

BI694

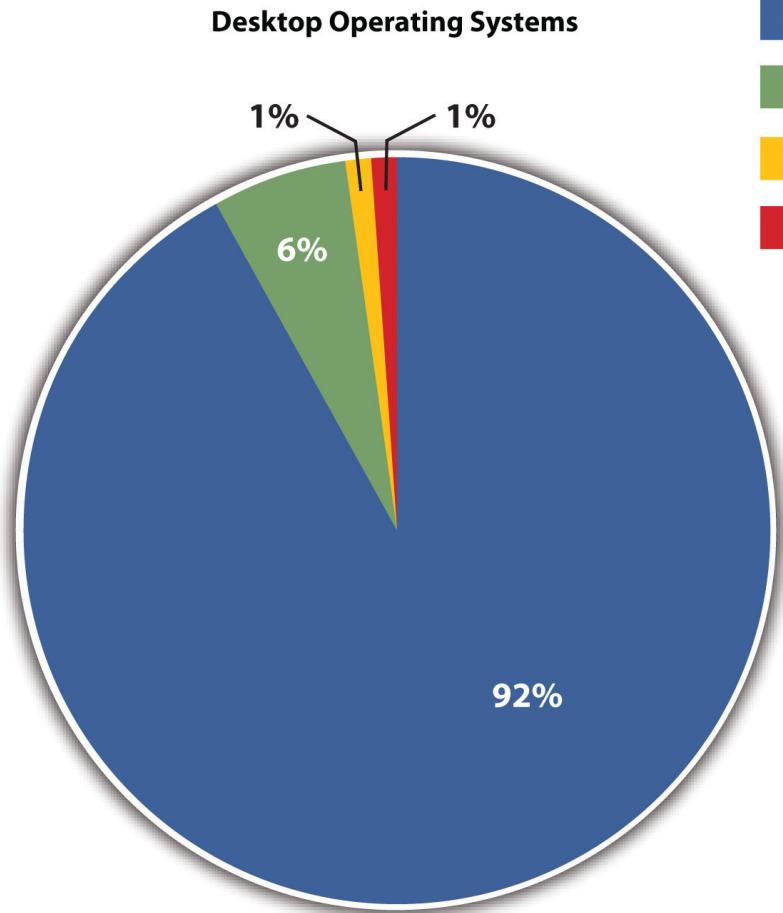
Bioinformatics & Phylogenetics

Winter Semester 2017

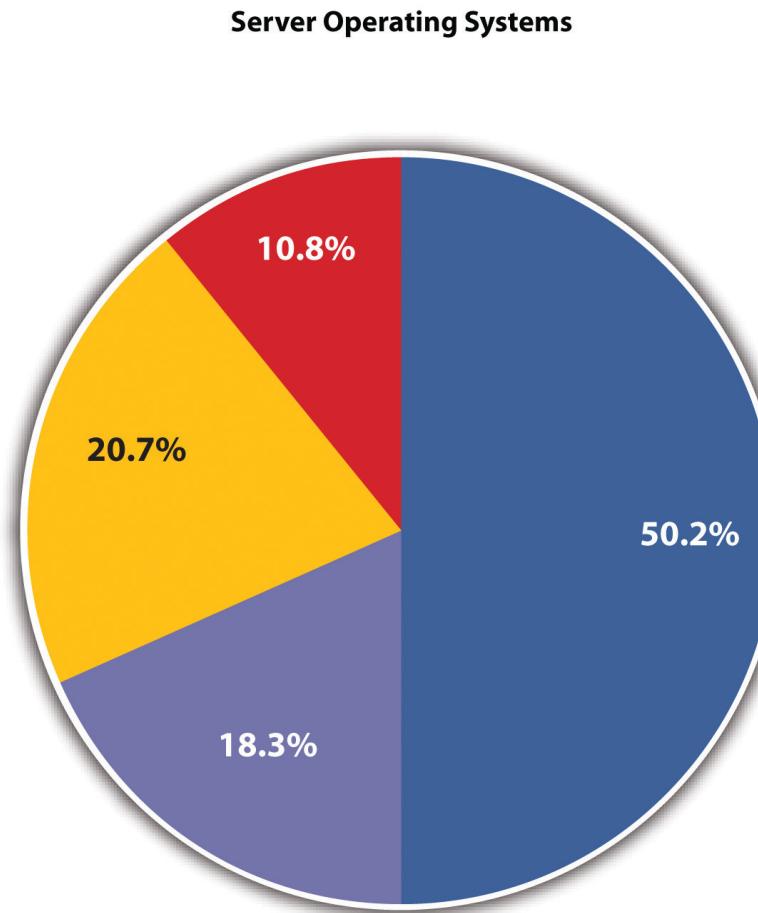
What is Bioinformatics?

“Bioinformatics employs a wide range of computational techniques including sequence and structural alignment, database design and data mining, macromolecular geometry, phylogenetic tree construction, prediction of protein structure and function, gene finding, and expression data clustering. The emphasis is on approaches integrating a variety of computational methods and heterogeneous data sources. Finally, bioinformatics is a practical discipline.”

Luscombe et al. (2001) *Methods of Information in Medicine*

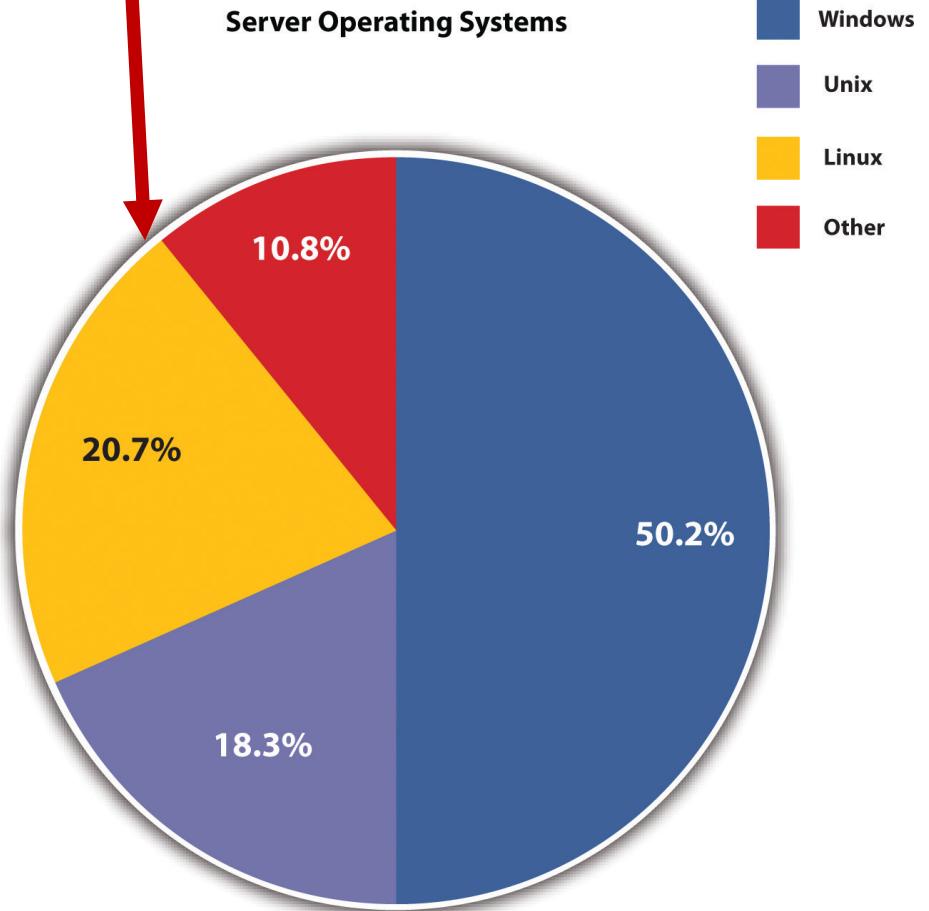
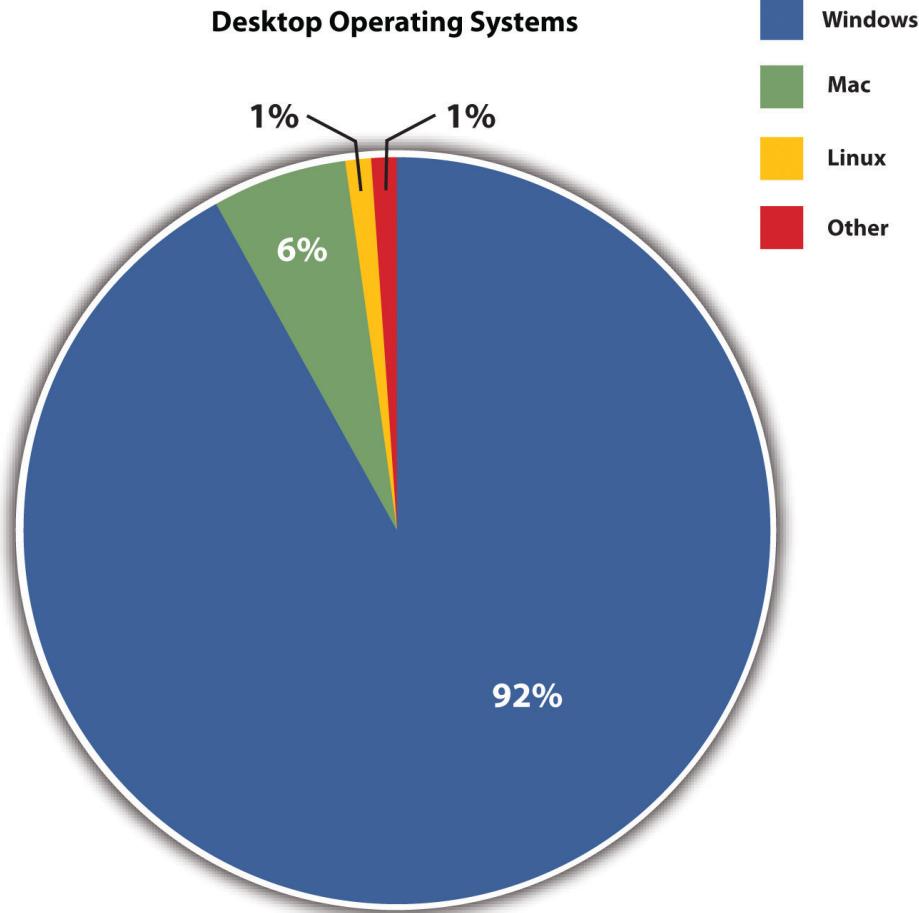


