

# GOBI 2025



## Investigating the Roots of Kleptoplasty

# Our Team



Matthäus



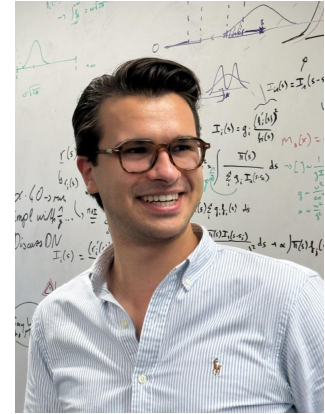
Lara



Rushali



Annika

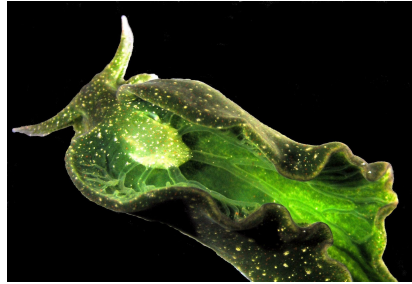


Max

# Our Organisms: Sea Slugs



Elysia Marginata



Elysia Chlorotica



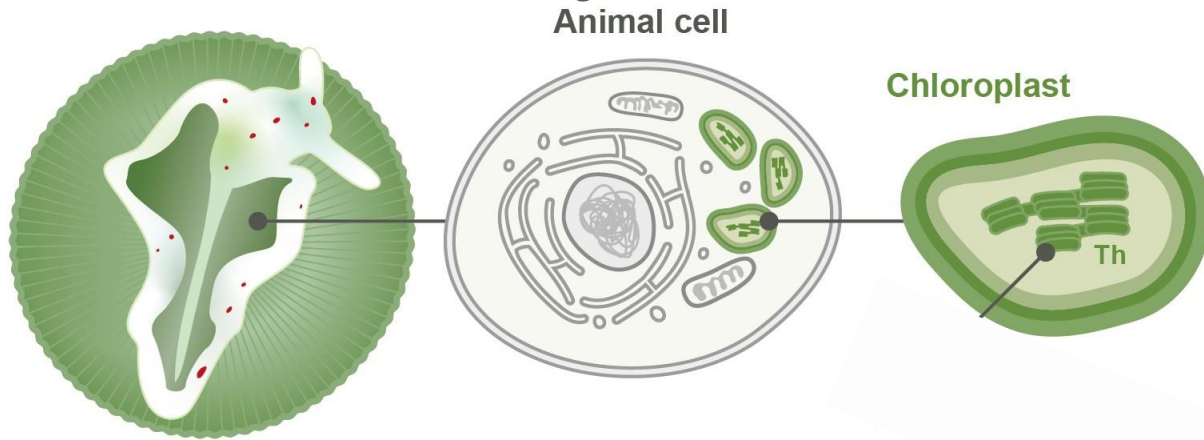
Plakobranchus  
Ocellatus



Elysia Timida

# Kleptoplasty in Sea Slugs

- some species within the sacoglossan sea slugs can retain photosynthetically active chloroplast from the algae they feed on
- many of the proteins required for the maintenance of functional chloroplasts are encoded in the nucleus of algae



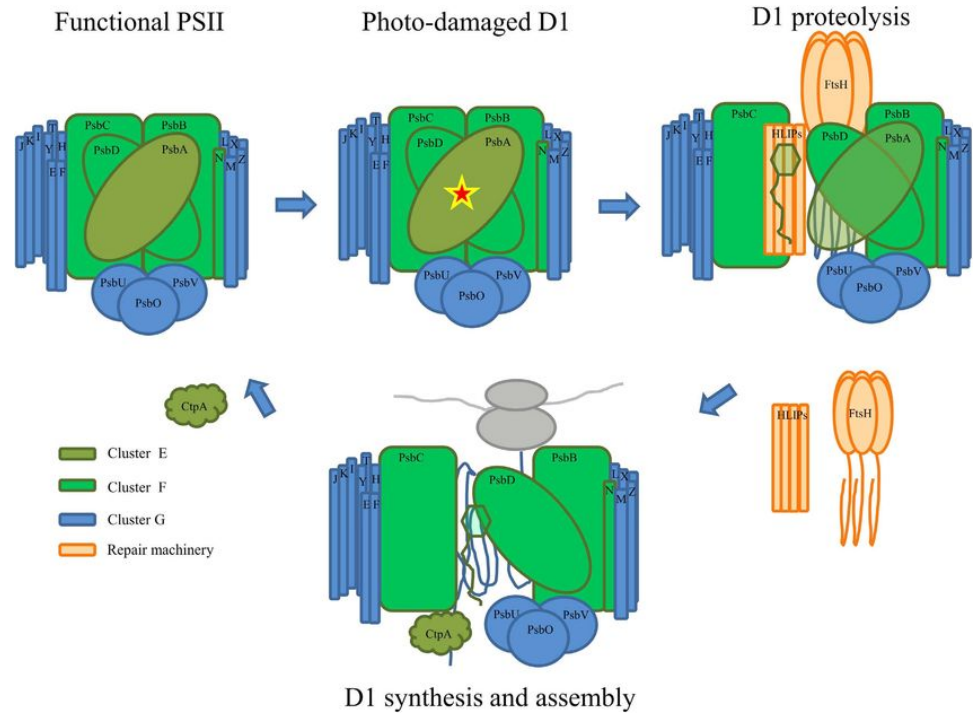
→ it is unclear how slugs are able to maintain the chloroplasts

# Theory: Algae Chloroplast are functionally independent

- kleptoplastic chloroplasts encode the *ftsH* gene themselves
  - *ftsH*: a quality control protease essential for Photosystem II repair
  - in land plants and some algae the *ftsH* is nuclear encoded
  - *ftsH* genes also in the slug genomes
- we decided to look more closely at *ftsH* genes in slugs both with and without kleptoplasty

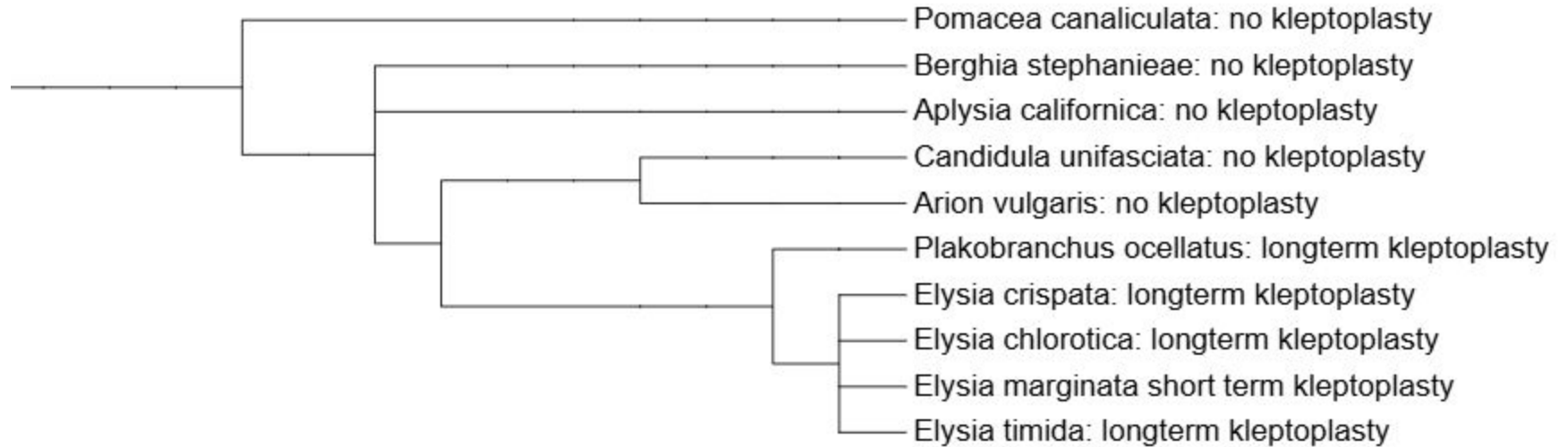
# ftsH

- **Description:** Encodes membrane-bound, ATP-dependent protease known as ftsH
- **Eukaryotes:** Quality control of chloroplast and mitochondrial proteins
- **Chloroplasts:** Proteolizes D1 protein of Photosystem 2
- D1 needs to be continuously replaced because degenerate D1 produces reactive oxygen species (ROS)





