Pandorina morum genome assembly, annotation, and analysis

by

Matthew George Heffel

B.S., Kansas State University, 2019

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Biology
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2020

Approved by:

Major Professor
Bradly J. S. C. Olson

Copyright

Abstract

The evolution of multicellularity is a major evolutionary transition that leads to increased organismal complexity and has occurred various times in multiple domains of life. Despite its common occurrence, the evolution of multicellularity is not yet well understood largely due to genetic signatures being lost due to deep divergence between unicellular and multicellular lineages. The volvocine algae have recently made the transition to multicellularity (200 MYA) and cover a large range of morphologies, including unicellular *Chlamydomonas*, undifferentiated multicellular *Gonium* (8-16 cells), multicellular isogamous *Pandorina* (8-16 cells), multicellular isogamous *Yamagishiella* (32 cells), multicellular anisogamous *Eudorina* (32 cells), and multicellular differentiated *Volvox* with germ-soma division of labor (>500 cells). Using modern sequencing techniques, here, the genome of *Pandorina morum* is sequenced, assembled, and annotated. Brief comparative genomics work shows gene orthology to related volvocine species as well as a common trend of progressive gene loss occurring at a higher rate than gene gain and organismal complexity increases. This work opens the door to targeted mutagenesis as transgenic work using both foreign genes in *Pandorina* and genes from *Pandorina* in related species.

Table of Contents

List of Figures

List of Tables

Chapter 1 - Introduction

The evolution of multicellularity has independently been observed more than 25 times in multiple domains of life and qualifies as one of the most important evolutionary transitions in life history[1]. However, the mechanisms behind this process are not yet well understood. Multicellularity represents a prerequisite for complex organismal body plans[1,2] and deciphering the enigmatic processes behind how and why multicellularity evolved will certainly aid in the understanding of cellular differentiation, sexual reproduction, and ontogenesis.

The multicellular state is no recent innovation, with fossil evidence unmistakably showing the existence of multicellular organisms dating ~3.5 billion years ago (BYA); further evidence shows a species of cyanobacteria had already developed cell type differentiation by ~2 BYA[3]. Such a large amount of time passing since the emergence of these organisms makes it almost impossible to study due to blurring of genetic signals by genetic drift and other evolutionary processes. An ideal model system to study the evolution of multicellularity would contain several desirable features: 1) an extant collection of closely related organisms ranging in complexity from single cellular to multicellular with complete division of labor, 2) is mostly comprised of organisms that can be readily obtained from nature and maintained and stored in a laboratory, 3) has already been studied in enough detail to distinguish features that unify the group and variants that distinguish each member, 4) generalizations can be made about occurrences in the natural world and how they influence members in the group, 5) is of such recent origin that the footprint of genetic changes are still relatively unobstructed by genetic drift throughout the genomes, 6) the group contains multiple instances of each transition in levels of complexity occurring, and 7) is capable of being manipulated with modern genetic methods[4].

The volvocine algae lineage is one such group that contains the desirable features of a model system, making it an ideal candidate for studying multicellular evolution. Members of the Order volvocales range in size from 10 µm to 5 mm and exhibit a wide range of developmental complexity, with each member of containing 2^n Chlamydomonas-like cells, where species differ in the value of *n*, degree of extracellular matrix expansion, and ability of each cell to participate in reproduction[4,5].Major volvocine species of interest are unicellular *Chlamydomonas*, undifferentiated multicellular *Gonium* (8-16 cells), multicellular isogamous *Pandorina* (8-16 cells), multicellular isogamous *Yamagishiella* (32 cells), multicellular anisogamous *Eudorina* (32 cells), and multicellular differentiated *Volvox* with germ-soma division of labor (>500 cells); however, many intermediate species have been discovered[6,7]. A relatively recent (~200 MYA) transition to multicellularity preserved the genetic signatures throughout the lineage's species[8]. Of the major volvocine species mentioned, all but *Yamagishiella*, have methods published for stable insertion and expression of a transgene into the nuclear genome[9,10,11,12,13].