Windows
Mac
Linux
Other

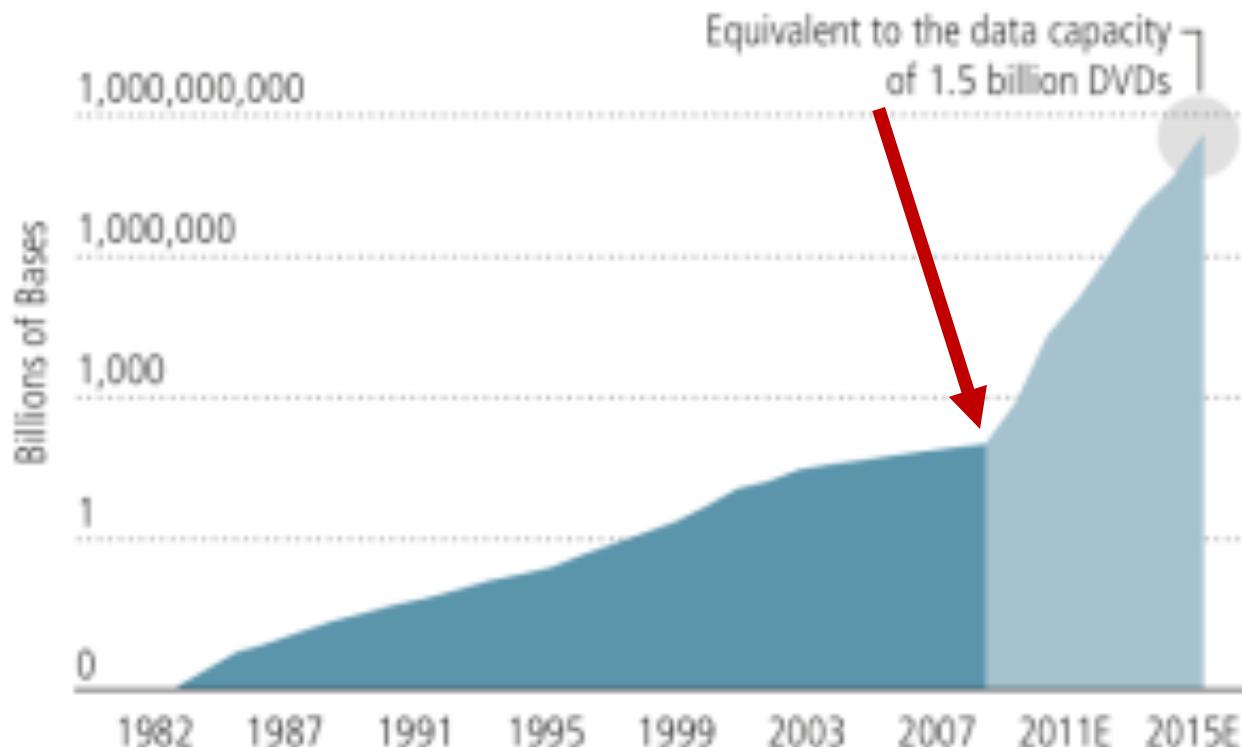


Windows
Unix
Linux
Other

Bioinformatics



Sequencing Data Generation Should Rise Rapidly, in Our View



As of February 8, 2011; estimates are subject to change.

Projections begin on January 1, 2009, and are in terms of bases from finished sequences, rather than the total number of bases generated (which is considerably higher, as multiple runs are needed to arrive at a complete sequence); DVD comparison assumes conversion rate of one base per byte.

