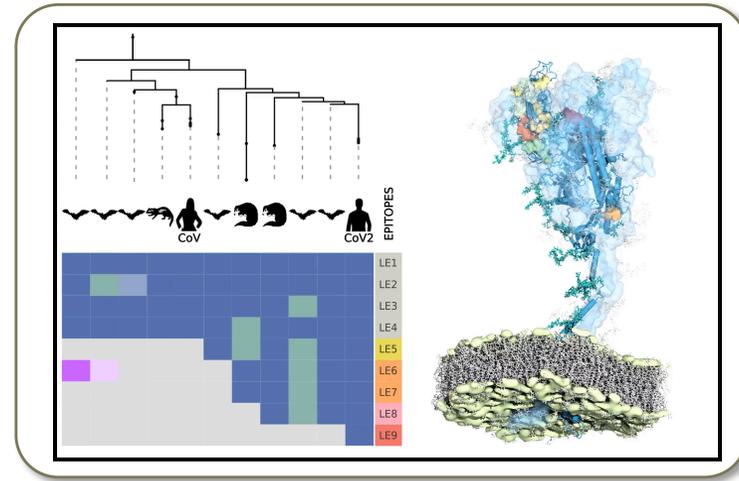
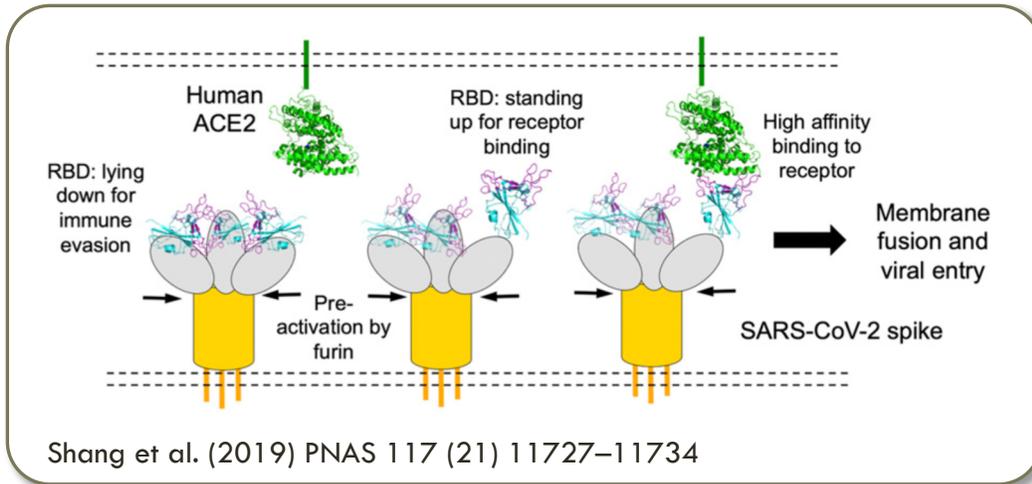
INCREASING RESOLUTION TO THE SUB-DOMAIN LEVEL — THE EPITOPE LANDSCAPE OF THE SPIKE



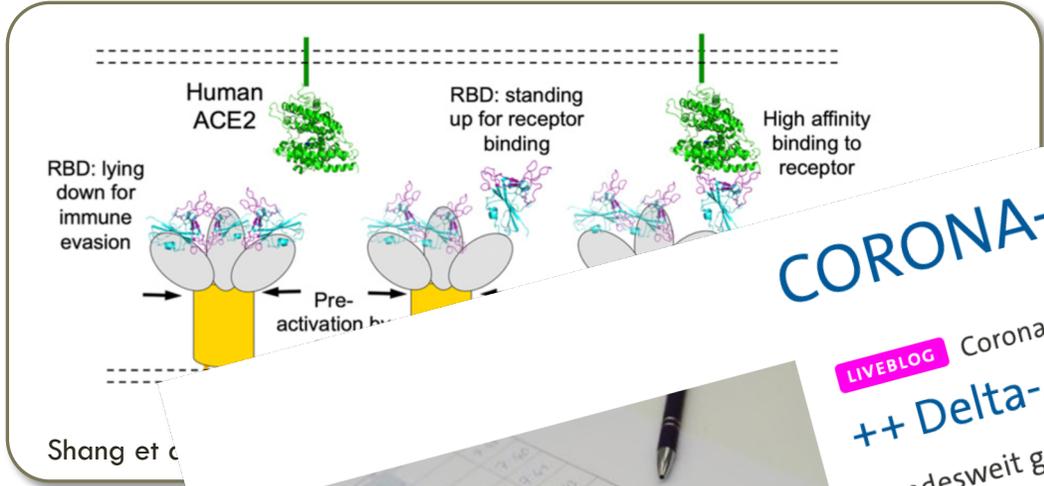
FURIN DEPENDENT DESTABILIZATION OF SPIKE A MASTER REGULATOR IN VIRUS – HOST INTERACTION



(FURIN DEPENDENT) DESTABILIZATION OF SPIKE EVOLUTIONARY FINE-TUNING SEEMS ONGOING



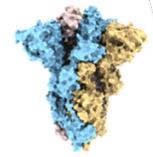
(FURIN DEPENDENT) DESTABILIZATION OF SPIKE EVOLUTIONARY FINE-TUNING SEEMS ONGOING



CORONA-PANDEMIE

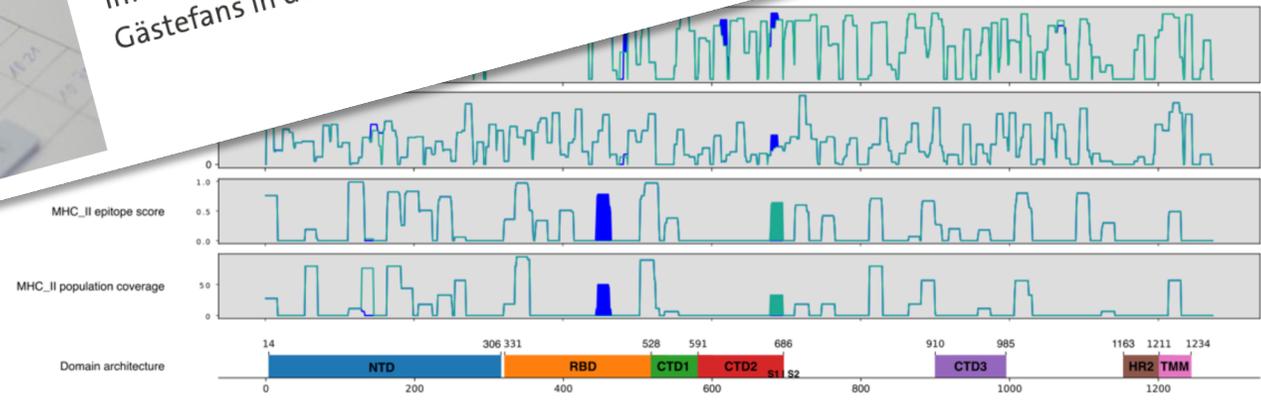
LIVEBLOG Coronavirus-Pandemie
++ Delta-Anteil steigt auf 74 Prozent ++
 Bundesweit gehen laut Robert Koch-Institut mittlerweile 74 Prozent aller Corona-Infektionen auf die Delta-Variante zurück. In der nächsten Bundesliga-Saison sind Gästefans in den Stadien zugelassen. Die Entwicklungen im Liveblog zum Nachlesen.

WT (D614G)



Closed
83%

Benton et al. (2021)



WIR SUCHEN UNTERSTÜTZUNG AUF ALLEN EBENEN