# Gastropod species we looked at



short term kleptoplasty

long term kleptoplasty

# ftsH Exons and gene size

Gene	Elysia chlorotica	E.crispata 1	E.crispata 2	P. ocellatus 1	P. ocellatus 2	P. ocellatus 3	P. ocellatus 4	E.timida 1	E.timida 2	E.timida 3
Exon	20	21	20	5	19	20	12	20	19	18
Protein length AA	778	744	771	773	421	784	280	776	761	767

Gene	E.marginata 1	E. marginata 2	E.marginata 3	E.marginata 4	E.marginata 5	C.unifasciat a 1	C.unifasciat a 2	C.unifasciat a 3	P. canaliculata 1	P. canaliculata 2	A.californica 1	A.california 2	A.california 3
Exon	22	7487	10	1	1	15	15	14	19	19	20	20	20
Protein length AA	650	513	933	261	374	738	652	585	748	779	816	746	765



# ftsH Gene Region in Gastropods

Organism/ Position	Elysia chloritica	E.crisparta 218530	E.crisparta g126370	E.Marginata GFR75316	E. Marginata GFR78877	E.Marginata GFR98292	E.Marginata_ GFR75213	E.Marginata GFR68670	P. ocellatus GFN86709	P. ocellatus GFN98809	P. ocellatu GFO05900	P. ocellatu GFO39444	E.timida_g38	E.timida_g39 82	E.timida_g13 96	C.unifasciata_ LOCUS16468	C.unifasciata_ LOCUS12164	C.unifasciata_ LOCUS16772	P. canaliculata LOC1125772 12	P. canaliculata paraplegin	A.californica_ paraplegin	A.california_A FG3 like Protein	A.california_Y ME1L	
5	YERD	HEBP2		SIRT5			GYRA		no blast match		NWD1		CEBPA/B	SLC52A3	K02A2.6-like	no blast match			ZNFX1	LOC1125764 75	LOC1018613 85	TAF4		
4	ACCC	PNKP		PDP1			GYRB		OSBP		no blast match		PPARG	SLC22A15	K02A2.6-like	no blast match			ZNFX1	LOC1125764 74	SLC6A19,	KCTD15		
3	BCCP	PNKP	scaffold end	EIF3			IF-2		ACNAT1	scaffold end	no blast match	scaffold end	PSMG2	no blast match	NPFFR2	RRBP1	scaffold end	scaffold end	ZNFX1	LOC1125764 65	SLC6A19	SMB1	scaffold end	
2	RSMA	FKH1	NWD1	SETMAR			RNR		EIF2AK2	no blast match	no blast match	NPY	LOC1319397 65	SLC25A13	UBE2Q2	ANKRD12	EMC4	LOC1125764 82	LOC1125764 82	ALPK1	TSNAXIP1	TMEM64	LOC1018618 46	
1	SYR	EIF2AK1	no blast match	UBE2Q2	scaffold end	scaffold end	adenylosuccin ate synthetase	scaffold end	LOC1060118 51	no blast match	no blast match	no blast match	ANKAR	LOC1060747 31	TIGD4	no blast match	BOP1	neo-calmoduli n-like	LOC1125772 82	ALPK1	UBC-25	NUDCD1	LOC1060132 30	
0	ftsH	ftsH	ftsH	ftsH	ftsH	ftsH	ftsH	ftsH	ftsH	ftsH	ftsH	ftsH	ftsH	ftsH	ftsH	ftsH	ftsH	ftsH	ftsH	AFG3-like protein 2	paraplegin- like	Paraplegin	AFG3-like protein	YME1L
1	scaffold end	WD40_zf-RIN G_UBOX	NOX2	scaffold end	scaffold end	scaffold end	no blast match	UROD	no blast match	EIF2AK2	MPD1	RPS28	no blast match	LOC1351137 35	EIF2AK2	OSGIN2	neo-calmoduli n-like	COPZ1	NOP9	ANKRD11	ANKRD11	TMEM170	shematin-like protein 1	
2		UNC50	no blast match				RRM1	HMBS	no blast match	UGT2A1	TTLL5	scaffold end	LOC1184780 16	SMARCB1	PPIL6	neo-calmoduli n-like	scaffold end	PFDN5	LOC1125771 07	CYPB	FOXC1	LOC1018594 92	LOC1018606 04	
3		ITFG1	no blast match				FBA1	Fe-S oxidoreductas e-like protein	endonuclease -reverse transcriptase	PPIL6	CYT1		DDO-like	PSME3	PNKP	scaffold end		GGACT	LOC1125771 08	CYPB	scaffold end	LOC1060120 95	ATAD2	
4		CFAP20	no blast match				ALDOB	ACAT1	scaffold end	PNKP	KDELC1		TLR4	LOC1319297 95	TRAF7			ABHD2	LOC1125531 54	SNRK		scaffold end	BRD4A	
5		no blast match	no blast match				scaffold end	COX1		TRAF7	PO21_19		DUS2-like	MRC1	TIP			BMPR1A	RMDN1	LOC1125533 48			scaffold end	

9

# ftsH Gen Region in chloroplast

Chlorella vulgaris	Cyanidioschyzon merolae	Gonium pectorale	Pycnococcus provasolii	Vaucheria litorea
psbL	psbB	psaJ	ycf4	trnW_cca
psbF	psbT	atpI	rnpB	rpl11
<i>cell division protein</i>	psbN	psbJ	trnW(cca)	rpl1
psbE	psbH	pasA	trnS(uga)	rpl12
no blast match	psaE	psbD	trnS(gcu)	psaE
ftsH	ftsH	ftsH	ftsH	ftsH
trnT	trnD	psbC	trnP(ugg)	psbB
cysA	trnS	trnH	rpl23	psbT
AccD	acpP	trnF	rpl2	psbN
psaI	ycf86	psaC	rps19	psbH
trnV	accB	petL	rps3	petF

no ftsh:

Arabidopsis thaliana

Bryopsis hypnoides

Klebsormidium nitens

# Low Exon Homology

we created an exon database (336 exons from 27 ftsH genes in 8 slug species)  
and used blastn with e-value cutoff of 0.05 to establish exon homology

All-vs-All and Org vs Exons and CDS vs Exons

→ only revealed these matches

E. Marginata: GFR753161	Exon 4	Exon 8	Exon 9		Exon 12		
P. Ocellatus: GFN89809	Exon 4				Exon 12	Exon 13	Exon 14
E. Timida: g1396	Exon 4			Exon 6			
E. Crispata: g218530		Exon 14	Exon 13	Exon 16		Exon 9	Exon 8