Source: Cisco Systems, GenBank and Alliance Bernstein

illumina®



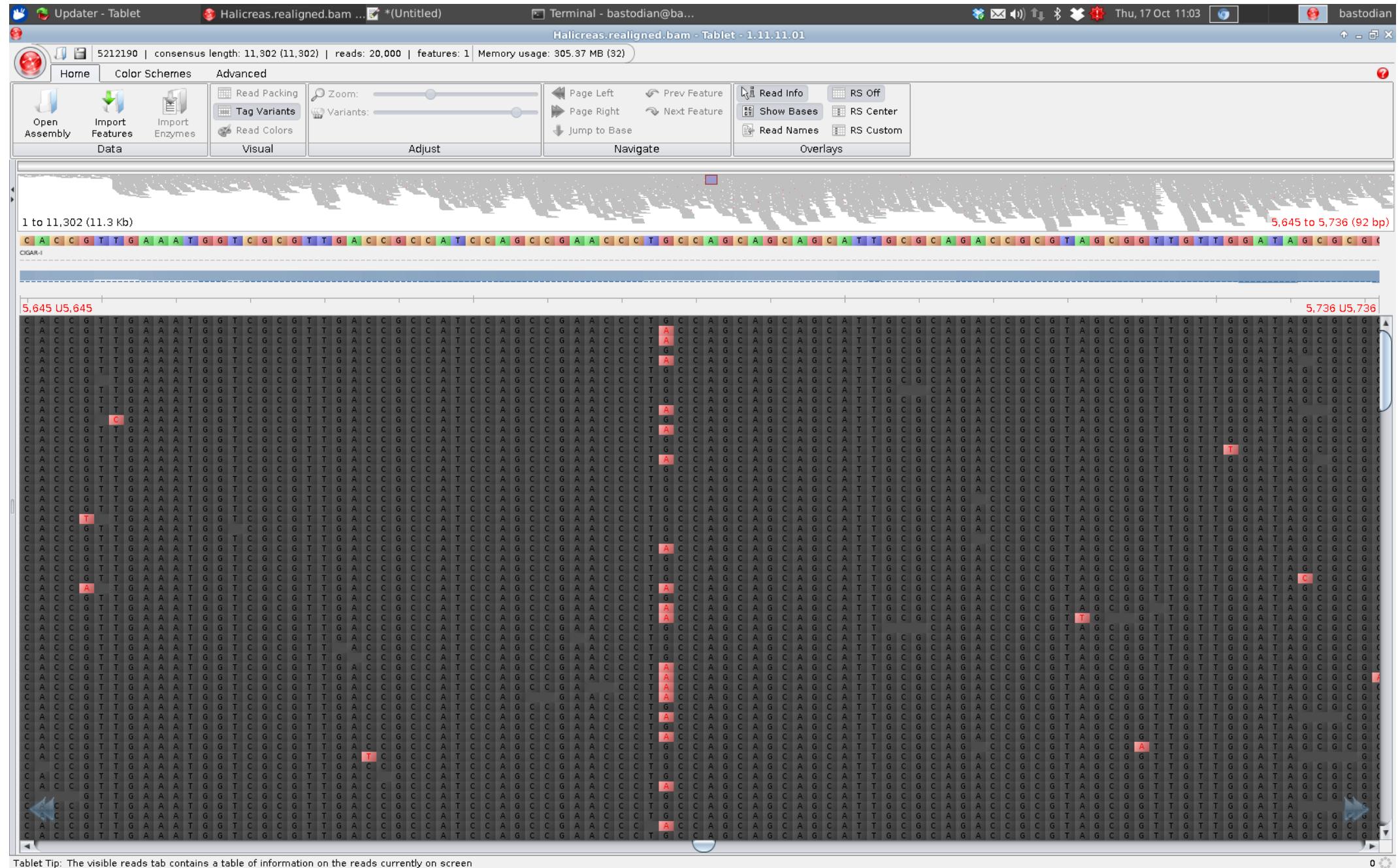
PACBIO®

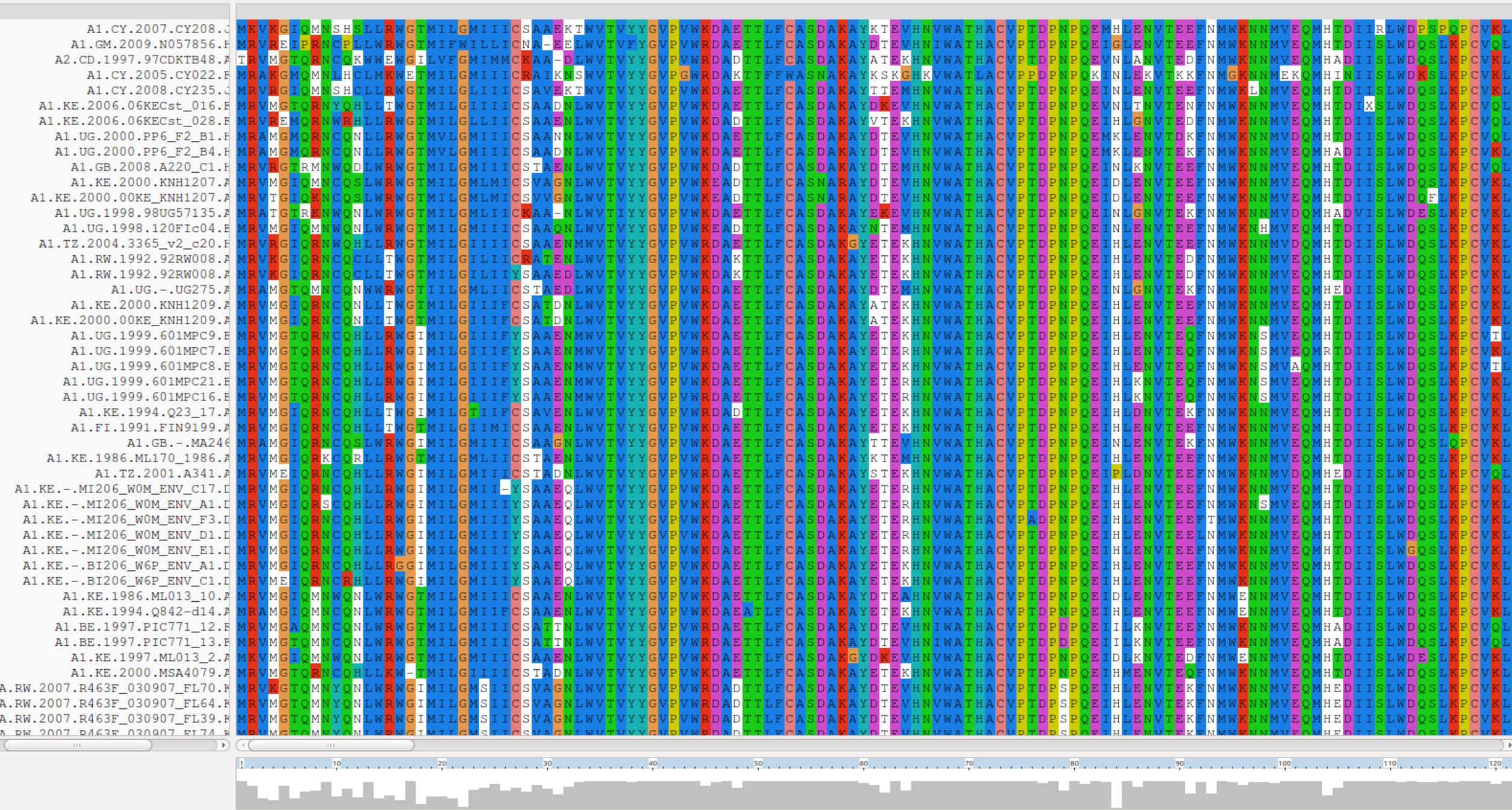
ion torrent