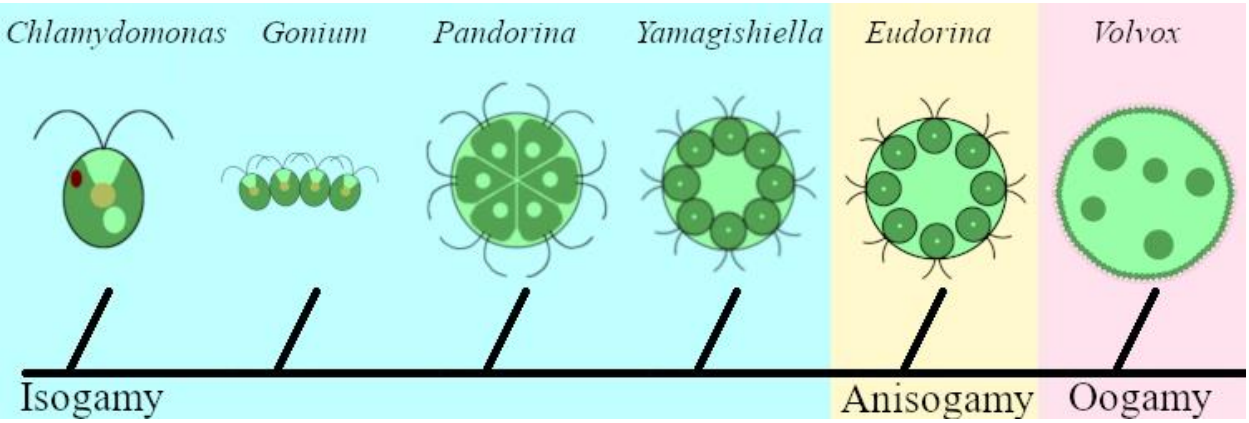

Figure 1: Diagram showing the stepwise evolution in the volvocales from isogamy to anisogamy to oogamy.

Pandorina morum is of intermediate developmental complexity between *Chlamydomonas* and *Volvox* and contains 8-16 cells held together by their bases at a centroid, anterior-posterior polarity (cell eyes spots are larger near the front of the organism and young

colonies bust out of cells in the back by vegetative reproduction), and two external facing flagella per cell, and has recently been shown capable of stably expressing nuclear transformation[11]. Distinct features of *Pandorina morum* are: the cells fill almost the entirety of the colonies volume, it is the smallest volvocine alga to exhibit a colonial boundary (the most conserved morphological trait of the volvocales), and it is isogamous (gametes are almost identical in morphology with the exception of plus/minus mating type allele expression)[4,]. These features make *Pandorina morum* an important member of the volvocine algae and an ideal candidate for genome sequencing, assembly, and annotation.

This work presents the newly assembled and annotated genome and transcriptome of *Pandorina morum* as well as a gene orthology report. *Pandorina morum* along with related species *Chlamydomonas reinhardtii*, *Gonium pectorale*, *Yamagishiella unicocca*, *Eudorina sp.*, and *Volvox carteri*, was subjected to comparative genomics analysis. This analysis supports, and is reinforced by related studies[14], that gene loss events progress during the transition to multicellularity and organismal complexity at a rate that is significantly greater than that of gene gain events. As gene loss has become prevalent as a signature of evolutionary change and complex development[15,16], these results suggest that the transition to multicellularity is driven by the removal of non-essential genes and the gain of novel gene interactions rather than entirely new genes themselves.

Chapter 2 - Methods

Strains Information

The strain sequenced and assembled here is *Pandorina morum 5*, a gift from Hisayoshi Nozaki, University of Tokyo, Japan. The *Chlamydomonas reinhardtii* 5.5 and *Volvox carteri* 2.1 genomes and transcriptomes were collected from Phytozome[17,18]. The remaining genomes were collected from GenBank being published previously[19,20].

Culture Conditions

Cultures were grown in SVM media REF under 400 µmol m-1 sec -1 of LED light with 0.5% CO2 bubbling at 30ºC. Cultures were grown with a 14h light, 10h dark light cycle.

Production of axenic *Pandorina morum*

Cultures of *Pandorina morum* were initially decontaminated by collecting colonies by centrifugation at 1000 **g** for 5 min at 25ºC and removal of the growth media. These colony pellets were overlaid with 50% Percoll® (Millipore-Sigma) in SVMA followed by illumination with 100 uE of light at the top of the tube. The top 1 mL of colonies that were able to phototax through this were then plated for single colonies on SVMA agar plates supplemented with 50 ug/mL carbenicillin, 50 ug/mL cefotaxime, 10 ug/mL trimethoprim and 20 ug/mL chloramphenicol. Cultures from single isolates were established in SVM.

DNA and RNA Isolation

Pandorina morum nucleic acids were extracted as follows:

Cells were harvested by centrifugation at 1000 **g** for 5 min at 25ºC and DNA extraction for all Illumina DNA sequencing was prepared as described for *Volvox* previously[18]. For jump

libraries, DNA was gel extracted with a Pippin prep system (Sage science) prior to library preparation. For PacBio sequencing, high molecular weight DNA was extracted using the Qiagen MagAttract HMW DNA kit (Qiagen, Cat. No. 67563).