# MACSE (Multiple Alignment of Coding SEquences)

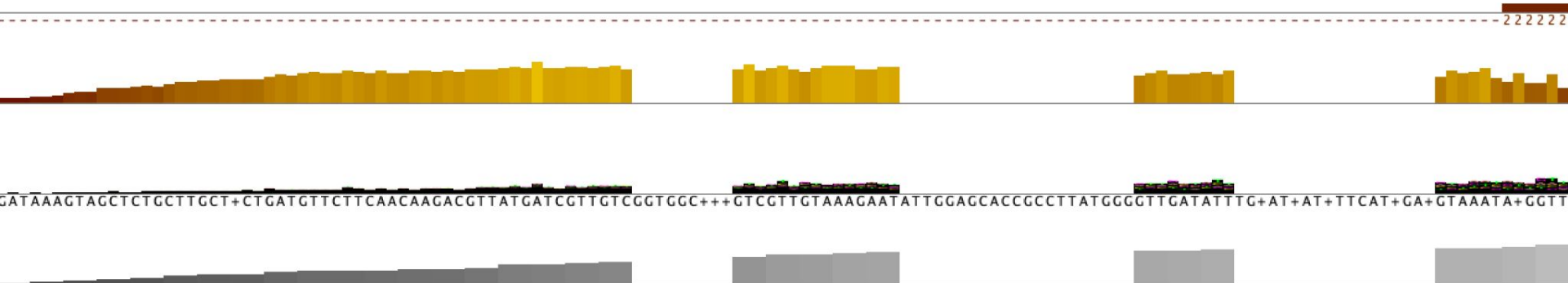
- codon-aware multiple sequence alignments of protein-coding DNA sequences
- takes into account the codon structure and reading frame → aligns sequences in a way that minimizes disruptions to the triplet codon organization
- Features:
  - Codon-aware
  - Frameshift
  - Designed for CDSs

# Alignment of all exons - highly conserved region

```

      940      950      960      970      980      990      1000      1010      1020      1030      1040      1050      1060      1070      1080
-----!CATTCCTGAGG!!TGAgaaaattctgaaagaaactcTTTCATGGTGATTTC-----TCTGTA CTGGCTGGA-----GTAGAGCAT-----GTAGAGAAGCAC
-----          !TGAAAACAACCTCCA-----TCAAATCTG!T!TCT-----ATTTCCGTC-----TTTAGTCTTGGT
-----CTCGCCTTTAAGAAGTATATTGTTGGTAGAAATCACT-----CCA!T!GAAATGTGG-----CTTTGTCCT-----CCAGGGTTACCA
-----          CTT-----TTGGATCGAATGTGT-----ATGATGTTA-----GGAGGAAGGATC
-----          CTT-----CTGGATCGCATGTGC-----ATGATGCTG-----GGTGGGAGAGTC
-----          ctc-----ctGGACCGCATGTGC-----ATGATGCTC-----GGTGGCAGAGTT
-----          CTC-----TTAGACAGAATGTGT-----ATGATGTTA-----GGTGGTCGGGTG
-----TACAACAGGACCGTAGCCCAGCTGAAGGCGTCCATGGACACTGCAATGGGTGGCCGGGCGGCGGAAGAGGTC-----ATATATGGC-----ATGGAGAAGGTG
TTTTATGAAACTATAGTCTGATTGACCTGTGGAATCGTAGTATGTCAAGttacctat-----ttttttattgagtcC-----GTA!A!ATG-----GTAACAATTGCA
-----CTGGCAACTACCTTATTGGAGAAAGAAGTTTTAGATTACGAC-----ATGCTGTGGCAACCG-----GCCGGATAT-----CCAACATTGTT
-----CTTGCTACCACTGTTGGAGAAGGAAGTGCTCAACTACGAC-----GATATAGCAGCTTTGATTGGAGCACCGCCTTATGGGACAAAGAAC-----AAGATAGAGCCA
-----CTTGCTACCACTGCTTGAGAAGGAGGTACTCAACTACGAC-----GACATCGAGAAACTC-----ATTGGTGCG-----CCCGCTATGGC
-----          GACATTGAGAAGCTG-----ATTGGGCCA-----CCACCCACGGG

```

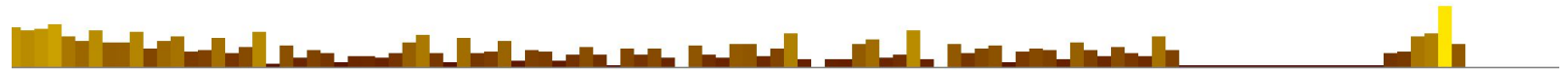


# Alignment of all exons - highly conserved region

```

      970      980      990      1000      1010      1020      1030      1040      1050      1060      1070
TAGAGAAGCACAGCGCTGCCAAGAGCCCCAAAGCAAGCATCTGTTTGTCAAGGTGGTCCGCCTCCAAGTTTCTTACCACCTTTGTC-----TCC!A!-----
TTAGTCTTTGGTAAACATTTGGAACATTTGacggaaaaagaagaaaaggacaaGAACAACCATCAGATCCATAGCCACCATGATCCA-----GGA!!G-----
CAGGGTTACCATCTTCTCGGCGACCAGGTGATAAAAGCCACACAAAGACCATAACAGATACAAATGTTGTAAGGAGCGAATAAGT-----TCA!!G-----
GAGGAAGGATCTCggaa caaat t t c t t t g g c A A A A T C A C A A C T G G T G C T A A G A T G A T C T G a g g a a a g t c a c a a a c a G C C T A T -----T C A C A G -----
GTGGGAGAGTCTCGGAGCAAATCTTCTTTGGGAAGGTGACCACAGGAGCTCAGGATGACCTCAGGAAGGTACACACAGACGGCCTAC-----TCCCAG-----
GTGGCAGAGTTTTCTGAGCAAATTTCTTTGGTAAAATCACAACCTGGTGCTCAGGATGATTTGAGGAAAGTCACGCAAAACAGCGTAC-----TCCCAG-----
GTGGTCCGGTGTCCGAGCAAATCTTTTTCGAGAAAATCACAACCTGGAGCCCAAGATGATTTGAGGAAAGTAACGCAGACAGCTTAT-----TCTCAG-----
TGGAGAAGGTGACCACTGGTGCATCCAACGATCTACAGATGGCCACACGCGTTGCCACTGCCATGGTCATGGTCAACGGCATGTCC-----GACAAA-----
TAACAATTGCACGAGCCTGTTTAGTTACTTTTTCCAAATCGCTGAGTGCCCGCTGTggata t t t a t t g a a a a t a a c t t c C T C T G C G -----G C G C G T -----C C G C
CAAACTATTGTTTACAGATGACATCAATAAAATCTAAACCCACGATCTCTGTGCTCCAGCTGTCTCATCTCAAGATTAAGGGGATT-----TCTGCA-----
AGATAGAGCCAGAGGGATGGAGAGGAATGGAGTCAGACACACAAGACAGTGACAGGCTACCAAGAAGCCTTACGTCACAAAACGA-----AAATAG-----
CCCGCTATGGCAAGAAGAATCTCATAGAACCACACGGCTGGGAGGGTATTATGCCAGAAAGCGCAGCAGCTTCTCACAAGAGAAAAG-----AAATGA-----
CACCCACGGAAGAAGCACATCATCGAGCCTCACGGCTGGGAAGGCATCATGCCGGAGCCAAATGTGGAGAAAAAATAACAGTAA-----