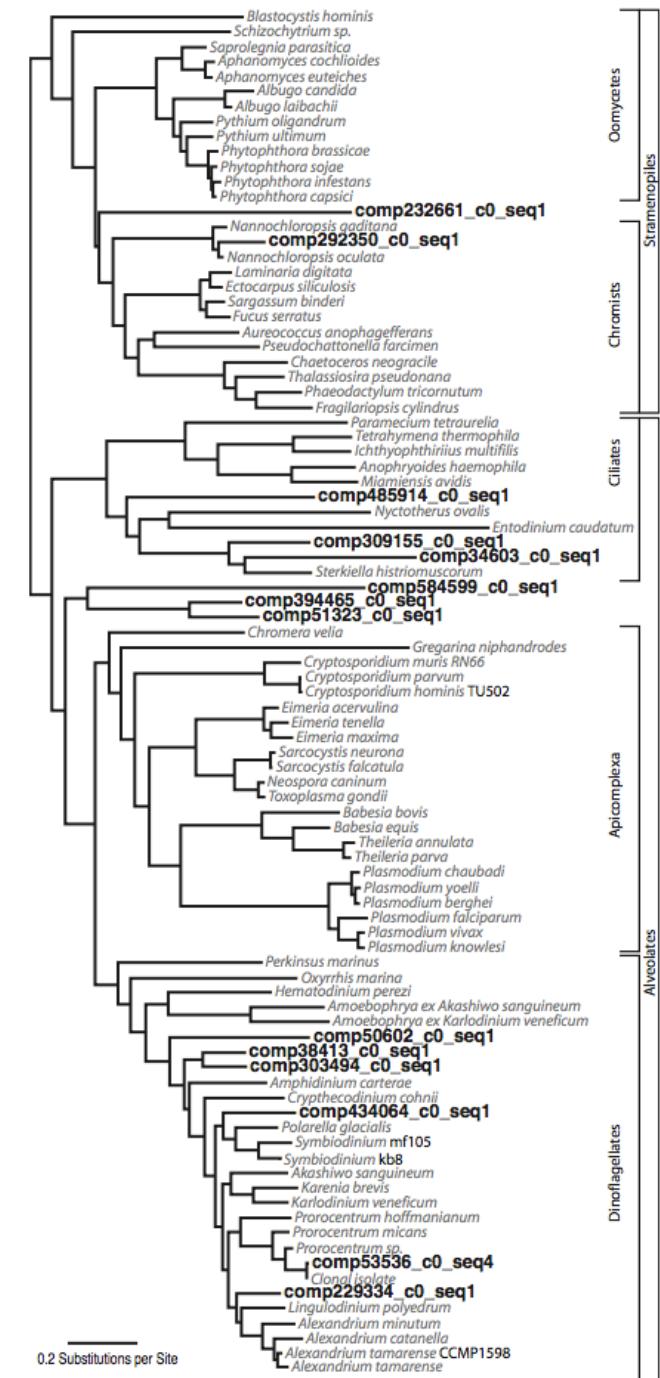
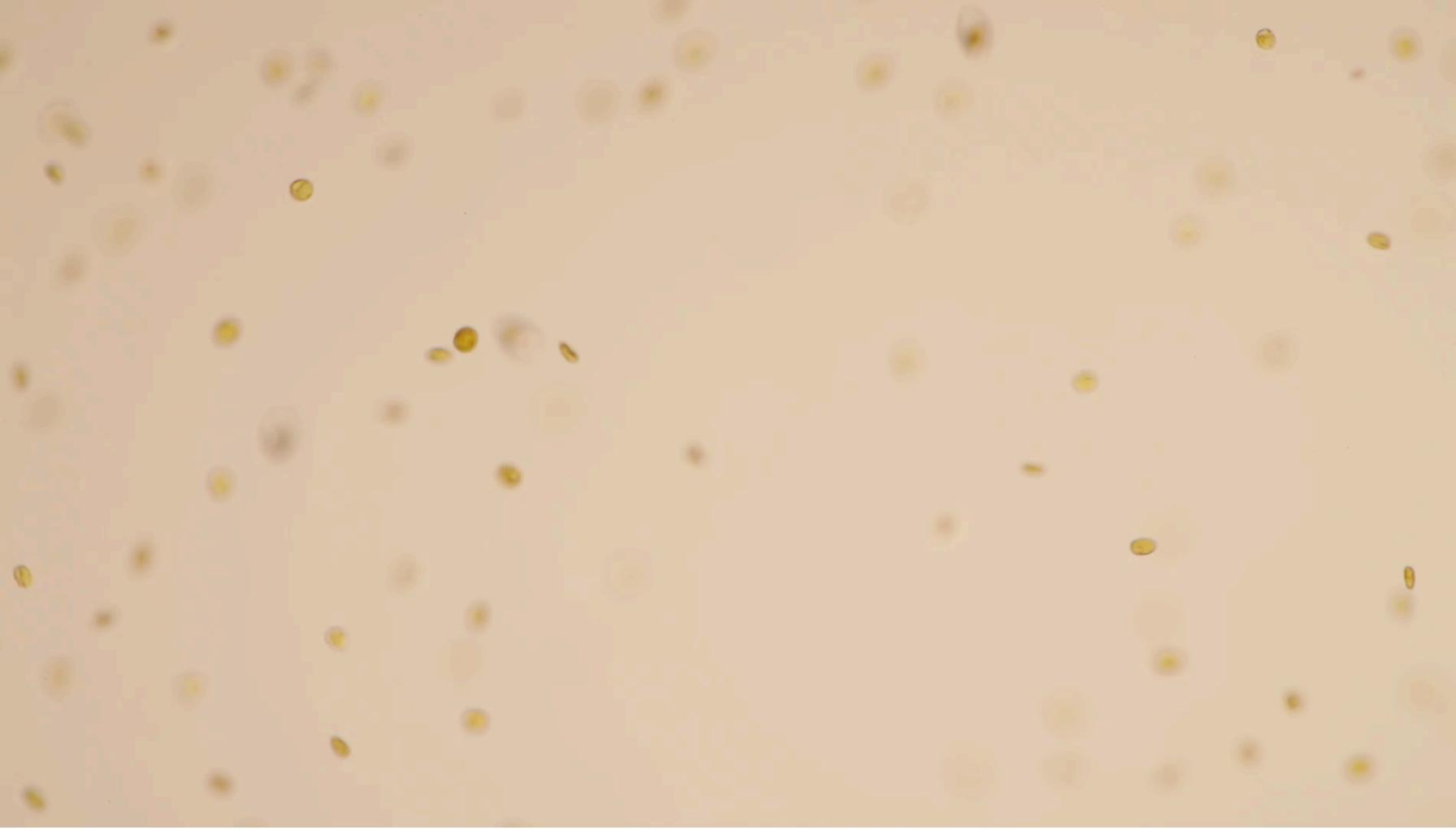


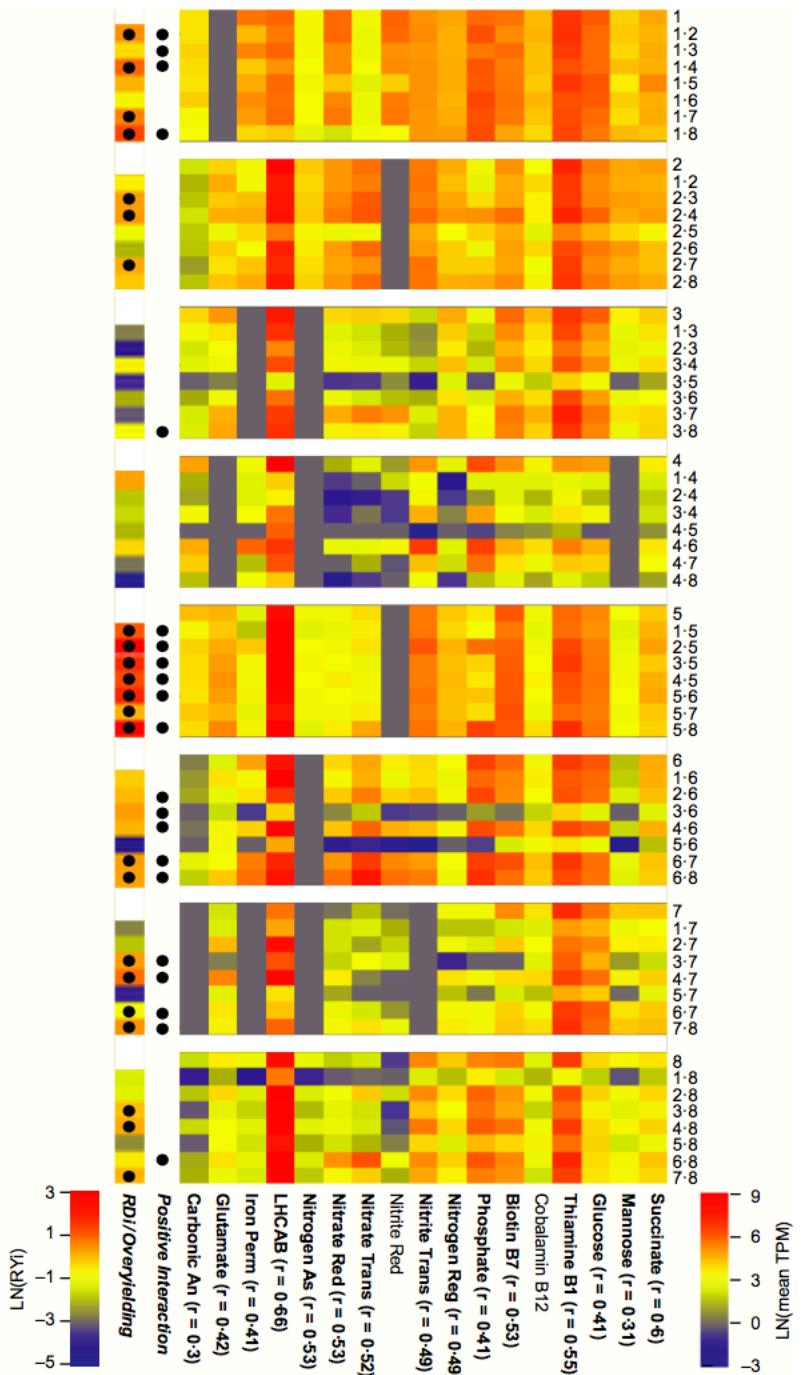
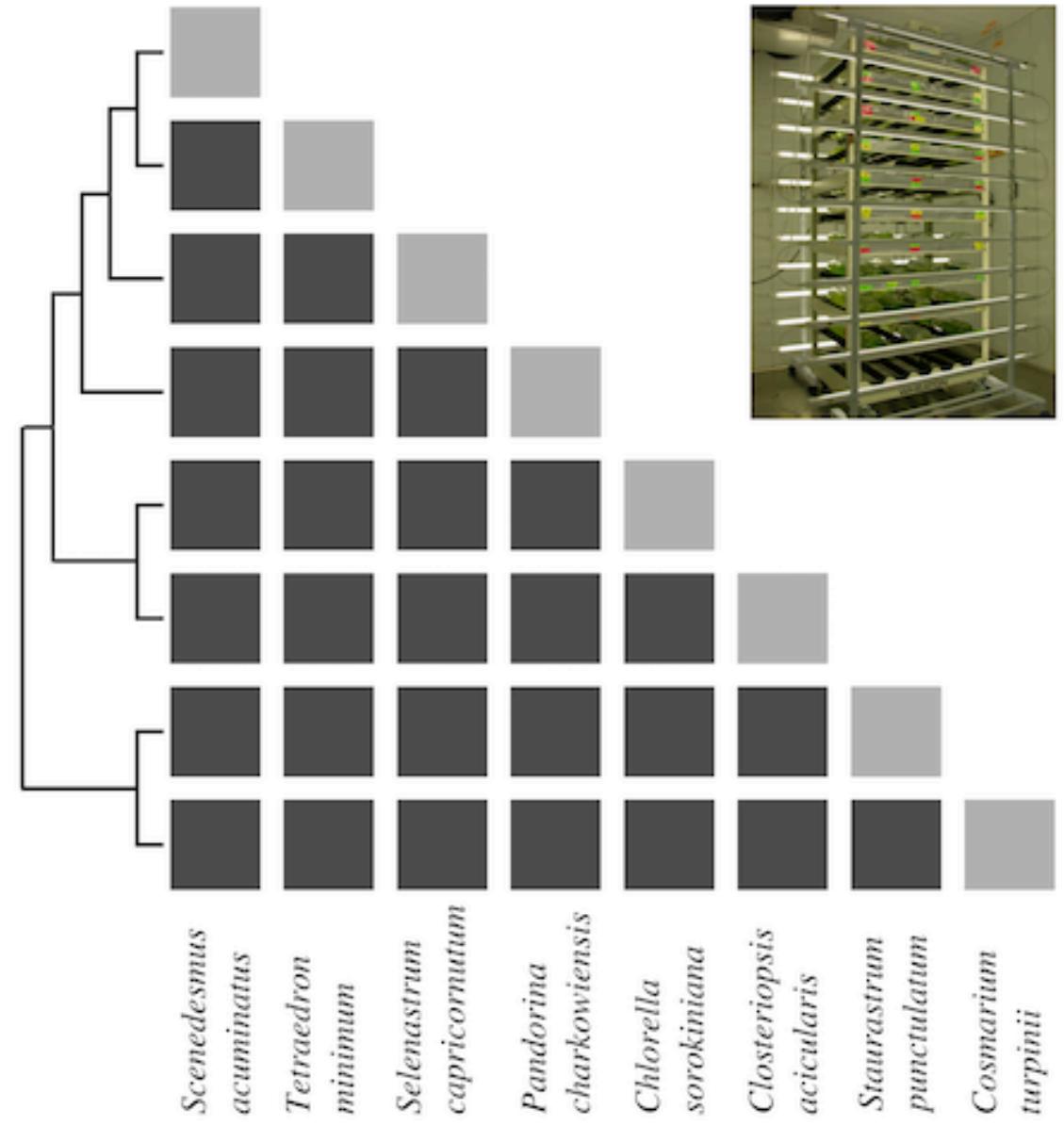
by life technologies™