For RNA sequencing, synchronously grown *Pandorina* cultures were collected hourly over a 24h light cycle (one complete life cycle). Tissue was collected by centrifugation at 1000 **g** for 5 min at 25ºC and flash frozen in liquid nitrogen. Tissue was then ground into a fine powder with a mortar and pestle under liquid nitrogen. Total RNA was extracted with a Zymo Plant Quick Plant RNA kit (Zymo Research, Cat. No. R2024). All RNA for sequencing was quality controlled by analysis on an Agilent Nano RNA 6000 kit (Agilent Technologies) where all RNA samples had a "RNA integrity number" (RIN) of 9.0 or greater.

DNA and RNA Sequencing

All Illumina DNA sequencing libraries were prepared with the TruSeq DNA PCR-Free High Throughput Library Prep Kit (Illumina, Cat. No. 20015963) using a barcoded approach. RNA-Seq libraries were prepared with the TruSeq Stranded mRNA Library Prep (Illumina, Cat. No. 20020595) using a barcoded approach. Sequencing and library preparation was done by the Genome Sequencing Core at the University of Kansas. All Illumina sequencing was paired-end, and was performed on HiSeq2500 equipment using FastRun mode. PacBio sequencing was performed at Genewiz on a PacBio Sequel using a SMRT Cell 1M.

In silico **decontamination of bacterial reads**

Despite antibiotic decontamination, *Pandorina morum* cultures showed evidence of being xenic after sequencing; hence, bacterial contamination reads needed to be identified and removed. The Illumina paired-end read data initially contained reads in separate groups by

sequencing run type, different insert sizes, short, 1kb, 2kb, 5kb, 10kb, and 15kb, as well as a no gel sequencing run. These paired-end read groups were assembled using the *de novo*, parallel, paired-end sequence assembler, ABySS[21] version 2.1.5 for *k*-mers 21 through 89 at increments of 4. The resulting scaffolds were then BLASTed[22,23] against NCBI's non-redundant nucleotide database (updated October 27, 2019). The BLAST output files were parsed for bacterial species' and their GenBank[24] reference IDs were collected. The bacterial species' genomes were downloaded using Biopython[25] version 1.74's Entrez module. The sets of paired-end read data was then mapped to the bacterial reference genomes using Bowtie2[26] version 2.3.4.2 and all of the reads that mapped to a bacterial reference with perfect alignment were discarded.

Assembly and Transcriptome Annotation

The remaining non-bacterial paired-end reads were merged into one group and reassembled using ABySS using *k*-mer sizes 21 through 89, again incrementing by 4. The output fasta files of scaffolds that resulted from this were modified to mimic long reads by assigning a perfect quality score to all the scaffolds, making it a fastq file. This allowed the scaffolds to be merged with Pacific Biosciences long reads. The long reads and scaffolds were assembled to span gaps and connect separate scaffolds using the single molecule sequences assembler, Canu[27] version 1.8. RNA sequencing data was then mapped to the final contig assembly from Canu using TopHat[28] version 2.2.1 to generate gene features (introns exons, UTRs). The Augustus software package[29] version 3.3.3 used the gene features and the assembled genome to identify transcripts.

Orthologous Group Identification and Analysis

The transcriptome generated from Augustus was analyzed for orthology to related species in the Volvocine algae lineage, *Chlamydomonas reinhardtii*, *Gonium pectorale*, *Yamagishiella unicocca*, *Eudorina sp.*, and *Volvox carteri*, using OrthoMCL[30] version 2.0.9 and MySQL version 8.0.19. PosiGene[31] version 0.1 was used on the same species, as well as *Chlorella sp.* as an outgroup, to determine distances to the latest common ancestor (LCA). The gene counts in orthologous groups per species from OrthoMCL as well as the LCA distances from PosiGene were used to create a linear regression for determining orthologous groups under contraction or expansion. Quartile analysis was done on the regression to identify orthologous groups found to be significantly contracting or expanding; outliers were considered as significantly contracting or expanding if they were beyond the first quartile minus 1.5 time the interquartile range or beyond the third quartile plus 1.5 times the interquartile range, respectively. The contracting and expanding orthologous groups were then functionally annotated by known members in the group from either *Chlamydomonas reinhardtii*, or *Volvox carteri* entries on Phytozome[17,18].