```



```

TAAATA+GGTTGTG+TCAAGATTGT+GCAGTTGATGAGACTGATGATGTCATG+AGGATGCTATTAATGAAAT+TTGATAATATTGTTGGAGTTGGGCC+GAA!AGTGC+C++

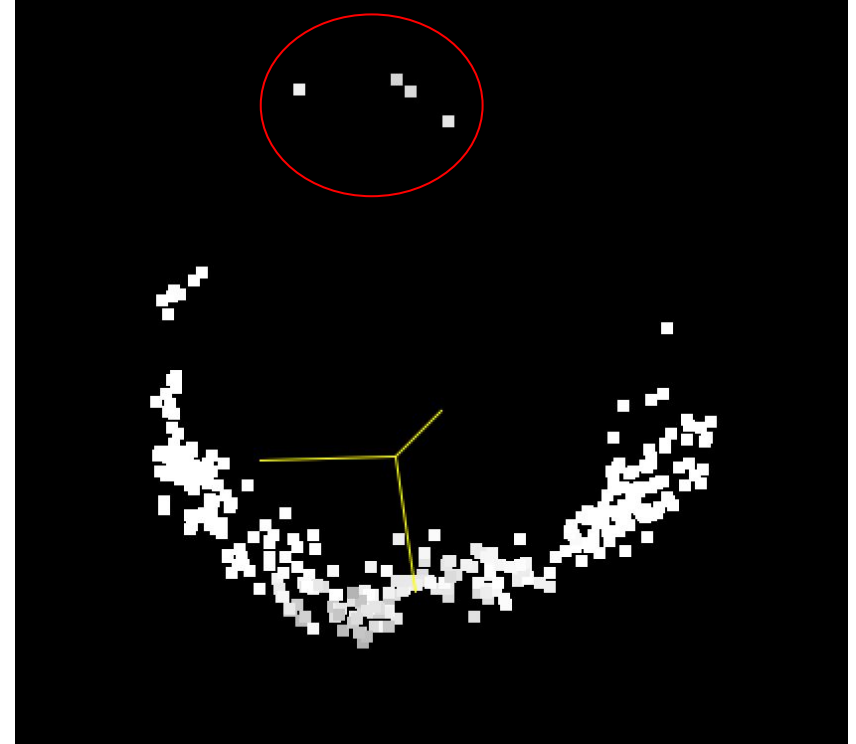
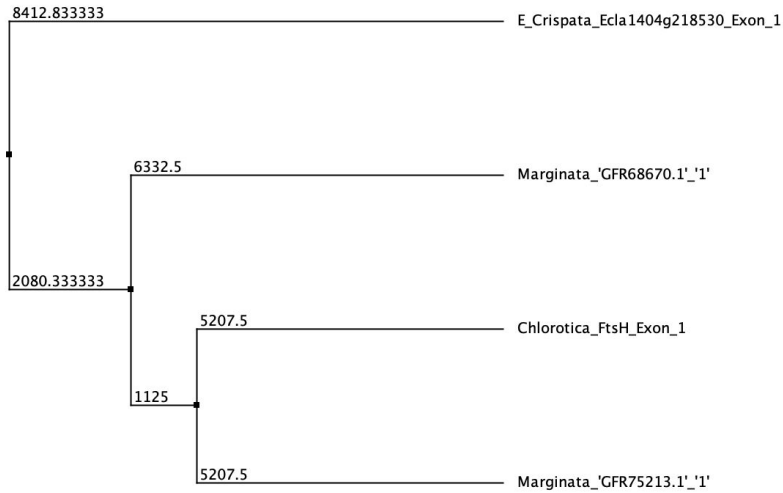
```



# Aligning all Exons → PaSiMap

In the circle:

2\* Marginata, Chlorotica, Crispata

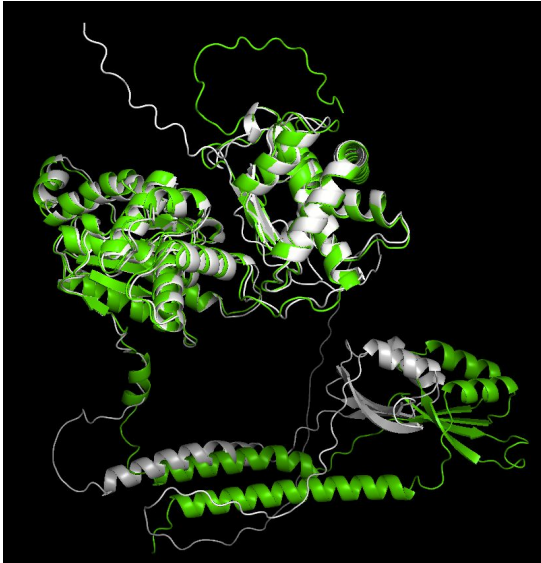




# Structure comparison of ftsH

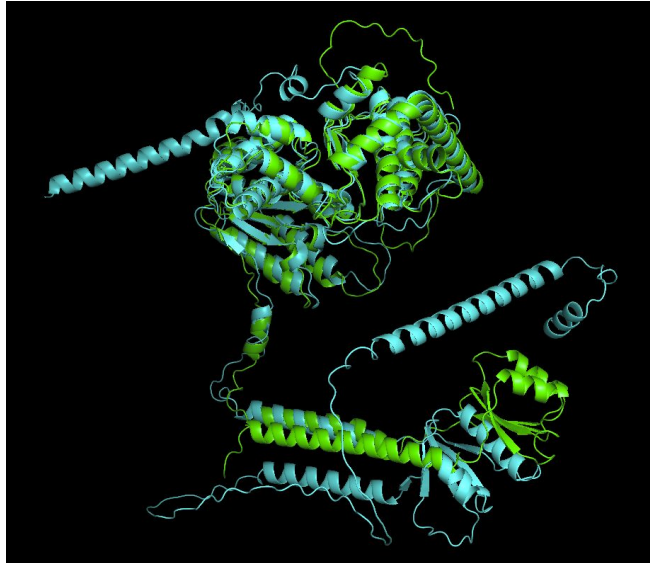
Vaucheria litorea vs

*A. thaliana* ftsH8



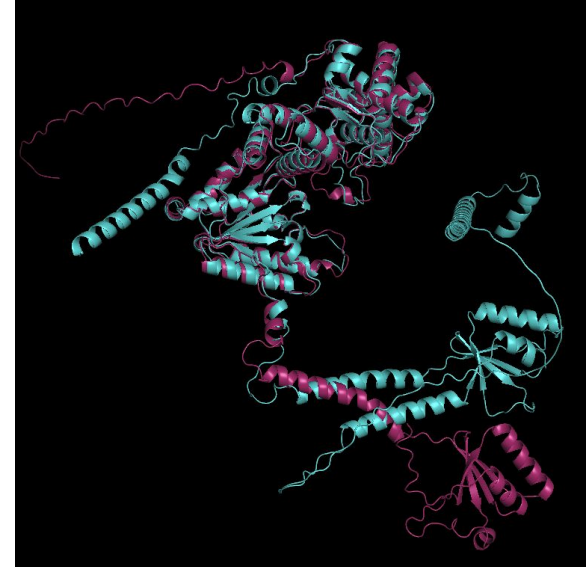
Vaucheria litorea vs

*Elysia chloroitca*

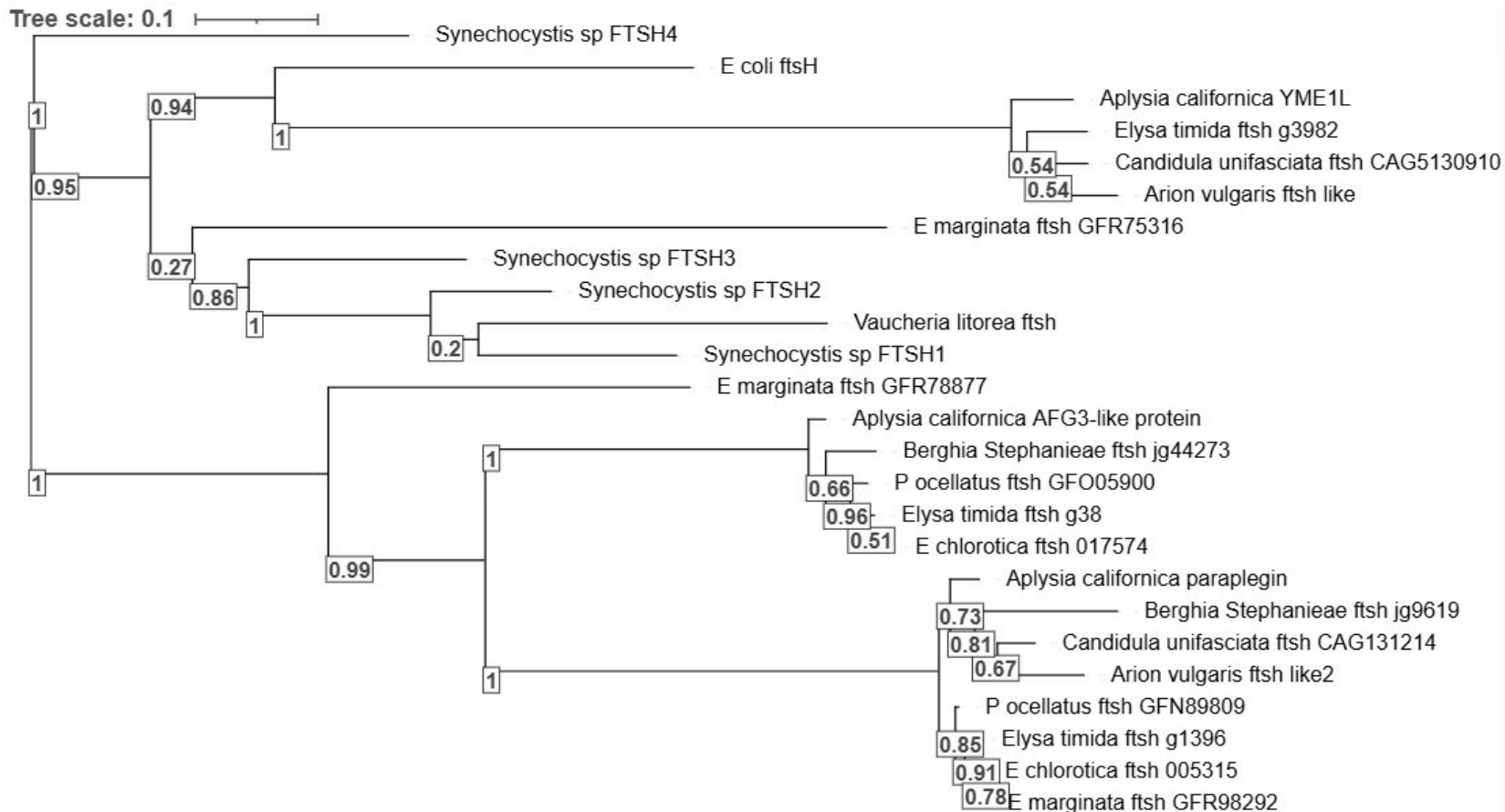


*Elysia marginata* vs

*Elysia chloroitca*

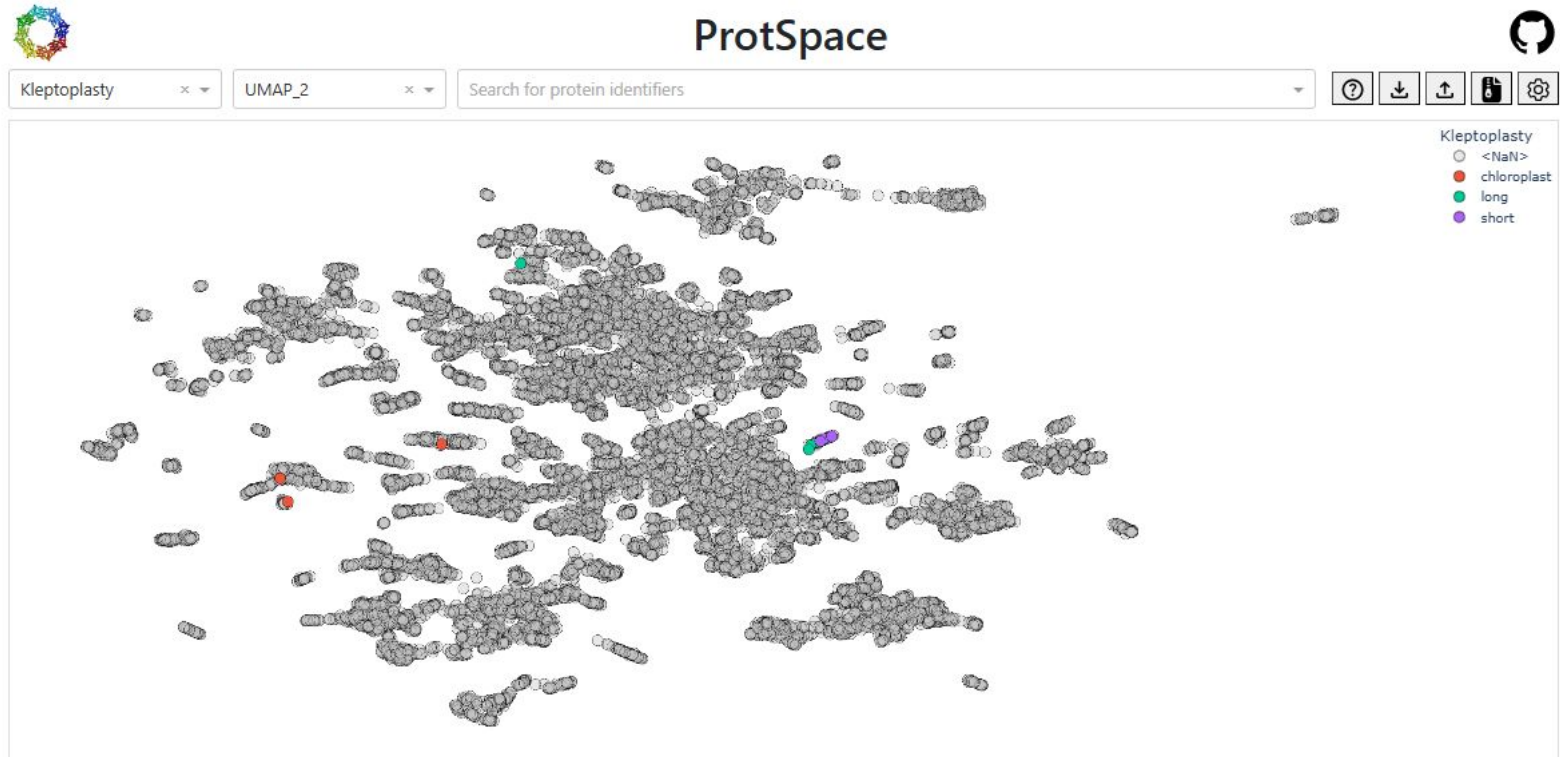


# ftsH protein tree



# Investigating the Protein

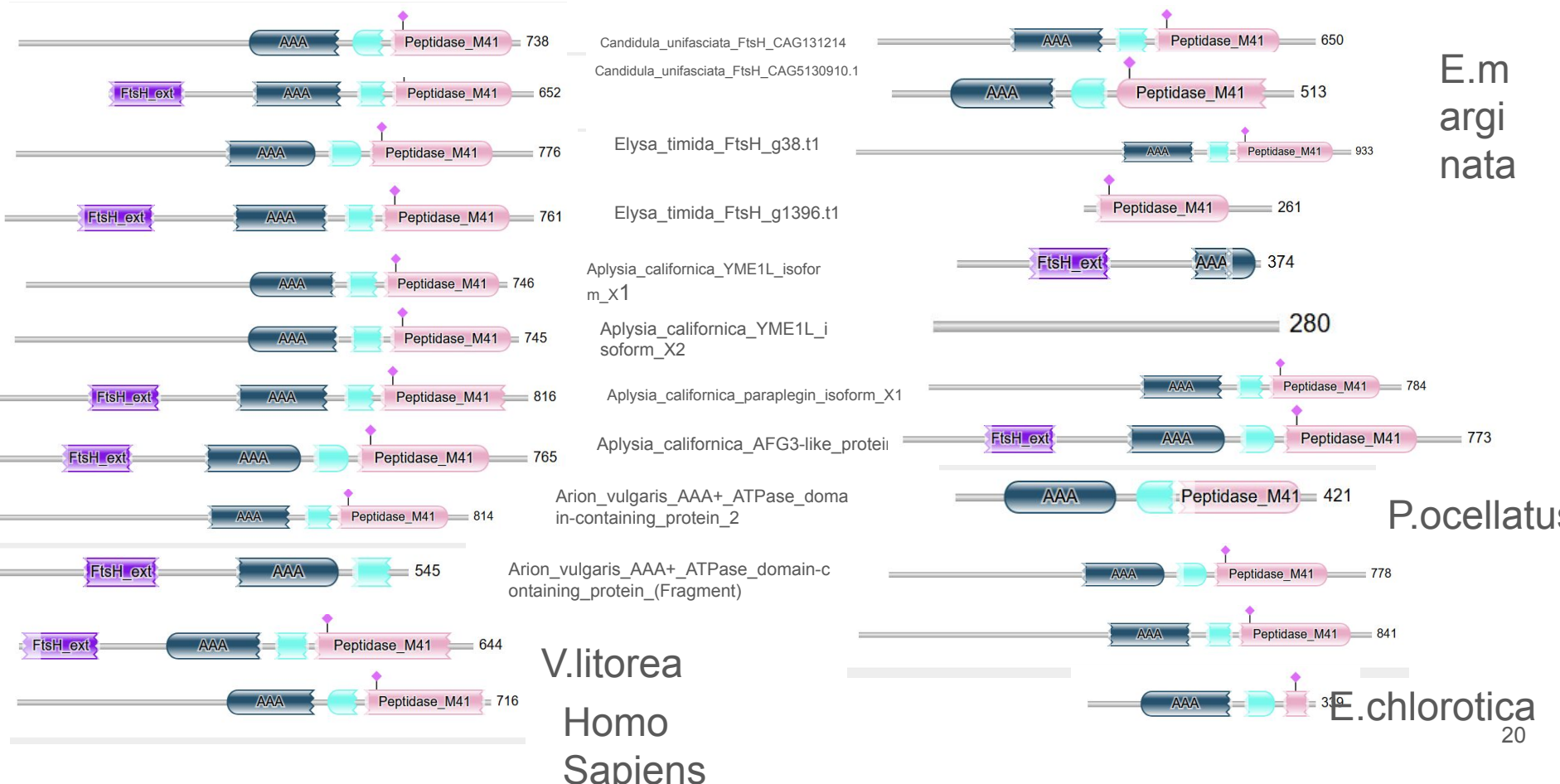
- Cluster Analysis kleptoplasty (UMAP)



# Domain Analysis:

- **AAA+ ATPase domain** (for substrate unfolding).
- **Zinc-binding protease domain** (for proteolysis)
- **M41 Peptidase Domain**: critical for D1 degradation

# Domains from InterproScan



# Conclusion

- Comparisons between slug and chloroplast ftsH showed similarity but no clear link to kleptoplasty was established
- Further analysis on intracellular location and protein-protein interactions needed to investigate ftsH's potential role in chloroplast maintenance

# Sources

## Detecting Sequence Signals in Targeting Peptides Using Deep Learning

José Juan Almagro Armenteros, Marco Salvatore, Ole Winther, Olof Emanuelsson, Gunnar von Heijne, Arne Elofsson, and Henrik Nielsen

*Life Science Alliance* **2** (5), e201900429. doi:[10.26508/lsa.201900429](https://doi.org/10.26508/lsa.201900429) (Open access)