What can we do with this?



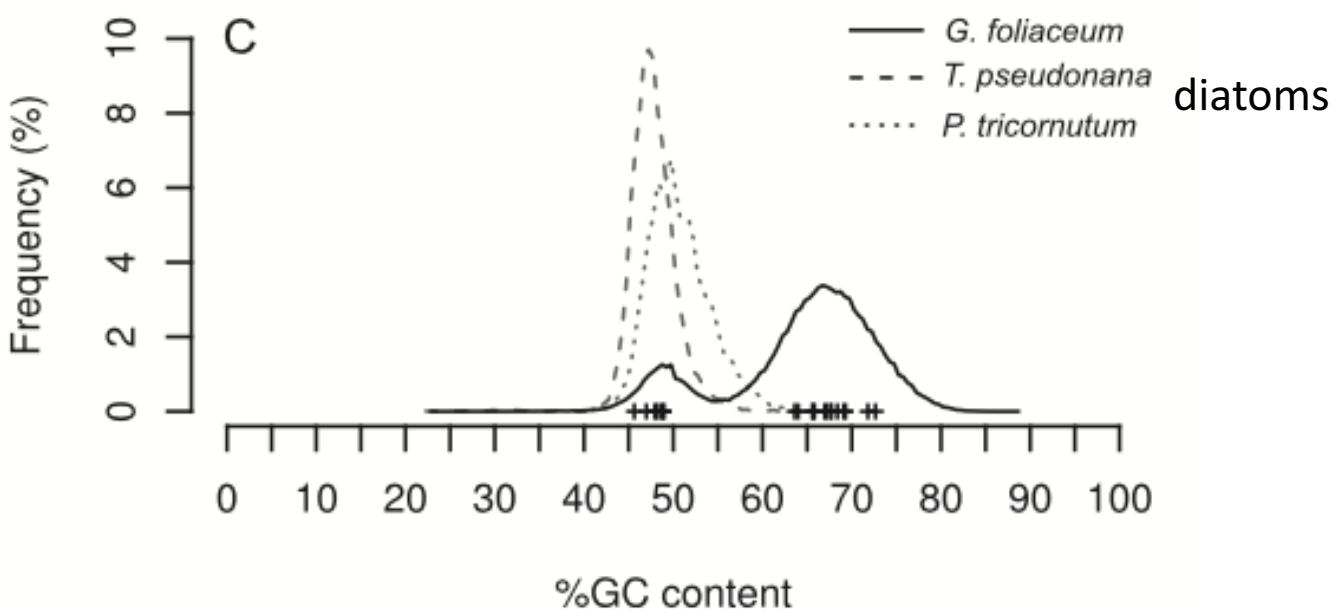
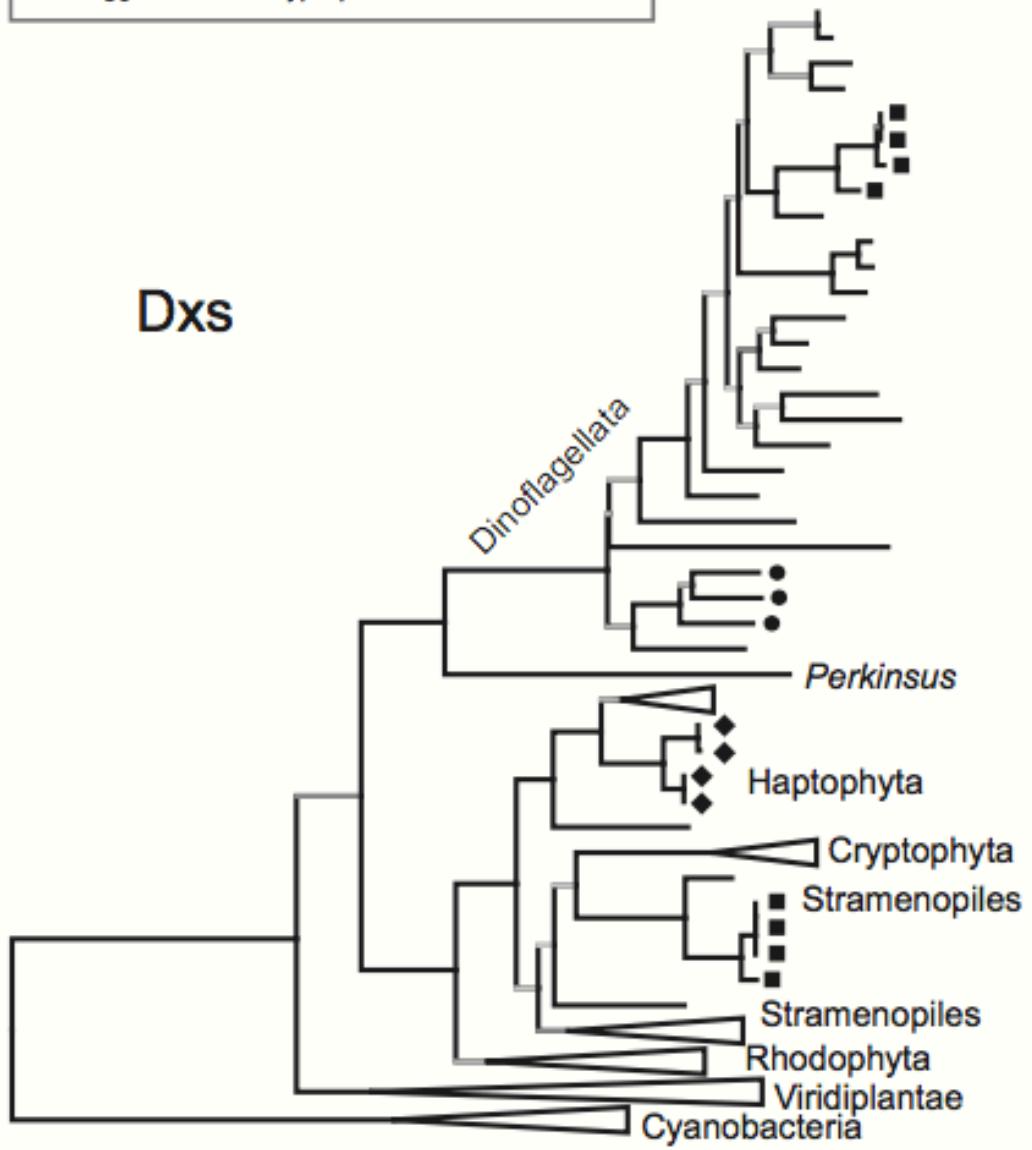