Chapter 3 - Results

	Chlamydomonas	*Gonium*	*Pandorina morum*	*Yamagishiella* Minus	*Yamagishiella* Plus	*Eudorina* Female	*Eudorina* Male	*Volvox*
Genome Length	111,100,715	148,806,172	140,344,465	140,837,241	134,234,618	184,032,255	168,620,790	131,163,211
N50	7,783,580	1,267,136	136,792	547,037	666,310	564,035	377,357	2,599,759
Coding Genome Size	43,185,003	26,339,654	28,783,307	42,867,180	43,250,373	40,243,092	43,655,475	32,509,373
% G and C content	64.1	64.5	62.8	62.9	62.8	63.4	63.2	56
Protein Coding Genes	17737	17948	15976	18180	18416	20744	22924	14247
Total Transcripts	19526	17984	16542	30755	31705	32233	38492	16075
Average Exons Per Transcript	8.62	6.94	8.37	7.06	6.93	6.38	5.73	7.91
Average Exon length	260.99	211.04	216.8	205.93	205.37	207.78	211.58	254.39
Average Intron length	279.17	408.89	458.7	348.37	348.28	396.73	382.3	399.5

Table 1: Genome assembly comparisons for the volvocine algae lineage

Table 1: Genome assembly comparisons for the volvocine algae lineage

Assembly and Annotation

Filtering out bacterial reads from the paired end reads resulted in 230,008,156 (69.92%) paired end reads from the original 328,965,000. Assembling these remaining reads yielded scaffolds with a peak N50 of 16,002 from *k*-mer 77. The maximum length scaffold assembled was 186,643 base pairs (bp) long and was in the *k*-mer 73 assembly. Upon being reassembled with the 1,111,336 PacBio long reads data, the scaffolds generated 2,399 contigs with an N50 of 136,792, a maximum contig length of 1,379,579, and a total genome length of 140,344,465 (140.3 mb) with 62.8 percent GC content. Tophat mapped 256,130,731 RNA sequencing reads to the newly assembled genome and Augustus identified 16,542 transcripts to 15,976 genes using

the Tophat mapping and the assembled *Pandorina* genome. The total length of the coding regions withing the genome is 28.7 mb, which is equal to 20.5 percent of the entire genome.

Orthologous Group Analysis and Comparison

OrthoMCL generated 15,204 total orthologous groups between the 6 species, *Chlamydomonas reinhardtii*, *Gonium pectorale*, *Pandorina morum*, Yamagishiella *unicocca*, *Eudorina sp.*, and *Volvox carteri*. Of these 15,204 orthologous groups, 7,155 (47.06%) were shared among all 6 genomes, 9227 (59.95%) were shared by at least 5 genomes, and 2744 (18.05%) were species specific. Of the 15,976 genes identified in *Pandorina*, OrthoMCL assigned 14,152 (88.86%) genes into 10,251 unique orthologous groups. The LCA distances from PosiGene are shown in Figure 2.