Abramson, J., Adler, J., Dunger, J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024). <https://doi.org/10.1038/s41586-024-07487-w>

Sebastian Bittrich, Joan Segura, Jose M Duarte, Stephen K Burley, Yana Rose, RCSB protein Data Bank: exploring protein 3D similarities via comprehensive structural alignments, *Bioinformatics*, Volume 40, Issue 6, June 2024, btae370, <https://doi.org/10.1093/bioinformatics/btae370>



# Conclusions

- evidence for and against role of ftsH in kleptoplasty

structural similarity between chloroplast ftsH and slug ftsH	no structural difference between slugs with
	ftsH encoded by chloroplast genomes of kleptoplastic algae

# To-do

- rerun ftsh-tree (schönere Namen + neue Sequenz (Pomacea canaliculata und chloroplasten) Lara
- rerun exon analysis Max
- chloroplasten gene regions plot Annika
- ftsh gene region Matthäus -> erledigt
- protspace Matthäus
- annotation chlorotica, (ftsh fragments, did something mutate to stop codon? are there fsth sequences (before) and after, that are not transcribed; psaA?) Lara, Max, Matthäus, (Annika)
- Plots verbessern, Einleitung Rushali

## nächstes Meeting:

- PP fertig machen & üben
- citations? -> Ja, lass mal unten auf die Folien in kleinen hellgrauen textboxen machen oder einfach nur in den appendix/Quellenverzeichnis ( MH )
- extra Folien für potentielle Fragen