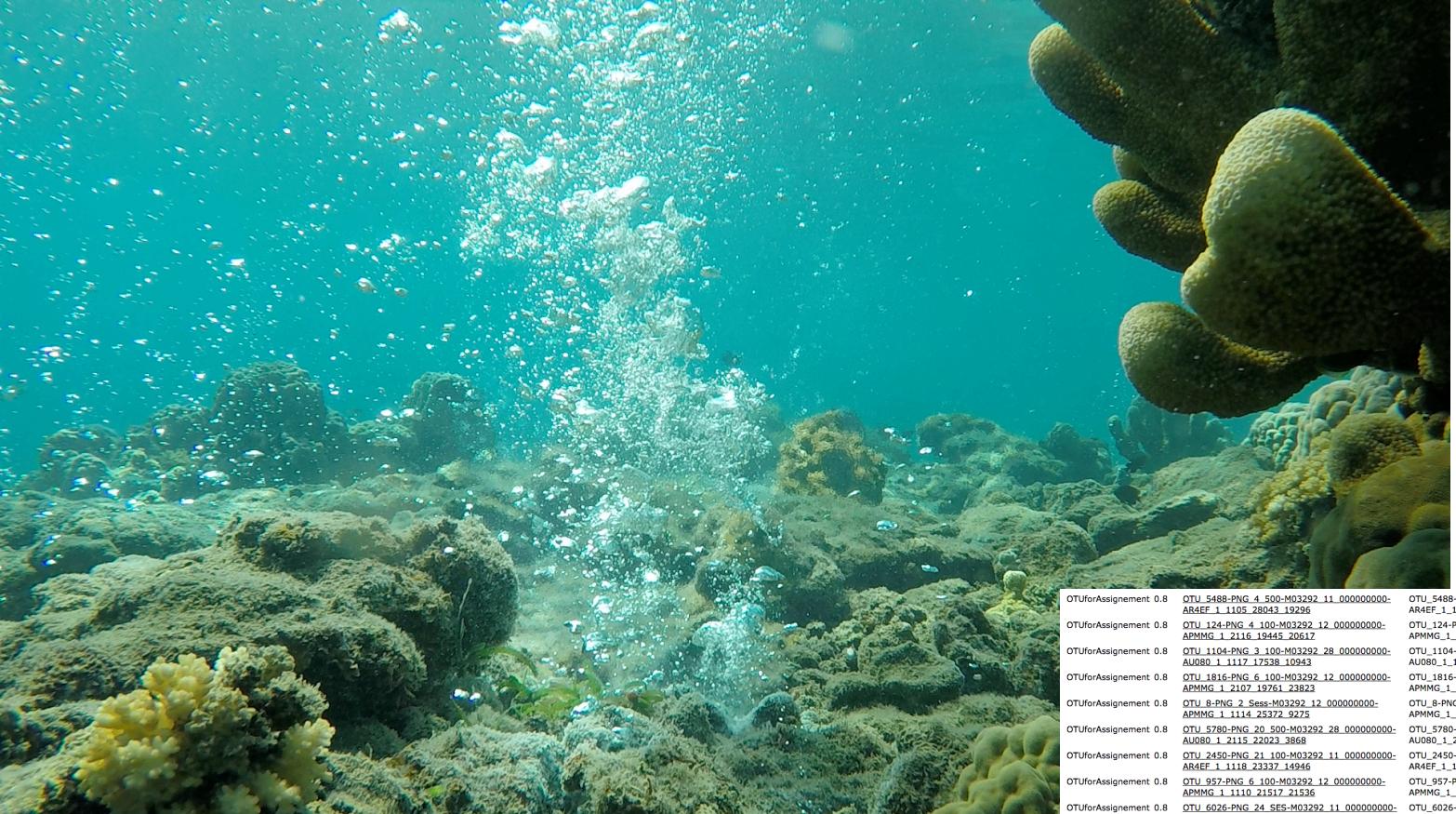






Dxs





OTUforAssignment_0.8	OTU_5488-PNG_4_100-M03292_11_000000000-	OTU_5488-PNG_4_100-M03292_11_000000000-	Rhodophyta	Florideophyceae	Ceramiales	Rhodomelaceae	Herposiphonia	NA	43	0.001	
OTUforAssignment_0.8	AR4EE_1_1105_28043_1929	OTU_1_1105_28043_1929	Nematoda	Enoplognathidae	Enoplognathidae	Enoplognathidae	Enoplognathidae	NA	50	NA	
OTUforAssignment_0.8	OTU_124-PNG_4_100-M03292_12_000000000-	OTU_124-PNG_4_100-M03292_12_000000000-	Arthropoda	Malacostraca	Amphipoda	NA	NA	NA	50	NA	
OTUforAssignment_0.8	APMMG_1_2116_19445_20617	APMMG_1_2116_19445_20617	Arthropoda	Malacostraca	Amphipoda	NA	NA	NA	50	NA	
OTUforAssignment_0.8	OTU_1104-PNG_3_100-M03292_28_000000000-	OTU_1104-PNG_3_100-M03292_28_000000000-	Annelida	Polychaeta	Phyllodocida	Syllidae	Syllis	Syllis.gracilis	25	0.001	
OTUforAssignment_0.8	AU080_1_1117_17538_10943	AU080_1_1117_17538_10943	Arthropoda	Maxillopoda	Harpacticoida	NA	NA	NA	50	0.001	
OTUforAssignment_0.8	OTU_1816-PNG_5_100-M03292_12_000000000-	OTU_1816-PNG_5_100-M03292_12_000000000-	Arthropoda	Maxillopoda	Harpacticoida	NA	NA	NA	50	0.002	
OTUforAssignment_0.8	APMMG_1_2107_19761_28232	APMMG_1_2107_19761_28232	Arthropoda	Maxillopoda	Harpacticoida	NA	NA	NA	50	0.001	
OTUforAssignment_0.8	OTU_8-PNG_2_Sess-M03292_12_000000000-	OTU_8-PNG_2_Sess-M03292_12_000000000-	Nemertea	NA	NA	NA	NA	NA	50	0.001	
OTUforAssignment_0.8	APMMG_1_1114_25372_925	APMMG_1_1114_25372_925	Arthropoda	Malacostraca	Amphipoda	NA	NA	NA	50	NA	
OTUforAssignment_0.8	OTU_5780-PNG_20_500-M03292_28_000000000-	OTU_5780-PNG_20_500-M03292_28_000000000-	Arthropoda	Malacostraca	Amphipoda	NA	NA	NA	50	NA	
OTUforAssignment_0.8	OTU_2450-PNG_21_100-M03292_11_000000000-	OTU_2450-PNG_21_100-M03292_11_000000000-	Arthropoda	Malacostraca	Amphipoda	NA	NA	NA	50	0.001	
OTUforAssignment_0.8	AR4EE_1_1118_23337_14946	AR4EE_1_1118_23337_14946	Arthropoda	Malacostraca	Amphipoda	NA	NA	NA	50	0.001	
OTUforAssignment_0.8	OTU_957-PNG_6_100-M03292_12_000000000-	OTU_957-PNG_6_100-M03292_12_000000000-	Arthropoda	Malacostraca	Amphipoda	NA	NA	NA	50	NA	
OTUforAssignment_0.8	APMMG_1_1110_21512_21536	APMMG_1_1110_21512_21536	Arthropoda	Malacostraca	Amphipoda	NA	NA	NA	50	NA	
OTUforAssignment_0.8	OTU_6026-PNG_24_SES-M03292_11_000000000-	OTU_6026-PNG_24_SES-M03292_11_000000000-	Arthropoda	Insecta	Hymenoptera	Megachilidae	Megachile	NA	41	0.001	
OTUforAssignment_0.8	AR4EE_1_1113_17947_8951	AR4EE_1_1113_17947_8951	Arthropoda	Insecta	Hymenoptera	Megachilidae	Megachile	NA	41	0.001	
OTUforAssignment_0.8	OTU_5343-PNG_3_Sess-M03292_12_000000000-	OTU_5343-PNG_3_Sess-M03292_12_000000000-	Arthropoda	Bacillariophyta	Bacillariophyceae	Naviculales	NA	NA	NA	15	0.001
OTUforAssignment_0.8	APMMG_1_1104_26602_5648	APMMG_1_1104_26602_5648	Arthropoda	Bacillariophyta	Bacillariophyceae	Naviculales	NA	NA	NA	15	0.001
OTUforAssignment_0.8	OTU_5502-PNG_29_SES-M03292_28_000000000-	OTU_5502-PNG_29_SES-M03292_28_000000000-	Arthropoda	Malacostraca	Amphipoda	NA	NA	NA	50	NA	
OTUforAssignment_0.8	AU080_1_2106_10214_9166	AU080_1_2106_10214_9166	Arthropoda	Malacostraca	Amphipoda	NA	NA	NA	50	NA	
OTUforAssignment_0.8	OTU_5831-PNG_2_Sess-M03292_12_000000000-	OTU_5831-PNG_2_Sess-M03292_12_000000000-	Arthropoda	Malacostraca	Decapoda	NA	NA	NA	11	NA	
OTUforAssignment_0.8	APMMG_1_2109_5744_15552	APMMG_1_2109_5744_15552	Arthropoda	Malacostraca	Decapoda	NA	NA	NA	11	NA	
OTUforAssignment_0.8	OTU_231-PNG_27_100-M03292_28_000000000-	OTU_231-PNG_27_100-M03292_28_000000000-	Arthropoda	Maxillopoda	Harpacticoida	NA	NA	NA	50	0.003	
OTUforAssignment_0.8	AU080_1_2119_5043_13176	AU080_1_2119_5043_13176	Arthropoda	Maxillopoda	Harpacticoida	NA	NA	NA	50	0.003	
OTUforAssignment_0.8	OTU_5070-PNG_5_100-M03292_12_000000000-	OTU_5070-PNG_5_100-M03292_12_000000000-	Arthropoda	Insecta	Hymenoptera	Megachilidae	Megachile	NA	50	0.001	
OTUforAssignment_0.8	APMMG_1_1117_9243_19627	APMMG_1_1117_9243_19627	Arthropoda	Insecta	Hymenoptera	Megachilidae	Megachile	NA	50	0.001	
OTUforAssignment_0.8	OTU_3171-PNG_24_500-M03292_12_000000000-	OTU_3171-PNG_24_500-M03292_12_000000000-	Arthropoda	Polychaeta	Phyllodocida	Syllidae	Branchiosyllis	Branchiosyllis.exilis	50	0.001	