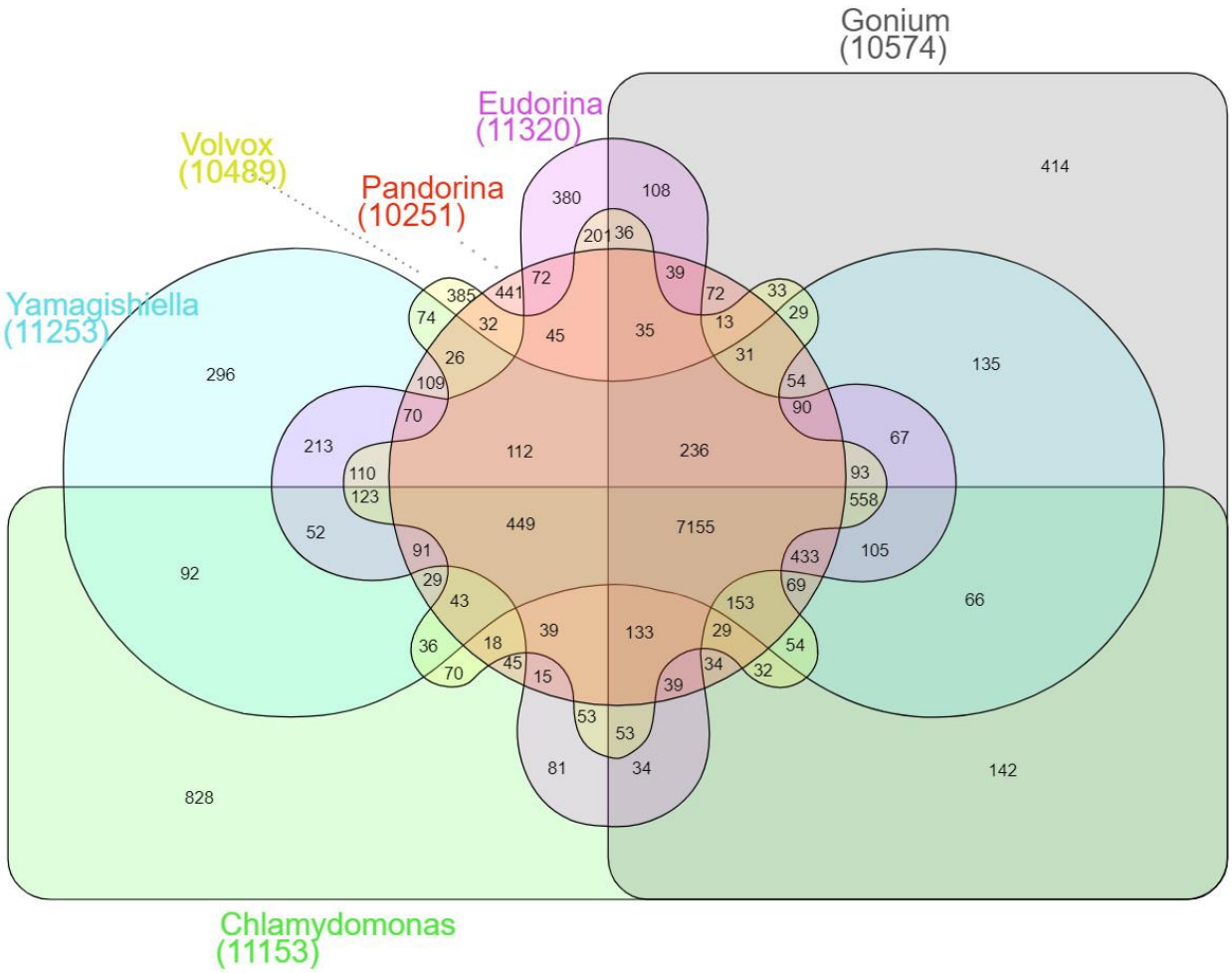

Figure 2: Venn Diagram of orthologous groups containing genes from each species[35]

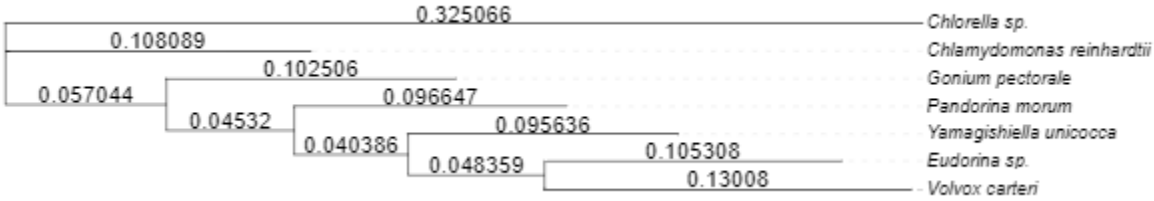

Figure 3: Distances to latest common ancestor calculated using Posigene shown in a phylogenetic tree[9]

Quartile analysis showed 330 gene families to be significantly contracting while were shown to 187 to be significantly expanding. This compares to similar work[14] using only *Chlamydomonas*, *Gonium*, *Yamagishiella*, *Eudorina*, and *Volvox*, which showed 229 gene families to be under significant contraction and 146 under significant expansion. Of these two quartile analyses, 221 families overlapped in the contracting group and 90 families overlapping in the expanding group.