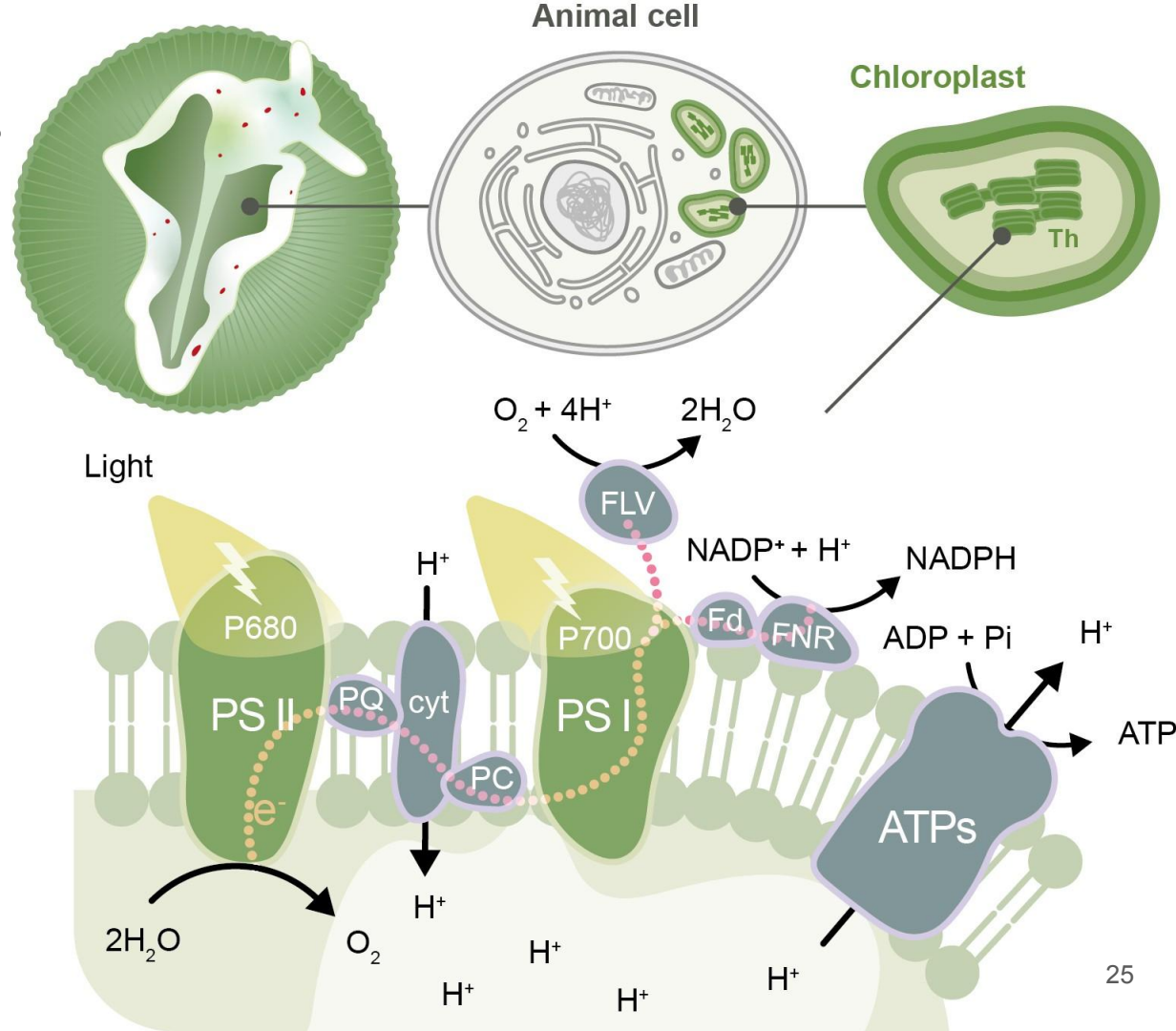
Idea MH: Add  
3D structure of  
ftsH from  
AlphaFold  
already here

idea MH: I.  
slide before  
What do  
FTSH do  
what do  
know about  
(show diagram  
from paper)

# What is Kleptoplasty ?

- Stealing Chloroplasts from algae
- IN SEA SLUGS:

Plastid retention and photosynthetic ability varies. But how do they maintain Chloroplasts?



# Original Hypothesis: Horizontal Gene Transfer

HGT between the algal nuclear genome and slug genomes

we compared slug genomes with algae genomes and found little evidence for this theory

# Which Genes did we consider that led to FTSH and did we arrive at them:

- Lists that we looked at
  - From Literature
  - From Metabolic approach

## Your BlastKOALA job

Query dataset: 9796 entries

KEGG database searched:

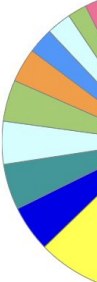
Job submitted: Fri Mar 14

Job completed: Fri Mar 14

## Annotation data [View](#) | [Download](#)

Summary 2557 entries

Functional category [View Pathway](#)



KEGG Mapper [Reconstruct](#)

# ftsH Gen Region in chloroplast

Chlorella vulgaris	Cyanidioschyzon merolae	Gonium pectorale	Pycnococcus provasolii	Vaucheria litorea
psbL	psbB	psaJ	ycf4	trnW_cca
psbF	psbT	atpI	rnpB	rpl11
cell division protein	psbN	psbJ	trnW(cca)	rpl1
psbE	psbH	pasA	trnS(uga)	rpl12
no blast match	psaE	psbD	trnS(gcu)	psaE
ftsH	ftsH	ftsH	ftsH	ftsH
trnT	trnD	psbC	trnP(ugg)	psbB
cysA	trnS	trnH	rpl23	psbT
AccD	acpP	trnF	rpl2	psbN
psaI	ycf86	psaC	rps19	psbH
trnV	accB	petL	rps3	petF

no protein product

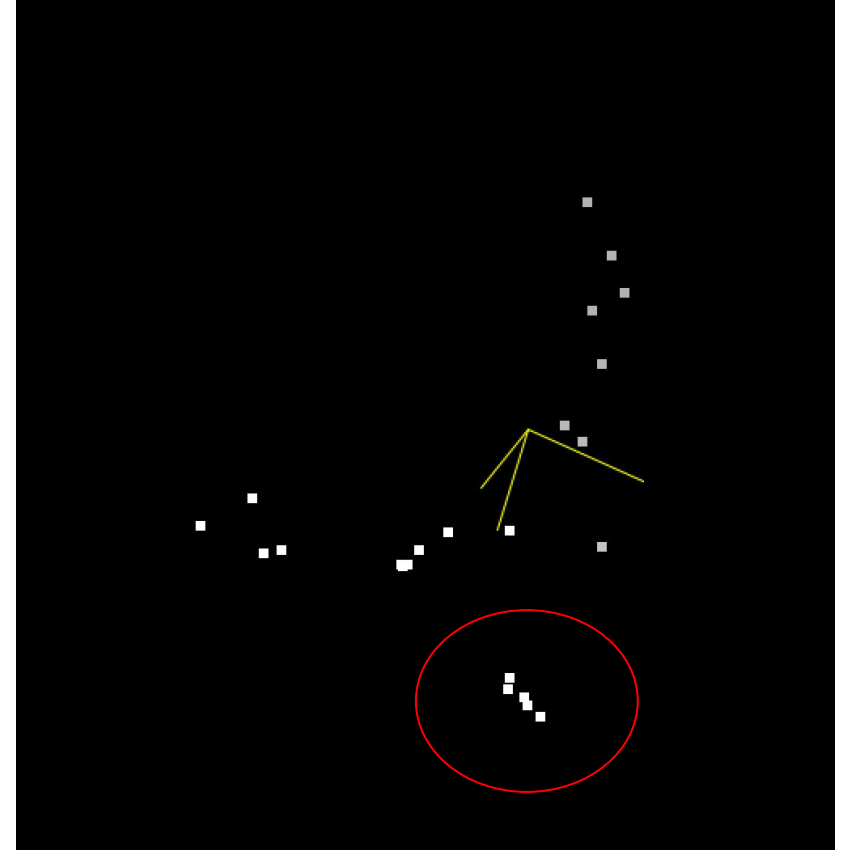
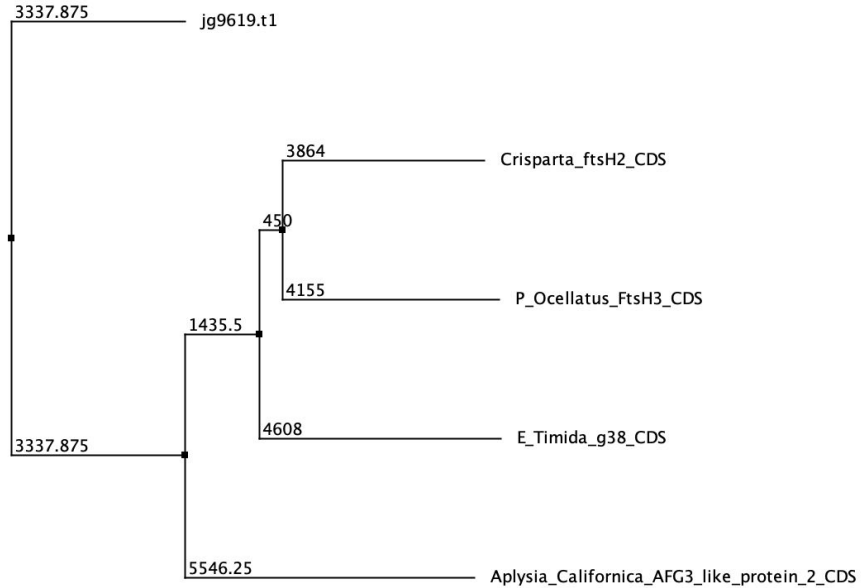
no ftsh:

Arabidopsis thaliana

Bryopsis hypnoides

Klebsormidium nitens

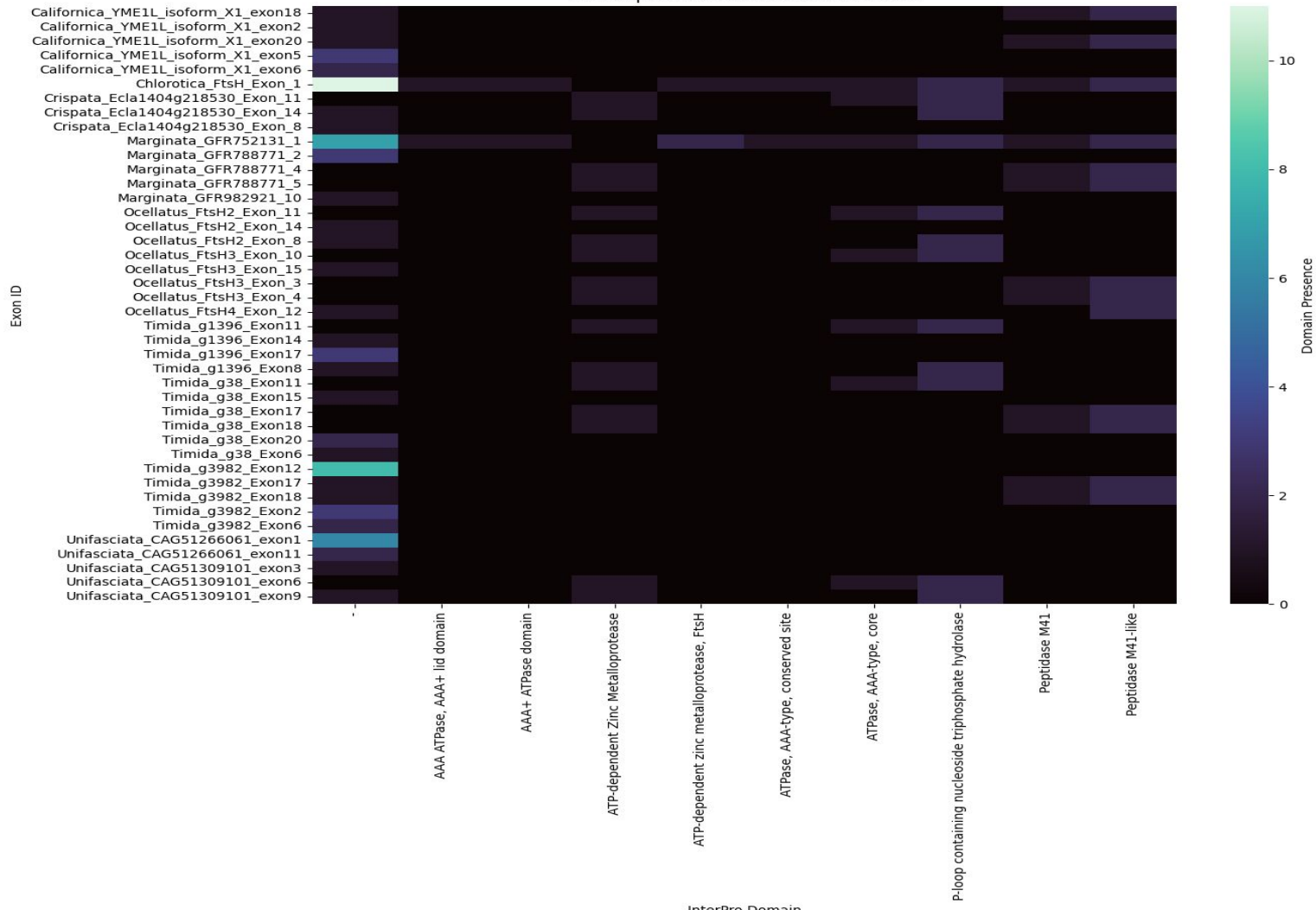
# Aligning all CDS



TODO: Add Parameter



Heatmap: Domain Presence Across Exons



# Comparing simulated protein structures

Vaucheria litorea vs	RMSD	TM-score	Identity %	RMSD	TM-score	Identity %
aplysia_californica_afg3_like_protein	3.45	0.75	41	4	0.64	39
aplysia_californica_paraplegin_isoform_x1	3.75	0.76	37	4.04	0.61	33
aplysia_californica_yme1l_isoform_x1	3.4	0.71	40	3.09	0.61	39
aplysia_californica_yme1l_isoform_x2	2.88	0.69	42	3.15	0.6	36
arion_vulgaris_aaa_atpase_domain_containing_protein_1	2.75	0.7	42	3.2	0.61	38
arion_vulgaris_aaa_atpase_domain_containing_protein_2	3.57	0.77	35	3.93	0.61	33
arion_vulgaris_aaa_atpase_domain_containing_protein_fragment	4.38	0.49	44	5.01	0.49	42
candidula_unifasciata_ftsh_cag131214	3.49	0.7	41	3.08	0.61	38
candidula_unifasciata_ftsh_cag5130910	3.42	0.78	37	3.2	0.76	35
candidula_unifasciata_ftsh_cag5131214	3.42	0.74	39	3.29	0.65	38
elysia_timida_ftsh_g1396	4	0.77	36	3.52	0.65	34
elysia_timida_ftsh_g38	3.31	0.72	42	5.69	0.58	39
elysia_timida_ftsh_g3982	3.11	0.7	41	1.46	0.59	44
elysia_chloroitca_ftsh1	3.77	0.79	40	3.37	0.65	38
elysia_chloroitca_ftsh2_fragment	4.07	0.72	36	5.06	0.56	32
elysia_chloroitca_ftsh3	2.96	0.45	55	4.4	0.45	53
elysia_marginata_ftsh1	2.32	0.73	43	6.98	0.69	40
elysia_marginata_ftsh2	2.4	0.71	47	2.93	0.7	47
elysia_marginata_ftsh3	3.45	0.74	37	5.49	0.5	31
elysia_marginata_ftsh4	2.18	0.32	32	2.66	0.32	32
elysia_marginata_ftsh5	6.14	0.26	17	7.77	0.26	19
homo_sapiens_cab51858_atp_dependent_metalloprotease_yme1l	3.13	0.69	41	3.1	0.61	36
p_ocellatus_ftsh1_catalytic_subunit	5.89	0.11	9	8.72	0.07	4
p_ocellatus_ftsh2	4	0.77	36	3.69	0.64	37
p_ocellatus_ftsh3	3.93	0.8	39	3.01	0.66	38
p_ocellatus_ftsh4_partial	2.14	0.41	48	14.83	0.25	29

# Investigating the Protein

- Cluster Analysis protein clustering by organism (UMAP) *E.chlorotica* marked