OTUforAssignment_0.8	APMMG_1_1104_16133_16910	APMMG_1_1104_16133_16910	Arthropoda	Polychaeta	Phyllodocida	Syllidae	Branchiosyllis	Branchiosyllis.exilis	50	0.001	
OTUforAssignment_0.8	OTU_5578-PNG_2_100-M03292_12_000000000-	OTU_5578-PNG_2_100-M03292_12_000000000-	Arthropoda	Maxillopoda	Harpacticoida	NA	NA	NA	33	0.002	
OTUforAssignment_0.8	APMMG_1_2102_144515_7550	APMMG_1_2102_144515_7550	Arthropoda	Maxillopoda	Harpacticoida	NA	NA	NA	33	0.002	
OTUforAssignment_0.8	OTU_5772-PNG_24_100-M03292_12_000000000-	OTU_5772-PNG_24_100-M03292_12_000000000-	Arthropoda	Mollusca	NA	NA	NA	NA	6	0.01	
OTUforAssignment_0.8	APMMG_1_1111_22534_17680	APMMG_1_1111_22534_17680	Arthropoda	Mollusca	NA	NA	NA	NA	6	0.01	
OTUforAssignment_0.8	OTU_2666-PNG_5_500-M03292_11_000000000-	OTU_2666-PNG_5_500-M03292_11_000000000-	Arthropoda	Malacostraca	Amphipoda	NA	NA	NA	8	NA	
OTUforAssignment_0.8	AR4EE_1_2110_5346_10308	AR4EE_1_2110_5346_10308	Arthropoda	Malacostraca	Amphipoda	NA	NA	NA	8	NA	
OTUforAssignment_0.8	OTU_481-PNG_19_500-M03292_28_000000000-	OTU_481-PNG_19_500-M03292_28_000000000-	Arthropoda	Malacostraca	Amphipoda	NA	NA	NA	11	NA	
OTUforAssignment_0.8	AU080_1_1108_22931_13092	AU080_1_1108_22931_13092	Arthropoda	Malacostraca	Amphipoda	NA	NA	NA	11	NA	
OTUforAssignment_0.8	OTU_5371-PNG_2_500-M03292_11_000000000-	OTU_5371-PNG_2_500-M03292_11_000000000-	Arthropoda	Gastropoda	Gastropoda	Cypraeidae	Cibularia	Cibularia.exmouhensis	48	0.001	
OTUforAssignment_0.8	AR4EE_1_2108_16392_4475	AR4EE_1_2108_16392_4475	Arthropoda	Gastropoda	Gastropoda	Cypraeidae	Cibularia	Cibularia.exmouhensis	48	0.001	
OTUforAssignment_0.8	OTU_6517-PNG_21_SES-M03292_28_000000000-	OTU_6517-PNG_21_SES-M03292_28_000000000-	Arthropoda	Maxillopoda	NA	NA	NA	NA	50	NA	
OTUforAssignment_0.8	AU080_1_1117_24843_15153	AU080_1_1117_24843_15153	Arthropoda	Maxillopoda	NA	NA	NA	NA	50	NA	
OTUforAssignment_0.8	OTU_5298-PNG_20_SES-M03292_12_000000000-	OTU_5298-PNG_20_SES-M03292_12_000000000-	Arthropoda	Demospongiae	Astrophorida	Geodidae	Geodia	NA	50	NA	
OTUforAssignment_0.8	APMMG_1_2107_9539_6873	APMMG_1_2107_9539_6873	Arthropoda	Demospongiae	Astrophorida	Geodidae	Geodia	NA	50	NA	
OTUforAssignment_0.8	OTU_6351-PNG_3_100-M03292_28_000000000-	OTU_6351-PNG_3_100-M03292_28_000000000-	Arthropoda	Maxillopoda	Calanoida	Clausocalanidæ	Clausocalanus	Clausocalanus.parapergens	50	0.001	
OTUforAssignment_0.8	AU080_1_1104_18522_6951	AU080_1_1104_18522_6951	Arthropoda	Maxillopoda	Calanoida	Clausocalanidæ	Clausocalanus	Clausocalanus.parapergens	50	0.001	
OTUforAssignment_0.8	OTU_5938-PNG_11_SES-M03292_28_000000000-	OTU_5938-PNG_11_SES-M03292_28_000000000-	Arthropoda	Maxillopoda	Rhodophyta	Florideophyceae	Gastropoda	Gastropoda	NA	0.001	
OTUforAssignment_0.8	AU080_1_1112_10265_4296	AU080_1_1112_10265_4296	Arthropoda	Maxillopoda	Rhodophyta	Florideophyceae	Gastropoda	Gastropoda	NA	0.001	
OTUforAssignment_0.8	OTU_3060-PNG_24_100-M03292_12_000000000-	OTU_3060-PNG_24_100-M03292_12_000000000-	Arthropoda	Mollusca	Gastropoda	Gastropoda	Gastropoda	Gastropoda	5	0.861	
OTUforAssignment_0.8	APMMG_1_2103_17450_9288	APMMG_1_2103_17450_9288	Arthropoda	Mollusca	Gastropoda	Gastropoda	Gastropoda	Gastropoda	5	0.861	
OTUforAssignment_0.8	OTU_5638-PNG_21_100-M03292_11_000000000-	OTU_5638-PNG_21_100-M03292_11_000000000-	Arthropoda	Malacostraca	Amphipoda	NA	NA	NA	10	NA	
OTUforAssignment_0.8	AR4EE_1_2107_11621_13508	AR4EE_1_2107_11621_13508	Arthropoda	Malacostraca	Amphipoda	NA	NA	NA	10	NA	
OTUforAssignment_0.8	OTU_3073-PNG_3_500-M03292_11_000000000-	OTU_3073-PNG_3_500-M03292_11_000000000-	Arthropoda	Malacostraca	Decapoda	NA	NA	NA	50	0.001	
OTUforAssignment_0.8	AR4EE_1_2117_6447_20994	AR4EE_1_2117_6447_20994	Arthropoda	Malacostraca	Decapoda	NA	NA	NA	50	0.001	
OTUforAssignment_0.8	OTU_5520-PNG_4_100-M03292_12_000000000-	OTU_5520-PNG_4_100-M03292_12_000000000-	Arthropoda	Malacostraca	Decapoda	NA	NA	NA	34	0.001	
OTUforAssignment_0.8	APMMG_1_1113_26984_19057	APMMG_1_1113_26984_19057	Arthropoda	Malacostraca	Decapoda	NA	NA	NA	34	0.001	