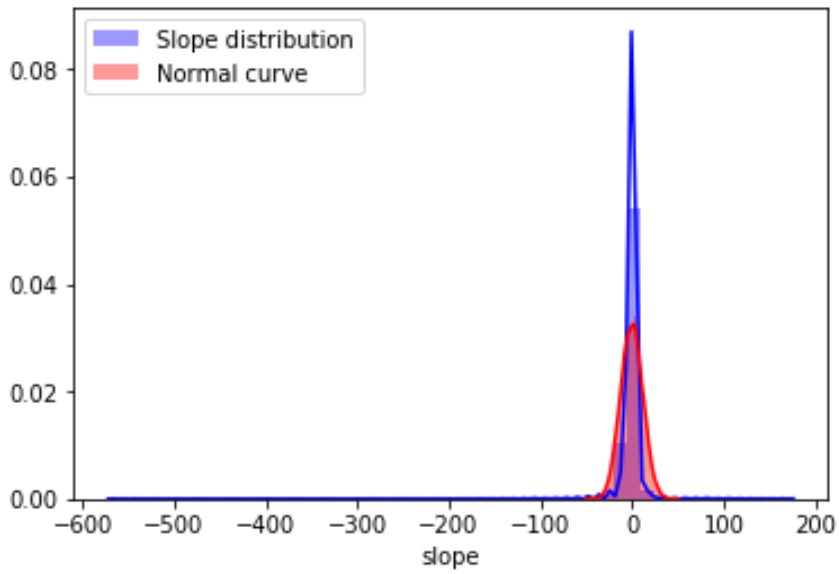

Figure 4: Orthologous group expansion/contraction slope compared to a normal curve

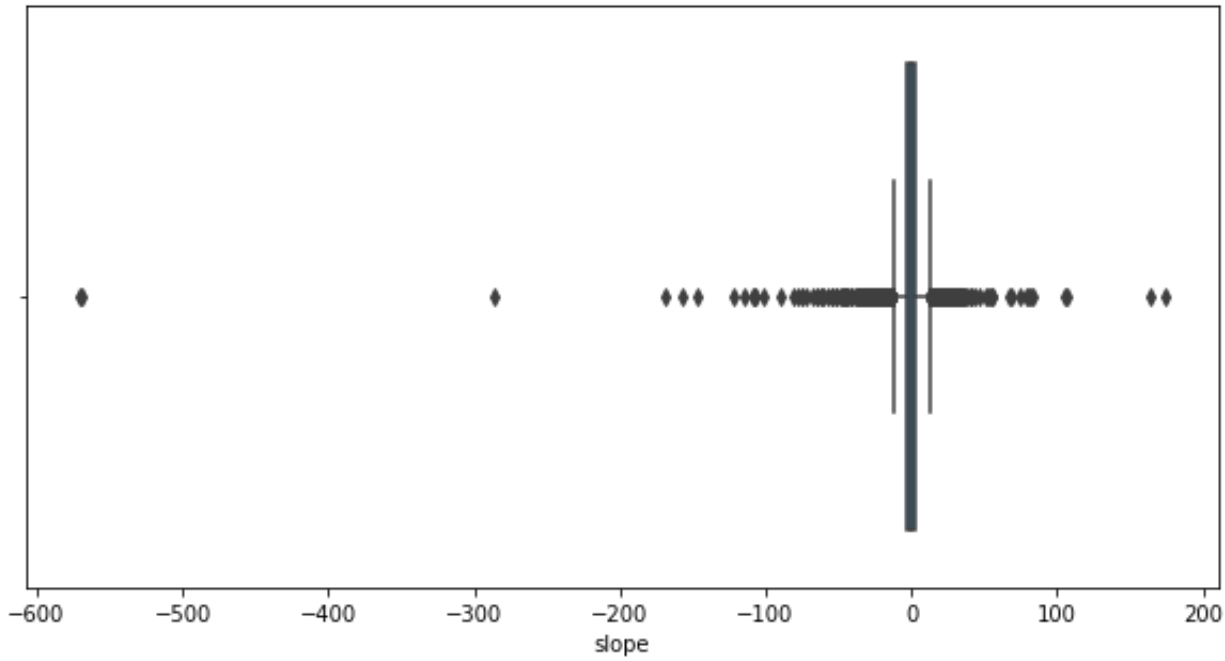

Figure 5: Diamonds represent orthologous groups that were 1.5*interquartile range beyond the first or third quartile

Chapter 4 - Discussion

Here the first version of the *Pandorina morum* nuclear genome is presented. It is known that despite morphological differences, the genomes of the volvocine algae are highly similar, indicating that few genetic alteration are required for multicellularity to occur[17,18]. *Pandorina morum* is the sixth species to be assembled and annotated in the volvocine algae and the addition of *P. morum* to current analysis continues the trend of gene loss being more prevalent than gene gain in the evolution of complexity in multicellularity.

Theoretical work has called a need for greater understanding in the origin of multicellularity and development of related morphological complexities in an integrated phylogenetic group[32,33,34,13]. The volvocine algae lineage has an incredibly recent transition to multicellularity, shows stepwise acquisition of more complex developmental features, and may well be our best chance at laying the groundwork for understanding the foundations of multicellular evolution.

Three major phases in morphological innovation of the volvocales, cell cycle regulation, increased organismal size, and germ/soma cell differentiation have had their genetic pathways identified with the sequencing of the *Chlamydomonas*, *Gonium*, and *Volvox* genomes[19]. As new species are sequenced and annotated more of the story of multicellular evolution in the volvocine algae lineage unfolds, allowing intricate dissection of multiple related genomes to further identify causative agents in more specific and minor innovations leading to multicellularity and increased complexity.

The assembled and annotated genome of *Pandorina morum* allows for future mutagenesis and transgenic work to be done with increased precision to study and better

understand the mechanisms behind the developmental complexity in *Pandorina* and the related volvocine algae. This includes gene knockout or replacement with an ortholog from a related alga to inspect loss of function impact or isoform functional discrepancies. Continued assembly and annotations of more volvocales genomes and mutant strains of those already genetically sequenced as well as further omics sequencing (epigenetics, epitranscriptomics, proteomics…) will set the stage for complex machine learning analysis to be done on the volvocine lineage which will hopefully uncover more key elements in the transition to multicellularity that were otherwise unrecognizable to the human eye.

References

1 Grosberg, R. K., & Strathmann, R. R. (2007). The Evolution of Multicellularity: A Minor Major Transition? *Annual Review of Ecology, Evolution, and Systematics, 38*(1), 621-654. doi:10.1146/annurev.ecolsys.36.102403.114735

2 Knoll, A. H. (2011). The Multiple Origins of Complex Multicellularity. *Annual Review of Earth and Planetary Sciences, 39*(1), 217-239. doi:10.1146/annurev.earth.031208.100209

3 Schopf, J. W. (1974). Paleobiology of the Precambrian: The Age of Blue-Green Algae. *Evolutionary Biology*, 1-43. doi:10.1007/978-1-4615-6944-2_1

4 Kirk, D. L. (2005). *Volvox Molecular-Genetic Origins of Multicellularity and Cellular Differentiation*. Cambridge, U.K.: Cambridge University Press.

5 Hanschen, E. R., Herron, M. D., Wiens, J. J., Nozaki, H., & Michod, R. E. (2018). Multicellularity Drives the Evolution of Sexual Traits. *The American Naturalist, 192*(3). doi:10.1086/698301

6 Kirk, D. L. (2005). A twelve-step program for evolving multicellularity and a division of labor. *BioEssays, 27*(3), 299-310. doi:10.1002/bies.20197

7 Coleman, A. W. (2012). A Comparative Analysis Of The Volvocaceae (Chlorophyta)1. *Journal of Phycology*, *48*(3), 491-513. doi:10.1111/j.1529-8817.2012.01168.x

8 Herron, M. D., Hackett, J. D., Aylward, F. O., & Michod, R. E. (2009). Triassic origin and early radiation of multicellular volvocine algae. *Proceedings of the National Academy of Sciences, 106*(9), 3254-3258. doi:10.1073/pnas.0811205106

9 Mayfield, S. P., & Kindle, K. L. (1990). Stable nuclear transformation of Chlamydomonas reinhardtii by using a C. reinhardtii gene as the selectable marker. *Proceedings of the National Academy of Sciences, 87*(6), 2087-2091. doi:10.1073/pnas.87.6.2087

10 Lerche, K., & Hallmann, A. (2009). Stable nuclear transformation of Gonium pectorale. *BMC Biotechnology, 9*(1), 64. doi:10.1186/1472-6750-9-64

11 Lerche, K., & Hallmann, A. (2014). Stable nuclear transformation of Pandorina morum. *BMC Biotechnology, 14*(1). doi:10.1186/1472-6750-14-65

12 Lerche, K., & Hallmann, A. (2013). Stable nuclear transformation of Eudorina elegans. *BMC Biotechnology, 13*(1). doi:10.1186/1472-6750-13-11

13 Schiedlmeier, B., Schmitt, R., Muller, W., Kirk, M. M., Gruber, H., Mages, W., & Kirk, D. L. (1994). Nuclear transformation of Volvox carteri. *Proceedings of the National Academy of Sciences, 91*(11), 5080-5084. doi:10.1073/pnas.91.11.5080

14 Jimenez-Marin et.al., in progress

15 Albalat, R., & Cañestro, C. (2016). Evolution by gene loss. *Nature Reviews Genetics, 17*(7), 379-391. doi:10.1038/nrg.2016.39