Part I. Introduction (week 1)

- 1) Course outline and installation of Linux on laptops using virtual machines

Part II. Introduction to Linux and the shell (weeks 2-3)

- 1) Command-line operations and handling text files in the shell
- 2) Pattern searches using regular expressions and simple pipelines
- 3) basics of shell scripting
- 4) Programming concepts (flow control with loops and conditionals)
- 5) Genetic sequence, multiple sequence alignment, and phylogenetic tree file formats
- 6) Installing software (compiling from source and package managers)

Part III. Programming in Python (weeks 4-5) → MIDTERM

- 1) Why use Python?
- 2) Reading and writing files
- 2) Variables, arrays, lists, and dictionaries (hash tables)
- 3) Mathematical and logical operators
- 4) Modules and libraries
- 5) Debugging scripts

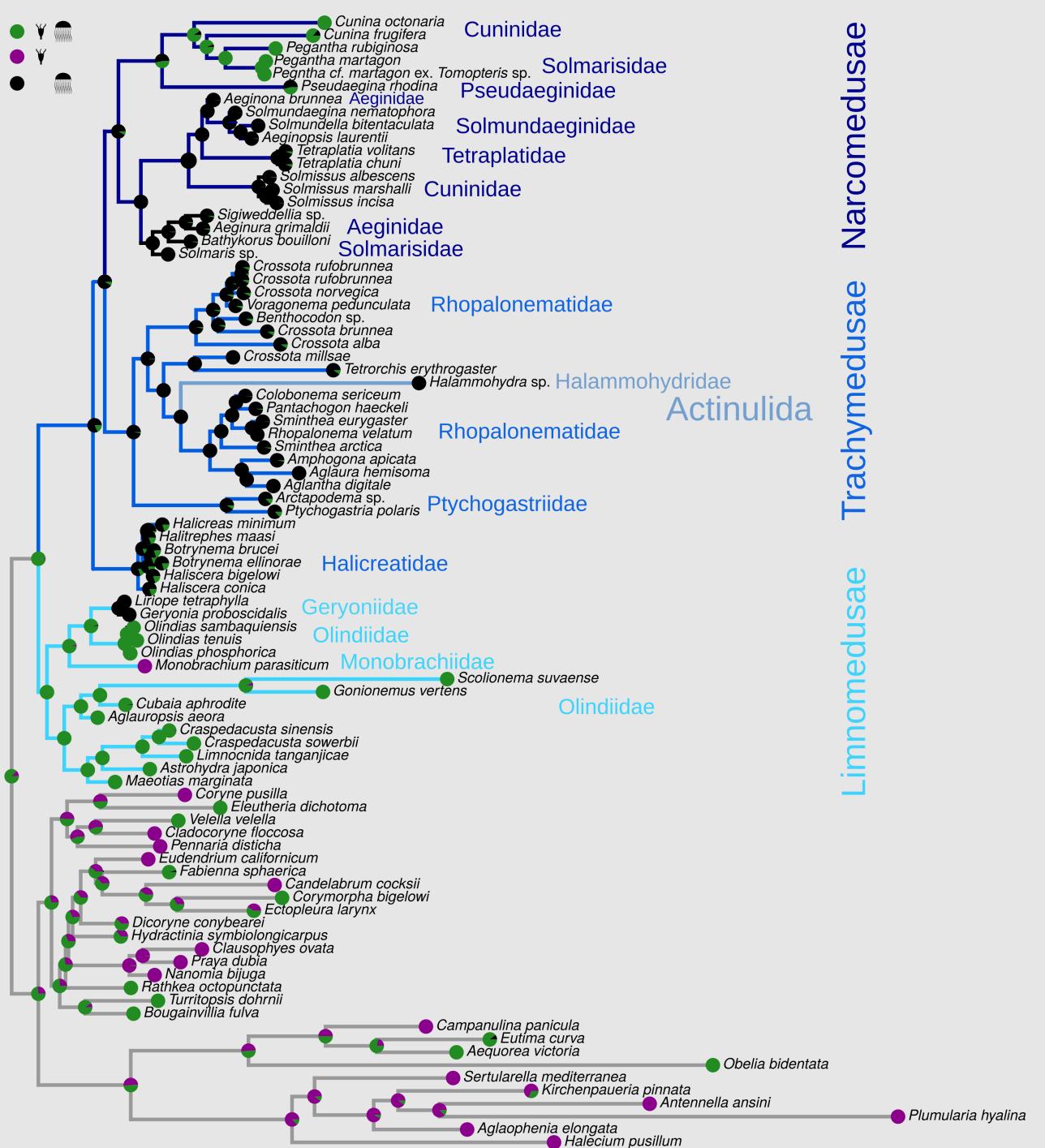
Part IV. Theory and practice of bioinformatic analyses (week 6-14) → FINAL EXAM

- 1) Working on remote servers
- 2) Sequence similarity searches using BLAST and hidden markov models
- 3) Public sequence databases
- 4) Finding related sequences using clustering
- 5) multiple sequence alignments
- 6) inferring phylogenetic trees using sequence data
- 7) identifying genetic variation in populations using single nucleotide polymorphisms (SNPs)
- 8) Genome (and transcriptome) assembly
- 9) Relational databases
- 10) Version control and backing up your work

- Meeting time
- Laptops
- Evaluation
- Anything else?

Email: Bastian.Bentlage@gmail.com

Office Hours: by appointment



Hydrodolina Traehylinia