16 Kvitek, D. J., & Sherlock, G. (2013). Whole Genome, Whole Population Sequencing Reveals That Loss of Signaling Networks Is the Major Adaptive Strategy in a Constant Environment. *PLoS Genetics, 9*(11). doi:10.1371/journal.pgen.1003972

17 Merchant, S. S., Prochnik, S. E., Vallon, O., Harris, E. H., Karpowicz, S. J., Witman, G. B., . . . Grossman, A. R. (2007). The Chlamydomonas Genome Reveals the Evolution of Key Animal and Plant Functions. *Science, 318*(5848), 245-250. doi:10.1126/science.1143609

18 Prochnik, S. E., Umen, J., Nedelcu, A. M., Hallmann, A., Miller, S. M., Nishii, I., . . . Rokhsar, D. S. (2010). Genomic Analysis of Organismal Complexity in the Multicellular Green Alga Volvox carteri. *Science, 329*(5988), 223-226. doi:10.1126/science.1188800

19 Hanschen, E. R., Marriage, T. N., Ferris, P. J., Hamaji, T., Toyoda, A., Fujiyama, A., . . . Olson, B. J. (2016). The Gonium pectorale genome demonstrates co-option of cell cycle regulation during the evolution of multicellularity. *Nature Communications, 7*(1). doi:10.1038/ncomms11370

20 Hamaji, T., Kawai-Toyooka, H., Uchimura, H., Suzuki, M., Noguchi, H., Minakuchi, Y., . . . Nozaki, H. (2018). Anisogamy evolved with a reduced sex-determining region in volvocine green algae. *Communications Biology, 1*(1). doi:10.1038/s42003-018-0019-5

21 Jackman, S. D., Vandervalk, B. P., Mohamadi, H., Chu, J., Yeo, S., Hammond, S. A., . . . Birol, I. (2017). ABySS 2.0: Resource-efficient assembly of large genomes using a Bloom filter. *Genome Research, 27*(5), 768-777. doi:10.1101/gr.214346.116

22 Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology, 215*(3), 403-410. doi:10.1016/s0022-2836(05)80360-2

23 Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: Architecture and applications. *BMC Bioinformatics, 10*(1), 421. doi:10.1186/1471-2105-10-421

24 National Center for Biotechnology Information (NCBI)[Internet]. Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information; [1988]

25 Cock, P. J., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., . . . Hoon, M. J. (2009). Biopython: Freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics, 25*(11), 1422-1423. doi:10.1093/bioinformatics/btp163

26 Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods, 9*(4), 357-359. doi:10.1038/nmeth.1923

27 Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., & Phillippy, A. M. (2017). Canu: Scalable and accurate long-read assembly via adaptivek-mer weighting and repeat separation. *Genome Research, 27*(5), 722-736. doi:10.1101/gr.215087.116

28 Trapnell, C., Pachter, L., & Salzberg, S. L. (2009). TopHat: Discovering splice junctions with RNA-Seq. *Bioinformatics, 25*(9), 1105-1111. doi:10.1093/bioinformatics/btp120

29 Stanke, M., Tzvetkova, A., & Morgenstern, B. (2006). *Genome Biology, 7*(Suppl 1). doi:10.1186/gb-2006-7-s1-s11

30 Li, L. (2003). OrthoMCL: Identification of Ortholog Groups for Eukaryotic Genomes. *Genome Research, 13*(9), 2178-2189. doi:10.1101/gr.1224503

31 Sahm, A., Bens, M., Platzer, M., & Szafranski, K. (2017). PosiGene: Automated and easy-to-use pipeline for genome-wide detection of positively selected genes. *Nucleic Acids Research, 45*(11). doi:10.1093/nar/gkx179

32 Shelton, D. E., & Michod, R. E. (2014). Group Selection and Group Adaptation During a Major Evolutionary Transition: Insights from the Evolution of Multicellularity in the Volvocine Algae. *Biological Theory, 9*(4), 452-469. doi:10.1007/s13752-014-0159-x

33 Libby, E., & Rainey, P. B. (2013). A conceptual framework for the evolutionary origins of multicellularity. *Physical Biology, 10*(3), 035001. doi:10.1088/1478-3975/10/3/035001

34 Rainey, P. B., & Kerr, B. (2010). Cheats as first propagules: A new hypothesis for the evolution of individuality during the transition from single cells to multicellularity. *BioEssays, 32*(10), 872-880. doi:10.1002/bies.201000039

35 Heberle, H.; Meirelles, G. V.; da Silva, F. R.; Telles, G. P.; Minghim, R. *InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams*. BMC Bioinformatics 16:169 (2015)